

PROBABILIDAD Y ESTADÍSTICA

para las ciencias sociales del comportamiento y la salud

Primera edición
en español



William Mendenhall III • Robert J. Beaver • Barbara M. Beaver

Probabilidad y estadística para las ciencias sociales del comportamiento y la salud



**EDICIÓN
EN ESPAÑOL**

William Mendenhall, III

University of Florida, Emérito

Robert J. Beaver

University of California, Riverside, Emérito

Barbara M. Beaver

University of California, Riverside, Emérito

Adaptación:

Haroldo Elorza Pérez-Tejada

Traducción:

Jorge Alberto Velázquez Arellano

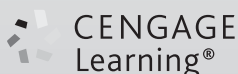
Revisión técnica:

M.I. Ángel Leonardo Bañuelos Saucedo

Profesor de Carrera Titular

Facultad de Ingeniería

Universidad Nacional Autónoma de México (UNAM)



Australia • Brasil • Corea • España • Estados Unidos • Japón • México • Reino Unido • Singapur

***Probabilidad y estadística para las ciencias
sociales del comportamiento y la salud
Primera edición en español***

William Mendenhall, III; Robert J. Beaver y
Barbara M. Beaver

Director Editorial para Latinoamérica:
Ricardo H. Rodríguez

**Editora de Adquisiciones para
Latinoamérica:**
Claudia C. Garay Castro

**Gerente de Manufactura para
Latinoamérica:**
Antonio Mateos Martínez

**Gerente Editorial en Español para
Latinoamérica:**
Pilar Hernández Santamarina

Gerente de Proyectos Especiales:
Luciana Rabuffetti

Coordinador de Manufactura:
Rafael Pérez González

Editora:
Ivonne Arciniega Torres

Diseño de portada:
Gloria Ivonne Álvarez López

Imagen de portada:
© AMV_80/Shutterstock.com

Composición tipográfica:
Tsuki Marketing S.A. de C.V.
Gerardo Larios García

© D.R. 2017 por Cengage Learning Editores, S.A. de C.V.,
una Compañía de Cengage Learning, Inc.
Corporativo Santa Fe
Av. Santa Fe núm. 505, piso 12
Col. Cruz Manca, Santa Fe
C.P. 05349, México, D.F.
Cengage Learning® es una marca registrada
usada bajo permiso.

DERECHOS RESERVADOS. Ninguna parte de
este trabajo amparado por la Ley Federal del
Derecho de Autor, podrá ser reproducida,
transmitida, almacenada o utilizada en
cualquier forma o por cualquier medio, ya sea
gráfico, electrónico o mecánico, incluyendo,
pero sin limitarse a lo siguiente: fotocopiado,
reproducción, escaneo, digitalización,
grabación en audio, distribución en Internet,
distribución en redes de información o
almacenamiento y recopilación en sistemas
de información a excepción de lo permitido
en el Capítulo III, Artículo 27 de la Ley Federal
del Derecho de Autor, sin el consentimiento
por escrito de la Editorial. Reg 703

Traducido del libro
Introduction to Probability and Statistics,
Fourteenth Edition
William Mendenhall, III; Robert J. Beaver y Barbara M.
Beaver
Publicado en inglés por Brooks/Cole, una compañía de
Cengage Learning © 2013, 2009
ISBN: 978-1-133-10375-2

Adaptado del libro *Introducción a la probabilidad y
estadística*, 14a. ed.
William Mendenhall, III; Robert J. Beaver
y Barbara M. Beaver
Publicado por Cengage Learning © 2015
ISBN: 978-607-519-876-7

Datos para catalogación bibliográfica:
Mendenhall, William, III; Robert J. Beaver y Barbara M.
Beaver
*Probabilidad y estadística para las ciencias sociales del
comportamiento y la salud*, primera edición en español
ISBN: 978-607-526-310-6

Visite nuestro sitio en:
<http://latinoamerica.cengage.com>

Prefacio

Probabilidad y estadística para las ciencias sociales del comportamiento y la salud, es una versión modificada de la decimocuarta edición de *Introducción a la probabilidad y estadística*, de William Mendenhall III, Robert J. Beaver y Barbara M. Beaver. Esta adaptación se realizó con el objetivo de que los lectores, ya sea como parte de su formación profesional o porque deben realizar un estudio o investigación, adquieran conocimientos de probabilidad y estadística aplicadas que les sean útiles en el manejo, la organización y el análisis de grandes cantidades de datos relacionados con las ciencias sociales y la salud.

Esta versión adaptada consta de 12 capítulos, dos apéndices, anexo, respuestas a ejercicios seleccionados y glosario. Se han conservado algunas características de la versión original. A continuación se detalla el contenido de la obra:

El capítulo 1 busca que los lectores sean capaces de ubicar la estadística dentro del contexto de la ciencia y la investigación, así como describir conjuntos de datos; se estudian los aspectos fundamentales de la ciencia, medidas de centro y variabilidad, medición estándar, mediciones de posición relativa y gráficas de caja.

En el capítulo 2 se estudia la probabilidad como herramienta estadística para crear distribuciones que servirán como modelos para variables aleatorias discretas y describir dichas variables usando una media y desviación estándar a fin de obtener conclusiones acerca de las poblaciones muestreadas.

En el capítulo 3 se presentan tres variables aleatorias discretas importantes: binomial, de Poisson e hipergeométrica que por lo general se usan para describir el número de sucesos de un evento especificado en un número fijo de intentos o una unidad fija de tiempo o espacio.

En el capítulo 4 se estudian las muestras y las estadísticas que las describen. Dichas estadísticas se usan para hacer inferencias sobre los parámetros correspondientes de las poblaciones.

En el capítulo 5 se presenta un método para estimar los parámetros poblacionales y se ilustran con ejemplos prácticos los conceptos básicos de estimación estadística de muestra grande para situaciones que involucran medias y proporciones poblacionales.

El capítulo 6 complementa las técnicas de muestra grande al presentar las pruebas de muestra pequeña y los intervalos de confianza para medias y varianzas poblacionales.

En el capítulo 7 se incluyen tres diseños experimentales diferentes; se usa el análisis de varianza para determinar el modo en que los diferentes factores experimentales afectan la respuesta promedio.

En el capítulo 8 se considera la situación en la que el valor medio de una variable y se relaciona con otra variable x ; es posible usar la información dada por x para estimar el valor promedio de y , y predecir valores de y para valores de x asignados previamente.

En el capítulo 9 se amplían los conceptos de regresión y correlación lineales a una situación en la que el valor promedio de una variable aleatoria y está relacionada con varias variables $x_1, x_2, \dots, x_k$, en modelos más flexibles que el modelo de recta.

En el capítulo 10 se estudian métodos para analizar los datos categóricos provenientes de numerosos tipos de estudios y experimentos que resultan en variables cualitativas y no cuantitativas.

En el capítulo 11 se presentan varias pruebas estadísticas útiles en la comparación de poblaciones para los numerosos tipos de datos que no satisfagan las suposiciones especificadas en los capítulos 5 y 7.

En el capítulo 12 se explican los principales supuestos conceptuales y estadísticos de la teoría de respuesta al ítem.

CARACTERÍSTICAS DE LA 14^a. ED. (ADAPTADA)

Se ha conservado algunas características de la decimocuarta edición de *Introducción a la probabilidad y estadística*:



NECESITO SABER...

Esta sección proporciona información consistente sobre definiciones, procedimientos o sugerencias paso a paso sobre la solución de problemas para cuestiones específicas.



NECESITO SABER...

Cómo calcular la probabilidad de un evento

1. Haga una lista de todos los eventos sencillos del espacio muestral.
2. Asigne una probabilidad apropiada a cada evento simple.
3. Determine cuáles eventos sencillos resultan en el evento de interés.
4. Sume las probabilidades de los eventos sencillos que resulten en el evento de interés.



MI CONSEJO

Son numerosas sugerencias breves y concisas que aparecen en los márgenes del texto



MI CONSEJO

Regla empírica $\Leftrightarrow$ datos en forma de montículo
Chebyshev $\Leftrightarrow$ datos en cualquier forma

¿Es aplicable el teorema de Chebyshev? Sí, porque es posible usarlo para cualquier conjunto de datos. De acuerdo con el teorema de Chebyshev,

- al menos 3/4 de las mediciones caerán entre 10.6 y 32.6.
- al menos 8/9 de las mediciones caerán entre 5.1 y 38.1.

Ejercicios y Ejercicios suplementarios. La variedad y el número de aplicaciones reales en los conjuntos de ejercicios es la mayor fortaleza de esta edición. Se han revisado los conjuntos de ejercicios para darle nuevas e interesantes situaciones del mundo real y conjuntos de datos reales, muchos de ellos extraídos de periódicos y revistas científicas recientes.

2.7

EJERCICIOS

TÉCNICAS BÁSICAS

2.1 Tiro de un dado Un experimento consiste en tirar un solo dado. Éstos son algunos eventos:

- A: observar un 2
- B: observar un número par
- C: observar un número mayor que 2

2.2 Un espacio muestral S está formado por cinco eventos sencillos con estas probabilidades:

$$P(E_1) = P(E_2) = .15 \quad P(E_3) = .4 \\ P(E_4) = 2P(E_5)$$

- a. Encuentre las probabilidades para los eventos sencillos $E_1, \dots, E_5$.

Repaso del capítulo (conceptos clave y fórmulas) Las secciones llamadas “Conceptos clave y fórmulas” aparecen en cada capítulo como un repaso a manera de esbozo del material cubierto en ese capítulo.

REPASO DEL CAPÍTULO

Conceptos y fórmulas clave

I. Planes muestrales y diseños experimentales

1. Muestreo aleatorio simple
 - a. Cada posible muestra de tamaño n es igualmente probable de ocurrir
 - b. Use una computadora o tabla de números aleatorios
 - c. Los problemas son: sin respuesta, baja cobertura y sesgo verbal
2. Otros planes muestrales con aleatorización
 - a. Muestreo aleatorio estratificado

III. Distribución muestral de la media muestral

1. Cuando muestras de tamaño n se sacan al azar de una población normal con media μ y varianza σ^2 , la media muestral $\bar{x}$ tiene una distribución normal con media μ y desviación estándar $\sigma/\sqrt{n}$
2. Cuando muestras de tamaño n se sacan al azar de una población no normal con media μ y varianza σ^2 , el teorema central del límite asegura que la media muestral $\bar{x}$ tendrá una distribución aproximadamente normal con media μ y desviación estándar $\sigma/\sqrt{n}$ cuando n es grande ($n \geq 30$).



TECNOLOGÍA ACTUAL

Los alumnos podrán usar la computadora para hacer análisis estadístico estándar y como una herramienta para reforzar y visualizar conceptos estadísticos. Tanto *MS Excel* como *MINITAB 16* (consistente con versiones anteriores de *MINITAB*) se usan en forma exclusiva como los paquetes de cómputo para análisis estadístico. Sin embargo, hemos elegido aislar las instrucciones para generar salidas de computadora en secciones individuales llamadas “Tecnología actual” al final de cada capítulo. Cada exposición usa ejemplos numéricos para guiar al estudiante a través de los comandos de *MS Excel* y las opciones necesarias para los procedimientos presentados en ese capítulo, y luego presenta los pasos y comandos equivalentes necesarios para producir los mismos resultados o similares usando *MINITAB*. Se han incluido capturas de pantalla tanto de *MS Excel* como de *MINITAB 16*, de modo que el estudiante pueda trabajar realmente en estas secciones como “minilaboratorios”. Si no necesita un conocimiento “práctico” de *MINITAB* o *MS Excel*, o si usted usa otro paquete de software, puede saltarse estas secciones y simplemente usar las salidas impresas como guías para la comprensión básica de las salidas impresas de computadora.



TECNOLOGÍA ACTUAL

El teorema central del límite en operación: *Microsoft Excel*

Microsoft Excel puede usarse para explorar la forma en que funciona el teorema central del límite en la práctica. Recuerde que, de acuerdo con el teorema central del límite, si se extraen muestras aleatorias de tamaño n de una población no normal con media μ y desviación estándar σ , entonces cuando n es grande, la distribución muestral de la media muestral $\bar{x}$ será aproximadamente normal con la misma media μ y con error estándar $\sigma/\sqrt{n}$. Probemos muestrear de una población no normal usando *Excel*.

En una hoja de cálculo nueva, genere 100 muestras de tamaño $n = 30$ de una distribución uniforme continua del intervalo (0, 10). Etiquete la columna A “Muestra” e introduzca los números 1 a 100 en esa columna. Luego seleccione **Data ► Data Analysis ► Random Number Generation**, para obtener el cuadro de diálogo de la figura 4.15a). Escriba 30 para el número de variables y 100 para el número de números aleatorios. En la lista desplegable “Distribution” seleccione “Uniform” con parámetros entre 0 y 10. Del mismo modo, la primera fila

RECONOCIMIENTOS

Los autores agradecen a Molly Taylor y al personal editorial de Cengage Learning por su paciencia, asistencia y cooperación en la preparación de esta edición. Un agradecimiento especial a Gary McClelland por las applets Java usadas en el texto.

También se agradece a los revisores de la decimocuarta edición Ronald C. Degges, Bob C. Denton, Dra. Dorothy M. French, Jungwon Mun, Kazuhiko Shinki, Florence P. Shu y a los revisores de la décimo tercera edición Bob Denton, Timothy Husband, Rob LaBorde, Craig McBride, Marc Sylvester, Kanapathi Thiru y Vitaly Voloshin. Deseamos agradecer a los autores y organizaciones por permitirnos reimprimir material selecto; se hacen reconocimientos siempre que tal material aparece en el texto.

Robert J. Beaver
Barbara M. Beaver

Contenido breve

1	DESCRIPCIÓN DE DATOS CON MEDIDAS NUMÉRICAS	1
2	PROBABILIDAD	54
3	ALGUNAS DISTRIBUCIONES DE PROBABILIDAD IMPORTANTES	116
4	DISTRIBUCIONES MUESTRALES	187
5	ESTIMACIÓN DE MUESTRAS GRANDES	224
6	INFERENCIA A PARTIR DE MUESTRAS PEQUEÑAS	262
7	ANÁLISIS DE VARIANZA	320
8	REGRESIÓN LINEAL Y CORRELACIÓN	377
9	ANÁLISIS DE REGRESIÓN MÚLTIPLE	428
10	ANÁLISIS DE DATOS CATEGÓRICOS	471
11	ESTADÍSTICA NO PARAMÉTRICA	521
12	TEORÍA DE LA RESPUESTA AL ÍTEM	587
	APÉNDICE A LOS ESCRITOS CIENTÍFICOS	611
	APÉNDICE B MATRICES	631
	ANEXO	643
	FUENTES DE DATOS	684
	RESPUESTAS A EJERCICIOS SELECCIONADOS	691
	GLOSARIO	701

Contenido

1

DESCRIPCIÓN DE DATOS CON MEDIDAS NUMÉRICAS 1

- 1.1 Aspectos fundamentales de la ciencia 2
 - Propósitos 2
 - Introducción 2
 - Relaciones entre estadística e investigación 7
 - Medición y estadística 9
 - Inferencia estadística y científica 11
 - Estadística e informe científico 13
- 1.2 Descripción de un conjunto de datos con medidas numéricas 14
- 1.3 Medidas de centro 14
 - Ejercicios 18
- 1.4 Medidas de variabilidad 20
 - Ejercicios 24
- 1.5 Sobre la significación práctica de la medición estándar 25
- 1.6 Una medición del cálculo de s 29
 - Ejercicios 31
- 1.7 Mediciones de posición relativa 34
- 1.8 El resumen de cinco números y la gráfica de caja 38
 - Ejercicios 41
- Repaso del capítulo 44
- Tecnología actual 45
- Ejercicios suplementarios 48

2

PROBABILIDAD 54

- 2.1 Introducción 55
- 2.2 El papel de la probabilidad en estadística 56
- 2.3 Eventos y el espacio muestral 56
- 2.4 Enfoques o escuelas de la probabilidad 59
- 2.5 Axiomas de probabilidad 60
- 2.6 Particiones 62
- 2.7 Cálculo de probabilidades con el uso de eventos sencillos 63
 - Ejercicios 66
- 2.8 Reglas útiles de conteo 68
 - Ejercicios 73

2.9	Relaciones de evento y reglas de probabilidad	74
	Cálculo de probabilidades para uniones y complementos	76
2.10	Independencia, probabilidad condicional y la regla de la multiplicación	79
2.11	Probabilidad condicional	81
	Ejercicios	84
2.12	Teorema de Bayes	87
	Ejercicios	90
2.13	¿Eventos mutuamente excluyentes o independientes?	91
2.14	Procesos estocásticos	93
	Cadenas de Markov	94
	Representación gráfica	94
	Representación matricial	96
	Resumen	107
	Ejercicios	107
	Ejercicios suplementarios	110

3

ALGUNAS DISTRIBUCIONES DE PROBABILIDAD IMPORTANTES 116

3.1	Introducción	117
3.2	Variables aleatorias discretas y sus distribuciones de probabilidad	117
	Variables aleatorias	117
	Distribuciones de probabilidad	117
	La media y desviación estándar para una variable aleatoria discreta	119
	Ejercicios	123
3.3	La distribución binomial de probabilidad	125
	Ejercicios	133
3.4	Distribuciones de probabilidad para variables aleatorias continuas	136
3.5	La distribución de probabilidad de Poisson	139
	Ejercicios	144
3.6	La distribución hipergeométrica de probabilidad	145
	Ejercicios	147
3.7	La distribución normal de probabilidad	148
3.8	Áreas tabuladas de la distribución normal de probabilidad	149
	La variable aleatoria normal estándar	149
	Cálculo de probabilidades para una variable aleatoria normal general	153
	Ejercicios	156
3.9	La aproximación de la distribución de probabilidad binomial a la normal	158
	Ejercicios	163
	Repaso del capítulo	164
	Tecnología actual	166

Ejercicios suplementarios 177

CASO PRÁCTICO: Un misterio: casos de cáncer cerca de un reactor 185

CASO PRÁCTICO: “¿Va a calificar por curva?” 186

4

DISTRIBUCIONES MUESTRALES 187

4.1 Introducción 188

4.2 Planes muestrales y diseños experimentales 188

Ejercicios 191

4.3 Estadísticas y distribuciones muestrales 192

4.4 El teorema central del límite 195

4.5 La distribución muestral de la media muestral 198

Error estándar 199

Ejercicios 202

4.6 La distribución muestral de la proporción muestral 204

Ejercicios 208

4.7 Una aplicación muestral: control estadístico de procesos (opcional) 209

Una gráfica de control para la media del proceso: la gráfica $\bar{x}$ 209

Una gráfica de control para la proporción de piezas defectuosas: la gráfica p 211

Ejercicios 213

Repaso del capítulo 215

Tecnología actual 216

Ejercicios suplementarios 219

CASO PRÁCTICO: Muestreo de la Ruleta de Monte Carlo 222

5

ESTIMACIÓN DE MUESTRAS GRANDES 224

5.1 Dónde hemos estado 225

5.2 A dónde vamos: inferencia estadística 225

5.3 Tipos de estimadores 226

5.4 Estimación puntual 227

Ejercicios 232

5.5 Estimación de intervalo 233

Construcción de un intervalo de confianza 234

Intervalo de confianza de muestra grande para una media poblacional μ 236

Interpretación del intervalo de confianza 237

Intervalo de confianza de muestra grande para una proporción poblacional p 239

Ejercicios 241

5.6 Estimación de la diferencia entre dos medias poblacionales 242

Ejercicios 245

- 5.7 Estimación de la diferencia entre dos proporciones binomiales 248
Ejercicios 250
- 5.8 Límites de confianza a una cola 252
- 5.9 Selección del tamaño muestral 253
Ejercicios 257
- Repaso del capítulo 258**
- Ejercicios suplementarios 259**

6

INFERENCIA A PARTIR DE MUESTRAS PEQUEÑAS 262

- 6.1 Introducción 263
- 6.2 Distribución t de Student 263
Suposiciones detrás de la distribución t de Student 266
- 6.3 Inferencias de muestra pequeña respecto a una media poblacional 267
Ejercicios 271
- 6.4 Inferencias de muestra pequeña para la diferencia entre dos medias poblacionales: muestras aleatorias independientes 274
Ejercicios 280
- 6.5 Inferencias de muestra pequeña para la diferencia entre dos medias: una prueba de diferencia en pares 283
Ejercicios 288
- 6.6 Inferencias respecto a la varianza poblacional 290
Ejercicios 296
- 6.7 Comparación de dos varianzas poblacionales 297
Ejercicios 303
- 6.8 Repaso de suposiciones de muestra pequeña 305
Repaso del capítulo 306
Tecnología actual 307
Ejercicios suplementarios 313

7

ANÁLISIS DE VARIANZA 320

- 7.1 El diseño de un experimento 321
- 7.2 ¿Qué es un análisis de varianza? 322
- 7.3 Las suposiciones para un análisis de varianza 322
- 7.4 El diseño completamente aleatorizado: una clasificación en una dirección 323
- 7.5 El análisis de varianza para un diseño completamente aleatorizado 324
División de la variación total en un experimento 324
Prueba de igualdad de las medias de tratamiento 327
Estimación de diferencias en las medias de tratamiento 329
Ejercicios 332

7.6	Clasificación de medias poblacionales	335
	Ejercicios	338
7.7	Diseño de bloque aleatorizado: una clasificación en dos direcciones	339
7.8	El análisis de varianza para un diseño de bloque aleatorizado	340
	División de la variación total en el experimento	340
	Prueba de igualdad de las medias de tratamiento y de bloque	343
	Identificación de diferencias en las medias de tratamiento y de bloque	345
	Algunos comentarios de precaución en bloque	346
	Ejercicios	347
7.9	El experimento factorial $a \times b$: una clasificación en dos vías	351
7.10	El análisis de varianza para un experimento factorial $a \times b$	353
	Ejercicios	357
7.11	Repaso de las suposiciones del análisis de varianza	361
	Gráficas residuales	361
7.12	Un breve repaso	363
	Repaso del capítulo	364
	Tecnología actual	365
	Ejercicios suplementarios	370
	CASO PRÁCTICO: ¡Cómo ahorrar dinero en comestibles!	376

REGRESIÓN LINEAL Y CORRELACIÓN 377

8.1	Introducción	378
8.2	Modelo probabilístico lineal simple	378
8.3	El método de mínimos cuadrados	381
8.4	Un análisis de varianza para regresión lineal	383
	Ejercicios	386
8.5	Prueba de la utilidad del modelo de regresión lineal	389
	Inferencias respecto a β , la pendiente de la recta de medias	390
	El análisis de varianza de la prueba F	393
	Medir la fuerza de la relación: el coeficiente de determinación	393
	Interpretación de los resultados de una regresión significativa	394
	Ejercicios	395
8.6	Herramientas de diagnóstico para verificar suposiciones de la regresión	398
	Términos de error dependientes	398
	Gráficas residuales	398
	Ejercicios	399
8.7	Estimación y predicción usando la recta ajustada	402
	Ejercicios	406
8.8	Análisis de correlación	408
	Ejercicios	412

- 8.9 Covarianza 414
- Repaso del capítulo 417
- Tecnología actual 418
- Ejercicios suplementarios 421
- CASO PRÁCTICO: ¿Su automóvil está “Hecho en EUA”? 426

9

ANÁLISIS DE REGRESIÓN MÚLTIPLE 428

- 9.1 Introducción 429
- 9.2 El modelo de regresión múltiple 429
- 9.3 Un análisis de regresión múltiple 430
 - El método de mínimos cuadrados 431
 - El análisis de varianza para regresión múltiple 432
 - Prueba de la utilidad del modelo de regresión 433
 - Interpretación de los resultados de una regresión significativa 434
 - Comprobación de suposiciones de regresión 436
 - Uso del modelo de regresión para estimación y predicción 436
- 9.4 Un modelo de regresión polinomial 437
 - Ejercicios 440
- 9.5 Uso de variables predictoras cuantitativas y cualitativas en un modelo de regresión 444
 - Ejercicios 450
- 9.6 Prueba de conjuntos de coeficientes de regresión 453
- 9.7 Interpretación de gráficas residuales 456
- 9.8 Análisis de regresión por pasos 457
- 9.9 Interpretación errónea de un análisis de regresión 458
 - Causalidad 458
 - Multicolinealidad 458
- 9.10 Pasos a seguir al construir un modelo de regresión múltiple 460
 - Repaso del capítulo 460
 - Tecnología actual 461
 - Ejercicios suplementarios 463

10

ANÁLISIS DE DATOS CATEGÓRICOS 471

- 10.1 Introducción 472
 - Estudios evaluativos, un enfoque actual 472
 - Los diferentes objetos de la evaluación 473
 - Estudios evaluativos: procedimientos generales 474
 - Áreas de interés del estudio evaluativo 477
 - Programas susceptibles de evaluación 477
 - Interpretación de los resultados 477

10.2	Una descripción del experimento	479
10.3	Estadístico ji cuadrada de Pearson	480
10.4	Prueba de probabilidades de celda especificada: la prueba de bondad de ajuste	481
	Ejercicios	483
10.5	Tablas de contingencia: una clasificación de dos vías	485
	La prueba de independencia ji cuadrada	485
	Ejercicios	490
10.6	Procedimiento <i>post hoc</i>	491
	Coeficiente $f_i(\Phi)$	493
	Coeficiente de contingencia (C)	496
	Prueba de significancia	498
	Coeficiente V de Kramer	499
10.7	Prueba exacta de Fisher	501
10.8	Prueba de McNemar	502
10.9	Comparación de varias poblaciones multinomiales: una clasificación de dos vías con totales de fila o columna fijos	504
	Ejercicios	507
10.10	La equivalencia de pruebas estadísticas	509
10.11	Otras aplicaciones del estadístico de prueba ji cuadrada	509
	Repaso del capítulo	511
	Tecnología actual	511
	Ejercicios suplementarios	515

ESTADÍSTICA NO PARAMÉTRICA 521

11.1	Introducción	522
11.2	La prueba de suma de rango de Wilcoxon: muestras aleatorias independientes	522
	Aproximación normal para la prueba de suma de rango de Wilcoxon	526
	Ejercicios	529
11.3	La prueba del signo para un experimento de dos poblaciones	531
	Aproximación normal para la prueba del signo	532
	Ejercicios	534
11.4	Una comparación de pruebas estadísticas	535
11.5	La prueba de rango con signo de Wilcoxon para un experimento de dos poblaciones	536
	Aproximación normal para la prueba de rango con signo de Wilcoxon	539
	Ejercicios	540
11.6	La prueba H de Kruskal-Wallis para diseños completamente aleatorizados	542
	Ejercicios	546
11.7	La prueba F_r de Friedman para diseños de bloque aleatorizados	548

11.8	Prueba de Nemenyi	551
	Ejercicios	552
11.9	Prueba de la mediana	554
11.10	Coeficiente de correlación de rango	556
11.11	Prueba de significancia de r_s	560
11.12	Coeficiente tau (τ) de Kendall	561
11.13	Coeficiente de concordancia (ω) de Kendall	564
	Prueba de significancia de ω	566
11.14	Coeficiente de correlación (r_{bp}) biserial de punto	566
	Prueba de significancia de r_{bp}	569
11.15	Prueba de Kappa	570
	Ejercicios	572
11.16	Resumen	575
	Repaso del capítulo	576
	Tecnología actual	577
	Ejercicios suplementarios	580
	CASO PRÁCTICO: ¿Cómo está su nivel de colesterol?	585

TEORÍA DE LA RESPUESTA AL ÍTEM 587

12.1	Introducción	588
12.2	Teoría clásica de los tests en la psicometría	588
	Supuestos básicos de la teoría de la puntuación verdadera	589
	Confiabilidad de un test	589
	Condiciones de paralelismo	590
	Características de los ítems en la TCT	590
	Principales limitaciones de la teoría clásica de los tests	591
12.3	¿Qué ofrece la teoría de la respuesta al ítem?	591
	Curva característica del ítem (CCI)	592
	Modelo ideal de Guttman y parámetros de un ítem	593
	Índice de dificultad	593
	Discriminación de un ítem	594
	Parámetro de la seudoadivinación	594
	Modelo de ojiva normal	594
	Reparametrización del modelo de ojiva normal	595
	Modelo logístico de un parámetro o modelo de Rasch	597
	Modelo logístico de dos parámetros	599
	Modelo logístico con tres parámetros	599
12.4	Principales supuestos de la TRÍ	600
	Unidimensionalidad del test	600
	Indeterminación de la escala de rasgo latente	600

12.5	Estimación de parámetros del examinado y los ítems	601
	Método de estimación de máxima verosimilitud	601
	Estimación de los parámetros: a y b	602
12.6	Función de información	602
	Usos de la función de información	603
	Función de información del test	603
12.7	Evaluación de bondad de ajuste del modelo	604
	Interpretación del índice de bondad de ajuste	604
12.8	Modelos politómicos de la teoría de la respuesta al ítem	605
	Modelos politómicos para categorías ordenadas	606
	Modelo de respuesta graduada	607
	Ventajas de los modelos politómicos	609
12.9	Resumen	610

APÉNDICE A LOS ESCRITOS CIENTÍFICOS 611

APÉNDICE B MATRICES 631

ANEXO 643

FUENTES DE DATOS 684

RESPUESTAS A EJERCICIOS SELECCIONADOS 691

GLOSARIO 701

Descripción de datos con medidas numéricas

OBJETIVOS GENERALES

Las gráficas son sumamente útiles para la descripción visual de un conjunto de datos, pero no siempre son la mejor herramienta cuando se desea hacer inferencias acerca de una población a partir de la información contenida en una muestra. Para este propósito, es mejor usar medidas numéricas para construir una imagen mental de los datos.

ÍNDICE DEL CAPÍTULO

- Aspectos fundamentales de la ciencia (1.1)
- Descripción de un conjunto de datos con medidas numéricas (1.2)
- Medidas de centro (1.3)
- Medidas de variabilidad (1.4)
- Sobre la significación práctica de la medición estándar (1.5)
- Una medición del cálculo de s (1.6)
- Mediciones de posición relativa (1.7)
- El resumen de cinco números y la gráfica de caja (1.8)



NECESITO SABER...

Cómo calcular cuartiles muestrales

Los muchachos de verano

¿Los campeones de béisbol de hoy son mejores que los de “ayer”? ¿Los jugadores de la Liga Nacional batean mejor que los de la Liga Americana? El estudio práctico del final de este capítulo contiene los promedios de bateo de campeones de las ligas mayores. Se pueden usar medidas numéricas descriptivas para contestar éstas y otras preguntas similares.

ASPECTOS FUNDAMENTALES DE LA CIENCIA

Propósitos

El objetivo central del presente capítulo es que el lector sea capaz de ubicar a la estadística dentro del contexto de la ciencia y la investigación.

De igual forma, al término del mismo el lector podrá:

- Reconocer que la ciencia ha facilitado el desarrollo de las teorías que exponen la realidad.
- Comprender la conceptualización del empirismo y del positivismo en el proceso de acumulación del conocimiento.
- Explicar la forma en la que las teorías constituyen simplemente la organización lógica de las leyes empíricas.
- Enunciar la forma que tiene el empirismo de entender la ciencia.
- Reconocer que toda teoría, todo modelo y toda ley científica son una conjetura acerca de cómo es la realidad.
- Relacionar la opinión de Popper de que toda ley, principio, teoría o modelo es una conjetura o suposición.
- Diferenciar el punto de vista de los positivistas y los justificacionistas con relación a la ciencia.
- Considerar el punto de vista de Popper acerca de que: “Lo que caracteriza al hombre de ciencia no es la posesión del conocimiento o de verdades irrefutables, sino la investigación desinteresada e incesante de la verdad”.

Introducción

La ciencia es una de las empresas más humanas y productivas que haya desarrollado el hombre. Si lo que caracteriza al ser humano es su excepcional inteligencia, la cual le ha dotado de lenguaje y le ha permitido servirse de él para crear una singular organización social, de insólita eficacia, para dominar la naturaleza, entonces la ciencia es el logro humano más perfecto y contundente, el cual señala la cúspide de los frutos de su intelecto, único en el Sistema Solar y tal vez en el universo mismo.

La ciencia basada en un proceso analítico y crítico produce el conocimiento que ha permitido una mejor comprensión de la realidad circundante. Asimismo, ha facultado al hombre para penetrar en los secretos más profundos del mundo, incluido el ser del hombre mismo. La ciencia ha facilitado el desarrollo de teorías que exponen la realidad, con base en un examen de la relación entre los intentos de explicación teórica, evidencia empírica y congruencia lógica, tanto interna a la explicación como en lo relativo a otras teorías con las que tienen vínculos conceptuales. Esto ha implicado que el científico pruebe sus teorías confrontándolas con la evidencia existente que, con el objeto de evaluar la teoría de que se trata, se acumula con procedimientos rigurosos. Asimismo, el científico está a la caza de inconsistencias internas en la lógica de las explicaciones, así como de las contradicciones entre las diversas teorías vinculadas.

La conceptualización del empirismo y del positivismo acerca de la naturaleza del proceso de acumulación de conocimiento se ha sustentado siempre en el proceso de inducción. Este principio señala, tal como lo plantea Hume, que si observa una cierta regularidad en los procesos naturales (incluida la naturaleza humana), entonces es posible generalizar a partir del establecimiento de una ley. De acuerdo con esta visión, el problema de la ciencia es observar cuidadosamente la naturaleza, evitando caer en errores debidos a la posible confusión de causas. El mejor modo de evitar el error es realizar una cuidadosa observación y medición del fenómeno y utilizar el método experimental para no confundir la verdadera causa de los fenómenos con otras que en apariencia los producen. De acuerdo con ellos, los hechos observados y establecidos prueban una cierta concepción de la realidad. Al ser entonces el proceso científico un proceso lineal y acumulativo, las teorías constituirían simplemente la organización lógica de las leyes empíricas y la explicación de varias de ellas por principios más generales, surgidos de la inducción. Ésta es la forma que tiene el empirismo de entender la ciencia y, con ciertas modificaciones, el positivismo.

Hume ya había planteado la naturaleza de las limitaciones lógicas del conocimiento inductivo: independientemente de cuántas observaciones se hayan hecho de una regularidad, esto no da ninguna “garantía lógica” de que volverá a ocurrir del mismo modo en la siguiente ocasión. La solución planteada por Karl Popper (1972) a este dilema se hizo en términos de postular que nunca se puede partir de ninguna certidumbre acerca de nada de lo que se cree. De acuerdo con él, toda teoría, todo modelo o toda ley científica, es una conjetura de cómo es la realidad; no importa que su origen sea la inducción, un conocimiento tácito, tal vez de carácter personal, o una especulación; la teoría es una conjetura, una suposición, una hipótesis acerca de la realidad.

Las teorías, dice este autor, basan su desarrollo en la confrontación crítica con los hechos y con la lógica. En sus palabras, “... ningún conjunto de enunciados contrastadores verdaderos podrá justificar la pretensión de que una teoría explicativa universal es verdadera”.[†] Sin embargo, afirma que: “suponiendo que los enunciados contrastadores sean verdaderos, con base en ellos a veces podemos justificar la pretensión de que una teoría explicativa universal es falsa”.^{††} Esto desplaza el énfasis de la investigación al sentido contrario de como lo plantea el punto de vista tradicional científico: no es posible probar que las teorías sean verdaderas, sólo es factible eliminar las falsas. Por ello, Popper dice: “el método de la ciencia es el método de las conjeturas audaces e ingeniosas seguidas por intentos rigurosos de refutarlas”.^{†††}

Esto hace de la ciencia una aventura fascinante, donde las teorías se tienen que construir; hay que inventarlas sobre la base de lo que ya se comprende del fenómeno en cuestión. No obstante, lo que hace a la ciencia más emocionante aún, es la posibilidad de someter las teorías a rigurosas pruebas de evidencia. Por un lado, esto otorga un grado mucho mayor de libertad, pero también un enorme sentido de responsabilidad.

De acuerdo con la opinión de Popper, toda ley, todo principio, toda teoría o todo modelo es una conjetura, una suposición. Las teorías no surgen, como supondrían los positivistas, mediante el proceso de inducción a partir de los datos, que, en todo caso, tan sólo proporcionan una inspiración inicial para la concepción de una teoría y no son una base empírica para el proceso lógico de la generalización por inducción. Los datos, cuando se generan *a posteriori*, sirven también para poner a prueba la elaboración de una ley o teoría, y si ésta resulta rechazada, es precisamente la naturaleza de las fallas la que podría servir de inspiración para el posterior planteamiento.

Las teorías se valoran por su poder explicativo y heurístico. Por tanto, son mejores las teorías que explican más hechos conocidos, las que tienen menos hechos que las contradicen y, sobre todo, las que permiten internarse en lo desconocido haciendo pronósticos no triviales y novedosos, sobre cuya base se les somete a pruebas rigurosas. El carácter riguroso de la

[†] K. R. Popper, *Conocimiento objetivo*, Tecnos, Madrid, 1974, p. 20.

^{††} *Ibid.*, p. 20.

^{†††} *Ibid.*, p. 83.

contrastación hace que las teorías cuantitativas sean mejores, permiten mayor precisión en la elaboración del pronóstico y, por tanto, en la prueba de ellas, ya que permite señalar con toda exactitud el grado de error de pronóstico y decidir si éste sólo se debe a un error de medición, o si se debe a una falla de la teoría.

Explicación y teoría

El papel de la teoría es explicar, proporcionar una comprensión de fenómenos, leyes, principios y cualquier otro tipo de *hecho* por medio de postulados generales, mecanismos internos, entes hipotéticos, procesos subyacentes o cualquier otro artificio intelectual; los que se combinan para proporcionar una estructura que dé cuenta racional de aquello que se pretende explicar. Es decir, las teorías tratan de dar sentido a aquello que explican, ubicándolo en la naturaleza y haciendo explícitas sus propiedades y relaciones con otros entes.

El propósito de la explicación es profundizar en la comprensión de los fenómenos. Por ejemplo, en el área de la química, Robert Boyle había desarrollado la distinción taxonómica entre elementos y compuestos. A partir de esa base, Proust elaboró la ley empírica de las proporciones constantes, la cual sostiene que los elementos tienen que combinarse en una determinada proporción de peso, para producir una reacción que genere un compuesto específico, sin que ninguno de los elementos que participaron en la reacción sobre, de modo que se requiere que estos elementos guarden una relación que se pueda expresar por medio de números enteros. Cuando esta proporción no se cumplía, la reacción no era completa y sobraban los elementos que tenían una proporción mayor a la estipulada. Esta ley empírica era suficiente para manejar coherentemente los fenómenos de la química que influían en las reacciones entre sustancias. Sin embargo, Dalton, un inglés, modesto profesor de primaria, introdujo una de las especulaciones más fructíferas en la historia de la humanidad: explicó esas regularidades numéricas suponiendo que la materia es discontinua y, retomando la idea de Leucipo y Demócrito, postuló la existencia de átomos para explicar esos hechos. De acuerdo con Dalton, los átomos de cada elemento se unen en combinaciones determinadas para formar moléculas de compuestos, las cuales son apiñamientos de átomos en estructuras determinadas. Es entonces el número de átomos de cada clase, que existe en cada molécula de un compuesto específico, lo que define la proporción de los elementos que deben entrar en la reacción para que no sobren átomos de un tipo u otro.

No ha existido una propuesta más fértil que ésta. Al poco tiempo, no sólo daba cuenta de los fenómenos conocidos de la química, sino que asimiló la ley de Boyle-Mariott de los gases a la explicación atómica, mediante la teoría cinética de los gases, que se basó en una aplicación de la mecánica newtoniana al movimiento de los átomos y las moléculas.

Como puede observar ahora, las teorías son instrumentos intelectuales muy poderosos que permiten dar sentido a la apabullante complejidad de la experiencia fenoménica, así como lidiar con la realidad por medio de la creación de un esquema conceptual de ésta, el que supone que es así en verdad. En este sentido, la ciencia es el instrumento intelectual más importante logrado por la humanidad, después de la invención de la escritura. La ciencia permite al hombre entender y anticipar el mundo que lo rodea, gracias al desarrollo de teorías que se asemejan cada vez más a la realidad, ya que, como lo señala Popper, las teorías van siendo, por selección natural, cada vez mejores mapas conceptuales de la realidad y cada vez más exactos y precisos. Las teorías se transforman en las mejores guías para la praxis humana, permitiendo el desarrollo de las poderosas tecnologías que caracterizan a la época moderna y haciendo factible el enorme éxito de la especie, por el que la humanidad ha logrado la población con la que actualmente cuenta.

Naturaleza de la investigación

La investigación se considera no sólo la parte creativa de la ciencia con la que se busca expandir el conocimiento y comprensión de la realidad, sino también la base que permitirá construir un mapa de ésta capaz de guiar al hombre en su búsqueda. Los mapas que proporciona la

ciencia no son únicamente esquemas descriptivos sino conceptuales-causales del mundo circundante; es decir, son guías en relación con las clases de objetos y eventos y sus conexiones causales recíprocas. Así, en función de esta situación el hombre avanza en su dominio cognoscitivo de la realidad.

La naturaleza de la ciencia y, por ende, de la investigación, han sido explicadas a través de la rama de la filosofía denominada *Filosofía de la ciencia*.[†] Esta disciplina es un esfuerzo del razonamiento humano por comprender cuál es el fundamento de esa actividad tan exitosa llamada ciencia. La filosofía, entendida como la reflexión sobre la naturaleza última de la realidad y de la existencia humana, lleva a un razonamiento acerca de la relación cognoscitiva existente entre el hombre y la realidad, que es la rama denominada *epistemología*. Dentro de esa reflexión se encuentra ubicado un análisis más específico del proceso de adquisición de conocimiento por medio de la ciencia.

La ciencia, como tal, surge en forma sistemática y organizada entre los griegos. La ciencia se desarrolló en el año 600 a.C. en las mentes inquietas e inquisitivas de investigadores de la naturaleza y de filósofos que buscaban la esencia de la realidad, incluida la naturaleza del conocimiento; desde la filosofía de la ciencia de Aristóteles, Platón, Demócrito, etc., hasta las contribuciones empíricas y teóricas concretas de Anaxágoras, Aristarco, Arquímedes, entre otros.

Sin embargo, no fue sino hasta que inició el Renacimiento cuando surgió de nuevo un concepto sistemático del proceder científico para el avance del conocimiento; es decir, una búsqueda activa de la verdad a través de la experiencia y la puesta a prueba empírica de las hipótesis, siendo un hecho que casi todo lo que distingue al mundo moderno de los siglos anteriores es atribuible a la ciencia. Ésta, como práctica, surge al lado y bajo el cobijo de la *filosofía empirista*. Cuatro astrónomos preeminentes en la creación de la ciencia: Copérnico, Kepler, Galileo y Newton, físicos además los dos últimos, impulsaron el surgimiento de ésta, al ayudar a abrir el camino a la investigación crítica como medio para avanzar en el conocimiento, lo que obtuvo sus logros más espectaculares en el siglo XVII.

Junto a quienes practicaban la ciencia como método empírico para abordar el conocimiento, surgían los filósofos empiristas, que fundamentaban el nuevo método de obtener conocimiento. Bacon, Hobbes, Locke, Berkeley y Hume instituyen el empirismo como el único camino al conocimiento, al establecer la experiencia empírica como la única posibilidad para conocer la verdad y la inducción como el método lógico que hacía posible esto al usar la inferencia como medio para el logro de conocimientos generales a partir de experiencias particulares. Ellos establecieron el conocimiento científico como un camino seguro a la verdad. Intentaban desarrollar un sistema de inferencia racional que hiciera posible la generalización a partir de experiencias particulares y concretas. Suponían también un carácter acumulativo de la ciencia; para ellos, los hechos son contactos objetivos con el mundo que, una vez establecidos, quedan de manera perenne en el acervo de conocimiento verdadero, siendo la ciencia un proceso de acumulación de hechos. En pocas palabras, con ellos, la concepción de la ciencia se desarrolla como la búsqueda en la experiencia empírica de un camino para una seguridad absoluta que justifique los conocimientos así desarrollados como productos permanentes de un método fehaciente.

Comte dio el siguiente paso en el desarrollo de una concepción de la ciencia. El desarrollo del *positivismo clásico* fue un avance en la concepción de la ciencia empírica y de un sistema metodológico para su ejercicio concreto.^{††} El positivismo considera a la experiencia como fuente de conocimiento, y los hechos generales o leyes son la única fuente de *certidumbre*. Encontramos a pensadores como Mach, Avenarius, Poincaré y Pearson, entre otros, como estructuradores de una filosofía que establecía a la ciencia sobre una base empírica que se pro-

[†] Se ha llegado al estudio de la naturaleza del conocimiento por una variedad de ramas de la filosofía y de las ciencias particulares, denominadas *epistemología*, *filosofía de la ciencia* y *metodología*. El carácter va de lo más general, en la epistemología, a lo más específico, en la metodología.

^{††} Comte fue, además, padre de la sociología, que desarrolla dentro del marco filosófico de su *método positivista* de hacer ciencia.

ponía como guía pragmática para enfrentar la vida. El Universo, incluyendo al hombre, estaría constituido por fenómenos que se conectan causalmente entre sí, conexiones que se podrían descubrir por medio de la inducción, controlada, en la medida de lo posible, por el método experimental. Las leyes y las teorías serían símbolos convencionales que reflejarían el orden en las relaciones dentro de la naturaleza.

Tanto el positivismo clásico como el empirismo mantienen una posición radical acerca del conocimiento. El conocimiento putativo no puede considerarse como verdadero a menos que se le pruebe, y la prueba consiste en ponerlo bajo la hegemonía de la autoridad epistemológica pertinente, en este caso, la experiencia empírica.

En la actualidad, el trabajo de filósofos con enfoques diferentes, aunque con un núcleo central de acuerdo fundamental, culmina el desarrollo de una filosofía de la ciencia empírica. Todos ellos usan la lógica y la lingüística como instrumentos para el desarrollo de una relación entre teoría y realidad, aunque el fundamento de la verdad empírica sigue siendo el criterio epistemológico último. Wittgenstein, Ayer, Carnap, Tarsky y Feigl, desde el positivismo lógico; Russell y Whitehead, desde una combinación de realismo crítico y filosofía analítica y Moore, Wittgenstein y Wisdom, desde la filosofía analítica, abordan la búsqueda de la verdad mediante variantes de un mismo esquema fundamental. Si la inferencia no puede demostrar su validez *absoluta* como método lógico para establecer conocimiento verdadero, es decir, no se le puede probar, el concepto de inducción se sustituye por uno de inducción probabilística. Se fusionan los conceptos de inducción y probabilidad, y es necesario probar el conocimiento en términos de probabilidades.

Este punto de vista de la ciencia prevaleció sin desafío hasta el siglo pasado, pero en la actualidad ha surgido con gran vigor la perspectiva de la ciencia, ya mencionada, llamada no justificacionista, que analiza el proceso de conocimiento científico sin recurrir al de la *justificación empírica* como base para el establecimiento de éste. Autores como Popper, Kuhn, Lakatos, Feyerabend y Weimer han jugado un papel muy importante para dar esa visión alternativa de la ciencia.

La visión de la investigación científica desarrollada por las filosofías empírica y positivista fue relativamente clara. Existen dos tipos de entes: los hechos y las teorías. Los primeros provienen del ingreso sensorial, mientras que las segundas son conjuntos de proposiciones que surgen de los hechos a partir de la inducción. El problema es sencillo: hay que probar las teorías asegurando que sus conceptos tengan una relación unívoca con los hechos establecidos por inducción.

Weimer llama *justificacionismo*[†] al denominador común de todas estas aproximaciones porque encuentra a la “metateoría” como la concepción de que hay una fuente de autoridad que produce una justificación incontrovertible para un método. En esto, afirma que tanto el racionalismo como el empirismo-positivismo parten de una misma posición fundamental; de lo que Dewey llamó *búsqueda de la certeza*. El racionalismo lo hace apelando a la autoridad del intelecto, mientras que el empirismo-positivismo a la del ingreso sensorial.

Popper señala que es precisamente esa búsqueda de una base firme e incontrovertible la fuente de los problemas. Hace un análisis sobre la reflexión de Hume acerca de la inducción y coincide con él en que no es posible que partiendo de la observación de una serie de casos reiterados de una relación determinada se llegue a una conclusión válida acerca de casos aún no observados; es decir, no se justifica desde el punto de vista lógico la inferencia. La solución que ofrece para no caer en un solipsismo estéril es que, si bien no es posible de modo alguno comprobar teorías, sí es factible refutarlas. Su solución para el funcionamiento de la ciencia puede resumirse en la idea de que la ciencia opera sobre la base de conjeturas que se someten a una prueba rigurosa ante la evidencia empírica y ante el análisis de la consistencia lógica. En esta perspectiva no *justificacionista*, la teoría no surge directamente de los datos a partir de

[†] El *no justificacionismo* se inicia propiamente a partir del trabajo seminal de Popper y Kuhn, quienes hacen una crítica devastadora del positivismo lógico desde el interior de éste.

un proceso de inducción, ya que cualquier proposición teórica, desde una simple ley empírica hasta un modelo teórico o una teoría, proviene de *una conjetura*. El origen puede ser, como se señaló anteriormente, cualquier posible fuente: la observación de una o varias regularidades, una especulación teórica, una analogía o algún otro proceso. Lo importante es, como ya se ha dicho, que las conjeturas científicas se ponen a prueba por medio de la crítica lógica y empírica (a diferencia de las conjeturas puramente especulativas en otros ámbitos). No obstante, si los hechos apoyan la teoría, no cabe pensar que la justifican, sólo que hasta ahora no la han refutado.

Justificación frente a confrontación

De acuerdo con Lakatos, un programa de investigación se juzga a partir de su comportamiento comparado con programas rivales. La conciencia de nuevas variables extrañas generalmente se da en torno a la competencia entre teorías rivales; el investigador no se percató de qué variables debe controlar hasta que otra explicación sugiere los aspectos que debe considerar con más cuidado para decidir cuál explicación es la que mejor da cuenta de los hechos. Lakatos asevera que no es tan importante el choque entre teoría y datos como la competencia entre las teorías rivales. La actitud rigurosa no implica la supresión instantánea de una teoría, sino la exploración seria y crítica de sus posibilidades frente a otras opciones de explicación. Tal como señala Weimer, “en la mayoría de los casos en la práctica científica actual, el medio más efectivo de crítica disponible para un investigador es permanecer comprometido con una posición para poder articularla plenamente y explorar sus consecuencias”.

¿De dónde surgen las teorías?

Como se ha visto, las teorías científicas son intentos de explicación de la realidad, confrontadas con los hechos de manera rigurosa, que compiten entre sí para tratar de encontrar la mejor manera de dar cuenta de los hechos. Son sistemas de creencias acerca del mundo, más explícitos, claros y precisos que otros conjuntos de creencias (la religión, el sentido común, las pseudociencias, etc.), y que son sometidos a una rigurosa prueba sistemática. Las teorías pueden tener una génesis muy diversa. Por una parte, se encuentra el conocimiento tácito de muchos aspectos de la realidad, donde el sentido común y el conocimiento personal son una fuente muy importante de hipótesis científicas. En la vida cotidiana observa casualmente muchos hechos que después lleva al laboratorio y examina con más cuidado. Con frecuencia, esas mismas observaciones inspiran los primeros intentos de explicación, que al desarrollarse pueden ser la base de una teoría. Otra fuente común son los accidentes en el proceso de investigación, que llevan a encontrar lo que no se busca y se le ha denominado *serendipia*. En otras ocasiones, las teorías surgen de una observación cuidadosa de los hechos, tal vez experimentales, y el desarrollo de una *inferencia* a partir de ellos, entendiendo que lo observado da claves para la construcción de la explicación. Otro origen frecuente de teorías es la observación de una discrepancia entre algunos hechos y una teoría. Esto puede llevar a una reflexión que dé lugar al desarrollo de una teoría alterna y resuelva el conflicto.

Relaciones entre estadística e investigación

El tema de este capítulo es examinar el papel que tiene la estadística en la investigación científica. La estadística es una rama de las matemáticas que se dedica a entender los fenómenos que tienen un cierto grado de azar. En la ciencia se enfrenta el problema de que los fenómenos son multicausales y existe una diversidad de aspectos de los que sólo se tiene un grado de control relativo. Frente a esta problemática, resulta útil emplear un método que permita lidiar con datos con una cierta dosis de incertidumbre. En realidad la estadística es un instrumento muy valioso para organizar la información científica y para tomar decisiones acerca de ella; sería imposible concebir la investigación científica moderna sin dicha estadística.

La investigación, con muy raras excepciones, se refiere a grupos de datos e incluso a grupos de objetos, plantas, animales o personas. Un investigador en astronomía puede tomar varios registros de la distancia a la que se encuentra la Luna o algún objeto lejano con una técnica específica (por ejemplo, usando un radar) para controlar el error de medida, y luego usar la estadística para decidir si su nueva medición es igual o diferente a la que tuvo usando un método más primitivo. Un psicólogo puede medir la ejecución de una tarea por tres grupos de sujetos en un experimento que difieran en la cantidad de alcohol que han ingerido, para ver el efecto sobre una tarea consistente en colocar palitos en agujeros hechos en una tabla. En este caso, es posible usar la estadística para establecer si hay diferencias entre estos grupos de sujetos.

Error de medida y error experimental

Existen dos conceptos de gran importancia en los que la estadística tiene un papel preponderante: los errores de medida y los experimentales. Ambos son importantes fuentes de problemas para el investigador y poderosas razones para usar la estadística en la investigación.

El error de medida es el que se comete al medir cualquier cosa a pesar del cuidado que se tenga. Por una variedad de razones es posible cometer dos tipos de error: el sistemático, que implica una falla regular en una dirección (por ejemplo, un metro un poco más grande de lo debido) o el aleatorio, que se refiere a inexactitudes de un instrumento al medir con él. El primero produce distorsiones de nuestros datos, que a la vez implican un error en nuestras conclusiones.

Los errores sistemáticos pueden radicar en fallas de calibración de los instrumentos. Los instrumentos de medición deben ser comparados con un estándar, el cual determina que el instrumento efectivamente arroja los valores adecuados a la escala que está usándose. Por ejemplo, el metro tiene como estándar de calibración una varilla de vanadio-iridio, colocada sobre un soporte especial en una cámara con temperatura y ambiente controlados que se encuentra en la Oficina de Pesos y Medidas en París, Francia. Los estándares de calibración de los diversos países se obtienen marcando otra varilla similar en sitios análogos a los de la varilla estándar y conservándolos en condiciones similares. Los instrumentos psicométricos (*tests*) se estandarizan (una forma de calibración) aplicándolos a una gran muestra de la población donde van a usarse (por ejemplo, la ciudad de La Plata o México), y luego se establecen las calificaciones estándar. Es decir, si se usa una prueba de inteligencia en México y se emplean estándares ingleses o argentinos, se estaría produciendo un error sistemático de medida.

Los errores sistemáticos también pueden ser causados por la influencia de alguna variable ajena que afecta el proceso de medición, por ejemplo, la presencia de un campo electromagnético cerca de un instrumento de medición con una aguja de bobina, como lo pudiera ser un sonómetro, o un efecto de una variable no adecuadamente controlada como el sexo o la clase social del encuestador en una prueba de personalidad.

Los errores aleatorios (de azar) son aquellos que se cometen por aspectos accidentales, tales como limitaciones perceptuales o inexactitud al momento de tomar una medida, como pudiera ser el caso de un error al leer una escala de manera distraída. Asimismo, los errores aleatorios también se deben a la influencia accidental, de carácter temporal, de otras variables, como el estado de ánimo de un sujeto al someterse a una prueba, las variaciones accidentales de la corriente eléctrica al medir con equipo electrónico que use la energía de la red eléctrica, o el efecto de la temperatura en el funcionamiento de un equipo.

La estadística permite lidiar con ambos tipos de error. El error sistemático se establece viendo si un grupo de medidas difiere de un estándar bien establecido; por ejemplo, verificar si los metros que se usan en Polonia difieren del metro en la Oficina de Pesos y Medidas en París. Para esto se usan ciertas formas de estadística inferencial.

El error aleatorio se anula a través de la estadística. Es posible comparar medidas con error y estimar el valor casi exacto de cierta medida gracias a la estadística.

Medición y estadística

La estadística se aplica sobre medidas obtenidas de los diversos objetos de estudio en diferentes condiciones. Por ejemplo, si desea verificar si un curso de capacitación para soluciones de problemas mejora la inteligencia de quienes lo cursaron, puede tener un grupo al cual le mide la inteligencia antes y después de llevar el citado curso; es decir, aplica la estadística sobre medidas tomadas de los casos, antes y después de la intervención.

Medir, según Torgerson, es asignar números a una propiedad de acuerdo con una regla.[†] Es decir, medir es una forma particular de observación en la cual se asignan números a las propiedades observadas. Es de notar que esta asignación no es del todo arbitraria, ya que usa una regla de asignación de números a los valores de la propiedad. Algo que es necesario comprender es que debe abstraer la dimensión, lo cual es más difícil si se trata de aspectos no observables directamente, como el nivel del metabolismo basal, el peso de los átomos o la inteligencia.

En la vida cotidiana, sin duda, aparecen numerosas formas de medir, como usar una báscula para pesar. El peso se refiere a estándares, como el gramo, que es el peso de un centímetro cúbico de agua a nivel del mar. La regla para pesar consiste en comparar el peso del objeto de interés con el de un estándar. El número (valor) es asignado de acuerdo con la regla de que el peso del objeto sea igual o un múltiplo del peso del estándar.

Las balanzas son, tal vez, las que permiten ver esto de modo más directo; porque una varilla suspendida horizontalmente por el centro de un postecillo indica que se encuentra equilibrada y, si cuelga en los extremos unos platillos de igual peso, el equilibrio no se altera. En esta balanza pone el objeto que quiere pesar (harina) y se asegura de que tiene un kilogramo colocando en uno de los platillos el estándar de un kilogramo y en el otro la harina. Si el equilibrio se mantiene, entonces tiene el peso deseado. Si no fuese así, tendría que agregar o quitar harina hasta lograr el equilibrio, o puede cambiar o combinar estándares.

Las básculas modernas tienen un plato de un lado, suspendido sobre el brazo de la báscula, y del otro lado, un brazo sobre el cual corre un peso estándar; el efecto del peso varía al correr el estándar sobre el brazo de la palanca.

Otro uso de la estadística en psicología y ciencias afines es el desarrollo de modelos psicométricos. Estos modelos se basan en una teoría que plantea que la respuesta a un problema, pregunta o algo similar, depende de diversas variables. Si selecciona una de esas variables para medirla, también puede escoger varios reactivos que supuestamente la midan, constituyendo una prueba o test con ellos. Usando estadísticas como la correlación y el análisis factorial, es posible ver qué tan efectivamente funciona cada reactivo (pregunta) en relación con la prueba e ir mejorándola de modo que obtenga una medida precisa, y que en efecto mida dicho atributo.

Si bien entrar en detalles en cuanto a la teoría psicométrica está fuera del alcance de este libro, esto da idea de la importancia de aprender estadística para poder después usar la psicometría.

Escalas de medición

Como ya se mencionó, medir es asignar números a propiedades de un objeto de acuerdo con reglas, pero las reglas que es posible usar son de muy diferentes tipos. Al asignar números aproveche las propiedades de los sistemas numéricos. Stevens definió cuatro tipos de escalas, de acuerdo con las propiedades del sistema numérico que se aprovechan por la regla que se usa para la asignación.

El primer tipo, llamado *escala nominal*, emplea nombres para los objetos. Éste sería el caso de usar el 0 para sexo femenino y el 1 para masculino (o viceversa) o usar números diferentes para las personas que escogen distintos tipos de cereal: 1 para los de “Corn flakes”, 2 para “Dulcereal”, etcétera.

[†] Medición numérica.
Medición categórica: nominal y ordinal.

El segundo tipo, denominado *escala ordinal*, asigna los números de acuerdo con la propiedad ordinal del sistema numérico: los valores están ordenados de menos a más, pero no hay una idea de igualdad en las distancias entre los números. La regla de correspondencia permite entonces asignar los valores numéricos a una propiedad del objeto de estudio, de modo que reflejen niveles crecientes de esa propiedad, sin que haya un compromiso de que las distancias en esa propiedad sean iguales. Por ejemplo, en una escala de actitudes puede asignar números: 1, 2, 3,..., etc., a los valores de una actitud. Es decir: “Indique el aprecio que tiene por el presidente de la República: 1. ninguno; 2. poco; 3. regular; 4. mucho”. En esta escala no es fácil decir que la distancia en aprecio entre el que responde 1 y el que responde 2 es igual a la que hay entre 3 y 4, pero sí apreciar que el valor 4 es mayor que el 3 en esa dimensión.

En el tercer tipo, la *escala de intervalo*, no sólo se usa el ordenamiento, sino que establece que las distancias que hay entre número y número son iguales. Por ejemplo, las temperaturas medidas por los termómetros permiten aseverar que la cantidad de incremento de temperatura es igual para distancias iguales en la escala. Por ejemplo, un incremento de 5 °C es igual, ya sea cuando se pasa de 0 a 5 °C o cuando se pasa de 10 a 15 °C.

En el último nivel de escala, la de *razón*, se usan las propiedades anteriores, pero, además, se tiene un cero que refleja la ausencia de la cualidad. Por ejemplo, en el caso anterior de la temperatura visto, las escalas hacen referencia a un cero que es arbitrario y no refleja la ausencia de la propiedad que se mide. El cero, en la escala Celsius, es el punto en que el hielo se derrite (o el agua se congela). En la escala Fahrenheit, la referencia es el alcohol en vez del agua. Ambos son ceros arbitrarios y por eso las escalas generan números negativos, es decir, hay temperaturas bajo cero. Por lo contrario, la escala Kelvin, sí hace referencia a un cero absoluto que implica la ausencia total de movimiento molecular y, por tanto, de temperatura.

Así, los diferentes tipos de escalas usan ciertas propiedades de los sistemas numéricos para generar un tipo de medidas que reflejen ciertas propiedades de la dimensión que se pretende reflejar con esa medida. Las escalas nominales, por ejemplo, sirven para medir cosas que tienen que ver con la pertenencia a grupos u otras formas de clasificar cosas o personas. En este caso, los números sólo sirven como nombres y es indistinto el orden que se use. Aquí sólo se utiliza la propiedad de identidad de los números.

Las escalas ordinales usan la propiedad ordinal, es decir, el hecho de que se siga una secuencia. De este modo, sabe que el 2 es mayor que el 1 o que el 11 es mayor que el 9, sin que eso implique que la distancia entre 9 y 11 tenga que ser mayor que entre 1 y 2, sólo se toma en cuenta el orden.

Las escalas intervalares usan la distancia entre números como algo válido, de manera que la distancia entre 3 y 5 es igual a la distancia entre 7 y 9, pero no hacen referencia a un cero absoluto, de modo que no puede decir que 8 es el doble de 4.

Las escalas de razón usan todas las propiedades de los números: identidad, orden, igualdad de las distancias y referencia al cero.

Limitaciones de las estadísticas por nivel de medida

El uso de la estadística se ve limitado por el tipo de medidas que usa. Por ejemplo, las medidas de razón y de intervalo utilizan los procedimientos más poderosos, llamados *paramétricos*. Existen otros procedimientos que se aplican a los casos de las medidas ordinales y nominales y se les denomina *no paramétricos*. Algunos de ellos utilizan las propiedades de orden como Kolmogorov-Smirnov o la U de Mann-Whitney y otras como la ji cuadrada, que se utilizan para analizar términos de la probabilidad de clases de eventos. Estos procedimientos se verán más adelante con todo detalle; lo importante ahora es percatarse que el tipo de medidas determina el tipo de estadística.

Inferencia estadística y científica

La estadística funciona para hacer inferencias de las distribuciones de las medidas de los fenómenos; partiendo de la suposición de que varias muestras pertenecen a la misma población; y cuando la población a la que pertenecen difiere de ellas, esto se refleja en las muestras.

Para entender mejor esto es preciso decir qué se entiende por población y por muestra. La población es la totalidad de sujetos de una condición que se está observando; es difícil de abarcar y a veces incluye sujetos inaccesibles, como los muertos. Pero incluso los vivos son difíciles de incluir en su totalidad, por ejemplo todos los seres humanos (mayores de 18 años) en el planeta Tierra. Ni siquiera todos los niños menores de 12 años con síndrome de Down. Lo más frecuente es no referirse a sujetos u objetos en sí, sino a alguna dimensión o variable de éstos, como puede ser la estatura, la inteligencia, etcétera.

Una muestra es un subconjunto de la población seleccionado al azar (esto es lo ideal), donde todos los miembros de la población tienen probabilidad de formar parte de ella. Esto en la práctica es muy difícil y costoso, cuando no imposible.

La estadística usa la distribución de probabilidad de los estadísticos muestrales (media, desviación estándar, varianza, etc.). Por ejemplo, la media aritmética, que se verá en el capítulo 2, es una medida global que identifica a un grupo de medidas. Es el valor en el punto central o de equilibrio y que, por tanto, representa al grupo. Las medias aritméticas de muestras aun del mismo tamaño varían entre sí, no siendo exactamente iguales. La frecuencia de estas medias se distribuye de acuerdo con una forma (función de probabilidad) por ejemplo la t de Student). Esta distribución es más alta donde se encuentra la verdadera media aritmética de la población y disminuye a medida que se aleja. Esto implica que cuando toma una muestra aleatoria, la media de ésta tiene una mayor probabilidad de ser igual a la de la población, pero hay una probabilidad pequeña de que difieran.

La inferencia estadística se basa en llegar a una conclusión a partir de una probabilidad de que las medias de dos grupos pertenezcan a la misma población. Si la probabilidad es lo bastante baja se concluye que las muestras no pertenecen a dicha población y que por tanto la razón por la cual los grupos difieren (por ejemplo, una manipulación experimental o la procedencia de grupos con características distintas) genera diferentes poblaciones en esa medida.

Por ejemplo, si supone que el alcohol afecta la comprensión de un texto puede usar una medida del grado de comprensión que tiene un lector. Esta medida puede ser una serie de preguntas acerca del texto (que deberán ser tratadas psicométricamente). Ahora, suponga que forma tres grupos de estudiantes de psicología: al primer grupo le da una bebida sin alcohol, al segundo le da una copa de tequila y al tercero dos copas a cada uno de ellos. Les sugiere leer el texto (cada uno tiene una copia del mismo) a continuación les aplica un cuestionario que mide comprensión de lectura. Si los tres grupos provienen de la misma población (de comprensión de dicho texto) por probabilidad las medias aritméticas serían todas parecidas; pero si el consumo de alcohol tuvo un efecto en la comprensión de la lectura, estas medias diferirán. La diferencia (obtenida mediante un análisis de varianza) determina la probabilidad de que éstos pertenezcan a una población homogénea; y cuando la probabilidad es lo suficientemente baja implica que la hipótesis alterna, esto es que los grupos difieren entre sí, no se rechaza.

Este tipo de inferencia, al igual que la inferencia no estadística que se mencionó anteriormente, se debe tomar con la reserva debida. Por experiencia profesional, tal vez, surgió la hipótesis de que el consumo de alcohol afecta la comprensión de textos. Esta hipótesis se pone a prueba en dicha investigación y deberá seguirse contrastando con diferentes muestras, condiciones, sujetos y lecturas.

Diseño experimental

El diseño experimental es simplemente el plan de investigación. Se trata de un plan para hacer que varíe de la manera más amplia posible la variable, o las variables (variables independientes), de la cual interesa ver su efecto sobre otra u otras variables (variables dependientes) para

establecer relaciones causales o, al menos, funcionales. Los experimentos están diseñados para poner a prueba rigurosa las hipótesis de investigación, las cuales se derivan de los diferentes planteamientos teóricos. De esta manera, varía aquello de lo que quiere observar su efecto sobre algo más.

En las ciencias del comportamiento, lo que interesa en la población son los estímulos, la situación, las variables de la conducta y las relacionadas con los procesos internos.

El desarrollo actual de la tecnología ha hecho posible medir y controlar aspectos muy complejos de los objetos de estudio. Aunque en la época de Galileo, por ejemplo, ya se tenían estas nociones acerca del diseño, no se podían observar muchas cosas porque no se contaba con el desarrollo científico y la consecuente tecnología para observar, medir y controlar muchos de ellos. De esta manera, la ciencia, mediante su propio desarrollo, genera métodos para producir y controlar los diferentes aspectos (variables) que son de su interés, potenciándose a sí misma.

En general un experimento trata de:

- Observar y medir lo más exactamente posible las variables dependientes, es decir, aquellas sobre las cuales quiere ver si hay un efecto causal de las independientes.
- Modificar amplia y sistemáticamente las variables independientes o causales, para ver si éstas afectan el fenómeno tal como se hipotetiza.
- Controlar las variables extrañas, es decir, aquellas que no entran en la hipótesis de investigación, pero que de algún modo podrían influir en los resultados, distorsionándolos. Estas variables son de tres tipos: *a)* la variación de error, debida a una falla en las medidas, la cual se corrige mejorando el proceso de medición; *b)* las que se controlan llevando a las variables a un estado constante que no afecte al fenómeno, y *c)* las intrínsecas al sujeto, que se controlan asignando a los sujetos al azar a cada situación o usándolos como su propio control, es decir, que el mismo sujeto pase por todas las condiciones experimentales.

Existen diseños más o menos estándar, producto del ingenio y dedicación de muchas generaciones de investigadores, lo que hace que generalmente no tenga que inventar nuevos diseños para lograr control y buenos efectos en las investigaciones. Aquí sólo se menciona el hecho, pero el lector tendrá que consultar un texto sobre diseño experimental, para mayores detalles.

Sin embargo, se señalan algunos de los diseños experimentales más comunes que tendrán características diferentes, según el nivel de medición a aplicar, tanto a las variables dependientes como a las independientes. El más simple y básico sería el diseño de dos grupos: experimental y control. Este diseño tiene en un grupo, el experimental, una condición que se supone afecta al proceso, y el segundo grupo, el control, carece de esa condición para dar un parámetro de comparación.

Otro diseño que es más refinado sería el llamado de k grupos. En este caso, en vez de manejar sólo dos condiciones, hay un número k de condiciones, tal que $k > 2$. Por lo general, una de las condiciones muestra la ausencia de la variable, sirviendo de grupo control.

Otro diseño muy popular es el factorial, ahí el sujeto es sometido a condiciones con más de una variable. En este caso, en lugar de un vector (una hilera de condiciones) con k grupos, hay una matriz, es decir, un cuadro, un cubo, etc., donde cada dimensión corresponde a una variable y cada cruce equivale a una cierta combinación de variables. En realidad, el diseño factorial es tan sólo un plan sistemático para producir todas las combinaciones posibles de una serie de factores.

La estadística permitirá obtener resultados en todos los casos, ayuda a discernir si las diferencias encontradas se deben al azar, causadas por variaciones naturales de los grupos, o son debidas al efecto de la variable de interés, la que está manipulando.

Diseño cuasiexperimental

Hay ocasiones en que no es posible controlar adecuadamente algunas variables. Por ejemplo, en un estudio sobre educación se deben tomar los grupos naturales y esto impide la asignación al azar. En estos casos, la estadística viene al rescate, permite tomar en cuenta el posible efecto de esas variables no controladas. Existen dos métodos experimentales: uno es el de análisis de covarianza, que requiere la medición de las variables extrañas potenciales y su introducción al modelo estadístico (en capítulos posteriores se verá cómo se logra esto). El otro método se refiere al uso de series temporales para extraer la varianza y las relaciones de los fenómenos en el tiempo. Éstos son métodos estadísticos que van más allá del alcance de este libro, pero son mencionados para dar una idea general.

Entonces, en los métodos cuasiexperimentales se tienen los mismos elementos que en los experimentales, es decir, maximizar la variación de la variable o variables independientes y controlar las extrañas, pero sólo una parte del control es experimental; la otra es estadística de las variables extrañas. Estos métodos son más adecuados para estudios en condiciones naturales.

Métodos cualitativos

En muchas ocasiones no hay manera de abstraer dimensiones y generar procedimientos para medirlas. En estos casos, los investigadores sólo clasifican los fenómenos y los atributos de éstos y tratan de establecer relaciones causales. Cuando alguien únicamente clasifica, está usando un método cualitativo y el nivel de medición es nominal.

La estadística es igualmente útil en este caso, ya que permite observar las frecuencias de cada clase y establecer relaciones entre éstas.

Estadística e informe científico

El informe es el acto de escribir los resultados de una investigación con el objeto de darlos a conocer, para que se publiquen. Incluye estándares técnicos para su organización y existen manuales de redacción, normativos tanto en su estructura como en su estilo. Uno muy conocido es el *Manual* de la APA (American Psychological Association), que es ya considerado un estándar internacional. Se trata de lograr que el informe sea ordenado, completo y bien organizado para que el lector no sólo se dé cuenta de los resultados, sino de sus implicaciones, del modo como se hicieron las cosas y qué se tendría que hacer para reproducir el estudio.

La estadística desempeña un papel al informar los resultados. Allí deberán mostrarse cuadros y gráficas, así como describir verbalmente lo que se obtuvo (sin interpretar los resultados, lo que viene más adelante en la discusión y las conclusiones). Es importante mostrar los datos y señalar qué diferencias fueron significativas estadísticamente.

Si no quiere leer sólo la información repetida en los libros de texto, sino también las investigaciones originales, debe consultar los artículos de las revistas especializadas. Esto es muy importante si uno quiere mantenerse al día en un campo, la información tarda entre 3 y 10 años en llegar a los libros. Para poder leer estos reportes y entenderlos, debe entender la estadística que usó el autor y qué significa; sólo así podrá seguir sus argumentos.

Gráficas

Las gráficas son un modo muy eficiente de mostrar resultados. Por lo general, los datos se muestran tanto en tablas (donde aparecen los números exactos), como en gráficas, las cuales permiten visualizar mejor la forma de los datos y el patrón que se da en ellos.

1.2

DESCRIPCIÓN DE UN CONJUNTO DE DATOS CON MEDIDAS NUMÉRICAS

Las gráficas ayudan a describir la forma básica de una distribución de datos. Sabemos que “una imagen vale por mil palabras” pero hay limitaciones para usarlas. Supongamos que usted necesita presentar sus datos a un grupo de personas y que el foco del proyector de imágenes se quema o que usted necesita describir sus datos por teléfono; no hay modo de ver las gráficas. Necesita entonces hallar otra forma de llevar la imagen mental de los datos a su audiencia.

Una segunda limitación es que las gráficas son un tanto imprecisas para usar en inferencia estadística. Por ejemplo, supongamos que desea usar un histograma muestral para hacer inferencias acerca de un histograma poblacional. ¿Cómo puede medir las similitudes y diferencias entre los dos histogramas en alguna forma concreta? Si son idénticas, usted podría decir que son las mismas, pero si son diferentes es difícil describir el grado de diferencia.

Una forma de superar estos problemas es usar **medidas numéricas**, que se calculan para una muestra o una población de mediciones. Se usan los datos para calcular un conjunto de **números** que llevarán una buena imagen mental de la distribución de frecuencia. Estas mediciones se llaman **parámetros** cuando se asocian con la población y se denominan **estadísticas** cuando se calculan a partir de mediciones muestrales.

Definición Las mediciones descriptivas numéricas asociadas con una población de mediciones se llaman **parámetros**; las calculadas a partir de mediciones muestrales reciben el nombre de **estadísticas**.

1.3

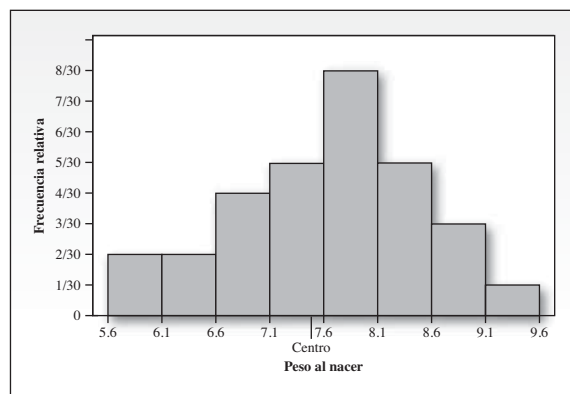
MEDIDAS DE CENTRO

Es posible usar gráficas de puntos, gráficas de tallo y hoja e histogramas para describir la distribución de un conjunto de mediciones en una variable cuantitativa x . El eje horizontal presenta los valores de x , y los datos están “distribuidos” a lo largo de esta recta horizontal. Una de las primeras mediciones numéricas importantes es una **medida de centro**, es decir, una medida a lo largo del eje horizontal que localiza el centro de la distribución.

Los datos de peso al nacer presentados en la tabla 1.1 van de un punto bajo de 5.6 a uno alto de 9.4, con el centro del histograma situado en la cercanía de 7.5 (véase la figura 1.1). Consideremos algunas reglas para localizar el centro de una distribución de mediciones.

FIGURA 1.1

Centro de los datos de peso al nacer



El promedio aritmético de un conjunto de mediciones es una medida de centro muy común y útil. Es frecuente que esta medida se conozca como la **media aritmética**, o simplemente media, de un conjunto de mediciones. Para distinguir entre la **media** para la muestra y la media para la población, usamos el símbolo $\bar{x}$ (x barra) para una media muestral y el símbolo μ (la letra griega mu minúscula) para la media de una población.

Definición La **media aritmética** o **promedio** de un conjunto de n mediciones es igual a la suma de las mediciones dividida entre n .

Como es frecuente que las fórmulas estadísticas comprendan la suma de números, usamos un símbolo para indicar el proceso de sumar. Suponga que hay n mediciones en la variable x y que las llamamos $x_1, x_2, \dots, x_n$. Para sumar las n mediciones, utilizamos esta notación abreviada:

$$\sum_{i=1}^n x_i \quad \text{que significa } x_1 + x_2 + x_3 + \dots + x_n$$

La letra griega mayúscula sigma (Σ) pide sumar los términos que aparezcan a su derecha, empezando con el número debajo de la sigma ($i = 1$) y terminando con el número arriba (n). No obstante, como las sumas típicas en cálculos estadísticos se hacen casi siempre sobre el conjunto total de n mediciones, se puede usar una notación más sencilla:

Σx_i que significa “la suma de todas las mediciones de x ”

Utilizando esta notación, escribimos la fórmula para la media muestral:

NOTACIÓN

Media muestral: $\bar{x} = \frac{\Sigma x_i}{n}$

Media poblacional: μ

EJEMPLO

1.1

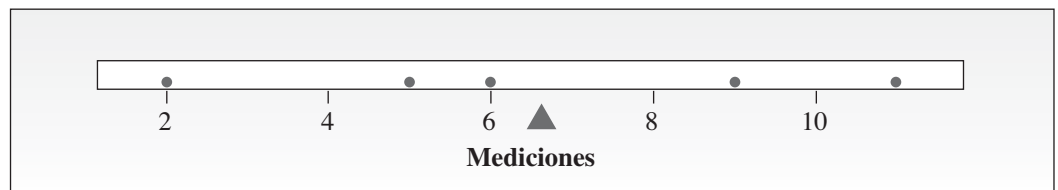
Trace una gráfica de puntos para las $n = 5$ mediciones 2, 9, 11, 5, 6. Encuentre la media muestral y compare su valor con lo que usted pudiera considerar el “centro” de estas observaciones en la gráfica de puntos.

Solución La gráfica de puntos de la figura 1.2 parece estar centrada entre 6 y 8. Para hallar la media muestral, calcule

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{2 + 9 + 11 + 5 + 6}{5} = 6.6$$

FIGURA 1.2

Gráfica de puntos para el ejemplo 1.1



La estadística $\bar{x} = 6.6$ es el punto de equilibrio o fulcro que se muestra en la gráfica de puntos. Éste aparece para marcar el centro de los datos.

TIP MI CONSEJO

Media = punto de equilibrio o fulcro

Recuerde que las muestras son mediciones tomadas de una población más grande que en general es desconocida. Un uso importante de la media muestral $\bar{x}$ es un estimador de la media poblacional desconocida μ . Los datos de peso al nacer en la tabla 1.1 son una muestra de una población más grande de peso al nacer y la distribución se muestra en la figura 1.1. La media de los 30 pesos al nacer es

$$\bar{x} = \frac{\sum x_i}{30} = \frac{227.2}{30} = 7.57$$

ilustrada en la figura 1.1; marca el punto de equilibrio de la distribución. La media de toda la población de pesos de recién nacidos es desconocida, pero si usted tuviera que calcular su valor, su mejor estimación sería 7.57. Aun cuando cambia la media muestral $\bar{x}$ de una muestra a otra, la media poblacional μ sigue igual.

Una segunda medida de tendencia central es la **mediana**, que es el valor de la posición media en el conjunto de mediciones ordenada de menor a mayor.

Definición La **mediana** m de un conjunto de n mediciones es el valor de x que cae en la posición media cuando las mediciones son ordenadas de menor a mayor.

EJEMPLO

1.2

Encuentre la mediana para el conjunto de mediciones 2, 9, 11, 5, 6.

Solución Ordene las $n = 5$ mediciones de menor a mayor:

2 5 6 9 11
 ↑

La observación de en medio, marcada con una flecha, es el centro del conjunto o sea $m = 6$.

EJEMPLO

1.3

Encuentre la mediana para el conjunto de mediciones 2, 9, 11, 5, 6, 27.

Solución Ordene las mediciones de menor a mayor:

2 5 6 9 11 27
 ↑

TIP MI CONSEJO

Casi 50% de las mediciones son más pequeñas, 50% son más grandes que la mediana

Ahora hay dos observaciones “de en medio”, mostradas en la caja. Para hallar la mediana, elija un valor a la mitad entre las dos observaciones de en medio:

$$m = \frac{6 + 9}{2} = 7.5$$

El valor $.5(n + 1)$ indica la **posición de la mediana** del conjunto ordenado de datos. Si la posición de la mediana es un número que termina en el valor **.5**, usted necesita promediar los dos valores adyacentes.

EJEMPLO

1.4

Para las $n = 5$ mediciones ordenadas del ejemplo 1.2, la posición de la mediana es $0.5(n + 1) = .5(6)$, y la mediana es la *tercera observación ordenada*, o $m = 6$. Para las $n = 6$ mediciones ordenadas del ejemplo 1.3, la posición de la mediana es $.5(n + 1) = .5(7) = 3.5$, y la mediana es *el promedio de las 3a. y 4a. observaciones ordenadas*, o $m = (6 + 9)/2 = 7.5$.

TIP MI CONSEJO

Simétrico:
media = mediana
Sesgada a la derecha:
media > mediana
Sesgada a la izquierda:
media < mediana

Aunque tanto la media como la mediana son buenas medidas del centro de una distribución, la mediana es menos sensible a valores extremos o *resultados atípicos*. Por ejemplo, el valor $x = 27$ en el ejemplo 1.3 es mucho mayor que las otras cinco mediciones. La mediana, $m = 7.5$, no es afectada por el resultado atípico, en tanto que el promedio muestral,

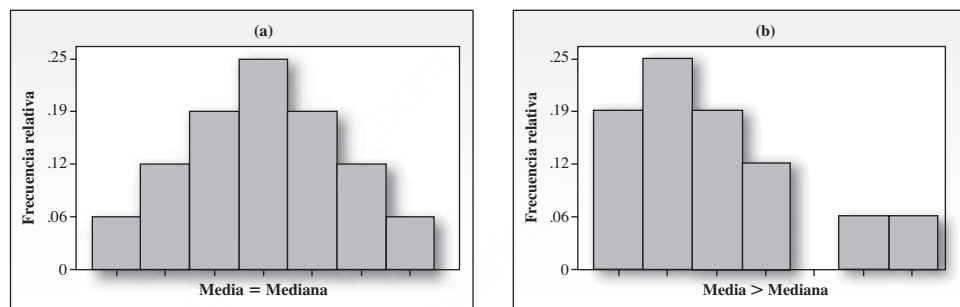
$$\bar{x} = \frac{\sum x_i}{n} = \frac{60}{6} = 10$$

sí es afectado; su valor no es representativo de las cinco observaciones restantes.

Cuando un conjunto de datos tiene valores extremadamente pequeños u observaciones muy grandes, la media muestral se traza hacia la dirección de las mediciones extremas (véase la figura 1.3).

FIGURA 1.3

Distribuciones de frecuencia relativa mostrando el efecto de valores extremos en la media y la mediana

**MI** APLET EN LÍNEA

Cómo afectan los valores extremos a la media y la mediana

Si una distribución está sesgada a la derecha, la media se corre a la derecha; si una distribución está sesgada a la izquierda, la media se corre a la izquierda. La mediana no es afectada por estos valores extremos porque los valores numéricos de las mediciones no se usan en este cálculo. Cuando una distribución es simétrica, la media y la mediana son iguales. Si una distribución está fuertemente sesgada por uno o más valores extremos, debe emplear la mediana en lugar de la media como medida de centro.

Otra forma de localizar el centro de una distribución es buscar el valor de x que se presenta con la frecuencia más alta. Esta medida del centro se denomina **moda**.

Definición La **moda** es la categoría o el valor de x que se presenta con más frecuencia. Cuando las mediciones en una variable continua se han agrupado como histograma de frecuencia o de frecuencia relativa, la clase con el pico más alto o frecuencia se llama **clase modal**, y el punto medio de esa clase se toma como la moda.

La moda por lo general se usa para describir conjuntos grandes de datos, mientras que la media y la mediana se utilizan para conjuntos de datos grandes y pequeños. De los datos de la tabla 1.1a), la moda de la distribución del número de visitas hechas semanalmente a Starbucks para 30 clientes es de 5. La clase modal y el valor de x que se presenta con la más alta frecuencia son iguales, como se muestra en la figura 1.4a).

Para los datos de peso al nacer de la tabla 1.1b), un peso de 7.7 libras al nacer se presenta cuatro veces y, por tanto, la moda para la distribución de pesos al nacer es 7.7. Usando el his-

TIP MI CONSEJO

Recuerde que puede haber varias modas o no haber ninguna (si cada observación se presenta sólo una vez)

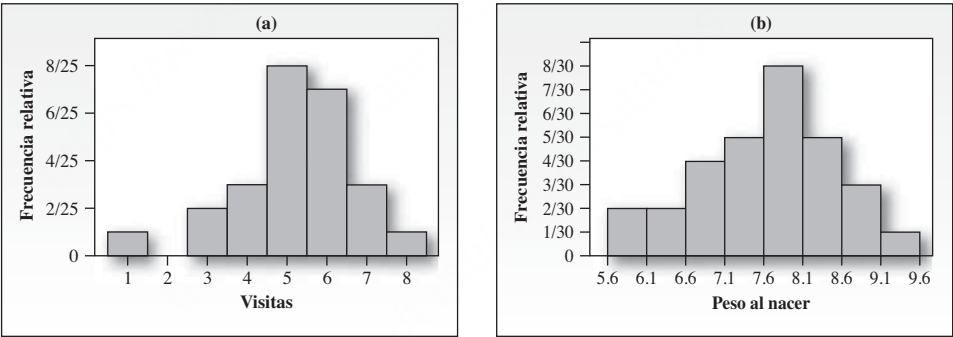
tograma para hallar la clase modal, se encuentra que la clase con el pico más alto es la quinta clase, de 7.6 a 8.1. Nuestra opción para la moda sería el punto medio de esta clase, o sea 7.85. Véase la figura 1.4b).

Es posible que una distribución de mediciones tenga más de una moda. Estas modas aparecerían como “picos locales” en la distribución de frecuencia relativa. Por ejemplo, si fuéramos a tabular la longitud de los peces sacados de un lago durante una temporada, podríamos obtener una *distribución bimodal*, posiblemente reflejando una mezcla de peces jóvenes y viejos en la población. A veces las distribuciones bimodales de tamaños o pesos reflejan una mezcla de mediciones tomadas en machos y hembras. En cualquier caso, un conjunto o distribución de mediciones puede tener más de una moda.

TABLA 1.1

Starbucks y datos de peso al nacer									
a) Datos de Starbucks					b) Datos de peso al nacer (en libras)				
6	7	1	5	6	7.2	7.8	6.8	6.2	8.2
4	6	4	6	8	8.0	8.2	5.6	8.6	7.1
6	5	6	3	4	8.2	7.7	7.5	7.2	7.7
5	5	5	7	6	5.8	6.8	6.8	8.5	7.5
3	5	7	5	5	6.1	7.9	9.4	9.0	7.8
					8.5	9.0	7.7	6.7	7.7

FIGURA 1.4
Histogramas de frecuencia
relativa para datos de
Starbucks y peso al nacer



1.3

EJERCICIOS

TÉCNICAS BÁSICAS

- 1.1** Nos dan $n = 5$ mediciones: 0, 5, 1, 1, 3.

a. Trace una gráfica de puntos para los datos. (SUGERENCIA: Si dos mediciones son iguales, ponga un punto arriba del otro.) Calcule el “centro” aproximado.

b. Encuentre la media, mediana y moda.

c. Localice las tres mediciones de centro en la gráfica de puntos en la parte **a**. Con base en las posiciones relativas de la media y mediana, ¿las mediciones son simétricas o son sesgadas?
- 1.2** Nos dan $n = 8$ mediciones: 3, 2, 5, 6, 4, 4, 3, 5.

a. Encuentre $\bar{x}$.

b. Encuentre m .

c. Con base en los resultados de las partes **a** y **b**, ¿las medidas son simétricas o sesgadas? Trace la gráfica de puntos para confirmar su respuesta.

1.3 Nos dan $n = 10$ mediciones: 3, 5, 4, 6, 10, 5, 6, 9, 2, 8.

a. Calcule $\bar{x}$.

b. Encuentre m .

c. Encuentre la moda.

APLICACIONES

DATOS
EX0104

1.4 Reproductores de DVD Un reproductor de discos de video es un aparato común en la mayoría de los hogares en Estados Unidos. De hecho, casi todas las familias los tienen y muchas tienen más de uno. Una muestra de 25 familias produjo las siguientes mediciones en x , el número de los reproductores de DVD en la casa:

a. La distribución de x , el número de los reproductores de DVD en una familia, ¿es simétrica o sesgada? Explique.

1	0	2	1	1
1	0	2	1	0
0	1	2	3	2
1	1	1	0	1
3	1	0	1	1

- b. Calcule el valor de la moda, el valor de x que se presenta con más frecuencia.
- c. Calcule la media, la mediana y la moda para estas mediciones.
- d. Trace un histograma de frecuencia relativa para el conjunto de datos. Localice la media, mediana y moda a lo largo del eje horizontal. ¿Las respuestas a las partes a y b son correctas?

DATOS
EX0105

1.5 Ingresos en Fortune 500 Diez de las compañías más grandes de Estados Unidos, seleccionadas al azar de *Fortune 500*, aparecen enseguida junto con sus ingresos (en millones de dólares):¹

Compañía	Ingresos (\$)	Compañía	Ingresos (\$)
General Motors	104,589	Target	65,357
IBM	95,758	Morgan Stanley	31,515
Bank of America	150,450	Johnson & Johnson	61,867
Home Depot	66,176	Apple	36,537
Boeing	68,281	Exxon Mobil	284,650

- a. Trace una gráfica de tallo y hoja para los datos. ¿Los datos están sesgados?
- b. Calcule el ingreso medio para estas 10 compañías. Calcule la mediana de los ingresos.
- c. ¿Cuál de las dos medidas de la parte b describe mejor el centro de los datos? Explique.

1.6 Orden de nacimiento y personalidad ¿El orden de nacimiento tiene algún efecto en la personalidad de un individuo? Un informe sobre un estudio, hecho por un investigador del MIT, indica que es más probable que los hijos nacidos después del primogénito pongan a prueba lo establecido, sean más abiertos a nuevas ideas y acepten más un cambio.² De hecho, el número de esta clase de hijos es creciente. Durante los años de la Depresión en el decenio de 1930, las familias promediaban 2.5 hijos (59% después del primogénito), mientras que los padres de familia de los *baby boomers* promediaban de tres a cuatro hijos (68% después del primogénito). ¿Qué quiere decir el autor con un promedio de 2.5 hijos?

DATOS
EX0107

1.7 Atunes Un artículo en *Consumer Reports* da el precio, un promedio estimado de una lata de 6 onzas (180 gramos) o un paquete de 7.06 onzas (210 gramos), para 14 marcas diferentes de atún claro empacado en agua, basado en precios pagados a nivel nacional en supermercados.³

.99	1.92	1.23	.85	.65	.53	1.41
1.12	.63	.67	.69	.60	.60	.66

- a. Encuentre el precio promedio para las 14 marcas diferentes de atún.
- b. Encuentre el precio mediano para las 14 marcas diferentes de atún.
- c. Con base en lo que encuentre en las partes a y b, ¿piensa usted que la distribución de precios está sesgada? Explique.

1.8 Salarios en deportes A medida que los equipos deportivos profesionales se vuelven negocios cada vez más rentables para sus propietarios, los salarios pagados a los jugadores también han aumentado. De hecho, a las superestrellas deportivas se les pagan salarios astronómicos por su talento. Si una compañía de administración deportiva le pide a usted que describa la distribución de salarios de jugadores, en varias categorías diferentes de deportes profesionales, ¿qué medida de centro elegiría? ¿Por qué?

1.9 Tiempo utilizado en una tarea En un experimento psicológico, fue registrado el tiempo que utilizaron 10 individuos en una tarea con una limitación de 5 minutos. Estas mediciones están en segundos:

175	190	250	230	240
200	185	190	225	265

- a. Encuentre el tiempo promedio utilizado en la tarea.
- b. Encuentre la mediana del tiempo utilizado en la tarea.
- c. Si usted está escribiendo un informe para describir estos datos, ¿qué medida de tendencia central usaría? Explique.

DATOS
EX0110

1.10 Wii de Nintendo El “Wii” es una consola de juegos interactiva popular entre muchos jugadores. Su costo puede variar radicalmente, dependiendo de donde se compre. El sitio web www.pricegrabber.com enumeró 14 vendedores en línea con diversos precios, que incluyen gastos de envío e impuestos:⁴

Vendedor	Precio (\$)	Vendedor	Precio (\$)
Buy.com	216.49	Dell	184.86
Sears	222.84	Kmart	222.84
Sam's Club	180.17	EagleDirectUSA	231.04
USA Sales	279.90	Wii4family	262.95
PalaceToys	280.98	QuickShip USA	299.48
Simbaoo7	289.97	BUY-IT-NOW	384.99
jandk425	433.00	SW Evolution	1024.24

- a. ¿Cuál es el precio promedio del Wii para estos 14 vendedores?
- b. ¿Cuál es la mediana del precio del Wii para estos 14 vendedores?
- c. Como consumidor, ¿le interesaría el precio promedio del Wii? ¿Cuáles otras variables serían importantes para usted?

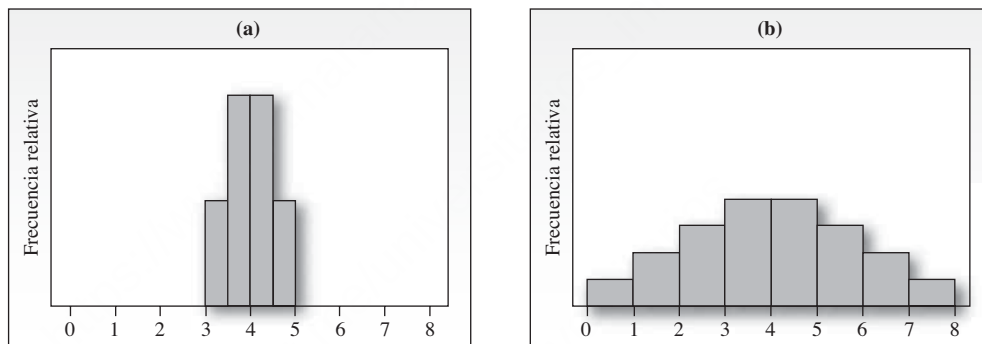
1.4

MEDIDAS DE VARIABILIDAD

Los conjuntos de datos pueden tener el mismo centro pero con aspecto diferente por la forma en que los números se *dispersan* desde el centro. Considere las dos distribuciones que se muestran en la figura 1.5. Ambas distribuciones están centradas en $x = 4$, pero hay una gran diferencia en la forma en que las mediciones se dispersan o *varían*. Las mediciones de la figura 1.5a) varían de 3 a 5; en la figura 1.5b) las mediciones varían de 0 a 8.

FIGURA 1.5

Variabilidad o dispersión de datos



La **variabilidad** o **dispersión** es una muy importante característica de datos. Por ejemplo, si usted fabrica tornillos, la variación extrema en los diámetros de los tornillos causaría un alto porcentaje de productos defectuosos. Por el contrario, si estuviera tratando de discriminar entre contadores buenos y malos, tendría problemas si el examen siempre produjera calificaciones con poca variación, lo cual hace muy difícil la discriminación.

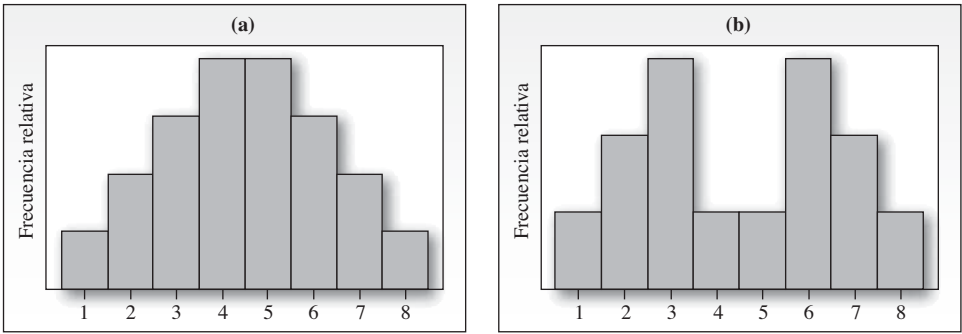
Las **medidas de variabilidad** pueden ayudarle a crear una imagen mental de la dispersión de los datos. Presentaremos algunas de las más importantes. La medida más sencilla de variación es el **rango**.

Definición El **rango**, R , de un conjunto de n mediciones se define como la diferencia entre la medición más grande y la más pequeña.

Por ejemplo, las mediciones 5, 7, 1, 2, 4 varían de 1 a 7. Por tanto, el rango es $7 - 1 = 6$. El rango es fácil de calcular, fácil de interpretar y es una medida adecuada de variación para conjuntos pequeños de datos. Pero, para conjuntos grandes, el rango no es una medida adecuada de variabilidad. Por ejemplo, las dos distribuciones de frecuencia relativa de la figura 1.6 tienen el mismo rango pero muy diferentes formas y variabilidad.

FIGURA 1.6

Distribuciones con igual rango y desigual variabilidad

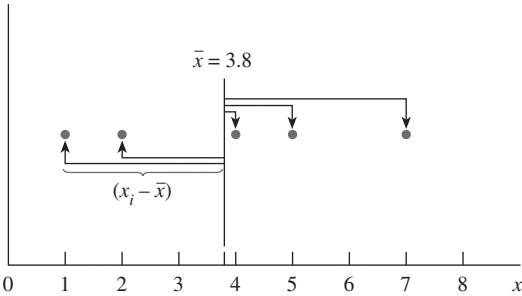


¿Hay una medida de variabilidad que sea más sensible que el rango? Considere, como ejemplo, las mediciones muestrales 5, 7, 1, 2, 4, mostradas como una gráfica de puntos en la figura 1.7. La media de estas cinco mediciones es

$$\bar{x} = \frac{\sum x_i}{n} = \frac{19}{5} = 3.8$$

FIGURA 1.7

Gráfica de puntos que muestra las desviaciones de puntos desde la media



como se indica en la gráfica de puntos. Las distancias horizontales entre cada punto (medición) y la media $\bar{x}$ ayudarán a medir la variabilidad. Si las distancias son grandes, los datos son más dispersos o *variables* que si las distancias son pequeñas. Si x_i es un punto particular (medición), entonces la **desviación** de esa medición desde la media es $(x_i - \bar{x})$. Las mediciones a la derecha de la media producen desviaciones positivas y, las de la izquierda, negativas. Los valores de x y las desviaciones para nuestro ejemplo se detallan en las columnas primera y segunda de la tabla 1.2.

TABLA 1.2

Cálculo de $\sum(x_i - \bar{x})^2$

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
5	1.2	1.44
7	3.2	10.24
1	-2.8	7.84
2	-1.8	3.24
4	.2	.04
19	0.0	22.80

Como las desviaciones en la segunda columna de la tabla contienen información sobre variabilidad, una forma para combinar las cinco desviaciones en una medida numérica es promediarlas. Desafortunadamente, el promedio no funcionará porque algunas de las desviaciones son positivas, algunas son negativas y la suma es siempre cero (a menos que se hayan

introducido errores redondeados en los cálculos). Observe que las desviaciones en la segunda columna de la tabla 1.2 suman cero.

Otra posibilidad sería no hacer caso de los signos de las desviaciones y calcular el promedio de sus valores absolutos.[†] Este método se ha usado como medida de variabilidad en el análisis exploratorio de datos y en el análisis de datos de series de tiempo. Preferimos, no obstante, superar la dificultad causada por los signos de las desviaciones al trabajar con su suma de cuadrados. De la suma de desviaciones cuadradas, se calcula una sola medida llamada **varianza**. Para distinguir entre la varianza de una *muestra* y la varianza de una *población*, usamos el símbolo s^2 para una varianza muestral y σ^2 (letra griega sigma minúscula) para una varianza de población. *La varianza será relativamente grande para datos muy variables y relativamente pequeña para datos menos variables.*

Definición La **varianza de una población** de N mediciones es el promedio de los cuadrados de las desviaciones de las mediciones alrededor de su media μ . La varianza poblacional se denota con σ^2 y está dada por la fórmula

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

La mayoría de las veces, no tendremos todas las mediciones de población disponibles pero necesitaremos calcular la *varianza de una muestra* de n mediciones.

Definición La **varianza de una muestra** de n mediciones es la suma de las desviaciones cuadradas de las mediciones alrededor de la media $\bar{x}$ dividida entre $(n - 1)$. La varianza muestral se denota con s^2 y está dada por la fórmula

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Para el conjunto de $n = 5$ mediciones muestrales presentadas en la tabla 1.2, el cuadrado de la desviación de cada medición se registra en la tercera columna. Sumando, tendremos

$$\sum (x_i - \bar{x})^2 = 22.80$$

y la varianza muestral es

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} = \frac{22.80}{4} = 5.70$$

La varianza se mide en términos del cuadrado de las unidades originales de medición. Si las mediciones originales están en pulgadas, la varianza se expresa en pulgadas cuadradas. Tomando la raíz cuadrada de la varianza, obtenemos la **desviación estándar**, que regresa la medida de variabilidad a las unidades originales de medición.

Definición La desviación estándar de un conjunto de mediciones es igual a la raíz cuadrada positiva de la varianza.

TIP MI CONSEJO

La varianza y la desviación estándar *no* pueden ser números negativos

[†] El valor absoluto de un número es su magnitud, sin atender su signo. Por ejemplo, el valor absoluto de -2 , representado por el símbolo $|2|$, es 2. El valor absoluto de 2, esto es, $|2|$, es 2.

NOTACIÓN n : número de mediciones en la muestra s^2 : varianza muestral $s = \sqrt{s^2}$: desviación muestral estándar N : número de mediciones en la población σ^2 : varianza poblacional $\sigma = \sqrt{\sigma^2}$: desviación poblacional estándar**TIP** MI CONSEJO

Si usted usa calculadora, asegúrese de elegir la tecla correcta para la desviación estándar de la muestra

Para el conjunto de $n = 5$ mediciones muestrales en la tabla 1.2, la varianza muestral es $s^2 = 5.70$, de modo que la desviación estándar de la muestra es $s = \sqrt{s^2} = \sqrt{5.70} = 2.39$. Cuanto más variable sea el conjunto de datos, mayor es el valor de s .

Para el pequeño conjunto de datos que empleamos, el cálculo de la varianza no es demasiado difícil. No obstante, para un conjunto más grande, los cálculos se vuelven muy tediosos. Casi todas las calculadoras científicas contienen programas que calcularán $\bar{x}$ y s o μ y σ , de modo que el trabajo computacional se reducirá para el usuario. La tecla de la muestra o media poblacional suele estar marcada con $\bar{x}$. La tecla de la desviación estándar de la muestra suele estar marcada s , s_x , o σ_{xn-1} , y la tecla de desviación estándar poblacional con σ , σ_x , o σ_{xn} . Al usar cualquier calculadora con estas teclas de función interna, ¡asegúrese de ver qué cálculo es realizado por cada tecla!

Si necesita calcular manualmente s^2 y s , es mucho más fácil usar la fórmula alternativa de cálculo dada a continuación. Esta forma computacional se denomina a veces **método breve para calcular s^2** .

FÓRMULA COMPUTACIONAL PARA CALCULAR s^2

$$s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n - 1}$$

Los símbolos $(\sum x_i)^2$ y $\sum x_i^2$ en la fórmula computacional son métodos breves para indicar la operación aritmética que es necesario efectuar. Usted sabe de la fórmula para la media muestral que $\sum x_i$ es la suma de todas las mediciones. Para hallar $\sum x_i^2$, eleve al cuadrado cada medición individual y luego súmelas.

$\sum x_i^2$ = Suma de cuadrados de las mediciones individuales

$(\sum x_i)^2$ = Cuadrado de la suma de las mediciones individuales

La *desviación estándar de la muestra*, s , es la raíz cuadrada positiva de s^2 .

EJEMPLO

1.5

Calcule la varianza y desviación estándar para las cinco mediciones de la tabla 1.3, que son 5, 7, 1, 2, 4. Use la fórmula computacional para s^2 y compare sus resultados con los obtenidos usando la definición original de s^2 .

TABLA 1.3

Tabla para cálculo simplificado de s^2 y s

x_i	x_i^2
5	25
7	49
1	1
2	4
4	16
19	95

TIP MI CONSEJO

¡No redondee resultados parciales conforme avanza!

Solución Las entradas en la tabla 1.2 son las mediciones individuales, x_i , y sus cuadrados, x_i^2 , junto con sus sumas. Usando la fórmula computacional para s^2 , tenemos

$$s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n - 1} = \frac{95 - \frac{(19)^2}{5}}{4} = \frac{22.80}{4} = 5.70$$

$$\text{y } s = \sqrt{s^2} = \sqrt{5.70} = 2.39, \text{ como antes.}$$

MI APLET EN LÍNEA

¿Por qué dividir entre $n - 1$?

Usted se puede preguntar por qué es necesario dividir entre $(n - 1)$ en lugar de n cuando calcula la varianza muestral. Así como empleamos la media muestral $\bar{x}$ para estimar la media poblacional μ , se usa la varianza muestral s^2 para calcular la varianza poblacional σ^2 . Resulta que la varianza muestral s^2 con $(n - 1)$ en el denominador proporciona estimaciones mejores de σ^2 de lo que daría un estimador calculado con n en el denominador. **Por esta razón, siempre dividimos entre $(n - 1)$ al calcular la varianza muestral s^2 y la desviación estándar de la muestra s .**

En este punto, usted ha aprendido a calcular la varianza y desviación estándar de un conjunto de mediciones. Recuerde estos puntos:

- El valor de s es siempre mayor que o igual a cero.
- Cuanto mayor sea el valor de s^2 o de s , mayor es la variabilidad del conjunto de datos.
- Si s^2 o s es igual a cero, todas las mediciones deben tener el mismo valor.
- Para medir la variabilidad en las mismas unidades que las observaciones originales, calculamos la desviación estándar $s = \sqrt{s^2}$.

Esta información le permite comparar varios conjuntos de datos respecto a sus ubicaciones y su variabilidad. ¿Cómo puede usar estas mediciones para decir algo más específico acerca de un solo conjunto de datos? El teorema y la regla que se presentan en la siguiente sección ayudarán a contestar esta pregunta.

1.4 EJERCICIOS

TÉCNICAS BÁSICAS

1.11 Nos dan $n = 5$ mediciones: 2, 1, 1, 3, 5.

a. Calcule la media muestral, $\bar{x}$.

b. Calcule la varianza muestral, s^2 , usando la fórmula dada por la definición.

c. Encuentre la desviación estándar de la muestra, s .

d. Encuentre s^2 y s usando la fórmula computacional.

Compare los resultados con los hallados en las partes **b** y **c**.

1.12 Consulte el ejercicio 1.11.

a. Use el método de entrada de datos en su calculadora científica para introducir las cinco mediciones. Recuerde las memorias apropiadas para hallar la media muestral y la desviación estándar.

b. Verifique que la calculadora dé los mismos valores para $\bar{x}$ y s como en el ejercicio 1.11, partes **a** y **c**.

1.13 Nos dan $n = 8$ mediciones: 4, 1, 3, 1, 3, 1, 2, 2.

a. Encuentre el rango.

b. Calcule $\bar{x}$.

c. Calcule s^2 y s usando la fórmula computacional.

d. Use el método de entrada de datos en su calculadora para hallar $\bar{x}$, s y s^2 . Verifique que sus respuestas sean iguales a las de las partes **b** y **c**.

1.14 Nos dan $n = 8$ mediciones: 3, 1, 5, 6, 4, 4, 3, 5.

a. Calcule el rango.

b. Calcule la media muestral.

c. Calcule la varianza muestral y la desviación estándar.

d. Compare el rango y la desviación estándar. ¿Cuántas desviaciones estándar es aproximadamente el rango?

porcentaje de óxido de hierro en cada una de las cinco muestras recolectadas en el sitio de Island Thorns fue:

1.28, 2.39, 1.50, 1.88, 1.51

a. Calcule el rango.

b. Calcule la varianza muestral y la desviación estándar usando la fórmula computacional.

c. Compare el rango y la desviación estándar. ¿Cuántas desviaciones estándar es aproximadamente el rango?

DATOS
EX0116

1.16 Estados de cuenta por consumo eléctrico en el sur de California

Los estados de cuenta mensuales por consumo eléctrico para una familia en Riverside, California, se registraron durante 12 meses consecutivos empezando en enero de 2010:

Mes	Cantidad (\$)	Mes	Cantidad (\$)
Enero	288.02	Julio	311.20
Febrero	230.60	Agosto	370.23
Marzo	216.85	Septiembre	368.57
Abril	243.74	Octubre	301.79
Mayo	236.96	Noviembre	271.99
Junio	288.57	Diciembre	298.12

a. Calcule el rango del pago de electricidad para el año 2010.

b. Calcule el promedio mensual de pago de electricidad en 2010.

c. Calcule la desviación estándar para el pago de electricidad para el mismo año.

APLICACIONES

1.15 Un hallazgo arqueológico Un artículo en *Archaeometry* contenía un análisis de 26 muestras de cerámica romano-británica hallada en cuatro hornos diferentes en el Reino Unido.⁵ Las muestras fueron analizadas para determinar su composición química. El

SOBRE LA SIGNIFICACIÓN PRÁCTICA DE LA DESVIACIÓN ESTÁNDAR

1.5

A continuación presentamos un útil teorema ideado por el matemático ruso Chebyshev. La demostración del teorema no es difícil, pero estamos más interesados en su aplicación que en demostrarlo.

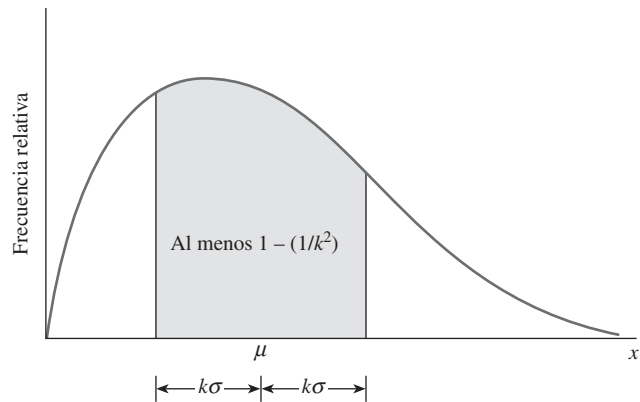
Teorema de Chebyshev

Dado un número k mayor que o igual a 1 y un conjunto de n mediciones, al menos $[1 - (1/k^2)]$ de las mediciones estarán dentro de k desviaciones estándar de su media.

El teorema de Chebyshev aplica a *cualquier conjunto de mediciones* y se usa para describir ya sea una muestra o una población. Usaremos la notación apropiada para poblaciones, pero usted debe percatarse de que con la misma facilidad podríamos usar la media y la desviación estándar para la muestra.

La idea comprendida en el teorema de Chebyshev se ilustra en la figura 1.8. Se construye un intervalo al medir una distancia $k\sigma$ a cualquier lado de la media μ . El número k puede ser cualquiera mientras sea mayor que o igual a 1. Entonces el teorema de Chebyshev expresa que al menos $[1 - (1/k^2)]$ del número total n de mediciones está en el intervalo construido.

FIGURA 1.8
Ilustración del teorema de Chebyshev



En la tabla 1.4 elegimos unos cuantos valores numéricos para k y calculamos $[1 - (1/k^2)]$.

TABLA 1.4 **Valores ilustrativos de $[1 - (1/k^2)]$**

k	$1 - (1/k^2)$
1	$1 - 1 = 0$
2	$1 - 1/4 = 3/4$
3	$1 - 1/9 = 8/9$

De los cálculos de la tabla 1.4, el teorema establece que:

- Al menos ninguna de las mediciones está en el intervalo $\mu - \sigma$ a $\mu + \sigma$.
- Al menos 3/4 de las mediciones están en el intervalo $\mu - 2\sigma$ a $\mu + 2\sigma$.
- Al menos 8/9 de las mediciones están en el intervalo $\mu - 3\sigma$ a $\mu + 3\sigma$.

Aun cuando el primer enunciado no es útil en absoluto, los otros dos valores de k dan valiosa información acerca de la proporción de mediciones que caen en ciertos intervalos. Los valores $k = 2$ y $k = 3$ no son los únicos valores de k a usarse; por ejemplo, la proporción de mediciones que caen dentro de $k = 2.5$ desviaciones estándar de la media es al menos $1 - [1/(2.5)^2] = .84$.

EJEMPLO 1.6

La media y la varianza de una muestra de $n = 25$ mediciones son 75 y 100, respectivamente. Use el teorema de Chebyshev para describir la distribución de mediciones.

Solución Nos dan $\bar{x} = 75$ y $s^2 = 100$. La desviación estándar es $s = \sqrt{100} = 10$. La distribución de mediciones está centrada alrededor de $\bar{x} = 75$, y el teorema de Chebyshev establece que:

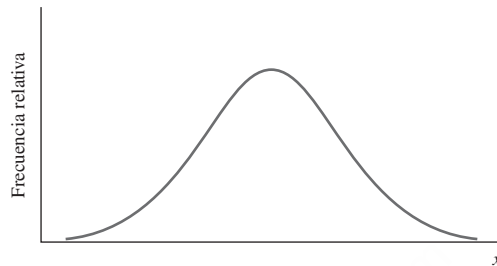
- Al menos 3/4 de las 25 mediciones están en el intervalo $\bar{x} \pm 2s = 75 \pm 2(10)$, esto es, 55 a 95.
- Al menos 8/9 de las mediciones están en el intervalo $\bar{x} \pm 3s = 75 \pm 3(10)$, esto es, 45 a 105.

Como el teorema de Chebyshev se aplica a *cualquier* distribución, es muy conservador. Ésta es la razón por la que hacemos hincapié en “al menos $1 - (1/k^2)$ ” en este teorema.

Otra regla para describir la variabilidad de un conjunto de datos no funciona para *todos* los conjuntos de datos, pero funciona muy bien para datos que “se apilan” en la conocida forma de montículo de la figura 1.9. Cuanto más cerca se encuentre la distribución a la curva en forma de montículo de la figura 1.9, más precisa será la regla. Como la distribución de datos en forma de montículo se presenta con frecuencia en la naturaleza, la regla se usa en numerosas ocasiones en aplicaciones prácticas. Por esta razón, se denomina **regla empírica**.

FIGURA 1.9

Distribución en forma de montículo



Regla empírica Dada una distribución de mediciones que tiene forma aproximada de montículo:

El intervalo $(\mu \pm \sigma)$ contiene aproximadamente 68% de las mediciones.

El intervalo $(\mu \pm 2\sigma)$ contiene aproximadamente 95% de las mediciones.

El intervalo $(\mu \pm 3\sigma)$ contiene aproximadamente 99.7% de las mediciones.

TIP MI CONSEJO

Recuerde estos tres números:

68—95—99.7

La distribución en forma de montículo o campana que se muestra en la figura 1.9 se conoce comúnmente como **distribución normal**.

EJEMPLO

1.7

En un estudio de tiempo efectuado en una planta manufacturera, el tiempo para completar una operación especificada se midió para cada uno de los $n = 40$ trabajadores. Se encontró que la media y la desviación estándar son 12.8 y 1.7, respectivamente. Describa los datos muestrales usando la regla empírica.

Solución Para describir los datos, calcule estos intervalos:

$$(\bar{x} \pm s) = 12.8 \pm 1.7 \quad \text{o} \quad 11.1 \text{ a } 14.5$$

$$(\bar{x} \pm 2s) = 12.8 \pm 2(1.7) \quad \text{o} \quad 9.4 \text{ a } 16.2$$

$$(\bar{x} \pm 3s) = 12.8 \pm 3(1.7) \quad \text{o} \quad 7.7 \text{ a } 17.9$$

De acuerdo con la regla empírica, usted espera que aproximadamente 68% de las mediciones caigan en el intervalo de 11.1 a 14.5, aproximadamente 95% caigan en el intervalo de 9.4 a 16.2, y aproximadamente 99.7% caigan en el intervalo de 7.7 a 17.9.

Si hay duda de que la distribución de mediciones tenga forma de montículo o si usted desea ser conservador por alguna razón, puede aplicar el teorema de Chebyshev y estar absolutamente seguro de sus afirmaciones. El teorema de Chebyshev dice que al menos $3/4$ de las mediciones caen en el intervalo de 9.4 a 16.2 y al menos $8/9$ en el intervalo de 7.7 a 17.9.

EJEMPLO

1.8

Los futuros profesores son capacitados para desarrollar planes de lecciones, en la suposición de que el plan escrito les ayudará a trabajar de manera satisfactoria en el salón de clases. En un estudio para evaluar la relación entre planes de lección escritos y su implementación en el salón de clases, se calificaron 25 planes de lección en una escala de 0 a 34 de acuerdo con una Lista de verificación de Plan de lección. Las 25 calificaciones se muestran en la tabla 1.5. Use el teorema de Chebyshev y la regla empírica (si aplica) para describir la distribución de estas calificaciones de evaluación.

TABLA 1.5

Calificaciones para evaluación de Plan de lección

26.1	26.0	14.5	29.3	19.7
22.1	21.2	26.6	31.9	25.0
15.9	20.8	20.2	17.8	13.3
25.6	26.5	15.7	22.1	13.8
29.0	21.3	23.5	22.1	10.2

Solución Use su calculadora o las fórmulas computacionales para verificar que $\bar{x} = 21.6$ y $s = 5.5$. Los intervalos apropiados están calculados y aparecen en la tabla 1.6. También hemos consultado las 25 mediciones originales y contado el número real de mediciones que caen en cada uno de estos intervalos. Estas frecuencias y frecuencias relativas se muestran en la tabla 1.6.

TABLA 1.6

Intervalos $\bar{x} \pm ks$ para los datos de la tabla 1.5

k	Intervalo $\bar{x} \pm ks$	Frecuencia en intervalo	Frecuencia relativa
1	16.1–27.1	16	.64
2	10.6–32.6	24	.96
3	5.1–38.1	25	1.00

TIP MI CONSEJO

Regla empírica $\Leftrightarrow$ datos en forma de montículo
Chebyshev $\Leftrightarrow$ datos en cualquier forma

¿Es aplicable el teorema de Chebyshev? Sí, porque es posible usarlo para cualquier conjunto de datos. De acuerdo con el teorema de Chebyshev,

- al menos 3/4 de las mediciones caerán entre 10.6 y 32.6.
- al menos 8/9 de las mediciones caerán entre 5.1 y 38.1.

Observe que en la tabla 1.6 el teorema de Chebyshev es verdadero para estos datos. De hecho, las proporciones de mediciones que caen en los intervalos especificados exceden el límite inferior dado por este teorema.

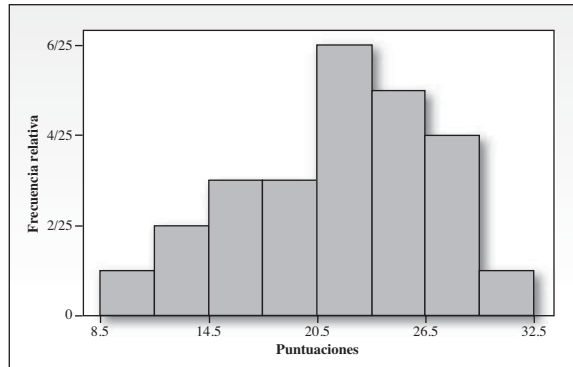
¿Es aplicable la regla empírica? Usted puede comprobarlo por sí mismo si traza una gráfica, ya sea una gráfica de tallo y hoja o un histograma. El histograma de frecuencia relativa de la figura 1.10 muestra que la distribución es *relativamente* en forma de montículo, de modo que la regla empírica debe funcionar *relativamente bien*. Esto es,

- aproximadamente 68% de las mediciones caerán entre 16.1 y 27.1.
- aproximadamente 95% de las mediciones caerán entre 10.6 y 32.6.
- aproximadamente 99.7% de las mediciones caerán entre 5.1 y 38.1.

Las frecuencias relativas de la tabla 1.6 se aproximan mucho a las especificadas por la regla empírica.

FIGURA 1.10

Histograma de frecuencia relativa para el ejemplo 1.8



USO DEL TEOREMA DE CHEBYSHEV Y LA REGLA EMPÍRICA

El teorema de Chebyshev se demuestra matemáticamente. Se aplica a cualquier conjunto de mediciones, muestra o población, grande o pequeño, en forma de montículo o sesgado.

El teorema de Chebyshev da un límite inferior a la fracción de mediciones a encontrar en un intervalo construido como $\bar{x} \pm ks$. ¡Al menos $1 - (1/k^2)$ de las mediciones caerán en este intervalo, y probablemente más!

La regla empírica es una “regla práctica” que se utiliza como herramienta descriptiva cuando los datos tienden a ser de forma más o menos de montículo (los datos tienden a apilarse cerca del centro de la distribución).

Cuando use estas dos herramientas para describir un conjunto de mediciones, el teorema de Chebyshev siempre se satisface pero es una estimación muy conservadora de la fracción de mediciones que caen en un intervalo particular. Si es apropiado usar la regla empírica (datos en forma de montículo), esta regla dará una estimación más precisa de la fracción de mediciones que caen en el intervalo.

1.6

UNA MEDICIÓN DEL CÁLCULO DE s

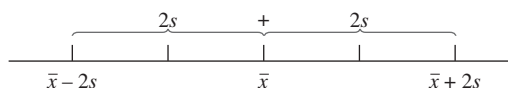
El teorema de Chebyshev y la regla empírica se usan para detectar errores burdos en el cálculo de s . En términos generales, estas dos herramientas le indican que *la mayoría de las veces* las mediciones caen dentro de *dos* desviaciones estándar de su media. Este intervalo está marcado en la figura 1.11, e implica que el rango total de mediciones, de la más pequeña a la más grande, debe estar en algún punto alrededor de cuatro desviaciones estándar. Esto es, desde luego, una aproximación muy burda pero es muy útil para localizar errores grandes en el cálculo de s . Si el rango, R , es de alrededor de cuatro desviaciones estándar, o $4s$, escriba

$$R \approx 4s \quad \text{o bien} \quad s \approx \frac{R}{4}$$

El valor calculado de s usando la fórmula de atajo debe ser de alrededor del mismo orden que la aproximación.

FIGURA 1.11

Aproximación de rango para s



EJEMPLO

1.9

Use la aproximación de rango para comprobar el cálculo de s para la tabla 1.3.

Solución El rango de las cinco mediciones, 5, 7, 1, 2, 4, es

$$R = 7 - 1 = 6$$

Entonces

$$s \approx \frac{R}{4} = \frac{6}{4} = 1.5$$

Esto es del mismo orden que el valor calculado $s = 2.39$.

TIP**MI CONSEJO**

$s \approx R/4$ sólo da un valor aproximado para s

La aproximación de rango *no* tiene la finalidad de dar un valor preciso para s . Más bien, su propósito es detectar errores burdos de cálculo; por ejemplo, no dividir la suma de cuadrados de desviaciones entre $(n - 1)$ o no tomar la raíz cuadrada de s^2 . Si usted comete uno de estos errores, su respuesta será muchas veces más grande que la aproximación de rango de s .

EJEMPLO

1.10

Use la aproximación de rango para determinar un valor aproximado para la desviación estándar para los datos de la tabla 1.5.

Solución El rango $R = 31.9 - 10.2 = 21.7$. Entonces

$$s \approx \frac{R}{4} = \frac{21.7}{4} = 5.4$$

Como el valor exacto de s es 5.5 para los datos de la tabla 1.5, la aproximación es muy cercana.

El rango para una muestra de n mediciones dependerá del tamaño muestral, n . Para valores más grandes de n , se espera un rango más grande de valores x . El rango para muestras grandes (por ejemplo $n = 50$ o más observaciones) puede ser hasta de $6s$, mientras que el rango para muestras pequeñas (por ejemplo $n = 5$ o menos) puede ser de sólo $2.5s$ o menor.

La aproximación de rango para s será mejor si se sabe que la muestra se toma de una distribución de datos en forma de montículo. Entonces, la s calculada no debe diferir de manera importante a partir del rango dividido entre la razón apropiada dada en la tabla 1.7.

TABLA 1.7

Divisor para la aproximación de rango de s

Número de mediciones	Razón esperada de rango para s
5	2.5
10	3
25	4

TÉCNICAS BÁSICAS

1.17 Un conjunto de $n = 10$ mediciones consta de los valores 5, 2, 3, 6, 1, 2, 4, 5, 1, 3.

- Use la aproximación de rango para estimar el valor de s para este conjunto. (SUGERENCIA: Use la tabla del final de la sección 1.5.)
- Use su calculadora para hallar el valor real de s . ¿El valor real es cercano a su estimación en la parte **a**?
- Trace una gráfica de puntos de este conjunto de datos. ¿Los datos tienen forma de montículo?
- ¿Usaría el teorema de Chebyshev para describir este conjunto de datos? ¿Por qué sí o por qué no?
- ¿Utilizaría la regla empírica para describir este conjunto de datos? ¿Por qué sí o por qué no?

1.18 Supongamos que usted desea crear una imagen mental del histograma de frecuencia relativa para un conjunto de datos grande formado por 1000 observaciones y sabe que la media y la desviación estándar del conjunto de datos son 36 y 3, respectivamente.

- Si está más o menos seguro que la distribución de frecuencia relativa de los datos tiene forma de montículo, ¿cómo representaría la distribución de frecuencia relativa? (SUGERENCIA: Use la regla empírica.)
- Si no tiene usted información previa respecto a la forma de la distribución de frecuencia relativa, ¿qué diría acerca del histograma de frecuencia relativa? (SUGERENCIA: Construya intervalos $\bar{x} \pm ks$ para varias opciones de k .)

1.19 Una distribución de mediciones tiene relativamente la forma de un montículo con media de 50 y desviación estándar de 10.

- ¿Qué proporción de las mediciones caerá entre 40 y 60?
- ¿Qué proporción de las mediciones caerá entre 30 y 70?
- ¿Qué proporción de las mediciones caerá entre 30 y 60?
- Si se elige al azar una medición de esta distribución, ¿cuál es la probabilidad de que sea mayor que 60?

1.20 Un conjunto de datos tiene una media de 75 y una desviación estándar de 5. Usted no sabe nada más acerca del tamaño del conjunto de datos o de la forma de la distribución de datos.

- ¿Qué diría acerca de la proporción de mediciones que caen entre 60 y 90?
- ¿Qué diría acerca de la proporción de mediciones que caen entre 65 y 85?

- ¿Qué diría acerca de la proporción de mediciones que sean menores que 65?

APLICACIONES

1.21 Emergencias de automovilistas El tiempo requerido para que el conductor de un automóvil responda a una situación particular de emergencia se registró para $n = 10$ conductores. Los tiempos (en segundos) fueron .5, .8, 1.1, .7, .6, .9, .7, .8, .7, .8.

- Busque en los datos y use el procedimiento de la sección 1.5 para hallar un valor aproximado para s . Use este valor para verificar sus cálculos en la parte **b**.
- Calcule la media muestral y la desviación estándar s . Compare con la parte **a**.

DATOS
EX0122

1.22 Empacar carne para hamburguesas Los datos que aparecen enseguida son los pesos (en libras) de 27 paquetes de carne molida de res en un exhibidor de supermercado:

1.08	.99	.97	1.18	1.41	1.28	.83
1.06	1.14	1.38	.75	.96	1.08	.87
.89	.89	.96	1.12	1.12	.93	1.24
.89	.98	1.14	.92	1.18	1.17	

- Construya una gráfica de tallo y hoja o un histograma de frecuencia relativa para mostrar la distribución de pesos. ¿La distribución es relativamente de forma de montículo?
- Encuentre la media y desviación estándar del conjunto de datos.
- Encuentre el porcentaje de mediciones en los intervalos $\bar{x} \pm s$, $\bar{x} \pm 2s$ y $\bar{x} \pm 3s$.
- ¿Cómo se comparan los porcentajes obtenidos en la parte **c** con los datos proporcionados por la regla empírica? Explique.
- ¿Cuántos de los paquetes pesan exactamente una libra? ¿Hay alguna explicación para esto?

1.23 Ritmo respiratorio ¿Es normal su ritmo respiratorio? En realidad, no hay un ritmo estándar de respiración para seres humanos. Puede variar desde sólo cuatro respiraciones por minuto hasta 70 o 75 para una persona que realice un ejercicio agotador. Suponga que los ritmos respiratorios en reposo para estudiantes universitarios tienen una distribución de frecuencia relativa en forma de montículo, con una media igual a 12 y una desviación estándar de 2.3 respiraciones por minuto.

¿Qué fracción de todos los estudiantes tendría ritmos respiratorios en los siguientes intervalos?

- 9.7 a 14.3 respiraciones por minuto.
- 7.4 a 16.6 respiraciones por minuto.
- Más de 18.9 o menos de 5.1 respiraciones por minuto.

DATOS
EX0124 **1.24 Muestras de mineral** Una geóloga recolectó 20 muestras diferentes de mineral, todas del mismo peso, y las dividió al azar en dos grupos. Ella midió el contenido de titanio (Ti) de las muestras usando dos métodos diferentes.

Método 1	Método 2
.011 .013 .013 .015 .014	.011 .016 .013 .012 .015
.013 .010 .013 .011 .012	.012 .017 .013 .014 .015

- Construya gráficas de tallo y hoja para los dos conjuntos de datos. Compare visualmente sus centros y sus rangos.
- Calcule las medias muestrales y desviaciones estándar para los dos conjuntos. ¿Los valores calculados confirman sus conclusiones visuales de la parte a?

DATOS
EX0125 **1.25 Números de credencial** Se pidió a un grupo de 70 estudiantes que registrara el último dígito de su número de credencial de la universidad.

1 6 9 1 5 9 0 2 8 4
0 7 3 4 2 3 5 8 4 2
3 2 0 0 2 1 2 7 7 4
0 0 9 9 5 3 8 4 7 4
6 6 9 0 2 6 2 9 5 8
5 1 7 7 7 8 7 5 1 8
3 4 1 9 3 8 6 6 6 6

- Trace un histograma de frecuencia relativa usando los valores 0 a 9 como los puntos medios de clase. ¿Cuál es la forma de la distribución? Con base en la forma, ¿cuál sería su mejor estimación para la media del conjunto de datos?
- Use la aproximación de rango para calcular el valor de s para este conjunto.
- Use su calculadora para hallar los valores reales de $\bar{x}$ y s . Compárelas con sus estimaciones en las partes a y b.

1.26 Números de credencial, continúa Consulte el conjunto de datos del ejercicio 1.25.

- Encuentre el porcentaje de mediciones en los intervalos $\bar{x} \pm s$, $\bar{x} \pm 2s$ y $\bar{x} \pm 3s$.
- ¿Cómo se comparan los porcentajes obtenidos en la parte a con los dados por la regla empírica? ¿Deben ser aproximadamente iguales? Explique.

1.27 Tiempos de supervivencia Un grupo de animales experimentales es infectado con una forma particular de bacterias, encontrándose que su tiempo de supervivencia es de 32 días con una desviación estándar de 36 días.

- Visualice la distribución de tiempos de supervivencia. ¿Piensa usted que la distribución es de forma relativamente de montículo, sesgada a la derecha o sesgada a la izquierda? Explique.
- ¿Dentro de qué límites esperaría usted que se encuentren al menos $3/4$ de las mediciones?

1.28 Tiempos de supervivencia, continúa Consulte el ejercicio 1.27. Utilice la regla empírica para ver por qué la distribución de tiempos de supervivencia no podría tener forma de montículo.

- Encuentre el valor de x que esté exactamente una desviación estándar debajo de la media.
- Si la distribución tiene en realidad forma de montículo, ¿aproximadamente qué porcentaje de las mediciones debe ser menor que el valor de x encontrado en la parte a?
- Como la variable que se mide es tiempo, ¿es posible hallar algunas mediciones que estén más de una desviación estándar debajo de la media?
- Use sus respuestas a las partes b y c para explicar por qué la distribución de datos no puede tener forma de montículo.

DATOS
EX0129 **1.29 Terreno maderero** Para calcular la cantidad de madera en un terreno maderero, un propietario determinó contar el número de árboles con diámetros mayores a 12 pulgadas en cuadrados de 50×50 pies seleccionados al azar. Se eligieron 70 de estos cuadrados y se contaron los árboles seleccionados de cada extensión. Los datos son:

7 8 7 10 4 8 6 8 9 10
9 6 4 9 10 9 8 8 7 9
3 9 5 9 9 8 7 5 8 8
10 2 7 4 8 5 10 7 7 7
9 6 8 8 8 7 8 9 6 8
6 11 9 11 7 7 11 7 9 13
10 8 8 5 9 9 8 5 9 8

- Construya un histograma de frecuencia relativa para describir los datos.
- Calcule la media muestral como estimación de μ , el número medio de árboles para todos los cuadrados de 50×50 pies del terreno.
- Calcule s para los datos. Construya los intervalos $\bar{x} \pm s$, $\bar{x} \pm 2s$ y $\bar{x} \pm 3s$. Calcule el porcentaje de cuadrados que caen en cada uno de los tres intervalos y compare con los correspondientes porcentajes dados por la regla empírica y el teorema de Chebyshev.

1.30 Atunes, otra vez Consulte el ejercicio 1.7 y el conjunto de datos EX0107. A continuación aparecen los precios de una lata de 6 onzas, o una bolsa de 7.06 onzas, para 14 marcas diferentes de atún claro elaborado en agua basados en precios pagados nacionalmente en supermercados.³

.99 1.92 1.23 .85 .65 .53 1.41
1.12 .63 .67 .69 .60 .60 .66

a. Use la aproximación de rango para hallar una estimación de s .

b. ¿Cómo se compara con el valor calculado de s ?

DATOS
EX0131

1.31 Old Faithful (El viejo fiel) Los datos siguientes son 30 tiempos de espera entre erupciones del géiser Old Faithful del parque nacional de Yellowstone.⁶

56 89 51 79 58 82 52 88 52 78 69 75 77 72 71
55 87 53 85 61 93 54 76 80 81 59 86 78 71 77

a. Calcule el rango.

b. Use la aproximación de rango para aproximar la desviación estándar de estas 30 mediciones.

c. Calcule la desviación estándar de la muestra s .

d. ¿Qué proporción de las mediciones se encuentra a no más de dos desviaciones estándar de la media? ¿Y a no más de tres desviaciones estándar de la media? ¿Estas proporciones concuerdan con las proporciones dadas en el teorema de Chebyshev?

1.32 Un hallazgo arqueológico, otra vez Consulte el ejercicio 1.15. El porcentaje de óxido de hierro en cada una de cinco muestras de cerámica recolectadas en el sitio de Island Thorns fue:

1.28 2.39 1.50 1.88 1.51

a. Use la aproximación de rango para hallar una estimación de s , utilizando un divisor apropiado de la tabla 1.7.

b. Calcule la desviación estándar s . ¿Qué tan cerca estuvo su estimación del valor real de s ?

CÁLCULO DE LA MEDIA Y DESVIACIÓN ESTÁNDAR PARA DATOS AGRUPADOS (OPCIONAL)

1.33 Suponga que algunas mediciones se presentan más de una vez y que los datos $x_1, x_2, \dots, x_k$ están dispuestos en una tabla de frecuencia como se muestran aquí:

Observaciones	Frecuencia f_i
x_1	f_1
x_2	f_2
$\vdots$	$\vdots$
x_k	f_k

Las fórmulas para la media y varianza para datos agrupados son

$$\bar{x} = \frac{\sum x_i f_i}{n}, \quad \text{donde } n = \sum f_i$$

y

$$s^2 = \frac{\sum x_i^2 f_i - \frac{(\sum x_i f_i)^2}{n}}{n - 1}$$

Observe que si cada uno de los valores se presenta una vez, estas fórmulas se reducen a las dadas en el texto. Aun cuando estas fórmulas para datos agrupados son básicamente de valor cuando tenemos un gran número de mediciones, demuestre su uso para la muestra 1, 0, 0, 1, 3, 1, 3, 2, 3, 0, 0, 1, 1, 3, 2.

a. Calcule $\bar{x}$ y s^2 directamente, usando las fórmulas para datos no agrupados.

b. La tabla de frecuencia para las $n = 15$ mediciones es como sigue:

x	f
0	4
1	5
2	2
3	4

Calcule $\bar{x}$ y s^2 usando las fórmulas para datos agrupados. Compare con sus respuestas de la parte a.

1.34 International Baccalaureate Los estudiantes de bachillerato en un programa International Baccalaureate (IB) son inscritos en cursos acelerados o avanzados y deben tomar exámenes IB en cada una de seis materias al terminar su penúltimo o último año. Los estudiantes son calificados en una escala de 1-7, con 1-2 malo, 3 mediocre, 4 promedio y 5-7 excelente. Durante su primer año de operación en la preparatoria John W. North en Riverside, California, 17 estudiantes de penúltimo año trataron de pasar el examen IB de economía, con estos resultados:

Calificación de examen	Número de estudiantes
7	1
6	4
5	4
4	4
3	4

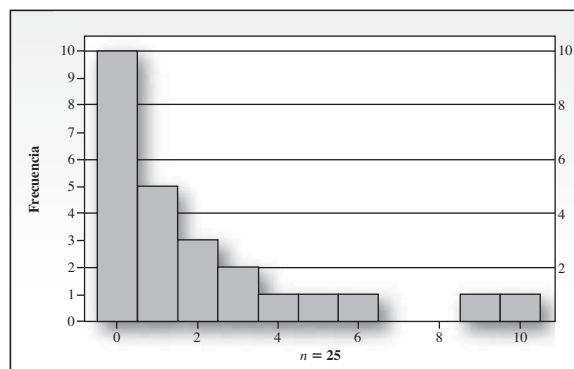
Calcule la media y desviación estándar para estas calificaciones.

1.35 Una distribución sesgada Para ilustrar la utilidad de la regla empírica, considere una distribución que está fuertemente sesgada a la derecha, como se muestra en la figura siguiente.

- Calcule $\bar{x}$ y s para los datos mostrados. (NOTA: Hay 10 ceros, cinco unos, y así sucesivamente.)
- Construya los intervalos $\bar{x} \pm s$, $\bar{x} \pm 2s$ y $\bar{x} \pm 3s$ y localícelos en la distribución de frecuencia.
- Calcule la proporción de las $n = 25$ mediciones que caen en cada uno de tres intervalos. Compare con el teorema de Chebyshev y la regla empírica. Observe que, aun cuando la proporción que cae en el intervalo $\bar{x} \pm s$ no concuerda cercanamente con la regla empírica, las proporciones que caen en los intervalos $\bar{x} \pm 2s$ y $\bar{x} \pm 3s$

concuerdan muy bien. Muchas veces esto es cierto, aun para distribuciones de datos que no tengan forma de montículo.

Distribución para el ejercicio 1.35



1.7

MEDICIONES DE POSICIÓN RELATIVA

A veces es necesario conocer la posición de una observación respecto a otras de un conjunto de datos. Por ejemplo, si usted se examina con un total de 30 puntos, podría desear saber cómo se compara su calificación de 30 con las calificaciones de los otros estudiantes del grupo. La media y desviación estándar de las calificaciones se pueden usar para calcular un **puntaje z**, que mide la posición relativa de una medición en un conjunto de datos.

Definición El **puntaje z muestral** es una medida de posición relativa definida por

$$\text{puntaje } z = \frac{x - \bar{x}}{s}$$

TIP MI CONSEJO

Puntaje z positivo $\Leftrightarrow x$ está arriba de la media
Puntaje z negativo $\Leftrightarrow x$ está debajo de la media

Un **puntaje z mide la distancia entre una observación y la media, medidas en unidades de desviación estándar**. Por ejemplo, suponga que la media y desviación estándar de los puntajes de examen (basados en un total de 35 puntos) son 25 y 4, respectivamente. El puntaje z para su calificación de 30 se calcula como sigue:

$$\text{puntaje } z = \frac{x - \bar{x}}{s} = \frac{30 - 25}{4} = 1.25$$

Su puntaje de 30 está a 1.25 desviaciones estándar arriba de la media ($30 = \bar{x} + 1.25s$).

El puntaje z es una valiosa herramienta para determinar si es probable que una observación particular se presente con frecuencia, o si es improbable y puede ser considerada como **resultado atípico**.

De acuerdo con el teorema de Chebyshev y la regla empírica,

- al menos 75% y más probablemente 95% de las observaciones están a no más de dos desviaciones estándar de su media: sus puntajes z están entre -2 y $+2$. *Las observaciones con puntajes z mayores a 2 en valor absoluto se presentan menos del 5% de veces para datos en forma de montículo y son consideradas un tanto improbables.*
- al menos 89% y más probablemente 99.7% de las observaciones están a no más de tres desviaciones estándar de su media: sus puntajes z están entre -3 y $+3$. *Las observaciones con puntajes z mayores a 3 en valor absoluto se presentan menos del 1% del tiempo para datos en forma de montículo y son consideradas muy poco probables.*

TIP MI CONSEJO

Los puntajes z mayores a 3 en valor absoluto son muy poco comunes

EJEMPLO

1.11

Considere esta muestra de n mediciones:

1, 1, 0, 15, 2, 3, 4, 0, 1, 3

La medición $x = 15$ parece ser extraordinariamente grande. Calcule el puntaje z para esta observación y exprese sus conclusiones.

Solución Calcule $\bar{x} = 3.0$ y $s = 4.42$ para las $n = 10$ mediciones. Entonces el puntaje z para el resultado atípico sospechoso, $x = 15$, se calcula como

$$\text{puntaje } z = \frac{x - \bar{x}}{s} = \frac{15 - 3}{4.42} = 2.71$$

En consecuencia, la medición $x = 15$ está 2.71 desviaciones estándar arriba de la media muestral, $\bar{x} = 3.0$. Aun cuando el puntaje z no excede de 3, está lo suficientemente cercano para que usted sospeche que $x = 15$ es un resultado atípico. Debe examinar el procedimiento de muestreo para ver si $x = 15$ es una observación defectuosa.

Un **percentil** es otra medida de posición relativa y se usa con más frecuencia para conjuntos grandes de datos. (Los percentiles no son muy útiles para conjuntos pequeños de datos.)

Definición Un conjunto de n mediciones de la variable x se ha dispuesto en orden de magnitud. El **p -ésimo percentil** es el valor de x que es mayor que $p\%$ de las mediciones y es menor que el restante $(100 - p)\%$.

EJEMPLO

1.12

Suponga que ha sido notificado que su calificación de 610, en el examen verbal de graduación, lo ha colocado en el 60o. percentil en la distribución de calificaciones. ¿Dónde está su calificación de 610 en relación con las calificaciones de los otros que tomaron el examen?

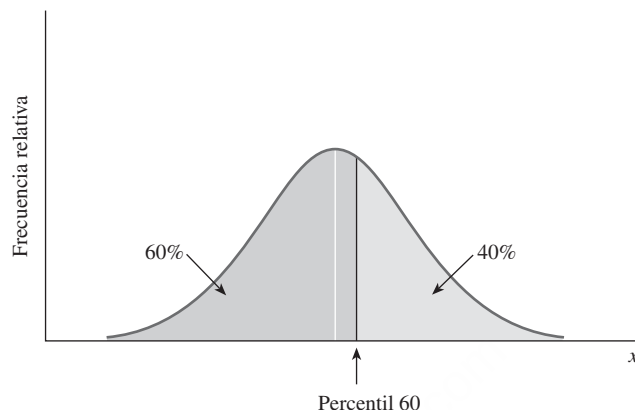
Solución Calificar en el 60o. percentil significa que 60% de todas las calificaciones del examen fueron más bajas que su calificación y 40% fueron más altas.

Para cualquier distribución de datos, sin importar su forma, el 60o. percentil para la variable x es un punto en el *eje horizontal* de la distribución de datos que es mayor que 60% de las mediciones y menor que las otras. Esto es, 60% de las mediciones son menores que el 60o.

percentil y 40% son mayores (véase la figura 1.12). Como el área total bajo la distribución es 100%, 60% del área está a la izquierda y 40% a la derecha del 60o. percentil. Recuerde que la mediana, m , de un conjunto de datos es la medición central; esto es, 50% de las mediciones son más pequeñas y 50% más grandes que la mediana. Por tanto, *¡la mediana es lo mismo que el 50o. percentil!*

FIGURA 1.12

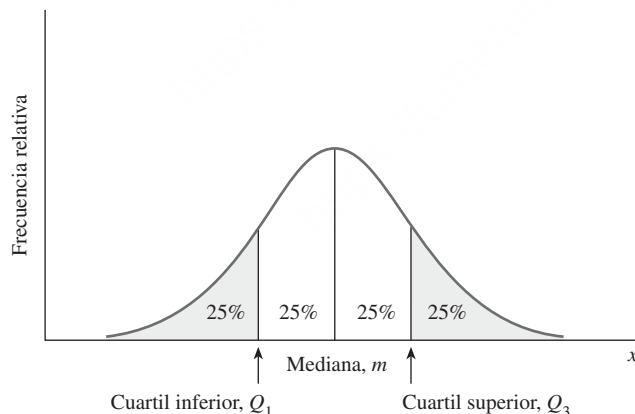
El 60o. percentil mostrado en el histograma de frecuencia relativa para un conjunto de datos



Los percentiles 25 y 75, llamados **cuartiles inferior y superior**, junto con la mediana (el 50o. percentil), localizan puntos que dividen los datos en cuatro conjuntos, cada uno conteniendo un número igual de mediciones. Veinticinco por ciento de las mediciones serán menores que el cuartil inferior (primero), 50% serán menores que la mediana (el segundo cuartil) y 75% serán menores que el cuartil superior (tercero). De este modo, la mediana y los cuartiles inferior y superior están ubicados en puntos en el eje x de modo que el área bajo el histograma de frecuencia relativa para los datos está dividida en cuatro áreas iguales, como se muestra en la figura 1.13.

FIGURA 1.13

Ubicación de cuartiles



Definición Un conjunto de n mediciones en la variable x se ha dispuesto en orden de magnitud. El **cuartil inferior (primer cuartil)**, Q_1 , es el valor de x que es mayor que un cuarto de las mediciones y es menor que los restantes tres cuartos. El **segundo cuartil** es la mediana. El **cuartil superior (tercer cuartil)**, Q_3 , es el valor de x que es mayor que tres cuartos de las mediciones y es menor que el restante un cuarto.

Para conjuntos de datos pequeños, con frecuencia es imposible dividir el conjunto en cuatro grupos, cada uno de los cuales contiene exactamente 25% de las mediciones. Por ejemplo, cuando $n = 10$, ¡usted necesitaría tener $2\frac{1}{2}$ mediciones en cada grupo! Aun cuando usted efectúe esta tarea (por ejemplo, si $n = 12$), hay muchos números que satisfarían la definición anterior y, por lo tanto, podrían considerarse “cuartiles”. Para evitar esta ambigüedad, usamos la siguiente regla para localizar cuartiles muestrales.

CÁLCULO DE CUARTILES MUESTRALES

- Cuando las mediciones están dispuestas en orden de magnitud, el **cuartil inferior**, Q_1 , es el valor de x en la posición $.25(n + 1)$, y el **cuartil superior**, Q_3 , es el valor de x en la posición $.75(n + 1)$.
- Cuando $.25(n + 1)$ y $.75(n + 1)$ no son enteros, los cuartiles se encuentran por interpolación, usando los valores de las dos posiciones adyacentes.[†]

EJEMPLO

1.13

Encuentre los cuartiles inferior y superior para este conjunto de mediciones:

16, 25, 4, 18, 11, 13, 20, 8, 11, 9

Solución Ordene las $n = 10$ mediciones de menor a mayor:

4, 8, 9, 11, 11, 13, 16, 18, 20, 25

Calcule

$$\text{Posición de } Q_1 = .25(n + 1) = .25(10 + 1) = 2.75$$

$$\text{Posición de } Q_3 = .75(n + 1) = .75(10 + 1) = 8.25$$

Como estas posiciones no son enteros, el cuartil inferior se toma como el valor $3/4$ de la distancia entre la segunda y tercera mediciones ordenadas, y el cuartil superior se toma como el valor $1/4$ de la distancia entre la octava y novena mediciones ordenadas. Por tanto,

$$Q_1 = 8 + .75(9 - 8) = 8 + .75 = 8.75$$

y

$$Q_3 = 18 + .25(20 - 18) = 18 + .5 = 18.5$$

Como la mediana y los cuartiles dividen la distribución de datos en cuatro partes, cada una de ellas conteniendo alrededor de 25% de las mediciones, Q_1 y Q_3 son las fronteras superior e inferior para el 50% central de la distribución. Podemos medir el rango de este “50% central” de la distribución usando una medida numérica llamada **rango intercuartil**.

Definición El **rango intercuartil (IQR)** para un conjunto de mediciones es la diferencia entre los cuartiles superior e inferior; esto es, $IQR = Q_3 - Q_1$.

Para los datos del ejemplo 1.13, $IQR = Q_3 - Q_1 = 18.50 - 8.75 = 9.75$. Usaremos el IQR junto con los cuartiles y la mediana en la siguiente sección para construir otra gráfica para describir conjuntos de datos.

[†] Esta definición de cuartiles es consistente con la empleada en el paquete *MINITAB 16* y *MS Excel 2010*. Algunos libros de texto emplean redondeo ordinario cuando buscan posiciones de cuartil, mientras que otros calculan cuartiles muestrales como las medianas de las mitades superior e inferior del conjunto de datos.



NECESITO SABER...

Cómo calcular cuartiles muestrales

1. Acomode el conjunto de datos en orden de magnitud de menor a mayor.
2. Calcule las posiciones de cuartil:
 - Posición de Q_1 : $.25(n + 1)$
 - Posición de Q_3 : $.75(n + 1)$
3. Si las posiciones son de enteros, entonces Q_1 y Q_3 son los valores del conjunto ordenado de datos que se encuentra en esas posiciones.
4. Si las posiciones del paso 2 no son de enteros, encuentre las dos mediciones en las posiciones justo arriba y justo debajo de la posición calculada. Calcule el cuartil al hallar un valor ya sea de un cuarto, un medio o tres cuartos de la distancia entre estas dos mediciones.

Muchas de las medidas numéricas que usted ha aprendido se encuentran fácilmente usando programas de cómputo o incluso calculadoras graficadoras. El comando **MINITAB Stat ► Basic Statistics ► Display Descriptive Statistics** o el comando de **Excel Data ► Data Analysis ► Descriptive Statistics** (véase la sección “Tecnología actual” al final de este capítulo) produce una salida que contiene la media, la desviación estándar, la mediana y los cuartiles inferior y superior, así como los valores de algunas otras estadísticas que todavía no examinamos. Los datos del ejemplo 1.13 produjeron la salida **MINITAB** que se muestra en la figura 1.14. Observe que los cuartiles son idénticos a los valores calculados manualmente en ese ejemplo.

FIGURA 1.14

Salida **MINITAB** para los datos del ejemplo 1.13

Estadística descriptiva: x

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
x	10	0	13.50	1.98	6.28	4.00	8.75	12.00	18.50	25.00

EL RESUMEN DE CINCO NÚMEROS Y LA GRÁFICA DE CAJA

1.8

La mediana y los cuartiles superior e inferior que se muestran en la figura 1.13 dividen los datos en cuatro conjuntos, cada uno de los cuales contiene igual número de mediciones. Si agregamos el número más grande (Máx) y el número más pequeño (Mín) del conjunto de datos a este grupo, tendremos un conjunto de números que da un rápido y aproximado resumen de la distribución de datos.

El **resumen de cinco números** consta del número más pequeño, el cuartil inferior, la mediana, el cuartil superior, y el número más grande, presentados en orden de menor a mayor:

Mín Q_1 Mediana Q_3 Máx

Por definición, un cuarto de las mediciones del conjunto de datos se ubica entre cada uno de los cuatro pares adyacentes de números.

El resumen de cinco números se utiliza para crear una gráfica sencilla llamada **gráfica de caja** a fin de describir visualmente la distribución de datos. De la gráfica de caja, rápidamente se puede detectar cualquier sesgo en la forma de la distribución y ver si hay algunos resultados atípicos en el conjunto de datos. Un resultado atípico aparece al trasponer dígitos cuando se registra una medición, al leer incorrectamente la carátula de un instrumento, por el mal funcionamiento de una pieza de equipo o por otros problemas.

Aun cuando no haya errores de registro o de observación, un conjunto de datos puede contener una o más mediciones válidas que, por una u otra razón, difieren marcadamente de las otras del conjunto. Estos resultados atípicos pueden causar una notable distorsión en medidas numéricas de uso común tales como $\bar{x}$ y s . De hecho, los atípicos pueden contener información importante no compartida con las otras mediciones del conjunto. Por tanto, los resultados atípicos aislados, si están presentes, son un paso importante en cualquier análisis preliminar de un conjunto de datos. La gráfica de caja está diseñada expresamente para este fin.

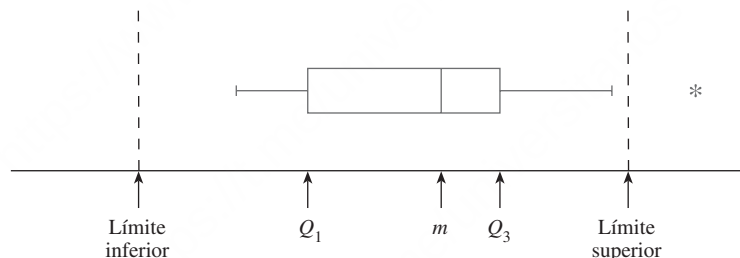
PARA CONSTRUIR UNA GRÁFICA DE CAJA

- Calcule la mediana, los cuartiles superior e inferior y el IQR para el conjunto de datos.
- Trace una recta horizontal que represente la escala de medición. Forme una caja justo arriba de la recta horizontal con los extremos derecho e izquierdo en Q_1 y Q_3 . Trace una recta vertical que pase por la caja en la ubicación de la mediana.

Se muestra una gráfica de caja en la figura 1.15.

FIGURA 1.15

Gráfica de caja



En la sección 1.6, el puntaje z dio fronteras para hallar mediciones extraordinariamente grandes o pequeñas. Buscamos puntajes z mayores que 2 o 3 en valor absoluto. La gráfica de caja usa el IQR para crear “límites” imaginarios para separar resultados atípicos del resto del conjunto de datos:

DETECCIÓN DE RESULTADOS ATÍPICOS—OBSERVACIONES QUE ESTÁN A MAYOR DISTANCIA:

- Límite inferior: $Q_1 - 1.5(\text{IQR})$
- Límite superior: $Q_3 + 1.5(\text{IQR})$

Los límites superior e inferior se muestran con líneas interrumpidas en la figura 1.15, pero no suelen ser trazadas en la gráfica de caja. Cualquier medición a mayor distancia del límite superior o inferior es un **resultado atípico**; el resto de las mediciones, dentro de los límites, no son inusuales. Por último, la gráfica de caja marca el rango del conjunto de datos usando “bigotes” para conectar las mediciones más pequeñas y más grandes (*excluyendo resultados atípicos*) a la caja.

PARA TERMINAR LA GRÁFICA DE CAJA

- Marque cualesquier **resultados atípicos** con un asterisco (*) en la gráfica.
- Prolongue rectas horizontales llamadas “bigotes” desde los extremos de la caja a las observaciones más pequeñas y más grandes *que no* sean resultados atípicos.

EJEMPLO

1.14

A medida que los consumidores estadounidenses tienen más cuidado con los alimentos que consumen, las procesadoras de alimentos tratan de ser competitivas al evitar cantidades excesivas de grasa, colesterol y sodio en los alimentos que venden. Los datos siguientes son las cantidades de sodio por rebanada (en miligramos) para cada una de ocho marcas de queso normal estadounidense. Construya una gráfica de caja para los datos y busque resultados atípicos.

340, 300, 520, 340, 320, 290, 260, 330

Solución Las $n = 8$ mediciones se ordenan primero de menor a mayor:

260, 290, 300, 320, 330, 340, 340, 520

Las posiciones de la mediana, Q_1 y Q_3 son

$$.5(n + 1) = .5(9) = 4.5$$

$$.25(n + 1) = .25(9) = 2.25$$

$$.75(n + 1) = .75(9) = 6.75$$

de modo que $m = (320 + 330)/2 = 325$, $Q_1 = 290 + .25(10) = 292.5$, y $Q_3 = 340$.

El rango intercuartil se calcula como

$$\text{IQR} = Q_3 - Q_1 = 340 - 292.5 = 47.5$$

Calcule los límites superior e inferior:

$$\text{Límite inferior: } 292.5 - 1.5(47.5) = 221.25$$

$$\text{Límite superior: } 340 + 1.5(47.5) = 411.25$$



APPLET EN LÍNEA

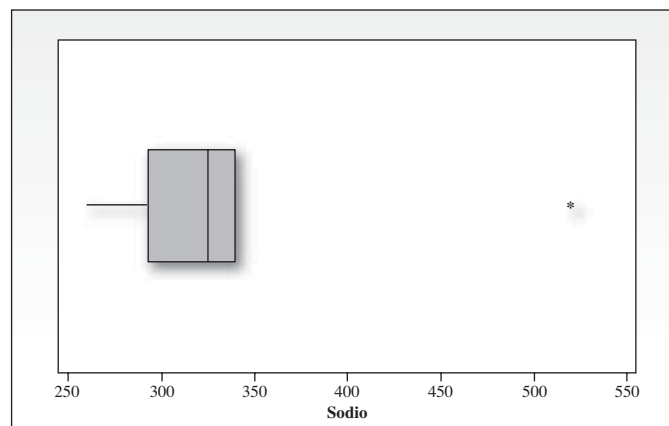
Construir una gráfica de caja

El valor $x = 520$, una marca de queso que contiene 520 miligramos de sodio, es el único *resultado atípico* que se encuentra fuera del límite superior.

La gráfica de caja para los datos se muestra en la figura 1.16. El resultado atípico está marcado con un asterisco (*). Una vez excluido el resultado atípico, encontramos (del conjunto ordenado de datos) que las mediciones más pequeña y más grande son $x = 260$ y $x = 340$. Éstos son los dos valores que forman los bigotes. Como el valor $x = 340$ es igual a Q_3 , no hay bigote en el lado derecho de la caja.

FIGURA 1.16

Gráfica de caja para el ejemplo 1.14



Utilice la gráfica de caja para describir la forma de una distribución de datos al ver la posición de la recta mediana comparada contra Q_1 y Q_3 , así como los extremos izquierdo y derecho de la caja. Si la mediana está cerca del centro de la caja, la distribución es más o menos simétrica, dando así intervalos de igual tamaño para contener los dos cuartos centrales de los datos. Si la recta mediana está a la izquierda del centro, la distribución está sesgada a la derecha; si la mediana está a la derecha del centro, la distribución está sesgada a la izquierda. También, para casi todas las distribuciones sesgadas, el bigote en el lado sesgado de la caja tiende a ser más largo que el bigote del otro lado.

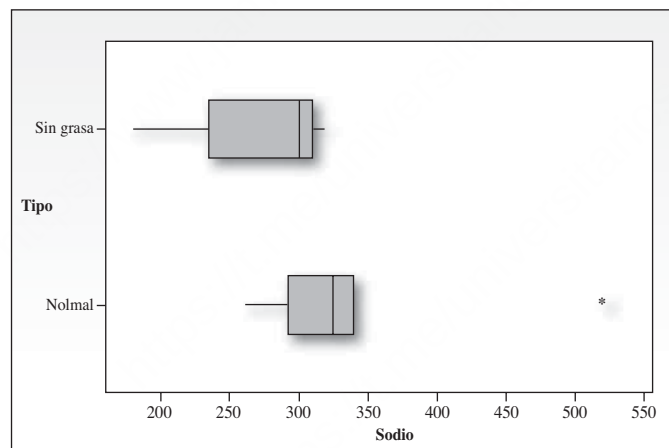
La figura 1.17 muestra dos gráficas de caja, una para el contenido de sodio de las ocho marcas de queso del ejemplo 1.14, y otra para cinco marcas de queso sin grasa con estos contenidos de sodio:

300, 300, 320, 290, 180

Examinemos el bigote largo del lado izquierdo de ambas gráficas y la posición de las rectas medianas. Ambas distribuciones están sesgadas a la izquierda; esto es, hay unas pocas mediciones inusualmente pequeñas. No obstante, los datos del queso normal también muestran una marca ($x = 520$) con una cantidad de sodio extraordinariamente grande. En general, aparece que el contenido de sodio de las marcas sin grasa es menor que la de las marcas regulares, pero la variabilidad del contenido de sodio para queso normal (excluyendo el resultado atípico) es menor que la de las marcas sin grasa.

FIGURA 1.17

Gráficas de caja para queso normal y sin grasa



1.8

EJERCICIOS

TÉCNICAS BÁSICAS

1.36 Dado el siguiente conjunto de datos: 8, 7, 1, 4, 6, 6, 4, 5, 7, 6, 3, 0.

- Encuentre el resumen de cinco números y el IQR.
- Calcule $\bar{x}$ y s .
- Calcule el puntaje z para las observaciones más pequeñas y más grandes. ¿Alguna de estas observaciones es extraordinariamente grande o pequeña?

1.37 Encuentre el resumen de cinco números y el IQR para estos datos:

19, 12, 16, 0, 14, 9, 6, 1, 12, 13, 10, 19, 7, 5, 8

1.38 Dado el siguiente conjunto de datos: 2.3, 1.0, 2.1, 6.5, 2.8, 8.8, 1.7, 2.9, 4.4, 5.1, 2.0

- Encuentre las posiciones de los cuartiles inferior y superior.
- Ordene los datos de menor a mayor y encuentre los cuartiles inferior y superior.
- Calcule el IQR.

1.39 Dado el siguiente conjunto de datos: .23, .30, .35, .41, .56, .58, .76, .80.

- Encuentre los cuartiles inferior y superior.
- Calcule el IQR.

c. Calcule los límites inferior y superior. ¿Hay resultados atípicos?

1.40 Construya una gráfica de caja para estos datos e identifique los resultados atípicos:

25, 22, 26, 23, 27, 26, 28, 18, 25, 24, 12

1.41 Construya una gráfica de caja para estos datos e identifique los resultados atípicos:

3, 9, 10, 2, 6, 7, 5, 8, 6, 6, 4, 9, 22

APLICACIONES

1.42 Si usted calificó en el 69o. percentil en un examen de conocimientos, ¿cómo se compara su calificación con otras?

DATOS
EX0143

1.43 Concentración de mercurio en

delfines Los científicos del medio ambiente están cada vez más preocupados por la acumulación de elementos tóxicos en mamíferos marinos, así como en la transferencia de dichos elementos a los descendientes de estos animales. El delfín de franjas (*Stenella coeruleoalba*), considerado el principal depredador en la cadena alimenticia marina, fue objeto de este estudio. Las concentraciones de mercurio (microgramos/gramo) en los hígados de 28 delfines de franjas machos fueron como sigue:

1.70	183.00	221.00	286.00
1.72	168.00	406.00	315.00
8.80	218.00	252.00	241.00
5.90	180.00	329.00	397.00
101.00	264.00	316.00	209.00
85.40	481.00	445.00	314.00
118.00	485.00	278.00	318.00

a. Calcule el resumen de cinco números para los datos.

b. Construya una gráfica de caja para los datos.

c. ¿Hay algún resultado atípico?

d. Si usted supiera que los primeros cuatro delfines tenían menos de tres años de edad, en tanto que los otros eran mayores de ocho años, ¿esta información ayudaría a explicar la diferencia en la magnitud de esas cuatro observaciones? Explique.

1.44 Carne para hamburguesa Los pesos (en libras) de los 27 paquetes de carne molida de res del ejercicio 1.22 (véase el conjunto de datos EX0122) se muestran a continuación, en orden de menor a mayor:

.75	.83	.87	.89	.89	.89	.92
.93	.96	.96	.97	.98	.99	1.06
1.08	1.08	1.12	1.12	1.14	1.14	1.17
1.18	1.18	1.24	1.28	1.38	1.41	

a. Confirme los valores de la media y desviación estándar, calculados en el ejercicio 1.22 como $\bar{x} = 1.05$ y $s = .17$.

b. Los dos paquetes de carne más grandes pesan 1.38 y 1.41 libras. ¿Estos dos paquetes son inusualmente pesados? Explique.

c. Construya una gráfica de caja para los pesos de paquetes. ¿Qué nos dice la posición de la recta mediana y la longitud de los bigotes acerca de la forma de la distribución?

DATOS
EX0145

1.45 Comparación de mariscales de campo de la NFL

¿Cómo se compara Aaron Rodgers, mariscal de campo de los ganadores del Súper Tazón de 2011, los Empacadores de Green Bay, con Drew Brees, mariscal de campo de los ganadores del Súper Tazón de 2010, los Santos de New Orleans? La tabla siguiente muestra el número de pases completados para cada atleta durante la temporada 2010 de fútbol de la NFL:⁷

Aaron Rodgers	Drew Brees
19	21
7	27
37	25
19	15
25	28
34	34
27	29
19	30
12	27
22	35
27	33
26	29
18	22
21	24
	23
	21
	24

a. Calcule los resúmenes de cinco números para el número de pases completos de Aaron Rodgers y Drew Brees.

b. Construya gráficas de caja para los dos conjuntos de datos. ¿Hay resultados atípicos? ¿Qué nos dicen las gráficas de caja acerca de las formas de las dos distribuciones?

c. Escriba un breve párrafo que compare el número de pases completos para los dos mariscales de campo.

1.46 Tiempos de supervivencia Altman y Bland informan de tiempos de supervivencia para pacientes con hepatitis activa, una mitad tratada con prednisona y la otra mitad no recibe tratamiento.⁸ Los tiempos de supervivencia (en meses) están adaptados de sus datos para los tratados con prednisona.

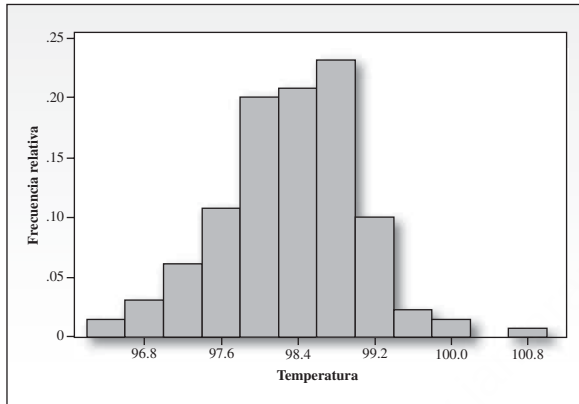
8	87	127	147
11	93	133	148
52	97	139	157
57	109	142	162
65	120	144	165

a. ¿Al observar estos datos, se puede decir si es más o menos simétrica? O bien, ¿es sesgada?

b. Calcule la media y mediana. Use estas medidas para determinar si los datos son o no simétricos o sesgados.

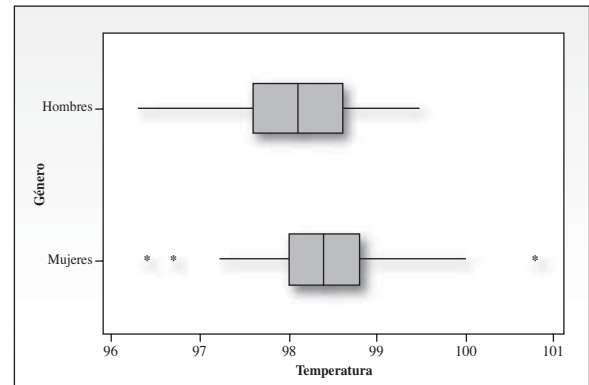
c. Trace una gráfica de caja para describir los datos. Explique por qué la gráfica de caja confirma lo concluido por usted en la parte b.

1.47 ¿Qué es normal? La temperatura corporal de 98.6 grados Fahrenheit como estándar en los seres humanos fue obtenida por un médico alemán en 1868. En un intento por verificar esta afirmación, Mackowiak, Wasserman y Levine⁹ tomaron las temperaturas de 148 personas sanas en un periodo de tres días. Un conjunto de datos, que estrechamente se compara con el del artículo de Mackowiak, fue obtenido por Allen Shoemaker y aparece en la *Journal of Statistics Education*.¹⁰ Las temperaturas corporales de 130 personas se muestran en el histograma de frecuencia relativa siguiente.



Además de registrar la temperatura corporal en grados Fahrenheit de las 130 personas, los datos incluyen su género. A continuación se presentan gráficas de caja para los dos grupos, hombres y mujeres:¹¹

Gráfica de caja para el ejercicio 1.47



¿Cómo describiría las similitudes y diferencias entre las temperaturas de hombres y mujeres en este conjunto de datos?

REPASO DEL CAPÍTULO

Conceptos clave y fórmulas

I. Medidas de centro de una distribución de datos

- Media aritmética (media) o promedio
 - Población: μ
 - Muestra de n mediciones: $\bar{x} = \frac{\sum x_i}{n}$
- Mediana; **posición** de la mediana = $.5(n + 1)$
- Moda
- Es preferida la mediana a la media si los datos son altamente sesgados.

II. Medidas de variabilidad

- Rango: $R = \text{máximo} - \text{mínimo}$
- Varianza
 - Población de N mediciones:

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

- Muestra de n mediciones:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n - 1}$$

- Desviación estándar
 - Población: $\sigma = \sqrt{\sigma^2}$
 - Muestra: $s = \sqrt{s^2}$
- Una aproximación burda para s se calcula como $s \approx R/4$. El divisor se puede ajustar de acuerdo con el tamaño muestral.

III. Teorema de Chebyshev y la regla empírica

- Use el teorema de Chebyshev para cualquier conjunto de datos, cualquiera que sea su forma o tamaño.
 - Al menos $1 - (1/k^2)$ de las mediciones se encuentra a no más de k desviaciones estándar de la media.
 - Éste es sólo un límite inferior; puede haber más mediciones en el intervalo.

- La regla empírica sólo se usa para conjuntos de datos en forma relativa de montículo. Aproximadamente 68%, 95% y 99.7% de las mediciones están dentro de una, dos y tres desviaciones estándar de la media, respectivamente.

IV. Mediciones de posición relativa

- Puntaje z muestral: $z = \frac{x - \bar{x}}{s}$
- p -ésimo percentil; $p\%$ de las mediciones son más pequeñas y $(100 - p)\%$ son más grandes.
- Cuartil inferior, Q_1 ; **posición** de $Q_1 = .25(n + 1)$
- Cuartil superior, Q_3 ; **posición** de $Q_3 = .75(n + 1)$
- Rango intercuartil: $IQR = Q_3 - Q_1$

V. El resumen de cinco números y gráficas de caja

- El resumen de cinco números:

Mín Q_1 Mediana Q_3 Máx

Un cuarto de las mediciones del conjunto de datos está entre cada uno de los cuatro pares adyacentes de números.

- Se usan gráficas de caja para detectar resultados atípicos y formas de distribuciones.
- Q_1 y Q_3 forman los extremos de la caja. La recta mediana está en el interior de la caja.
- Se usan límites superiores e inferiores para hallar resultados atípicos, observaciones que están fuera de estos límites.
 - Límite inferior:** $Q_1 - 1.5(IQR)$
 - Límite superior:** $Q_3 + 1.5(IQR)$
- Los **resultados atípicos** están marcados en la gráfica de caja con un asterisco (*).
- Los **bigotes** están conectados a la caja desde las observaciones más pequeña y más grande que *no sean* resultados atípicos.
- Las distribuciones sesgadas por lo general tienen un bigote largo *en la dirección del sesgo* y la recta mediana se traza *alejándose de la dirección del sesgo*.



Medidas descriptivas numéricas en *Excel*

MS Excel proporciona la mayoría de las estadísticas descriptivas básicas presentadas en este capítulo usando un solo comando en la pestaña **Data**. Otras estadísticas descriptivas pueden calcularse usando el comando **Function** en la pestaña **Formulas**.

EJEMPLO

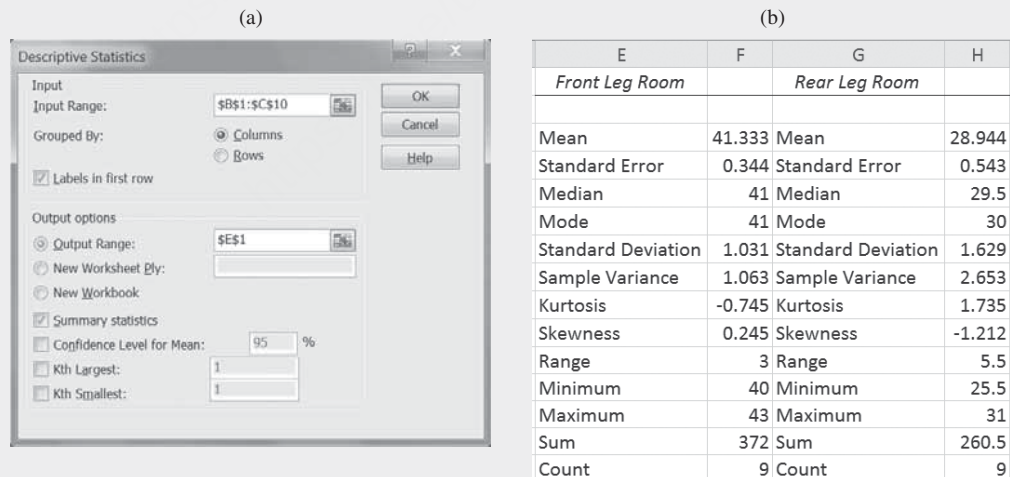
1.15

Los siguientes datos son los espacios (en pulgadas) frontal y trasero para las piernas de nueve vehículos minivans diferentes:¹²

Marca y modelo	Espacio frontal para las piernas	Espacio trasero para las piernas
Acura MDX	41.0	28.5
Buick Enclave	41.5	30.0
Chevy TrailBlazer	40.0	25.5
Chevy Tahoe Hybrid V8 CVT	41.0	27.5
GMC Terrain 1LT 4-cyl	43.0	31.0
Honda CR-V	41.0	29.5
Hyundai Tucson	42.5	29.5
Kia Sportage	40.0	29.0
Lexus GX	42.0	30.0

1. En vista que los datos involucran dos variables y una tercera variable etiquetadora, introduzca los datos en las primeras tres columnas de una hoja de cálculo de *Excel*, usando las etiquetas en la tabla. Seleccione **Data ► Data Analysis ► Descriptive Statistics**, y resalte o escriba el **Input range** (los datos en la segunda y tercera columnas) en el cuadro de diálogo Descriptive Statistics (figura 1.18a)). Introduzca una ubicación Output (Salida), asegúrese de comprobar que los cuadros “Labels in First Row” y “Summary Statistics” están seleccionados y haga clic en **OK**. Los estadísticos de resumen (figura 1.18b)) aparecerán en la ubicación seleccionada en la hoja de cálculo.

FIGURA 1.18



2. Observe que algunas de las celdas en la hoja de cálculo se traslapan. Para ajustar esto, resalte las columnas afectadas y haga clic en la pestaña **Home**. En el grupo **Cells**, elija **Format ► AutoFit Column Width**. Quizá desee modificar la apariencia de la salida por

la disminución de la precisión decimal en ciertas celdas. Resalte las celdas apropiadas y haga clic en el icono **Decrease Decimal**  (pestaña **Home**, grupo **Number**) para modificar la salida. Hemos desplegado la precisión a tres lugares decimales.

- Observe que no se proporcionan los cuartiles muestrales, Q_1 y Q_3 , en la salida de *Excel* en la figura 1.18b). Puede calcular los cuartiles usando el comando de función. Coloque el cursor en una celda vacía y seleccione **Formulas ► More Functions ► Statistical ► QUARTILE.EXC**. Resalte las celdas apropiadas en el cuadro marcado “Array” e introduzca un entero (0 = mín, 1 = primer cuartil, 2 = mediana, 3 = tercer cuartil o 4 = máx) en el cuadro marcado “Quart”. El cuartil (calculado usando el método de este libro) aparecerá en la celda que eligió. Se usará un método alternativo para calcular los cuartiles si selecciona **Formulas ► More Functions ► Statistical ► QUARTILE.INC**. (NOTA: Esta función se llama **QUARTILE** en *Excel 2007* y versiones anteriores.) Usando los dos cuartiles, puede calcular el IQR y construir una gráfica de caja en forma manual.

Medidas numéricas descriptivas en MINITAB

El *MINITAB* da casi todas las estadísticas descriptivas básicas presentadas en este capítulo usando un solo comando en los menús desplegables.

EJEMPLO

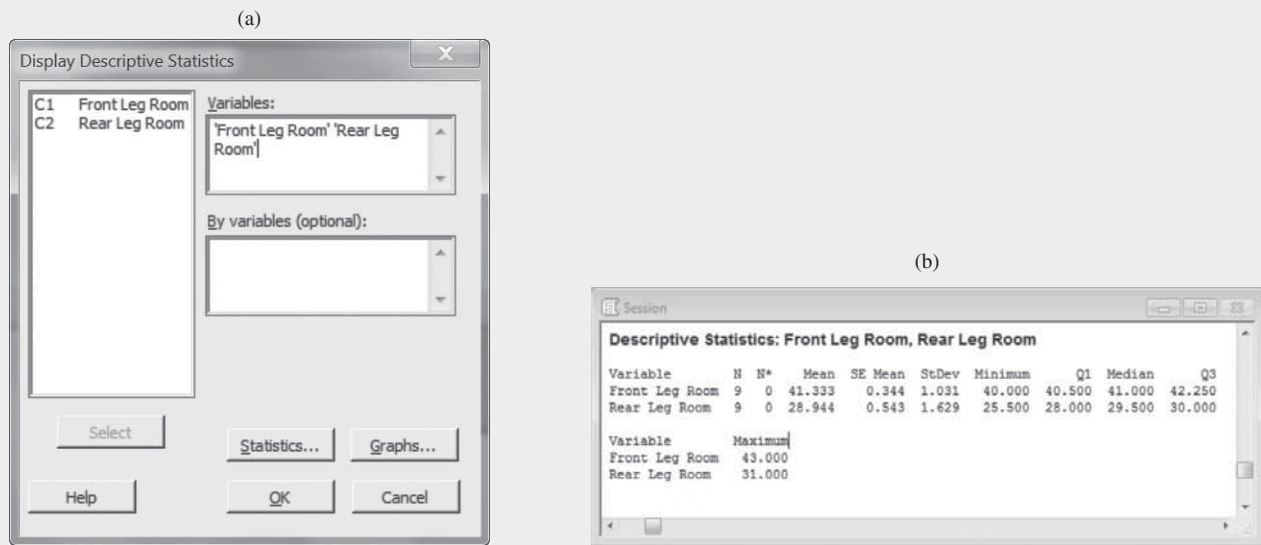
1.16

Los siguientes datos son los espacios para las piernas frontal y trasero (en pulgadas) para nueve vehículos minivans diferentes:¹²

Marca y modelo	Espacio frontal para las piernas	Espacio trasero para las piernas
Acura MDX	41.0	28.5
Buick Enclave	41.5	30.0
Chevy TrailBlazer	40.0	25.5
Chevy Tahoe Hybrid V8 CVT	41.0	27.5
GMC Terrain 1LT 4-cyl	43.0	31.0
Honda CR-V	41.0	29.5
Hyundai Tucson	42.5	29.5
Kia Sportage	40.0	29.0
Lexus GX	42.0	30.0

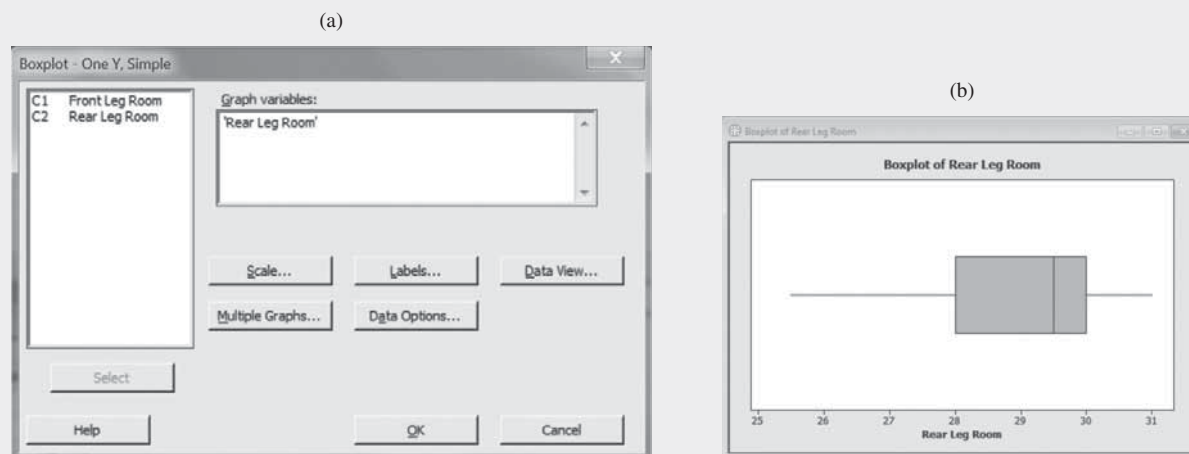
- En vista que los datos involucran dos variables y una tercera variable etiquetadora, introduzca los datos en las primeras tres columnas de una hoja de cálculo de *MINITAB*, usando las etiquetas en la tabla. Usando los menús desplegables, seleccione **Stat ► Basic Statistics ► Display Descriptive Statistics**. El cuadro de diálogo se muestra en la figura 1.19a).

FIGURA 1.19



- Ahora haga clic en el cuadro Variables y **seleccione** ambas columnas de la lista de la izquierda. (Puede dar un clic en la opción **Graphs** y elegir una de varias gráficas si lo desea. También puede hacer clic en la opción **Statistics** para seleccionar las estadísticas que desee ver en pantalla.) Haga clic en **OK**. En la ventana Session aparecerá una pantalla de estadísticas descriptivas para ambas columnas (véase la figura 1.19b)). Si lo desea imprima esta salida usando **File ► Print Session Window**.
- Para examinar la distribución de las dos variables y buscar resultados atípicos, cree gráficas de caja usando el comando **Graph ► Boxplot ► One Y ► Simple**. Haga clic en **OK**. Seleccione la columna de mediciones apropiada del cuadro de Diálogo (véase la figura 1.20a)). Puede cambiar la presentación de la gráfica de caja en varias formas. **Scale ► Axes and Ticks** le permitirán trasponer los ejes y orientar la gráfica de caja en sentido horizontal, cuando aplique un puntaje en la caja “Transpose value and category scales”. **Multiple Graphs** da opciones de impresión para múltiples gráficas de caja. **Labels** permite poner notas, títulos y notas al pie en la gráfica. Si ya ha introducido datos en la hoja de trabajo como distribución de frecuencia (valores en una columna, frecuencias en otra), las **Data Options** permitirán leer los datos en ese formato. La gráfica de caja de los espacios traseros para las piernas se muestra en la figura 1.20b).
- Guarde esta hoja de cálculo en un archivo llamado “Espacio para las piernas” antes de salir de *MINITAB*.

FIGURA 1.20



Ejercicios suplementarios

DATOS
EX0148 **1.48 Pasas** Se contaron el número de pasas en cada una de las 14 minicajas (tamaño de 1/2 onza) de una marca genérica y de la marca Sunmaid.

Aquí se presentan los dos conjuntos de datos:

Marca genérica	Sunmaid
25 26 25 28	25 29 24 24
26 28 28 27	28 24 28 22
26 27 24 25	25 28 30 27
26 26	28 24

- ¿Cuáles son la media y desviación estándar para la marca genérica?
- ¿Cuáles son la media y desviación estándar para la marca Sunmaid?
- Compare los centros y variabilidades de las dos marcas usando los resultados de las partes **a** y **b**.

1.49 Pasas, continúa Consulte el ejercicio 1.48.

- Encuentre la mediana, los cuartiles superior e inferior y el IQR para cada uno de los dos conjuntos de datos.
- Construya dos gráficas de caja en la misma escala horizontal para comparar los dos conjuntos de datos.

- Trace dos gráficas de tallo y hoja para describir las formas de los dos conjuntos de datos. ¿Las gráficas de caja de la parte **b** verifican estos resultados?
- Si suponemos que ninguna de las cajas de pasas se llena bien (es decir, todas pesan aproximadamente 1/2 onza), ¿qué dicen los resultados de usted acerca del número promedio de pasas para las dos marcas?

DATOS
EX0150 **1.50 Televidentes** El número de horas vistas de televisión por familia, así como las horas de mayor audiencia, son dos factores que afectan el ingreso por publicidad en televisión. Una muestra aleatoria de 25 familias en una zona particular produjo las siguientes estimaciones de horas vistas por familia:

3.0	6.0	7.5	15.0	12.0
6.5	8.0	4.0	5.5	6.0
5.0	12.0	1.0	3.5	3.0
7.5	5.0	10.0	8.0	3.5
9.0	2.0	6.5	1.0	5.0

- Revise los datos y use el rango para encontrar un valor aproximado de s . Use este valor para verificar sus cálculos de la parte **b**.

- b. Calcule la media muestral $\bar{x}$ y la desviación estándar de la muestra s . Compare s con el valor aproximado obtenido en la parte a.
- c. Encuentre el porcentaje de las horas vistas de televisión por familia, que caiga en el intervalo $\bar{x} \pm 2s$. Compare con el correspondiente porcentaje dado por la regla empírica.

1.51 Una enfermedad recurrente Se registraron los tiempos (en meses) entre el comienzo de una enfermedad particular y su recurrencia:

2.1	4.4	2.7	32.3	9.9
9.0	2.0	6.6	3.9	1.6
14.7	9.6	16.7	7.4	8.2
19.2	6.9	4.3	3.3	1.2
4.1	18.4	.2	6.1	13.5
7.4	.2	8.3	.3	1.3
14.1	1.0	2.4	2.4	18.0
8.7	24.0	1.4	8.2	5.8
1.6	3.5	11.4	18.0	26.7
3.7	12.6	23.1	5.6	.4

- a. Encuentre el rango.
- b. Use la aproximación del rango para hallar un valor aproximado de s .
- c. Calcule s para los datos y compárela con su aproximación de la parte b.

1.52 Una enfermedad recurrente, continúa Consulte el ejercicio 1.51.

- a. Examine los datos y cuente el número de observaciones que caen en los intervalos $\bar{x} \pm s$, $\bar{x} \pm 2s$ y $\bar{x} \pm 3s$.
- b. ¿Los porcentajes que caen en estos intervalos concuerdan con el teorema de Chebyshev? ¿Y con la regla empírica?
- c. ¿Por qué la regla empírica no sería apropiada para describir estos datos?

1.53 Una enfermedad recurrente, otra vez Encuentre la mediana, así como los cuartiles inferior y superior, para los datos sobre los tiempos hasta la recurrencia de una enfermedad del ejercicio 1.51. Utilice estas medidas descriptivas para construir una gráfica de caja para los datos. Use la gráfica de caja para describir la distribución de datos.

1.54 Atunes, otra vez Consulte el ejercicio 1.7. A continuación se reproducen aquí los precios de una lata de 6 onzas o una bolsa de 7.06 onzas, para 14 marcas diferentes de atún claro empacado en agua, con base en precios pagados nacionalmente en supermercados.³

.99	1.92	1.23	.85	.65	.53	1.41
1.12	.63	.67	.69	.60	.60	.66

- a. Calcule el resumen de cinco números.
- b. Construya una gráfica de caja para los datos. ¿Hay algún resultado atípico?

- c. El valor $x = 1.92$ se ve mayor en comparación con los otros precios. Use un puntaje z para determinar si ésta es una marca inusualmente cara de atún.

1.55 Electrólisis Un químico analítico desea usar electrólisis para determinar el número de moles de iones de cobre en un volumen determinado de solución. La solución se dividió en $n = 30$ partes de .2 mililitros cada una y se probó cada una de las partes. Se encontró que el número promedio de moles de iones de cobre para las $n = 30$ partes fue de .17 moles; la desviación estándar fue de .01 mol.

- a. Describa la distribución de las mediciones para las $n = 30$ partes de la solución usando el teorema de Chebyshev.
- b. Describa la distribución de las mediciones para las $n = 30$ partes de la solución usando la regla empírica. (¿Espera usted que la regla empírica sea apropiada para describir estos datos?)
- c. Suponga que el químico sólo usó $n = 4$ partes de la solución para el experimento y obtuvo las lecturas .15, .19, .17 y .15. ¿La Regla empírica sería apropiada para describir las $n = 4$ mediciones? ¿Por qué?

1.56 Cloroformo De acuerdo con la EPA, el cloroformo, que en su estado gaseoso es sospechoso de ser un agente cancerígeno, está presente en pequeñas cantidades en todas las 240 mil fuentes públicas de agua del país (Estados Unidos). Si la media y desviación estándar de las cantidades de cloroformo presentes en las fuentes de agua son 34 y 53 microgramos por litro, respectivamente, describa la distribución para la población de todas las fuentes públicas de agua.

1.57 Exámenes de aprovechamiento Se encontró que las calificaciones de un examen de aprovechamiento de matemáticas para 400 estudiantes tenían una media y una varianza iguales a 600 y 4900, respectivamente. Si la distribución de las puntuaciones del examen tiene forma de montículo, aproximadamente: ¿cuántas de las calificaciones caerían en el intervalo de 530 a 670?, ¿cuántas calificaciones se esperaría que cayeran en el intervalo de 460 a 740?

1.58 Sueño y el estudiante universitario ¿Cuánto tiempo duerme un universitario durante una noche normal en la escuela? A un grupo de 10 estudiantes universitarios se le pidió que informaran del número de horas que durmió la noche previa, con los siguientes resultados:

7, 6, 7.25, 7, 8.5, 5, 8, 7, 6.75, 6

- a. Encuentre la media y la desviación estándar del número de horas de sueño para estos 10 estudiantes.
- b. Calcule el puntaje z para el máximo valor ($x = 8.5$).
¿Es éste un estudiante universitario que duerme más de lo normal?

- c. ¿Cuál es la medición que se informa con más frecuencia? ¿Cuál es el nombre de esta medida del centro?
- d. Construya una gráfica de caja para los datos. ¿La gráfica confirma sus resultados de la parte b? [SUGERENCIA: Como el puntaje z y la gráfica de caja son dos métodos no relacionados para detectar resultados atípicos y usan diferentes tipos de estadísticas, no necesariamente tienen que producir (pero por lo común lo hacen) los mismos resultados.]

DATOS
EX0159

1.59 Rendimiento en millas A continuación se muestran las millas por galón (mpg), para cada uno de los 20 automóviles medianos seleccionados de una línea de producción durante el mes de marzo.

23.1	21.3	23.6	23.7
20.2	24.4	25.3	27.0
24.7	22.7	26.2	23.2
25.9	24.7	24.4	24.2
24.9	22.2	22.9	24.6

- a. ¿Cuáles son el máximo y mínimo de millas por galón? ¿Cuál es el rango?
- b. Construya un histograma de frecuencia relativa para estos datos. ¿Cómo describiría usted la forma de la distribución?
- c. Encuentre la media y la desviación estándar.
- d. Ordene los datos de menor a mayor. Encuentre los puntajes z para las observaciones máxima y mínima. ¿Los consideraría usted como resultados atípicos? ¿Por qué sí o por qué no?
- e. ¿Cuál es la mediana?
- f. Encuentre los cuartiles inferior y superior.

1.60 Rendimiento en millas, continúa Consulte el ejercicio 1.59. Construya una gráfica de caja para los datos. ¿Hay algún resultado atípico? ¿Esta conclusión concuerda con sus resultados del ejercicio 1.59?

1.61 Agua de mar contaminada La contaminación causada por petróleo en mares y océanos estimula el crecimiento de algunos tipos de bacterias. Un conteo de microorganismos que se originan en el petróleo (bacterias por 100 mililitros) en 10 partes de agua de mar dieron estas lecturas:

49, 70, 54, 67, 59, 40, 61, 69, 71, 52

- a. Calcule el valor de s usando la aproximación de rango.
- b. Calcule $\bar{x}$ y s y compare con la aproximación de rango de la parte a.
- c. Construya una gráfica de caja para los datos y úsela para describir la distribución de datos.

1.62 Exámenes de aptitud escolar Los exámenes verbales y de aptitud escolar de matemáticas de un Consejo Universitario se califican en una escala de 200 a 800.

Parece razonable suponer que una distribución de todas las calificaciones de examen, ya sea verbal o de matemáticas, tiene forma de montículo. Si σ es la desviación estándar de una de estas distribuciones, ¿cuál es el valor máximo (aproximadamente) que tomaría σ ? Explique.

1.63 Rosas de tallo largo Una variedad de rosas de tallo largo tiene una distribución normal aproximada, con una longitud media de tallo de 15 pulgadas y desviación estándar de 2.5 pulgadas.

- a. Si uno acepta como “rosas de tallo largo” sólo las rosas con una longitud de tallo mayor que 12.5 pulgadas, ¿qué porcentaje de esas rosas sería inaceptable?
- b. ¿Qué porcentaje de esas rosas tendría una longitud de tallo entre 12.5 y 20 pulgadas?

DATOS
EX0164

1.64 Medicina para hipertensión Una compañía farmacéutica desea saber si un medicamento experimental que se está probando en sus laboratorios tiene algún efecto en la presión sanguínea sistólica. A 15 personas seleccionadas al azar se les dio el medicamento y se registraron sus presiones sanguíneas sistólicas (en milímetros).

172	148	123
140	108	152
123	129	133
130	137	128
115	161	142

- a. Calcule el valor de s usando la aproximación de rango.
- b. Calcule $\bar{x}$ y s para las 15 presiones sanguíneas.
- c. Encuentre dos valores, a y b , para que al menos 75% de las mediciones caigan entre a y b .

1.65 Derechos madereros A una compañía interesada en derechos madereros, para cierto terreno de pinos allioti, se le indica que el diámetro medio de estos árboles es de 14 pulgadas con una desviación estándar de 2.8 pulgadas. Suponga que la distribución de diámetros tiene forma aproximada de montículo.

- a. ¿Qué fracción de los árboles tendrá diámetros entre 8.4 y 22.4 pulgadas?
- b. ¿Qué fracción de los árboles tendrá diámetros mayores que 16.8 pulgadas?

DATOS
EX0166

1.66 Ambivalencia social Los siguientes datos representan las puntuaciones de ambivalencia social para 15 personas, medidas por un examen psicológico. (Cuanta más alta la calificación, más fuerte es la ambivalencia.)

9	13	12
14	15	11
10	4	10
8	19	13
11	17	9

- Calcule el valor de s usando la aproximación de rango.
- Calcule $\bar{x}$ y s para las 15 calificaciones de ambivalencia social.
- ¿Qué fracción de las calificaciones en realidad están en el intervalo $\bar{x} \pm 2s$?

1.67 Comerciales en TV La duración media de anuncios comerciales en televisión en una red televisiva determinada es de 75 segundos, con una desviación estándar de 20 segundos. Suponga que las duraciones están distribuidas normalmente en forma aproximada.

- ¿Cuál es la probabilidad aproximada de que un comercial dure menos de 35 segundos?
- ¿Cuál es la probabilidad aproximada de que un comercial dure más de 55 segundos?

DATOS
EX0168

1.68 Parásitos en zorros Una muestra aleatoria de 100 zorros fue examinada por un equipo de veterinarios para determinar la prevalencia de un tipo particular de parásito. Contando el número de parásitos por zorro, los veterinarios encontraron que 69 zorros no tenían parásitos, 17 tenían un parásito, y así sucesivamente. A continuación tenemos una tabulación de frecuencia de los datos:

Número de parásitos, x	0	1	2	3	4	5	6	7	8
Número de zorros, f	69	17	6	3	1	2	1	0	1

- Construya un histograma de frecuencia relativa para x , el número de parásitos por zorro.
- Calcule $\bar{x}$ y s para la muestra.
- ¿Qué fracción de las cuentas de parásitos cae dentro de dos desviaciones estándar de la media? ¿Dentro de tres desviaciones estándar? ¿Estos resultados concuerdan con el teorema de Chebyshev? ¿Y con la regla empírica?

1.69 Profesores universitarios Considere una población formada por el número de profesores por universidad en pequeños colegios. Suponga que el número de profesores por universidad tiene un promedio $\mu = 175$ y una desviación estándar $\sigma = 15$.

- Use el teorema de Chebyshev para hacer un enunciado acerca del porcentaje de universidades que tienen entre 145 y 205 profesores.
- Suponga que la población está normalmente distribuida. ¿Qué fracción de universidades tiene más de 190 profesores?

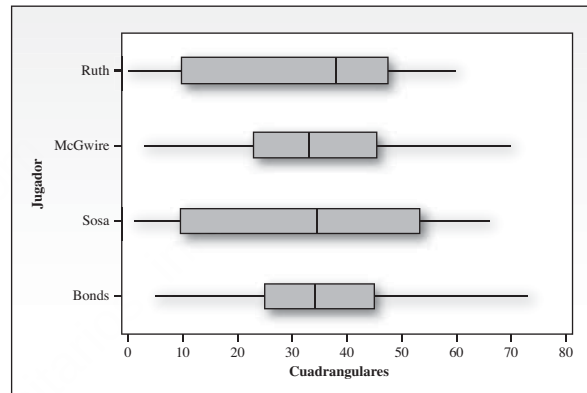
DATOS
EX0170

1.70 ¿Es precisa? De los datos siguientes, un estudiante calculó que s es .263. ¿En qué situación podríamos dudar de su precisión? ¿Cuál es el valor correcto (al centésimo más cercano)?

17.2 17.1 17.0 17.1 16.9 17.0 17.1 17.0 17.3 17.2
17.1 17.0 17.1 16.9 17.0 17.1 17.3 17.2 17.4 17.1

DATOS
EX0171

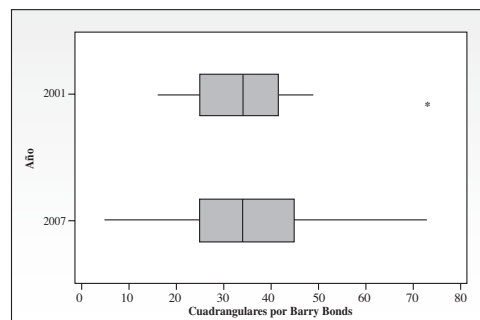
1.71 Reyes de cuadrangulares En el verano de 2001, Barry Bonds empezó su búsqueda de romper el récord de Mark McGwire de 70 cuadrangulares conectados en una sola temporada. Al terminar la temporada de beisbol de 2003 de las ligas mayores, se registró el número de cuadrangulares conectados por temporada por cada uno de cuatro superestrellas de ligas mayores en su carrera y a continuación se presentan en las gráficas de caja:¹³



Escriba un párrafo corto que compare los patrones de bateo de cuadrangulares de estos cuatro jugadores.

DATOS
EX0172

1.72 Barry Bonds En las temporadas que siguieron a la de 2001 en la que implantó récord, Barry Bonds conectó 46, 45, 45, 5, 26 y 28 cuadrangulares, respectivamente, hasta que se retiró de la liga mayor de beisbol en 2007 (www.espn.com).¹³ A continuación aparecen dos gráficas de caja, la primera de los cuadrangulares de Bonds en 2001 y la segunda que incluye los años 2002-2007.



Las estadísticas empleadas para construir estas gráficas de caja se dan en la tabla.

Años	Mín	Q_1	Mediana	Q_3	IQR	Máx	n
2001	16	25.00	34.00	41.50	16.5	73	16
2007	5	25.00	34.00	45.00	20.0	73	22

- a. Calcule los límites superiores para estas dos gráficas de caja.
- b. Explique por qué el número récord de cuadrangulares es un resultado atípico en la gráfica de caja de 2001, pero no en la gráfica de caja de 2007.

1.73 Tamiz metabólico En una maternidad se realizó a 50 bebés, desde 0 hasta 36 días de nacidos, un tamiz metabólico neonatal para la detección temprana de hipotiroidismo congénito.

```

0  0  0  0  0  0  0  0  0  0
0  0  1  1  1  1  1  1  2  2
2  3  3  3  4  4  5  5  5  5
6  8  9  9 10 16 17 17 19 19
19 20 20 21 22 23 25 25 28 36

```

- a. ¿Cuál es la edad promedio en días?
- b. ¿Cuál es la edad mediana en días?
- c. Con base en los resultados de las partes a y b, ¿cómo describiría usted la distribución?
- d. Construya una gráfica de caja para el conjunto de días. ¿Hay algún resultado atípico? ¿La gráfica de caja confirma su descripción de la forma de la distribución?

1.74 Instantáneas A continuación aparecen unos cuantos datos publicados como Snapshots (Instantáneas) en *USA Today*.

- a. Alrededor de 12% de los voluntarios de Estados Unidos dedican más de 5 horas a la semana en el voluntariado.¹⁴
- b. Cincuenta y ocho por ciento de todos los automóviles en operación tienen al menos 8 años de antigüedad.¹⁵
- c. Veintidós por ciento de todos los fanáticos están dispuestos a pagar \$75 o más por un boleto para uno de las 100 giras de conciertos principales.¹⁶

Identifique la variable x que se mide, y cualesquier percentiles que pueda usted determinar de esta información.

DATOS
EX0175

1.75 Patrones de respiración Psicólogos investigadores están interesados en averiguar si los patrones de respiración de una persona son afectados por un tratamiento experimental particular. Para determinar los patrones respiratorios generales de las $n = 30$ personas en el estudio, los investigadores recolectaron algunas mediciones de línea de base, es decir, el total de ventilación en litros de aire por minuto ajustados al tamaño del cuerpo, para cada persona antes del tratamiento. Los datos se muestran a continuación, junto con algunas herramientas descriptivas generadas por *MINITAB* y *MS Excel*.

```

5.23  4.79  5.83  5.37  4.35  5.54  6.04  5.48  6.58  4.82
5.92  5.38  6.34  5.12  5.14  4.72  5.17  4.99  4.51  5.70
4.67  5.77  5.84  6.19  5.58  5.72  5.16  5.32  4.96  5.63

```

Estadísticas descriptivas: litros

Variable	N	N*	Mean	SE Mean	StDev
Liters	30	0	5.3953	0.0997	0.5462
Minimum	Q1	Median	Q3	Variable	Maximum
4.3500	4.9825	5.3750	5.7850	Liters	6.5800

Gráfica de tallo y hoja: litros

Stem-and-leaf of Liters N = 30
Leaf Unit = 0.10

```

1  4  3
2  4  5
5  4  677
8  4  899
12 5 1111
(4) 5 2333
14 5 455
11 5 6777
7  5 889
4  6 01
2  6  3
1  6  5

```

Estadísticas descriptivas de MS Excel

Litros	
Mean	5.3953
Standard Error	0.0997
Median	5.3750
Mode	#N/A
Standard Deviation	0.5462
Sample Variance	0.2983
Kurtosis	20.4069
Skewness	0.1301
Range	2.23
Minimum	4.35
Maximum	6.58
Sum	161.86
Count	30

- a. Haga un resumen de las características de la distribución de datos usando la salida de computadora.
- b. ¿La regla empírica da una buena descripción de la proporción de mediciones que caen dentro de dos o tres desviaciones estándar de la media? Explique.
- c. ¿Qué tan grande o pequeña tiene que ser una medición de ventilación antes que sea considerada como poco común?

DATOS
EX0176

1.76 Ordenamiento de objetos Los datos siguientes son tiempos de respuesta en segundos para $n = 25$ estudiantes de primer año para ordenar tres objetos por tamaño.

```

5.2  3.8  5.7  3.9  3.7
4.2  4.1  4.3  4.7  4.3
3.1  2.5  3.0  4.4  4.8
3.6  3.9  4.8  5.3  4.2
4.7  3.3  4.2  3.8  5.4

```

- a. Encuentre la media y la desviación estándar para estos 25 tiempos de respuesta.
- b. Ordene los datos de menor a mayor.
- c. Encuentre los puntajes z para los tiempos de respuesta mínimo y máximo. ¿Hay alguna razón para creer que estos tiempos son extraordinariamente grandes o pequeños? Explique.

1.77 Ordenamiento de objetos, continúa Consulte el ejercicio 1.76.

- a. Encuentre el resumen de cinco números para este conjunto de datos.
- b. Construya una gráfica de caja para los datos.
- c. ¿Hay algunos tiempos de respuesta extraordinariamente grandes o pequeños identificados por la gráfica de caja?
- d. Construya una gráfica de tallo y hoja para los tiempos de respuesta. ¿Cómo describiría usted la forma de la distribución? ¿La forma de la gráfica de caja confirma este resultado?

Probabilidad

OBJETIVOS GENERALES

Ahora que usted ya ha aprendido a describir un conjunto de datos, ¿cómo usaría los datos muestrales para sacar conclusiones acerca de las poblaciones muestreadas? En esta técnica interviene una herramienta estadística llamada *probabilidad* y, para usarla correctamente, primero debe entender cómo funciona. La primera parte de este capítulo le presentará los conceptos básicos con ejemplos sencillos.

Las variables que medimos en el capítulo 1 se redefinen ahora como variables aleatorias, con valores que dependen de la selección de la probabilidad de los elementos de la muestra. Usando la probabilidad como herramienta, usted podrá crear distribuciones de probabilidad que servirán como modelos para variables aleatorias discretas y describirá estas variables aleatorias usando una media y desviación estándar semejantes a las del capítulo 1.

ÍNDICE DEL CAPÍTULO

- Introducción (2.1)
- El papel de la probabilidad en estadística (2.2)
- Eventos y el espacio muestral (2.3)
- Enfoques o escuelas de la probabilidad (2.4)
- Axiomas de probabilidad (2.5)
- Particiones (2.6)
- Cálculo de probabilidades con el uso de eventos sencillos (2.7)
- Reglas útiles de conteo (2.8)
- Relaciones de evento y reglas de probabilidad (2.9)
- Independencia, probabilidad condicional y la regla de la multiplicación (2.10)
- Probabilidad condicional (2.11)
- Teorema de Bayes (2.12)
- ¿Eventos mutuamente excluyentes o independientes? (2.13)
- Procesos estocásticos (2.14)



NECESITO SABER...

Cómo calcular la probabilidad de un evento
La diferencia entre eventos mutuamente excluyentes y eventos independientes

INTRODUCCIÓN

En la interpretación del concepto de probabilidad se han seguido tres escuelas. La primera, **clásica o de juegos**, la considera una actividad que pudiera conducir a los resultados previstos, sin olvidar que el efecto particular observado en cualquiera de sus etapas también está generado por el azar. En el siglo xvii, esta escuela tenía como principal interés determinar la probabilidad de éxito en los juegos de azar. En este caso, los experimentos en cuestión eran tirar un dado, girar la ruleta, manejar las cartas, etc. Los matemáticos de ese siglo, B. Pascal, P. Ferment y P.S. Laplace, entre otros, fueron quienes más contribuyeron a formalizar los conceptos de probabilidad. Considera “espacios muestrales equiprobables”, por lo que espera *resultados igualmente probables e implica resultados a priori*.

La segunda escuela es la **subjetivista o bayesiana**, la cual utiliza los resultados *a priori* para calcular la probabilidad de ocurrencia de varios estados del universo de eventos (también llamado “espacio de eventos”), en una etapa particular del experimento que se trata.

El conocimiento *a priori* es la forma en que surgirá la distribución de probabilidades, dadas las consideraciones de la naturaleza del experimento que va a realizarse. Por otra parte, el enfoque subjetivista permite al investigador asignar probabilidades o eventos particulares (algunos posibles resultados del experimento que se realiza), el grado de confianza en algo, está en la mente y varía de persona a persona.

Como se mencionó, la escuela subjetivista, también conocida como **bayesiana** (en este libro, “bayesiana” significa el deseo de incorporar probabilidades subjetivas en el análisis estadístico), es contraria a los autores de la escuela clásica que no deseaban incorporar presentimientos subjetivos en probabilidad a la estructura formal del análisis estadístico, pues no creían que de esta manera el concepto de probabilidad pudiera aplicarse a la verosimilitud de ocurrencia de varios acontecimientos en un experimento sin repetición relativa.

La tercera escuela, también llamada de **probabilidad estadística** o de **regularidad estadística**; al estudiar el fenómeno en condiciones constantes (o casi), cuando el tamaño de la muestra tiende a infinito, las proporciones en que el fenómeno ocurre son muy estables. No se puede predecir el resultado al estudiar uno o pocos elementos, pero **si n es grande es posible la predicción, con un error mínimo**, “*Ley de los grandes números*” sustentada por Bernoulli (1713). Ésta implica una regularidad estadística; al estudiar un fenómeno aleatorio muchas veces, en condiciones casi constantes (población) los diferentes resultados ocurren con una proporción casi constante (estable) a esta proporción se le llama probabilidad. Es decir, que la constancia de proporciones es igual a las probabilidades. Quetelet la llama **regularidad estadística**, que es la base de la **probabilidad frecuentista**. Aunque la definición de los resultados de interés (espacio muestral) y las condiciones de estudio (población) es *subjetiva*; sin embargo, los valores en los que se estabilizan las frecuencias relativas o probabilidades son *objetivas*. Para entender, describir y predecir fenómenos aleatorios, se pretende conocer dichas probabilidades. En este *enfoque objetivo*, la variación impredecible de la evidencia es tal que la frecuencia relativa con la que pertenece a cualquier conjunto en el dominio de la evidencia, tiende a estabilizarse alrededor de un valor idealizado, la **probabilidad**, siempre y cuando el número de observaciones se hiciese infinitamente grande.

Debido a que el observador tiene una ignorancia total o parcial, es decir, una incertidumbre acerca del resultado de una investigación, su actitud estará eventualmente en desventaja al adivinar qué ocurrirá. En un sentido, la probabilidad refleja la incertidumbre acerca del resultado de un experimento, por lo que aquella puede pensarse como el lenguaje matemático de ésta, la cual es la misma para el investigador como para el juego de azar.

2.2

EL PAPEL DE LA PROBABILIDAD EN ESTADÍSTICA

La probabilidad y la estadística están relacionadas en una forma importante. La probabilidad se emplea como *herramienta*; permite que usted evalúe la confiabilidad de sus conclusiones acerca de la población cuando tenga sólo información muestral. Considere estas situaciones:

- Cuando lance al aire una sola moneda, verá ya sea cara (H) o cruz (T). Si lanza la moneda varias veces al aire, generará un número infinitamente grande de caras o cruces, es decir, toda la población. ¿Qué aspecto tiene esta población? Si la moneda es imparcial, entonces la población debe contener 50% de H y 50% de T. Ahora lance al aire la moneda una vez más. ¿Cuál es la probabilidad de que caiga una cara? Casi todos dirían que la oportunidad o “probabilidad” es $1/2$.
- Ahora suponga que no está usted seguro de que la moneda sea imparcial, esto es, no sabe con certeza si la composición de la población es 50–50 y decide hacer un experimento sencillo. Lanza al aire la moneda $n = 10$ veces y observa 10 caras consecutivas. ¿Concluiría que la moneda es imparcial? Es probable que no, porque si así fuera, observar 10 caras en fila sería muy *improbable*; esto es, la “probabilidad” sería muy pequeña. Es más *probable* que la moneda esté “cargada”.

Al igual que en el ejemplo de lanzar al aire una moneda, los expertos en estadística usan la probabilidad en dos formas. Cuando la población es *conocida*, se usa para describir la probabilidad de observar un resultado muestral en particular. Cuando la población es *desconocida* y sólo se dispone de una muestra de esa población, la probabilidad se usa para hacer enunciados acerca de la composición de la población, es decir, hacer inferencias estadísticas.

En los capítulos siguientes se verán numerosas formas diferentes para calcular probabilidades. Supondrá que la población es *conocida* y calculará la probabilidad de observar varios resultados muestrales. Una vez que empiece a usar la probabilidad para inferencia estadística en el capítulo 5, la población será *desconocida* y usted usará su conocimiento de probabilidad para hacer inferencias confiables a partir de información muestral. Empecemos con algunos ejemplos sencillos para ayudarle a captar conceptos básicos de probabilidad.

2.3

EVENTOS Y EL ESPACIO MUESTRAL

Se obtienen datos al observar ya sea eventos no controlados en la naturaleza o situaciones controladas en un laboratorio. Usamos el término **experimento** para describir cualquiera de los dos métodos de recolección de datos.

Definición Un **experimento** es el proceso mediante el cual se obtiene una observación (o medición).

La observación o medición generada por un experimento produce o no un valor numérico. A continuación veamos algunos ejemplos de experimentos:

- Registrar la calificación de un examen
- Medir la cantidad de lluvia diaria
- Entrevistar a un dueño de casa para obtener su opinión sobre un reglamento para distribuir por zonas un área verde

- Probar una tarjeta de circuito impreso para determinar si es un producto defectuoso o aceptable
- Lanzar al aire una moneda y observar el lado que aparece

Cuando se realiza un experimento, lo que observamos es un resultado llamado **evento simple**, con frecuencia denotado por la mayúscula E con un subíndice.

Definición Un **evento simple** es el resultado que se observa en una sola repetición del experimento.

EJEMPLO

2.1



Experimento: Lance un dado y observe el número que aparece en la cara superior. Haga una lista de los eventos sencillos del experimento.

Solución Cuando el dado se lanza una vez, hay seis posibles resultados. Hay los eventos sencillos citados a continuación:

Evento E_1 : observar un 1	Evento E_4 : observar un 4
Evento E_2 : observar un 2	Evento E_5 : observar un 5
Evento E_3 : observar un 3	Evento E_6 : observar un 6
-

Ahora podemos definir un **evento** como un conjunto de eventos sencillos, a menudo denotado por una letra mayúscula.

Definición Un **evento** es un conjunto de eventos sencillos.

EJEMPLO
continúa

2.1

Definimos los eventos A y B para el experimento de lanzar al aire un dado:

- A : observar un número impar
 B : observar un número menor que 4

Como el evento A se presenta si la cara superior es 1, 3 o 5, es un conjunto de tres eventos sencillos y escribimos $A = \{E_1, E_3, E_5\}$. Del mismo modo, el evento B ocurre si la cara superior es 1, 2 o 3 y está definido como una serie o conjunto de estos tres eventos sencillos: $B = \{E_1, E_2, E_3\}$.

A veces, cuando ocurre un evento, significa que no puede ocurrir otro.

Definición Dos eventos son **mutuamente excluyentes** si, cuando ocurre un evento, el otro no puede ocurrir y viceversa.

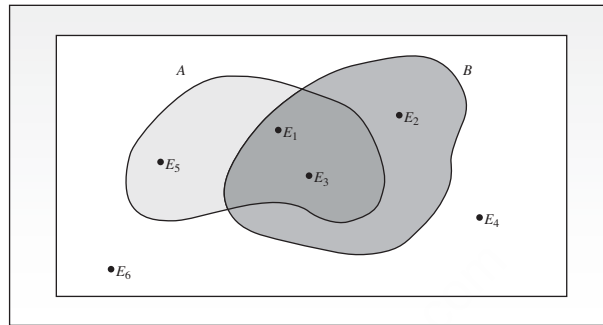
En el experimento de lanzar un dado, los eventos A y B *no* son mutuamente excluyentes, porque tienen dos resultados en común, si el número de la cara superior del dado es 1 o 3. Ambos eventos, A y B , ocurrirán si se observa E_1 o E_3 cuando se realiza el experimento. En contraste, los seis eventos simples $E_1, E_2, \dots, E_6$ forman un conjunto de todos los resultados mutuamente excluyentes del experimento. Cuando el experimento se realiza una vez, puede ocurrir uno y sólo uno de estos eventos sencillos.

Definición El conjunto de todos los eventos sencillos se denomina espacio muestral, S .

En ocasiones es útil visualizar un experimento usando una imagen llamada **diagrama de Venn**, que se ilustra en la figura 2.1. La caja exterior representa el *espacio muestral*, que contiene todos los *eventos sencillos*, representados por puntos marcados. Como un evento es un conjunto de uno o más eventos sencillos, los puntos apropiados están circulados y marcados con la letra del evento. Para el experimento de lanzar un dado, el espacio muestral es $S = \{E_1, E_2, E_3, E_4, E_5, E_6\}$ o bien, de un modo más simple, $S = \{1, 2, 3, 4, 5, 6\}$. Los eventos $A = \{1, 3, 5\}$ y $B = \{1, 2, 3\}$ están circulados en el diagrama de Venn.

FIGURA 2.1

Diagrama de Venn para el tiro de un dado

**EJEMPLO**

2.2

Experimento: Lance al aire una sola moneda y observe el resultado. Éstos son los eventos sencillos:



E_1 : observar una cara (H)

E_2 : observar una cruz (T)

El espacio muestral es $S = \{E_1, E_2\}$, o bien, en forma más sencilla, $S = \{H, T\}$.

EJEMPLO

2.3

Experimento: Registre el tipo de sangre de una persona. Los cuatro posibles resultados mutuamente excluyentes son estos eventos sencillos:

E_1 : sangre tipo A

E_2 : sangre tipo B

E_3 : sangre tipo AB

E_4 : sangre tipo O

El espacio muestral es $S = \{E_1, E_2, E_3, E_4\}$, o $S = \{A, B, AB, O\}$.

Algunos experimentos se generan en etapas y el espacio muestral se representa en un **diagrama de árbol**. Cada nivel de ramificación sucesivo del árbol corresponde a un paso requerido para generar el resultado final.

EJEMPLO

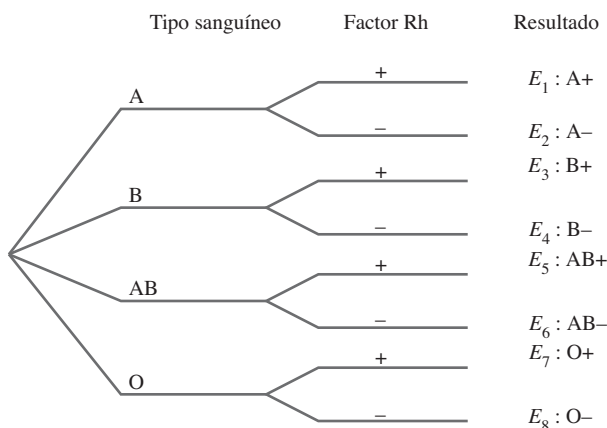
2.4

Un médico registra el tipo sanguíneo y factor Rh de una persona. Haga una lista de los eventos sencillos del experimento.

Solución Por cada persona, es necesario un procedimiento de dos etapas para registrar las dos variables de interés. El diagrama de árbol se muestra en la figura 2.2. Los ocho eventos sencillos del diagrama de árbol forman el espacio muestral, $S = \{A+, A-, B+, B-, AB+, AB-, O+, O-\}$

FIGURA 2.2

Diagrama de árbol para el ejemplo 2.4



Una forma alternativa para exhibir los eventos sencillos es usar una **tabla de probabilidad**, como se ilustra en la tabla 2.1. Los renglones y las columnas registran los posibles resultados en las etapas primera y segunda, respectivamente, y los eventos sencillos aparecen en las celdas de la tabla.

TABLA 2.1**Tabla de probabilidad para el ejemplo 2.4**

Factor Rh	Tipo sanguíneo			
	A	B	AB	O
Negativo	A-	B-	AB-	O-
Positivo	A+	B+	AB+	O+

2.4

ENFOQUES O ESCUELAS DE LA PROBABILIDAD

Aún no existe una interpretación única para definir *probabilidad*. Los estadísticos, filósofos y científicos en la materia no han podido homogeneizar el concepto, por lo que existen tres interpretaciones más usuales:

- Frecuencia relativa*. Este enfoque es una aproximación empírica del concepto de probabilidad; existe cuando se tiene un número muy grande de observaciones o repeticiones del mismo fenómeno o evento. A continuación, se enunciará una definición un poco más formal.

Si un experimento se ejecuta n veces en las mismas condiciones y hay x resultados, $x \leq n$ en que ocurrió un evento (hecho, suceso), entonces una estimación de la probabilidad de ese evento es la razón x/n , siempre que $n \rightarrow \infty$ (n es grande).

Algunos autores manejan el concepto de eventos “equiprobables” para decir que los posibles resultados de un experimento son igualmente probables o tienen las mismas condiciones.

EJEMPLO

2.5

La probabilidad de obtener águila al tirar una moneda es igual a la razón del número de modos en los que puede ocurrir el evento (uno: un águila) al número total de formas que el resultado del experimento pueda acontecer (dos: águila o sol), siempre que éste se repita n veces ($n \rightarrow \infty$).

Así pues, la probabilidad de que caiga águila es $\frac{1}{2}$. Es importante notar que la posibilidad obtenida al usar esta regla se refiere a la frecuencia de ocurrencia.

- b) *Teoría clásica*. La definición clásica de probabilidad fue dada por Laplace y desde entonces se ha repetido en casi todos los libros sobre teoría de probabilidades. En su forma primitiva dice: *probabilidad es la razón del número de casos favorables al número total de casos igualmente posibles*.

Algunos autores manejan el concepto de probabilidad *a priori* (antes de); es decir, que se determina una vez que se conoce la naturaleza del experimento.

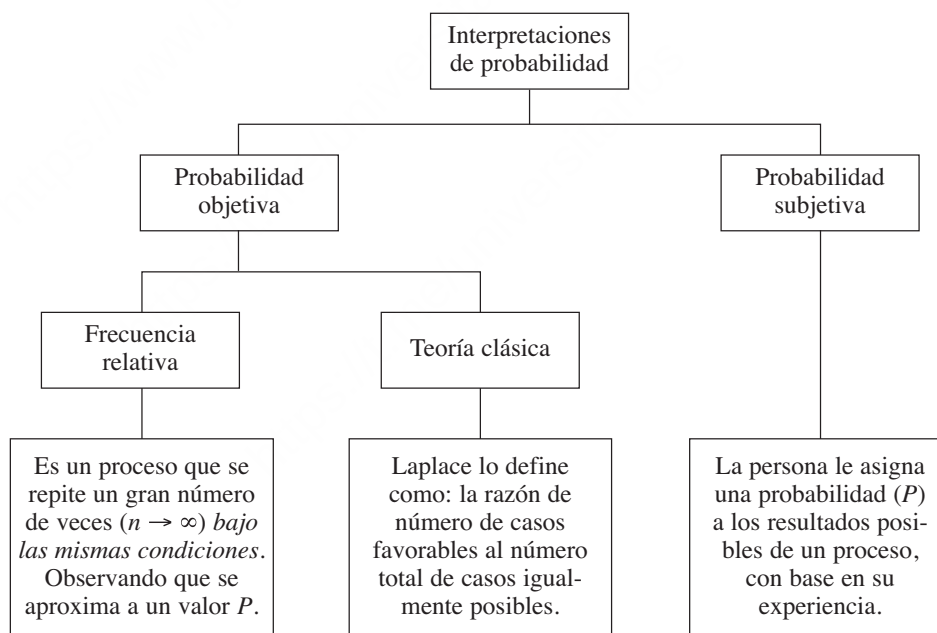
- c) *Probabilidad subjetiva*. Las probabilidades obtenidas mediante el enfoque de frecuencia relativa reciben el nombre de *probabilidades objetivas*, ya que se derivan de hechos. Sin embargo, hay numerosas situaciones en las que no se puede emplear el enfoque objetivo, es decir, situaciones en las que los resultados no son igualmente probables y no se dispone fácilmente de datos históricos. En este caso, se debe hacer una evaluación *subjetiva* de probabilidad.

Las probabilidades subjetivas son el resultado de un esfuerzo por cuantificar las expectativas o creencias respecto a algo.

En suma, la probabilidad subjetiva es una evaluación personal de la posibilidad de *que* ocurra un evento. (Para una mejor apreciación de las definiciones de probabilidad, véase la figura 2.3.)

FIGURA 2.3

Esquema de las definiciones de probabilidad



2.5

AXIOMAS DE PROBABILIDAD

A todo evento en el espacio muestral se le asigna una probabilidad, que se da en términos de números reales, y debe cumplir con las siguientes condiciones:

- i) $P(E) \geq 0$, para toda $E \subset \Omega$.
No existen probabilidades negativas.

ii) $P(E \cup F) = P(E) + P(F)$ para todo E y $F \subset \Omega$ y $E \cap F = \emptyset$.

La probabilidad del conjunto universal es igual a uno, o sea, la suma de todas y cada

una de las probabilidades de que se compone el espacio muestral, $\sum_{i=1}^n P(E_i) = 1$.

iii) $P(\Omega) = 1$ para todo Ω .

La probabilidad de ocurrencia de un evento deberá estar siempre en un intervalo $[0, 1]$.

EJEMPLO

2.6

Se lanza una vez un dado no cargado.[†]

El espacio muestral es:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Es el conjunto de todos los posibles resultados (Ω).

Sea E_1 = Obtener un 1 en la tirada.

$P(E_1)$ = Probabilidad de obtener un 1 en la tirada.

$$P(E_1) = \frac{1}{6}$$

Como E_1 consta de un solo elemento del espacio muestral, se le considera un *evento simple*.

Al considerar todo el espacio muestral, se tiene:

$$\sum_{i=1}^6 P(E_i) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{6}{6} = 1$$

$P(E_2)$ = Probabilidad de obtener un 2.

Hasta $P(E_6)$ = Probabilidad de obtener un 6.

De ahí que el experimento sí cumple con los axiomas de probabilidad.

Sea E_7 = Obtener un 2 o un 5 en la tirada.

$P(E_7)$ = Probabilidad de obtener un 2 o un 5 en la tirada de un dado no cargado:

$$P(E_7) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

Como E_7 está formado por los elementos del espacio muestral, entonces se le llama *evento compuesto*.

Teoremas: $P(\emptyset) = 0$ y $P(E^c) = 1 - P(E)$

Ahora, se desea calcular la siguiente probabilidad:

$P(\overline{E_7}) = P(E_7^c) = P(E_7')$ = Probabilidad de no obtener un 2 o un 5 en la tirada de un dado no cargado:

$$P(E_7^c) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$$

[†] Con $n \rightarrow \infty$ a priori, subjetivamente se da por equiprobable.

Es decir, excluya los eventos E_2 y E_5

$$P(E_2) = \frac{1}{6}$$

$$P(E_5) = \frac{1}{6}$$

Una forma alterna de solución es:

$$P(E'_7) = 1 - P(E_7)$$

$$= 1 - \frac{1}{3} = \frac{3}{3} - \frac{1}{3} = \frac{2}{3}$$

$$\Rightarrow P(E'_7) = \frac{2}{3}$$

2.6

PARTICIONES

La cardinalidad de un conjunto $n(E)$ = número de elementos de E ; $n(\emptyset) = 0$.

Sean $A_1, A_2, A_3, \dots, A_n$ subconjuntos del conjunto Ω . La colección de subconjuntos, $A_1, A_2, \dots, A_n$ forman una partición de Ω si y sólo si

- i) $A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n = \Omega$
- ii) $A_i \cap A_j = \emptyset$ para toda $i \neq j$

Para mayor claridad, la definición anterior se analizará con este ejemplo:

Una empresa está constituida por 200 empleados, de los cuales 110 son hombres y 90 mujeres.

Ω = Todos los empleados de la empresa, $n(\Omega) = 200$

A_1 = Todos los hombres que trabajan en la compañía $n(A_1) = 110$

A_2 = Todas las mujeres que trabajan en la compañía $n(A_2) = 90$

Ahora, se verificará que en verdad A_1 y A_2 forman una partición de Ω .

- i) $A_1 \cup A_2 = \Omega$
 $\{\text{hombres}\} \cup \{\text{mujeres}\} = \Omega$
 sí cumple i).

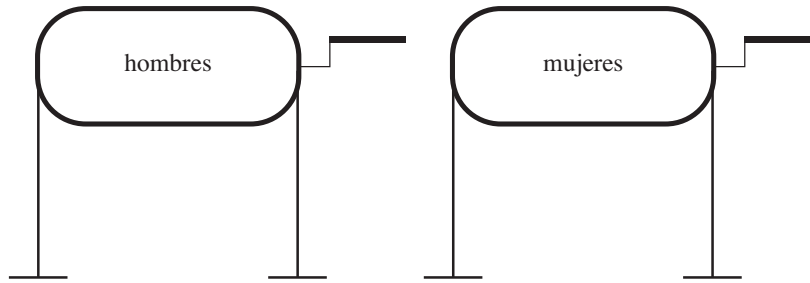
- ii) $A_1 \cap A_2 = \emptyset$
 $n(A_1 \cup A_2) = n(A_1) + n(A_2) - n(A_1 \cap A_2) = 110 + 90 - 0 = 200$

Como no existen elementos en común, la intersección es igual a vacío.

Así, los subconjuntos A_1 y A_2 son una partición del conjunto Ω . Es importante hacer notar que si no cumple una de las dos condiciones, deja de ser partición.

Suponga que en la fiesta de fin de año se desea dar premios a algunos de los empleados. Para ello se tienen dos tómbolas, una para hombres y otra para mujeres (figura 2.4).

FIGURA 2.4



Sean $P(A_1)$ = Probabilidad de que un hombre obtenga el premio

$$P(A_1) = \frac{1}{110}$$

$P(A_2)$ = Probabilidad de que una mujer obtenga el premio

$$P(A_2) = \frac{1}{90}$$

La definición y el ejemplo explicado en la presente sección, constituyen la base para entender algunos conceptos que se analizarán más adelante.

Si realiza una sola rifa, entonces la probabilidad que salga premiado un hombre es $\left(\frac{110}{200}\right)$ y la probabilidad de que salga premiada una mujer es $\left(\frac{90}{200}\right)$:

$$\left(\frac{90}{200}\right) + \left(\frac{110}{200}\right) = \frac{200}{200} = 1$$

CÁLCULO DE PROBABILIDADES CON EL USO DE EVENTOS SENCILLOS

2.7

La probabilidad de un evento A es una medida de nuestra creencia de que el evento A ocurrirá. Una manera práctica de interpretar esta medida es con el concepto de *frecuencia relativa*. Recuerde del capítulo 1 que si un experimento se realiza n veces, entonces la frecuencia relativa de un suceso particular, por ejemplo A , es

$$\text{Frecuencia relativa} = \frac{\text{Frecuencia}}{n}$$

donde la frecuencia es el número de veces que ocurre el evento A . Si hacemos que el número n de repeticiones del experimento se haga cada vez más grande ($n \rightarrow \infty$), en última instancia se generará toda la población. En ésta, la frecuencia relativa del evento A se define como la **probabilidad del evento A** ; esto es,

$$P(A) = \lim_{n \rightarrow \infty} \frac{\text{Frecuencia}}{n}$$

Puesto que $P(A)$ se comporta como una frecuencia relativa, $P(A)$ debe ser una proporción que se encuentre entre 0 y 1; $P(A) = 0$ si el evento A nunca ocurre, y $P(A) = 1$ si el evento A ocurre siempre. Cuanto más cercano sea $P(A)$ a 1, es más probable que A ocurra.



Por ejemplo, si se lanza un dado balanceado de seis caras un número infinito de veces, se esperaría que la frecuencia relativa para cualquiera de los seis valores, $x = 1, 2, 3, 4, 5, 6$, fuera $1/6$. Sobra decir que sería muy lento, si no imposible, repetir un experimento un número infinito de veces. Por esta razón, hay métodos alternativos para calcular probabilidades que hacen uso del concepto de frecuencia relativa.

Una consecuencia importante de la definición de frecuencia relativa de una probabilidad involucra a eventos sencillos. Como los eventos sencillos son mutuamente excluyentes, sus probabilidades deben satisfacer dos condiciones.

REQUISITOS PARA PROBABILIDADES DE UN EVENTO SIMPLE

- Cada probabilidad debe estar entre 0 y 1, inclusive.
- La suma de las probabilidades de todos los eventos sencillos en S es igual a 1.

Quando es posible escribir los eventos sencillos asociados con un experimento y determinar sus probabilidades respectivas, podemos hallar la probabilidad de un evento A como sigue:

Definición La **probabilidad de un evento A** es igual a la suma de las probabilidades de los eventos sencillos contenidos en A .

EJEMPLO 2.7

Lance al aire dos monedas imparciales y registre el resultado. Encuentre la probabilidad de observar exactamente una cara en los dos tiros.

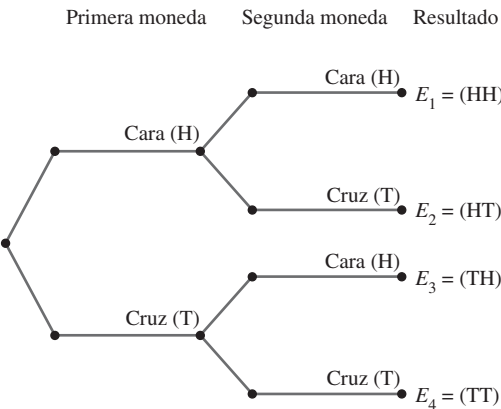


TIP MI CONSEJO
Las probabilidades deben estar entre 0 y 1

Solución Para poner en una lista los eventos sencillos en el espacio muestral, se usa un diagrama de árbol como se presenta en la figura 2.5. Las letras H y T significan que usted observó una cara (H) o una cruz (T), respectivamente, en un tiro en particular. Para asignar probabilidades a cada uno de los cuatro eventos sencillos, hay que recordar que las monedas son imparciales. Por tanto, cualquiera de los cuatro eventos sencillos es tan probable como cualquier otro. Como la suma de los cuatro eventos sencillos debe ser 1 cada uno debe tener una probabilidad $P(E_i) = 1/4$. Los eventos sencillos del espacio muestral se muestran en la tabla 2.2, junto con sus *probabilidades igualmente posibles*. Para hallar $P(A) = P(\text{observar exactamente una cara})$, es necesario hallar todos los eventos sencillos que resulten en el evento A , es decir E_2 y E_3 :

$$\begin{aligned} P(A) &= P(E_2) + P(E_3) \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \end{aligned}$$

FIGURA 2.5
Diagrama de árbol para el ejemplo 2.7



TIP MI CONSEJO
Las probabilidades de todos los eventos sencillos deben totalizar 1

TABLA 2.2

Eventos sencillos y sus probabilidades

Evento	Primera moneda	Segunda moneda	$P(E_i)$
E_1	H	H	$1/4$
E_2	H	T	$1/4$
E_3	T	H	$1/4$
E_4	T	T	$1/4$

EJEMPLO

2.8

Las proporciones de fenotipos sanguíneos A, B, AB y O en la población de raza caucásica en Estados Unidos se publican como .40, .11, .04 y .45, respectivamente.¹ Si se elige al azar una persona de este origen étnico en la población, ¿cuál es la probabilidad de que tenga sangre tipo A o tipo AB?

Solución Los cuatro eventos sencillos, A, B, AB y O *no* tienen probabilidades igualmente posibles. Sus probabilidades se encuentran usando el concepto de frecuencia relativa como

$$P(A) = .40 \quad P(B) = .11 \quad P(AB) = .04 \quad P(O) = .45$$

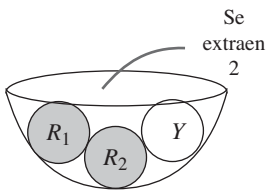
El evento de interés está formado por dos eventos sencillos, de modo que

$$\begin{aligned} P(\text{la persona es tipo A o tipo AB}) &= P(A) + P(AB) \\ &= .40 + .04 = .44 \end{aligned}$$

EJEMPLO

2.9

Un plato contiene un dulce amarillo y dos rojos. Usted cierra los ojos, elige dos dulces del plato, uno por uno y anota sus colores. ¿Cuál es la probabilidad de que ambos dulces sean rojos?



Solución Como no se dan probabilidades se debe hacer una lista de los eventos sencillos del espacio muestral. La selección de los dulces en dos etapas sugiere un diagrama de árbol, que se muestra en la figura 2.6. Hay dos dulces rojos en el plato, de modo que se usan las letras R_1 , R_2 y Y para indicar que se ha seleccionado el primero rojo, el segundo rojo o el dulce amarillo, respectivamente. Como usted cerró los ojos cuando eligió los dulces, las seis opciones deben ser *igualmente probables* y se les asigna la probabilidad $1/6$. Si A es el evento de que ambos dulces sean rojos, entonces

$$A = \{R_1R_2, R_2R_1\}$$

Entonces,

$$\begin{aligned} P(A) &= P(R_1R_2) + P(R_2R_1) \\ &= \frac{1}{6} + \frac{1}{6} = \frac{1}{3} \end{aligned}$$

FIGURA 2.6

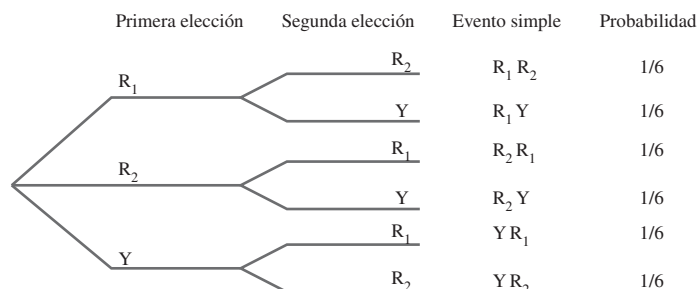
Diagrama de árbol para el ejemplo 2.9

TIP MI CONSEJO

Un diagrama de árbol ayuda a hallar eventos sencillos

Rama = paso hacia el resultado

Ramas siguientes $\Rightarrow$ lista de eventos sencillos





NECESITO SABER...

Cómo calcular la probabilidad de un evento

1. Haga una lista de todos los eventos sencillos del espacio muestral.
2. Asigne una probabilidad apropiada a cada evento simple.
3. Determine cuáles eventos sencillos resultan en el evento de interés.
4. Sume las probabilidades de los eventos sencillos que resulten en el evento de interés.

En su cálculo, siempre debe tener cuidado de satisfacer estas dos condiciones:

- Incluir todos los eventos sencillos en el espacio muestral.
- Asignar probabilidades realistas a los eventos sencillos.

Cuando el espacio muestral es grande, es fácil omitir sin intención algunos de los eventos sencillos. Si esto ocurre, o si sus probabilidades asignadas son erróneas, sus respuestas no serán útiles en la práctica.

Una forma de determinar el número requerido de eventos sencillos es usar las reglas de conteo presentadas en la siguiente sección. Estas reglas se aplican para resolver problemas más complejos, que generalmente comprenden un gran número de eventos sencillos.

2.7 EJERCICIOS

TÉCNICAS BÁSICAS

2.1 Tiro de un dado Un experimento consiste en tirar un solo dado. Éstos son algunos eventos:

- A: observar un 2
- B: observar un número par
- C: observar un número mayor que 2
- D: observar A y B
- E: observar A o B o ambos
- F: observar A y C

- a. Haga una lista de eventos sencillos del espacio muestral.
- b. Haga una lista de eventos sencillos en cada uno de los eventos A al F.
- c. ¿Qué probabilidades debe asignar a los eventos sencillos?
- d. Calcule las probabilidades de los seis eventos A al F sumando las probabilidades apropiadas de evento simple.

2.2 Un espacio muestral S está formado por cinco eventos sencillos con estas probabilidades:

$$P(E_1) = P(E_2) = .15 \quad P(E_3) = .4 \\ P(E_4) = 2P(E_5)$$

- a. Encuentre las probabilidades para los eventos sencillos E_4 y E_5 .
- b. Encuentre las probabilidades para estos dos eventos:
 $A = \{E_1, E_3, E_4\}$
 $B = \{E_2, E_3\}$
- c. Haga una lista de eventos sencillos que se encuentren en el evento A o en el evento B o en ambos.
- d. Haga una lista de eventos sencillos que se encuentren en el evento A y en el B.

2.3 Un espacio muestral contiene 10 eventos sencillos: $E_1, E_2, \dots, E_{10}$. Si $P(E_1) = 3P(E_2) = .45$ y los eventos sencillos restantes son igualmente probables, encuentre las probabilidades de estos eventos restantes.

2.4 Tiros libres Una jugadora de baloncesto acierta en 70% de sus tiros libres. Cuando lanza un par de tiros libres, los cuatro eventos sencillos posibles y tres de sus probabilidades asociadas se dan en la tabla:

Evento simple	Resultado del primer tiro libre	Resultado del segundo tiro libre	Probabilidad
1	Encesta	Encesta	.49
2	Encesta	Falla	?
3	Falla	Encesta	.21
4	Falla	Falla	.09

- Encuentre la probabilidad de que la jugadora enceste en el primer tiro y falle en el segundo.
- Encuentre la probabilidad de que la jugadora enceste en al menos uno de los dos tiros libres.

2.5 Cuatro monedas Un frasco contiene cuatro monedas: una de cinco, una de 10, una de 25 y una de 50 centavos. Se seleccionan al azar tres monedas del frasco.

- Haga una lista de los eventos simples en S .
- ¿Cuál es la probabilidad de que la selección contenga la moneda de 50 centavos?
- ¿Cuál es la probabilidad de que la suma total sacada sea igual a 60 centavos o más?

2.6 ¿Preescolar o no? El primer día de clase de jardín de niños, el maestro selecciona al azar uno de sus 25 estudiantes y registra su género, y también si había asistido a preescolar.

- ¿Cómo describiría usted el experimento?
- Construya un diagrama de árbol para este experimento. ¿Cuántos eventos simples hay ahí?
- La tabla siguiente muestra la distribución de los 25 estudiantes de acuerdo con su género y experiencia preescolar. Use la tabla para asignar probabilidades a los eventos simples de la parte **b**.

	Hombre	Mujer
Preescolar	8	9
Sin preescolar	6	2

- ¿Cuál es la probabilidad de que el estudiante seleccionado al azar sea hombre? ¿Cuál es la probabilidad de que sea mujer y no haya asistido a preescolar?

2.7 El problema de la urna Un tazón contiene tres pelotas rojas y dos amarillas. Dos de ellas se seleccionan al azar y se registran sus colores. Use un diagrama de árbol para hacer una lista de los 20 eventos simples del experimento, considerando el orden en el que se sacan las pelotas.

2.8 El problema de la urna, continúa Consulte el ejercicio 2.7. Se selecciona al azar una pelota del tazón que contiene tres pelotas rojas y dos amarillas. Se anota su color y se devuelve al tazón antes de seleccionar una segunda pelota. Haga una lista de los otros cinco eventos simples que deben agregarse al espacio muestral del ejercicio 2.7.

APLICACIONES

2.9 ¿Necesita anteojos? Un estudio clasificó un gran número de adultos de acuerdo con si se determinó que necesitaban anteojos para corregir su vista para leer y si los usaban cuando leían. Las proporciones que caen en las cuatro categorías se muestran en la tabla siguiente. (Observe que una pequeña proporción, .02, de adultos usaba anteojos cuando, de hecho, no se determinó que los necesitaran.)

Se determinó que necesitaban anteojos	Usaba anteojos para leer	
	Sí	No
Sí	.44	.14
No	.02	.40

Si se selecciona un solo adulto de este grupo grande, encuentre la probabilidad de cada evento:

- Se determinó que el adulto necesita anteojos.
- El adulto necesita anteojos para leer pero no los usa.
- El adulto usa anteojos para leer, los necesite o no.

2.10 Ruleta El juego de ruleta usa una rueda que contiene 38 buchacas. Treinta y seis buchacas numeradas 1, 2,..., 36 y las dos restantes están marcadas 0 y 00. Se hace girar la rueda y una buchaca se identifica como la “ganadora”. Suponga que la observancia de una buchaca es igualmente probable que cualquier otra.

- Identifique los eventos simples en un solo giro de la rueda de la ruleta.
- Asigne probabilidades a los eventos simples.
- Sea A el evento que usted observa ya sea 0 o 00. Haga una lista de los eventos simples del evento A y encuentre $P(A)$.
- Suponga que usted apostó a los números del 1 al 18. ¿Cuál es la probabilidad de que uno de sus números sea el ganador?

2.11 Miembros de un jurado Tres personas son seleccionadas al azar para reportarse como miembros de un jurado. El presidente del jurado toma nota del género de cada persona.

- a. Defina el experimento.
- b. Haga una lista de los eventos simples en S .
- c. Si es igualmente probable que cada persona sea hombre o mujer, ¿qué probabilidad le asigna usted a cada evento simple?
- d. ¿Cuál es la probabilidad de que sólo una de las tres sea hombre?

e. ¿Cuál es la probabilidad de que las tres sean mujeres?

2.12 Miembros de un jurado II

Consulte el ejercicio 2.11. Suponga que hay seis prospectos para miembros de jurado, cuatro hombres y dos mujeres, que podrían ser elegidos para ocupar un asiento en el jurado en un caso penal. Se seleccionan al azar dos miembros de jurado de estos seis para ocupar los dos asientos restantes del jurado.

- a. Haga una lista de los eventos simples del experimento. (SUGERENCIA: Hay 15 eventos simples si se ignora el orden de selección de los dos miembros de jurado.)
- b. ¿Cuál es la probabilidad de que ambos miembros de jurado elegidos sean mujeres?

2.13 Probadores de té Una compañía de alimentos planea efectuar un experimento para comparar su marca de té con la de dos competidores. Se contrata una sola persona para probar y clasificar cada una de las tres marcas de té, que no tienen marca excepto por símbolos de identificación A , B y C .

- a. Defina el experimento.
- b. Haga una lista de eventos simples en S .
- c. Si el probador no tiene capacidad para distinguir una diferencia en gusto entre los té, ¿cuál es la probabilidad de que clasifique el té tipo A como el más deseable? ¿Como el menos deseable?

2.14 Carrera de 100 metros Cuatro corredores igualmente calificados, A , B , C y D , corren un *sprint* de 100 metros y se registra el orden de llegadas.

- a. ¿Cuántos eventos simples hay en el espacio muestral?
- b. Si los corredores están igualmente calificados, ¿qué probabilidad debe usted asignar a cada evento simple?
- c. ¿Cuál es la probabilidad de que D gane la carrera?
- d. ¿Cuál es la probabilidad de que D gane y A se coloque en segundo lugar?
- e. ¿Cuál es la probabilidad de que C termine en último lugar?

2.15 Moscas de la fruta En un experimento de genética, el investigador apareó dos moscas de la fruta *Drosophila* y observó los rasgos de 300 descendientes. Los resultados se muestran en la tabla.

Color de ojos	Tamaño de alas	
	Normal	Miniatura
Normal	140	6
Bermellón	3	151

Se selecciona al azar uno de estos descendientes y se le observan los dos rasgos genéticos.

- a. ¿Cuál es la probabilidad de que la mosca tenga color normal de ojos y tamaño normal de alas?
- b. ¿Cuál es la probabilidad de que la mosca tenga ojos bermellón?
- c. ¿Cuál es la probabilidad de que la mosca tenga ojos bermellón o alas miniatura, o ambos?

2.8

REGLAS ÚTILES DE CONTEO

Suponga que un experimento comprende un gran número N de eventos simples y que usted sabe que todos esos eventos son *igualmente probables*. Entonces cada evento simple tiene una probabilidad $1/N$ y la probabilidad de un evento A se calcula como

$$P(A) = \frac{n_A}{N}$$

donde n_A es el número de eventos simples que resultan en el evento A . En esta sección presentamos tres reglas sencillas que se usan para contar ya sea N , el número de eventos simples del espacio muestral, o n_A , el número de eventos simples del evento A . Una vez que haya obtenido estas cuentas, puede hallar $P(A)$ sin en realidad hacer una lista de todos los eventos simples.

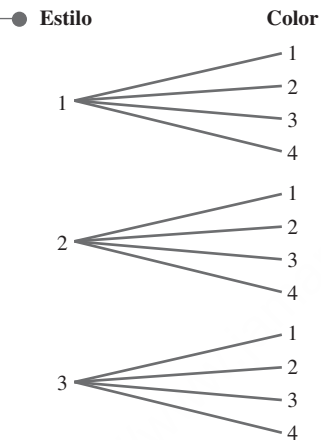
LA REGLA mn

Considere un experimento que se realiza en dos etapas. Si la primera etapa se efectúa en m formas y, para cada una de éstas, la segunda etapa se logra en n formas, entonces hay mn formas para efectuar el experimento.

Por ejemplo, supongamos que usted ordena un automóvil en uno de tres estilos y en uno de cuatro colores de pintura. Para averiguar cuántas opciones hay disponibles, considere primero elegir uno de los $m = 3$ estilos y luego seleccionar uno de los $n = 4$ colores de pintura. Con el uso de la regla mn , como se muestra en la figura 2.7, tiene $mn = (3)(4) = 12$ posibles opciones.

FIGURA 2.7

Combinaciones de estilo y color

**EJEMPLO**

2.10

Se tiran dos dados. ¿Cuántos eventos simples hay en el espacio muestral S ?



APPLET EN LÍNEA

Tirar dados

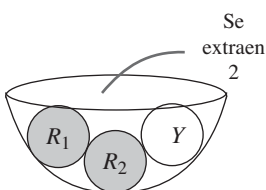
Solución El primer dado puede caer en una de $m = 6$ formas, y el segundo en una de $n = 6$ formas. Como el experimento comprende dos etapas, que forman los pares de números que se muestran en las dos caras, el número total de eventos simples en S es

$$mn = (6)(6) = 36$$

EJEMPLO

2.11

Un plato contiene un dulce amarillo y dos rojos. Se seleccionan dos dulces del plato, uno por uno, registrando sus colores. ¿Cuántos eventos simples hay en el espacio muestral S ?



Solución El primer dulce se elige en $m = 3$ formas. Como un dulce ya no está ahora, el segundo dulce se elige en $n = 2$ formas. El número total de eventos simples es

$$mn = (3)(2) = 6$$

Estos seis eventos simples aparecen en el ejemplo 2.9.

Podemos extender la regla mn para un experimento que se realiza en más de dos etapas.

LA REGLA *mn* EXTENDIDA

Si un experimento se realiza en k etapas, con n_1 formas para efectuar la primera etapa, n_2 formas para efectuar la segunda etapa,..., y n_k formas para efectuar la k -ésima etapa, entonces el número de formas para efectuar el experimento es

$$n_1 n_2 n_3 \cdots n_k$$

EJEMPLO

2.12

¿Cuántos eventos simples hay en el espacio muestral cuando se lanzan al aire tres monedas?



Solución Cada moneda puede caer en una de dos formas. Por tanto, el número de eventos simples es

$$(2)(2)(2) = 8$$

EJEMPLO

2.13

El conductor de un camión podría tomar tres rutas de la ciudad A a la ciudad B , cuatro de la ciudad B a la C y tres de la ciudad C a la D . Si, cuando viaja de A a D , el conductor debe ir de A a B a C a D , ¿cuántas rutas posibles de A a D hay?

Solución Sean

$$n_1 = \text{número de rutas de } A \text{ a } B = 3$$

$$n_2 = \text{número de rutas de } B \text{ a } C = 4$$

$$n_3 = \text{número de rutas de } C \text{ a } D = 3$$

Entonces, el número total de formas para construir una ruta completa, tomando una alterna desde cada uno de los tres grupos, $(A \text{ a } B)$, $(B \text{ a } C)$ y $(C \text{ a } D)$, es

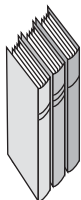
$$n_1 n_2 n_3 = (3)(4)(3) = 36$$

Una segunda y útil regla de conteo se sigue de la regla *mn* y comprende **ordenamientos** o **permutaciones**. Por ejemplo, suponga que usted tiene tres libros, A , B y C , pero tiene espacio sólo para dos en su estante. ¿En cuántas formas puede usted seleccionar y acomodar los dos libros? Hay tres opciones para los dos libros, A y B , A y C , o B y C , pero cada uno de los pares puede acomodarse en dos formas en el estante. Todas las permutaciones de los dos libros, seleccionados de tres, aparecen en la tabla 2.3. Entonces la Regla *mn* implica que hay seis formas, porque el primer libro se elige en $m = 3$ formas y el segundo en $n = 2$ formas, de modo que el resultado es $mn = 6$.

TABLA 2.3

Permutaciones de dos libros seleccionados de tres

Combinaciones de dos	Reordenamiento de combinaciones
AB	BA
AC	CA
BC	CB



¿En cuántas formas puede usted acomodar los tres libros en su estante? Hay las seis permutaciones:

$$\begin{array}{lll} ABC & ACB & BAC \\ BCA & CAB & CBA \end{array}$$

Como el primer libro se elige en $n_1 = 3$ formas, el segundo en $n_2 = 2$ formas, y el tercero en $n_3 = 1$ forma, el número total de ordenamientos es $n_1 n_2 n_3 = (3)(2)(1) = 6$.

En lugar de aplicar la regla mn cada vez, usted puede hallar el número de ordenamientos usando una fórmula general que involucra una *notación factorial*.

UNA REGLA DE CONTEO PARA PERMUTACIONES

El número de formas en que podemos acomodar n objetos distintos, tomándolos una cantidad r a la vez, es

$$P_r^n = \frac{n!}{(n-r)!}$$

donde $n! = n(n-1)(n-2) \cdots (3)(2)(1)$ y $0! \doteq 1$.

Debido a que se seleccionan r objetos, éste es un experimento de r -etapas. El primer objeto se elige en n formas, el segundo en $(n-1)$ formas, el tercero en $(n-2)$ formas y el r -ésimo en $(n-r+1)$ formas. Podemos simplificar esta engorrosa notación usando la regla de conteo para permutaciones porque

$$\begin{aligned} \frac{n!}{(n-r)!} &= \frac{n(n-1)(n-2) \cdots (n-r+1)(n-r) \cdots (2)(1)}{(n-r) \cdots (2)(1)} \\ &= n(n-1) \cdots (n-r+1) \end{aligned}$$

UN CASO ESPECIAL: ORDENAR n OBJETOS

El número de formas para ordenar todo un conjunto de n objetos distintos es $P_n^n = n!$

EJEMPLO

2.14

Tres billetes de lotería se extraen de un total de 50. Si los billetes se distribuyeran a cada uno de tres empleados en el orden en que se sacan, el orden será importante. ¿Cuántos eventos simples están asociados con el experimento?

Solución El número total de eventos simples es

$$P_3^{50} = \frac{50!}{47!} = 50(49)(48) = 117\,600$$

EJEMPLO

2.15

Una máquina está compuesta de cinco partes que se ensamblan en cualquier orden. Se realiza una prueba para determinar el tiempo necesario para cada orden de ensamblaje. Si cada orden se proba una vez, ¿cuántas pruebas deben efectuarse?

Solución El número total de pruebas es

$$P_5^5 = \frac{5!}{0!} = 5(4)(3)(2)(1) = 120$$

Cuando contamos el número de permutaciones de los dos libros elegidos para su librero, empleamos un método sistemático:

- Primero contamos el número de *combinaciones* o pares de libros a elegir.
- A continuación contamos el número de formas para ordenar en el librero los dos libros elegidos.

A veces el orden o acomodo de los objetos no es importante, sino sólo los objetos que se eligen. En este caso, se usa una regla de conteo para **combinaciones**. Por ejemplo, quizá no nos importe el orden en que los libros se coloquen en el librero, sino sólo cuáles libros podemos poner en el librero. Cuando se selecciona una comisión de cinco personas de un grupo de 12 estudiantes, el orden de la selección no es importante porque los cinco estudiantes serán miembros iguales de la comisión.

UNA REGLA DE CONTEO PARA COMBINACIONES

El número de combinaciones distintas de n objetos diferentes que se forma, tomando r de ellos a un tiempo, es

$$C_r^n = \frac{n!}{r!(n-r)!}$$

El número de *combinaciones* y el número de *permutaciones* están relacionados:

$$C_r^n = \frac{P_r^n}{r!}$$

Se observa que C_r^n resulta cuando se divide el número de permutaciones entre $r!$, el número de formas de reacomodar cada grupo distinto de r objetos elegidos del total n .

EJEMPLO

2.16

Se compra una tarjeta de circuito impreso elegida entre cinco proveedores. ¿En cuántas formas se pueden elegir tres proveedores entre los cinco?

Solución Como es sólo importante saber cuáles tres se han elegido, no el orden de selección, el número de formas es

$$C_3^5 = \frac{5!}{3!2!} = \frac{(5)(4)}{2} = 10$$

El siguiente ejemplo ilustra el uso de reglas de conteo para resolver un problema de probabilidad.

EJEMPLO

2.17

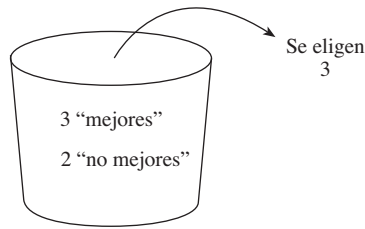
Cinco fabricantes producen cierto aparato electrónico, cuya calidad varía de un fabricante a otro. Si seleccionáramos al azar tres fabricantes, ¿cuál es la probabilidad de que la selección contenga exactamente dos de los tres mejores?

Solución Los eventos simples de este experimento están formados por todas las posibles combinaciones de tres fabricantes, elegidos de un grupo de cinco. De estos cinco, tres han sido designados como “mejores” y dos como “no mejores”. Piense en un plato de dulces que contenga tres dulces rojos y dos amarillos, de los cuales usted selecciona tres, como se ilustra en la figura 2.8. El número total de eventos simples N se puede contar como el número de formas para elegir tres de los cinco fabricantes, es decir

$$N = C_3^5 = \frac{5!}{3!2!} = 10$$

FIGURA 2.8

Ilustración para el ejemplo 2.17



Como los fabricantes se seleccionan al azar, cualquiera de estos 10 eventos simples será *igualmente probable*, con probabilidad $1/10$. Pero, ¿cuántos de estos eventos simples resultan en el evento?

A: exactamente dos de los “mejores” tres

Cuente n_A , el número de eventos en A, en dos pasos porque el evento A ocurrirá cuando seleccione dos de los “mejores” tres y uno de los dos “no mejores”. Existen

$$C_2^3 = \frac{3!}{2!1!} = 3$$

formas de efectuar la primera etapa y

$$C_1^2 = \frac{2!}{1!1!} = 2$$

formas de efectuar la segunda etapa. Aplicando la regla *mn*, encontramos que hay $n_A = (3)(2) = 6$ de los 10 eventos sencillos en el evento A y $P(A) = n_A/N = 6/10$.

Existen muchas otras reglas de conteo además de las tres presentadas en esta sección. Si usted está interesado en este tema, consulte uno de los numerosos libros de texto sobre matemáticas combinatorias.

2.8

EJERCICIOS

TÉCNICAS BÁSICAS

2.16 Usted tiene *dos* grupos de objetos muy diferentes, 10 en el primer grupo y ocho en el segundo. Si selecciona un objeto de cada grupo, ¿cuántos pares diferentes puede formar?

2.17 Usted tiene *tres* grupos de objetos muy diferentes, cuatro en el primer grupo, siete en el segundo y tres en el tercero. Si selecciona un objeto de cada grupo, ¿cuántas ternas diferentes formaría?

2.18 Permutaciones Evalúe las siguientes *permutaciones*. (SUGERENCIA: Su calculadora científica debe tener una función que permita calcular permutaciones y combinaciones con gran facilidad.)

a. P_3^5 b. P_9^{10} c. P_6^6 d. P_1^{20}

2.19 Combinaciones Evalúe estas *combinaciones*:

a. C_3^5 b. C_9^{10} c. C_6^6 d. C_1^{20}

2.20 Seleccionar personas ¿En cuántas formas se pueden seleccionar cinco personas de un grupo de ocho si el orden de selección es importante?

2.21 Seleccionar personas, otra vez ¿En cuántas formas se pueden seleccionar dos personas de un grupo de 20 si el orden de selección no es importante?

2.22 Dados Se tiran tres dados. ¿Cuántos eventos simples hay en el espacio muestral?

2.23 Monedas Se tiran al aire cuatro monedas. ¿Cuántos eventos simples hay en el espacio muestral?

2.24 Un problema de urna, otra vez Se seleccionan tres pelotas de una caja que contiene 10 de ellas. El orden de selección no es importante. ¿Cuántos eventos simples hay en el espacio muestral?

APLICACIONES

2.25 Itinerarios Un vendedor está preparando un itinerario para visitar seis ciudades principales. La distancia recorrida, y por tanto el costo del viaje, dependerá del orden en el que planee su ruta. ¿Cuántos itinerarios diferentes (y costos de viaje) son posibles?

2.26 Un juego de cartas Tres estudiantes juegan a las cartas. Deciden elegir al primero para jugar a seleccionar cada uno una carta del mazo de 52 y buscar la de mayor

valor y palo. Ordenan los palos de menor a mayor: tréboles, diamantes, corazones y espadas.

- a. Si la carta se devuelve al mazo después de que cada estudiante elige, ¿cuántas configuraciones son posibles de las tres selecciones?
- b. ¿Cuántas configuraciones hay en las que cada estudiante toma una carta diferente?
- c. ¿Cuál es la probabilidad de que los tres estudiantes elijan exactamente la misma carta?
- d. ¿Cuál es la probabilidad de que los tres estudiantes seleccionen cartas diferentes?

2.27 Comida en el restaurante Andree Un restaurante francés ofrece un menú especial de verano en el cual, por un costo fijo por comida, se elige una de dos ensaladas, una de dos entradas y uno de dos postres. ¿Cuántas comidas diferentes hay?

2.28 Jugador de póquer Se seleccionan cinco de un mazo de 52 cartas para una mano de póquer.

- a. ¿Cuántos eventos simples hay en el espacio muestral?
- b. Una *escalera real* es una mano que contiene el A, K, Q, J y 10, todas del mismo palo. ¿Cuántas formas hay para obtener una escalera real?
- c. ¿Cuál es la probabilidad de recibir una escalera real?

2.29 Póquer II Consulte el ejercicio 2.28. Usted tiene una mano de póquer con cuatro de una clase.

- a. ¿Cuántas manos de póquer posibles puede recibir?
- b. ¿En cuántas formas puede recibir cuatro cartas del mismo valor de cara y una carta de las otras 48 disponibles?
- c. ¿Cuál es la probabilidad de recibir cuatro de una clase?

2.30 Encuesta en un hospital Se efectuará un estudio en un hospital para determinar las actitudes de las enfermeras hacia diversos procedimientos administrativos.

Si se selecciona una muestra de 10 enfermeras entre un total de 90, ¿cuántas muestras diferentes se pueden seleccionar? (SUGERENCIA: ¿El orden es importante para determinar la conformación de la muestra a seleccionar para el estudio?)

2.31 Carrera de 100 metros, otra vez Consulte el ejercicio 2.14, en el cual A, B, C y D corren un *sprint* de 100 metros. Suponga que todos los corredores están igualmente calificados, de modo que cualquier orden de terminación es igualmente probable. Use la regla *mn* o permutaciones para contestar estas preguntas:

- a. ¿Cuántos órdenes de terminación son posibles?
- b. ¿Cuál es la probabilidad de que D gane el *sprint*?
- c. ¿Cuál es la probabilidad de que D gane y A obtenga el segundo lugar?
- d. ¿Cuál es la probabilidad de que C termine en último lugar?

2.32 ¿Sesgo en el género? Una mujer presentó una demanda por discriminación de género ante un consejo asesor en relaciones humanas formado por ocho miembros. El consejo, compuesto por cinco mujeres y tres hombres, votó 5-3 a favor de la demandante, las cinco mujeres votaron por la demandante y los tres hombres en contra. ¿El consejo fue afectado por sesgo de género? Es decir, si el voto a favor de la demandante fue 5-3 y los miembros del consejo no mostraron sesgo debido al género, ¿cuál es la probabilidad de que el voto se dividiera junto con las líneas de género (cinco mujeres a favor, tres hombres en contra)?

2.33 Estudio intensivo Una estudiante se prepara para un examen al estudiar una lista de 10 problemas; ella resuelve seis. Para el examen, el profesor selecciona cinco problemas al azar de la lista de 10. ¿Cuál es la probabilidad de que la estudiante resuelva cinco problemas en el examen?

RELACIONES DE EVENTO Y REGLAS DE PROBABILIDAD

2.9

En ocasiones el evento de interés se forma como una combinación de algunos otros eventos. Sean A y B dos eventos definidos en el espacio muestral S . Aquí hay tres relaciones importantes entre eventos.

Definición La **unión** de los eventos A y B , denotada por $A \cup B$, es el evento en que ocurren A o B o ambos.

Definición La **intersección** de eventos A y B , denotada por $A \cap B$, es el evento en que ocurren A y B .[†]

Definición El **complemento** de un evento A , denotado por A^c , es el evento en que A no ocurre.

Las figuras 2.9, 2.10 y 2.11 muestran representaciones del diagrama de Venn de $A \cup B$, $A \cap B$ y A^c , respectivamente. Cualquier evento simple en el área sombreada es un posible resultado que aparece en el evento apropiado. Una forma de hallar las probabilidades de la unión, la intersección o el complemento es sumar las probabilidades de todos los eventos simples asociados.

FIGURA 2.9

Diagrama de Venn de $A \cup B$

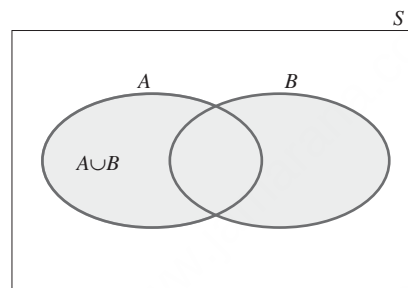
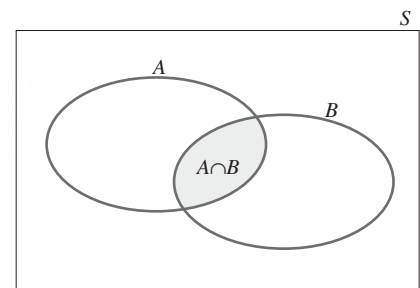


FIGURA 2.10

Diagrama de Venn de $A \cap B$

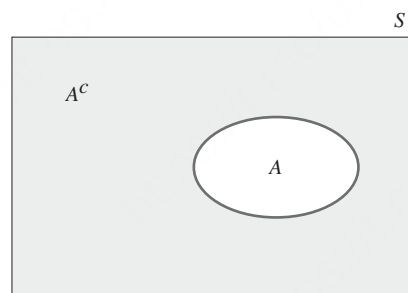


TIP MI CONSEJO

Intersección $\Leftrightarrow$ "ambos...y" o sólo "y".
Unión $\Leftrightarrow$ "uno de dos... o ambos" o sólo "o"

FIGURA 2.11

El complemento de un evento



EJEMPLO

2.18

Se tiran al aire dos monedas imparciales y se registra el resultado. Éstos son los eventos de interés:

A : observar al menos una cara

B : observar al menos una cruz

Defina los eventos A , B , $A \cap B$, $A \cup B$ y A^c como conjuntos de eventos simples, y encuentre sus probabilidades.

Solución Recuerde del ejemplo 2.7 que los eventos simples para este experimento son

E_1 : HH (cara en primera moneda, cara en segunda)

E_2 : HT

E_3 : TH

E_4 : TT

[†] Algunos autores usan la notación AB .

y que cada evento simple tiene probabilidad $1/4$. El evento A , al menos una cara, se presenta si ocurre E_1, E_2 o E_3 , de modo que

$$A = \{E_1, E_2, E_3\} \quad P(A) = \frac{3}{4}$$

y

$$A^c = \{E_4\} \quad P(A^c) = \frac{1}{4}$$

Del mismo modo,

$$B = \{E_2, E_3, E_4\} \quad P(B) = \frac{3}{4}$$

$$A \cap B = \{E_2, E_3\} \quad P(A \cap B) = \frac{1}{2}$$

$$A \cup B = \{E_1, E_2, E_3, E_4\} \quad P(A \cup B) = \frac{4}{4} = 1$$

Observe que $(A \cup B) = S$, el espacio muestral, y entonces es seguro que ocurra.

El concepto de uniones e intersecciones puede ampliarse a más de dos eventos. Por ejemplo, la unión de tres eventos A, B y C , que se escriben como $A \cup B \cup C$, es el conjunto de eventos simples que están en A o B o C o en cualquier combinación de esos eventos. Análogamente, la intersección de los tres eventos A, B y C , que se escribe como $A \cap B \cap C$, es el conjunto de eventos simples que son comunes a los tres eventos A, B y C .

Cálculo de probabilidades para uniones y complementos

Cuando podemos escribir el evento de interés en la forma de una unión, un complemento o una intersección, hay reglas de probabilidad especiales que simplifican nuestros cálculos. La primera regla se refiere a *uniones* de eventos.

REGLA DE LA ADICIÓN

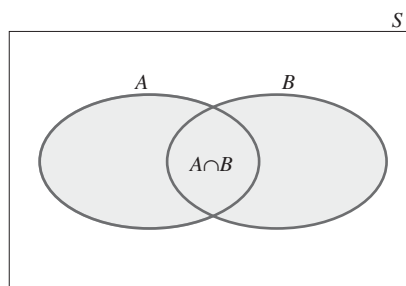
Dados dos eventos, A y B , la probabilidad de su unión, $A \cup B$, es igual a

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Observe en el diagrama de Venn en la figura 2.12 que la suma $P(A) + P(B)$ cuenta dos veces los eventos simples que son comunes para ambos: A y B . La resta de $P(A \cap B)$ da el resultado correcto.

FIGURA 2.12

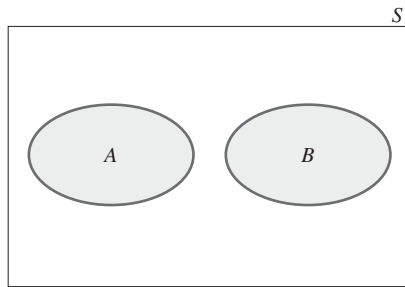
Regla de la adición



Cuando dos eventos A y B son **mutuamente excluyentes**, **disjuntos** o **ajenos**, significa que cuando ocurre A , B no puede ocurrir, y viceversa. Esto significa que la probabilidad de que ambos ocurran, $P(A \cap B)$, debe ser cero. La figura 2.13 es una representación de un diagrama de Venn de dos de estos eventos sin ningún evento simple en común.

FIGURA 2.13

Dos eventos ajenos o mutuamente excluyentes

**TIP** MI CONSEJO

Recuerde, mutuamente excluyente $\Leftrightarrow$
 $P(A \cap B) = 0$

Cuando dos eventos A y B son **mutuamente excluyentes**, entonces $P(A \cap B) = 0$ y la regla de la adición se simplifica a

$$P(A \cup B) = P(A) + P(B)$$

La segunda regla se refiere a *complementos* de eventos. Se observa del diagrama de Venn de la figura 2.11 que A y A^c son mutuamente excluyentes y que $A \cup A^c = S$, todo el espacio muestral. Se deduce que

$$P(A) + P(A^c) = 1 \text{ y } P(A^c) = 1 - P(A)$$

REGLA PARA COMPLEMENTOS

$$P(A^c) = 1 - P(A)$$

EJEMPLO

2.19

Una compañía de exploración petrolera planea perforar dos pozos de exploración. Se emplea evidencia del pasado para evaluar los posibles resultados de la tabla 2.4.

TABLA 2.4**Resultados para el experimento de perforación petrolífera**

Evento	Descripción	Probabilidad
A	Ningún pozo produce petróleo ni gas	.80
B	Exactamente un pozo produce petróleo o gas	.18
C	Ambos pozos producen petróleo o gas	.02

Encuentre $P(A \cup B)$ y $P(B \cup C)$.

Solución Por su definición, los eventos A , B y C son mutuamente excluyentes en forma conjunta porque el suceso de un evento impide que ocurra cualquiera de los otros dos. Por tanto,

$$P(A \cup B) = P(A) + P(B) = .80 + .18 = .98$$

y

$$P(B \cup C) = P(B) + P(C) = .18 + .02 = .20$$

El evento $A \cup B$ se describe como el evento que *a lo sumo* un pozo produce petróleo o gas, y $B \cup C$ describe el evento que *al menos* un pozo produce gas o petróleo.

EJEMPLO

2.20

En una encuesta telefónica hecha a 1 000 empleados, se les preguntó su opinión acerca del costo de una educación universitaria. Quienes respondieron se clasificaron de acuerdo con si actualmente tenían un hijo en la universidad y si pensaban que la carga de un préstamo para la mayoría de los estudiantes universitarios es demasiado alta, la cantidad correcta o muy poca. Las proporciones de quienes contestaron se muestran en la **tabla de probabilidad** de la tabla 2.5. Suponga que un entrevistado se elige al azar de este grupo.

TABLA 2.5

Tabla de probabilidad

	Demasiado alta (A)	Cantidad correcta (B)	Muy poca (C)
Hijo en universidad (D)	.35	.08	.01
Sin hijo en universidad (E)	.25	.20	.11

1. ¿Cuál es la probabilidad de que el entrevistado tenga un hijo en la universidad?
2. ¿Cuál es la probabilidad de que el entrevistado no tenga un hijo en la universidad?
3. ¿Cuál es la probabilidad de que el entrevistado tenga un hijo en la universidad o piense que la carga de un préstamo es demasiado alta, o ambos?

Solución La tabla 2.5 da las probabilidades para los seis eventos simples. Por ejemplo, la entrada en la esquina superior izquierda de la tabla es la probabilidad de que un entrevistado tenga un hijo en la universidad y *además* piense que la carga de un préstamo es demasiado alta ($A \cap D$).

1. El evento de que un entrevistado tenga un hijo en la universidad ocurrirá, cualquiera que sea su respuesta a la pregunta acerca de la carga por el préstamo. Esto es, el evento D consta de los eventos simples del primer renglón:

$$P(D) = .35 + .08 + .01 = .44$$

En general, las probabilidades de eventos marginales como D y A se encuentran al sumar las probabilidades en el renglón o columna apropiados.

2. El evento de que el entrevistado no tiene un hijo en la universidad es el complemento del evento D denotado por D^c . La probabilidad de D^c se encuentra como

$$P(D^c) = 1 - P(D)$$

Usando el resultado del punto 1, tenemos

$$P(D^c) = 1 - .44 = .56$$

3. El evento de interés es $P(A \cup D)$. Usando la Regla de la adición,

$$\begin{aligned} P(A \cup D) &= P(A) + P(D) - P(A \cap D) \\ &= .60 + .44 - .35 \\ &= .69 \end{aligned}$$

INDEPENDENCIA, PROBABILIDAD CONDICIONAL Y LA REGLA DE LA MULTIPLICACIÓN

En el ejemplo 2.20, pudimos usar la regla de la adición para calcular $P(A \cup D)$ porque se pudo hallar $P(A \cap D)$ directamente de la tabla de probabilidad. En ocasiones, sin embargo, la probabilidad de intersección se desconoce. En esta situación, hay una regla de probabilidad que se usa para calcular la probabilidad de la intersección de varios eventos. Esta regla depende del concepto estadístico importante de **eventos independientes o dependientes**.

Definición Se dice que dos eventos, A y B , son **independientes** si y sólo si la probabilidad del evento B no está influida o cambiada por el suceso del evento A , o viceversa.

Daltonismo Suponga que un investigador observa el género de una persona y si ésta no distingue los colores rojo y verde. ¿Cambia la probabilidad de que una persona sea daltónica, dependiendo de si es hombre o no? Defina dos eventos:

A : la persona es hombre
 B : la persona es daltónica

En este caso, como el daltonismo es una característica relacionada con el sexo masculino, la probabilidad de que un hombre sea daltónico será mayor que la probabilidad de que una persona seleccionada de la población general sea daltónica. La probabilidad del evento B , que una persona sea daltónica, depende de si ha ocurrido o no el evento A , que la persona sea hombre. Decimos que A y B son *eventos dependientes*.

Tirar dados Por otra parte, considere tirar un solo dado dos veces y defina dos eventos:

A : observar un 2 en el primer tiro
 B : observar un 2 en el segundo tiro

Si el dado es imparcial, la probabilidad del evento A es $P(A) = 1/6$. Considere la probabilidad del evento B . Ya sea que el evento A haya ocurrido o no, la probabilidad de observar un 2 en el segundo tiro todavía es $1/6$. Podríamos escribir:

$$P(B \text{ dado que } A \text{ ocurrió}) = 1/6$$

$$P(B \text{ dado que } A \text{ no ocurrió}) = 1/6$$

Como la probabilidad del evento B no ha cambiado por el suceso del evento A , decimos que A y B son *eventos independientes*.

La probabilidad de un evento A , dado que el evento B ha ocurrido, se denomina **probabilidad condicional de A , dado que B ha ocurrido**, denotada por $P(A|B)$. La barra vertical se lee “dada” y los eventos que aparecen a la derecha de la barra son aquellos que se sabe han ocurrido. Usaremos estas probabilidades para calcular la probabilidad de que *ambos*, A y B ocurran cuando se realice el experimento.

REGLA GENERAL DE LA MULTIPLICACIÓN

La probabilidad de que A y B ocurran cuando el experimento se realiza es

$$P(A \cap B) = P(A)P(B|A)$$

o

$$P(A \cap B) = P(B)P(A|B)$$

EJEMPLO

2.21

En un experimento de preferencia de color, ocho juguetes se ponen en un recipiente. Los juguetes son idénticos excepto por el color, dos son rojos y seis son verdes. Se pide a un niño que elija dos juguetes *al azar*. ¿Cuál es la probabilidad de que el niño elija los dos juguetes rojos?

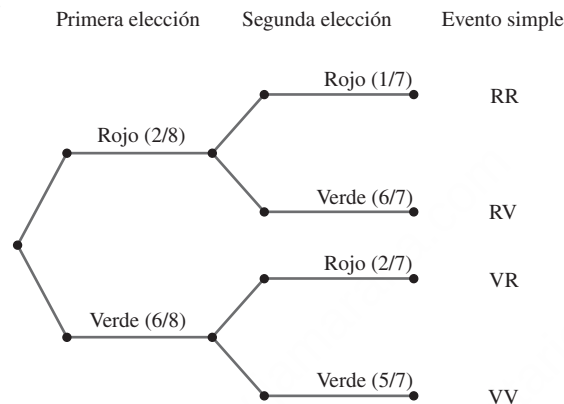
Solución Se visualiza el experimento usando un diagrama de árbol como se muestra en la figura 2.14. Defina los eventos siguientes:

R: se elige juguete rojo

V: se elige juguete verde

FIGURA 2.14

Diagrama de árbol para el ejemplo 2.21



El evento A (ambos juguetes son rojos) se construye como la intersección de dos eventos:

$$A = (\text{R en la primera selección}) \cap (\text{R en la segunda selección})$$

Como sólo hay dos juguetes rojos en el recipiente, la probabilidad de elegir el rojo en la primera selección es $2/8$. No obstante, una vez que haya sido seleccionado este juguete rojo, la probabilidad del rojo en la segunda selección es *dependiente* del resultado de la primera selección (véase la figura 2.14). Si la primera selección fue un juguete rojo, la probabilidad de elegir un segundo juguete rojo es sólo $1/7$ porque hay sólo un juguete rojo entre los siete restantes. Si la primera selección fue verde, la probabilidad de elegir rojo en la segunda selección es $2/7$ porque hay dos juguetes rojos entre los siete restantes. Usando esta información y la regla de la multiplicación, se puede hallar la probabilidad del evento A .

$$\begin{aligned} P(A) &= P(\text{R en la primera selección} \cap \text{R en la segunda selección}) \\ &= P(\text{R en la primera selección}) P(\text{R en la segunda selección} | \text{R en la primera}) \\ &= \left(\frac{2}{8}\right)\left(\frac{1}{7}\right) = \frac{2}{56} = \frac{1}{28} \end{aligned}$$

La solución en el ejemplo 2.21 fue posible sólo debido a que se sabía que $P(\text{R en la segunda selección} | \text{R en la primera selección})$. Si no conoce la probabilidad condicional, $P(A|B)$, puede calcularla usando la regla de la multiplicación en una forma ligeramente diferente. Sólo *reordene los términos* de la regla de la multiplicación.

PROBABILIDAD CONDICIONAL

PROBABILIDADES CONDICIONALES

La probabilidad condicional del evento A , dado que el evento B ha ocurrido, es

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{si } P(B) \neq 0$$

La probabilidad condicional del evento B , dado que el evento A ha ocurrido, es

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \text{si } P(A) \neq 0$$

Observe que, en esta forma, ¡usted necesita conocer $P(A \cap B)$!

Daltonismo, continúa Suponga que en la población general hay 51% de hombres y 49% de mujeres, y que las proporciones de hombres y mujeres daltónicos se muestran en la siguiente tabla de probabilidad:

	Hombres (B)	Mujeres (B^C)	Total
Daltónico (A)	.04	.002	.042
No daltónico (A^C)	.47	.488	.958
Total	.51	.49	1.00

Si una persona se selecciona al azar de esta población y se encuentra que es hombre (evento B), ¿cuál es la probabilidad de que el hombre sea daltónico (evento A)? Si sabemos que el evento B ha ocurrido, debemos restringir nuestra atención a sólo 51% de la población que es de hombres. La probabilidad de ser daltónico, dado que la persona es hombre, es 4% de 51%, o sea

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{.04}{.51} = .078$$

¿Cuál es la probabilidad de ser daltónico, dado que la persona es mujer? Ahora estamos restringidos a sólo el 49% de la población que es de mujeres y

$$P(A|B^C) = \frac{P(A \cap B^C)}{P(B^C)} = \frac{.002}{.49} = .004$$

Observe que la probabilidad del evento A cambió, dependiendo de si el evento B ocurrió. Esto indica que estos dos eventos son *dependientes*.

Cuando dos eventos son **independientes**, es decir, si la probabilidad del evento B es igual, ya sea que el evento A haya o no ocurrido, entonces el evento A no afecta al evento B y por tanto

$$P(B|A) = P(B)$$

Ahora se puede simplificar la regla de la multiplicación.

LA REGLA DE LA MULTIPLICACIÓN PARA EVENTOS INDEPENDIENTES

Si dos eventos A y B son independientes, la probabilidad de que ocurran A y B es

$$P(A \cap B) = P(A)P(B)$$

Del mismo modo, si A , B y C son eventos mutuamente independientes (todos los pares de eventos son independientes), entonces la probabilidad de que A , B y C ocurran es

$$P(A \cap B \cap C) = P(A)P(B)P(C)$$

Tiros de monedas en juegos de fútbol Un equipo de fútbol interviene en dos periodos de tiempo extra durante un juego determinado, de modo que hay tres tiros de monedas al aire. Si la moneda es imparcial, ¿cuál es la probabilidad de que pierdan los tres tiros?

Solución Si la moneda es imparcial, el evento se describe en tres pasos:

A : perder el primer tiro

B : perder el segundo tiro

C : perder el tercer tiro

Como los tiros son independientes y como $P(\text{gana}) = P(\text{pierde}) = .5$ para cualquiera de los tres tiros,

$$P(A \cap B \cap C) = P(A)P(B)P(C) = (.5)(.5)(.5) = .125$$

¿Cómo se verifica si los dos eventos son independientes o dependientes? La solución más fácil es redefinir el concepto de **independencia** en un modo más formal.

VERIFICACIÓN DE INDEPENDENCIA

Se dice que dos eventos A y B son **independientes** si y sólo si

$$P(A \cap B) = P(A)P(B)$$

o bien,

$$P(B|A) = P(B) \text{ o, en forma equivalente, } P(A|B) = P(A)$$

De otro modo, se dice que los eventos son **dependientes**.

EJEMPLO

2.22

Tire al aire dos monedas y observe el resultado. Defina estos eventos:

A : cara en la primera moneda

B : cruz en la segunda moneda

¿Los eventos A y B son independientes?

Solución De los ejemplos anteriores, sabemos que $S = \{HH, HT, TH, TT\}$. Utilice estos cuatro eventos simples para hallar

$$P(A) = \frac{1}{2}, P(B) = \frac{1}{2} \text{ y } P(A \cap B) = \frac{1}{4}$$

Como $P(A)P(B) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$ y $P(A \cap B) = \frac{1}{4}$, tenemos $P(A)P(B) = P(A \cap B)$

y los dos eventos deben ser independientes.

TIP MI CONSEJO

Recuerde,
independencia $\Leftrightarrow$
 $P(A \cap B) = P(A)P(B)$

EJEMPLO

2.23

Consulte la tabla de probabilidad del ejemplo 2.20, que se reproduce a continuación:

	Demasiado alta (A)	Cantidad correcta (B)	Muy poco (C)
Hijo en universidad (D)	.35	.08	.01
Sin hijo en universidad (E)	.25	.20	.11

¿Los eventos D y A son independientes? Explique.

Solución

- Utilice la tabla de probabilidad para hallar $P(A \cap D) = .35$, $P(A) = .60$, y $P(D) = .44$. Entonces

$$P(A)P(D) = (.60)(.44) = .264 \text{ y } P(A \cap D) = .35$$

Como estas dos probabilidades no son iguales, los eventos A y D son *dependientes*.

- Alternativamente, calcule

$$P(A|D) = \frac{P(A \cap D)}{P(D)} = \frac{.35}{.44} = .80$$

Ya que $P(A|D) = .80$ y $P(A) = .60$, de nuevo nos lleva a la conclusión de que los eventos A y D son *dependientes*.

- De manera equivalente,

$$P(D|A) = \frac{P(A \cap D)}{P(A)} = \frac{.35}{.60} = .58$$

mientras que $P(D) = .44$. Una vez más vemos que A y D son eventos dependientes.

**NECESITO SABER...**

La diferencia entre eventos mutuamente excluyentes y eventos independientes

Muchos estudiantes encuentran difícil decir la diferencia entre eventos *mutuamente excluyentes* y eventos *independientes*.

- Cuando dos eventos son *mutuamente excluyentes* o *ajenos*, no pueden ocurrir los dos cuando se realice el experimento. Una vez ocurrido el evento B , el evento A no puede ocurrir, de modo que $P(A|B) = 0$, o viceversa. El suceso del evento B ciertamente afecta la probabilidad de que el evento A pueda ocurrir.
- Por tanto, los eventos mutuamente excluyentes deben ser dependientes.
- Cuando dos eventos son mutuamente excluyentes o ajenos,
 $P(A \cap B) = 0$ y $P(A \cup B) = P(A) + P(B)$.
- Cuando dos eventos son independientes,
 $P(A \cap B) = P(A)P(B)$ y $P(A \cup B) = P(A) + P(B) - P(A)P(B)$.

El uso de reglas de probabilidad para calcular la probabilidad de un evento requiere alguna experiencia e ingenio. Usted necesita expresar el evento de interés como una unión o intersección (o la combinación de ambas) de dos o más eventos cuyas probabilidades son conocidas o se calculan con facilidad. Es frecuente que esto lo haga de diferentes formas; la clave es encontrar la combinación correcta.

EJEMPLO

2.24

Se sacan dos cartas de un mazo de 52. Calcule la probabilidad de que el par incluya un as y un 10.

Solución Considere el evento de interés:

A : sacar un as y un 10

Entonces $A = B \cup C$, donde

B : sacar el as en el primer saque y el 10 en el segundo

C : sacar el 10 en el primer saque y el as en el segundo

Los eventos B y C se eligen como mutuamente excluyentes y también como intersecciones de eventos con probabilidades conocidas; esto es,

$$B = B_1 \cap B_2 \text{ y } C = C_1 \cap C_2$$

donde

B_1 : sacar un as en el primer saque

B_2 : sacar un 10 en el segundo saque

C_1 : sacar un 10 en el primer saque

C_2 : sacar un as en el segundo saque

Aplicando la Regla de la multiplicación, tenemos

$$\begin{aligned} P(B_1 \cap B_2) &= P(B_1)P(B_2|B_1) \\ &= \left(\frac{4}{52}\right)\left(\frac{4}{51}\right) \end{aligned}$$

y

$$P(C_1 \cap C_2) = \left(\frac{4}{52}\right)\left(\frac{4}{51}\right)$$

Entonces, aplicando la Regla de la adición,

$$\begin{aligned} P(A) &= P(B) + P(C) \\ &= \left(\frac{4}{52}\right)\left(\frac{4}{51}\right) + \left(\frac{4}{52}\right)\left(\frac{4}{51}\right) = \frac{8}{663} \end{aligned}$$

Con todo cuidado verifique cada composición para asegurarse que en realidad es igual al evento de interés.

2.11

EJERCICIOS

TÉCNICAS BÁSICAS

2.34 Un experimento resultaría en uno de cinco eventos simples igualmente probables, $E_1, E_2, \dots, E_5$. Los eventos A , B y C se definen como sigue:

$$A: E_1, E_3 \quad P(A) = .4$$

$$B: E_1, E_2, E_4, E_5 \quad P(B) = .8$$

$$C: E_3, E_4 \quad P(C) = .4$$

Encuentre las probabilidades asociadas con los siguientes eventos, haciendo una lista de los eventos simples en cada uno.

a. A^c

d. $A \cup B$

g. $A \cup B \cup C$

b. $A \cap B$

e. $B|C$

h. $(A \cap B)^c$

c. $B \cap C$

f. $A|B$

2.35 Consulte el ejercicio 2.34. Use la definición de un evento complementario para hallar estas probabilidades:

a. $P(A^c)$

b. $P((A \cap B)^c)$

¿Los resultados concuerdan con los obtenidos en el ejercicio 2.34?

2.36 Consulte el ejercicio 2.34. Use la definición de probabilidad condicional para hallar estas probabilidades:

- a. $P(A|B)$ b. $P(B|C)$

¿Los resultados concuerdan con los obtenidos en el ejercicio 2.34?

2.37 Consulte el ejercicio 2.34. Use las Reglas de la adición y de la multiplicación para hallar estas probabilidades:

- a. $P(A \cup B)$ b. $P(A \cap B)$ c. $P(B \cap C)$

¿Los resultados concuerdan con los obtenidos en el ejercicio 2.34?

2.38 Consulte el ejercicio 2.34.

- a. ¿Los eventos A y B son independientes?
b. ¿Los eventos A y B son mutuamente excluyentes?

2.39 Suponga que $P(A) = .1$ y $P(B) = .5$.

- a. Si $P(A|B) = .1$, ¿qué es $P(A \cap B)$?
b. Si $P(A|B) = .1$, ¿ A y B son independientes?
c. Si $P(A \cap B) = 0$, ¿ A y B son independientes?
d. Si $P(A \cup B) = .65$, ¿ A y B son mutuamente excluyentes?

2.40 Dados Un experimento consiste en tirar un solo dado y observar el número de puntos que aparecen en la cara superior. Los eventos A , B y C están definidos como sigue:

A : observar un número menor que 4

B : observar un número menor o igual a 2

C : observar un número mayor que 3

Encuentre las probabilidades asociadas con los eventos citados a continuación, usando ya sea el método de evento simple o las reglas y definiciones de esta sección.

- a. S b. $A|B$ c. B
d. $A \cap B \cap C$ e. $A \cap B$ f. $A \cap C$
g. $B \cap C$ h. $A \cup C$ i. $B \cup C$

2.41 Consulte el ejercicio 2.40.

- a. ¿Los eventos A y B son independientes? ¿Mutuamente excluyentes?
b. ¿Los eventos A y C son independientes? ¿Mutuamente excluyentes?

2.42 Se lanzan dos dados imparciales.

- a. ¿Cuál es la probabilidad de que la suma de los números de puntos mostrados en las caras superiores sea igual a 7? ¿A 11?
b. ¿Cuál es la probabilidad de que tire “dobles”; es decir, que ambos dados tengan el mismo número en la cara superior?
c. ¿Cuál es la probabilidad de que ambos dados muestren un número impar?

2.43 Suponga que $P(A) = .4$ y $P(B) = .2$. Si los eventos A y B son independientes, encuentre estas probabilidades:

- a. $P(A \cap B)$ b. $P(A \cup B)$

2.44 Suponga que $P(A) = .3$ y $P(B) = .5$. Si los eventos A y B son mutuamente excluyentes, encuentre estas probabilidades:

- a. $P(A \cap B)$ b. $P(A \cup B)$

2.45 Suponga que $P(A) = .4$ y $P(A \cap B) = .12$.

- a. Encuentre $P(B|A)$.
b. ¿Los eventos A y B son mutuamente excluyentes?
c. Si $P(B) = .3$, ¿los eventos A y B son independientes?

2.46 Un experimento resulta en uno o ambos de los eventos A y B con las probabilidades que se muestran en esta tabla de probabilidad:

	A	A^c
B	.34	.46
B^c	.15	.05

Encuentre las siguientes probabilidades:

- a. $P(A)$ b. $P(B)$ c. $P(A \cap B)$
d. $P(A \cup B)$ e. $P(A|B)$ f. $P(B|A)$

2.47 Consulte el ejercicio 2.46.

- a. ¿Los eventos A y B son mutuamente excluyentes? Explique.
b. ¿Los eventos A y B son independientes? Explique.

APLICACIONES

2.48 Fondo monetario para donaciones Suponga que un conjunto de propuestas de investigación fue evaluado por un grupo de expertos en cuanto a si las propuestas merecían ser financiadas. Cuando estas mismas propuestas se enviaron a un segundo grupo independiente de expertos, la decisión para financiar se revirtió en 30% de los casos. Si la probabilidad es .2 de que una propuesta sea juzgada por el primer grupo como digna de ser financiada, ¿cuáles son las probabilidades de estos eventos?

- a. Una propuesta digna es aprobada por ambos grupos.
b. Una propuesta digna es desaprobada por ambos grupos.
c. Una propuesta digna es aprobada por un grupo.

2.49 Drogadictos delincuentes Un estudio acerca de drogadictos delincuentes, después de un tratamiento por abuso en el consumo de drogas, sugiere que la probabilidad de que sean condenados en un periodo no mayor a dos años después del tratamiento, quizá dependa de la educación del delincuente. Las proporciones del número total de casos que caen en cuatro categorías de educación/condena se muestran en la tabla siguiente.

Educación	Situación no mayor a 2 años después del tratamiento		Totales
	Condenado	No condenado	
10 años o más	.10	.30	.40
9 años o menos	.27	.33	.60
Totales	.37	.63	1.00

Suponga que se selecciona un delincuente del programa de tratamiento. Aquí tenemos los eventos de interés:

A: el delincuente tiene 10 años o más de educación

B: el delincuente es condenado no más de dos años después de terminar su tratamiento

Encuentre las probabilidades apropiadas para estos eventos:

- a. A b. B c. $A \cap B$
d. $A \cup B$ e. A^c f. $(A \cup B)^c$
g. $(A \cap B)^c$ h. A dado que B ha ocurrido
i. B dado que A ha ocurrido

2.50 Use las probabilidades del ejercicio 2.49 para demostrar que estas igualdades son verdaderas:

- a. $P(A \cap B) = P(A)P(B|A)$
b. $P(A \cap B) = P(B)P(A|B)$
c. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

2.51 El problema del cumpleaños Dos personas entran a un cuarto y se registran sus cumpleaños (omitendo sus años).

- a. Identifique la naturaleza de los eventos simples en S .
b. ¿Cuál es la probabilidad de que las dos personas tengan un par específico de cumpleaños?
c. Identifique los eventos simples en el evento A : ambas personas tienen el mismo cumpleaños.
d. Encuentre $P(A)$. e. Encuentre $P(A^c)$.

2.52 El problema del cumpleaños, continúa Si n personas entran a un cuarto, encuentre estas probabilidades:

- A: ninguna de las personas tienen el mismo cumpleaños
B: al menos dos de las personas tienen el mismo cumpleaños

Resuelva para

- a. $n = 3$ b. $n = 4$

[NOTA: Sorprendentemente, $P(B)$ aumenta con rapidez cuando n aumenta. Por ejemplo, para $n = 20$, $P(B) = .411$; para $n = 40$, $P(B) = .891$.]

2.53 Líneas de inspección Cierta artículo manufacturado es inspeccionado visualmente por dos inspectores diferentes. Cuando un artículo defectuoso pasa por la línea de producción, la probabilidad de que logre pasar por el primer inspector es .1. De aquellos que pasan al primer inspector, el segundo inspector “perderá” cinco

de 10. ¿Qué fracción de artículos defectuosos logra pasar por ambos inspectores?

2.54 Fumar y cáncer Un estudio realizado en pobladores de cierta región mostró que 20% eran fumadores. La probabilidad de muerte debida a cáncer pulmonar, dado que una persona fumaba, era alrededor de 10 veces la probabilidad de muerte debida a cáncer pulmonar de una persona que no fumaba. Si la probabilidad de muerte debida a cáncer pulmonar en la región es .006, ¿cuál es la probabilidad de muerte debida a cáncer pulmonar dado que una persona es fumadora?

2.55 Detectores de humo Un sistema detector de humo utiliza dos aparatos, A y B . Si hay humo, la probabilidad de que éste sea detectado por el aparato A es .95; por el aparato B , .98; y por ambos aparatos, .94.

- a. Si hay humo, encuentre la probabilidad de que éste sea detectado por el aparato A o el B o por ambos aparatos.
b. Encuentre la probabilidad de que el humo no sea detectado.

2.56 Genética de plantas Gregor Mendel sugirió en 1865 una teoría de la herencia basada en la ciencia de la genética. Él identificó individuos heterocigotos de flores de color que tenían dos alelos (un r = alelo recesivo de color blanco y uno R = alelo dominante de color rojo). Cuando estos individuos se apareaban, observó que 3/4 de los descendientes tenían flores rojas y 1/4 tenían flores blancas. La tabla siguiente resume este apareamiento; cada padre da uno de sus alelos para formar el gen del descendiente.

Padre 1	Padre 2	
	r	R
r	rr	rR
R	Rr	RR

Suponemos que es igualmente probable que cada padre dé cualquiera de los dos alelos y que, si uno de ellos o los dos alelos de un par es dominante (R), el descendiente tendrá flores rojas.

- a. ¿Cuál es la probabilidad de que un descendiente en este apareamiento tenga al menos un alelo dominante?
b. ¿Cuál es la probabilidad de que un descendiente tenga al menos un alelo recesivo?
c. ¿Cuál es la probabilidad de que un descendiente tenga un alelo recesivo, dado que el descendiente tiene flores rojas?

2.57 Lesiones en fútbol Durante la temporada inaugural de la liga mayor de fútbol soccer en Estados Unidos, los equipos médicos documentaron 256 lesiones que causaron la pérdida de tiempo de participación a jugadores. Los resultados de esta investigación, publicados en *The*

American Journal of Sports Medicine, se muestran en la tabla siguiente.²

Gravedad	Práctica (P)	Juego (G)	Total
Menor (A)	66	88	154
Moderado (B)	23	44	67
Grave (C)	12	23	35
Total	101	155	256

Si un individuo se saca al azar de este grupo de 256 jugadores de futbol soccer, encuentre las siguientes probabilidades:

- a. $P(A)$ b. $P(G)$ c. $P(A \cap G)$
d. $P(G|A)$ e. $P(G|B)$ f. $P(G|C)$
g. $P(C|P)$ h. $P(B^c)$

2.58 Elegir pareja Es frecuente que hombres y mujeres no estén de acuerdo sobre cómo piensan seleccionar una pareja. Suponga que una encuesta hecha a 1000 personas de entre 20 y 30 años dio las siguientes respuestas, a la pregunta de si es más importante para su futura pareja ser capaz de comunicar sus sentimientos (F) de lo que es para esa persona vivir bien (G).

	Sentimientos (F)	Vivir bien (G)	Totales
Hombres (M)	.35	.20	.55
Mujeres (W)	.36	.09	.45
Totales	.71	.29	1.00

Si se selecciona al azar una persona de este grupo de 1000, calcule las siguientes probabilidades:

- a. $P(F)$ b. $P(G)$ c. $P(F|M)$
d. $P(F|W)$ e. $P(M|F)$ f. $P(W|G)$

2.59 Kobe y Lamar Dos estrellas de los *Lakers de Los Ángeles* son muy diferentes cuando se trata de tiros libres. ESPN.com menciona que Kobe Bryant encesta 85% de sus tiros libres mientras Lamar Odum encesta 62% de sus tiros libres.³ Suponga que los tiros libres son independientes y que cada jugador lanza dos tiros libres durante un juego de práctica.

- a. ¿Cuál es la probabilidad de que Kobe enceste sus dos tiros libres?
b. ¿Cuál es la probabilidad de que Lamar enceste exactamente uno de sus dos tiros libres?
c. ¿Cuál es la probabilidad de que Kobe enceste sus dos tiros libres y que Lamar no enceste ninguno de los suyos?

2.12

TEOREMA DE BAYES

Daltonismo Reconsideremos el experimento referente a daltonismo visto en la sección 2.10. Observe que los dos eventos

B : la persona seleccionada es un hombre

B^c : la persona seleccionada es una mujer

tomados juntos conforman el espacio muestral S , formado de hombres y mujeres. Como los daltónicos pueden ser hombres o mujeres, el evento A , que es que una persona sea daltónica, está formado de los eventos simples que estén en A y además en B y de los eventos simples que estén en A y además en B^c . Ya que estas dos son *intersecciones mutuamente excluyentes*, puede escribir el evento A como

$$A = (A \cap B) \cup (A \cap B^c)$$

y

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^c) \\ &= .04 + .002 = .042 \end{aligned}$$

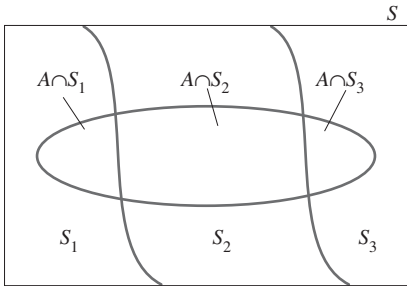
Suponga ahora que el espacio muestral se divide en k subpoblaciones, $S_1, S_2, S_3, \dots, S_k$, que, al igual que en el ejemplo de daltonismo, son **mutuamente excluyentes y exhaustivos**; esto es, tomados juntos conforman todo el espacio muestral. De un modo semejante, se puede expresar un evento A como

$$A = (A \cap S_1) \cup (A \cap S_2) \cup (A \cap S_3) \cup \dots \cup (A \cap S_k)$$

Entonces

$$P(A) = P(A \cap S_1) + P(A \cap S_2) + P(A \cap S_3) + \dots + P(A \cap S_k)$$

FIGURA 2.15
Descomposición del
evento A



Avance un paso más y use la regla de la multiplicación para escribir $P(A \cap S_i)$ como $P(S_i)P(A|S_i)$, para $i = 1, 2, \dots, k$. El resultado se conoce como la **ley de probabilidad total**.

LEY DE PROBABILIDAD TOTAL

Dado un conjunto de eventos $S_1, S_2, S_3, \dots, S_k$ que son mutuamente excluyentes y exhaustivos y un evento A, la probabilidad del evento A se expresa como

$$P(A) = P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + P(S_3)P(A|S_3) + \dots + P(S_k)P(A|S_k)$$

EJEMPLO 2.25

Los zapatos tenis ya no son sólo para jóvenes. De hecho, casi todos los adultos tienen varios pares. La tabla 2.6 da la fracción de adultos estadounidenses de 20 años de edad o más que tienen cinco o más pares de zapatos tenis en buen estado, junto con la fracción de adultos estadounidenses de 20 años o más en cada uno de los cinco grupos de edad.⁴ Use la Ley de probabilidad total para determinar la probabilidad incondicional de un adulto de 20 años de edad o más que tenga cinco o más pares de zapatos tenis en buen estado.

TABLA 2.6 Tabla de probabilidad

	Grupos y edades				
	G_1	G_2	G_3	G_4	G_5
	20–24	25–34	35–49	50–64	≥ 65
Fracción con ≥ 5 pares	.26	.20	.13	.18	.14
Fracción de adultos de 20 años o más	.09	.18	.30	.25	.18

Solución Sea A el evento de que una persona seleccionada al azar de entre una población de adultos de 20 años de edad y mayores tenga cinco o más pares de zapatos tenis en buen estado. Con $G_1, G_2, \dots, G_5$ represente el evento de que la persona seleccionada pertenezca a cada uno de los cinco grupos de edades, respectivamente. Como los cinco grupos son *exhaustivos*, se expresa el evento A como

$$A = (A \cap G_1) \cup (A \cap G_2) \cup (A \cap G_3) \cup (A \cap G_4) \cup (A \cap G_5)$$

Usando la ley de probabilidad total, encuentre la probabilidad de A como

$$\begin{aligned} P(A) &= P(A \cap G_1) + P(A \cap G_2) + P(A \cap G_3) + P(A \cap G_4) + P(A \cap G_5) \\ &= P(G_1)P(A|G_1) + P(G_2)P(A|G_2) + P(G_3)P(A|G_3) \\ &\quad + P(G_4)P(A|G_4) + P(G_5)P(A|G_5) \end{aligned}$$



De las probabilidades de la tabla 2.6,

$$\begin{aligned} P(A) &= (.09)(.26) + (.18)(.20) + (.30)(.13) + (.25)(.18) + (.18)(.14) \\ &= .0234 + .0360 + .0390 + .0450 + .0252 = .1686 \end{aligned}$$

La *probabilidad incondicional* de que una persona, seleccionada de dicha población de adultos de 20 años de edad y mayores, tenga al menos cinco pares de zapatos tenis en buen estado es de alrededor de .17. Observe que la Ley de probabilidad total es un promedio ponderado de las probabilidades dentro de cada grupo, con pesos .09, .18, .30, .25 y .18, que refleja los tamaños relativos de los grupos.

Con frecuencia es necesario hallar la probabilidad condicional de un evento B , dado que un evento A ha ocurrido. Una de estas situaciones ocurre al hacer exámenes de selección, que solían estar asociados principalmente con exámenes médicos de diagnóstico pero que ahora están encontrando aplicaciones en varios campos de actividad. Se emplea equipo automático de prueba para inspeccionar piezas en procesos de alto volumen de producción. Los exámenes de esteroides en atletas, los exámenes caseros de embarazo y los exámenes para detectar SIDA son algunas otras aplicaciones. Los exámenes de selección se evalúan sobre la probabilidad de un falso negativo o un falso positivo y estas dos son *probabilidades condicionales*.

Un **falso positivo** es el evento de que el examen sea positivo para una condición determinada, dado que la persona no tiene la condición. Un **falso negativo** es el evento de que el examen sea negativo para una condición determinada, dado que la persona tiene la condición. Se evalúan estas probabilidades condicionales usando una fórmula derivada por el probabilista Thomas Bayes.

El experimento comprende seleccionar una muestra de una de k subpoblaciones que sean mutuamente excluyentes y exhaustivas. Cada una de estas subpoblaciones, denotada por $S_1, S_2, \dots, S_k$, tiene una probabilidad de selección $P(S_1), P(S_2), P(S_3), \dots, P(S_k)$, llamadas *probabilidades previas*. Se observa un evento A en la selección. ¿Cuál es la probabilidad de que la muestra provino de la subpoblación S_i , dado que A ha ocurrido?

De la sección 2.6 se sabe que $P(S_i|A) = [P(A \cap S_i)]/P(A)$, que se puede reescribir como $P(S_i|A) = [P(S_i)P(A|S_i)]/P(A)$. Usando la Ley de probabilidad total para reescribir $P(A)$, tenemos

$$P(S_i|A) = \frac{P(S_i)P(A|S_i)}{P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + P(S_3)P(A|S_3) + \dots + P(S_k)P(A|S_k)}$$

Es frecuente que estas nuevas probabilidades se conozcan como *probabilidades posteriores*, es decir, probabilidades de las subpoblaciones (también llamadas *estados de naturaleza*) que se han actualizado después de observar la información muestral contenida en el evento A . Bayes sugirió que si las probabilidades previas son desconocidas, se pueden tomar como $1/k$, lo cual implica que cada uno de los eventos S_1 a S_k es igualmente probable.

TEOREMA DE BAYES

Con $S_1, S_2, \dots, S_k$ representemos k subpoblaciones mutuamente excluyentes y exhaustivas con probabilidades previas $P(S_1), P(S_2), \dots, P(S_k)$. Si ocurre un evento A , la probabilidad posterior de S_i dado A es la probabilidad condicional

$$P(S_i|A) = \frac{P(S_i)P(A|S_i)}{\sum_{j=1}^k P(S_j)P(A|S_j)}$$

para $i = 1, 2, \dots, k$.

EJEMPLO

2.26

Encuentre la probabilidad de que una persona seleccionada tuviera 65 años de edad o más, dado que la persona poseía al menos cinco pares de zapatos tenis en buen estado.

Solución Es necesario encontrar la probabilidad condicional dada por

$$P(G_5|A) = \frac{P(A \cap G_5)}{P(A)}$$

Ya se ha calculado $P(A) = .1686$ usando la Ley de probabilidad total. Por tanto,

$$\begin{aligned} P(G_5|A) &= \frac{P(G_5)P(A|G_5)}{P(G_1)P(A|G_1) + P(G_2)P(A|G_2) + P(G_3)P(A|G_3) + P(G_4)P(A|G_4) + P(G_5)P(A|G_5)} \\ &= \frac{(.18)(.14)}{(.09)(.26) + (.18)(.20) + (.30)(.13) + (.25)(.18) + (.18)(.14)} \\ &= \frac{.0252}{.1686} = .1495 \end{aligned}$$

En este caso, la probabilidad posterior de .15 es un poco mayor que la probabilidad previa de .13 (de la tabla 2.6). Este grupo *a priori* fue el segundo más pequeño y sólo una pequeña proporción de este segmento tenía cinco o más pares de zapatos tenis en buen estado.

¿Cuál es la probabilidad posterior de quienes tienen de 35 a 49 años de edad? Para este grupo de adultos, tenemos

$$\begin{aligned} P(G_3|A) &= \frac{(.30)(.13)}{(.09)(.26) + (.18)(.20) + (.30)(.13) + (.25)(.18) + (.18)(.14)} \\ &= \frac{.0390}{.1686} = .2313 \end{aligned}$$

Esta probabilidad posterior de .23 es considerablemente menor que la probabilidad previa de .30. En efecto, este grupo fue *a priori* el mayor segmento de la población muestreada, pero, al mismo tiempo, la proporción de individuos de este grupo que tenía al menos cinco pares de zapatos tenis en buen estado era la más pequeña de cualquiera de los grupos. Estos dos hechos tomados juntos causan un ajuste hacia abajo de casi un tercio de la probabilidad *a priori* de .30.

2.12

EJERCICIOS

TÉCNICAS BÁSICAS

2.60 Teorema de Bayes Se selecciona una muestra de una de dos poblaciones, S_1 y S_2 , con probabilidades $P(S_1) = .7$ y $P(S_2) = .3$. Si la muestra se ha seleccionado de S_1 , la probabilidad de observar un evento A es $P(A|S_1) = .2$. Del mismo modo, si la muestra se ha seleccionado de S_2 , la probabilidad de observar A es $P(A|S_2) = .3$.

- Si se selecciona una muestra al azar de una de las dos poblaciones, ¿cuál es la probabilidad de que ocurra el evento A ?
- Si la muestra se selecciona al azar y se observa el evento A , ¿cuál es la probabilidad de que la muestra sea seleccionada de la población S_1 ? ¿Y de la población S_2 ?

2.61 Teorema de Bayes II Si se realiza un experimento, puede ocurrir uno y sólo uno de los tres eventos mutuamente excluyentes S_1 , S_2 y S_3 , con estas probabilidades:

$$P(S_1) = .2 \quad P(S_2) = .5 \quad P(S_3) = .3$$

Las probabilidades de que ocurra un cuarto evento A , dado que ocurre el evento S_1 , S_2 o S_3 , son

$$P(A|S_1) = .2 \quad P(A|S_2) = .1 \quad P(A|S_3) = .3$$

Si se observa el evento A , encuentre $P(S_1|A)$, $P(S_2|A)$ y $P(S_3|A)$.

2.62 Ley de probabilidad total Una población se divide en dos subgrupos que se presentan con probabilidades de 60% y 40%, respectivamente. Un evento A ocurre 30% de las veces en el primer subgrupo y 50% en el segundo subgrupo. ¿Cuál es la probabilidad incondicional del evento A , cualquiera que sea el subgrupo de donde provenga?

APLICACIONES

2.63 Error de un trabajador Una máquina operada por un trabajador produce un artículo defectuoso con probabilidad .01 si el trabajador sigue exactamente las instrucciones de operación de la máquina y

con probabilidad .03 si no las sigue. Si él sigue las instrucciones 90% de las veces, ¿qué proporción de todos los artículos producidos por la máquina será defectuosa?

2.64 Seguridad en un aeropuerto Suponga que, en una ciudad en particular, el aeropuerto A maneja 50% de todo el tráfico aéreo y los aeropuertos B y C manejan 30% y 20%, respectivamente.

Los porcentajes de detección de armas en los tres aeropuertos son .9, .8 y .85, respectivamente. Si se encuentra un pasajero en uno de los aeropuertos llevando un arma por la puerta de abordar, ¿cuál es la probabilidad de que el pasajero esté usando el aeropuerto A ? ¿Y el aeropuerto C ?

¿EVENTOS MUTUAMENTE EXCLUYENTES O INDEPENDIENTES?

2.13

¿Dos eventos pueden ser mutuamente excluyentes e independientes al mismo tiempo?

Si los eventos A y B son independientes, se cumple $P(B) = P(B|A)$. Esto significa que si el número de elementos del conjunto B es 10% del número de elementos del conjunto universal; entonces, el número de elementos del subconjunto $(A \cap B)$ debe ser también igual a 10% del número de elementos del subconjunto A . Esto quiere decir que $(A \cap B)$ debe contener algunos elementos. $P(B) = P(B|A)$ implica que:

$$\frac{n(B)}{n(\Omega)} = \frac{n(\text{diagrama})}{n(A)}; \text{ donde diagrama representa } A \cap B$$

Por otro lado, si los eventos A y B son mutuamente excluyentes, no pueden ocurrir asociadamente y el subconjunto $(A \cap B)$ debe ser un conjunto nulo (vacío). Es evidente que ambas situaciones no pueden ocurrir al mismo tiempo; por consiguiente, *dos eventos no pueden ser mutuamente excluyentes e independientes al mismo tiempo.*[†]

Otra incongruencia es que, cuando los eventos A y B son independientes $P(B|A) = P(B)$, de cualquier forma; entonces, los eventos A y B son mutuamente excluyentes: $P(B|A) = 0$, $P(B \cap A) = 0$.

Otro modo de percibir la relación es pensar en términos de una tabla que describa los tipos de relación que pueden existir entre dos eventos (tabla 2.7).

La ocurrencia o no ocurrencia de un evento de ninguna forma afecta la probabilidad condicional de ocurrencia del otro evento, cuando son independientes:

$$P(B) = P(B|A) = P(B|\bar{A})$$

Considere un espacio muestral que consta de 100 elementos parecidos y dos subconjuntos: el subconjunto A con 20 elementos y el B con 40.

1. Si A y B son independientes

$$\left. \begin{array}{l} \text{Aunque sean } E \text{ el evento ser hombre y } F \\ \text{el de ser mujer} \end{array} \right\} \Rightarrow \begin{array}{l} E \cap F = \emptyset \\ \text{y} \\ E \text{ y } F \text{ son independientes} \end{array}$$

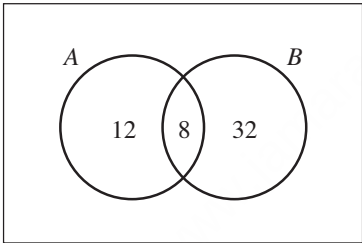
[†] Sólo cuando $P(B) = 0$.

$$P(A) = P(A \mid B) = P(A \mid \overline{B}) = \frac{20}{100} = \frac{8}{40} = \frac{12}{60}$$

TABLA 2.7 Tipos de relaciones que pueden existir entre dos eventos

Mutuamente excluyente	Dependiente	
	No mutuamente excluyente	Independiente
La ocurrencia de un evento en forma automática causa que la probabilidad condicional del segundo evento sea cero: $P(B \mid A) = 0$	La ocurrencia o no de un evento afecta la probabilidad condicional del otro evento que se da, aunque no incluye la ocurrencia de los otros eventos; por ejemplo, no causa que la probabilidad condicional del otro evento sea cero: $P(B) \neq P(B \mid A) \neq P(B \mid \overline{A}) \neq 0$	La ocurrencia o no ocurrencia de un evento no afecta la probabilidad condicional de ocurrencia del otro: $P(B) = P(B \mid A) = P(B \mid \overline{A})$

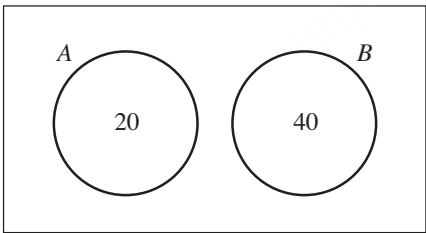
FIGURA 2.16



2. Si A y B son mutuamente excluyentes

$$P(A) = \frac{200}{100} \neq P(A \mid B) = \frac{0}{40} = 0$$

FIGURA 2.17



3. Si A y B son eventos dependientes, pero no mutuamente excluyentes, se tiene

$$P(A) = \frac{20}{100} \neq P(A \mid B),$$

donde puede haber una intersección y el número de elementos contenidos en $A \cap B$ es cualquier número entero no negativo hasta 20, excepto 0 (lo cual significaría que no hay intersección), y que los eventos son mutuamente excluyentes, y ocho (lo cual significaría que los eventos son independientes). Se observa que el número de elementos en la intersección nunca excede el número de elementos en el más pequeño de los dos subconjuntos, en este caso 20.

En suma, se presentan los siguientes casos:

1. Considere dos eventos, A y B , cuyas probabilidades $P(A)$ y $P(B) \neq 0$, y son mutuamente excluyentes:

$$\begin{aligned} \text{Por definición, } A \cap B &= \emptyset \\ P(A \cap B) &= 0 \end{aligned}$$

Para que A y B sean independientes deben cumplir $P(A) \times P(B) = P(A \cap B)$, pero como $P(A) \neq 0$, y $P(B) \neq 0$

$$P(A) \times P(B) \neq 0 \Rightarrow P(A \cap B) \neq P(A) \times P(B)$$

En suma, A y B no pueden ser independientes, cuando son mutuamente excluyentes.

2. Sean A y B dos eventos independientes; por definición se tiene:

$$P(A) \times P(B) = P(A \cap B),$$

ya que $P(A) \neq 0$, $P(B) \neq 0$ y $P(A \cap B) \neq 0$

$A \cap B \neq \emptyset$, por tanto, A y B no son mutuamente excluyentes.

2.14

PROCESOS ESTOCÁSTICOS

Estado: cada una de las situaciones o maneras mutuamente excluyentes, en las cuales se puede encontrar un sistema.

EJEMPLO

2.27

1. Una computadora puede estar en alguno de los siguientes estados: *en funcionamiento*, *descompuesta* o *en reparación*.

EJEMPLO

2.28

2. Una persona puede encontrarse en alguna de estas situaciones: *enferma*, *sana* o *muerta*.

Estado absorbente: situación a la cual ha llegado un sistema y de la cual ya no puede pasar a otro estado. Además, un estado absorbente sólo puede tener entradas, pero no salidas. En el ejemplo anterior, cuando la persona llega al estado de muerte, ya no puede pasar a ningún otro. Por tanto, la muerte es un estado absorbente.

Etapas: periodo que dura el desarrollo de una acción o proceso (establecido previamente por el investigador). Al término de cada etapa se observa el sistema, para determinar el estado en el cual se encuentra. El periodo puede abarcar una unidad convencional (segundo, día, mes, año, entre otros) o el tiempo que toma un lance del juego, o bien, la acción que corresponda según el sistema en estudio.

Transición: cualquier cambio de estado es una transición. En el ejemplo de la computadora, la transición es cuando pasa de estar en funcionamiento a descomponerse, o viceversa.

Probabilidad de transición: es la probabilidad de que se produzca el cambio de un estado a otro. Lo ideal es que ésta, que puede determinarse, permanezca constante en el transcurso de la vida de un sistema.

Vector de probabilidad: vector que representa la probabilidad de que el sistema se encuentre en un estado particular. En el ejemplo de la computadora, el vector de probabilidad inicial es $(1, 0, 0)$, donde la probabilidad 1 señala la computadora funcional y que hay probabilidad 0 de que esté descompuesta o en reparación. Un vector de probabilidad, que podría presentarse en algún momento de la vida del sistema, es $(0.6, 0.1, 0.3)$. Esto significa que la computadora tiene una probabilidad de 0.6 (60%) de estar en funcionamiento, de 0.1 (10%) de estar descompuesta y de 0.3 (30%) de estar en reparación.

Proceso estocástico: sucesión de las diferentes etapas de un sistema, fenómeno o actividad. El estado en que se encuentra el sistema en cada una de las etapas ocurre en forma aleatoria.

Cadena de Markov: proceso estocástico mediante el cual el estado de un sistema puede predecirse con cierta probabilidad a partir del conocimiento de su estado anterior. La probabilidad de transición entre los estados deberá permanecer constante en el transcurso de la vida del sistema.

Cadena de Markov absorbente: cadena de Markov que tiene uno o varios estados absorbentes y es posible llegar a ellos. Cuando el sistema llega a uno de éstos, ya no puede salir de él.

Cadenas de Markov

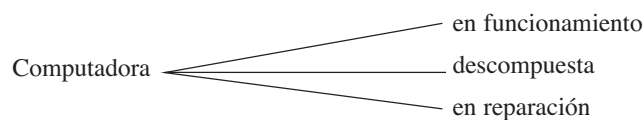
Es un experimento que consiste en una secuencia de ensayos, y cumple con las siguientes propiedades:

1. El resultado de cada ensayo (etapa) es uno de un conjunto finito de posibles resultados, llamados estados.
2. La probabilidad de cada resultado (estado) depende exclusivamente del ensayo (etapa) anterior.
3. La probabilidad de pasar de un resultado a otro permanecerá constante en el transcurso de todos los ensayos.

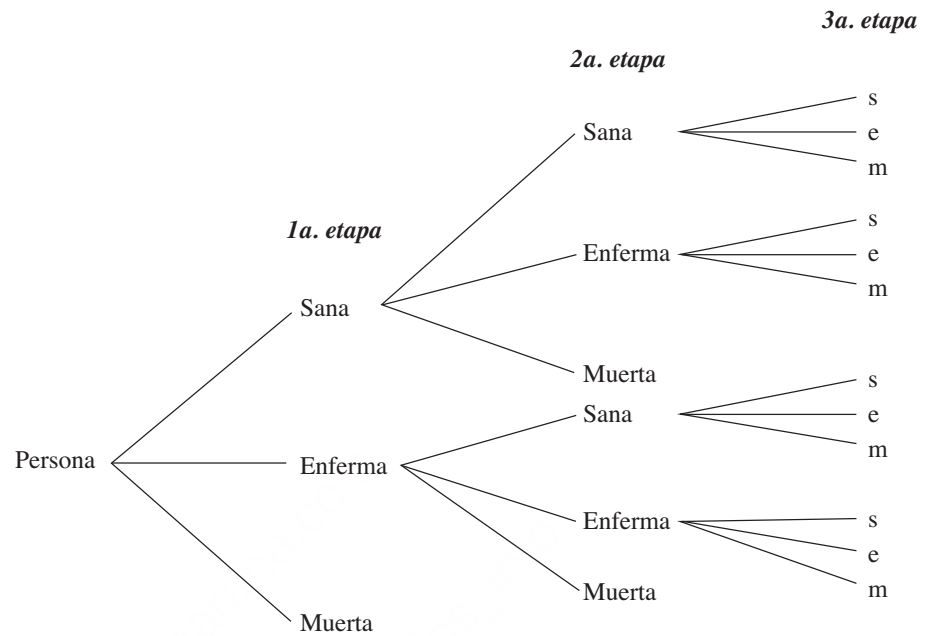
Para una definición más formal de cadena de Markov, véase Winston, Wayne L. (2005). *Investigación de operaciones. Aplicaciones y algoritmos* (4a. ed.). México : Cengage Learning Editores.

Representación gráfica

Para representar gráficamente los posibles cambios de los estados y sus etapas en un proceso estocástico, pueden utilizarse tanto los diagramas de árbol como de transición. Así, el diagrama de árbol que corresponde al ejemplo de la computadora es el siguiente:

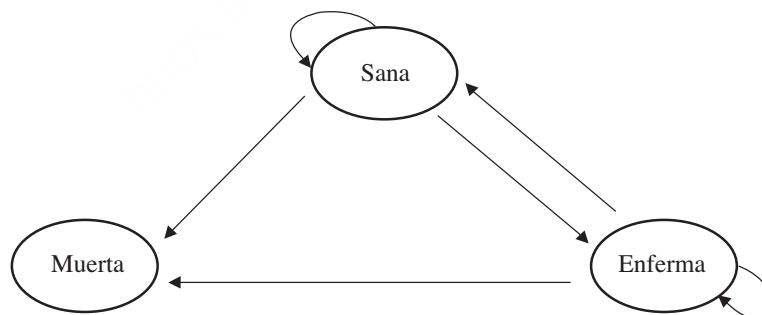


En el ejemplo de la persona, el diagrama es el siguiente:

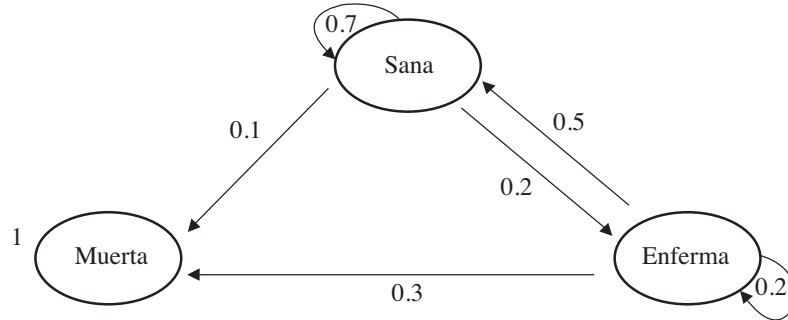


En la medida que se requieran mostrar más etapas, la construcción e interpretación del diagrama de árbol resultan menos prácticas, ya que aumentan el grado de complejidad y de confusión.

Para solucionar estos problemas, existe otro tipo de gráfica, denominada diagrama de transición, que es muy fácil de construir y lógica de entender:



En el diagrama, las flechas indican las posibles transiciones de un estado a otro. Además, para indicar las probabilidades de transición, en el diagrama se pone dicho valor junto a la flecha, por ejemplo:



El diagrama indica que una persona sana tiene en la siguiente etapa una probabilidad de 0.7 de estar sana, de 0.2 de estar enferma y de 0.1 de estar muerta. Por otra parte, una persona enferma se encontrará en la siguiente etapa sana, enferma o muerta, con sus respectivas probabilidades de 0.5, 0.2 y 0.3. Por último, una persona muerta seguirá muerta en la siguiente etapa; por tanto, la probabilidad de permanecer muerta asciende a 1, mientras que la de estar sana o enferma en la siguiente etapa es 0.

Representación matricial

Un proceso estocástico completo puede representarse de modo más conveniente por una matriz, donde los renglones representan los estados iniciales y las columnas los finales. Debido a que los elementos de la matriz son las probabilidades de transición, a aquélla se le suele conocer como *matriz de transición* T , y puede formarse para cualquier cadena de Markov:

Sea:

$$\begin{array}{cc}
 & \text{Estado final} \\
 & \begin{array}{ccc} A & B & C \end{array} \\
 \begin{array}{c} \text{Estado inicial} \\ A \\ B \\ C \end{array} & \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix} = T
 \end{array}$$

Las características de la matriz de transición son las siguientes:

- Es estocástica, donde la suma de los elementos de cada uno de los renglones o columnas constituye la unidad.
- Es cuadrada, debido a que el número de estados iniciales es igual al de estados finales.
- $T^5 T_1$ representa la etapa 1 y T^n representa la n -ésima etapa del proceso estocástico.

EJEMPLO

2.29

La ruina de un jugador

Dos oponentes (un jugador y la banca) participan en un juego que consiste en arrojar un dado no cargado de seis caras. El jugador escoge los números pares (2, 4, 6) o los nones (1, 3, 5), por lo que cada uno de los oponentes tiene probabilidad de ganar de 50% (0.5).

Sea el caso cuando el jugador cuenta con dos pesos, en cada lance tiene que apostar uno y puede ganar o perder. Por otra parte, la banca cuenta con tres pesos. El jugador suspenderá cuando pierda sus dos pesos o cuando gane los tres de la banca. ¿Qué probabilidad hay de que el jugador participe más de seis veces?

Solución:

Los posibles estados en que puede estar el jugador son:

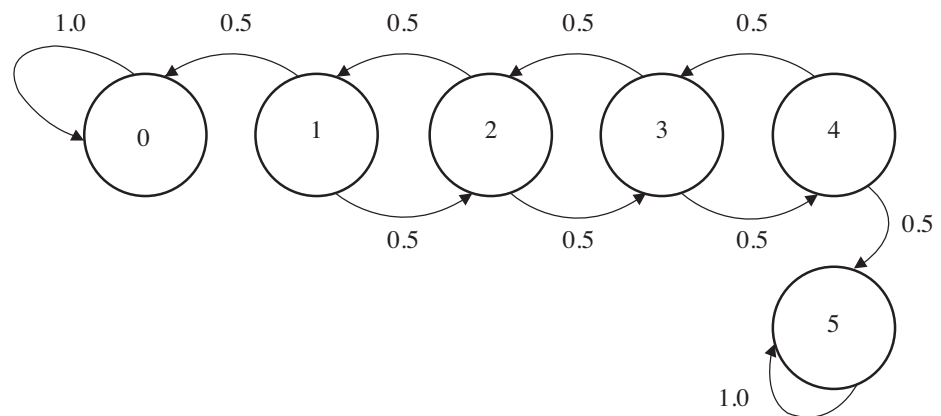
- 0 pesos (cuando pierde todo)
- 1 peso (cuando pierde un peso)
- 2 pesos (cuando inicia el juego)
- 3 pesos (cuando gana un peso a la banca)
- 4 pesos (cuando gana dos pesos a la banca)
- 5 pesos (cuando gana los tres pesos a la banca)

Cada una de estas situaciones posibles representan el número de pesos que tiene el jugador y se denominarán, respectivamente, estados 0, 1, 2, 3, 4 y 5.

En cada jugador “etapa” (gane o pierda), éste cambiará de estado. Los cambios de un estado a otro se llaman *transiciones*.

Como el jugador cuenta con dos pesos, se dice que inicia en el estado 2; a la siguiente jugada, se encontrará en el estado 3, si gana, o en el 1, si pierde. Si se halla en el estado 1, todavía le queda un peso por apostar (recuerde que sólo podrá abandonar el juego cuando gane o pierda todo). Si apuesta ese peso, puede ganar o perder en la siguiente jugada; en el primer caso se encontrará nuevamente en el estado 2, en el segundo, pasará al estado 0, y debe abandonar el juego, pues perdió todo. Del estado 2 puede transitar al 3, si gana, y, de ahí, al estado 4 y al 5, si se mantiene ganando.

Estas transiciones se notan claramente en el siguiente diagrama de transición del proceso, donde se aprecia que el jugador puede caer en los estados 0, 1, 2, 3, 4 o 5 (donde tiene 0, 1, 2, 3, 4 o 5 pesos):



En el diagrama se muestran las posibles transiciones y sus probabilidades asociadas entre los estados; por ejemplo, si el jugador se encuentra en el estado 2, hay una probabilidad de 0.5 de pasar al estado 1 (si pierde) y una probabilidad de 0.5 de pasar al 3 (si gana). Los estados 0 y 5 son en este caso los absorbentes, ya que cuando el jugador llega al estado 0 carece de dinero para jugar y, por tanto, acaba el juego, mientras que si llega al 5, significa que ganó los tres pesos de la banca y el juego termina.

Los estados absorbentes se identifican fácilmente en un diagrama de transición, por tener sólo entradas pero no salidas hacia otros estados.

Una vez que se tiene el diagrama de transición, se puede construir la matriz de transición:

$$T = \begin{array}{c|cccccc} \text{Hacia:} & & 0 & 1 & 2 & 3 & 4 & 5 \\ \text{De:} & 0 & \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \\ & 1 & & & & & & \\ & 2 & & & & & & \\ & 3 & & & & & & \\ & 4 & & & & & & \\ & 5 & & & & & & \end{array}$$

Ahora, vea cómo se interpreta la matriz:

- Se puede pasar del estado 0 hacia el 0 con una probabilidad de 1, ya que es un estado absorbente. Si el jugador no tiene dinero, así permanecerá.
- La probabilidad de pasar del estado 0 a cualquier otro estado es 0, ya que una vez sin dinero, el jugador no puede apostar y entonces no podrá ganar.
- La probabilidad de pasar del estado 1 al estado 0 es 0.5, ya que, si el jugador tiene un peso, la probabilidad de perderlo y quedar con 0 pesos es de 0.5.
- La probabilidad de tener un peso y conservarlo en el siguiente lance, es decir, pasar del estado 1 al 1 es 0, ya que al jugar, o gana un peso y pasa a 2, o pierde un peso y pasa a 0.
- La probabilidad de tener un peso y pasar a dos en el siguiente lance (o sea, el estado 2) es de 0.5, ya que hay una probabilidad de 0.5 de ganar.
- La probabilidad de pasar de 1 a 3, 4 o 5 es 0, ya que el jugador sólo puede ganar un peso; entonces, como máximo tendrá dos pesos en la próxima jugada. De manera análoga se interpretan los renglones que parten de 2, 3 y 4.
- Si se tienen cinco pesos, significa que el jugador ganó los tres del contrincante; entonces, la probabilidad de conservarlos es de 1, ya que el juego terminó al no haber más apuestas.

En una matriz de transición es tan fácil identificar los estados absorbentes como en los diagramas de transición: los estados que tengan 1 en la diagonal principal son los absorbentes.

En este caso, el estado 0 y el estado 5 tienen 1 en la diagonal principal, es decir, de 0 a 0 hay probabilidad de 1, y de 5 a 5 también 1. Por tanto, los estados 0 y 5 son absorbentes, tal como se había observado en el diagrama de transición.

En el problema, interesa saber qué probabilidad existe de que se jueguen más de seis lances. Con este fin habría que determinar la probabilidad de que el jugador esté en el estado 1, 2, 3 o 4 en el juego 6, o sea, en la transición 6 de la matriz (T^6). Entonces, se eleva la matriz T a la sexta potencia:

$$\begin{aligned}
 T &= \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix} \\
 T^2 &= \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.25 & 0 & 0.25 & 0 & 0 \\ 0.25 & 0 & 0.5 & 0 & 0.25 & 0 \\ 0 & 0.25 & 0 & 0.5 & 0 & 0.25 \\ 0 & 0 & 0.25 & 0 & 0.25 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix} \\
 T^3 &= \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.625 & 0 & 0.25 & 0 & 0.125 & 0 \\ 0.25 & 0.25 & 0 & 0.375 & 0 & 0.125 \\ 0.125 & 0 & 0.375 & 0 & 0.25 & 0.25 \\ 0 & 0.125 & 0 & 0.25 & 0 & 0.625 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix} \\
 T^4 &= \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.625 & 0.125 & 0 & 0.1875 & 0 & 0.0625 \\ 0.375 & 0 & 0.3125 & 0 & 0.1875 & 0.125 \\ 0.125 & 0.1875 & 0 & 0.3125 & 0 & 0.3750 \\ 0.0675 & 0 & 0.1875 & 0 & 0.125 & 0.625 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix} \\
 T^5 &= \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.6875 & 0 & 0.1563 & 0 & 0.0937 & 0.0625 \\ 0.3750 & 0.1563 & 0 & 0.25 & 0 & 0.2187 \\ 0.2187 & 0 & 0.25 & 0 & 0.1563 & 0.3750 \\ 0.0675 & 0.0937 & 0 & 0.1563 & 0 & 0.6870 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix} \text{ y finalmente } T^6 = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.6875 & 0.0781 & 0 & 0.1250 & 0 & 0.1094 \\ 0.4531 & 0 & 0.2031 & 0 & 0.125 & 0.2188 \\ 0.2188 & 0.1250 & 0 & 0.2031 & 0 & 0.4531 \\ 0.1094 & 0 & 0.1250 & 0 & 0.0781 & 0.6875 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix}
 \end{aligned}$$

La matriz T^6 indica la probabilidad que hay de pasar de un estado inicial a otro cualquiera, después de seis transiciones (en este caso, juegos).

Como se sabe que el jugador comienza con dos pesos, centre la atención en el renglón del estado 2:

Hacia:	0	1	2	3	4	5
De: 2	[0.4531	0	0.2031	0	0.125	0.2188]

Es decir:

- La probabilidad de que comience con dos pesos y tenga cero después de seis juegos es 0.4531
- La probabilidad de que comience con dos pesos y tenga uno después de seis juegos es 0
- La probabilidad de que comience con dos pesos y tenga dos después de seis juegos es 0.2031
- La probabilidad de que comience con dos pesos y tenga tres después de seis juegos es 0
- La probabilidad de que comience con dos pesos y tenga cuatro después de seis juegos es 0.125
- La probabilidad de que comience con dos pesos y tenga cinco después de seis juegos es 0.2188

Como se desea conocer la probabilidad de que se jueguen más de seis lances, deben considerarse únicamente los estados que permitirán al jugador proseguir: 1, 2, 3 y 4, ya que si cuenta con 1, 2, 3 o 4 pesos, continuará apostando. Los estados 0 y 5 no se toman en cuenta, ya que si se llega a cualquiera de éstos, el juego se detiene, al no tener dinero para apostar o al haber ganado todo el del contrincante.

Con ello, si se suman las probabilidades que permiten seguir jugando después de seis lances (estados 1, 2, 3 y 4), resulta:

$$0 + 0.2031 + 0 + 0.125 = 0.3281$$

Así, la probabilidad de que el jugador juegue más de seis veces es 0.3281

EJEMPLO

2.30

En una escuela preparatoria, cuyo ciclo escolar es de tres años, se realiza un estudio estadístico. De éste, se obtienen las siguientes proporciones (porcentajes) respecto de la situación escolar del alumnado: 15% de los estudiantes que ingresan en primer grado tuvieron que recursarlo; 60% pasaron a segundo grado, y 25% se dieron de baja. La situación de los estudiantes que cursan el segundo grado es la siguiente: 13% lo recursaron, 80% pasaron y 7% se dieron de baja. De los alumnos de tercer grado, 22% tuvieron que recursarlo, 8% se dieron de baja y 70% se graduaron.

La administración de la escuela desea saber la probabilidad de que un alumno se gradúe o se dé de baja, considerando el grado que cursa. Actualmente, el primer grado lo cursan 1 000 alumnos; el segundo, 820, y el tercero, 640. Si la Secretaría de Educación Pública recomienda a la escuela que se gradúen cuando menos 650 estudiantes, conservando el excelente nivel académico de la institución, calcule cuántos alumnos se graduarán de los 1000 que ingresan. Adicionalmente, si este número es menor que 650, determine el número de alumnos que tienen que ingresar para cumplir con la recomendación de las autoridades educativas.

Solución:

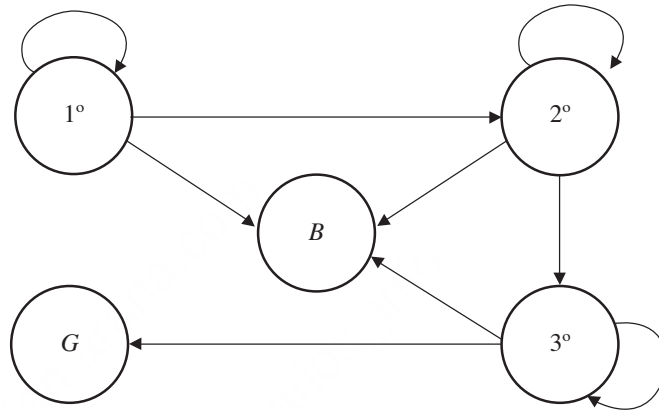
Paso 1. Construir el diagrama de transición que represente la situación en la que pueda encontrarse un alumno, considerando los estados siguientes:

- Primer año
- Segundo año
- Tercer año
- Baja
- Graduado

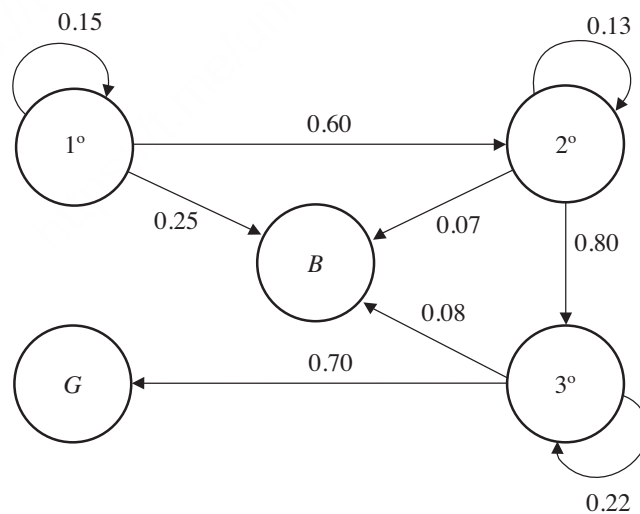
Las posibles transiciones son:

- Pasar al siguiente grado (siempre que esté en primero o segundo año)
- Repetir el grado
- Darse de baja
- Graduarse (si está en tercer año)

El diagrama de transición tiene la siguiente forma:



Considerando las probabilidades estimadas anteriormente, se tiene:



Paso 2. Construir la matriz de transición.

$$T = \begin{matrix} & \begin{matrix} 1^\circ & 2^\circ & 3^\circ & B & G \end{matrix} \\ \begin{matrix} 1^\circ \\ 2^\circ \\ 3^\circ \\ B \\ G \end{matrix} & \begin{pmatrix} 0.15 & 0.60 & 0 & 0.25 & 0 \\ 0 & 0.13 & 0.80 & 0.07 & 0 \\ 0 & 0 & 0.22 & 0.08 & 0.70 \\ 0 & 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & 0 & 1.0 \end{pmatrix} \end{matrix}$$

Paso 3. Determinar la matriz estacionaria. Se encuentra que a partir de T^{10} , la matriz se estaciona.

$$T^{10} = \begin{matrix} & \begin{matrix} 1^\circ & 2^\circ & 3^\circ & B & G \end{matrix} \\ \begin{matrix} 1^\circ \\ 2^\circ \\ 3^\circ \\ B \\ G \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0.4175 & 0.5825 \\ 0 & 0 & 0 & 0.1748 & 0.8252 \\ 0 & 0 & 0 & 0.1026 & 0.8974 \\ 0 & 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & 0 & 1.0 \end{pmatrix} \end{matrix}$$

Paso 4. De esta matriz se obtiene la probabilidad de que un estudiante se dé de baja o se gradúe, considerando el año que cursa. Para un alumno que cursa el primer año, la probabilidad de darse de baja asciende a 41.75% (0.4175), mientras que la de graduarse es 58.25% (0.5825). Si está en segundo año, las probabilidades del alumno son: 17.48% (0.1748) y 82.52% (0.8252) de darse de baja y graduarse, respectivamente. Por otra parte, un estudiante de tercer grado tiene una probabilidad de 10.26% (0.1026) de darse de baja y 89.74% (0.8974) de graduarse.

Paso 5. Actualmente ingresan 1000 alumnos en primer grado y la probabilidad de que se gradúen asciende a 58.25%. Así, egresarán $1000 \times 0.5825 = 582.5$ (583 alumnos), un número de graduados inferior al de 650, que recomienda la Secretaría de Educación Pública. ¿Qué acción podría emprenderse para incrementar el número de graduados?

Paso 6. Una solución es incrementar el número de alumnos de nuevo ingreso. En ese caso, si de 1000 estudiantes que ingresan, sólo 583 se gradúan, ¿cuántos alumnos deben inscribirse para que egresen al menos 650?
Si 1000 es a 582.5, ¿cuánto es a 650?

$$1000 \rightarrow 582.5$$

$$? \rightarrow 650$$

$$(1000 \times 650) \div 582.5 = 1115.88 \text{ (aproximadamente 1116)}$$

Con 1116 alumnos que se acepten en el primer grado se espera que se gradúen 650 ($1116 \times 58.25\% = 650.07$).

EJEMPLO

2.31

Los locales de un nuevo centro comercial están a punto de ofrecerse en alquiler. Asimismo, existe un centro comercial en servicio con características similares. El administrador decide implantar una norma según la cual el arrendatario deberá pagar un depósito por tres meses de alquiler. Así, el local podrá encontrarse en alguna de las siguientes situaciones:

- a) Al día en sus pagos.
- b) Atrasado con un pago.
- c) Atrasado con dos pagos.
- d) Atrasado con tres pagos.
- e) Disponible para ofrecerse en alquiler.

En el último inciso, pueden presentarse dos casos:

- 1. Se deben tres pagos y ninguno se efectúa posteriormente.
- 2. El arrendatario decide desocupar el local (en el caso de que adeude uno, dos o tres meses o, incluso, cuando se halle “al día” en el pago).

El centro comercial dispone de 130 locales, de los cuales 85 fueron alquilados a partir del día de la inauguración.

El administrador desea saber en qué situación se encontrarán estos 85 locales después de un año, es decir, el porcentaje en que estarán:

- a) Al día con sus pagos.
- b) Atrasado con uno, dos o tres pagos.
- c) Desocupado.

Asimismo, los 45 locales no alquilados el día de la inauguración, el administrador desea conocer el porcentaje que estará ocupado un año después.

Por otra parte, la experiencia durante varios años con el centro comercial en servicio ha proporcionado las estadísticas que siguen:

- De los locales al día con los pagos, 63% se mantiene así el siguiente mes, 35% deben una mensualidad y 2% estarán desocupados.
- De los locales que deben una mensualidad, 22% estarán al día el siguiente mes, 28% seguirán atrasados con un alquiler, 45% estarán atrasados con dos y 5% se desocuparán.
- De aquellos que están atrasados con dos pagos, 10% se pondrán al día el siguiente mes, 37% estarán atrasados únicamente con un pago, 33% con dos, 11% con tres y 9% estarán desocupados.
- De donde el pago se atrasó tres meses, sólo 6% se pondrán al día el siguiente mes, 34% deberán una mensualidad, 27% deberán dos pagos, 17% tres y 16% tendrán que desocupar, pues están atrasados con tres mensualidades y no realizan pago alguno.
- De los locales desocupados, 60% seguirán así al siguiente mes, mientras que 40% se ocuparán.

Además, el administrador desea determinar: el momento en que se detendrá el proceso, el porcentaje de los locales que permanecerá al día en sus pagos, el porcentaje que permanecerá desocupado, así como el porcentaje que no estará al corriente.

Solución:

Paso 1. Construir el diagrama de transición que represente la situación en la que puede encontrarse un local. Los estados respectivos son:

- Al día en sus pagos (0).
- Atrasado con un pago (1).
- Atrasado con dos pagos (2).
- Atrasado con tres pagos (3).
- Desocupado (D).

Las transiciones posibles entre los estados son:

a) Si un local está al corriente, al siguiente mes puede:

- Permanecer al día.
- No pagar y, por tanto, atrasarse con un pago.
- Desocuparse.

b) Si un local está atrasado con un pago, al siguiente mes puede:

- Estar al día (si paga el mes atrasado y el mes correspondiente).
- Seguir atrasado con un pago (si sólo paga un mes).
- Atrasarse con dos meses (si no efectúa pago alguno).
- Desocuparse.

c) Si un local está atrasado con dos pagos, al siguiente mes puede:

- Ponerse al día (si paga tres meses).
- Estar atrasado con un mes (si paga sólo dos).
- Seguir atrasado con dos meses (si paga uno solamente).
- Estar atrasado con tres meses (si no paga).
- Desocuparse.

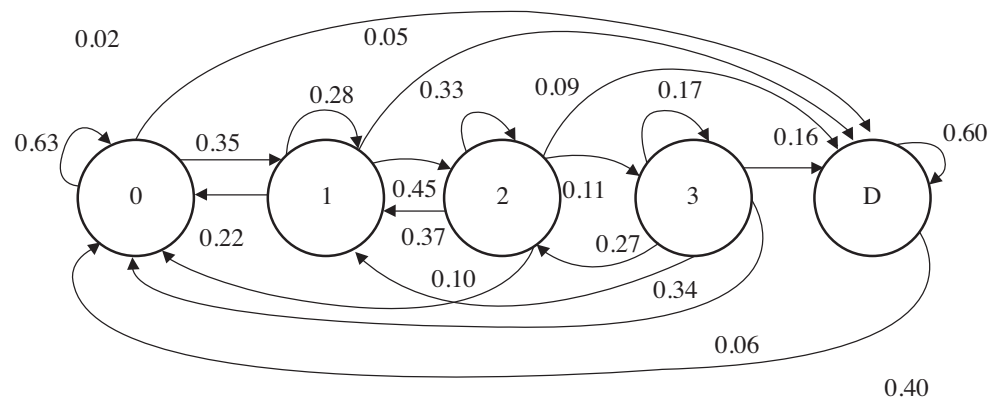
d) Si un local está atrasado con tres pagos, al siguiente mes puede:

- Ponerse al día (si paga cuatro meses).
- Estar atrasado con un mes (si paga tres).
- Estar atrasado con dos meses (si paga dos).
- Estar atrasado con tres meses (si paga uno).
- Desocuparse (si no paga).

e) Si un local está desocupado, al siguiente mes puede:

- Seguir desocupado.
- Estar ocupado.

El diagrama de transición respectivo es el siguiente:



Paso 2. Construir la matriz de transición:

$$T = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & D \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ D \end{matrix} & \begin{pmatrix} 0.63 & 0.35 & 0 & 0 & 0.02 \\ 0.22 & 0.28 & 0.45 & 0 & 0.05 \\ 0.10 & 0.37 & 0.33 & 0.11 & 0.09 \\ 0.06 & 0.34 & 0.27 & 0.17 & 0.16 \\ 0.40 & 0 & 0 & 0 & 0.60 \end{pmatrix} \end{matrix}$$

Paso 3. Obtener T^{12} , ya que esta matriz representará la situación de los locales en el mes 12:

$$T^{12} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & D \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ D \end{matrix} & \begin{pmatrix} 0.3569 & 0.2939 & 0.2088 & 0.0277 & 0.1126 \\ 0.3574 & 0.2938 & 0.2084 & 0.0276 & 0.1127 \\ 0.3576 & 0.2938 & 0.2083 & 0.0276 & 0.1127 \\ 0.3577 & 0.2939 & 0.2083 & 0.0276 & 0.1126 \\ 0.3576 & 0.2972 & 0.2086 & 0.0276 & 0.1120 \end{pmatrix} \end{matrix}$$

El administrador puede obtener de esta matriz el porcentaje de los 85 locales ocupados al momento de la inauguración. Estos locales comenzaron a partir de la inauguración, por lo que estaban al día (estado 0). En este caso, había que considerar el renglón 0:

$$D \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & D \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ D \end{matrix} & \begin{pmatrix} 0.3569 & 0.2939 & 0.2088 & 0.0277 & 0.1126 \end{pmatrix} \end{matrix}$$

- a) 35.69% de los 85 locales estarán al día en un año, esto es, aproximadamente 30 locales ($85 \times 0.3569 = 30.3365$)

- b) 29.39% estarán atrasados con un pago en un año, lo que da 25 locales aproximadamente ($85 \times 0.2939 = 24.9815$)
- 20.88% estarán atrasados dos pagos en un año, aproximadamente 18 locales ($85 \times 0.2088 = 17.748$)
- 2.77% estarán atrasados con tres pagos dentro de un año, es decir, dos locales aproximadamente ($85 \times 0.0277 = 2.3545$)
- c) 11.26% estarán desocupados dentro de un año, o sea, 10 locales aproximadamente ($85 \times 0.1126 = 9.571$)

En resumen, de los 85 locales, después de un año:

- 35.69% estarán al día (alrededor de 30 locales).
- 53.04% estarán atrasados con uno o varios pagos (aproximadamente 45 locales).
- 11.26% estarán desocupados (10 locales aproximadamente).

Ahora, el administrador desea saber qué porcentaje estará ocupado un año después de los 45 locales desocupados al momento de la inauguración. Por tanto, habrá que considerar el renglón D de T^{12} :

	0	1	2	3	D
D	0.3576	0.2942	0.2086	0.0276	0.1120

Los locales que estarán ocupados son los que se encuentran al corriente o deben algún pago, esto es, $35.76 + 29.42 + 20.86 + 2.76 = 88.8$ por ciento.

Una manera más simple de obtener este porcentaje es restar al 100% de los locales desocupados en la inauguración el porcentaje de locales desocupados después de un año, es decir, $100 - 11.20 = 88.8$ por ciento.

De los 45 locales, 88.8% equivale a 40 aproximadamente ($45 \times 0.888 = 39.96$). Esto es, alrededor de 40 de los 45 locales desocupados en la inauguración se habrá ocupado después de un año.

Por último, al administrador le interesa establecer el momento en que se estaciona el proceso, con el fin de obtener el porcentaje de locales al día, atrasados y desocupados. Para obtener esto, se continúa elevando la matriz hasta que se estacione; así, resulta que la matriz se estaciona en T^{20} :

$$T^{20} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & D \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ D \end{matrix} & \begin{pmatrix} 0.3573 & 0.2939 & 0.2085 & 0.0276 & 0.1126 \\ 0.3573 & 0.2939 & 0.2085 & 0.0276 & 0.1126 \\ 0.3573 & 0.2939 & 0.2085 & 0.0276 & 0.1126 \\ 0.3573 & 0.2939 & 0.2085 & 0.0276 & 0.1126 \\ 0.3573 & 0.2939 & 0.2085 & 0.0276 & 0.1126 \end{pmatrix} \end{matrix}$$

De aquí, se obtiene que a partir del mes 20:

- 35.73% de los locales estarán al día.
- 29.39% estarán atrasados con un pago.
- 20.85% estarán atrasados con dos pagos.
- 2.76% se atrasarán con tres pagos.
- 11.26% permanecerán desocupados.

Resumen

Tanto en situaciones profesionales como cotidianas, se toman decisiones enfrentando la incertidumbre o ignorancia.

El objetivo de analizar el concepto de probabilidad, es poder expresar de manera congruente, significativa y confiable la forma de concebir la incertidumbre respecto de la ocurrencia de los eventos. La tarea de definir la probabilidad puede considerarse mediante procedimientos diferentes pero, al mismo tiempo, relacionados:

En la interpretación de la probabilidad, se discuten tres enfoques:

- a) Clásica o de juegos, considera espacios muestrales equiprobables, que en la naturaleza no existen; por lo que no es posible obtener resultados igualmente probables, y esto además implica resultados a priori y por consiguiente es subjetiva.
- b) El enfoque subjetivo permite que el investigador asigne probabilidades, basados en su confianza y que está particularmente en su mente; por lo que varía de persona a persona, pero es en definitiva contraria a los seguidores de la escuela clásica, que no deseaban incorporar presentimientos subjetivos en probabilidad.
- c) La regularidad estadística, estudia los resultados de un fenómeno en condiciones constantes y en una muestra lo bastante grande, para que las proporciones de ocurrencia de dicho fenómeno sean constantes.

Ciertos principios de conteo que consideran permutaciones o combinaciones, ayudan a encontrar los resultados de un evento o fenómeno, sin necesidad de enumerar todos los posibles resultados.

Un conjunto de dos o más eventos es descrito, en la mayor parte de los casos, como independiente o dependiente. Si los eventos son dependientes, pueden ser o no mutuamente excluyentes. Algunas leyes y propiedades que rigen la probabilidad clásica, permiten calcular probabilidades a partir de otras probabilidades.

El teorema de Bayes, por tanto, es una de las leyes más importantes de la probabilidad; aunque aquí se ha desarrollado un concepto del particular, en el supuesto de que en el espacio muestral todos los resultados son igualmente probables, en realidad estas mismas propiedades y leyes no se aplican únicamente al modelo clásico.

Las distribuciones de probabilidad para las variables aleatorias pueden derivarse matemáticamente utilizando la teoría de probabilidad, o aproximarse de manera empírica repitiendo el experimento un gran número de veces y construyendo un histograma de la frecuencia relativa de los resultados. Estas distribuciones son modelos de la población y se caracterizan por la media y la desviación estándar.

Por último, no son las probabilidades numéricas el principal objeto de la teoría; su propósito es descubrir leyes generales y elaborar modelos teóricos satisfactorios.

2.14

EJERCICIOS

TÉCNICAS BÁSICAS

2.65 proyectiles dirigidos El porcentaje de falla para un sistema de control de proyectiles dirigidos es 1 en 1000. Suponga que un sistema de control duplicado, pero completamente independiente, se instala en cada proyectil para que, si el primero falla, el segundo tome el control. La confiabilidad de un proyectil es la probabilidad de que no falle. ¿Cuál es la confiabilidad del proyectil modificado?

2.66 Desgarres de la ACL/MCL *The American Journal of Sports Medicine* publicó un estudio de 810 jugadoras universitarias de rugby que tienen historias clínicas de lesiones en rodillas. Para estas atletas, las dos lesiones de rodilla comunes investigadas fueron torceduras del ligamento cruzado medio (MCL) y desgarres del ligamento cruzado anterior (ACL).⁵ Para las jugadoras de

las posiciones de defensas, se encontró que 39% tenían torceduras del MCL y 61% tenían desgarres del ACL. Para las jugadoras de las posiciones de delanteras, se encontró que 33% de ellas tenían torceduras del MCL y 67% tenían torceduras del ACL. Como un equipo de rugby está formado por ocho delanteras y siete defensas, se puede suponer que 47% de las jugadoras con lesiones en rodillas son defensas y 53% son delanteras.

- Encuentre la probabilidad incondicional de que una jugadora de rugby seleccionada al azar de este grupo haya experimentado una torcedura del MCL.
- Dado que se ha seleccionado una jugadora que tiene una torcedura del MCL, ¿cuál es la probabilidad de que la jugadora sea delantera?
- Dado que se ha seleccionado una jugadora que tiene un desgarre del ACL, ¿cuál es la probabilidad de que la jugadora sea defensa?

2.67 MRI Un artículo de *The American Journal of Sports Medicine* comparó los resultados de una evaluación de resonancia magnética (MRI), contra una evaluación de cirugía artroscópica de desgarres de cartílago, en dos sitios en las rodillas de 35 pacientes. Los exámenes de $2 \times 35 = 70$ produjeron las clasificaciones que se muestran en la tabla siguiente.⁶ Los desgarres reales fueron confirmados por examen de cirugía artroscópica.

	Desgarres	No desgarres	Total
MRI Positiva	27	0	27
MRI Negativa	4	39	43
Total	31	39	70

- ¿Cuál es la probabilidad de que un sitio seleccionado al azar tenga un desgarre y haya sido identificado como desgarre por la MRI?
- ¿Cuál es la probabilidad de que un sitio seleccionado al azar no tenga desgarre y haya sido identificado como que sí lo tiene?
- ¿Cuál es la probabilidad de que un sitio seleccionado al azar tenga desgarre y no haya sido identificado por la MRI?
- ¿Cuál es la probabilidad de una MRI positiva, dado que existe el desgarre?
- ¿Cuál es la probabilidad de un falso negativo, es decir, una MRI negativa, dado que existe el desgarre?

2.68 El juego en pares Cada uno de dos hombres tiran al aire una moneda. Obtienen un “par” si ambas monedas son caras o si ambas son cruces. Suponga que el tiro se repite tres veces.

- ¿Cuál es la probabilidad de tres pares?
- ¿Cuál es la probabilidad de que los seis tiros (tres para cada hombre) resulten en cruces?

- El tiro de monedas al aire da un modelo para muchos experimentos prácticos. Suponga que los tiros de monedas representan las respuestas dadas por dos estudiantes a tres preguntas específicas de verdadero o falso en un examen. Si los dos estudiantes dieron tres pares por respuestas, ¿la baja probabilidad determinada en la parte **a** sugiere una confabulación?

2.69 Negociaciones de contrato La experiencia ha demostrado que, 50% de las veces, una negociación particular entre empresa y sindicato llevó a un acuerdo antes de transcurridas dos semanas, 60% de las veces el fondo de huelga del sindicato era adecuado para apoyar una huelga y 30% de las veces ambas condiciones quedaron satisfechas. ¿Cuál es la probabilidad de un acuerdo de contrato dado que el fondo sindical para huelga es adecuado para soportar una huelga? ¿El acuerdo de un contrato antes de transcurridas dos semanas depende de si el fondo sindical de huelga es adecuado para soportar una huelga?

2.70 Permanencia en un trabajo Suponga que la probabilidad de permanecer 10 años o más con una compañía particular es $1/6$. Un hombre y una mujer empiezan a trabajar en la compañía el mismo día.

- ¿Cuál es la probabilidad de que el hombre trabaje ahí menos de 10 años?
- ¿Cuál es la probabilidad de que el hombre y la mujer trabajen ahí menos de 10 años? (Suponga que no tienen relación y sus tiempos de servicio son independientes entre sí.)
- ¿Cuál es la probabilidad de que uno u otro o ambos trabajen 10 años o más?

2.71 Tiempos de espera Suponga que en un supermercado particular la probabilidad de esperar 5 minutos o más en la fila para pagar es .2. En un día determinado, un hombre y su esposa deciden hacer compras individualmente en el mercado, cada uno saliendo en diferentes cajas de pago. Ambos llegan al mostrador al mismo tiempo.

- ¿Cuál es la probabilidad de que el hombre espere menos de 5 minutos para salir?
- ¿Cuál es la probabilidad de que el hombre y su esposa salgan en menos de 5 minutos? (Suponga que los tiempo de salida para los dos son eventos independientes.)
- ¿Cuál es la probabilidad de que uno o el otro o ambos esperen 5 minutos o más?

2.72 Control de calidad Un plan de control de calidad exige aceptar un lote grande de cojinetes para cigüeñal si se saca una muestra de siete y ninguno es defectuoso.

¿Cuál es la probabilidad de aceptar el lote si ninguno del lote es defectuoso? ¿Y si 1/10 son defectuosos? ¿Y si 1/2 son defectuosos?

2.73 Transporte colectivo Sólo 40% de todas las personas de una comunidad está a favor de desarrollar un sistema de transporte colectivo. Si se seleccionan al azar cuatro ciudadanos de la comunidad, ¿cuál es la probabilidad de que los cuatro estén a favor del sistema de transporte colectivo? ¿Y de que ninguno esté a favor de él?

2.74 Mediciones de presión sanguínea Un médico investigador comparó la efectividad de dos medicamentos A y B para la presión sanguínea, administrando los dos a cada uno de cuatro pares de gemelos idénticos. El medicamento A se dio a un miembro de un par; el medicamento B se dio al otro. Si, de hecho, no hay diferencia en los efectos de los medicamentos, ¿cuál es la probabilidad de que el descenso en la lectura de la presión sanguínea para el medicamento A sobrepase la caída correspondiente en la lectura del medicamento B , para los cuatro pares de gemelos? Suponga que el medicamento B creó un descenso más pronunciado en la presión sanguínea que el medicamento A , para cada uno de los cuatro pares de gemelos. ¿Piensa usted que esto es suficiente evidencia para indicar que el medicamento B es más eficaz para bajar la presión sanguínea que el medicamento A ?

2.75 Exámenes de sangre Para reducir el costo de detectar una enfermedad, los exámenes de sangre se realizan en una muestra agrupada de sangre tomada de un grupo de n personas. Si no hay indicio de la enfermedad presente en la muestra sanguínea de grupo, ninguno tiene la enfermedad. Si el análisis de la muestra sanguínea de grupo indica que la enfermedad está presente, cada individuo debe someterse a un examen de sangre. Los exámenes individuales son realizados en secuencia. Si, entre un grupo de cinco personas, una de ellas tiene la enfermedad, ¿cuál es la probabilidad de que seis exámenes de sangre (incluyendo el examen de grupo) se requieran para detectar a la persona enferma? Si dos personas tienen la enfermedad, ¿cuál es la probabilidad de que se requieran seis exámenes para localizar a ambas personas enfermas?

2.76 Horario flexible Una encuesta para determinar la disponibilidad de horarios de trabajo flexibles en el mercado laboral de California proporcionó la siguiente información para 220 empresas ubicadas en dos ciudades de California.

Ciudad	Horario flexible		
	Disponible	No disponible	Total
A	39	75	114
B	25	81	106
Totales	64	156	220

Se selecciona al azar una compañía de este grupo de 220 compañías.

- ¿Cuál es la probabilidad de que la compañía esté ubicada en la ciudad A ?
- ¿Cuál es la probabilidad de que la compañía esté ubicada en la ciudad B y ofrezca horarios flexibles de trabajo?
- ¿Cuál es la probabilidad de que la compañía no tenga horarios flexibles?
- ¿Cuál es la probabilidad de que la compañía esté ubicada en la ciudad B , dado que la compañía tiene horarios flexibles?

2.77 Experimento para reconocer colores Se realiza un experimento como sigue: los colores rojo, amarillo y azul se proyectan en una pantalla durante un breve periodo. Una persona ve los colores y se le pide elegir el que piense que duró más tiempo. El experimento se repite tres veces con la misma persona.

- Si todos los colores se proyectaron durante el mismo tiempo, encuentre la distribución de probabilidad para x , el número de veces que la persona eligió el color rojo. Suponga que sus tres selecciones son independientes.
- Construya el histograma de probabilidad para la variable aleatoria x .

2.78 Política en una orquesta El consejo de directores de una orquesta sinfónica principal ha votado por crear una comisión de músicos con el fin de manejar quejas de empleados. El consejo estará formado por el presidente y vicepresidente del consejo sinfónico y dos representantes de la orquesta. Los dos representantes de la orquesta serán seleccionados al azar de una lista de seis voluntarios, compuesta de cuatro hombres y dos mujeres.

- Encuentre la distribución de probabilidad para x , el número de mujeres elegidas como representantes de la orquesta.
- ¿Cuál es la probabilidad de que ambos representantes de la orquesta sean mujeres?
- Encuentre la media y varianza para la variable aleatoria x .

2.79 Independencia y mutuamente excluyentes

Suponga que $P(A) = .3$ y $P(B) = .4$.

- Si $P(A \cap B) = .12$, ¿ A y B son independientes? Justifique su respuesta.
- Si $P(A \cup B) = .7$, ¿qué es $P(A \cap B)$? Justifique su respuesta.
- Si A y B son independientes, ¿qué es $P(A|B)$?
- Si A y B son mutuamente excluyentes, ¿qué es $P(A|B)$?

Ejercicios suplementarios

2.80 En una caja hay 10 canicas rojas, cinco blancas y cinco azules. Si R es el conjunto de todas las canicas rojas, B el de las blancas y A el de las azules, calcule:

- a) N
- b) $P(R)$
- c) $P(B)$
- d) $P(A)$

2.81 Al ingresar al primer año de la escuela primaria, se les realiza un examen médico a niños de 6 años de edad. Suponga que hay 10% de probabilidad de que un niño tenga deficiencia visual específicamente en un ojo, mientras que la probabilidad de que sea en ambos ojos es 7%.

- a) ¿La deficiencia en ambos ojos resulta ser un evento independiente?
- b) Calcular la probabilidad de que ambos ojos presenten deficiencia.
- c) Calcular la probabilidad de que uno de los dos ojos esté afectado.

2.82 En un cesto hay 30 gatitos. Si va a sacarse uno al azar, y M es el evento “sacar una gatita” y H es “sacar un gatito”, calcule:

- a) N
- b) $n(M)$
- c) $n(H)$
- d) $P(M)$
- e) $P(H)$

2.83 Un experimento consiste en el lanzamiento de una moneda y un dado simultáneamente; si E_1 es el hecho de que salga “águila” en el lanzamiento de la moneda y E_2 es el hecho de obtener 3 o 6 en el lanzamiento del dado, explique qué significado tiene cada uno de los siguientes eventos:

- a) $\bar{E}_1$
- b) $\bar{E}_2$
- c) $P(E_1 \cap \bar{E}_2)$

2.84 En el supermercado existen cuatro marcas diferentes de agua purificada para beber, en presentación de un litro y medio. Si considera que éstas tienen la misma calidad y una persona desea adquirir una sin considerar las características en especial (A, B, C, D):

- a) Describa el espacio muestral y sus probabilidades correspondientes.
- b) Para el caso de que tenga que adquirir tres litros, ¿cuál será el espacio muestral y las probabilidades correspondientes si:

- i) los tres litros deben ser de marcas diferentes?
- ii) las marcas se pueden repetir?

2.85 Dados $S = \{1, 2, 3\}$, $A = \{1\}$, $B = \{3\}$, $C = \{2\}$, encuentre:

- a) $P(C)$
- b) $P(A \cup B)$
- c) $P(\bar{A})$
- d) $P(\bar{A} \cap \bar{B})$
- e) $P(\bar{A} \cup \bar{B})$
- f) $P(B \cup C)$

2.86 Dado un evento en el que $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{2}$, $P(A \cup B) = \frac{2}{3}$,

- a) $P(\bar{A})$
- b) $P(\bar{B})$
- c) $P(A \cap B)$
- d) $P(\bar{A} \cap \bar{B})$
- e) $P(\bar{A} \cup \bar{B})$
- f) $P(A \cap \bar{B})$
- g) $P(\overline{A \cup B})$
- h) $P(\bar{A} \cup B)$

2.87 Después de que 100 ratones corrieron por un laberinto, se detectaron los siguientes datos: 50 ratones eran machos, 50 fueron previamente entrenados, 42 tomaron a la izquierda en la primera oportunidad, 21 eran machos previamente entrenados, siete machos siguieron a la izquierda, 30 ratones previamente entrenados tomaron a la izquierda y 5 machos previamente entrenados se fueron a la izquierda.

- a) Ubique los datos en un diagrama de Venn.
- b) Ubique los datos en una tabla de doble entrada.
- c) Calcule la probabilidad de que sea hembra, esté previamente entrenada y haya tomado a la izquierda.
- d) Calcule la probabilidad de que sea macho, no esté entrenado y tomará a la izquierda.
- e) Calcule la probabilidad de que sea hembra, esté entrenada y no haya tomado a la izquierda.

2.88 Se realizó una entrevista a 885 amas de casa y se recabó la información siguiente:

600 veían telenovelas

400 series policiacas

620 programas deportivos

195 telenovelas y series policiacas

190 series policiacas y programas deportivos

500 telenovelas y programas deportivos
y 150 ven los tres programas.

Calcular la probabilidad de:

- a) Que vean únicamente telenovelas.
- b) Que vean telenovelas y series policiacas.
- c) Que no vea ningún programa.
- d) Que vean series policiacas y programas deportivos.
- e) Que vean series policiacas y programas deportivos, pero no telenovelas, o telenovelas y deportes, pero no policiacas.

2.89 A dos pacientes del mismo hospital, un hombre y una mujer, se les ha diagnosticado cáncer en el estómago. Los médicos especialistas pronostican que vivirán, cuando menos un año más, y la probabilidad es la siguiente: para el hombre 30% y para la mujer 40%.

Si se considera que los eventos son independientes, calcule que:

- a) Los dos vivan
- b) El hombre no viva
- c) La mujer no viva
- d) Ni la mujer ni el hombre vivan
- e) Al menos uno viva
- f) Ambos no vivan
- g) Ubique las probabilidades en un diagrama de Venn-Euler

2.90 Un noticiero anuncia el pronóstico del tiempo. Se estima la probabilidad de que llueva hoy $P(E_2) = 0.20$, de que llueva mañana $P(E_1) = 0.22$ y de que llueva hoy y mañana $P(E_1 \cap E_2) = 0.14$

a) Complete la siguiente tabla:

	Llueva mañana	No llueva mañana	Total
	E_1	$\bar{E}_1$	
Llueva hoy E_2	$(E_1 \cap E_2)$ 0.14		0.20
No llueva hoy $\bar{E}_2$		0.22	1

- b) Si la probabilidad de que llueva mañana (dado que llueva hoy), es $P(E_1 | E_2) = 0.7$ y de que no llueva mañana (dado que no llueva hoy), es 0.9, calcule la probabilidad de que llueva mañana dado que no llueva hoy, y de que llueva mañana dado que llueva hoy. Utilice un diagrama de árbol.
- c) Calcule la probabilidad de que llueva mañana.

2.91 En un examen de estadística, aplicado a 300 alumnos, se obtuvieron las calificaciones de la tabla siguiente:

Eventos	Hombres (E_6)	Mujeres (E_7)	Total
E_1 : MB	20	5	25
E_2 : B	15	50	65
E_3 : S	5	90	95
E_4 : NA	0	80	80
E_5 : NP	0	35	35

- a) Obtenga $P(E_1)$
- b) Obtenga $P(E_1 \cup E_6)$
- c) Obtenga $P(E_1 \cap E_6)$
- d) Obtenga $P(E_6 | \Omega)$
- e) $P(E_1 \cup E_6 | \Omega)$
- f) $P(E_1 | E_6)$
- g) $P(E_6 | E_1)$

donde $\Omega = \text{Universo}$

2.92 En el departamento de fotocopiado de una universidad existen tres copiadoras que fueron adquiridas al mismo tiempo, con las mismas características técnicas para una gran demanda de trabajo. Este tipo de copiadora está fuera de servicio 10% del tiempo de uso (por mantenimiento y reparación). Suponga la posibilidad de que ninguna de las fotocopadoras, cuando están fuera de servicio, dependa de la condición actual de las otras dos. El funcionamiento de cada una es independiente entre sí. Calcule la probabilidad de que:

- a) Las tres fotocopadoras estén fuera de servicio.
- b) La número 1 esté fuera de servicio, pero la 2 y 3 sigan funcionando.
- c) Las número 1 y 2 estén fuera de servicio y la 3 siga funcionando.
- d) Una de las tres esté fuera de servicio.
- e) Dos de las tres estén fuera de servicio.
- f) Al menos una esté fuera de servicio.
- g) Por lo menos dos estén fuera de servicio, donde: $x =$ fotocopadora 1 funcionando.

Obviamente, $\bar{x}$, $\bar{y}$, $\bar{z}$ significan que no funciona la fotocopadora, respectivamente.

2.93 Considerando el ejercicio 2.81, sea A el ojo derecho afectado y B el ojo izquierdo afectado.

- a) Calcular la probabilidad de que el ojo izquierdo esté afectado, dado que el ojo derecho también lo esté.
- b) Calcular la probabilidad de que el ojo izquierdo esté afectado, dado que el ojo derecho no.
- c) Si el riesgo relativo de B , dado A , se define como $p(B | A) / p(B | \bar{A})$, calcule el riesgo relativo de que el ojo izquierdo esté afectado porque el ojo derecho también lo esté.

2.94 La probabilidad de que una persona con síndrome de inmunodeficiencia adquirida tenga una reacción positiva es del 89%; la probabilidad de que una persona sin SIDA tenga una reacción positiva es del 2%; en una comunidad del

continente africano, el 20% tiene SIDA. Calcule todos los eventos posibles (aplique el teorema de Bayes).

2.95 Sean dos eventos A y B con su no ocurrencia $\bar{A}$ y $\bar{B}$ (los eventos A y B no son mutuamente excluyentes).

- Ubique las probabilidades en una tabla de dos entradas.
- Deduzca la probabilidad condicional correspondiente.
- Deduzca el teorema de Bayes correspondiente.

2.96 En una empresa los empleados tienen las características siguientes:

317 son hombres

316 son casados

25 son mujeres casadas sin profesión

72 son hombres casados sin profesión

83 son hombres profesionales solteros

15 son mujeres profesionales solteras

125 son hombres profesionales casados

49 son mujeres solteras sin profesión

- Calcule el número total de empleados.
- ¿Cuál es la probabilidad de que sea profesional?

2.97 Con base en su experiencia un médico ha recabado la siguiente información, relativa a las enfermedades de sus pacientes: 5% creen tener cáncer y lo tienen, 45% creen tener cáncer y no lo tienen, 10% no creen tenerlo y no lo tienen y, finalmente, 40% creen no tenerlo y es cierto. De entre los pacientes del médico, estos porcentajes implican las siguientes probabilidades para un paciente seleccionado al azar. Calcule:

- $P(\text{lo tenga} \mid \text{crea})$.
- $P(\text{lo tenga} \mid \text{no crea})$.
- $P(\text{lo crea} \mid \text{no tenga})$.
- $P(\text{lo crea} \mid \text{lo tenga})$.

2.98 Suponga que la ciencia médica ha desarrollado una prueba para el diagnóstico del cáncer que tiene 95% de exactitud, tanto en los que tienen cáncer como entre los que no lo tienen. Se sabe que el 0.005 de la población realmente tiene cáncer.

- Calcule la probabilidad de que determinado individuo tenga cáncer si la prueba dice que lo tiene.
- ¿Cuál es la probabilidad de que determinado individuo no tenga cáncer si la prueba dice que lo tiene?
- ¿Cuál es la probabilidad de que determinado individuo no tenga cáncer si la prueba dice que no lo tiene?
- ¿Cuál es la probabilidad de que determinado individuo tenga cáncer si la prueba dice que no lo tiene?

2.99 Un apostador de carreras de caballos estima que la probabilidad de que gane hoy es de 30%, la probabilidad de que gane con el mismo caballo la siguiente semana es

de 40%, y de que gane hoy y la siguiente semana es de 8%. Calcule todos los eventos posibles.

2.100 Se selecciona al azar un estudiante de una escuela preparatoria, en la cual 3% del total padece insomnio. Se aplica una prueba psicológica para detectar dicho padecimiento; la probabilidad de que el resultado sea positivo, dado que tiene insomnio, es de 95%; la probabilidad de que una persona sin insomnio, pero con resultados positivos, es de 2%, calcule todos los eventos posibles.

2.101 Un apostador de carreras de galgos estima que la probabilidad de que gane hoy es de 0.40, la probabilidad de que gane la siguiente semana es de 0.30 y de que gane hoy y la siguiente semana es 0.08. Calcule todos los eventos posibles.

2.102 Se aplica una prueba de aptitud para candidatos a ingresar en una universidad. Dicho examen lo aprueba el 60% de los aspirantes. De los que exentaron, 80% termina exitosamente el semestre; se selecciona una muestra aleatoria de los aspirantes que no aprobaron, pero cursan el semestre, y 50% de este grupo lo termina con éxito. Calcule todos los eventos posibles.

2.103 Veinte por ciento de los empleados de una compañía son profesionales. De éstos, 75% están asignados a la dirección general. De los que no son profesionales 20% está también en la dirección general.

- Elabore un diagrama de árbol.
- ¿Cuál es la probabilidad de que sea profesional y esté asignado a la dirección general?
- ¿Cuál es la probabilidad de que no sea profesional y no esté asignado a la dirección general?
- ¿Cuál es la probabilidad de que no sea profesional y esté asignado a la dirección general?
- ¿Cuál es la probabilidad de que sea profesional, dado que está asignado a la dirección general?
- ¿Cuál es la probabilidad de que esté asignado a la dirección general, dado que es profesional?
- ¿Cuál es la probabilidad de que no esté asignado a la dirección general, dado que es profesional?
- ¿Cuál es la probabilidad de que no sea profesional, dado que está asignado a la dirección general?
- ¿Cuál es la probabilidad de que esté asignado a la dirección general, dado que no es profesional?

2.104 Una psicóloga aplica un examen de aptitud para un trabajo técnico. Su experiencia es que la probabilidad de que un candidato pueda aprobar el examen es de 0.6. Si determinado candidato aprueba el examen, la probabilidad de que realice el trabajo satisfactoriamente es de 0.80. Si no pasa el examen, la probabilidad de que no realice el trabajo satisfactoriamente es de 0.40.

- Haga un diagrama de árbol.

- b) ¿Cuál es la probabilidad de que pase el examen, dado que realiza satisfactoriamente el trabajo?
- c) ¿Cuál es la probabilidad de que pase el examen, dado que no realiza satisfactoriamente su trabajo?
- d) ¿Cuál es la probabilidad de que realice satisfactoriamente su trabajo, dado que pasó el examen?
- e) ¿Cuál es la probabilidad de que no realice su trabajo satisfactoriamente, dado que pasó el examen?
- f) ¿Cuál es la probabilidad de que realice su trabajo satisfactoriamente, dado que no pasó el examen?
- g) ¿Cuál es la probabilidad de que no realice su trabajo satisfactoriamente, dado que no pasó el examen?

2.105 Las estudiantes de la Facultad de Psicología fueron clasificadas en altas, bajas, hermosas, listas, orgullosas y tímidas.

22 son hermosas, listas y altas

17 son hermosas, listas y bajas

13 son altas, orgullosas y listas

4 son orgullosas, altas y tímidas

18 son orgullosas, listas y bajas

11 son hermosas, tímidas y altas

7 son bajas, hermosas y tímidas

5 son tímidas, bajas y orgullosas

a) Haga una tabla de dos entradas.

b) Calcule todos los eventos posibles y sus probabilidades correspondientes.

2.106 En un vuelo de cierta línea aérea viajan 18 muchachos, 178 hombres, 10 mexicanos del sexo masculino, dos muchachos mexicanos, 26 personas de nacionalidad mexicana y 14 muchachas extranjeras, sin contar la tripulación.

a) Haga una tabla de dos entradas.

b) Describa todos los eventos posibles y sus probabilidades correspondientes.

2.107 En un hospital se desea realizar un estudio sobre colesterol en 1045 hombres sanos de 40 a 60 años de edad. A continuación, se presenta una tabla con los datos obtenidos.

Niveles de colesterol en hombres sanos (de 40 a 60 años)

Colesterol (mg/100 ml)				Frecuencia relativa (porcentaje) ^a	Frecuencia acumulada (porcentaje) ^a
LIR	LSR	MC	f		
119.5	139.5	129.5	10	1	1
139.5	159.5	149.5	21	2	3
159.5	179.5	169.5	37	3.5	6.5
179.5	199.5	189.5	97	9.3	15.8
199.5	219.5	209.5	152	14.5	30.3
219.5	239.5	229.5	206	19.7	50.0

239.5	259.5	249.5	195	18.7	68.7
259.5	279.5	269.5	131	12.5	81.2
279.5	299.5	289.5	96	9.2	90.4
299.5	319.5	309.5	47	4.5	94.9
319.5	339.5	329.5	30	2.9	97.8
339.5	359.5	249.5	13	1.2	99.0
359.5	379.5	369.5	6	0.6	99.6
379.5	399.5	389.5	4	0.4	100%
				100%	

^a Valores aproximados.

a) ¿Cuál es la probabilidad de que una persona tomada al azar tenga un nivel de colesterol por debajo de 160 mg/100 ml?

b) ¿Cuál es la probabilidad de que una persona tomada al azar tenga un nivel de colesterol mayor de 340 mg/100 ml?

c) ¿Cuál es la probabilidad de que una persona tomada al azar tenga un nivel menor de 160 mg/100 ml o mayor de 340 mg/100 ml?

2.108 En un centro de salud se desea hacer un estudio sobre las necesidades médicas y dentales que existen en su población, considerando también el sector donde laboran. Los resultados obtenidos se muestran en la tabla siguiente.

	Necesidades		
	Dental	Médica	Total
Público	470	280	750
Privado	110	140	250
Total	580	420	1000

a) ¿Cuál es la probabilidad de que una persona tomada al azar trabaje en el sector público?

b) ¿Cuál es la probabilidad de que una persona tomada al azar necesite atención médica?

c) Calcule la probabilidad de $P(A \cup B)$ (según los incisos anteriores).

d) ¿Cuál es la probabilidad de seleccionar una persona al azar que necesite servicio dental y trabaje en el sector privado?

2.109 Dadas las matrices siguientes, construya su diagrama de transición:

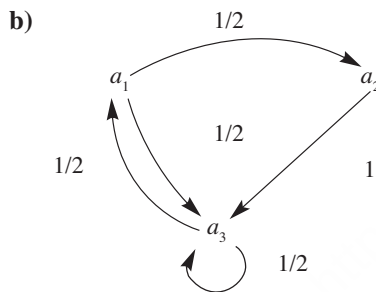
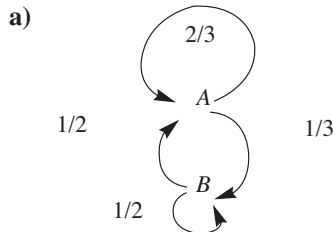
a)
$$\begin{bmatrix} 0.667 & 0.333 \\ 0.250 & 0.750 \end{bmatrix}$$

b)
$$\begin{bmatrix} 0.167 & 0.500 & 0.333 \\ 0.500 & 0.250 & 0.250 \\ 0.417 & 0.250 & 0.333 \end{bmatrix}$$

c)
$$\begin{bmatrix} 0.500 & 0.300 & 0.200 & 0.000 \\ 0.000 & 0.000 & 0.500 & 0.500 \\ 0.33\bar{3} & 0.33\bar{3} & 0.33\bar{3} & 0.33\bar{3}^a \\ 0.250 & 0.750 & 0.000 & 0.000 \end{bmatrix}$$

d)
$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

2.110 Dado el diagrama de transición, encuentre su matriz:



2.111 Cualquier empleado del área A tiene una probabilidad de permanecer ahí de 0.3 y de ser cambiado a B de 0.7; un empleado de B tiene 0.5 de probabilidad de ser transferido a A y 0.5 de ser transferido a C. Ningún empleado permanece en B por una segunda semana consecutiva. En el caso de los empleados del área C, la probabilidad es de 0.2 de ser transferidos a A, 0.3 de ser transferidos a B y 0.5 para permanecer en ella.

Considere cada reubicación del personal como un ensayo en el cual el resultado está controlado por probabilidades. Primero vea la naturaleza de las condiciones que controlan las transferencias; observe que la probabilidad de ser transferido a cada una de las áreas depende solamente del área a la cual es asignado el candidato en el tiempo de la transferencia. No es necesario considerar la cantidad de tiempo que el candidato ha permanecido en cada área en el pasado, asimismo el único efecto de asignación anterior es que la ubicación de la semana previa está en el área actual.

Note que esta política describe una secuencia de transferencias que se habrán de efectuar con el mismo

conjunto de probabilidades. El patrón de transferencia no varía.

a) Construya la matriz estocástica.

b) Haga el diagrama de transición.

c) Realice un diagrama de árbol.

2.112 Alina y Kitzia son dos jugadoras de tenis del mismo nivel y de características parecidas. Son escogidas al azar para sostener algunos juegos; sin embargo, cada vez que Alina gana un juego, adquiere mayor confianza en sí misma y la probabilidad de que gane el siguiente juego es 0.60. Por otro lado, cada vez que Alina pierde un juego se deprime y la probabilidad de que gane el siguiente disminuye a 0.30

a) ¿Cuál es la probabilidad de que Alina gane en el primer juego?

b) Construya la matriz de transición para Alina.

c) ¿Cuál es la probabilidad de que gane en el segundo juego?

d) ¿Cuál es la probabilidad de que gane en el tercer juego?

e) Construya un diagrama de árbol asignando las probabilidades correspondientes.

2.113 Se recopilieron datos durante 15 temporadas de lluvia en el norte de México. Una buena descripción de la ocurrencia de lluvias está dada por una cadena de Markov de orden dos. Los datos obtenidos fueron los siguientes:

		Día posterior	
		E_1	E_2
Día anterior	E_1	1049	350
	E_2	351	687

a) Calcule la matriz de transición, si E_1 = estado seco y E_2 = mojado.

b) Haga un diagrama de transición.

2.114 Por informaciones recientes se sabe que en la ciudad de México y zona metropolitana, 0.1% de sus habitantes transfirieron su residencia a provincia en un año; en cambio, 0.8% de los habitantes de provincia se mudaron a la ciudad de México y zona metropolitana. Si dichos porcentajes permanecen constantes, determine la distribución de probabilidad de los residentes de provincia (75%) y de la ciudad de México con su zona metropolitana (25%).

$$A = \begin{pmatrix} 0.25 & 0.75 \end{pmatrix}$$

^a El número $0.33\bar{3}$ significa que el número 3 se repite.

- a) Construya la matriz de transición.
- b) A los 3 años.
- c) A los 10 años.

2.115 Suponga que la probabilidad de que un padre con intolerancia a la leche herede a su hijo dicha deficiencia enzimática, es de 0.7; la probabilidad de otro padre sin ninguna deficiencia es de 0.4:

- a) ¿Cuál es la probabilidad de que el padre intolerante tenga un bisnieto similar a él?
- b) ¿Cuál es la distribución de probabilidad?
- c) ¿Depende de que los abuelos sufran la misma intolerancia?

2.116 Una ensambladora de automóviles promueve un modelo (E_1) de auto. El resultado de dicha promoción es que 75% de las personas que prefirieron ese modelo el primer año de su fabricación, adquieran el modelo reciente al siguiente año. De las personas que adquirieron otra marca (E_2) el año anterior, 35% cambia al modelo E_1 al siguiente año.

¿Qué porcentaje de automovilistas adquirirán el modelo E_1 después de dos años, si 50% lo adquiere ahora?

- a) Construya la matriz de transición.
- b) Dibuje un diagrama de árbol.
- c) Concluya.

2.117 En un estudio realizado con mujeres en la ciudad de México, se presentó la siguiente tabla de proporción respecto de la actitud de aceptación hacia el aborto:

Escolaridad de las hijas			
Escolaridad de las madres	Primaria	Bachillerato	Profesional
Primaria	0.486	0.699	0.011
Bachillerato	0.247	0.503	0.054
Profesional	0.068	0.484	0.448

- a) ¿Cuál será la actitud de la siguiente generación?
- b) ¿Cuál será la actitud de la quinta generación?

2.118 En una familia se observó, durante la primera generación, una enfermedad hereditaria (VIH). Suponga que cada síntoma puede ser considerado una unidad de tiempo, en una cadena de Markov con cuatro síntomas: sarcoma de Kaposi (S), cuadro gripal (G), neumocistitis (N), pérdida de peso (P).

A continuación se muestra la matriz de probabilidad de cada síntoma en la primera generación:

	S	G	N	P
S	0.75	0.10	0.10	0.05
G	0.05	0.75	0.15	0.05
N	0.20	0.40	0.30	0.10
P	0.10	0.30	0.20	0.40

Si comienza sus observaciones (primera generación) con sarcoma de Kaposi (S), calcule la probabilidad de que:

- a) Aparezca un cuadro clínico gripal en la próxima generación.
- b) El cuadro gripal aparezca en la segunda generación y sarcoma de Kaposi en la tercera.
- c) El mismo cuadro aparezca en la segunda o en la tercera generación, o en ambas.
- d) Aparezca primero en la tercera generación.

Algunas distribuciones de probabilidad importantes

OBJETIVOS GENERALES

Las variables aleatorias discretas se emplean en numerosas aplicaciones prácticas. En este capítulo presentamos tres variables aleatorias discretas importantes: la binomial, la de Poisson y la hipergeométrica. Es frecuente que estas variables aleatorias se usen para describir el número de sucesos de un evento especificado en un número fijo de intentos o una unidad fija de tiempo o espacio. Usted aprendió acerca de variables aleatorias discretas y sus distribuciones de probabilidad. En este capítulo estudiaremos las variables aleatorias continuas y sus distribuciones de probabilidad, así como una variable aleatoria continua muy importante, la normal. Usted aprenderá cómo calcular probabilidades normales y, en ciertas condiciones, cómo usar la distribución normal de probabilidad para aproximar la distribución binomial de probabilidad. En los capítulos siguientes verá la forma en que la distribución normal de probabilidad desempeña un papel central en la inferencia estadística.

ÍNDICE DEL CAPÍTULO

- Introducción (3.1)
- Variables aleatorias discretas y sus distribuciones de probabilidad (3.2)
- La distribución binomial de probabilidad (3.3)
- Distribuciones de probabilidad para variables aleatorias continuas (3.4)
- La distribución de probabilidad de Poisson (3.5)
- La distribución hipergeométrica de probabilidad (3.6)
- La distribución normal de probabilidad (3.7)
- Áreas tabuladas de la distribución normal de probabilidad (3.8)
- La aproximación de la distribución de probabilidad binomial a la normal (3.9)



NECESITO SABER...

Cómo utilizar la tabla 1 para calcular probabilidades binomiales
 Cómo usar la tabla 2 para calcular probabilidades de Poisson
 Cómo utilizar la tabla 3 para calcular probabilidades bajo la curva normal estándar
 Cómo calcular probabilidades binomiales usando la aproximación normal

Un misterio: casos de cáncer cerca de un reactor

¿El reactor nuclear Pilgrim I es responsable del aumento en los casos de cáncer en el área circundante? Surgió una controversia política cuando el Departamento de Salud Pública de Massachusetts encontró un número anormalmente grande de casos en una franja costera de cuatro millas de ancho un poco al norte del reactor nuclear de Plymouth, Massachusetts. El caso práctico, que aparece al final de este capítulo, examina cómo esta pregunta puede contestarse usando una de las distribuciones discretas de probabilidad presentadas aquí.

“¿Va a calificar por curva?”

“Calificar por curva” no necesariamente significa que recibirá una calificación más alta en un examen, ¡aunque a muchos estudiantes les gustaría pensar que sí! Calificar por curva en realidad se refiere a un método para asignar las calificaciones con las letras A, B, C, D o F usando proporciones fijas de las calificaciones correspondientes a cada una de las calificaciones con letra. Una de dichas técnicas para calificar por curva supone que la distribución de las calificaciones es aproximadamente normal y usa estas proporciones.

Calificación por letra	A	B	C	D	F
Proporción de calificaciones	10%	20%	40%	20%	10%

En el caso práctico al final de este capítulo se examinará ésta y otras proporciones asignadas para calificar por curva.

3.1

INTRODUCCIÓN

Es posible encontrar ejemplos de *variables aleatorias discretas* en numerosas situaciones cotidianas y en casi todas las disciplinas académicas. No obstante, hay tres distribuciones discretas de probabilidad que sirven de *modelos* para un gran número de estas aplicaciones. En este capítulo estudiamos las distribuciones de probabilidad binomial, de Poisson e hipergeométrica y discutimos su utilidad en diferentes situaciones físicas.

3.2

VARIABLES ALEATORIAS DISCRETAS
Y SUS DISTRIBUCIONES DE PROBABILIDAD

En el capítulo 1 las *variables* se definieron como características que cambian o varían con el tiempo y para diferentes personas u objetos en consideración. Las *variables cuantitativas* generan datos numéricos, en tanto que las *variables cualitativas* generan datos categóricos. No obstante, incluso las variables cualitativas generan datos numéricos si las categorías se codifican numéricamente para formar una escala. Por ejemplo, si se lanza al aire una sola moneda, el resultado cualitativo podría registrarse como “0” si es cara o como “1” si es cruz.

Variables aleatorias

Una variable x valuada numéricamente varía o cambia, dependiendo del resultado particular del experimento que se mida. Por ejemplo, suponga que se tira un dado y se mide x , el número observado en la cara superior. La variable x puede tomar cualquiera de seis valores: 1, 2, 3, 4, 5, 6, dependiendo del resultado *aleatorio* del experimento. Por esta razón, la variable x se conoce como **variable aleatoria**.

Definición Una variable x es **variable aleatoria** si el valor que toma, correspondiente al resultado de un experimento, es una probabilidad o evento aleatorio.

Se consideran numerosos ejemplos de variables aleatorias:

- x = Número de defectos en una pieza de mueble seleccionada al azar
- x = Calificación de examen de aptitud escolar (SAT) para un aspirante universitario seleccionado al azar
- x = Número de llamadas telefónicas recibidas por una línea directa de intervención en crisis durante un periodo seleccionado al azar

Las variables aleatorias cuantitativas se clasifican ya sea como *discretas* o como *continuas*, de acuerdo con los valores que x tome. Es importante distinguir entre variables aleatorias discretas y continuas, porque se usan técnicas diferentes para describir sus distribuciones. Nos concentramos en variables aleatorias discretas en el resto de este capítulo; las variables aleatorias continuas son el tema del capítulo 6.

Distribuciones de probabilidad

En los capítulos 1 y 2 usted aprendió a construir la *distribución de frecuencia relativa* para un conjunto de mediciones numéricas en una variable x . La distribución dio esta información acerca de x :

- ¿Qué valores de x se presentaron?
- ¿Con qué frecuencia se presentó cada valor de x ?

Usted también aprendió a usar la media y desviación estándar para medir el centro y variabilidad de este conjunto de datos.

En este capítulo, definimos la *probabilidad* como el valor limitando de la frecuencia relativa cuando el experimento se repite una y otra vez. Ahora definimos la **distribución de probabilidad** para una variable aleatoria x como la *distribución de frecuencia relativa* construida para toda la población de mediciones.

Definición La **distribución de probabilidad** para una variable aleatoria discreta es una fórmula, tabla o gráfica que da los posibles valores de x , y la probabilidad $p(x)$ asociada con cada valor de x .

Los valores de x representan eventos numéricos mutuamente excluyentes. Sumar $p(x)$ sobre todos los valores de x es equivalente a sumar las probabilidades de todos los eventos simples y por tanto es igual a 1.

REQUISITOS PARA UNA DISTRIBUCIÓN DE PROBABILIDAD DISCRETA

- $0 \leq p(x) \leq 1$
- $\sum p(x) = 1$

EJEMPLO 3.1



Lance al aire dos monedas imparciales y sea x igual al número observado de caras. Encuentre la distribución de probabilidad para x .

Solución Los eventos simples para este experimento con sus respectivas probabilidades se muestran en la tabla 3.1. Como $E_1 = HH$ resulta en dos caras, este evento simple resulta en el valor $x = 2$. Del mismo modo, el valor $x = 1$ se asigna a E_2 , y así sucesivamente.

TABLA 3.1 Eventos simples y probabilidades al lanzar al aire dos monedas

Evento simple	Moneda		$P(E_i)$	x
	1	2		
E_1	H	H	1/4	2
E_2	H	T	1/4	1
E_3	T	H	1/4	1
E_4	T	T	1/4	0

Para cada valor de x , se calcula $p(x)$ al sumar las probabilidades de los eventos simples en ese evento. Por ejemplo, cuando $x = 0$, ocurre el evento simple E_4 , de modo que

$$p(0) = P(E_4) = \frac{1}{4}$$

y cuando $x = 1$,

$$p(1) = P(E_2) + P(E_3) = \frac{1}{2}$$

Los valores de x y sus probabilidades respectivas, $p(x)$, aparecen en la tabla 3.2. Observe que las probabilidades totalizan 1.

TABLA 3.2

Distribución de probabilidad para x (x = número de caras)

Eventos simples		
x	en x	$p(x)$
0	E_4	$1/4$
1	E_2, E_3	$1/2$
2	E_1	$1/4$
$\Sigma p(x) = 1$		



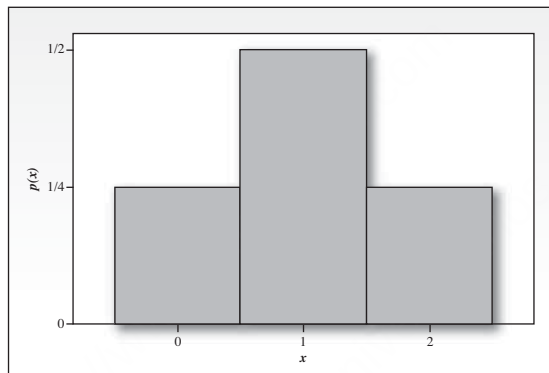
APPLET EN LÍNEA

Lanzar monedas

La distribución de probabilidad de la tabla 3.2 se grafica para formar el **histograma de probabilidad** en la figura 3.1.[†] Los tres valores de la variable aleatoria x se encuentran en el eje horizontal, y las probabilidades $p(x)$ están en el eje vertical (sustituyendo a las frecuencias relativas. Como el ancho de cada barra es 1, el área bajo la barra es la probabilidad de observar el valor particular de x y el área total es igual a 1.

FIGURA 3.1

Histograma de probabilidad para el ejemplo 3.1



La media y desviación estándar para una variable aleatoria discreta

La distribución de probabilidad para una variable aleatoria discreta luce muy semejante a la distribución de frecuencia relativa. La diferencia es que la distribución de frecuencia relativa describe una *muestra* de n mediciones, en tanto que la distribución de probabilidad se construye como un modelo para *toda la población* de mediciones. Así como la media $\bar{x}$ y la desviación estándar s midieron el centro y dispersión de los datos muestrales, usted calculará medidas similares para describir el centro y dispersión de la población.

La media poblacional, que mide el valor promedio de x en la población, también se denomina **valor esperado** de la variable aleatoria x , y se escribe como $E(x)$. Es el valor que *se esperaría* observar en *promedio* si el experimento se repite una y otra vez. La fórmula para calcular la media poblacional es más fácil de entender por ejemplo. Lance otra vez al aire esas dos monedas imparciales, y sea x el número de caras observado. Construimos esta distribución de probabilidad para x :

x	0	1	2
$p(x)$	$1/4$	$1/2$	$1/4$

[†]La distribución de probabilidad de la tabla 3.8 también se puede presentar usando una fórmula, que se da en la sección 3.2.

Suponga que el experimento se repite un gran número de veces, por ejemplo $n = 4\,000\,000$ de veces. Intuitivamente, se esperaría observar alrededor de un millón de ceros, dos millones de números 1 y un millón de números dos. Entonces el valor promedio de x sería igual a

$$\begin{aligned}\frac{\text{Suma de las medidas}}{n} &= \frac{1\,000\,000(0) + 2\,000\,000(1) + 1\,000\,000(2)}{4\,000\,000} \\ &= \left(\frac{1}{4}\right)(0) + \left(\frac{1}{2}\right)(1) + \left(\frac{1}{4}\right)(2)\end{aligned}$$

Observe que el primer término de esta suma es $(0)p(0)$, el segundo es igual a $(1)p(1)$ y el tercero es $(2)p(2)$. El valor promedio de x , entonces, es

$$\sum xp(x) = 0 + \frac{1}{2} + \frac{2}{4} = 1$$

Este resultado da alguna justificación intuitiva para la definición del valor esperado de una variable aleatoria x discreta.

Definición Sea x una variable aleatoria discreta con distribución de probabilidad $p(x)$. La media o **valor esperado de x** está dada como

$$\mu = E(x) = \sum xp(x)$$

donde los elementos se suman sobre todos los valores de la variable aleatoria x .

Podríamos usar un argumento similar para justificar las fórmulas para la **varianza poblacional** σ^2 y la **desviación estándar de la población** σ . Estas medidas numéricas describen la dispersión o variabilidad de la variable aleatoria usando el “promedio” o “valor esperado” de $(x - \mu)^2$, el cuadrado de las desviaciones de los valores x desde su media μ .

Definición Sea x una variable aleatoria discreta con distribución de probabilidad $p(x)$ y media. La **varianza de x** es

$$\sigma^2 = E[(x - \mu)^2] = \sum (x - \mu)^2 p(x)$$

donde la sumatoria es sobre todos los valores de la variable aleatoria x .[†]

Definición La **desviación estándar σ de una variable aleatoria x** es igual a la raíz cuadrada positiva de su varianza.

EJEMPLO

3.2

Una tienda de electrónica vende un modelo particular de computadora portátil. Hay sólo cuatro computadoras en existencia y la gerente se pregunta cuál será la demanda de hoy para este modelo particular. Ella se enteró en el departamento de marketing de que la distribución de probabilidad para x , la demanda diaria para la laptop, es como se muestra en la tabla. Encuentre la media, varianza y desviación estándar de x . ¿Es probable que cinco o más clientes deseen comprar una laptop hoy?

x	0	1	2	3	4	5
$p(x)$	.10	.40	.20	.15	.10	.05

[†]Se puede demostrar (prueba omitida) que $\sigma^2 = \sum (x - \mu)^2 p(x) = \sum x^2 p(x) - \mu^2$. Este resultado es análogo a la fórmula de computación para la suma de cuadrados de las desviaciones dadas en el capítulo 1.

Solución La tabla 3.3 muestra los valores de x y $p(x)$, junto con los términos individuales empleados en las fórmulas para μ y σ^2 . La suma de los valores en la tercera columna es

$$\mu = \sum xp(x) = (0)(.10) + (1)(.40) + \dots + (5)(.05) = 1.90$$

en tanto que la suma de los valores en la quinta columna es

$$\begin{aligned}\sigma^2 &= \sum (x - \mu)^2 p(x) \\ &= (0 - 1.9)^2(.10) + (1 - 1.9)^2(.40) + \dots + (5 - 1.9)^2(.05) = 1.79\end{aligned}$$

y

$$\sigma = \sqrt{\sigma^2} = \sqrt{1.79} = 1.34$$

TABLA 3.3

Cálculos para el ejemplo 3.2

x	$p(x)$	$xp(x)$	$(x - \mu)^2$	$(x - \mu)^2 p(x)$
0	.10	.00	3.61	.361
1	.40	.40	.81	.324
2	.20	.40	.01	.002
3	.15	.45	1.21	.1815
4	.10	.40	4.41	.441
5	.05	.25	9.61	.4805
Totales	1.00	$\mu = 1.90$		$\sigma^2 = 1.79$

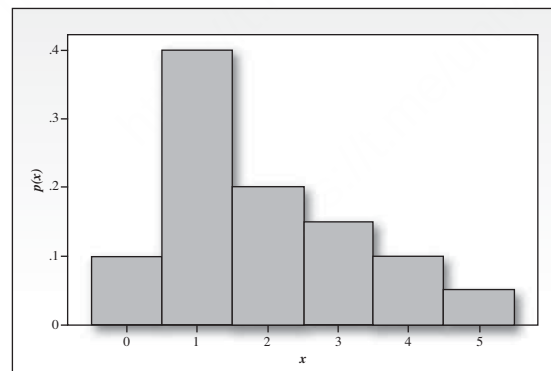
La gráfica de la distribución de probabilidad se muestra en la figura 3.2. Como la distribución tiene más o menos la forma de montículo, aproximadamente 95% de todas las mediciones deben estar a no más de *dos* desviaciones estándar de la media, es decir,

$$\mu \pm 2\sigma \Rightarrow 1.90 \pm 2(1.34) \quad \text{o} \quad -.78 \text{ a } 4.58$$

Como $x = 5$ está fuera de este intervalo, se dice que es improbable que cinco o más clientes deseen comprar una laptop hoy. De hecho, $P(x \geq 5)$ es exactamente .05, o sea 1 vez en 20.

FIGURA 3.2

Distribución de probabilidad para el ejemplo 3.2



EJEMPLO

3.3

En una lotería que se realiza a beneficio de una institución local de caridad, se deben vender 8000 boletos a \$10 cada uno. El premio es un automóvil de \$24 000. Si usted compra dos boletos, ¿cuál es su ganancia esperada?

Solución Su ganancia x toma uno de dos valores. O bien perderá \$20 (es decir, su “ganancia” será -\$20) o ganará \$23 980, con probabilidades de 7 998/8 000 y 2/8 000, respectivamente. La distribución de probabilidad para la ganancia x se muestra en la tabla:

x	$p(x)$
-\$20	7 998/8 000
\$23 980	2/8 000

La ganancia esperada será

$$\begin{aligned}\mu &= \sum xp(x) \\ &= (-\$20)\left(\frac{7\,998}{8\,000}\right) + (\$23\,980)\left(\frac{2}{8\,000}\right) = -\$14\end{aligned}$$

Recuerde que el valor esperado de x es el promedio de la población teórica que resultaría si en la lotería se repitiera un número infinitamente grande de veces. Si se hiciera esto, su ganancia promedio o esperada por boleto de lotería sería una pérdida de \$14.

EJEMPLO

3.4

Determine la prima anual para una póliza de seguro de \$10 000 que cubre un evento que, en un largo tiempo, ha ocurrido a razón de 2 veces en 100. Sea x igual a la ganancia financiera anual para la compañía de seguros, que resulte de la venta de la póliza, y sea C igual a la prima anual desconocida. Calcule el valor de C tal que la ganancia esperada $E(x)$ iguale a cero. Entonces C es la prima requerida para que haya punto de equilibrio. Para esto, la compañía sumaría los costos administrativos y la utilidad.

Solución El primer paso en la solución es determinar los valores que la ganancia x toma y luego determinar $p(x)$. Si el evento no ocurre durante el año, la compañía de seguros ganará la prima de $x = C$ dólares. Si el evento ocurre, la ganancia será negativa; esto es, la compañía perderá \$10 000 menos la prima de C dólares ya recolectada. Entonces $x = -(10\,000 - C)$ dólares. Las probabilidades asociadas con estos dos valores de x son $98/100$ y $2/100$, respectivamente. La distribución de probabilidad para la ganancia se muestra en la tabla:

$x = \text{Ganancia}$	$p(x)$
C	$98/100$
$-(10\,000 - C)$	$2/100$

Como la compañía desea una prima de seguro C tal que, a largo plazo (para muchas pólizas similares), la ganancia media sea igual a cero, se puede establecer el valor esperado de x igual a cero y despejar C . Entonces

$$\begin{aligned}\mu &= E(x) = \sum xp(x) \\ &= C\left(\frac{98}{100}\right) + [-10\,000 + C]\left(\frac{2}{100}\right) = 0 \\ \text{o} \\ \frac{98}{100}C + \frac{2}{100}C - 200 &= 0\end{aligned}$$

Despejando C de esta ecuación, se obtiene $C = \$200$. Por tanto, si la compañía de seguros cobró una prima anual de \$200, el promedio de ganancia calculada para un gran número de pólizas similares sería igual a cero. La prima real sería igual a \$200 más los costos administrativos y la utilidad.

El método para calcular el valor esperado de x para una variable aleatoria continua es similar a lo que acabamos de hacer, pero en la práctica requiere el uso de cálculo. No obstante, los resultados básicos respecto a expectativas son los mismos para variables aleatorias continuas y discretas. Por ejemplo, sin considerar si x es continua o discreta, $\mu = E(x)$ y $\sigma^2 = E[(x - \mu)^2]$.

3.2 EJERCICIOS

TÉCNICAS BÁSICAS

3.1 ¿Discretas o continuas? Identifique las siguientes variables aleatorias como discretas o continuas:

- El número total de puntos anotados en un juego de fútbol.
- La duración en estante de un medicamento particular.
- La altura de la marea del océano en un lugar determinado.
- Longitud de una perca americana de 2 años de edad.
- El número de choques que casi ocurren de aviones en el aire en un año.

3.2 ¿Discretas o continuas? II Identifique las siguientes variables aleatorias como discretas o continuas:

- Aumento en tiempo de vida alcanzado por un paciente de cáncer como resultado de una cirugía.
- Resistencia a la ruptura (en libras por pulgada cuadrada) de un cable de acero de una pulgada de diámetro.
- Número de venados muertos por año en una reservación estatal de fauna silvestre.
- Número de cuentas vencidas en una tienda de departamentos en un tiempo particular.
- Su presión sanguínea.

3.3 Distribución de probabilidad I Una variable aleatoria x tiene esta distribución de probabilidad:

x	0	1	2	3	4	5
$p(x)$	.1	.3	.4	.1	?	.05

- Encuentre $p(4)$.
- Construya un histograma de probabilidad para describir $p(x)$.
- Encuentre μ , σ^2 , y σ .
- Localice el intervalo $\mu \pm 2\sigma$ en el eje x del histograma. ¿Cuál es la probabilidad de que x caiga en este intervalo?
- Si seleccionáramos un número muy grande de valores de x de la población, ¿la mayoría caería en el intervalo $\mu \pm 2\sigma$? Explique.

3.4 Distribución de probabilidad II Una variable aleatoria x puede tomar cinco valores: 0, 1, 2, 3, 4. Una parte de la distribución de probabilidad se muestra aquí:

x	0	1	2	3	4
$p(x)$	.1	.3	.3	?	.1

- Encuentre $p(3)$.
- Construya un histograma de probabilidad para $p(x)$.
- Calcule la media poblacional, varianza y desviación estándar.

d. ¿Cuál es la probabilidad de que x sea mayor que 2?

e. ¿Cuál es la probabilidad de que x sea 3 o menor?

3.5 Visitas de tienda Represente con x el número de veces que un cliente acude a una tienda en un periodo de una semana. Suponga que ésta es la distribución de probabilidad de x :

x	0	1	2	3
$p(x)$	.1	.4	.4	.1

Encuentre el valor esperado de x , el número promedio de veces que un cliente acude a la tienda.

3.6 Si lanza un par de dados, la suma T de los números que aparecen en las caras superiores de los dados puede asumir el valor de un entero en el intervalo $2 \leq T \leq 12$.

- Encuentre la distribución de probabilidad para T .
Presente esta distribución de probabilidad en una tabla.
- Construya un histograma de probabilidad para $P(T)$.
¿Cómo describiría la forma de esta distribución?

APLICACIONES

3.7 ¿Mensajes de texto mientras se conduce? La proporción de adultos (18 años o más) que admiten enviar mensajes de texto mientras conducen es 47%.¹ Suponga que selecciona al azar tres conductores adultos y les pregunta si envían mensajes de texto mientras conducen.

- Encuentre la distribución de probabilidad para x , el número de conductores en la muestra que admiten enviar mensajes de texto mientras conducen.
- Construya un histograma de probabilidad para $p(x)$.
- ¿Cuál es la probabilidad de que exactamente uno de los tres conductores envíe mensajes de texto mientras conduce?
- ¿Cuáles son la media y la desviación estándar poblacionales para la variable aleatoria x ?

3.8 ¿Sesgo de género? Una compañía tiene cinco solicitantes para dos puestos de trabajo: dos mujeres y tres hombres. Suponga que los cinco solicitantes son igualmente calificados y que no hay preferencia para elegir cualquier género. Sea x igual al número de mujeres seleccionadas para ocupar los dos puestos de trabajo.

- Encuentre $p(x)$.
- Construya un histograma de probabilidad para x .

3.9 Perforación de pozos petroleros La experiencia ha demostrado que, en promedio, sólo uno de cada 10 pozos produce petróleo. Sea x el número de perforaciones hasta el primer éxito (se encuentra petróleo). Suponga que las perforaciones representan eventos independientes.

- Encuentre $p(1)$, $p(2)$ y $p(3)$.
- Dé una fórmula para $p(x)$.
- Grafique $p(x)$.

3.10 ¿Alguien juega tenis? Dos jugadores profesionales de tenis, A y B , están programados para jugar un partido: el ganador del partido es el primero en ganar tres sets de un total de cinco. El evento en que A gane algún set es independiente del evento de que gane cualquier otro y la probabilidad de que gane cualquier set es igual a .6. Sea x igual al número total de sets del partido; esto es, $x = 3, 4$ o 5 . Encuentre $p(x)$.

3.11 Tenis, otra vez En el ejercicio 3.10 encontró la distribución de probabilidad para x , el número de sets requeridos para jugar un partido como el mejor de cinco sets, dado que la probabilidad de que A gane cualquier set —llamemos a esto $P(A)$ — es .6.

- Encuentre el número esperado de sets necesario para completar el partido para $P(A) = .6$.
- Encuentre el número esperado de sets necesario para completar el partido cuando los jugadores sean de igual capacidad, es decir, $P(A) = .5$.
- Encuentre el número esperado de sets necesario para completar el partido cuando los jugadores difieran en mucho en capacidad, es decir, por ejemplo, $P(A) = .9$.
- ¿Cuál es la relación entre $P(A)$ y $E(x)$, el número esperado de sets requeridos para completar el partido?

3.12 Prueba de la FDA La duración máxima de patente para un nuevo medicamento es 17 años. Restando el tiempo requerido por la FDA para probar y aprobar el medicamento proporciona la vida real de patente del medicamento, es decir, el tiempo que una compañía tiene para recuperar costos de investigación y desarrollo y obtener una utilidad. Suponga que la distribución de tiempos de vida de patente para nuevos medicamentos es como se muestra a continuación:

Años, x	3	4	5	6	7	8
$p(x)$	.03	.05	.07	.10	.14	.20

Años, x	9	10	11	12	13
$p(x)$	.18	.12	.07	.03	.01

- Encuentre el número esperado de años de vigencia de patente para un nuevo medicamento.
- Encuentre la desviación estándar de x .
- Encuentre la probabilidad de que x caiga en el intervalo $\mu \pm 2\sigma$.

3.13 Descanso para tomar café La mayoría de las personas que bebe café se da un poco de tiempo para hacerlo y muchas toman más de un descanso al día. La tabla siguiente, adaptada de un Snapshot de *USA Today*, muestra la distribución de probabilidad para x , el número de descansos por día que se dan quienes beben café.²

x	0	1	2	3	4	5
$p(x)$	.28	.37	.17	.12	.05	.01

- ¿Cuál es la probabilidad de que una persona que toma café, seleccionada al azar, no se dé un descanso para tomar café durante el día?
- ¿Cuál es la probabilidad de que una persona que toma café, seleccionada al azar, se dé más de dos descansos para tomar café durante el día?
- Calcule la media y desviación estándar para la variable aleatoria x .
- Encuentre la probabilidad de que x caiga en el intervalo $\mu \pm 2\sigma$.

3.14 Actuarios El director de una empresa está considerando tomar una póliza de seguro para cubrir posibles pérdidas en que incurriría al vender un nuevo producto. Si el producto es un completo fracaso, el director incurrirá en una pérdida de \$800 000; si es sólo un éxito moderado, incurrirá en una pérdida de \$250 000. Los actuarios de seguros han determinado que las probabilidades de que ese producto sea un fracaso o sólo tenga un éxito moderado son .01 y .05, respectivamente. Suponiendo que el director de la empresa esté dispuesto a ignorar todas las otras posibles pérdidas, ¿qué prima debería cobrar la compañía de seguros por una póliza para no tener pérdida ni ganancia?

3.3

LA DISTRIBUCIÓN BINOMIAL DE PROBABILIDAD

El experimento de lanzar al aire una moneda es un ejemplo sencillo de una importante variable aleatoria discreta llamada **variable aleatoria binomial**. Muchos experimentos prácticos resultan en datos similares a que salgan cara o cruz al tirar la moneda. Por ejemplo, considere las encuestas políticas que se emplean para predecir las preferencias de los votantes en elecciones. Cada votante entrevistado se puede comparar a una moneda porque es probable que el votante esté a favor de nuestro candidato (una “cara”) o no (una “cruz”). Casi siempre, la proporción de votantes que están a favor de nuestro candidato no es igual a $1/2$, es decir, la moneda no es imparcial. De hecho, la encuesta está diseñada exactamente para determinar la proporción de votantes que están a favor de nuestro candidato.

Veamos aquí algunas otras situaciones semejantes al experimento de lanzar al aire una moneda:

- Un sociólogo está interesado en la proporción de maestros de escuela primaria que sean varones.
- Una comerciante en bebidas gaseosas está interesada en la proporción de consumidores de refrescos de cola que prefieren su marca.
- Un genetista está interesado en la proporción de la población que posee un gen vinculado a la enfermedad de Alzheimer.

Cada persona muestreada es análoga a lanzar al aire una moneda, pero la probabilidad de una “cara” no es necesariamente igual a $1/2$. Aun cuando estas situaciones tienen diferentes objetivos prácticos, todas exhiben las características comunes del **experimento binomial**.

Definición Un **experimento binomial** tiene estas cinco características:

1. El experimento consiste en n intentos idénticos.
2. Cada intento produce uno de dos resultados. Por falta de un mejor nombre, el resultado uno se llama éxito, S, y el otro, fracaso, F.
3. La probabilidad de éxito en un solo intento es igual a p y es el mismo de un intento tras otro. La probabilidad de fracaso es igual a $(1 - p) = q$.
4. Los intentos son independientes.
5. Estamos interesados en x , el número de éxitos observado durante los n intentos, para $x = 0, 1, 2, \dots, n$.

EJEMPLO

3.5

Suponga que hay alrededor de un millón de adultos en un condado y una proporción desconocida p están a favor de limitar el periodo de función de políticos. Se elegirá una muestra de mil adultos en forma tal que cada uno, del millón de adultos, tenga igual probabilidad de ser seleccionado y a cada uno se le preguntará si él o ella está a favor de limitar el periodo. (El objetivo final de esta encuesta es estimar la proporción desconocida p , un problema que veremos en el capítulo 8.) ¿Este experimento es binomial?

Solución ¿El experimento tiene las cinco características binomiales?

1. Un “intento” es la selección de un solo adulto de entre el millón de electores de la ciudad. Esta muestra consta de $n = 1000$ intentos idénticos.
2. Como cada elector estará o no a favor de limitar el periodo, hay dos resultados que representan los “éxitos” y “fracasos” del experimento binomial.[†]

[†]Aun cuando es tradicional que los dos posibles resultados de un intento se denominen “éxito” y “fracaso”, el resultado llamado “éxito” no necesita ser visto como éxito en el uso ordinario de la palabra.

3. La probabilidad de éxito, p , es la probabilidad de que un adulto esté a favor del límite del periodo. ¿Esta probabilidad sigue igual para cada uno de los electores de la muestra? Para todos los fines prácticos, la respuesta es sí. Por ejemplo, si 500 000 electores de la población están a favor de limitar el periodo, entonces la probabilidad de un “éxito” cuando se elija al primer elector es $500\,000/1\,000\,000 = 1/2$. Cuando se elija al segundo elector, la probabilidad p cambia ligeramente, dependiendo de la primera selección. Esto es, habrá 499 999 o 500 000 éxitos que queden entre los 999 999 adultos. En cualquiera de estos casos, p es todavía más o menos igual a $1/2$.
4. La independencia de los intentos está garantizada debido al grupo grande de adultos del que se toma la muestra. La probabilidad de que un elector esté a favor de limitar el periodo no cambia, dependiendo de las respuestas de las personas previamente elegidas.
5. La variable aleatoria x es el número de electores de la muestra que estén a favor de limitar el periodo.

Debido a que el estudio satisface las cinco características razonablemente bien, para todos los fines prácticos se le considera un experimento binomial.

EJEMPLO

3.6

Un paciente llena una receta para un régimen de 10 días de dos píldoras diarias. Sin que lo sepa el farmacéutico ni el paciente, las 20 pastillas están formadas por 18 píldoras del medicamento prescrito y dos píldoras que son el equivalente genérico del medicamento prescrito. El paciente selecciona dos píldoras al azar para la dosis del primer día. Si verificamos la selección y registramos el número de píldoras que son genéricas, ¿es éste un experimento binomial?

Solución Verifique de nuevo el procedimiento de muestra para las características de un experimento binomial.

1. Un “intento” es la selección de una píldora de entre las 20 de la receta. Este experimento consta de $n = 2$ intentos.
2. Cada intento resulta en uno de dos resultados. O bien, la píldora es genérica (llame “éxito” a esto) o no lo es (un “fracaso”).
3. Como las píldoras de una botella de receta se consideran “mezcladas” al azar, la probabilidad incondicional de sacar una píldora genérica en un intento determinado sería $2/20$.
4. La condición de independencia entre intentos *no* está satisfecha, porque la probabilidad de sacar una píldora genérica en el segundo intento depende del primer intento. Por ejemplo, si la primera píldora sacada es genérica entonces hay sólo una píldora genérica en las restantes 19. Por tanto,

$$P(\text{genérica en intento 2} | \text{genérica en intento 1}) = 1/19$$

Si la primera selección *no* resulta en una píldora genérica, entonces hay todavía dos píldoras genéricas en las restantes 19, y la probabilidad de un “éxito” (una píldora genérica) cambia a

$$P(\text{genérica en el intento 2} | \text{no genérica en el intento 1}) = 2/19$$

Por tanto, los intentos son dependientes y el muestreo no representa un experimento binomial.

Considere la diferencia entre estos dos ejemplos. Cuando la muestra (los n intentos idénticos) vinieron de una población grande, la probabilidad de éxito p siguió siendo más o menos la misma de un intento a otro. Cuando el tamaño poblacional N era pequeño, la probabilidad de éxito p cambió en forma considerable de un intento a otro, y el experimento *no fue* binomial.

REGLA PRÁCTICA

Si el tamaño muestral es grande respecto al tamaño poblacional, en particular si $n/N \geq .05$, entonces el experimento resultante no es binomial.

Si lanzamos al aire dos monedas “honestas” y construimos la distribución de probabilidad para x , el número de caras, un experimento binomial con $n = 2$ y $p = .5$. La distribución binomial general de probabilidad se construye en la misma forma, pero el procedimiento se complica cuando n se hace grande. Afortunadamente, las probabilidades $p(x)$ siguen un modelo general. Esto nos permite usar una sola fórmula para hallar $p(x)$ para cualquier valor dado de x .

LA DISTRIBUCIÓN BINOMIAL DE PROBABILIDAD

Un experimento binomial consta de n intentos idénticos con probabilidad p de éxito en cada intento. La probabilidad de k éxitos en n intentos es

$$P(x = k) = C_k^n p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

para valores de $k = 0, 1, 2, \dots, n$. El símbolo C_k^n es igual a,

$$\frac{n!}{k!(n-k)!}$$

donde $n! = n(n-1)(n-2) \cdots (2)(1)$ y $0! \equiv 1$.

Las fórmulas generales para μ , σ^2 y σ se usan para obtener las siguientes fórmulas más sencillas para la media y la desviación estándar binomiales.

MEDIA Y DESVIACIÓN ESTÁNDAR PARA LA VARIABLE ALEATORIA BINOMIAL

La variable aleatoria x , el número de éxitos en n intentos, tiene una distribución de probabilidad con este centro y dispersión:

Media:	$\mu = np$
Varianza:	$\sigma^2 = npq$
Desviación estándar:	$\sigma = \sqrt{npq}$

EJEMPLO

3.7

Encuentre $P(x = 2)$ para una variable aleatoria binomial con $n = 10$ y $p = .1$.

Solución $P(x = 2)$ es la probabilidad de observar 2 éxitos y 8 fracasos en una secuencia de 10 intentos. Se podrían observar 2 éxitos primero, seguidos de 8 fracasos consecutivos:

E, E, F, F, F, F, F, F, F, F

Como p es la probabilidad de éxito y q es la probabilidad de fracaso, esta secuencia particular tiene probabilidad

$$ppqqqqqqqq = p^2 q^8$$

Sin embargo, puede también resultar muchas otras secuencias en $x = 2$ éxitos. La fórmula binomial utiliza C_2^{10} para contar el número de secuencias y da la probabilidad exacta cuando se usa la fórmula binomial con $k = 2$:

TIP MI CONSEJO

$n! =$
 $n(n-1)(n-2) \cdots (2)(1)$
 Por ejemplo,
 $5! = 5(4)(3)(2)(1) = 120$
 y $0! \equiv 1$

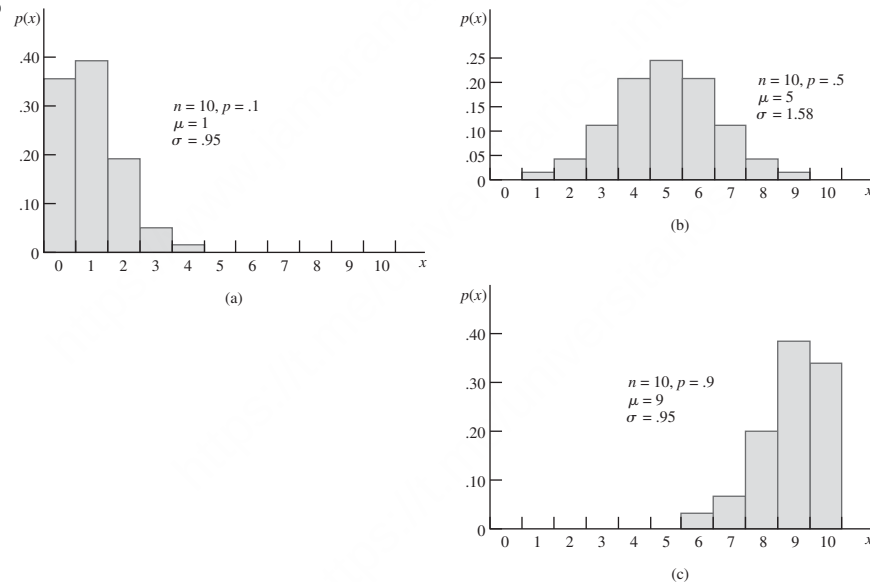
$$\begin{aligned}
 P(x = 2) &= C_2^{10}(.1)^2(.9)^{10-2} \\
 &= \frac{10!}{2!(10-2)!}(.1)^2(.9)^8 = \frac{10(9)}{2(1)}(.01)(.430467) = .1937
 \end{aligned}$$

Se podría repetir el procedimiento del ejemplo 3.7 para cada valor de x (0, 1, 2, ..., 10) y encontrar todos los valores de $p(x)$ necesarios para construir un histograma de probabilidad para x . Éste sería un trabajo largo y tedioso, pero la gráfica resultante se vería como la figura 3.3a). Se puede verificar la altura de la barra para $x = 2$ y encontrar $p(2) = P(x = 2) = .1937$. La gráfica está sesgada a la derecha; esto es, casi todo el tiempo se observarán valores pequeños de x . La media o “punto de equilibrio” está alrededor de $x = 1$; de hecho, se utiliza la fórmula para hallar la media exacta:

$$\mu = np = 10(.1) = 1$$

Las figuras 3.3b) y 3.3c) muestran las otras dos distribuciones binomiales con $n = 10$ pero con diferentes valores de p . Vea las formas de estas distribuciones. Cuando $p = .5$, la distribución es exactamente simétrica alrededor de la media, $\mu = np = 10(.5) = 5$. Cuando $p = .9$, la distribución es la “imagen espejo” de la distribución para $p = .1$ y está sesgada a la izquierda.

FIGURA 3.3
Distribuciones de probabilidad binomial



EJEMPLO

3.8

Durante un largo periodo se ha observado que un jugador profesional de baloncesto puede hacer un tiro libre en un intento determinado con una probabilidad igual a .8. Suponga que él lanza cuatro tiros libres.

1. ¿Cuál es la probabilidad de que enceste exactamente dos tiros libres?
2. ¿Cuál es la probabilidad de que enceste al menos un tiro libre?

Solución Un “intento” es un solo tiro libre y se define un “éxito” como una canasta y un “fracaso” como una falla, de modo que $n = 4$ y $p = .8$. Si se supone que la probabilidad del jugador de encestar el tiro libre no cambia de un tiro a otro, entonces el número x de veces que enceste el tiro libre es una *variable aleatoria binomial*.

$$\begin{aligned}
 1. \quad P(x = 2) &= C_2^4(.8)^2(.2)^2 \\
 &= \frac{4!}{2!2!}(.64)(.04) = \frac{4(3)(2)(1)}{2(1)(2)(1)}(.64)(.04) = .1536
 \end{aligned}$$

La probabilidad es .1536 de que encesté exactamente dos tiros libres.

$$\begin{aligned}
 2. \quad P(\text{al menos uno}) &= P(x \geq 1) = p(1) + p(2) + p(3) + p(4) \\
 &= 1 - p(0) \\
 &= 1 - C_0^4(.8)^0(.2)^4 \\
 &= 1 - .0016 = .9984
 \end{aligned}$$

Aun cuando usted podría calcular $P(x = 1)$, $P(x = 2)$, $P(x = 3)$ y $P(x = 4)$ para hallar esta probabilidad, usar el complemento del evento hace más fácil su trabajo; es decir,

$$P(x \geq 1) = 1 - P(x < 1) = 1 - P(x = 0)$$

¿Considera alguna razón por la que su suposición de intentos independientes podría ser errónea? Si el jugador aprende de su intento previo (es decir, ajusta su tiro de acuerdo con su último intento), entonces su probabilidad p de encestar el tiro libre puede cambiar, posiblemente aumentar, de un tiro a otro. Los intentos *no* serían independientes y el experimento *no* sería binomial.

TIP MI CONSEJO

Use la tabla 1 del Anexo en lugar de la fórmula binomial siempre que sea posible. ¡Esta es una forma más fácil!

Calcular probabilidades binomiales puede ser tedioso incluso para valores relativamente pequeños de n . Cuando n se hace grande, es casi imposible sin ayuda de una calculadora o computadora. Por fortuna, tenemos estas dos herramientas. Las tablas de **probabilidades binomiales acumulativas** generadas por computadora se presentan en la tabla 1 del Anexo, para valores de n que van de 2 a 25 y para valores seleccionados de p . Estas probabilidades también pueden ser generadas si se usa el *MINITAB*, *MS Excel* o los applets Java en el sitio web CourseMate.

Las probabilidades binomiales *acumulativas* difieren de las probabilidades binomiales *individuales* que se calcularon con la fórmula binomial. Una vez que usted encuentre la columna de probabilidades para los valores correctos de n y p en la tabla 1, el renglón marcado como k proporciona la suma de todas las probabilidades binomiales de $x = 0$ a $x = k$. La tabla 3.5 muestra parte de la tabla 1 para $n = 5$ y $p = .6$. Si se observa el renglón marcado $k = 3$, se encuentra

$$P(x \leq 3) = p(0) + p(1) + p(2) + p(3) = .663$$

TABLA 3.4 Parte de la tabla 1 del Anexo para $n = 5$

k	p													k
	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	
0	—	—	—	—	—	—	—	.010	—	—	—	—	—	0
1	—	—	—	—	—	—	—	.087	—	—	—	—	—	1
2	—	—	—	—	—	—	—	.317	—	—	—	—	—	2
3	—	—	—	—	—	—	—	.663	—	—	—	—	—	3
4	—	—	—	—	—	—	—	.922	—	—	—	—	—	4
5	—	—	—	—	—	—	—	1.000	—	—	—	—	—	5

Si la probabilidad que necesita calcular no está en esta forma, deberá considerar una forma para reescribir su probabilidad y hacer uso de las tablas.

EJEMPLO

3.9

Use la tabla binomial acumulativa para $n = 5$ y $p = .6$ para hallar las probabilidades de estos eventos:

1. Exactamente tres éxitos
2. Tres o más éxitos

Solución

1. Si encuentra $k = 3$ en la tabla 3.5, el valor presentado es

$$P(x \leq 3) = p(0) + p(1) + p(2) + p(3)$$

Como usted desea sólo $P(x = 3) = p(3)$, debe restar la probabilidad no deseada:

$$P(x \leq 2) = p(0) + p(1) + p(2)$$

que se encuentra en la tabla 3.5 con $k = 2$. Entonces

$$\begin{aligned} P(x = 3) &= P(x \leq 3) - P(x \leq 2) \\ &= .663 - .317 = .346 \end{aligned}$$

2. Para hallar $P(\text{tres o más éxitos}) = P(x \geq 3)$ usando la tabla 3.1, se debe emplear el complemento del evento de interés. Escriba

$$P(x \geq 3) = 1 - P(x < 3) = 1 - P(x \leq 2)$$

Puede hallar $P(x \leq 2)$ en la tabla 3.5 con $k = 2$. Entonces

$$\begin{aligned} P(x \geq 3) &= 1 - P(x \leq 2) \\ &= 1 - .317 = .683 \end{aligned}$$

EJEMPLO

3.10

Consulte el ejemplo 3.5 y la variable aleatoria binomial x con $n = 5$ y $p = .6$. Use la tabla binomial acumulativa para encontrar las probabilidades binomiales restantes, $p(0)$, $p(1)$, $p(2)$, $p(4)$ y $p(5)$. Construya el histograma de probabilidad para la variable aleatoria x y describa su forma y ubicación.

Solución

1. Puede encontrar $P(x = 0)$ directamente de la tabla 3.5 con $k = 0$. Esto es, $p(0) = .010$.
2. Las otras probabilidades pueden encontrarse restando entradas sucesivas en la tabla 3.5. Entonces

$$P(x = 1) = P(x \leq 1) - P(x = 0) = .087 - .010 = .077$$

$$P(x = 2) = P(x \leq 2) - P(x \leq 1) = .317 - .087 = .230$$

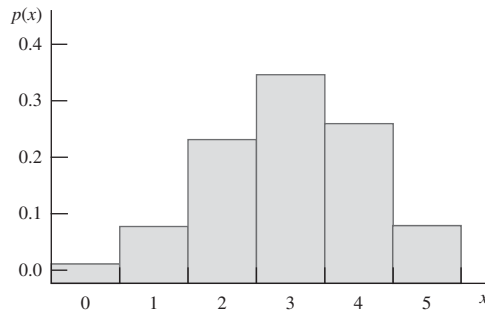
$$P(x = 4) = P(x \leq 4) - P(x \leq 3) = .922 - .663 = .259$$

$$P(x = 5) = P(x \leq 5) - P(x \leq 4) = 1.000 - .922 = .078$$

El histograma de probabilidad se muestra en la figura 3.4. La distribución tiene una forma relativamente de montículo, con un centro alrededor de 3.

FIGURA 3.4

Distribución de probabilidad binomial para el ejemplo 3.10.



NECESITO SABER...

Cómo utilizar la tabla 1 para calcular probabilidades binomiales

1. Encuentre los valores necesarios de n y p . Aísle la columna apropiada en la tabla 1.
2. La tabla 1 da $P(x \leq k)$ en la fila marcada k . Reescriba la probabilidad que necesita de modo que quede de esta forma.
 - Haga una lista de los valores de x en su evento.
 - De la lista, escriba el evento ya sea como la diferencia de dos probabilidades:

$$P(x \leq a) - P(x \leq b) \text{ para } a > b$$

o como el complemento del evento:

$$1 - P(x \leq a)$$

o justo como el evento mismo:

$$P(x \leq a) \text{ o } P(x < a) = P(x \leq a - 1)$$

EJEMPLO

3.11



APPLET EN LÍNEA

Cálculo de probabilidades binomiales

Se probó un régimen formado por una dosis diaria de vitamina C para determinar su eficacia para prevenir el resfriado común. Diez personas que estuvieron siguiendo el régimen prescrito fueron observadas durante un año. Ocho pasaron el invierno sin un resfriado. Suponga que la probabilidad de pasar el invierno sin un resfriado es .5 cuando no se sigue el régimen de vitamina C. ¿Cuál es la probabilidad de observar ocho o más sobrevivientes, dado que el régimen es ineficiente para aumentar la resistencia a resfriados?

Solución Si se supone que el régimen de vitamina C es ineficiente, entonces la probabilidad p de sobrevivir el invierno sin un resfriado es .5. La distribución de probabilidad para x , el número de sobrevivientes, es

$$p(x) = C_{10}^x (.5)^x (.5)^{10-x}$$

Usted ya ha aprendido varias formas de hallar $P(8 \text{ o más sobrevivientes}) = P(x \geq 8)$. Obtendrá los mismos resultados con cualquiera de esos métodos; elija el más conveniente para su problema particular.

1. La fórmula binomial:

$$\begin{aligned} P(8 \text{ o más}) &= p(8) + p(9) + p(10) \\ &= C_8^{10} (.5)^{10} + C_9^{10} (.5)^{10} + C_{10}^{10} (.5)^{10} \\ &= .055 \end{aligned}$$

2. *Las tablas binomiales acumulativas:* Encuentre la columna correspondiente a $p = .5$ en la tabla para $n = 10$:

$$\begin{aligned} P(8 \text{ o más}) &= P(x \geq 8) = 1 - P(x \leq 7) \\ &= 1 - .945 = .055 \end{aligned}$$

3. *Salida del MINITAB o MS Excel:* Las salidas que se muestran en las figuras 3.5a) y 3.5b) dan la **función acumulativa de distribución**, que son las mismas probabilidades que encontré en las tablas acumulativas binomiales. **La función de densidad de probabilidad** da las probabilidades binomiales individuales, que encontré usted usando la fórmula binomial.

FIGURA 3.5(a)

Salida de MINITAB para el ejemplo 3.11

● **Función de distribución acumulativa**

Binomial with n = 10 and p = 0.5

x	P(X ≤ x)
0	0.00098
1	0.01074
2	0.05469
3	0.17187
4	0.37695
5	0.62305
6	0.82813
7	0.94531
8	0.98926
9	0.99902
10	1.00000

Función de densidad de probabilidad

Binomial with n = 10 and p = 0.5

x	P(X = x)
0	0.000977
1	0.009766
2	0.043945
3	0.117188
4	0.205078
5	0.246094
6	0.205078
7	0.117188
8	0.043945
9	0.009766
10	0.000977

FIGURA 3.5(b)

Salida de Excel para el ejemplo 3.11

x	P(X ≤ x)	P(X = x)
0	0.000977	0.000977
1	0.010742	0.009766
2	0.054688	0.043945
3	0.171875	0.117188
4	0.376953	0.205078
5	0.623047	0.246094
6	0.828125	0.205078
7	0.945313	0.117188
8	0.989258	0.043945
9	0.999023	0.009766
10	1	0.000977

Usando la función acumulativa de distribución, calcule

$$\begin{aligned} P(x \geq 8) &= 1 - P(x \leq 7) \\ &= 1 - .94531 = .05469 \end{aligned}$$

O bien, usando la función de densidad de probabilidad, calcule

$$\begin{aligned} P(x \geq 8) &= p(8) + p(9) + p(10) \\ &= .043945 + .009766 + .000977 = .05469 \end{aligned}$$

EJEMPLO

3.12

¿Preferiría usted presentar un examen de opción múltiple o uno de memoria? Si no sabe nada del material, tendrá una calificación de cero en un examen de memoria, pero, si le dan cinco opciones por cada pregunta, ¿tiene al menos una probabilidad en cinco de adivinar correctamente! Si un examen de opción múltiple contiene 100 preguntas, cada una con cinco posibles

respuestas, ¿cuál es la calificación esperada para un estudiante que está adivinando en cada pregunta? ¿Dentro de qué límites caerán las calificaciones de “no saber”?

Solución Si x es el número de respuestas correctas en el examen de 100 preguntas, la probabilidad de una respuesta correcta, p , es una en cinco, de modo que $p = .2$. Como el estudiante selecciona respuestas al azar, las $n = 100$ respuestas son independientes y la calificación esperada para esta variable aleatoria binomial es

$$\mu = np = 100(.2) = 20 \quad \text{respuestas correctas}$$

Para evaluar la dispersión o variabilidad de las calificaciones, se puede calcular

$$\sigma = \sqrt{npq} = \sqrt{100(.2)(.8)} = 4$$

Entonces, usando su conocimiento de variación a partir del teorema de Chebyshev y la regla empírica, puede hacer estos enunciados:

- Una gran proporción de las calificaciones estará a no más de dos desviaciones estándar de la media, o sea de $20 - 8 = 12$ a $20 + 8 = 28$.
- Casi todas las calificaciones estarán a no más de tres desviaciones estándar de la media, o sea de $20 - 12 = 8$ a $20 + 12 = 32$.

La opción de “adivinar” da al estudiante una mejor calificación que la de cero en el examen de memoria, pero el estudiante todavía no pasará el examen. ¿Qué otras opciones tiene el estudiante?

3.3

EJERCICIOS

TÉCNICAS BÁSICAS

3.15 Considere una variable aleatoria binomial con $n = 8$ y $p = .7$. Sea x el número de éxitos en la muestra.

- a. Encuentre la probabilidad de que x sea 3 o menor.
- b. Encuentre la probabilidad de que x sea 3 o mayor.
- c. Encuentre $P(x < 3)$.
- d. Encuentre $P(x = 3)$.
- e. Encuentre $P(3 \leq x \leq 5)$.

3.16 Considere una variable aleatoria binomial con $n = 9$ y $p = .3$. Sea x el número de éxitos en la muestra.

- a. Encuentre la probabilidad de que x sea exactamente 2.
- b. Encuentre la probabilidad de que x sea menor que 2.
- c. Encuentre $P(x > 2)$.
- d. Encuentre $P(2 \leq x \leq 4)$.

3.17 Evalúe estas probabilidades binomiales:

- a. $C_2^8(.3)^2(.7)^6$
- b. $C_0^4(.05)^0(.95)^4$
- c. $C_3^{10}(.5)^3(.5)^7$
- d. $C_1^7(.2)^1(.8)^6$

3.18 Evalúe estas probabilidades binomiales:

- a. $C_0^8(.2)^0(.8)^8$
- b. $C_1^8(.2)^1(.8)^7$
- c. $C_2^8(.2)^2(.8)^6$
- d. $P(x \leq 1)$ cuando $n = 8$, $p = .2$
- e. $P(\text{dos éxitos o menos})$

3.19 Sea x una variable aleatoria binomial con $n = 7$, $p = .3$. Encuentre estos valores:

- a. $P(x = 4)$
- b. $P(x \leq 1)$
- c. $P(x > 1)$
- d. $\mu = np$
- e. $\sigma = \sqrt{npq}$

3.20 Use la fórmula para la distribución binomial de probabilidad para calcular los valores de $p(x)$ y construya el histograma de probabilidad para x cuando $n = 6$ y $p = .2$. [SUGERENCIA: Calcule $P(x = k)$ para siete valores diferentes de k .]

3.21 Consulte el ejercicio 3.20. Construya el histograma de probabilidad para una variable aleatoria binomial x con $n = 6$ y $p = .8$. Use los resultados del ejercicio 3.20; no calcule de nuevo todas las probabilidades.

3.22 Si x tiene una distribución binomial con $p = .5$, ¿la forma de la distribución de probabilidad será simétrica, sesgada a la izquierda o sesgada a la derecha?

3.23 Sea x una variable aleatoria binomial con $n = 10$ y $p = .4$. Encuentre estos valores:

- a. $P(x = 4)$
- b. $P(x \geq 4)$
- c. $P(x > 4)$
- d. $P(x \leq 4)$
- e. $\mu = np$
- f. $\sigma = \sqrt{npq}$

3.24 Use la tabla 1 del Anexo para hallar la suma de las probabilidades binomiales de $x = 0$ a $x = k$ para estos casos:

- a. $n = 10, p = .1, k = 3$
 b. $n = 15, p = .6, k = 7$
 c. $n = 25, p = .5, k = 14$

3.25 Use la tabla 1 del Anexo para evaluar las siguientes probabilidades para $n = 6$ y $p = .8$:

- a. $P(x \geq 4)$ b. $P(x = 2)$
 c. $P(x < 2)$ d. $P(x > 1)$

Verifique estas respuestas usando los valores de $p(x)$ calculados en el ejercicio 3.21.

3.26 Encuentre $P(x \leq k)$ en cada caso:

- a. $n = 20, p = .05, k = 2$
 b. $n = 15, p = .7, k = 8$
 c. $n = 10, p = .9, k = 9$

3.27 Use la tabla 1 del Anexo para hallar lo siguiente:

- a. $P(x < 12)$ para $n = 20, p = .5$
 b. $P(x \leq 6)$ para $n = 15, p = .4$
 c. $P(x > 4)$ para $n = 10, p = .4$
 d. $P(x \geq 6)$ para $n = 15, p = .6$
 e. $P(3 < x < 7)$ para $n = 10, p = .5$

3.28 Encuentre la media y la desviación estándar para una distribución binomial con estos valores:

- a. $n = 1000, p = .3$ b. $n = 400, p = .01$
 c. $n = 500, p = .5$ d. $n = 1600, p = .8$

3.29 Encuentre la media y la desviación estándar para una distribución binomial con $n = 100$ y estos valores de p :

- a. $p = .01$ b. $p = .9$ c. $p = .3$
 d. $p = .7$ e. $p = .5$

3.30 En el ejercicio 3.29, se calcularon la media y la desviación estándar para una variable aleatoria binomial para un tamaño muestral fijo, $n = 100$, y para valores diferentes de p . Grafique los valores de la desviación estándar para los cinco valores de p dados en el ejercicio 3.29. ¿Para qué valores de p la desviación estándar parece ser un máximo?

3.31 Sea x una variable aleatoria binomial con $n = 20$ y $p = .1$.

- a. Calcule $P(x \leq 4)$ usando la fórmula binomial.
 b. Calcule $P(x \leq 4)$ usando la tabla 1 del Anexo.
 c. Use la salida *Excel* siguiente para calcular $P(x \leq 4)$. Compare los resultados de las partes a, b y c.
 d. Calcule la media y la desviación estándar de la variable aleatoria x .
 e. Use los resultados de la parte d para calcular los intervalos $\mu \pm \sigma$, $\mu \pm 2\sigma$, y $\mu \pm 3\sigma$. Encuentre la

probabilidad de que una observación caiga en cada uno de estos intervalos.

- f. ¿Los resultados de la parte e son consistentes con el teorema de Chebyshev? ¿Con la regla empírica? ¿Por qué?

Salida *Excel* para el ejercicio 3.76: Binomial con $n = 20$ y $p = .1$

x	$p(x)$	x	$p(x)$
0	0.1216	11	7E-07
1	0.2702	12	5E-08
2	0.2852	13	4E-09
3	0.1901	14	2E-10
4	0.0898	15	9E-12
5	0.0319	16	3E-13
6	0.0089	17	8E-15
7	0.0020	18	2E-16
8	0.0004	19	2E-18
9	0.0001	20	1E-20
10	0.0000		

APLICACIONES

3.32 El problema de la urna Un frasco contiene cinco pelotas: tres rojas y dos blancas. Se eligen dos pelotas al azar sin devolverlas al frasco, y se registra el número x de pelotas rojas. Explique por qué x es o no una variable aleatoria binomial. (SUGERENCIA: Compare las características de este experimento con las de un experimento binomial dado en esta sección.) Si el experimento es binomial, dé los valores de n y p .

3.33 El problema de la urna, continúa Consulte el ejercicio 3.32. Suponga que el muestreo fue realizado con reposición. Esto es, suponga que la primera pelota se seleccionó del frasco, se observó y después fue devuelta al frasco, y que entonces las pelotas se mezclaron antes de seleccionar la segunda pelota. Explique por qué x , el número de pelotas rojas observado, es o no una variable aleatoria binomial. Si el experimento es binomial, dé los valores de n y p .

3.34 Clima en Buenos Aires Un meteorólogo registró el número de días de lluvia durante un periodo de 30 días. Si la variable aleatoria x se define como el número de días de lluvia, ¿ x tiene una distribución binomial? Si no es así, ¿por qué no? Si es así, ¿se conocen los valores de n y de p ?

3.35 Telemercadeo Una empresa de investigación de mercado contrata operadores para realizar encuestas por teléfono. La computadora marca al azar un número telefónico y la operadora pregunta a quien conteste si tiene tiempo para contestar algunas preguntas. Sea x el número de llamadas telefónicas hechas hasta que el primer entrevistado está dispuesto a contestar las preguntas de la operadora. ¿Es éste un experimento binomial? Explique.

3.36 Calificaciones del SAT (examen de aptitud

escolar) En 2010, el promedio combinado de calificaciones del SAT (lectura analítica, matemáticas y escritura) para estudiantes que van hacia la universidad en Estados Unidos fue 1 509 de 2 400. Suponga que aproximadamente 45% de todos los graduados de preparatoria presentan este examen y que 100 son seleccionados al azar en todo Estados Unidos.³ ¿Cuál de las siguientes variables aleatorias tiene una distribución binomial aproximada? Si es posible, dé los valores para n y p .

- El número de estudiantes que presentaron el SAT.
- Las calificaciones de los 100 estudiantes en el SAT.
- El número de estudiantes que calificaron arriba del promedio del SAT.
- El tiempo que tomó a cada estudiante para completar el SAT.

3.37 Sistemas de seguridad El sistema de seguridad de una casa está diseñado para tener 99% de confiabilidad. Suponga que nueve casas equipadas con este sistema experimentan un intento de robo. Encuentre las probabilidades de estos eventos:

- Al menos una alarma se activó.
- Más de siete alarmas se activaron.
- Ocho o menos alarmas se activaron.

3.38 Tipos de sangre En cierta población, 85% de las personas tienen sangre tipo Rh positivo. Suponga que dos personas de esta población se casan. ¿Cuál es la probabilidad de que ambos sean Rh negativo, lo cual hace inevitable que sus hijos sean Rh negativo?

3.39 Colores de autos La preferencia por el color de un auto cambia con los años y de acuerdo con el modelo particular que seleccione el cliente. En un año reciente, suponga que 10% de todos los autos de lujo que se vendieron eran negros. Si 25 autos de ese año y tipo se seleccionan al azar, encuentre las siguientes probabilidades:

- Al menos cinco autos son negros.
- A lo sumo seis autos son negros.
- Más de cuatro autos son negros.
- Exactamente cuatro autos son negros.
- Entre tres y cinco autos (inclusive) son negros.
- Más de 20 autos no son negros.

3.40 Cuentas del médico Unos registros muestran que 30% de los pacientes ingresados en una clínica médica dejan de pagar sus cuentas y que, en última instancia, esas cuentas son condonadas. Suponga que $n = 4$ nuevos pacientes representan una selección aleatoria de entre un

gran conjunto de prospectos de pacientes atendidos por la clínica. Encuentre estas probabilidades:

- Las cuentas de todos los pacientes tendrán finalmente que condonarse.
- Una tendrá que condonarse.
- Ninguna tendrá que condonarse.

3.41 Cuentas del médico II Consulte el ejercicio 3.40, donde 30% de los pacientes admitidos dejan de pagar sus cuentas y que las deudas al final son condonadas. Suponga que la clínica trata a 2 000 pacientes diferentes durante un periodo de un año, y sea x el número de deudas condonadas.

- ¿Cuál es la media del número (esperado) de deudas que se deberán condonar?
- Encuentre la varianza y la desviación estándar de x .
- ¿Qué puede decir sobre la probabilidad de que x excederá de 700? (SUGERENCIA: Use los valores de μ y σ , junto con el teorema de Chebyshev.)

3.42 Infestación de la mosca blanca Suponga que 10% de los campos en una región agrícola determinada están infestados con la mosca blanca de la remolacha. Se seleccionan de manera aleatoria 100 campos de esta región y se inspeccionan para ver si están infestados.

- ¿Cuál es el número promedio de campos muestreados que están infestados de la mosca blanca?
- ¿Dentro de qué límites esperaría usted hallar el número de campos infestados, con probabilidad aproximada de 95%?
- ¿Qué podría usted concluir si encuentra que $x = 25$ campos estuvieran infestados? ¿Es posible que una de las características de un experimento binomial no se satisfaga en este experimento? Explique.

3.43 Preferencias de color en ratones En un experimento de psicología la investigadora diseña un laberinto en el que el ratón debe elegir uno de dos caminos, ya sea el de color rojo o el azul, en cada una de las 10 intersecciones. Al final del laberinto, el ratón es recompensado con alimento. La investigadora cuenta el número de veces que el ratón elige el camino rojo. Si usted fuera la investigadora, ¿cómo usaría este conteo para determinar si el ratón tiene alguna preferencia por un color?

3.44 Dolor de espalda Seis de cada 10 personas adultas dicen que el dolor de la espalda baja limita en forma considerable sus actividades atléticas.⁴ Se le preguntó a una muestra al azar de $n = 8$ adultos si el dolor de la espalda baja era un factor que limitaba sus actividades atléticas. La salida impresa del MINITAB muestra las probabilidades acumulativas e individuales para una variable aleatoria binomial con $n = 8$ y $p = .6$.

Salida impresa de MINITAB para el ejercicio 3.89

Función acumulativa de distribución

Binomial with n = 8
and p = 0.6

x	P(X ≤ x)
0	0.00066
1	0.00852
2	0.04981
3	0.17367
4	0.40591
5	0.68461
6	0.89362
7	0.98320
8	1.00000

Función de densidad de probabilidad

Binomial with n = 8
and p = 0.6

x	P(X = x)
0	0.000655
1	0.007864
2	0.041288
3	0.123863
4	0.232243
5	0.278692
6	0.209019
7	0.089580
8	0.016796

- Use la fórmula binomial para hallar la probabilidad de que los ocho indiquen que el dolor de la espalda baja era un factor limitante en sus actividades atléticas.
- Confirme los resultados de la parte a usando la salida impresa.
- ¿Cuál es la probabilidad de que a lo sumo siete individuos digan que el dolor de la espalda baja es un factor limitante en sus actividades atléticas?

3.45 Prueba del gusto por el PTC La prueba del gusto por el PTC (feniltiocarbamida) es un ejercicio favorito

para toda clase de genética humana. Se ha establecido que un solo gen determina la característica y que 70% de los estadounidenses son “degustadores”, en tanto que 30% son “no degustadores”. Suponga que se eligen al azar 20 estadounidenses y se someten a la prueba del gusto del PTC.

- ¿Cuál es la probabilidad de que 17 o más sean “degustadores”?
- ¿Cuál es la probabilidad de que 15 o menos sean “degustadores”?

3.46 El mejor amigo del hombre Según la Sociedad protectora de animales de Estados Unidos, hay aproximadamente 77.5 millones de perros con dueño en Estados Unidos y alrededor del 40% de los hogares estadounidenses tienen al menos un perro.⁵ Suponga que la cifra 40% es correcta y que se seleccionan al azar 15 hogares para un estudio sobre la propiedad de mascotas.

- ¿Cuál es la probabilidad de que exactamente ocho de los hogares tengan al menos un perro?
- ¿Cuál es la probabilidad de que a lo sumo cuatro de los hogares tengan al menos un perro?
- ¿Cuál es la probabilidad de que más de 10 hogares tengan al menos un perro?

DISTRIBUCIONES DE PROBABILIDAD PARA VARIABLES ALEATORIAS CONTINUAS

3.4

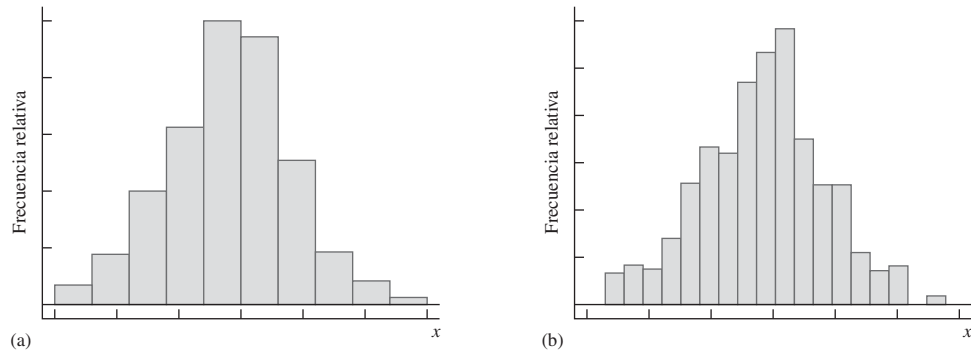
Cuando una variable aleatoria x es discreta, se asigna una probabilidad positiva a cada uno de los valores que x pueda tomar y obtener la distribución de probabilidad para x . La suma de todas las probabilidades asociada con los diferentes valores de x es 1, pero no todos los experimentos resultan en variables aleatorias que sean discretas.

Las **variables aleatorias continuas** (por ejemplo estaturas y pesos, lapso de vida útil de un producto en particular o un error experimental de laboratorio) pueden tomar los infinitamente numerosos valores correspondientes a los puntos en un intervalo de una recta. Si se trata de asignar una probabilidad positiva a cada uno de estos innumerables valores, las probabilidades ya no sumarán 1, como cuando se trata de variables aleatorias discretas. Por tanto, se debe usar un método diferente para generar la distribución de probabilidad para una variable aleatoria continua.

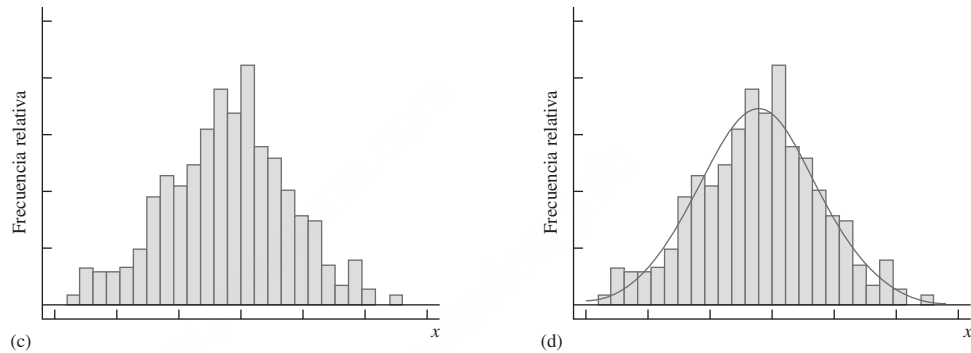
Suponga que usted tiene un conjunto de mediciones en una variable aleatoria continua y que crea un histograma de frecuencia relativa para describir su distribución. Para un pequeño número de mediciones, se usa un pequeño número de clases; entonces, a medida que se recolecten más y más mediciones, se usan más clases y se reduce el ancho de clase. El perfil del histograma cambiará ligeramente, casi todo el tiempo haciéndose cada vez menos irregular, como se muestra en la figura 3.6. Cuando el número de mediciones se hace muy grande y los anchos de clase se hacen muy angostos, el histograma de frecuencia relativa aparece cada vez más como la curva suave que aparece en la figura 3.6d). Esta curva suave describe la **distribución de probabilidad de la variable aleatoria continua**.

FIGURA 3.6

Histogramas de frecuencia relativa para tamaños muestrales cada vez más crecientes

**FIGURA 3.6**

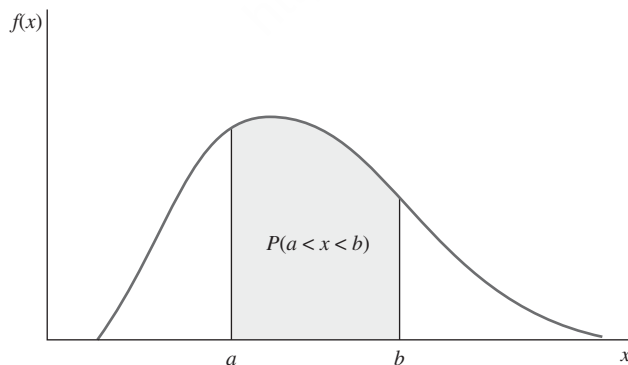
Continuación



¿Cómo se puede crear un modelo para esta distribución de probabilidad? Una variable aleatoria continua puede tomar cualquiera de un número infinito de valores de la recta real, en forma semejante al número infinito de granos de arena en una playa. La distribución de probabilidad se crea al distribuir una unidad de probabilidad a lo largo de la recta, igual que como se puede distribuir un puñado de arena. La probabilidad, es decir granos de arena o de mediciones, se apilarán en ciertos lugares y el resultado es la distribución de probabilidad mostrada en la figura 3.7. La profundidad o **densidad** de la probabilidad, que varía con x , se describe por una fórmula matemática $f(x)$, llamada **distribución de probabilidad** o **función de densidad de probabilidad** para la variable aleatoria x .

FIGURA 3.7

La distribución de probabilidad $f(x)$; $P(a < x < b)$ es igual al área sombreada bajo la curva

**TIP** MI CONSEJO

Para variables aleatorias continuas,
área = probabilidad

Recuerde que, para variables aleatorias discretas, 1) la suma de todas las probabilidades $p(x)$ es igual a 1, y 2) la probabilidad de que x quede en un intervalo determinado es la suma de todas las probabilidades en ese intervalo. Las variables aleatorias continuas tienen algunas características paralelas que se enumeran a continuación.

- El área bajo una distribución continua de probabilidad es igual a 1.
- La probabilidad de que x caiga en un intervalo particular, por ejemplo de a a b , es igual al área bajo la curva entre los dos puntos a y b . Ésta es el área sombreada de la figura 3.7.

TIP MI CONSEJO

El área bajo la curva es igual a 1

También hay una diferencia importante entre variables aleatorias discretas y continuas. Considere la probabilidad de que x sea igual a algún valor en particular, por ejemplo a . Como no hay área arriba de un solo punto, por ejemplo $x = a$, en la distribución de probabilidad para una variable aleatoria continua, nuestra definición implica que la probabilidad es 0.

- $P(x = a) = 0$ para variables aleatorias continuas.
- Esto implica que $P(x \geq a) = P(x > a)$ y $P(x \leq a) = P(x < a)$.
- Esto *no* es cierto en general para variables aleatorias discretas.

¿Cómo se elige el modelo, es decir la distribución de probabilidad $f(x)$, apropiado para un experimento dado? Existen muchos tipos de curvas continuas para modelar. Algunas son de forma de montículo, como la de la figura 3.6d), pero otras no lo son. En general, trate de elegir un modelo que satisfaga estos criterios:

- Se ajusta al cuerpo de datos acumulado.
- Permite hacer las mejores inferencias posibles usando los datos.

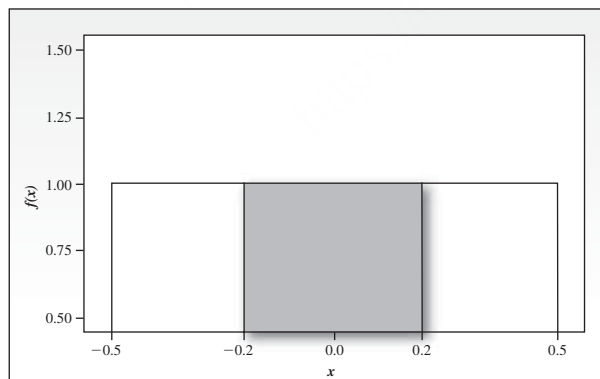
EJEMPLO

3.13

La *variable aleatoria uniforme** se emplea para modelar el comportamiento de una variable aleatoria continua cuyos valores estén uniforme o equitativamente distribuidos en un intervalo dado. Por ejemplo, es probable que el error x introducido al redondear una observación a la pulgada más cercana tenga una distribución uniforme en el intervalo de $-.5$ a $.5$. La función de densidad de probabilidad $f(x)$ sería “plana” como se muestra en la figura 3.8. La altura del rectángulo está fija en 1, de modo que el área total bajo la distribución de probabilidad es 1.

FIGURA 3.8

Una distribución uniforme de probabilidad



¿Cuál es la probabilidad de que el error de redondeo sea menor que .2 en magnitud?

Solución Esta probabilidad corresponde al área bajo la distribución entre $x = -.2$ y $x = .2$. Como la altura del rectángulo es 1,

$$P(-.2 < x < .2) = [.2 - (-.2)] \times 1 = .4$$

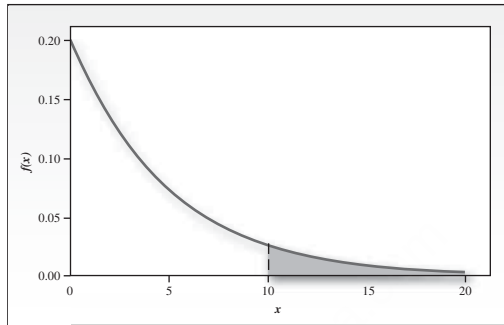
EJEMPLO

3.14

La *variable aleatoria exponencial** se utiliza para modelar variables aleatorias continuas tales como tiempos de espera o vidas útiles asociadas con componentes electrónicos. Por ejemplo, el tiempo de espera en una caja de pago de un supermercado tiene una distribución exponencial con un tiempo de espera promedio de 5 minutos. La función de densidad de probabilidad $f(x) = .2e^{-.2x}$ se muestra en la figura 3.9. Para hallar áreas bajo esta curva, se usa el hecho de que $P(x > a) = e^{-.2a}$ para $a > 0$. ¿Cuál es la probabilidad de que usted tenga que esperar 10 minutos o más en la caja de pago del supermercado?

FIGURA 3.9

Una distribución de probabilidad exponencial



Solución La probabilidad que se va a calcular es el área sombreada en la figura 3.7. Use la fórmula general para $P(x > a)$ para hallar

$$P(x > 10) = e^{-.2(10)} = .135$$

Su modelo no siempre puede ajustar perfectamente la situación experimental, pero debe tratar de elegir un modelo que *mejor se ajuste* al histograma de frecuencia relativa poblacional. Cuanto más se aproxime el modelo a la realidad, mejores serán las inferencias. Por fortuna, muchas variables aleatorias continuas tienen distribuciones de frecuencia de forma de montículo, por ejemplo los datos de la figura 3.1d). La **distribución normal de probabilidad** da un buen modelo para describir este tipo de datos.

LA DISTRIBUCIÓN DE PROBABILIDAD DE POISSON

3.5

Otra variable aleatoria discreta que tiene numerosas aplicaciones prácticas es la **variable aleatoria de Poisson**. Su distribución de probabilidad da un buen modelo para datos que representan el número de sucesos de un evento especificado en una unidad determinada de tiempo o espacio. A continuación vemos algunos ejemplos de experimentos para los cuales la variable aleatoria x puede ser modelada por la variable aleatoria de Poisson:

- El número de llamadas recibidas por un especialista de soporte técnico durante un tiempo determinado
- El número de bacterias por volumen pequeño de fluido
- El número de llegadas de clientes al mostrador de una caja de pago durante un minuto determinado
- El número de descomposturas de una máquina durante un día determinado

* La función de densidad de probabilidad (*fdp*) para la variable aleatoria uniforme general x es $f(x) = 1/(b - a)$ para $a \leq x \leq b$ y con media $\mu = (a + b)/2$ y $\sigma^2 = (b - a)^2/12$. Para la *variable aleatoria exponencial* x , la *fdp* es $f(x) = \lambda e^{-\lambda x}$ para $x \geq 0$, $\lambda > 0$, con $\mu = 1/\lambda$ y $\sigma^2 = 1/\lambda^2$.

- El número de accidentes de tránsito en una sección de autopista durante un tiempo determinado

En cada ejemplo, x representa el número de eventos que ocurren en un periodo o espacio, durante el cual se espera que ocurra un promedio de μ de estos eventos. Las únicas suposiciones necesarias, cuando uno usa la distribución de Poisson para modelar experimentos como éstos, son que las cuentas o eventos ocurren aleatoria e **independientemente** unos de otros. La fórmula para la distribución de probabilidad de Poisson, así como su media y varianza, se dan a continuación.

LA DISTRIBUCIÓN DE PROBABILIDAD DE POISSON

Sea μ el número promedio de veces que ocurre un evento en cierto periodo o espacio. La probabilidad de k sucesos de este evento es

$$P(x = k) = \frac{\mu^k e^{-\mu}}{k!}$$

para valores de $k = 0, 1, 2, 3, \dots$. La media y la desviación estándar de la variable aleatoria de Poisson x son

$$\text{Media: } \mu \quad \text{Desviación estándar: } \sigma = \sqrt{\mu}$$

TIP MI CONSEJO

Utilice la fórmula de Poisson o la tabla 2 para calcular probabilidades de Poisson

El símbolo $e = 2.71828 \dots$ se evalúa usando su calculadora científica, que debe tener una función como e^x . Para cada valor de k , puede obtener las probabilidades individuales para la variable aleatoria de Poisson, igual que como lo hizo para la variable aleatoria binomial.

EJEMPLO

3.15

El número promedio de accidentes de tránsito en cierto cruce de carretera es dos por semana. Suponga que el número de accidentes sigue una distribución de Poisson con $\mu = 2$.

1. Encuentre la probabilidad de que no haya accidentes en este cruce de carretera durante un periodo de 1 semana.
2. Encuentre la probabilidad de que haya tres accidentes como máximo en esta sección de carretera durante un periodo de 2 semanas.

Solución

1. El número promedio de accidentes por semana es $\mu = 2$. Por tanto, la probabilidad de que no haya accidentes en esta sección de carretera durante 1 semana determinada es

$$P(x = 0) = p(0) = \frac{2^0 e^{-2}}{0!} = e^{-2} = .135335$$

2. Durante un periodo de 2 semanas, el número promedio de accidentes en esta sección de carretera es $2(2) = 4$. La probabilidad de que a lo sumo haya tres accidentes durante un periodo de 2 semanas es

$$P(x \leq 3) = p(0) + p(1) + p(2) + p(3)$$

donde

$$\begin{aligned} p(0) &= \frac{4^0 e^{-4}}{0!} = .018316 & p(2) &= \frac{4^2 e^{-4}}{2!} = .146525 \\ p(1) &= \frac{4^1 e^{-4}}{1!} = .073263 & p(3) &= \frac{4^3 e^{-4}}{3!} = .195367 \end{aligned}$$

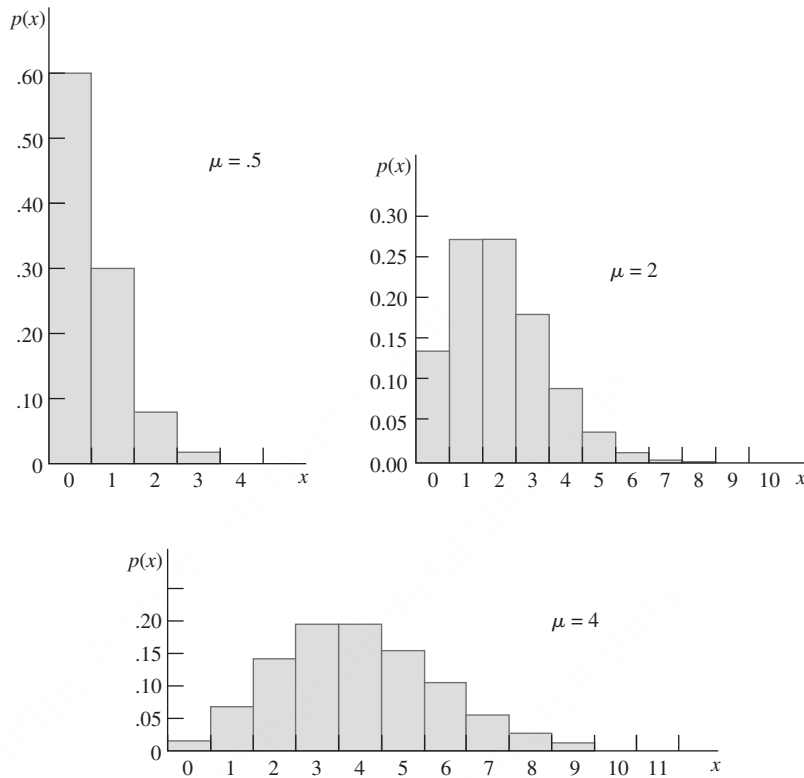
Por tanto,

$$P(x \leq 3) = .018316 + .073263 + .146525 + .195367 = .433471$$

Una vez que se han calculado los valores para $p(x)$, puede usarlos para construir un histograma de probabilidad para la variable aleatoria x . Las gráficas de la distribución de probabilidad de Poisson para $\mu = .5, 2$ y 4 se muestran en la figura 3.10.

FIGURA 3.10

Distribuciones de probabilidad de Poisson para $\mu = .5, 2$ y 4



De manera alternativa, pueden usarse **tablas acumulativas de Poisson** (tabla 2 del Anexo) o probabilidades acumulativas o individuales generadas por *MINITAB* o *MS Excel*. Estas dos opciones son por lo general más convenientes que hacer el cálculo en forma manual. Los procedimientos son semejantes a los empleados para la variable aleatoria binomial.



NECESITO SABER...

Cómo usar la tabla 2 para calcular probabilidades de Poisson

1. Encuentre el valor necesario de μ . Aísle la columna apropiada en la tabla 2.
2. La tabla 2 da $P(x \leq k)$ en la fila marcada k . Reescriba la probabilidad que necesite de modo que esté en esta forma.
 - Haga una lista de los valores de x en su evento.
 - De la lista, escriba el evento ya sea como la diferencia de dos probabilidades:

$$P(x \leq a) - P(x \leq b) \text{ para } a > b$$

o el complemento del evento:

$$1 - P(x \leq a)$$

o tan solo el evento mismo:

$$P(x \leq a) \text{ o } P(x < a - 1)$$

EJEMPLO

3.16

Consulte el ejemplo 3.15, donde se calcularon probabilidades para una distribución de Poisson con $\mu = 2$ y $\mu = 4$. Use la tabla de Poisson acumulativa para encontrar las probabilidades de estos eventos:

1. No hubo accidentes durante un periodo de una semana.
2. Cuando mucho tres accidentes durante un periodo de dos semanas.

Solución

Una parte de la tabla 2 del Anexo se muestra en la figura 3.11.

FIGURA 3.11

Porción de la tabla 2 en el Anexo

k	μ				
	2.0	2.5	3.0	3.5	4.0
0	.135	.082	.055	.033	.018
1	.406	.287	.199	.136	.092
2	.677	.544	.423	.321	.238
3	.857	.758	.647	.537	.433
4	.947	.891	.815	.725	.629
5	.983	.958	.916	.858	.785
6	.995	.986	.966	.935	.889
7	.999	.996	.988	.973	.949
8	1.000	.999	.996	.990	.979
9		1.000	.999	.997	.992
10			1.000	.999	.997
11				1.000	.999
12					1.000

1. Del ejemplo 3.15, el número promedio de accidentes en un periodo de una semana es $\mu = 2.0$. Por consiguiente, la probabilidad de que no haya accidentes en un periodo de una semana puede leerse directamente de la tabla 2 en la columna marcada “2.0” como $P(x = 0) = p(0) = .135$.
2. El número promedio de accidentes en un periodo de dos semanas es $2(2) = 4$. Por consiguiente, la probabilidad de cuando mucho tres accidentes en un periodo de dos semanas se encuentra en la tabla 2, indexando $\mu = 4.0$ y $k = 3$ como $P(x \leq 3) = .433$.

Ambas probabilidades coinciden con el cálculo realizado en el ejemplo 3.15, correcto hasta tres lugares decimales.

En la sección 3.8 utilizamos las tablas binomiales acumulativas para simplificar el cálculo de probabilidades binomiales. Desafortunadamente, en situaciones prácticas, con frecuencia n es grande y no se dispone de tablas.

TIP MI CONSEJO

Se pueden estimar probabilidades binomiales con la de Poisson cuando n es grande y p es pequeña

LA APROXIMACIÓN DE POISSON A LA DISTRIBUCIÓN BINOMIAL

La distribución de probabilidad de Poisson da una aproximación sencilla, fácil de calcular y precisa a probabilidades binomiales cuando n es grande y $\mu = np$ es pequeña, de preferencia con $np < 7$. Una aproximación apropiada para valores más grandes de $\mu = np$ se dará en el tema de la distribución normal.

EJEMPLO

3.17

Suponga que una compañía de seguros de vida asegura la vida de 5 000 hombres de 42 años de edad. Si estudios actuariales muestran que la probabilidad de que cualquier hombre de 42 años muera en un año determinado es .001, encuentre la probabilidad exacta de que la compañía tendrá que pagar $x = 4$ reclamaciones durante un año determinado.

Solución La probabilidad exacta está dada por la distribución binomial como

$$P(x = 4) = p(4) = \frac{5\,000!}{4!4\,996!}(.001)^4(.999)^{4\,996}$$

para la cual no se dispone de tablas binomiales. Calcular $P(x = 4)$ sin ayuda de una calculadora científica o una computadora sería muy lento, pero la distribución de Poisson se puede usar para dar una buena aproximación para $P(x = 4)$. Calculando $\mu = np = (5\,000)(.001) = 5$ y sustituyendo en la fórmula para la distribución de probabilidad de Poisson, tenemos

$$p(4) \approx \frac{\mu^4 e^{-\mu}}{4!} = \frac{5^4 e^{-5}}{4!} = \frac{(625)(.006738)}{24} = .175$$

El valor de $p(4)$ podría obtenerse también usando la tabla 2 del Anexo con $\mu = 5$ como

$$p(4) = P(x \leq 4) - P(x \leq 3) = .440 - .265 = .175$$

EJEMPLO

3.18

Una fabricante de podadoras para el pasto compra a un proveedor motores de 1 hp y 2 ciclos en lotes de 1 000. Entonces se equipa cada una de las podadoras producidas por su planta con uno de los motores. La historia muestra que la probabilidad de que cualquier motor del proveedor resulte no satisfactorio es .001. En un embarque de 1 000 motores, ¿cuál es la probabilidad de que ninguno sea defectuoso? ¿Son tres? ¿Son cuatro?

Solución Éste es un experimento binomial con $n = 1\,000$ y $p = .001$. El número esperado de motores defectuosos en un embarque de $n = 1\,000$ motores es $\mu = np = (1\,000)(.001) = 1$. Como éste es un experimento binomial con $np < 7$, la probabilidad de x motores defectuosos en este embarque puede aproximarse con

$$P(x = k) = p(k) = \frac{\mu^k e^{-\mu}}{k!} = \frac{1^k e^{-1}}{k!} = \frac{e^{-1}}{k!}$$

Por lo tanto,

$$p(0) \approx \frac{e^{-1}}{0!} = \frac{.368}{1} = .368$$

$$p(3) \approx \frac{e^{-1}}{3!} = \frac{.368}{6} = .061$$

$$p(4) \approx \frac{e^{-1}}{4!} = \frac{.368}{24} = .015$$

Las probabilidades individuales de Poisson para $\mu = 1$, junto con las probabilidades binomiales individuales para $n = 1000$ y $p = .001$ y $x = 0, 1, \dots, 10$, fueron generadas por *MS Excel* y se muestran en la figura 3.12. Las probabilidades individuales, aun cuando se calculan con fórmulas totalmente diferentes, son casi iguales. Las probabilidades binomiales exactas están en la sección izquierda de la figura 3.12, y las aproximaciones de Poisson están a la derecha.

FIGURA 3.12

Salida impresa de Excel de probabilidades binomiales y de Poisson

x	Binomial $p(x)$	x	Poisson $p(x)$
0	0.3677	0	0.3679
1	0.3681	1	0.3679
2	0.1840	2	0.1839
3	0.0613	3	0.0613
4	0.0153	4	0.0153
5	0.0030	5	0.0031
6	0.0005	6	0.0005
7	0.0001	7	0.0001
8	0.0000	8	0.0000
9	0.0000	9	0.0000
10	0.0000	10	0.0000

3.5 EJERCICIOS

TÉCNICAS BÁSICAS

3.47 Considere una variable aleatoria de Poisson con $\mu = 2.5$. Use la fórmula de Poisson para calcular las siguientes probabilidades:

- a. $P(x = 0)$ b. $P(x = 1)$
c. $P(x = 2)$ d. $P(x \leq 2)$

3.48 Considere una variable aleatoria de Poisson con $\mu = 3$. Use la fórmula de Poisson para calcular las siguientes probabilidades:

- a. $P(x = 0)$ b. $P(x = 1)$
c. $P(x > 1)$

3.49 Considere una variable aleatoria de Poisson con $\mu = 3$. Use la tabla 2 para encontrar las siguientes probabilidades:

- a. $P(x \leq 3)$ b. $P(x > 3)$
c. $P(x = 3)$ d. $P(3 \leq x \leq 5)$

3.50 Considere una variable aleatoria de Poisson con $\mu = 0.8$. Use la tabla 2 para encontrar las siguientes probabilidades:

- a. $P(x = 0)$ b. $P(x \leq 2)$
c. $P(x > 2)$ d. $P(2 \leq x \leq 4)$

3.51 Sea x una variable aleatoria de Poisson con media $\mu = 2$. Calcule estas probabilidades:

- a. $P(x = 0)$ b. $P(x = 1)$
c. $P(x > 1)$ d. $P(x = 5)$

3.52 Sea x una variable aleatoria de Poisson con media $\mu = 2.5$. Use la tabla 2 del Anexo para calcular estas probabilidades:

- a. $P(x \geq 5)$ b. $P(x < 6)$
c. $P(x = 2)$ d. $P(1 \leq x \leq 4)$

3.53 Poisson vs. binomial Sea x una variable aleatoria binomial con $n = 20$ y $p = .1$.

- a. Calcule $P(x \leq 2)$ usando la tabla 1 del Anexo para obtener la probabilidad binomial exacta.

- b. Use la aproximación de Poisson para calcular $P(x \leq 2)$.

- c. Compare los resultados de las partes a y b. ¿Es precisa la aproximación?

3.54 Poisson vs binomial II Para ilustrar qué tan bien la distribución de probabilidad de Poisson aproxima la distribución binomial de probabilidad, calcule los valores aproximados de Poisson para $p(0)$ y $p(1)$ para una distribución binomial de probabilidad con $n = 25$ y $p = .05$. Compare las respuestas contra los valores exactos obtenidos de la tabla 1 del Anexo.

APLICACIONES

3.55 Seguridad en un aeropuerto El incremento del número de vuelos regionales cortos en aeropuertos importantes ha aumentado la preocupación por la seguridad en el aire. Un aeropuerto de la región este ha registrado un promedio mensual de cinco accidentes a punto de ocurrir en aterrizajes y despegues en los últimos 5 años.

- a. Encuentre la probabilidad de que durante un mes determinado no haya accidentes a punto de ocurrir en aterrizajes y despegues en el aeropuerto.
b. Encuentre la probabilidad de que durante un mes determinado haya cinco accidentes a punto de ocurrir.
c. Encuentre la probabilidad de que haya al menos cinco accidentes a punto de ocurrir durante un mes particular.

3.56 Cuidados intensivos El número x de personas ingresadas a una unidad de cuidados intensivos en un hospital particular, en cualquier día, tiene una distribución de probabilidad de Poisson con media igual a cinco personas por día.

- a. ¿Cuál es la probabilidad de que el número de personas ingresadas a una unidad de cuidados intensivos, en un día particular, sea dos? ¿Menor o igual a dos?
b. ¿Es probable que x exceda de 10? Explique.

3.57 Propenso a accidentes De acuerdo con un estudio realizado por el Instituto Nacional de Pediatría de la Ciudad de México, los niños que se lesionan dos o más veces tienden a sufrir estas lesiones durante un tiempo relativamente limitado, por lo general un año o menos. Si el número promedio de lesiones por año para niños en edad escolar es de dos, ¿cuáles son las probabilidades de estos eventos?

- Un niño en edad escolar sufrirá dos lesiones durante el año.
- Un niño en edad escolar sufrirá dos o más lesiones durante el año.
- Un niño en edad escolar sufrirá a lo sumo una lesión durante el año.

3.58 Propenso a accidentes, continúa Consulte el ejercicio 3.57.

- Calcule la media y la desviación estándar para x , el número de lesiones por año sufridas por un niño en edad escolar.
- ¿Dentro de qué límites esperaría usted que caiga el número de lesiones por año?

3.59 Bacterias en muestras de agua Si una gota de agua se pone en un portaobjetos y se examina bajo un

microscopio, el número x de un tipo particular de bacteria presente se ha encontrado que tiene una distribución de probabilidad de Poisson. Suponga que la cantidad máxima permisible por espécimen de agua para este tipo de bacteria es cinco. Si la cantidad media de su suministro de agua es de dos y usted prueba una sola muestra, ¿es probable que la cantidad exceda la cantidad máxima permisible? Explique.

3.60 Brotes de *E. coli* Un brote de infecciones por *E. coli* ocurrió en agosto de 2010 en tres guarderías infantiles en el estado de Washington. Hubo ocho casos confirmados y seis sospechosos de *E. coli*, con más de 70 niños aguardando los resultados de sus exámenes.⁶ Se reportó que los brotes de infecciones por *E. coli* para 2009 fueron menos de 1 por 100 000, menos de los 2.5 por 100 000 reportados antes.⁷ Usando la tasa de 1 por 100 000, encuentre las siguientes probabilidades.

- ¿Cuál es la probabilidad de que se reporten cuando mucho dos brotes por 100 000 en todo Estados Unidos este año?
- ¿Cuál es la probabilidad de que se reporten más de tres brotes por 100 000 en todo Estados Unidos este año?

LA DISTRIBUCIÓN HIPERGEOMÉTRICA DE PROBABILIDAD

3.6

Supongamos que usted está seleccionando una muestra de elementos de una población y que registra si cada elemento posee o no cierta característica. Usted está registrando la típica información de “éxito” o “fracaso” que se encuentra en el experimento binomial. El estudio de la muestra del ejemplo 3.5, así como el muestreo para ver si hay defectos en el ejemplo 3.6, son ilustraciones prácticas de estas situaciones de muestreo.

Si el número de elementos de la población es grande respecto al número en la muestra (como en el ejemplo 3.30), la probabilidad de seleccionar un éxito en un solo intento es igual a la proporción p de éxitos en la población. Debido a que la población es grande respecto al tamaño muestral, esta probabilidad permanecerá constante (para todos los fines prácticos) de un intento a otro y el número x de éxitos en la muestra seguirá una distribución binomial de probabilidad. No obstante, si el número de elementos en la población es pequeño respecto al tamaño muestral ($n/N \geq .05$), la probabilidad de un éxito para un intento determinado depende de los resultados de intentos precedentes. Entonces el número x de éxitos sigue lo que se conoce como una **distribución hipergeométrica de probabilidad**.

Es fácil visualizar la **variable hipergeométrica aleatoria x** si se considera un tazón que contenga M esferas rojas y $N - M$ esferas blancas, para un total de N esferas en el tazón. Usted selecciona n esferas del tazón y registra x , el número de esferas rojas que vea. Si ahora define un “éxito” como una esfera roja, tendrá un ejemplo de la variable aleatoria x hipergeométrica.

La fórmula para calcular la probabilidad de exactamente k éxitos en n intentos se da a continuación.

LA DISTRIBUCIÓN HIPERGEOMÉTRICA DE PROBABILIDAD

Una población contiene M éxitos y $N - M$ fracasos. La probabilidad de exactamente k éxitos en una muestra aleatoria de tamaño n es

$$P(x = k) = \frac{C_k^M C_{n-k}^{N-M}}{C_n^N}$$

para valores de k que dependen de N , M y n con

$$C_n^N = \frac{N!}{n!(N-n)!}$$

La media y la varianza de una variable aleatoria hipergeométrica son muy semejantes a las de una variable aleatoria binomial con una corrección para el tamaño finito de población:

$$\mu = n \left(\frac{M}{N} \right)$$

$$\sigma^2 = n \left(\frac{M}{N} \right) \left(\frac{N-M}{N} \right) \left(\frac{N-n}{N-1} \right)$$

EJEMPLO

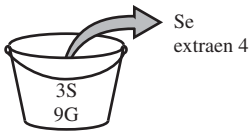
3.19

Una caja tiene 12 botellas de vino, 3 de las cuales contienen vino echado a perder. Se selecciona al azar una muestra de 4 botellas de la caja.

1. Encuentre la distribución de probabilidad para x , el número de botellas de vino echado a perder de la muestra.
2. ¿Cuáles son la media y la varianza de x ?

Solución Para este ejemplo, $N = 12$, $n = 4$, $M = 3$ y $(N - M) = 9$. Entonces

$$p(x) = \frac{C_x^3 C_{4-x}^9}{C_4^{12}}$$

TIP MI CONSEJO

S = V (vinagre)
G = B (buen estado)

1. Los valores posibles para x son 0, 1, 2 y 3, con probabilidades

$$p(0) = \frac{C_0^3 C_4^9}{C_4^{12}} = \frac{1(126)}{495} = .25$$

$$p(1) = \frac{C_1^3 C_3^9}{C_4^{12}} = \frac{3(84)}{495} = .51$$

$$p(2) = \frac{C_2^3 C_2^9}{C_4^{12}} = \frac{3(36)}{495} = .22$$

$$p(3) = \frac{C_3^3 C_1^9}{C_4^{12}} = \frac{1(9)}{495} = .02$$

2. La media está dada por

$$\mu = 4 \left(\frac{3}{12} \right) = 1$$

y la varianza es

$$\sigma^2 = 4 \left(\frac{3}{12} \right) \left(\frac{9}{12} \right) \left(\frac{12-4}{11} \right) = .5455$$

EJEMPLO

3.20

Un producto industrial particular se envía en lotes de 20. Es costoso hacer pruebas para determinar si un artículo está defectuoso; por tanto, el fabricante muestrea la producción en lugar de usar un plan de inspección del 100%. Un plan de muestreo construido para reducir al mínimo el número de artículos defectuosos, enviados a los clientes, exige muestrear cinco artículos de cada lote y rechazar el lote si se observa más de una pieza defectuosa. (Si el lote es rechazado, entonces se verifica cada artículo del lote.) Si un lote contiene cuatro artículos defectuosos, ¿cuál es la probabilidad de que sea aceptado?

Solución Sea x el número de artículos defectuosos en la muestra. Entonces $n = 20$, $M = 4$, $(N - M) = 16$ y $n = 5$. El lote será rechazado si $x = 2, 3$ o 4 . Entonces

$$\begin{aligned} P(\text{aceptar el lote}) &= P(x \leq 1) = p(0) + p(1) = \frac{C_0^4 C_5^{16}}{C_5^{20}} + \frac{C_1^4 C_4^{16}}{C_5^{20}} \\ &= \frac{\left(\frac{4!}{0!4!}\right)\left(\frac{16!}{5!11!}\right)}{\frac{20!}{5!15!}} + \frac{\left(\frac{4!}{1!3!}\right)\left(\frac{16!}{4!12!}\right)}{\frac{20!}{5!15!}} \\ &= \frac{91}{323} + \frac{455}{969} = .2817 + .4696 = .7513 \end{aligned}$$

3.6

EJERCICIOS

TÉCNICAS BÁSICAS

3.61 Evalúe estas probabilidades:

a. $\frac{C_1^2 C_1^1}{C_2^3}$ b. $\frac{C_0^4 C_2^2}{C_2^6}$ c. $\frac{C_2^2 C_1^2}{C_3^4}$

3.62 Sea x el número de éxitos observados en una muestra de $n = 4$ elementos seleccionados de una población de $N = 8$. Suponga que de los $N = 8$ elementos, 5 se consideran “éxitos”.

- Encuentre la probabilidad de observar todos los éxitos.
- Encuentre la probabilidad de observar un éxito.
- Encuentre la probabilidad de observar cuando mucho dos éxitos.

3.63 Evalúe estas probabilidades:

a. $\frac{C_1^3 C_1^2}{C_2^5}$ b. $\frac{C_2^4 C_1^3}{C_3^7}$ c. $\frac{C_4^5 C_0^3}{C_4^8}$

3.64 Sea x el número de éxitos observados en una muestra de $n = 5$ elementos seleccionados de $N = 10$. Suponga que, de los $N = 10$ elementos, 6 se consideran “éxitos”.

- Encuentre la probabilidad de no observar éxitos.
- Encuentre la probabilidad de observar al menos dos éxitos.

- Encuentre la probabilidad de observar exactamente dos éxitos.

3.65 Sea x una variable aleatoria hipergeométrica con $N = 15$, $n = 3$ y $M = 4$.

- Calcule $p(0)$, $p(1)$, $p(2)$ y $p(3)$.
- Construya el histograma de probabilidad para x .
- Use las fórmulas dadas en la sección 3.6 para calcular $\mu = E(x)$ y σ^2 .
- ¿Qué proporción de la población de mediciones cae en el intervalo $(\mu \pm 2\sigma)$? ¿En el intervalo $(\mu \pm 3\sigma)$? ¿Estos resultados concuerdan con los dados por el teorema de Chebyshev?

APLICACIONES

3.66 Chips de computadora defectuosos Una pieza de equipo electrónico contiene seis chips de computadora, dos de los cuales están defectuosos. Se seleccionan al azar tres chips de computadora para inspeccionarlos y se registra el número de los defectuosos. Encuentre la distribución de probabilidad para x , el número de chips de computadora defectuosos.

3.67 ¿Sesgo relacionado con el género? Una compañía tiene cinco solicitantes para dos puestos de trabajo: dos mujeres y tres hombres. Suponga que los cinco solicitantes están igualmente calificados y que no se da preferencia para elegir género alguno. Sea x igual al número de mujeres escogido para ocupar las dos posiciones.

- Escriba la fórmula para $p(x)$, la distribución de probabilidad de x .
- ¿Cuáles son la media y la varianza de esta distribución?
- Construya un histograma de probabilidad para x .

3.68 Diplomas de enseñanza En el sur de California, un creciente número de personas que buscan un diploma de enseñanza están eligiendo internados pagados en lugar de los tradicionales programas estudiantiles de enseñanza. Un grupo de ocho candidatos para tres posiciones locales de enseñanza estaba formado por cinco candidatos, que se habían inscrito en internados pagados y tres candidatos que se habían inscrito en programas tradicionales estudiantiles de enseñanza. Supongamos que los ocho candidatos están igualmente calificados para las posiciones. Represente con x el número de candidatos capacitados en un internado que son contratados para estas tres posiciones.

- ¿La x tiene una distribución binomial o una distribución hipergeométrica? Argumente su respuesta.
- Encuentre la probabilidad de que tres candidatos capacitados en internado sean contratados para estas posiciones.
- ¿Cuál es la probabilidad de que ninguno de los tres contratados sea capacitado en internado?
- Encuentre $P(x \leq 1)$.

3.69 Tratamiento a semillas Es frecuente que las semillas sean tratadas con un fungicida para protegerlas de ambientes mal drenados, húmedos. En un intento a pequeña escala, antes de un experimento a gran escala para determinar qué dilución del fungicida aplicar, se plantaron en suelo arcilloso cinco semillas tratadas y cinco no tratadas y se registró el número de plantas que brotaron de las semillas tratadas y de las no tratadas. Suponga que la dilución no fue eficaz y sólo brotaron cuatro plantas. Represente con x el número de plantas que germinaron de semillas tratadas.

- Encuentre la probabilidad de que $x = 4$.
- Encuentre $P(x \leq 3)$.
- Encuentre $P(2 \leq x \leq 3)$.

3.7

LA DISTRIBUCIÓN NORMAL DE PROBABILIDAD

Las distribuciones de probabilidad continua pueden tomar varias formas, pero un gran número de variables aleatorias observadas en la naturaleza poseen una distribución de frecuencia que tiene más o menos la forma de montículo, o bien, como diría un estadístico, es aproximadamente una distribución normal de probabilidad. La fórmula o función de densidad de probabilidad (f_{dp}) que genera esta distribución se muestra a continuación.

DISTRIBUCIÓN NORMAL DE PROBABILIDAD

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} \quad -\infty < x < \infty$$

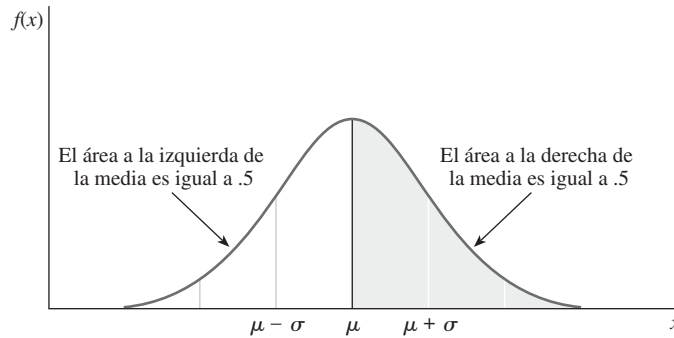
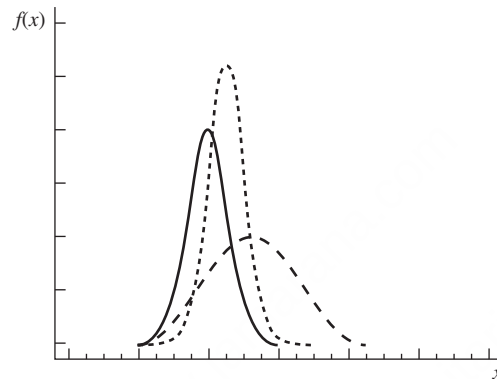
Los símbolos e y π son constantes matemáticas dadas en forma aproximada por 2.7183 y 3.1416, respectivamente; μ y σ ($\sigma > 0$) son parámetros que representan la media poblacional y desviación estándar, respectivamente.

La gráfica de una distribución normal de probabilidad con media μ y desviación estándar σ se muestran en la figura 3.13. La media μ localiza el *centro* de la distribución, y la distribución es *simétrica* alrededor de su media μ . Como el área total bajo la distribución normal de probabilidad es igual a 1, la simetría implica que el área a la derecha de μ es .5 y el área a la izquierda de μ es también .5.

La *forma* de la distribución está determinada por σ , la desviación estándar de la población. Como se puede observar en la figura 3.14, los valores grandes de σ reducen la altura de la curva y aumentan la dispersión; los valores pequeños de σ aumentan la altura de la curva y

FIGURA 3.13

Distribución normal de probabilidad

**FIGURA 3.14**Distribuciones normales de probabilidad que difieren con valores de μ y σ 

APPLET EN LÍNEA

Visualizar curvas normales

reducen la dispersión. La figura 3.14 muestra tres distribuciones normales de probabilidad con diferentes medias y desviaciones estándar. Observe las diferencias en forma y ubicación.

Raras veces se encuentra una variable con valores que sean infinitamente pequeños ($-\infty$) o infinitamente grandes ($+\infty$). Aun así, muchas variables aleatorias *positivas* (por ejemplo estaturas, pesos y tiempos) tienen distribuciones que son bien aproximadas por una distribución normal. De acuerdo con la regla empírica, casi todos los valores de una variable aleatoria normal se encuentran en el intervalo $\mu \pm 3\sigma$. Mientras los valores dentro de tres desviaciones estándar de la media sean *positivos*, la distribución normal da un buen modelo para describir los datos.

ÁREAS TABULADAS DE LA DISTRIBUCIÓN NORMAL DE PROBABILIDAD

3.8

Para hallar la probabilidad de que una variable aleatoria normal x se encuentre en el intervalo de a a b , necesitamos hallar el área bajo la curva normal entre los puntos a y b (véase figura 3.7). No obstante (véase la figura 3.14), hay un número infinitamente grande de distribuciones normales, uno para cada media y desviación estándar diferentes. Una tabla separada de áreas para cada una de estas curvas es obviamente impráctica; en cambio, usamos un procedimiento de estandarización que nos permite usar la misma tabla para todas las distribuciones normales.

La variable aleatoria normal estándar

Una variable aleatoria normal x está **estandarizada** al expresar su valor como el número de desviaciones estándar (σ) que se encuentran a la izquierda o a la derecha de su media μ . Éste es realmente sólo un cambio en las unidades de medida que usamos, ¡como si midiéramos en

pulgadas en lugar de centímetros! La variable aleatoria normal estandarizada, z , se define como

$$z = \frac{x - \mu}{\sigma}$$

TIP

MI CONSEJO

El área bajo la curva z es igual a 1

o bien, lo que es equivalente,

$$x = \mu + z\sigma$$

De la fórmula para z , podemos sacar estas conclusiones:

- Cuando x es menor que la media μ , el valor de z es negativo.
- Cuando x es mayor que la media μ , el valor de z es positivo.
- Cuando $x = \mu$, el valor de $z = 0$.

La distribución de probabilidad para z , ilustrada en la figura 3.15, se denomina **distribución normal estandarizada** porque su media es 0 y su desviación estándar es 1. Los valores de z del lado izquierdo de la curva son negativos, en tanto que los del lado derecho son positivos. El área bajo la curva normal estándar a la izquierda de un valor especificado de z , por ejemplo z_0 , es la probabilidad $P(z \leq z_0)$. Esta **área acumulativa** está registrada en la tabla 3 del Anexo y se muestra como el área sombreada en la figura 3.15. Una versión abreviada de la tabla 3 se da en la tabla 3.5. Observe que la tabla contiene valores positivos y negativos de z . La columna izquierda de la tabla da el valor de z correcto al décimo lugar; el segundo lugar decimal para z , correspondiente a las centenas, se da en el renglón superior.

FIGURA 3.15
Distribución normal estandarizada

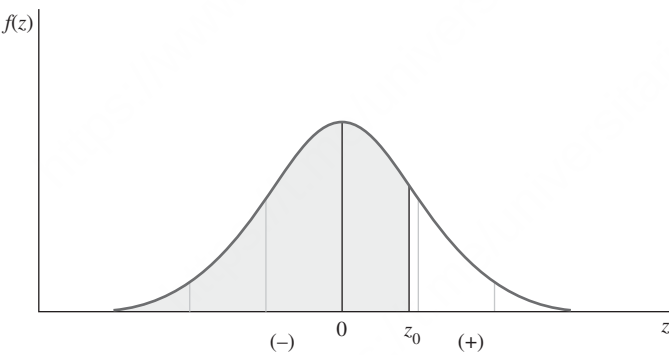


TABLA 3.5
Versión abreviada de la tabla 3 del Anexo
Tabla 3. Áreas bajo la curva normal

z	.00	.01	.02	.03	...	.09
-3.4	.0003	.0003	.0003	.0003		
-3.3	.0005	.0005	.0005	.0004		
-3.2	.0007	.0007	.0006	.0006		
-3.1	.0010	.0009	.0009	.0009		
-3.0	.0013	.0013	.0013	.0012	...	.0010
-2.9	.0019	.	.	.		
-2.8	.0026	.	.	.		
-2.7	.0035	.	.	.		
-2.6	.0047					
-2.5	.0062					
.	.					
.	.					
.	.					
-2.0	.0228					
.	.					
.	.					
.	.					

Tabla 3. Áreas bajo la curva normal (continúa)

z	.00	.01	.02	.03	.04	...	.09
0.0	.5000	.5040	.5080	.5120	.5160		
0.1	.5398	.5438	.5478	.5517	.5557		
0.2	.5793	.5832	.5871	.5910	.5948		
0.3	.6179	.6217	.6255	.6293	.6331		
0.4	.6554	.6591	.6628	.6664	.6700	...	.6879
0.5	.6915	.	.	.			
0.6	.7257	.	.	.			
0.7	.7580	.	.	.			
0.8	.7881						
0.9	.8159						
.	.						
.	.						
.	.						
2.0	.9772						

EJEMPLO**3.21**

Encuentre $P(z \leq 1.63)$. Esta probabilidad corresponde al área a la izquierda de un punto $z = 1.63$ desviaciones estándar a la derecha de la media (véase figura 3.16).

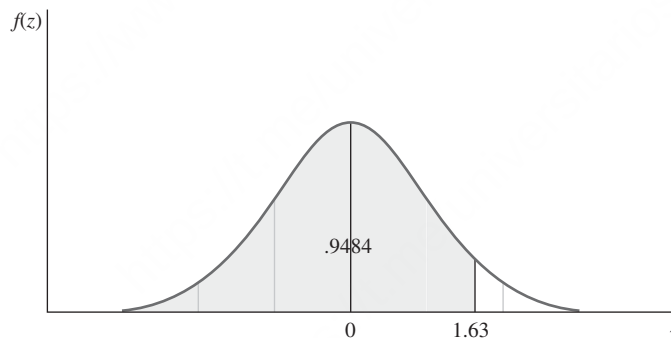
TIP**MI CONSEJO**

$$P(z \leq 1.63) = P(z < 1.63)$$

Solución El área está sombreada en la figura 3.16. Como la tabla 3 del Anexo da áreas bajo la curva normal a la izquierda de un valor especificado de z , sólo se necesita hallar el valor tabulado para $z = 1.63$. Baje por la columna izquierda de la tabla hasta $z = 1.6$ y en sentido horizontal en la parte superior de la tabla hasta la columna marcada .03. La intersección de esta combinación de renglón y columna da el área .9484, que es $P(z \leq 1.63)$.

FIGURA 3.16

Área bajo la curva normal estándar para el ejemplo 3.3



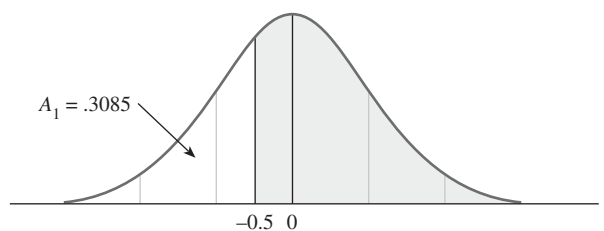
Las áreas a la izquierda de $z = 0$ se encuentran usando valores negativos de z .

EJEMPLO**3.22**

Encuentre $P(z \geq -.5)$. Esta probabilidad corresponde al área a la *derecha* de un punto $z = -.5$ de desviación estándar a la izquierda de la media (véase figura 3.17).

FIGURA 3.17

Área bajo la curva normal estándar para el ejemplo 3.4



Solución El área dada en la tabla 3 del Anexo es el área a la izquierda de un valor especificado de z . Haciendo un índice de $z = -0.5$ en la tabla 3, podemos hallar que el área A_1 a la izquierda de -0.5 es $.3085$.

Como el área bajo la curva es 1, encontramos

$$P(z \geq -0.5) = 1 - A_1 = 1 - .3085 = .6915.$$

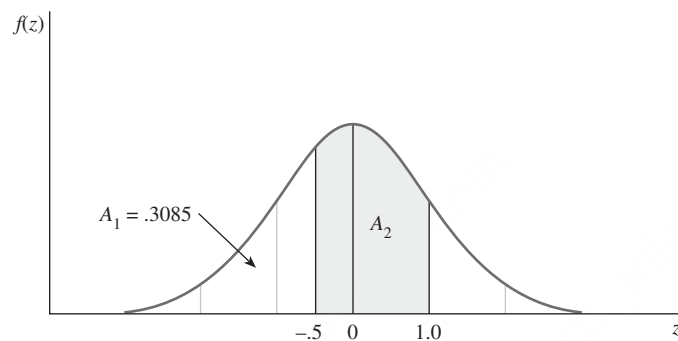
EJEMPLO

3.23

Encuentre $P(-0.5 \leq z \leq 1.0)$. Esta probabilidad es el área entre $z = -0.5$ y $z = 1.0$, como se muestra en la figura 3.18.

FIGURA 3.18

Área bajo la curva normal estándar para el ejemplo 3.22



Solución El área que se pide es el área sombreada A_2 en la figura 3.18. De la tabla 3 del Anexo, se puede hallar el área a la izquierda de $z = -0.5$ ($A_1 = .3085$) y el área a la izquierda de $z = 1.0$ ($A_1 + A_2 = .8413$). Para hallar el área marcada A_2 , restamos las dos entradas:

$$A_2 = (A_1 + A_2) - A_1 = .8413 - .3085 = .5328$$

Esto es, $P(-0.5 \leq z \leq 1.0) = .5328$.

**NECESITO SABER...**

Cómo utilizar la tabla 3 para calcular probabilidades bajo la curva normal estándar

- Para calcular el área a la izquierda de un valor z , encuentre el área directamente de la tabla 3.
- Para calcular el área a la derecha de un valor z , encuentre el área en la tabla 3 y réstela de 1.
- Para calcular el área entre dos valores de z , encuentre las dos áreas en la tabla 3 y reste un área de la otra.

EJEMPLO

3.24

Encuentre la probabilidad de que una variable aleatoria normalmente distribuida quede dentro de estos intervalos:

1. Una desviación estándar de su media
2. Dos desviaciones estándar de su media

Solución

- Como la variable aleatoria normal estándar z mide la distancia desde la media en unidades de desviaciones estándar, es necesario hallar

$$P(-1 \leq z \leq 1) = .8413 - .1587 = .6826$$

Recuerde que usted calcula el área entre dos valores z al restar las entradas tabuladas para los dos valores.

- Al igual que en la parte 1, $P(-2 \leq z \leq 2) = .9772 - .0228 = .9544$.

Estas probabilidades concuerdan con valores aproximados de 68% y 95% en la regla empírica del capítulo 1.

EJEMPLO

3.25

Encuentre el valor de z , llámelo z_0 , tal que .95 del área se encuentre a no más de $\pm z_0$ desviaciones estándar de la media.

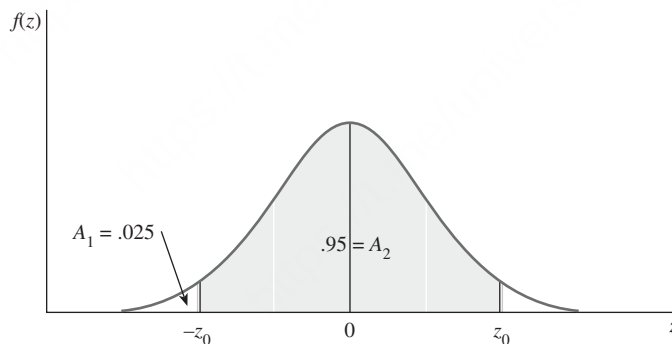
TIP**MI CONSEJO**

Conocemos el área.
Trabaje de adentro hacia afuera de la tabla

Solución El área sombreada de la figura 3.19 es el área que se encuentra a no más de $\pm z_0$ desviaciones estándar de la media, que necesita ser igual a .95. Las “áreas de cola” bajo la curva no están sombreadas y tienen un área combinada de $1 - .95 = .05$. Debido a la simetría de la curva normal, estas dos áreas de cola tienen la misma área, de modo que $A_1 = .05/2 = .025$ en la figura 3.19. Por tanto, el *área acumulativa* entera a la izquierda de $-z_0$ es igual a $A_1 = .025$. Esta área se encuentra en el interior de la tabla 3 del Anexo en el renglón correspondiente a $z = -1.9$ y la columna .06. En consecuencia, $-z_0 = -1.96$ o $z_0 = 1.96$. Observe que este resultado es muy cercano al valor aproximado, $z = 2$, que se usa en la regla empírica.

FIGURA 3.19

Área bajo la curva normal estándar para el ejemplo 3.24



Cálculo de probabilidades para una variable aleatoria normal general

La mayoría de las veces, las probabilidades en las que estamos interesados contienen x , una variable aleatoria normal con media μ y desviación estándar σ . Entonces se debe *estandarizar* el intervalo de interés escribiéndolo como el intervalo equivalente en términos de z , la variable aleatoria normal estándar. Una vez hecho esto, la probabilidad de interés es el área que se encuentra usando la *distribución estándar normal de probabilidad*.

EJEMPLO

3.26

Sea x una variable aleatoria normalmente distribuida con una media de 10 y una desviación estándar de 2. Encuentre la probabilidad de que x se encuentre entre 11 y 13.6.

Solución El intervalo de $x = 11$ a $x = 13.6$ se debe estandarizar usando la fórmula para z . Cuando $x = 11$,

$$z = \frac{x - \mu}{\sigma} = \frac{11 - 10}{2} = .5$$

y cuando $x = 13.6$,

$$z = \frac{x - \mu}{\sigma} = \frac{13.6 - 10}{2} = 1.8$$

TIP MI CONSEJO

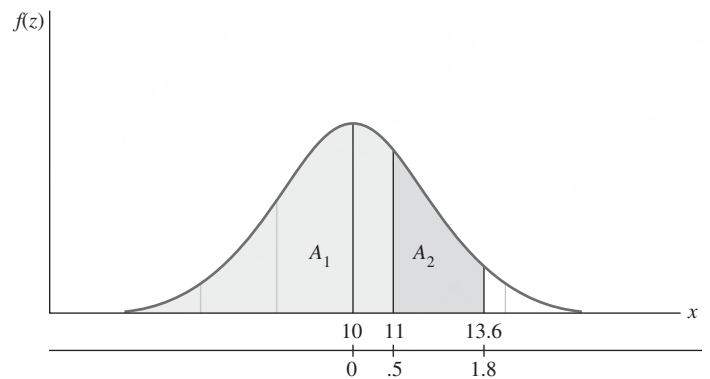
Siempre trace una figura; ¡esto ayuda!

La probabilidad deseada es, por tanto, $P(.5 \leq z \leq 1.8)$, el área que está entre $z = .5$ y $z = 1.8$, como se muestra en la figura 3.20. De la tabla 3 del Anexo, se encuentra que el área a la izquierda de $z = .5$ es .6915, y el área a la izquierda de $z = 1.8$ es .9641. La probabilidad deseada es la diferencia entre estas dos probabilidades, es decir,

$$P(.5 \leq z \leq 1.8) = .9641 - .6915 = .2726$$

FIGURA 3.20

Área bajo la curva normal estándar para el ejemplo 3.25

**MI** APLET EN LÍNEA

¿Distribuciones de probabilidad normales y probabilidades normales y puntuaciones z ?

EJEMPLO

3.27

Estudios realizados demuestran que el uso de gasolina para autos compactos vendidos en Estados Unidos está normalmente distribuido, con una media de 35.5 millas por galón (mpg) y una desviación estándar de 4.5 mpg. ¿Qué porcentaje de compactos recorre 40 mpg o más?

Solución La proporción de compactos que recorren 40 mpg o más está dada por el área sombreada en la figura 3.21. Para resolver este problema, primero se debe hallar el valor z correspondiente a $x = 40$ calculando

$$z = \frac{x - \mu}{\sigma} = \frac{40 - 35.5}{4.5} = 1.0$$

El área A_1 a la izquierda de $z = 1.0$, es .8413 (de la tabla 3 del Anexo). Entonces la proporción de compactos que recorren 40 mpg o más es igual a

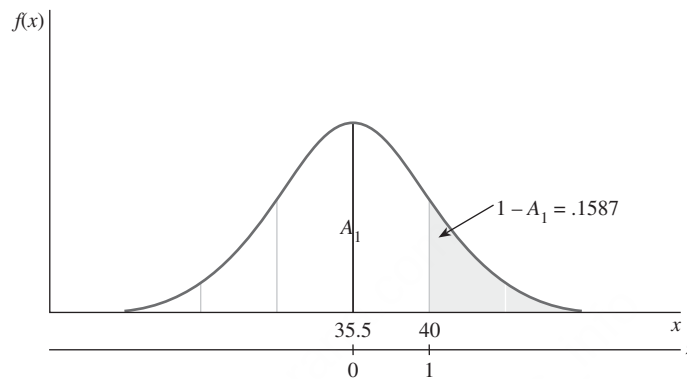
$$P(x \geq 40) = 1 - P(z < 1) = 1 - .8413 = .1587$$

El porcentaje que rebasa los 40 mpg es

$$100(.1587) = 15.87\%$$

FIGURA 3.21

Área bajo la curva normal estándar para el ejemplo 3.26



EJEMPLO

3.28

Consulte el ejemplo 3.27. En tiempos de escasez de recursos energéticos, una ventaja competitiva se da a un fabricante de automóviles que puede producir un auto que tiene una economía de consumo de combustible considerablemente mejor que los autos de los competidores. Si un fabricante desea desarrollar un auto compacto que supere 95% de los compactos actuales en economía de combustible, ¿cuál debe ser el porcentaje de uso de gasolina para el nuevo auto?

Solución El porcentaje x de uso de gasolina tiene una distribución normal con una media de 35.5 mpg y una desviación estándar de 4.5 mpg. Usted necesita hallar un valor particular, por ejemplo x_0 , tal que

$$P(x \leq x_0) = .95$$

Éste es el 95o. percentil de la distribución del porcentaje x de uso de gasolina. Como la única información que tenemos acerca de las probabilidades normales es en términos de la variable z aleatoria normal estándar, empiece por estandarizar el valor de x_0 :

$$z_0 = \frac{x_0 - 35.5}{4.5}$$

Como el valor de z_0 corresponde a x_0 , también debe tener un área de .95 a su izquierda, como se muestra en la figura 3.22. Si usted ve al interior de la tabla 3 del Anexo, encontrará que el área .9500 está exactamente a la mitad entre las áreas para $z = 1.64$ y $z = 1.65$. Por tanto, z_0 debe estar exactamente a la mitad entre 1.64 y 1.65, es decir

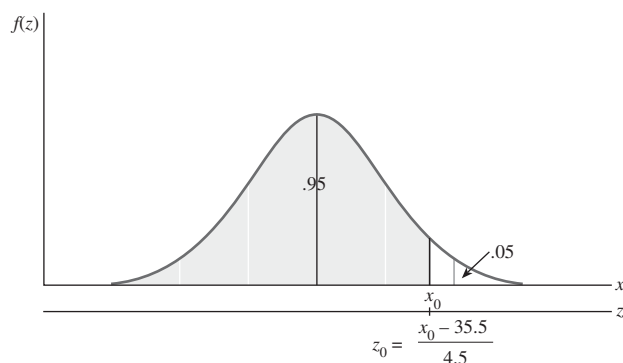
$$z_0 = \frac{x_0 - 35.5}{4.5} = 1.645$$

Al despejar x_0 resulta

$$x_0 = \mu + z_0\sigma = 35.5 + (1.645)(4.5) = 42.9$$

FIGURA 3.22

Área bajo la curva normal estándar para el ejemplo 3.27



El nuevo auto compacto del fabricante debe recorrer, por tanto, 42.9 mpg para superar 95% de los autos compactos actualmente disponibles en el mercado en Estados Unidos.

3.8 EJERCICIOS

TÉCNICAS BÁSICAS

3.70 Considere una variable aleatoria normal estándar con $\mu = 0$ y desviación estándar $\sigma = 1$. Use la tabla 3 para encontrar las siguientes probabilidades:

- a. $P(z < 2)$ b. $P(z > 1.16)$
c. $P(-2.33 < z < 2.33)$ d. $P(z < 1.88)$

3.71 Encuentre estas probabilidades asociadas con la variable aleatoria normal estándar z :

- a. $P(z > 5)$ b. $P(-3 < z < 3)$
c. $P(z < 2.81)$ d. $P(z > 2.81)$

3.72 Calcule el área bajo la curva normal estándar a la izquierda de estos valores:

- a. $z = 1.6$ b. $z = 1.83$
c. $z = .90$ d. $z = 4.18$

3.73 Calcule el área bajo la curva normal estándar entre estos valores:

- a. $z = -1.4$ y $z = 1.4$ b. $z = -3.0$ y $z = 3.0$

3.74 Encuentre las siguientes probabilidades para la variable aleatoria normal estándar z :

- a. $P(-1.43 < z < .68)$ b. $P(.58 < z < 1.74)$
c. $P(-1.55 < z < -.44)$ d. $P(z > 1.34)$
e. $P(z < -4.32)$

3.75 Encuentre las siguientes probabilidades para la variable aleatoria normal estándar z :

- a. $P(z < 2.33)$ b. $P(z < 1.645)$
c. $P(z > 1.96)$ d. $P(-2.58 < z < 2.58)$

3.76 a. Encuentre una z_0 tal que $P(z > z_0) = .025$.

b. Encuentre una z_0 tal que $P(z < z_0) = .9251$.

3.77 Encuentre una z_0 tal que $P(-z_0 < z < z_0) = .8262$.

3.78 a. Encuentre una z_0 que tenga área .9505 a su izquierda.

b. Encuentre una z_0 que tenga área .05 a su izquierda.

3.79 a. Encuentre una z_0 tal que $P(-z_0 < z < z_0) = .90$.

b. Encuentre una z_0 tal que $P(-z_0 < z < z_0) = .99$.

3.80 Encuentre los siguientes *percentiles* para la variable aleatoria normal estándar z :

- a. 90o. percentil b. 95o. percentil
c. 98o. percentil d. 99o. percentil

3.81 Una variable aleatoria normal x tiene media $\mu = 10$ y desviación estándar $\sigma = 2$. Encuentre la probabilidad asociada con cada uno de los siguientes intervalos.

- a. $x > 13.5$ b. $x < 8.2$ c. $9.4 < x < 10.6$

3.82 Una variable aleatoria normal x tiene media $\mu = 1.2$ y desviación estándar $\sigma = .15$. Encuentre la probabilidad asociada con cada uno de los siguientes intervalos.

- a. $1.00 < x < 1.10$ b. $x > 1.38$
c. $1.35 < x < 1.50$

3.83 Una variable aleatoria normal x tiene una media μ desconocida y desviación estándar $\sigma = 2$. Si la probabilidad que x exceda de 7.5 es .8023, encuentre μ .

3.84 Una variable aleatoria normal x tiene media de 35 y desviación estándar 10. Encuentre un valor de x que tenga

área .01 a su derecha. Éste es el 99o. percentil de su distribución normal.

3.85 Una variable aleatoria normal x tiene media de 50 y desviación estándar 15. ¿Sería poco común ver el valor $x = 0$? Explique su respuesta.

3.86 Una variable aleatoria normal x tiene una media y desviación estándar desconocidas. La probabilidad de que x exceda de 4 es .9772 y la probabilidad de que x exceda de 5 es .9332. Encuentre μ y σ .

APLICACIONES

3.87 Flujo sanguíneo cerebral El flujo sanguíneo cerebral (FSC) en los cerebros de personas sanas está normalmente distribuido con una media de 74 y una desviación estándar de 16.

- ¿Qué proporción de personas sanas tendrán lecturas de FSC entre 60 y 80?
- ¿Qué proporción de personas sanas tendrán lecturas de FSC arriba de 100?
- Si una persona tiene una lectura FSC debajo de 40, es clasificada como en riesgo de sufrir un accidente cerebrovascular. ¿A qué proporción de personas sanas se les diagnosticará erróneamente como “en riesgo”?

3.88 Distancias de frenado Para un auto que corre a 30 millas por hora (mph), la distancia necesaria de frenado hasta detenerse por completo está normalmente distribuida con una media de 50 pies y una desviación estándar de 8 pies. Suponga que usted está viajando a 30 mph en una zona residencial y un auto se atraviesa en forma abrupta en su camino, a una distancia de 60 pies.

- Si usted aplica los frenos, ¿cuál es la probabilidad de que frene hasta detenerse a los 40 pies o menos? ¿Y en no más de 50?
- Si la única forma de evitar una colisión es frenar hasta detenerse por completo, ¿cuál es la probabilidad de que evite la colisión?

3.89 Capacidades en elevadores Supongamos que usted debe establecer reglas respecto al número máximo de personas que pueden ocupar un elevador. Un estudio indica que si ocho personas ocupan el elevador, la distribución de probabilidad del peso total de las ocho personas tiene una distribución aproximadamente normal con una media igual a 1200 libras y una desviación estándar de 99 libras. ¿Cuál es la probabilidad de que el peso total de ocho personas exceda de 1300 libras? ¿Y de 1500 libras?

3.90 Una mina de fosfato La descarga de sólidos suspendidos desde una mina de fosfato está normalmente distribuida, con una descarga media diaria de 27 miligramos por litro (mg/l) y una desviación estándar de

14 mg/l. ¿Qué proporción de días excederá de 50 mg/l la descarga diaria?

3.91 Girasoles Un experimentador que publicó en *Annals of Botany* investigó si los diámetros de tallos del girasol dicotiledónea cambiaría, dependiendo de si se dejó a la planta balancearse libremente en el viento o se le había sostenido artificialmente.⁸ Suponga que los diámetros de tallos no soportados en la base, de una especie particular de girasol, tienen una distribución normal con un diámetro promedio de 35 milímetros (mm) y una desviación estándar de 3 mm.

- ¿Cuál es la probabilidad de que una planta de girasol tenga un diámetro de base de más de 40 mm?
- Si dos plantas de girasol se seleccionan al azar, ¿cuál es la probabilidad de que ambas plantas tengan un diámetro de base de más de 40 mm?
- ¿Dentro de qué límites esperaría usted que se encuentren los diámetros de base, con probabilidad .95?
- ¿Qué diámetro representa el 90o. percentil de la distribución de diámetros?

3.92 Frecuencia respiratoria El número de veces x que un humano adulto respira por minuto, cuando está en reposo, tiene una distribución de probabilidad aproximadamente normal, con la media igual a 16 y la desviación estándar igual a 4. Si una persona se selecciona al azar y se registra el número x de respiraciones por minuto cuando está en reposo, ¿cuál es la probabilidad de que x exceda de 22?

3.93 Pronósticos económicos Un método para obtener pronósticos económicos es usar una propuesta de consensos. Se obtiene un pronóstico de cada uno de un número grande de analistas y el promedio de estos pronósticos individuales es el pronóstico de consenso. Suponga que los pronósticos individuales de la tasa de interés preferente de enero de 2013, hechos por analistas económicos, están normalmente distribuidos en forma aproximada con la media igual a 4.75% y una desviación estándar igual a 0.2%. Si se selecciona al azar un solo analista de este grupo, ¿cuál es la probabilidad de que el pronóstico del analista de la tasa preferente tome estos valores?

- Rebase de 4.25%.
- Sea menor que 4.375%.

3.94 Bacterias en agua potable Suponga que los números de un tipo particular de bacterias, en muestras de 1 mililitro (ml) de agua potable, tienden a estar normalmente distribuidos en forma aproximada con una media de 85 y una desviación estándar de 9. ¿Cuál es la probabilidad de que una muestra determinada de 1 ml contenga más de 100 bacterias?

3.95 Carga de granos Un cargador de granos se puede ajustar para descargar granos en cantidades que están normalmente distribuidas, con media de μ bushels y desviación estándar de 25.7 bushels. Si una compañía desea usar el cargador para llenar contenedores de 2000 bushels de grano y llenar de más sólo un contenedor en 100, ¿a qué valor de μ debe la compañía ajustar el cargador?

3.96 ¿Cuántas palabras? Un editor ha descubierto que los números de palabras contenidos en un nuevo manuscrito están normalmente distribuidos, con una media igual a 20 000 palabras más de las especificadas en el contrato del autor y una desviación estándar de 10 000 palabras. Si el editor desea estar casi seguro (digamos con una probabilidad de .95) que el manuscrito tenga menos de 100 000 palabras, ¿qué número de palabras debe especificar el editor en el contrato?

3.97 Compras en un centro comercial Una encuesta dice que más del doble de compradores salen de compras los fines de semana que durante la semana.⁹ No sólo eso, porque esos compradores ¡también gastan más dinero en sus compras en sábados y domingos! Suponga que la cantidad de dinero gastada en centros comerciales, entre las 4 p.m. y las 6 p.m. los domingos tiene una distribución normal con media de \$85 y una desviación estándar de

\$20. Se selecciona al azar un comprador un domingo entre las 4 p.m. y las 6 p.m. y se le pregunta sobre su forma de gastar.

- ¿Cuál es la probabilidad de que él haya gastado más de \$95 en el centro comercial?
- ¿Cuál es la probabilidad de que él haya gastado entre \$95 y \$115 en el centro comercial?
- Si dos compradores se seleccionan al azar, ¿cuál es la probabilidad de que ambos compradores hayan gastado más de \$115 en el centro comercial?

3.98 Frecuencia de pulsaciones ¿Cuál es una frecuencia de pulsaciones *normal*? Eso depende de varios factores. Las frecuencias de pulsaciones entre 60 y 100 pulsaciones por minuto se consideran normales para niños mayores de 10 años de edad y en adultos.¹⁰ Suponga que estas frecuencias de pulsaciones están distribuidas normalmente en forma aproximada, con una media de 78 y una desviación estándar de 12.

- ¿Qué proporción de adultos tendrá frecuencias de pulsaciones entre 60 y 100?
- ¿Cuál es el 95o. percentil para las frecuencias de pulsaciones de adultos?
- ¿Una frecuencia de pulsaciones de 110 sería considerada poco común? Explique

LA APROXIMACIÓN DE LA DISTRIBUCIÓN DE PROBABILIDAD BINOMIAL A LA NORMAL

3.9

Aprendimos tres formas de calcular probabilidades para la variable binomial aleatoria x :

- Con el uso de la fórmula binomial, $P(x = k) = C_k^n p^k q^{n-k}$
- Con el uso de las tablas acumulativas binomiales
- Con el uso de *MS Excel* y *MINITAB*

La fórmula binomial produce cálculos largos y las tablas están disponibles sólo para ciertos valores de n y p . Hay otra opción cuando $np < 7$; las probabilidades de Poisson se pueden usar para aproximar $P(x = k)$. Cuando esta aproximación *no funciona* y n es grande, la distribución normal de probabilidad da otra aproximación para probabilidades binomiales.

LA APROXIMACIÓN DE LA DISTRIBUCIÓN DE PROBABILIDAD BINOMIAL A LA NORMAL

Sea x la variable binomial aleatoria con n intentos y probabilidad p de éxito. La distribución de probabilidad de x se aproxima mediante el uso de una curva normal con

$$\mu = np \quad y \quad \sigma = \sqrt{npq}$$

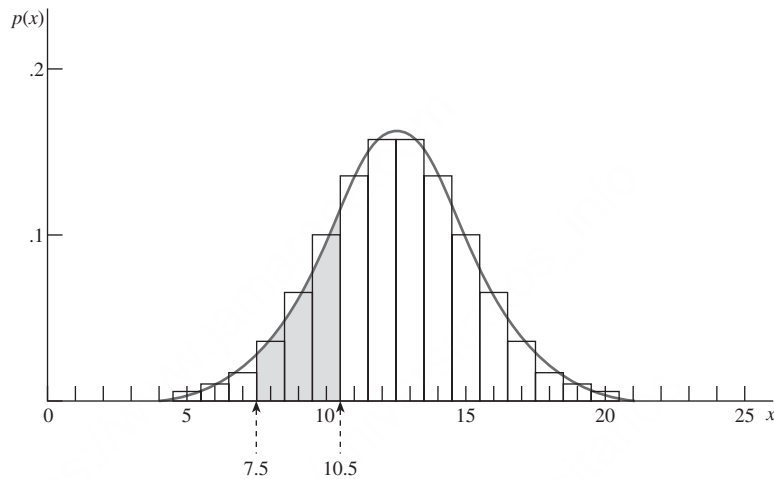
Esta aproximación es adecuada mientras n sea grande y p no esté demasiado cerca de 0 o 1.

Como la distribución normal es continua, el área bajo la curva en cualquier punto individual es igual a 0. Recuerde que este resultado se aplica sólo a variables aleatorias continuas. Como la variable binomial aleatoria x es una variable aleatoria discreta, la probabilidad de que x tome algún valor específico, por ejemplo $x = 11$, no necesariamente será igual a 0.

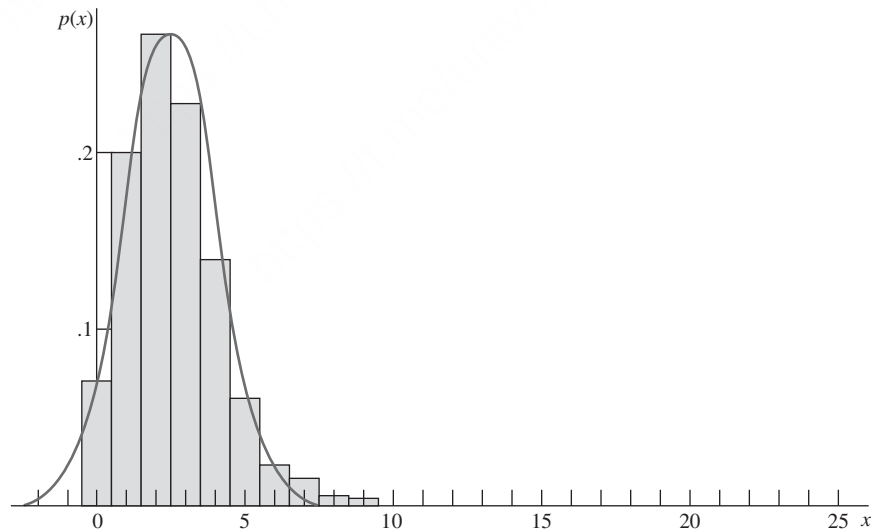
Las figuras 3.23 y 3.24 muestran los histogramas binomiales de probabilidad para $n = 25$ con $p = .5$ y $p = .1$, respectivamente. La distribución en la figura 3.21 es exactamente simétrica. Si usted superpone una curva normal con la misma media, $\mu = np$, y la misma desviación estándar, $\sigma = \sqrt{npq}$, sobre la parte superior de las barras, “se ajusta” muy bien; esto es, las áreas bajo la curva son casi iguales a las áreas bajo las barras. No obstante, cuando la probabilidad de éxito, p , es pequeña y la distribución está sesgada, como se muestra en la figura 3.24, la curva normal simétrica ya no ajusta muy bien. Si trata de usar las áreas de la curva normal para aproximar el área bajo las barras, la aproximación no será muy buena.

FIGURA 3.23

Distribución binomial de probabilidad para $n = 25$ y $p = .5$ y la distribución normal de aproximación con $\mu = 12.5$ y $\sigma = 2.5$

**FIGURA 3.24**

Distribución de probabilidad binomial y la distribución normal de aproximación para $n = 25$ y $p = .1$

**EJEMPLO****3.29**

Use la curva normal para aproximar la probabilidad de que $x = 8, 9$ o 10 para una variable aleatoria binomial con $n = 25$ y $p = .5$. Compare esta aproximación a la probabilidad binomial exacta.

Solución Se puede hallar la probabilidad binomial exacta para este ejemplo porque hay tablas binomiales acumulativas para $n = 25$. De la tabla 1 del Anexo,

$$P(x = 8, 9, \text{ o } 10) = P(x \leq 10) - P(x \leq 7) = .212 - .022 = .190$$

Para usar la aproximación normal, primero encuentre la media y desviación estándar apropiadas para la curva normal:

$$\mu = np = 25(.5) = 12.5$$

$$\sigma = \sqrt{npq} = \sqrt{25(.5)(.5)} = 2.5$$

TIP MI CONSEJO

Sólo use la corrección de continuidad ¡si x tiene una distribución **binomial**!

La probabilidad que usted necesita corresponde al área de los tres rectángulos que se encuentran en $x = 8, 9$ y 10 . El área equivalente bajo la curva normal se encuentra entre $x = 7.5$ (el lado inferior del rectángulo para $x = 8$) y $x = 10.5$ (el lado superior del rectángulo para $x = 10$). Esta área está sombreada en la figura 3.23.

Para hallar la probabilidad normal, siga los procedimientos de la sección 3.3. Primero estandarice el punto extremo de cada intervalo:

$$z = \frac{x - \mu}{\sigma} = \frac{7.5 - 12.5}{2.5} = -2.0$$

$$z = \frac{x - \mu}{\sigma} = \frac{10.5 - 12.5}{2.5} = -.8$$

A continuación la probabilidad aproximada (sombreada en la figura 3.23) se encuentra de la tabla 3 del Anexo:

$$P(-2.0 < z < -.8) = .2119 - .0228 = .1891$$

Se puede comparar la aproximación, .1891, con la probabilidad real, .190. ¡Son muy cercanas!

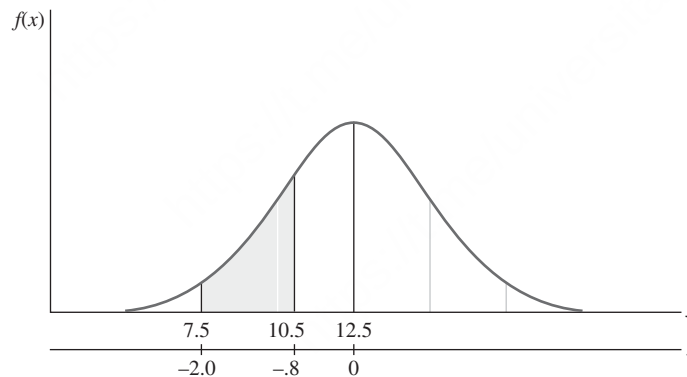
FIGURA 3.25

Área bajo la curva normal para el ejemplo 3.28



APPLET EN LÍNEA

Aproximación normal a probabilidades binomiales



Usted debe tener cuidado de no excluir la mitad de los dos rectángulos de probabilidad extremos, cuando use la aproximación normal a la distribución binomial de probabilidad. Este ajuste, llamado **corrección de continuidad**, ayuda a considerar el hecho de que usted se aproxima a una *variable aleatoria discreta* con una *continua*. ¡Si olvida la corrección, su aproximación no será muy buena! Utilice esta corrección sólo para *probabilidades binomiales*; no trate de usarla cuando la variable aleatoria ya es continua, por ejemplo una estatura o un peso.

¿Cómo saber cuándo es apropiado usar la aproximación normal a probabilidades binomiales? La aproximación normal funciona bien cuando el histograma binomial es casi simétrico.

Esto ocurre cuando la distribución binomial no está “agrupada” cerca de 0 o n , es decir, cuando se puede dispersar al menos dos desviaciones estándar desde su media sin exceder sus límites, 0 y n . Usando este criterio, se puede deducir esta sencilla regla práctica:

REGLA PRÁCTICA

La aproximación normal a las probabilidades binomiales será adecuada si

$$np > 5 \quad \text{y} \quad nq > 5$$



NECESITO SABER...

Cómo calcular probabilidades binomiales usando la aproximación normal

- Encuentre los valores necesarios de n y p . Calcule $\mu = np$ y $\sigma = \sqrt{npq}$.
- Escriba la probabilidad que necesite en términos de x y localice el área apropiada en la curva.
- Corrija el valor de x en $\pm .5$ para incluir todo el bloque de probabilidad para ese valor. Ésta es la *corrección de continuidad*.
- Convierta los valores x necesarios a valores z usando

$$z = \frac{x \pm .5 - np}{\sqrt{npq}}$$

- Use la tabla 3 del Anexo para calcular la probabilidad aproximada.

EJEMPLO

3.30

La confiabilidad de un fusible eléctrico es la probabilidad de que un fusible, elegido al azar de la producción, funcione bajo sus condiciones de diseño. Se probó una muestra aleatoria de 1000 fusibles y se observaron $x = 27$ defectuosos. Calcule la probabilidad aproximada de observar 27 o más defectuosos, suponiendo que la confiabilidad de un fusible es .98.

Solución La probabilidad de observar uno defectuoso cuando un solo fusible se prueba es $p = .02$, dado que la confiabilidad del fusible es .98. Entonces

$$\begin{aligned}\mu &= np = 1000(.02) = 20 \\ \sigma &= \sqrt{npq} = \sqrt{1000(.02)(.98)} = 4.43\end{aligned}$$

La probabilidad de 27 o más fusibles defectuosos, dada $n = 1000$, es

$$P(x \geq 27) = p(27) + p(28) + p(29) + \cdots + p(999) + p(1000)$$

Es apropiado usar la aproximación normal a la probabilidad binomial porque

$$np = 1000(.02) = 20 \quad \text{y} \quad nq = 1000(.98) = 980$$

TIP MI CONSEJO

Si np y nq son mayores que 5, se puede usar la aproximación normal

son mayores que 5. El área normal empleada para aproximar $P(x \geq 27)$ es el área bajo la curva normal a la derecha de 26.5, de modo que todo el rectángulo para $x = 27$ está incluido. Entonces, el valor z correspondiente a $x = 26.5$ es

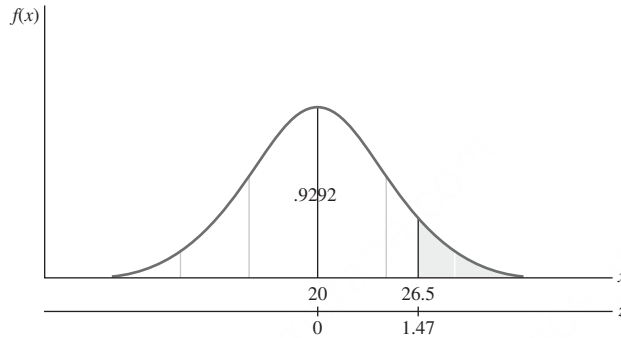
$$z = \frac{x - \mu}{\sigma} = \frac{26.5 - 20}{4.43} = \frac{6.5}{4.43} = 1.47$$

y el área a la izquierda de $z = 1.47$ es .9292, como se muestra en la figura 3.26. Como el área total bajo la curva es 1, usted tiene

$$P(x \geq 27) \approx P(z \geq 1.47) = 1 - .9292 = .0708$$

FIGURA 3.26

Aproximación normal a la binomial para el ejemplo 3.30



EJEMPLO

3.31

Una productora de cervezas estaba completamente segura de que su marca tenía 10% de participación en el mercado. En un estudio de mercado que comprendía 2500 consumidores de cerveza, $x = 211$ expresaron una preferencia por su marca. Si la cifra de 10% es correcta, encuentre la probabilidad de observar 211 o menos consumidores que prefieren la marca de cerveza de esta productora.

Solución Si la productora tiene razón, entonces la probabilidad de que un consumidor prefiera su marca de cerveza es $p = .10$. Entonces

$$\mu = np = 2500(.10) = 250$$

$$\sigma = \sqrt{npq} = \sqrt{2500(.10)(.90)} = 15$$

La probabilidad de observar 211 o menos que prefieran esta marca es

$$P(x \leq 211) = p(0) + p(1) + \cdots + p(210) + p(211)$$

La aproximación normal a esta probabilidad es el área a la izquierda de 211.5 bajo una curva normal con una media de 250 y una desviación estándar de 15. Primero calculamos

$$z = \frac{x - \mu}{\sigma} = \frac{211.5 - 250}{15} = -2.57$$

Entonces

$$P(x \leq 211) \approx P(z < -2.57) = .0051$$

La probabilidad de observar un valor de muestra de 211 o menos cuando $p = .10$ es tan pequeña que se puede concluir que ha ocurrido una de dos cosas: o se ha observado una muestra poco común aun cuando en realidad $p = .10$, o bien, la muestra refleja que el valor real de p es menor que .10 y quizá más cercano a la proporción muestral observada, $211/2500 = .08$.

3.9

EJERCICIOS

TÉCNICAS BÁSICAS

3.99 Considere una variable aleatoria binomial x con $n = 25$ y $p = .6$.

- ¿En este caso puede usarse la aproximación normal para aproximar probabilidades? ¿Por qué sí o por qué no?
- ¿Cuáles son la media y la desviación estándar de x ?
- Usando la corrección para continuidad, aproxime $P(x > 9)$.

3.100 Considere una variable aleatoria binomial x con $n = 45$ y $p = .05$.

- ¿Tanto np como nq son más grandes que 5?
- Con base en su respuesta de la parte **a**, ¿se puede usar la aproximación normal para aproximar las probabilidades binomiales asociadas con x ? De no ser así, ¿hay otra aproximación posible que pudiéramos usar?

3.101 Sea x la variable aleatoria binomial con $n = 25$ y $p = .3$.

- ¿La aproximación normal es apropiada para esta variable aleatoria binomial?
- Encuentre la media y desviación estándar para x .
- Use la aproximación normal para hallar $P(6 \leq x \leq 9)$.
- Use la tabla 1 del Anexo para hallar la probabilidad exacta $P(6 \leq x \leq 9)$. Compare los resultados de las partes **c** y **d**. ¿Qué tan cercana es la aproximación hecha por usted?

3.102 Sea x la variable aleatoria binomial con $n = 15$ y $p = .5$.

- ¿La aproximación normal es apropiada?
- Encuentre $P(x \geq 6)$ usando la aproximación normal.
- Encuentre $P(x > 6)$ usando la aproximación normal.
- Encuentre las probabilidades exactas de las partes **b** y **c**, y compare usted éstas con sus aproximaciones.

3.103 Sea x la variable aleatoria binomial con $n = 100$ y $p = .2$. Encuentre aproximaciones a estas probabilidades:

- $P(x > 22)$
- $P(x \geq 22)$
- $P(20 < x < 25)$
- $P(x \leq 25)$

3.104 Sea x la variable aleatoria binomial con $n = 25$ y $p = .2$.

- Use la tabla 1 del Anexo para calcular $P(4 \leq x \leq 6)$.
- Encuentre μ y σ para la distribución binomial de probabilidad y use la distribución normal para aproximar la probabilidad $P(4 \leq x \leq 6)$. Observe que

este valor es una buena aproximación al valor exacto de $P(4 \leq x \leq 6)$ aun cuando $np = 5$.

3.105 Suponga que una variable aleatoria x tiene una distribución binomial correspondiente a $n = 20$ y $p = .30$. Use la tabla 1 del Anexo para calcular estas probabilidades:

- $P(x = 5)$
- $P(x \geq 7)$

3.106 Consulte el ejercicio 3.50. Use la aproximación normal para calcular $P(x = 5)$ y $P(x \geq 7)$. Compare con los valores exactos obtenidos de la tabla 1 del Anexo.

3.107 Considere un experimento binomial con $n = 20$ y $p = .4$. Calcule $P(x \geq 10)$ usando cada uno de estos métodos:

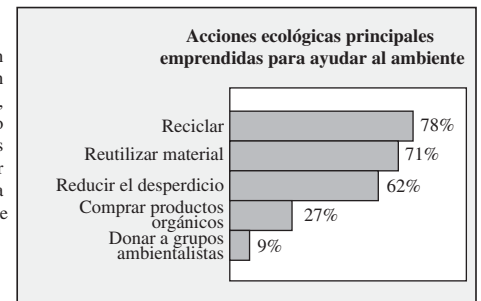
- Tabla 1 del Anexo.
- La aproximación normal a la distribución binomial de probabilidad.

3.108 Encuentre la aproximación normal a $P(355 \leq x \leq 360)$ para una distribución binomial de probabilidad con $n = 400$ y $p = .9$.

APLICACIONES

3.109 ¿Cómo puedo ayudar? ¿Está usted ayudando a salvar el ambiente? Un estudio encontró que alrededor de 78% de los habitantes de un municipio creen que reciclar la basura hace la mayor diferencia en la protección del ambiente.

Los que respondieron la encuesta dijeron que reciclar la basura, reducir el desperdicio y reutilizar materiales hace la mayor diferencia para la protección del ambiente



Suponga que a una muestra aleatoria de $n = 50$ adultos se le realizó una encuesta y se le preguntó si creían que el reciclaje hacía la mayor diferencia en la protección de nuestro ambiente. Supongamos que la cifra de 78% es, de hecho, correcta. ¿Cuáles son las probabilidades para los siguientes eventos?

- ¿Menos de 30 individuos creen que reciclar hace la mayor diferencia?

- b. ¿Más de 40 individuos creen que reciclar hace la mayor diferencia?
- c. ¿Más de 10 individuos creen que reciclar *no* hace la mayor diferencia?

3.110 Defectos genéticos Datos recolectados durante un largo periodo demuestran que se presenta un defecto genético particular en 1 de cada 1000 niños. Los registros de una clínica médica presentan $x = 60$ niños con el defecto en un total de 50 mil examinados. Si los 50 000 niños fueran una muestra aleatoria de la población de niños representada por registros del pasado, ¿cuál es la probabilidad de observar un valor de x igual a 60 o más? ¿Diría usted que la observación de $x = 60$ niños con defectos genéticos representa un evento raro?

3.111 No presentada Es frecuente que líneas aéreas y hoteles concedan reservaciones que rebasan la capacidad, para reducir al mínimo las pérdidas debidas a las que no se concretan. Suponga que los registros de un hotel indican que, en promedio, 10% de sus prospectos de huéspedes no reclaman su reservación. Si el hotel acepta 215 reservaciones y hay sólo 200 cuartos en el hotel, ¿cuál es la probabilidad de que todos los huéspedes que lleguen a pedir un cuarto lo reciban?

3.112 Cáncer de pulmones La compilación de grandes masas de datos sobre cáncer de pulmones muestra que alrededor de 1 de cada 40 adultos adquiere la enfermedad. Se sabe que los trabajadores de cierta ocupación trabajan en un ambiente de aire contaminado que puede causar un mayor porcentaje de cáncer de pulmón. Una muestra aleatoria de $n = 400$ trabajadores revela 19 con casos

identificables de cáncer de pulmón. ¿Estos datos dan suficiente evidencia para indicar un porcentaje más alto de cáncer de pulmón para estos trabajadores que para el promedio nacional?

3.113 El factor Rh En cierta población, 15% de las personas tienen tipo de sangre Rh negativo. Un banco de sangre que da servicio a esta población recibe 92 donadores en un día particular.

- a. ¿Cuál es la probabilidad de que 10 o menos tengan Rh negativo?
- b. ¿Cuál es la probabilidad de que 15 a 20 (inclusive) de los donadores tengan Rh negativo?
- c. ¿Cuál es la probabilidad de que más de 80 de los donadores tengan Rh negativo?

3.114 Dijimos, “¡descansen!” Un estudio realizado por una universidad reporta que 80% de hombres y 62% de mujeres emplean más de 40 horas a la semana en el trabajo. Suponga que estos porcentajes son correctos para todos los empleados de gobierno y que se selecciona una muestra aleatoria de 50 empleados.

- a. ¿Cuál es el número promedio de mujeres que emplean más de 40 horas a la semana en el trabajo?
- b. ¿Cuál es la desviación estándar para el número de mujeres que emplean más de 40 horas a la semana en el trabajo?
- c. Suponga que en nuestra muestra de 50 mujeres hay 25 que trabajan más de 40 horas a la semana. ¿Considera usted que esto es un suceso poco común? Explique.

REPASO DEL CAPÍTULO

Conceptos clave y fórmulas

I. La variable aleatoria binomial

- Cinco características:** n intentos independientes idénticos, cada uno resultando ya sea en *éxito* (E) o en *fracaso* (F); la probabilidad de éxito es p y es constante de un intento a otro; y x es el número de éxitos en n intentos.
- Cálculo de probabilidades binomiales**
 - Fórmula: $P(x = k) = C_k^n p^k q^{n-k}$
 - Tablas binomiales acumulativas
 - Probabilidades individuales y acumulativas usando *MINITAB* y *MS Excel*

- Media de la variable aleatoria binomial: $\mu = np$
- Varianza y desviación estándar: $\sigma^2 = npq$ y $\sigma = \sqrt{npq}$

II. La variable aleatoria de Poisson

- El número de eventos que ocurren en un periodo o espacio, durante el cual se espera que ocurra un promedio de μ eventos.
- Cálculo de probabilidades de Poisson**
 - Fórmula: $P(x = k) = \frac{\mu^k e^{-\mu}}{k!}$

- b. Tablas acumulativas de Poisson
- c. Probabilidades individuales y acumulativas usando *MINITAB* y *MS Excel*
- 3. Media de la variable aleatoria de Poisson:
 $E(x) = \mu$
- 4. Varianza y desviación estándar:
 $\sigma^2 = \mu$ y $\sigma = \sqrt{\mu}$
- 5. Las probabilidades binomiales se pueden aproximar con probabilidades de Poisson cuando $np < 7$, usando $\mu = np$

III. La variable aleatoria hipergeométrica

- 1. El número de éxitos en una muestra de tamaño n de una población finita que contiene M éxitos y $N - M$ fracasos.
- 2. Fórmula para la probabilidad de k éxitos en n intentos:

$$P(x = k) = \frac{C_k^M C_{n-k}^{N-M}}{C_n^N}$$

- 3. Media de la variable aleatoria hipergeométrica:

$$\mu = n \left(\frac{M}{N} \right)$$

- 4. Varianza y desviación estándar:

$$\sigma^2 = n \left(\frac{M}{N} \right) \left(\frac{N-M}{N} \right) \left(\frac{N-n}{N-1} \right) \text{ y } \sigma = \sqrt{\sigma^2}$$

IV. Experimentos y espacio muestral

- 1. Experimentos, eventos, eventos mutuamente excluyentes, eventos simples
- 2. El espacio muestral
- 3. Diagramas de Venn, diagramas de árbol, tablas de probabilidad

V. Probabilidades

- 1. Definición de probabilidad de frecuencia relativa
- 2. Propiedades de probabilidades
 - a. Cada probabilidad está entre 0 y 1.
 - b. La suma de todas las probabilidades de evento simple es igual a 1.
- 3. $P(A)$, la suma de las probabilidades para todos los eventos simples en A .

VI. Reglas de conteo

- 1. Regla mn ; regla mn extendida
- 2. Permutaciones: $P_r^n = \frac{n!}{(n-r)!}$
- 3. Combinaciones: $C_r^n = \frac{n!}{r!(n-r)!}$

VII. Relaciones de eventos

- 1. Uniones e intersecciones
- 2. Eventos
 - a. Ajenos o mutuamente excluyentes:
 $P(A \cap B) = 0$
 - b. Complementarios: $P(A) = 1 - P(A^c)$
- 3. Probabilidad condicional: $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- 4. Eventos independientes y dependientes
- 5. Regla de la adición:
 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- 6. Regla de la multiplicación:
 $P(A \cap B) = P(A)P(B|A)$
- 7. Ley de probabilidad total
- 8. Teorema de Bayes

VIII. Variables aleatorias discretas y distribuciones de probabilidad

- 1. Variables aleatorias, discretas y continuas
- 2. Propiedades de distribuciones de probabilidad
 - a. $0 \leq p(x) \leq 1$
 - b. $\sum p(x) = 1$
- 3. Media o valor esperado de una variable aleatoria discreta: $\mu = \sum xp(x)$
- 4. Varianza y desviación estándar de una variable aleatoria discreta: $\sigma^2 = \sum (x - \mu)^2 p(x)$ y $\sigma = \sqrt{\sigma^2}$

IX. Distribuciones continuas de probabilidad

- 1. Variables aleatorias continuas.
- 2. Distribuciones de probabilidad o funciones de densidad de probabilidad:
 - a. Las curvas son lisas
 - b. El área bajo la curva es igual a 1
 - c. El área bajo la curva entre a y b representa la probabilidad de que x caiga entre a y b
 - d. $P(x = a) = 0$ para variables aleatorias continuas

X. La distribución normal de probabilidad

- 1. Es simétrica alrededor de su media μ .
- 2. Forma determinada por su desviación estándar σ .

XI. La distribución normal estándar

- 1. La variable aleatoria normal estándar z tiene media 0 y desviación estándar 1.
- 2. Cualquier variable aleatoria normal x puede ser

transformada a una variable aleatoria normal estándar usando

$$z = \frac{x - \mu}{\sigma}$$

3. Convierta valores necesarios de x a z .
4. Use la tabla 3 del Anexo para calcular probabilidades normales estándar.

5. Varios valores z importantes tienen áreas de cola derecha como sigue:

Área de cola derecha	.005	.01	.025	.05	.10
Valor z	2.58	2.33	1.96	1.645	1.28



TECNOLOGÍA ACTUAL

Probabilidades binomiales y de Poisson en *Microsoft Excel*

Para una variable aleatoria que tiene, ya sea una distribución de probabilidad binomial o una de Poisson, *MS Excel* ha sido programado para calcular ambas probabilidades exactas, $P(x = k)$, para un valor dado de k o probabilidades acumulativas, $P(x \leq k)$, para un valor dado de k . Usted debe especificar cuál distribución está usando y los parámetros necesarios: n y p para la distribución binomial y μ para la distribución de Poisson.

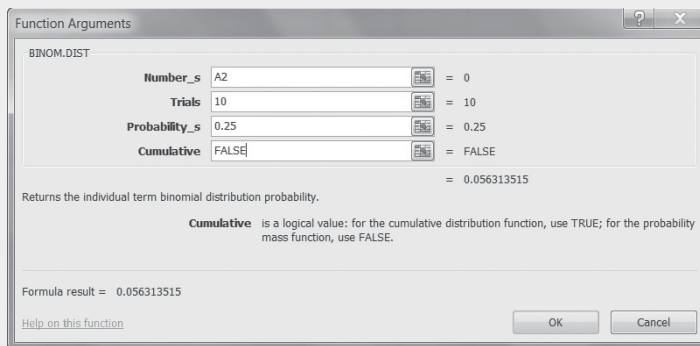
Probabilidades binomiales

1. Considere una distribución binomial con $n = 10$ y $p = .25$. El valor de p no aparece en la tabla 1 del Anexo, pero puede usar *Excel* para generar la distribución de probabilidad entera al igual que las probabilidades acumulativas al introducir los números 0–10 en la columna A.

Una forma de introducir rápido un conjunto de enteros consecutivos en una columna es hacer lo siguiente:

- Nombre las columnas A, B y C como “ x ”, “ $P(x = k)$ ” y $P(x \leq k)$, respectivamente.
 - Introduzca los dos primeros valores de x , 0 y 1, para crear un patrón en la columna A.
 - Use el ratón para resaltar los primeros dos enteros. Luego sujete el asa cuadrada que está en la esquina inferior derecha del área resaltada. Arrastre el asa hacia abajo para continuar el patrón.
 - Mientras arrastra, verá aparecer un entero en un pequeño rectángulo. Suelte el ratón cuando tenga el número deseado de enteros —en este caso, 10.
2. Una vez que se han introducido los valores necesarios de x , coloque el cursor en la celda correspondiente a $p(0)$, celda B2 en la hoja de cálculo. Seleccione el icono **Insert Function** . En la lista desplegable, seleccione la categoría **Statistical**, seleccione la función **BINOM.DIST** y haga clic en **OK**. (NOTA: Esta función se llama **BINOMDIST** en *Excel 2007* y versiones anteriores.) Aparecerá el cuadro de diálogo que se muestra en la figura 3.27.

FIGURA 3.27



- Debe teclear o seleccionar números o ubicaciones de celda para cada uno de los cuatro cuadros. Cuando coloque el cursor en el cuadro, verá una descripción de la entrada necesaria para ese cuadro. Introduzca la dirección de la celda correspondiente a $x = 0$ (celda A2) en el primer cuadro, el valor de n en el segundo cuadro, el valor de p en el tercer cuadro y la palabra FALSE en el cuarto cuadro para calcular $P(x = k)$.
- La probabilidad resultante está marcada como “Formula result = .056313515” en la parte inferior del cuadro, y cuando haga clic en **OK**, aparecerá la probabilidad $P(x = 0)$ en la celda B2. Para obtener las otras probabilidades, simplemente coloque el cursor en la celda B2, sujete el asa cuadrada en la esquina inferior derecha de la celda y arrastre el asa hacia abajo para copiar la fórmula en las otras nueve celdas. *MS Excel* ajustará en forma automática la ubicación de las celdas en la fórmula conforme las copie.
- Si desea generar las probabilidades acumulativas, $P(x \leq k)$, coloque el cursor en la celda correspondiente para $P(x \leq 0)$, celda C2 en la hoja de cálculo. Luego seleccione **Insert Function ► Statistical ► BINOM.DIST** y haga clic en **OK**. Continúe como en los pasos 3 y 4, pero teclee TRUE en la cuarta línea del cuadro de diálogo para calcular $P(x \leq k)$. La salida resultante se muestra en la figura 3.28.

FIGURA 3.28

	A	B	C
1	x	$P(x = k)$	$P(x \leq k)$
2	0	0.0563	0.0563
3	1	0.1877	0.2440
4	2	0.2816	0.5256
5	3	0.2503	0.7759
6	4	0.1460	0.9219
7	5	0.0584	0.9803
8	6	0.0162	0.9965
9	7	0.0031	0.9996
10	8	0.0004	1.0000
11	9	0.0000	1.0000
12	10	0.0000	1.0000

- ¿Cuál valor k es tal que sólo 5% de los valores de x exceden este valor (y 95% son menores o iguales a k)? Coloque el cursor en una celda vacía, seleccione **Insert Function ► Statistical ► BINOM.INV** y haga clic en **OK**. (NOTA: Esta función es nueva en *Excel* 2010.) El cuadro de diálogo resultante calculará lo que en ocasiones se llama **probabilidad acumulativa inversa**. Teclee 10 en el primer cuadro, .25 en el segundo y .95 en el tercero. Cuando haga clic en **OK** aparecerá el número 5 en la celda vacía. Éste es el

valor más pequeño de k para el cual $P(x \leq k)$ es mayor o igual a .95. Consulte la figura 3.28 y observe que $P(x \leq 5) = .9803$ de modo que $P(x > 5) = 1 - .9803 = .0197$. Por tanto, si observó un valor de $x = 5$, ésta sería una observación poco común.

Probabilidades de Poisson

1. Los procedimientos para calcular probabilidades y distribuciones de probabilidad individuales o acumulativas para la variable aleatoria de Poisson son parecidos a los usados para la distribución binomial.
2. Para hallar probabilidades de Poisson $P(x = k)$ o $P(x \leq k)$, seleccione **Insert Function ► Statistical ► POISSON.DIST**, y haga clic en **OK**. (NOTA: Esta función se llama **POISSON** en *Excel 2007* y versiones anteriores.) Introduzca los valores para k , μ y **FALSE/TRUE** antes de hacer clic en **OK**.
3. No hay **probabilidad acumulativa inversa** como en el caso de la distribución binomial.

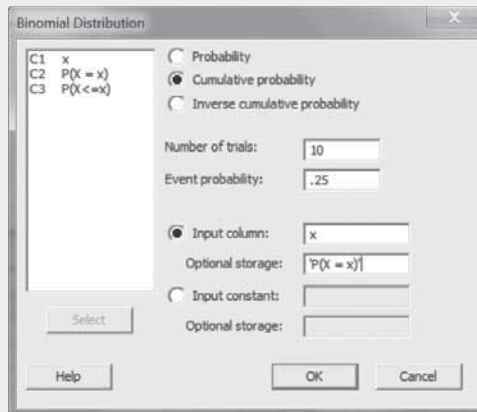
Probabilidades binomiales y de Poisson en MINITAB

Para una variable aleatoria que tiene una distribución de probabilidad binomial o una de Poisson, *MINITAB* ha sido programado para calcular ya sea probabilidades exactas, $P(X = x)$, para un valor determinado de x , o las probabilidades acumulativas, $P(X \leq x)$, para un valor determinado de x . (NOTA: *MINITAB* usa la notación “ X ” para la variable aleatoria y “ x ” para un valor particular de la variable aleatoria.) El usuario debe especificar cuál distribución está usando y los parámetros necesarios: n y p para la distribución binomial y μ para la distribución de Poisson.

Probabilidades binomiales

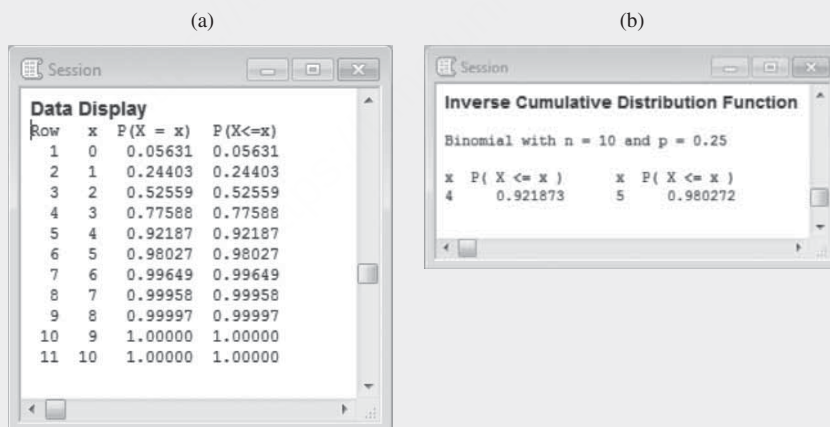
1. Considere una distribución binomial con $n = 10$ y $p = .25$. El valor de p no aparece en la tabla 1 del Anexo, pero puede usar *MINITAB* para generar la distribución de probabilidad entera al igual que las probabilidades acumulativas introduciendo los números 0 a 10 en la columna A.
2. Una forma de introducir rápidamente un conjunto de enteros consecutivos en una columna es hacer lo siguiente:
 - Nombrar las columnas C1, C2 y C3 como “ x ”, “ $P(X = x)$ ” y $P(X \leq x)$, respectivamente.
 - Introduzca los primeros dos valores de x , 0 y 1, para crear un patrón en la columna C1.
 - Use el ratón para seleccionar los primeros dos enteros. Luego sujete el asa cuadrada de la esquina inferior izquierda del área seleccionada. Arrastre el asa hacia abajo para continuar con el patrón.
 - Notará que aparece un entero en un pequeño cuadro. Suelte el ratón cuando tenga el número deseado de enteros, en este caso 10.
3. Una vez introducidos los valores necesarios de x , use **Calc ► Probability Distributions ► Binomial** para generar el cuadro de diálogo que se muestra en la figura 3.29.

FIGURA 3.29



4. Teclee el número de intentos y el valor de p (Event probability) (probabilidad de evento) en los cuadros apropiados, seleccione C1 ('x') para la columna de entradas y C2 (' $P(X = x)$ ') para el almacenamiento Optional. Asegúrese de que el botón de radio marcado "Probability" está seleccionado. La distribución de probabilidad para x aparecerá en la columna C2 cuando haga clic en **OK**. (NOTA: Si no selecciona una columna para el almacenamiento Optional, los resultados se mostrarán en la ventana Session.)
5. Si desea generar las probabilidades acumulativas, $P(x \leq k)$, una vez más use **Calc ► Probability Distributions ► Binomial** para generar el cuadro de diálogo. Esta vez, seleccione el botón de radio marcado "Cumulative probability" y seleccione C3 ($P(X \leq x)$) para el almacenamiento Optional en el cuadro de diálogo (figura 3.29). La distribución de probabilidad acumulativa aparecerá en la columna C3 cuando haga clic en **OK**. Puede desplegar ambas distribuciones en la ventana Session usando **Data ► Display Data**, seleccionando las columnas C1–C3 y haciendo clic en **OK**. Los resultados se muestran en la figura 3.30a).

FIGURA 3.30



6. ¿Cuál valor x es tal que sólo 5% de los valores de la variable aleatoria X exceden este valor (y 95% son menores o iguales a x)? De nuevo, use **Calc ► Probability Distributions ► Binomial** para generar el cuadro de diálogo. Esta vez, seleccione el botón de radio marcado "Inverse cumulative probability" e introduzca la probabilidad .95 en el cuadro "Input constant" (figura 3.29). Cuando haga clic en **OK**, los valores de x en ambos lados de la "marca .95" aparecerán en la ventana Session como se muestra en la figura 3.30b). Por tanto, si observó un valor de $x = 5$, ésta sería una observación inusual, debido a que $P(x > 5) = 1 - .980272 = .019728$.

Probabilidades de Poisson

1. Los procedimientos para calcular probabilidades y distribuciones de probabilidad individuales o acumulativas para la variable aleatoria de Poisson son parecidos a los usados para la distribución binomial.
2. Para encontrar probabilidades de Poisson $P(X = x)$ o $P(X \leq x)$, use **Calc ► Probability Distributions ► Poisson** para generar el cuadro de diálogo. Introduzca el valor para la media μ , elija el botón de radio apropiado e introduzca ya sea una columna o una constante para indicar el valor o valores de X para los cuales desea calcular una probabilidad antes de hacer clic en **OK**.
3. La **probabilidad acumulativa inversa** calcula los valores de x tales que $P(X \leq x) = C$, donde C es una probabilidad constante, entre 0 y 1. Siga los pasos descritos para la variable aleatoria binomial en el paso 6 anterior.

EJEMPLO

3.32

Suponga que tiene esta distribución de probabilidad discreta general:

x	0	1	3	5
$p(x)$	.25	.35	.25	.15

1. Introduzca los valores de x y $p(x)$ en las columnas A y B de una hoja de cálculo nueva de *Excel*. Luego cree dos columnas: la columna C (llamada " $x \cdot p(x)$ ") y la columna D (llamada " $x^2 \cdot p(x)$ "). Ahora use el comando **Function**  para llenar las columnas C y D. En *Excel*, un signo "igual" indica que se va a escribir una ecuación (o a insertar una función). Por tanto, en la celda C2, escriba:

$$=A2*B2$$
Luego, para copiar esta fórmula en las tres celdas restantes en la columna C, tan sólo haga clic en la celda C2, tome con el ratón el cuadrado en la esquina inferior derecha de la celda y arrástrelo hasta la celda C5.
2. Para llenar la columna D, escriba la siguiente ecuación en la celda D2: $=A2*A2*B2$ y luego copie esta fórmula a las tres celdas restantes en la columna D como se explicó antes.
3. Por último, usamos las primeras tres celdas en la columna F para escribir los nombres "Media" (Mean), "Varianza" (Variance) y "Desviación estándar" (Std Dev). Una vez más, use los comandos de ecuación (o use la función de insertar). En la celda G1 (Media), escriba: $=SUM(C2:C5)$; en la celda G2 (Varianza), escriba: $=SUM(D2:D5)-(G1*G1)$; y en la celda G3 (Desviación estándar), escriba: $=SQRT(G2)$. La hoja de cálculo resultante se muestra en la figura 3.31.

FIGURA 3.31

	A	B	C	D	E	F	G
1	x	p(x)	$x \cdot p(x)$	$x^2 \cdot p(x)$		Mean	1.85
2	0	0.25	0	0		Variance	2.9275
3	1	0.35	0.35	0.35		Std Dev	1.710994
4	3	0.25	0.75	2.25			
5	5	0.15	0.75	3.75			

Distribuciones discretas de probabilidad en MINITAB

Aun cuando el *MINITAB* no ayuda a resolver los tipos de problemas generales de probabilidad presentados en este capítulo, es útil para graficar la distribución de probabilidad $p(x)$ para una variable general aleatoria discreta x cuando se conocen las probabilidades, y para calcular la media, varianza y desviación estándar de la variable aleatoria x . Se usará *MINITAB* para calcular las probabilidades exactas para tres casos especiales: las variables binomiales, las variables de Poisson y las variables aleatorias normales.

EJEMPLO

3.33

Supongamos que usted tiene esta distribución general de probabilidad:

x	0	1	3	5
$p(x)$	.25	.35	.25	.15

1. Introduzca los valores de x y $p(x)$ en las columnas C1 y C2 de una hoja de trabajo *MINITAB* nueva. En las cajas grises un poco abajo de C3, C4 y C5, respectivamente, escriba los nombres “Media”, “Varianza” y “Desviación estándar”. Ahora use el comando **Calc ► Calculator** para calcular μ , σ^2 y σ y para guardar los resultados en las columnas C3–C5 de la hoja de trabajo.
2. Use el mismo método para los tres parámetros. En el cuadro de diálogo Calculator, seleccione “Media” como la columna en la cual guardar μ . En el cuadro Expression, use la lista de Funciones, las teclas de la calculadora y la lista de variables de la izquierda para resaltar, seleccionar y crear la expresión para la media (véase la figura 3.32a):

$\text{SUM}('x' * p(x))$

¡*MINITAB* multiplicará el elemento de cada renglón en C1 veces por el correspondiente elemento de renglón en C2, sumará los productos resultantes y guardará el resultado en C3! Compruebe manualmente el resultado si lo desea.

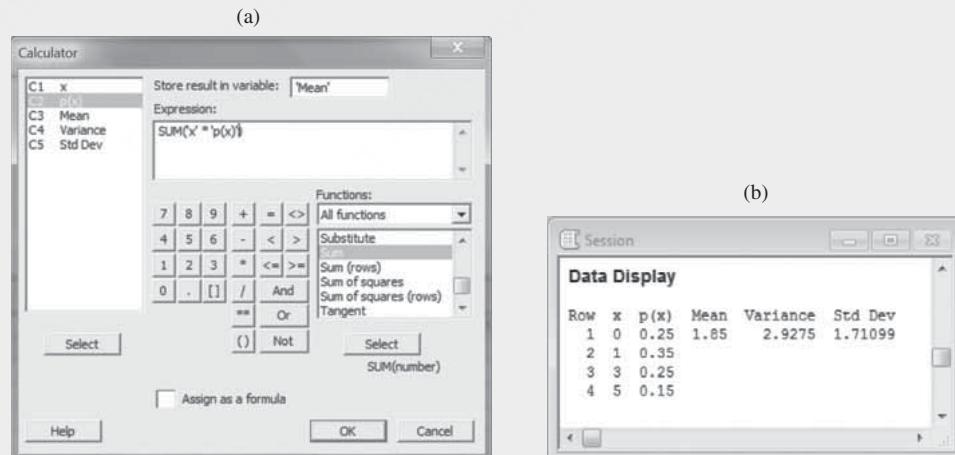
3. Las fórmulas para la varianza y desviación estándar se seleccionan de un modo semejante:

Varianza: $\text{SUM}((x - \text{Media})^2 * p(x))$

Desviación estándar: $\text{SQRT}(\text{Varianza})$

4. Para ver la forma tabular de la distribución de probabilidad y los tres parámetros, use **Data ► Display Data** y seleccione las cinco columnas. Haga clic en **OK** y los resultados se exhibirán en la ventana Session, como se muestra en la figura 3.32b).

FIGURA 3.32

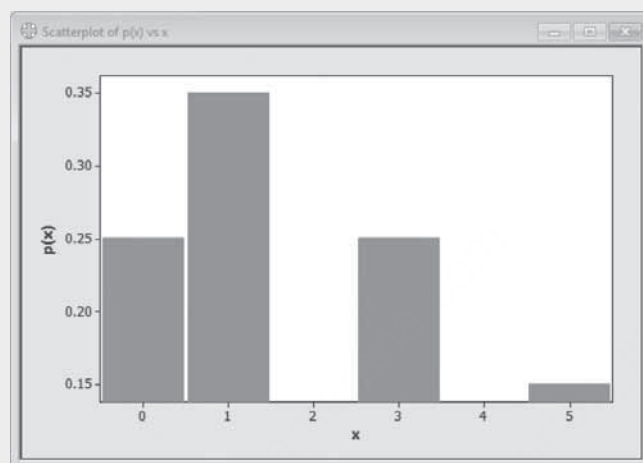


El histograma de probabilidad se grafica usando el comando *MINITAB* **Graph ► Scatterplot ► Simple ► OK**. En el cuadro de diálogo Scatterplot, seleccione ‘p(x)’ para **Y variables** y ‘x’ para **X variables**. Para exhibir las barras discretas de probabilidad, haga clic en **Data View**, quite la marca de la caja marcada “Symbols”, y ponga marca en la caja “Project Lines”. Haga doble clic en **OK** para ver la gráfica. Verá una sola recta proyectada en cada uno de los cuatro valores de x . Si desea que la gráfica se vea más como los histogramas discretos de probabilidad, coloque el cursor en una de las rectas, haga clic en el lado derecho del ratón y elija “Edit Project Lines”. Bajo la ficha “Attributes”, seleccione **Custom** y cambie el

tamaño de recta a **75**. Haga clic en **OK**. Si el ancho de la barra no es satisfactorio, reajuste el tamaño de la recta. Por último, haga clic derecho en el eje X, elija “Edit X Scale” y seleccione **-.5** y **5.5** para las **Scale Ranges** mínimas y máximas. Haga clic en **OK**. El histograma de probabilidad se muestra en la figura 3.33.

Localice la media en la gráfica. ¿Está en el centro de la distribución? Si usted marca dos desviaciones estándar en cualquier lado de la media, ¿la mayoría de los posibles valores de x cae en este intervalo?

FIGURA 3.33



Distribuciones discretas de probabilidad en *MS Excel*

Aunque *MS Excel* no puede ayudarle a resolver los tipos de problemas de probabilidad generales que se presentan en este capítulo, es útil para calcular la media, varianza y desviación estándar de la variable aleatoria x . Se usará *Excel* para calcular probabilidades exactas para tres casos especiales: las variables binomiales, las variables de Poisson y las variables aleatorias normales.

Probabilidades normales en *Microsoft Excel*

Cuando la variable aleatoria de interés tiene una distribución de probabilidad normal, se pueden generar las siguientes probabilidades usando las funciones que se presentan a continuación:

- **NORM.DIST** y **NORM.S.DIST**. Generan probabilidades acumulativas: $P(x \leq x_0)$ para una variable aleatoria normal general o $P(z \leq z_0)$ para una variable aleatoria normal estándar. (NOTA: Estas funciones se llaman **NORMDIST** y **NORMSDIST** en *Excel 2007* y versiones anteriores.)
- **NORM.INV** y **NORM.S.INV**. Generan probabilidades acumulativas inversas: el valor x_0 tal que el área a su izquierda bajo la distribución de probabilidad normal general es igual a a , o el valor de z_0 tal que el área a su izquierda bajo la distribución de probabilidad normal estándar es igual a a . (NOTA: Estas funciones se llaman **NORMINV** y **NORMSINV** en *Excel 2007* y versiones anteriores.)

Se debe especificar cuál distribución normal se está usando y, si es una variable aleatoria normal general, los valores para la media μ y la desviación estándar σ . También deben especificarse los valores para x_0 , z_0 o a , dependiendo de la función que se esté usando.

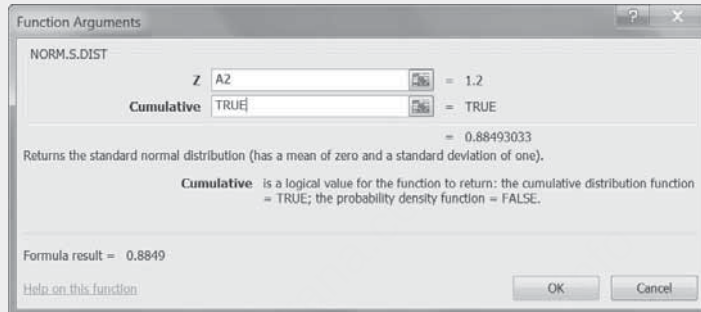
EJEMPLO

3.34

Para una variable aleatoria normal estándar z , encuentre $P(1.2 < z < 1.96)$. Encuentre el valor z_0 con área .025 a su derecha.

1. Nombre las columnas A y B de una hoja de cálculo de Excel como “ z_0 ” y “ $P(z \leq z_0)$ ”, respectivamente. Luego introduzca los dos valores para z_0 (1.2 y 1.96) en las celdas A2 y A3. Para generar probabilidades acumulativas para estos dos valores, primero coloque el cursor en la celda B2. Seleccione **Insert Function** ► **Statistical** ► **NORM.S.DIST** y haga clic en **OK**. Aparecerá el cuadro de diálogo que se muestra en la figura 3.34.

FIGURA 3.34



2. Introduzca la ubicación del primer valor de z_0 (celda A2) en el primer cuadro y la palabra TRUE en el segundo cuadro. La probabilidad resultante está marcada como “Formula result = .8849” en la parte inferior del cuadro, y cuando haga clic en **OK**, aparecerá la probabilidad $P(z \leq 1.2)$ en la celda B2. Para obtener la otra probabilidad, simplemente coloque el cursor en la celda B2, tome el asa cuadrada en la esquina inferior derecha de la celda y arrástrela hacia abajo para copiar la fórmula en la otra celda y obtener $P(z \leq 1.96) = .9750$. *MS Excel* ajusta de manera automática la ubicación de la celda en la fórmula al copiarla.
3. Para hallar $P(1.2 < z < 1.96)$, recuerde que la probabilidad acumulativa es el área a la izquierda del valor dado de z . Por tanto,

$$P(1.2 < z < 1.96) = P(z < 1.96) - P(z < 1.2) = .9750 - .8849 = .0901.$$

¡Si lo desea, puede verificar este cálculo usando la tabla 3 en el Anexo!

4. Para calcular probabilidades acumulativas inversas, coloque el cursor en una celda vacía, seleccione **Insert Function** ► **Statistical** ► **NORM.S.INV** y haga clic en **OK**. Se necesita un valor z_0 con área .025 a su derecha, o área .975 a su izquierda. Introduzca **.975** en el cuadro marcado “Probability” y observe el dato “Formula Result = 1.959963985”, el cual aparecerá en la celda vacía cuando haga clic en **OK**. Este valor, cuando se redondea a dos lugares decimales, es el familiar $z_0 = 1.96$ usado en el ejemplo 3.24.

EJEMPLO

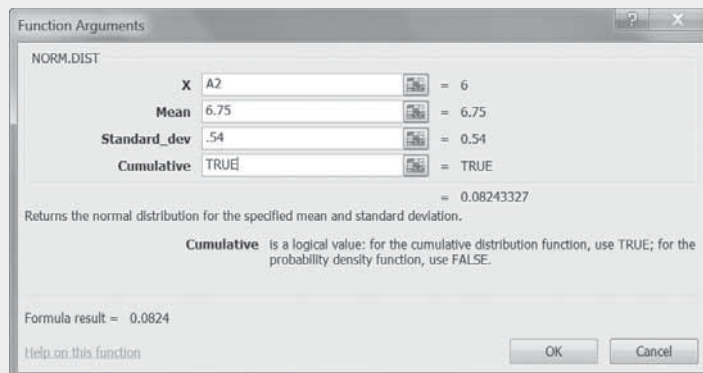
3.35

Suponga que los pesos promedio al nacer de bebés nacidos en hospitales del sector salud son aproximadamente normales con media de 6.75 libras y desviación estándar de 0.54 libras. ¿Qué proporción de bebés nacidos en estos hospitales pesa entre 6 y 7 libras? Encuentre el 95o. percentil de estos pesos al nacer.

1. Nombre las columnas A y B de una hoja de cálculo de Excel como “ x_0 ” y “ $P(x \leq x_0)$ ”, respectivamente. Luego introduzca los dos valores para x_0 (6 y 7) en las celdas A2 y A3. Proceda como en el ejemplo 3.32, esta vez seleccionando **Insert Function** ► **Statisti-**

cal ► **NORM.DIST** y haga clic en **OK**. Aparecerá el cuadro de diálogo mostrado en la figura 3.35.

FIGURA 3.35

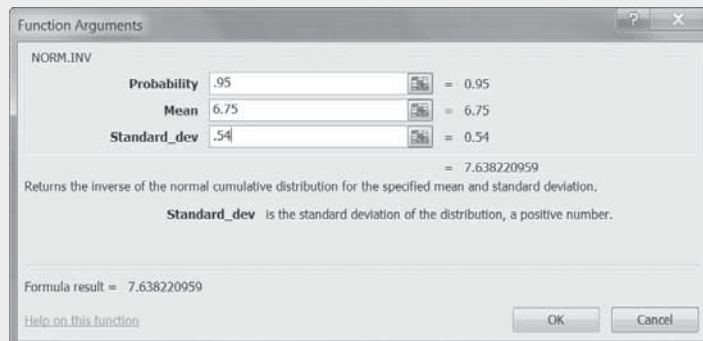


- Introduzca la ubicación del primer valor de x_0 (celda A2) en el primer cuadro, la media y desviación estándar apropiadas en el segundo y tercer cuadros, y la palabra TRUE en el cuarto cuadro. La probabilidad resultante está marcada como “Formula result = .0824” en la parte inferior del cuadro, y cuando haga clic en **OK** aparecerá la probabilidad $P(x \leq 6)$ en la celda B2. Para obtener la otra probabilidad, simplemente coloque el cursor en la celda B2, tome el asa cuadrada en la esquina inferior derecha de la celda y arrástrela hacia abajo para copiar la fórmula en la otra celda y obtener $P(x \leq 7) = .6783$.
- Por último, use los valores calculados por Excel para calcular

$$P(6 < x < 7) = P(x < 7) - P(x < 6) = .6783 - .0824 = .5959.$$

- Para calcular el 95o. percentil, coloque el cursor en una celda vacía, seleccione **Insert Function ► Statistical ► NORM.INV** y haga clic en **OK**. Se necesita un valor x_0 con área .95 a su izquierda. Introduzca .95 en el cuadro marcado “Probability”, la media y desviación estándar apropiadas (véase la figura 3.36) y observe el dato “Formula Result = 7.638220959”, el cual aparecerá en la celda vacía cuando haga clic en **OK**.

FIGURA 3.36



Es decir, 95% de todos los bebés nacidos en estos hospitales pesan 7.638 libras o menos. ¿Consideraría que un bebé que pese 9 libras es inusualmente grande?

Probabilidades normales en MINITAB

Cuando la variable aleatoria de interés tiene una distribución normal de probabilidad, se pueden generar cualquiera de estas probabilidades:

- Probabilidades acumulativas: $P(X \leq x)$ para un valor dado de x . (NOTA: MINITAB usa la notación “X” para la variable aleatoria y “x” para un valor particular de la variable aleatoria.)
- Probabilidades acumulativas inversas, el valor de x tal que el área a su izquierda bajo la distribución normal de probabilidad es igual a a .

Usted debe especificar cuál distribución normal está usando y los valores para la media μ y la desviación estándar σ . Usted tiene la opción de especificar sólo un valor de x (o de a) o varios valores de x (o de a), que deben guardarse en una columna (por ejemplo en C1) de la hoja de trabajo de MINITAB.

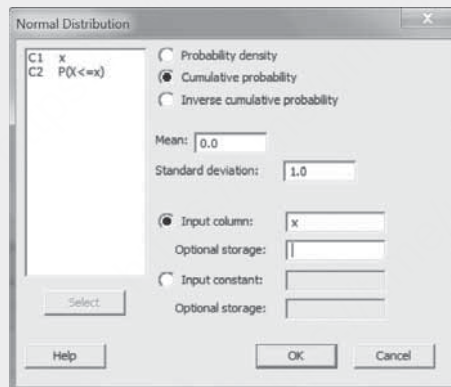
EJEMPLO

3.36

Para una variable aleatoria normal estándar z , encontrar $P(1.2 < z < 1.96)$. Encuentre el valor z_0 con área .025 a su derecha.

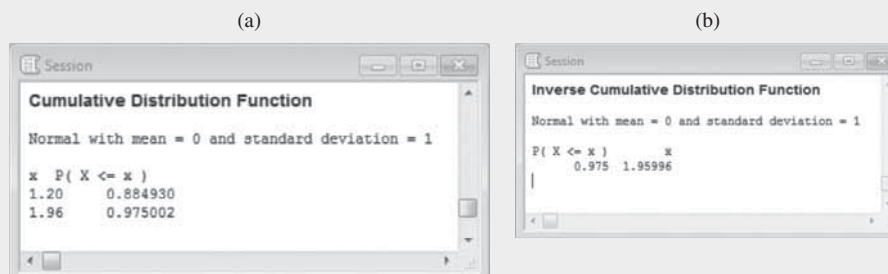
1. Nombre las columnas C1 y C2 de una hoja de cálculo de MINITAB como “x” y “P(X≤x)”, respectivamente. Luego introduzca los dos valores para x (1.2 y 1.96) en las primeras dos celdas de la columna C1. Para generar probabilidades acumulativas para estos dos valores, seleccione **Calc ► Probability Distributions ► Normal** y aparecerá el cuadro de diálogo que se muestra en la figura 3.37.

FIGURA 3.37



2. En forma predeterminada, MINITAB elige $\mu = 0$ y $\sigma = 1$ como la media y la desviación estándar de la distribución z normal estándar, así que sólo necesitará introducir la columna Input (C1) y asegurarse de que el botón de radio marcado “Cumulative probability” está seleccionado. Si no especifica una columna para “Optional storage”, MINITAB desplegará los resultados en la ventana Session, que se muestra en la figura 3.38a).

FIGURA 3.38



3. Para encontrar $P(1.2 < z < 1.96)$, recuerde que la probabilidad acumulativa es el área a la izquierda del valor dado de z . Por tanto,

$$P(1.2 < z < 1.96) = P(z < 1.96) - P(z < 1.2) = .975002 - .884930 = .090072.$$

¡Si lo desea, puede comprobar este cálculo usando la tabla 3 en el Anexo!

4. Para calcular las probabilidades acumulativas inversas, seleccione **Calc ► Probability Distributions ► Normal** y haga clic en el botón de radio marcado “Inverse cumulative probability”, que se muestra en la figura 3.37. Se necesita un valor z_0 con área .025 a su derecha, o área .975 a su izquierda. Introduzca **.975** en el cuadro marcado “Input constant” y haga clic en **OK**. El valor de z_0 aparecerá en la ventana Session, mostrada en la figura 3.38b). Este valor, cuando se redondea a dos lugares decimales, es el familiar $z_0 = 1.96$ usado en el ejemplo 3.24.

EJEMPLO

3.37

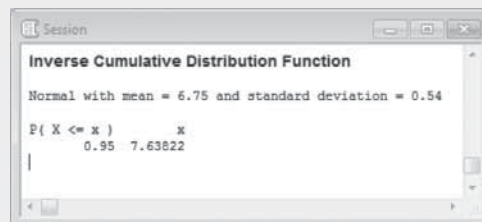
Suponga que los pesos promedio al nacer de bebés nacidos en hospitales del sector salud es aproximadamente normal con media de 6.75 libras y desviación estándar de .54 libras. ¿Qué proporción de bebés nacidos en estos hospitales pesa entre 6 y 7 libras? Encuentre el 95avo percentil de estos pesos al nacer.

1. Introduzca los dos valores para x (6 y 7) en las primeras dos celdas de la columna C1. Proceda como en el ejemplo 3.36, seleccionando de nuevo **Calc ► Probability Distributions ► Normal**. Esta vez, introduzca los valores para la media ($\mu = 6.75$) y la desviación estándar ($\sigma = .54$) en los cuadros apropiados y seleccione la columna C1 (“x”) para la columna Input. Asegúrese que el botón de radio marcado “Cumulative probability” está seleccionado y haga clic en **OK**. En la ventana Session, verá que $P(x \leq 7) = .678305$ y $P(x \leq 6) = .082433$.
2. Por último, use los valores calculados por *MINITAB* para calcular

$$P(6 < x < 7) = P(x < 7) - P(x < 6) = .678305 - .082433 = .595872.$$

3. Para calcular el 95o. percentil, seleccione **Calc ► Probability Distributions ► Normal**, introduzca los valores para la media ($\mu = 6.75$) y la desviación estándar ($\sigma = .54$) en los cuadros apropiados y asegúrese que el botón de radio marcado “Inverse cumulative probability” está seleccionado. Se necesita un valor x_0 con área .95 a su izquierda. Introduzca **.95** en el cuadro marcado “Input constant” y haga clic en **OK**. En la ventana Session verá el 95o. percentil, como se muestra en la figura 3.39.

FIGURA 3.39



Es decir, 95% de todos los bebés nacidos en estos hospitales pesan 7.63822 libras o menos. ¿Consideraría que un bebé que pesa 9 libras es inusualmente grande?

Ejercicios suplementarios

3.115 Haga una lista de las cinco características del experimento binomial.

3.116 ¿Bajo qué condiciones se usa la variable aleatoria de Poisson para aproximar las probabilidades asociadas con la variable aleatoria binomial? ¿Qué aplicaciones tiene la distribución de Poisson que no sea estimar ciertas probabilidades binomiales?

3.117 ¿Bajo qué condiciones usaría la distribución hipergeométrica de probabilidad para evaluar la probabilidad de x éxitos en n intentos?

3.118 Tiro de una moneda Una moneda equilibrada se lanza al aire tres veces. Sea x igual al número observado de caras.

- Use la fórmula para la distribución binomial de probabilidad para calcular las probabilidades asociadas con $x = 0, 1, 2$ y 3 .
- Construya la distribución de probabilidad.
- Encuentre la media y desviación estándar de x , usando estas fórmulas:

$$\mu = np$$

$$\sigma = \sqrt{npq}$$

- Usando la distribución de probabilidad de la parte **b**, encuentre la fracción de las mediciones de población que estén a no más de una desviación estándar de la media. Repita para dos desviaciones estándar. ¿Cómo concuerdan los resultados con el teorema de Chebyshev y la regla empírica?

3.119 Monedas, continúa Consulte el ejercicio 3.118. Suponga que la moneda está en verdad “cargada” y la probabilidad de que salga una “cara” es igual a $p = .1$. Siga las instrucciones de las partes **a**, **b**, **c** y **d**. Observe que la distribución de probabilidad pierde su simetría y se vuelve sesgada cuando p no es igual a $1/2$.

3.120 Porcentajes para sobrevivir al cáncer El porcentaje de sobrevivencia de 10 años al cáncer en la vejiga es alrededor del 50%. Si 20 personas que tienen cáncer en la vejiga reciben tratamiento adecuado para esa enfermedad, ¿cuál es la probabilidad de que

- al menos 1 sobrevivirá los 10 años?
- al menos 10 sobrevivirán los 10 años?
- al menos 15 sobrevivirán los 10 años?

3.121 Recolección de basura El comisionado de un municipio declara que 80% de la población de la ciudad está a favor de que la recolección de basura sea por contrato a una empresa privada (en contraste a la recolección por empleados del municipio). Para verificar la afirmación del 80%, usted muestrea al azar 25 personas

y encuentra que x , el número de personas que apoyan lo declarado por el comisionado, es 22.

- ¿Cuál es la probabilidad de observar al menos 22 que apoyen lo dicho por el comisionado si, en efecto, $p = .8$?
- ¿Cuál es la probabilidad de que x sea exactamente igual a 22?
- Con base en los resultados de la parte **a**, ¿qué concluiría acerca de la afirmación de que 80% de todas las personas de la ciudad está a favor de la recolección privada? Explique.

3.122 Enteros A 20 personas se les pide seleccionar un número de 0 a 9. Ocho de ellas eligen un 4, 5 o 6.

- Si la selección de cualquier número es tan probable como cualquier otra, ¿cuál es la probabilidad de observar ocho o más selecciones de los números 4, 5 o 6?
- ¿Qué conclusiones se sacarían de los resultados de la parte **a**?

3.123 Problemas psicosomáticos Una psiquiatra dice que 80% de las personas que van a consulta tienen problemas de naturaleza psicosomática. Ella decide seleccionar al azar 25 pacientes para probar su teoría.

- Suponiendo que la teoría de la psiquiatra es verdadera, ¿cuál es el valor esperado de x , el número de los 25 pacientes que tienen problemas psicosomáticos?
- ¿Cuál es la varianza de x , suponiendo que la teoría es verdadera?
- Encuentre $P(x \leq 14)$. (Use las tablas y suponga que la teoría es verdadera.)
- Con base en la probabilidad de la parte **c**, si sólo 14 de los 25 muestreados tenían problemas psicosomáticos, ¿a qué conclusiones llegaría acerca de la teoría de la psiquiatra? Explique.

3.124 Canas en la universidad ¡Los planteles universitarios están envejeciendo! De acuerdo con una encuesta, uno de cada cuatro estudiantes tiene 30 años de edad o más. Suponga que la cifra de 25% es precisa, que su universidad es representativa de universidades en general, y que su muestreo $n = 200$ estudiantes, registrándose x , el número de estudiantes de 30 años de edad o más.

- ¿Cuáles son la media y la desviación estándar de x ?
- Si hay 35 estudiantes en su muestra que tengan 30 años de edad o más, ¿estaría usted dispuesto a suponer que la cifra de 25% es representativa de su plantel? Explique.

3.125 Probabilidad de lluvia Para comprobar la precisión de cierto meteorólogo, se verificaron registros

sólo para aquellos días cuando el meteorólogo predijo lluvia “con 30% de probabilidad”. Una verificación de 25 de esos días indicó que llovió en 10 de los 25.

- Si el meteorólogo es preciso, ¿cuál es el valor apropiado de p , la probabilidad de lluvia en uno de los 25 días?
- ¿Cuáles son la media y la desviación estándar de x , el número de días en los que llovió, suponiendo que el meteorólogo es preciso?
- Calcule el puntaje z para el valor observado, $x = 10$. [SUGERENCIA: recuerde de la sección 1.6 que el puntaje z es $=(x - \mu)/\sigma$.]
- ¿Estos datos no concuerdan con el pronóstico de un “30% de probabilidad de lluvia”? Explique.

3.126 ¿Qué hay para desayunar? Se lleva a cabo un experimento de empaque al colocar, uno junto a otro, dos diseños diferentes de paquetes para desayuno en un estante de supermercado. En un día determinado, 25 clientes compraron un paquete del supermercado. Sea x igual al número de compradores que eligieron el segundo diseño de paquete.

- Si no hay preferencia por alguno de los dos diseños, ¿cuál es el valor de p , la probabilidad de que un comprador elija el segundo diseño de paquete?
- Si no hay preferencia, use los resultados de la parte **a** para calcular la media y la desviación estándar de x .
- Si 5 de los 25 clientes eligieron el diseño del primer paquete y 20 el segundo diseño, ¿qué concluiría usted acerca de la preferencia de los compradores por el segundo diseño de paquete?

3.127 Densidad de la planta Un modelo para competencia de plantas supone que hay una zona de agotamiento de recursos alrededor del vástago de cada planta. Dependiendo del tamaño de las zonas y la densidad de las plantas, las zonas de agotamiento de recursos pueden traslaparse con las de otros vástagos de la cercanía. Cuando las semillas se dispersan al azar en una superficie amplia, el número de vecinos que un vástago pueda tener por lo general sigue una distribución Poisson con una media igual a la densidad de semillas por unidad de área. Suponga que la densidad de vástagos es cuatro por metro cuadrado (m^2).

- ¿Cuál es la probabilidad de que un vástago determinado no tenga vecinos dentro de $1 m^2$?
- ¿Cuál es la probabilidad de que un vástago tenga a lo sumo tres vecinos por m^2 ?
- ¿Cuál es la probabilidad de que un vástago tenga cinco o más vecinos por m^2 ?
- Use el hecho de que la media y la varianza de una variable aleatoria de Poisson son iguales, para hallar

la proporción de vecinos que caerían en el intervalo $\mu \pm 2\sigma$. Comente sobre este resultado.

3.128 Genética de plantas Una peonía con pétalos rojos fue cruzada con otra planta que tenía pétalos veteados. La probabilidad de que un descendiente de esta cruce tenga flores rojas es .75. Sea x el número de plantas con pétalos rojos que resulte de 10 semillas de esta cruce que se recolectaron y germinaron.

- ¿La variable aleatoria x tiene una distribución binomial? Si no es así, ¿por qué? Si es así, ¿cuáles son los valores de n y p ?
- Encuentre $P(x \geq 9)$.
- Encuentre $P(x \leq 1)$.
- ¿Sería inusual observar una planta con pétalos rojos y las nueve plantas restantes con pétalos veteados? Si realmente se presentan estos resultados, ¿qué conclusiones podría sacar?

3.129 Características dominantes Los alelos de color negro (B) y blanco (b) de plumas de pollos muestran dominancia incompleta; los individuos con el par de genes Bb tienen plumas “azules”. Cuando un individuo que es homocigoto dominante (BB) para esta característica se aparea con un individuo que es homocigoto recesivo (bb) para esta característica, $1/2$ llevarán el par Bb de genes. Sea x el número de pollos con plumas “azules” en una muestra de $n = 20$ pollos que resulta de este tipo de cruces.

- ¿La variable aleatoria x tiene una distribución binomial? Si no es así, ¿por qué? Si es así, ¿cuáles son los valores de n y p ?
- ¿Cuál es el número medio de pollos con plumas “azules” en la muestra?
- ¿Cuál es la probabilidad de observar menos de cinco pollos con plumas “azules”?
- ¿Cuál es la probabilidad de que el número de pollos con plumas “azules” sea mayor o igual a 10, pero menor o igual a 12?

3.130 Tiros al aire de monedas en fútbol Durante la temporada de fútbol de 1992, los Carneros de Los Ángeles (ahora Carneros de San Luis) tenían una insólita fila de pérdidas en los tiros al aire de monedas. De hecho, perdieron la decisión 11 semanas consecutivas.¹¹

- El gerente del sistema de cómputo de los Carneros dijo que las probabilidades en contra de perder 11 tiros al hilo son 2047 a 1. ¿Está en lo correcto?
- Después de que estos resultados se publicaron, los Carneros perdieron la decisión en los siguientes dos juegos, para un total de 13 pérdidas consecutivas. ¿Cuál

es la probabilidad de que esto ocurra si, de hecho, la moneda era imparcial?

3.131 Diabetes en niños La diabetes dependiente de insulina (IDD) en niños se presenta con más frecuencia en descendientes de personas del norte de Europa. Los rangos de incidencia van desde lo bajo, 1-2 casos en 100 000 por año, hasta lo alto con más de 40 por 100 000 en áreas de Finlandia.¹² Supongamos que una región de Europa tiene una incidencia de 5 casos en 100 000 por año.

- ¿La distribución del número de casos de la IDD en esta región puede ser aproximada por una distribución de Poisson? Si es así, ¿cuál es la media?
- ¿Cuál es la probabilidad de que el número de casos de la IDD en esta región sea menor o igual a 3 en 100 000?
- ¿Cuál es la probabilidad de que el número de casos sea mayor o igual a 3, pero menor o igual a 7 en 100 000?
- ¿Esperaría usted observar 10 o más casos de la IDD en 100 000 en esta región en un año determinado? ¿Por qué?

3.132 ¿Problemas con su nuevo teléfono

inteligente? Un estudio nuevo realizado por Square Trade indica que los teléfonos inteligentes tienen 50% más probabilidad de fallar que los teléfonos simples en un periodo de tres años.¹³ De las fallas de los teléfonos inteligentes, 30% se relacionan con componentes internos que no funcionan y, en general, hay 31% de posibilidades de que su teléfono inteligente falle a lo largo de tres años. Suponga que los teléfonos inteligentes son enviados en cajas de $N = 50$ teléfonos. Antes del envío se seleccionan $n = 10$ teléfonos de cada caja y si ninguno de los teléfonos seleccionados está defectuoso, ésta se envía. Si se encuentra que uno o más están defectuosos, la caja entera se examina.

- ¿Cuál es la distribución de probabilidad de x , el número de teléfonos defectuosos relacionados con componentes internos que no funcionan en la muestra de $n = 10$ teléfonos?
- ¿Cuál es la probabilidad de que la caja sea enviada si dos de los $N = 50$ teléfonos inteligentes en la caja tienen componentes internos defectuosos?
- ¿Cuál es la probabilidad de que la caja sea enviada si contiene cuatro teléfonos defectuosos? y ¿si tiene seis defectuosos?

3.133 Chocolate oscuro A pesar de informes de que el chocolate oscuro es benéfico para el corazón, 47% de los adultos todavía prefieren el chocolate con leche que el chocolate oscuro.¹⁴ Suponga que una muestra aleatoria de $n = 5$ adultos se selecciona y se les pregunta si prefieren el chocolate con leche que el oscuro.

- ¿Cuál es la probabilidad de que los cinco adultos digan que prefieren el chocolate con leche que el oscuro?

- ¿Cuál es la probabilidad de que exactamente tres de los cinco adultos digan que prefieren el chocolate con leche que el oscuro?

- ¿Cuál es la probabilidad de que al menos un adulto prefiera el chocolate con leche que el oscuro?

3.134 Enfermedad de Tay-Sachs La enfermedad de Tay-Sachs es un mal genético que por lo general es fatal para los niños. Si los dos padres son portadores de la enfermedad, la probabilidad de que sus descendientes la desarrollen es aproximadamente .25. Suponga que un esposo y su esposa son portadores de la enfermedad y que la esposa ha estado embarazada en tres ocasiones. Si la ocurrencia de la enfermedad de Tay-Sachs en uno de los hijos es independiente de la ocurrencia en cualquiera de los otros, ¿cuáles son las probabilidades de estos eventos?

- Los tres hijos desarrollarán la enfermedad de Tay-Sachs.
- Sólo un hijo desarrollará la enfermedad de Tay-Sachs.
- El tercer hijo desarrollará la enfermedad de Tay-Sachs, dado que los primeros dos no la tuvieron.

3.135 La prueba del triángulo Un procedimiento que se usa con frecuencia para controlar la calidad de productos alimenticios de marca utiliza un panel de cinco “degustadores”. Cada miembro del panel prueba tres muestras, dos de las cuales son de lotes del producto que se sabe tiene el sabor deseado y la otra proviene del último lote. Cada degustador selecciona la muestra que es diferente de las otras dos. Suponga que el último lote tiene el sabor deseado y que no hay comunicación entre los degustadores.

- Si el último lote tiene el mismo sabor que los otros dos, ¿cuál es la probabilidad de que el degustador lo elija como el diferente?
- ¿Cuál es la probabilidad de que exactamente uno de los degustadores elija el último lote como el diferente?
- ¿Cuál es la probabilidad de que al menos uno de los degustadores elija el último lote como el diferente?

3.136 Problemas en aves de corral Una investigación preliminar informó que alrededor de 30% de las aves criadas en la localidad estaban infectadas con un parásito intestinal que, aunque no riesgoso para quienes consumen las aves de corral, disminuye los porcentajes acostumbrados de crecimiento en peso de las aves. Se agregó al alimento de las aves un suplemento dietético que se pensaba era eficaz contra este parásito. Se examinaron veinticinco aves después de tomar el suplemento durante por lo menos dos semanas y se encontró que tres aves todavía estaban infectadas con el parásito.

- Si el suplemento dietético es ineficaz, ¿cuál es la probabilidad de observar tres o menos aves infectadas con el parásito intestinal?

- b. Si de hecho el suplemento alimenticio era eficaz y redujo el porcentaje de infección al 10%, ¿cuál es la probabilidad de observar tres o menos aves infectadas?

3.137 Descomposturas de máquinas En una planta de procesamiento y empaque de alimentos, en promedio, dos máquinas empacadoras se descomponen por semana. Suponga que las descomposturas semanales de máquinas siguen una distribución de Poisson.

- a. ¿Cuál es la probabilidad de que no haya descomposturas de máquinas en una semana determinada?
- b. Calcule la probabilidad de que no haya más de dos descomposturas de máquinas en una semana determinada.

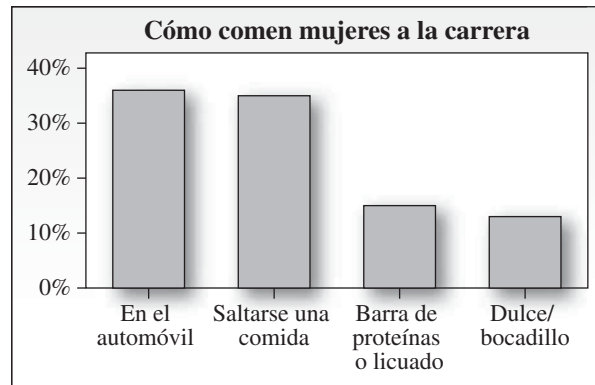
3.138 Estresado A un individuo se le enseña a hacer un trabajo de dos maneras. Estudios realizados han demostrado que cuando es sometido a presión mental y se le pide efectuar el trabajo, el individuo casi siempre vuelve al método que aprendió primero, sin considerar si es más fácil o más difícil. Si la probabilidad de que un sujeto regrese al primer método aprendido es .8 y seis sujetos son probados, ¿cuál es la probabilidad de que al menos cinco de los individuos vuelvan al primer método aprendido cuando se les pida efectúen su trabajo bajo estrés?

3.139 ¡Terremotos! Suponga que uno entre 10 propietarios de casa en el estado de California ha invertido en un seguro contra terremotos. Si 15 propietarios se seleccionan al azar para ser entrevistados,

- a. ¿Cuál es la probabilidad de que al menos uno tenga seguro contra terremotos?
- b. ¿Cuál es la probabilidad de que cuatro o más tengan seguro contra terremotos?
- c. ¿Dentro de qué límites esperaría que baje el número de propietarios asegurados contra terremotos?

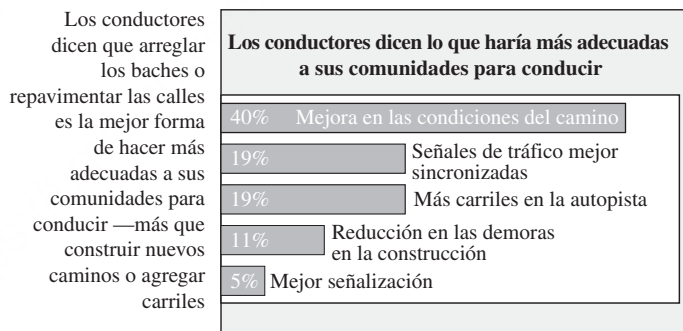
3.140 Comer a la carrera ¿Cómo sobrevive usted cuando no hay tiempo para comer, ya sea un bocado rápido, nada de alimento o una barra de proteína, o dulce? Una encuesta indica que 36% de mujeres entre 25 y 55 años de edad dijeron que, cuando están demasiado ocupadas para comer, adquieren una comida rápida. Se selecciona al azar una muestra de 100 mujeres entre 25 y 55 años de edad.

- a. ¿Cuál es el número promedio de mujeres que dicen que comen comida rápida cuando están demasiado ocupadas para comer?
- b. ¿Cuál es la desviación estándar para el número de mujeres que dicen que comen comida rápida cuando están demasiado ocupadas para comer?



- c. Si 49 de las mujeres de la muestra dijeron que comen comida rápida cuando están demasiado ocupadas para comer, ¿esto sería una ocurrencia poco común? Explique.

3.141 Calles aptas para conducir Según una Instantánea (Snapshot) de *USA Today*, los conductores dicen que arreglar o repavimentar las calles es la mejor forma de hacer que se pueda conducir en sus comunidades —mejor que construir caminos nuevos o agregar carriles.¹⁵ Suponga que se seleccionan al azar $n = 15$ conductores y x es el número de quienes dicen que mejorar las condiciones de los caminos haría más adecuadas a sus comunidades para conducir. Sea $p = .4$ cuando se encuentran probabilidades asociadas con cualquiera de los siguientes resultados:



- a. ¿Cuál es la distribución de probabilidad para x ?
- b. ¿Cuál es $P(x \leq 4)$?
- c. Encuentre la probabilidad de que x exceda de 5.
- d. ¿Cuál es el valor más grande de c para el cual $P(x \leq c) \leq .5$?

3.142 ¿Problemas de crédito? La recesión ha causado que muchas personas usen mucho menos sus tarjetas de crédito. De hecho, en Estados Unidos, 60% de los consumidores dicen que están comprometidos a vivir

con menos tarjetas de crédito.¹⁶ Se selecciona al azar una muestra de $n = 400$ consumidores con tarjetas de crédito.

- ¿Cuál es el número promedio de consumidores en la muestra que dicen estar comprometidos a vivir con menos tarjetas de crédito?
- ¿Cuál es la desviación estándar del número en la muestra de los que dicen que están comprometidos a vivir con menos tarjetas de crédito?
- ¿Dentro de qué rango esperaría encontrar el número en la muestra de los que dicen estar comprometidos a vivir con menos tarjetas de crédito?
- Si sólo 200 de la muestra de consumidores dicen estar comprometidos a vivir con menos tarjetas de crédito, ¿consideraría esto inusual? Explique. ¿Qué conclusión podría sacar de esta información muestral?

3.143 Cirugías exitosas Se dice que un nuevo procedimiento quirúrgico es exitoso en 80% de las veces. Suponga que la operación se efectúa cinco veces y se asume que los resultados son independientes entre sí. ¿Cuáles son las probabilidades de estos eventos?

- Las cinco operaciones son exitosas.
- Exactamente cuatro son exitosas.
- Menos de dos son exitosas.

3.144 Cirugía, continúa Consulte el ejercicio 3.143. Si menos de dos operaciones fueron exitosas, ¿qué impresión tendría usted del trabajo del equipo de cirujanos?

3.145 Falla en un motor Suponga que los cuatro motores de un avión comercial están arreglados para operar de manera independiente y que la probabilidad de falla en vuelo de un solo motor es .01. ¿Cuál es la probabilidad de los siguientes eventos en un vuelo determinado?

- No se observan fallas.
- Se observa no más de una falla.

3.146 Jugar en tragamonedas Una máquina tragamonedas tiene tres ranuras; cada una muestra una cereza, un limón, una estrella o una barra cuando se hace girar. El jugador gana si las tres ranuras muestran los mismos tres objetos. Si es igualmente probable que cada uno de los cuatro objetos aparezca en un giro determinado, ¿cuál es su probabilidad de ganar?

3.147 Sopladores de silbato “Sopladores de silbato” es el nombre que se da a empleados que informan de un fraude corporativo, robo u otras actividades nada éticas y tal vez delictivas cometidas por compañeros de trabajo o por su empleador. Aun cuando hay protección legal para los “sopladores de silbato”, se ha informado que alrededor de 23% de quienes informan de fraude sufren represalias como degradación o calificación de mal rendimiento. Suponga que la probabilidad de que un empleado no

informe de un caso de fraude es .69. Encuentre la probabilidad de que un trabajador que observe un caso de fraude informe del mismo y después sufra alguna forma de represalia.

3.148 Aspirina Dos tabletas para resfriado se colocan accidentalmente en una caja que contiene dos aspirinas. Las cuatro tabletas son idénticas en apariencia. Se selecciona al azar una tableta de la caja y es tomada por el primer paciente. A continuación se selecciona una tableta al azar de las tres tabletas restantes y es tomada por el segundo paciente. Defina los siguientes eventos como conjuntos específicos de eventos simples:

- El espacio muestral S .
- El evento A que el primer paciente obtuvo una tableta para resfriado.
- El evento B de que exactamente uno de los dos pacientes obtuvo una tableta para resfriado.
- El evento C de que ningún paciente obtuvo una tableta para resfriado.

3.149 Consulte el ejercicio 3.148. Resumiendo las probabilidades de eventos simples, encuentre $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup B)$, $P(C)$, $P(A \cap C)$, y $P(A \cup C)$.

3.150 Reproductores de discos de video Un minorista vende dos estilos de grabadoras de video digitales (DVR) de alto precio que la experiencia indica que tienen igual demanda (50% de todos los potenciales compradores prefieren el estilo 1 y 50% están a favor del estilo 2). Si el minorista tiene en existencia cuatro de cada uno, ¿cuál es la probabilidad de que los cuatro primeros clientes que buscan una DVR compren el mismo estilo?

3.151 Calcule el área bajo la curva normal estándar a la izquierda de estos valores:

- $z = -.90$
- $z = 2.34$
- $z = 5.4$

3.152 Calcule el área bajo la curva normal estándar entre estos valores:

- $z = -2.0$ y $z = 2.0$
- $z = -2.3$ y $z = -1.5$

3.153 Encuentre las siguientes probabilidades para la variable aleatoria normal estándar z :

- $P(-1.96 \leq z \leq 1.96)$
- $P(z > 1.96)$
- $P(z < -1.96)$

3.154 a. Encuentre una z_0 tal que $P(z > z_0) = .9750$.

b. Encuentre una z_0 tal que $P(z > z_0) = .3594$.

3.155 a. Encuentre una z_0 tal que $P(-z_0 \leq z \leq z_0) = .95$.

b. Encuentre una z_0 tal que $P(-z_0 \leq z \leq z_0) = .98$.

c. Encuentre una z_0 tal que $P(-z_0 \leq z \leq z_0) = .90$.

d. Encuentre una z_0 tal que $P(-z_0 \leq z \leq z_0) = .99$.

3.156 Una variable aleatoria normal x tiene una media $\mu = 5$ y una desviación estándar $\sigma = 2$. Encuentre las probabilidades asociadas con los siguientes intervalos:

- a. $1.2 < x < 10$ b. $x > 7.5$ c. $x \leq 0$

3.157 Sea x una variable aleatoria binomial con $n = 36$ y $p = .54$. Use la aproximación normal para hallar:

- a. $P(x \leq 25)$ b. $P(15 \leq x \leq 20)$ c. $P(x > 30)$

3.158 Use la tabla 3 del Anexo para calcular el área bajo la curva normal estándar a la izquierda de lo siguiente:

- a. $z = 1.2$ b. $z = -.9$
c. $z = 1.46$ d. $z = -.42$

3.159 Encuentre las siguientes probabilidades para la variable aleatoria normal estándar:

- a. $P(.3 < z < 1.56)$ b. $P(-.2 < z < .2)$

3.160 a. Encuentre la probabilidad de que z sea mayor que $-.75$.

b. Encuentre la probabilidad de que z sea menor que 1.35 .

3.161 Encuentre z_0 tal que $P(z > z_0) = .5$.

3.162 Encuentre la probabilidad de que z se encuentre entre $z = -1.48$ y $z = 1.48$.

3.163 Encuentre z_0 tal que $P(-z_0 < z < z_0) = .5$. ¿Cuáles percentiles representan $-z_0$ y z_0 ?

3.164 Barrenas Se estima que la vida media de las barrenas perforadoras de pozos petroleros es de 75 horas. Suponga que una compañía de exploración petrolera compra barrenas, que tienen una vida útil que está distribuida normalmente en forma aproximada, con una media igual a 75 horas y una desviación estándar igual a 12 horas.

- a. ¿Qué proporción de barrenas de la compañía fallarán antes de 60 horas de uso?
b. ¿Qué proporción durará al menos 60 horas?
c. ¿Qué proporción tendrá que cambiarse después de más de 90 horas de uso?

3.165 Diámetros de cojinetes La operación de una máquina produce cojinetes cuyos diámetros están normalmente distribuidos, con media y desviación estándar igual a .498 y .002, respectivamente. Si las especificaciones requieren que el diámetro del cojinete sea igual a .500 de pulgada $\pm .004$ de pulgada, ¿qué fracción de la producción sería inaceptable?

3.166 Control remoto de puertas de garaje Casi todos los usuarios de controles remotos de puertas de garaje activan sus controles a distancias que están normalmente distribuidas, con una media de 30 pies y una desviación estándar de 11 pies. Para reducir al mínimo la interferencia

con otros aparatos a control remoto, se especifica al fabricante que limite la distancia de operación a 50 pies. ¿Qué porcentaje del tiempo tratarán los usuarios de operar el control fuera de su límite de operación?

3.167 ¿Cuánto dura el examen? Se ha encontrado que la duración promedio de tiempo, requerida para completar un examen de conocimientos en una universidad, es igual a 70 minutos con una desviación estándar de 12 minutos. ¿Cuándo debe terminarse el examen si se desea permitir tiempo suficiente para que 90% de los estudiantes lo completen? (Suponga que el tiempo necesario para completar el examen está normalmente distribuido.)

3.168 Servicio a automóviles El tiempo necesario para correr una prueba de 5 000 millas y dar servicio a un automóvil tiene una media igual a 1.4 horas y una desviación estándar de .7 horas. Suponga que el departamento de servicio planea dar servicio a 50 automóviles por día de 8 horas y que, para hacerlo, no debe emplear más de un promedio de 1.6 horas por automóvil. ¿Qué proporción de todos los días tendrá que trabajar tiempo extra el departamento de servicio?

3.169 Televidentes Una agencia de publicidad ha expresado que 20% de todos los televidentes ven un programa particular. En una muestra aleatoria de 1000 televidentes, $x = 184$ estaban viendo el programa. ¿Estos datos presentan suficiente evidencia para contradecir lo dicho por la agencia de publicidad?

3.170 Llenar vasos de refresco Una máquina que expende refrescos puede ser regulada para descargar un promedio de μ onzas por vaso. Si las onzas de líquido están normalmente distribuidas, con desviación estándar igual a .3 de onza, dé el ajuste para μ de modo que vasos de 8 onzas ($\frac{1}{4}$ de litro) se rebosen sólo 1% de las veces.

3.171 Bombillas eléctricas Una planta fabricante utiliza bombillas eléctricas cuyas duraciones están normalmente distribuidas, con media y desviación estándar igual a 500 y 50 horas, respectivamente. Para reducir al mínimo el número de bombillas que se queman durante las horas de operación, todas las bombillas se cambian después de un periodo determinado de operación. ¿Con qué frecuencia deben cambiarse las bombillas si deseamos que no más de 1% de ellas se quemen entre periodos de cambio?

3.172 No presentadas Una línea aérea encuentra que 5% de las personas que hacen reservaciones en cierto vuelo no se presentan para el vuelo. Si la aerolínea vende 160 boletos para un vuelo que tiene sólo 155 asientos, ¿cuál es la probabilidad de que un asiento esté disponible para toda persona que tenga una reservación y planea volar?

3.173 Genética en plantas En el ejercicio 3.128, una cruce entre dos plantas, una de ellas con pétalos rojos y una con pétalos veteados, produjeron plantas descendientes con pétalos rojos 75% de las veces. Suponga que 100 semillas de esta cruce se recolectaron y germinaron, y x , el número de plantas con pétalos rojos, se registró.

- ¿Cuál es la distribución de probabilidad exacta para x ?
- ¿Es apropiado aproximar la distribución de la parte a usando la distribución normal? Explique.
- Use un método apropiado para hallar la probabilidad aproximada de que entre 70 y 80 (inclusive) plantas descendientes tengan flores rojas.
- ¿Cuál es la probabilidad de que 53 o menos de plantas descendientes tengan flores rojas? ¿Es éste un suceso poco común?
- Si en realidad se observaron 53 de 100 plantas descendientes con flores rojas y si estábamos seguros que la proporción genética 3:1 era correcta, ¿qué otra explicación podría darse para este suceso poco común?

3.174 ¿Proveedores A o B? Un comprador de relevadores eléctricos hace compras a dos proveedores, A y B. El proveedor A suministra dos de cada tres relevadores empleados por la compañía. Si se seleccionan al azar 75 relevadores de los empleados por la compañía, encuentre la probabilidad de que a lo sumo 48 de esos relevadores provengan del proveedor A. Suponga que la compañía utiliza un gran número de relevadores.

3.175 Bocado y TV Los psicólogos creen que comer en exceso puede estar asociado con estados emocionales (estar molesto o aburrido) e indicios ambientales (ver televisión, leer, etcétera). Para probar esta teoría, suponga que se han seleccionado al azar 60 personas y emparejadas en peso y género. Durante 2 semanas, se pidió a un miembro de cada pareja que pasara las tardes leyendo novelas de su interés, mientras el otro miembro pasa las tardes viendo televisión. Se registró la cantidad de calorías de todos los bocadillos y refrescos consumidos por las tardes por cada persona, siendo $x = 19$ el número de parejas para las cuales la ingesta de calorías de los televidentes fue mayor que la de los lectores. Si no hay diferencia en los efectos de televisión y lectura en ingesta de calorías, la probabilidad p de que la ingesta de calorías de un miembro de cada pareja exceda la del otro miembro es .5. ¿Estos datos dan suficiente evidencia para indicar una diferencia entre los efectos de ver televisión y leer en cuanto a ingesta de calorías se refiere? (SUGERENCIA: Calcule el puntaje z para el valor observado, $x = 19$.)

3.176 Tiempos de gestación *The Biology Data Book* informa que el tiempo de gestación para bebés humanos promedia 278 días, con una desviación estándar de 12 días.¹⁷ Suponga que estos tiempos de gestación están normalmente distribuidos.

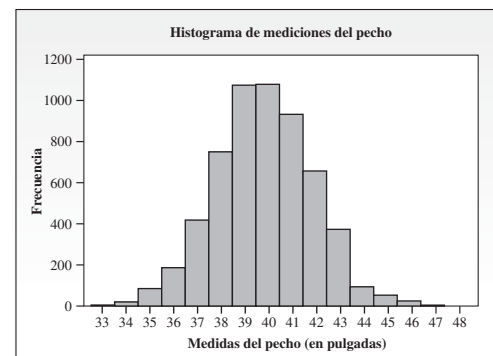
- Encuentre los cuartiles superior e inferior para los tiempos de gestación.
- ¿Sería poco común tener un bebé después de sólo 6 meses de gestación? Explique.

3.177 ¿Introvertido o extrovertido? Un examen psicológico para introvertidos-extrovertidos produjo calificaciones que tenían una distribución normal con una media y una desviación estándar de 75 y 12, respectivamente. Si deseamos designar el *más alto* 15% a extrovertidos, ¿cuál sería la calificación apropiada para elegir como el punto de corte?

3.178 ¿Distribución normal? Las medidas del pecho de 5738 milicianos escoceses de principios del siglo XIX se dan a continuación.¹⁸ El tamaño del pecho se mide en pulgadas, y cada observación reporta el número de soldados con ese tamaño de pecho.

Número de soldados	Pecho	Número de soldados	Pecho
3	33	934	41
18	34	658	42
81	35	370	43
185	36	92	44
420	37	50	45
749	38	21	46
1073	39	4	47
1079	40	1	48

Observe la normalidad aproximada del histograma de las 5738 mediciones de pecho.



- La media de esta distribución es $\bar{x} = 39.83$ y la desviación estándar es $s = 2.05$. ¿Cuál es el 95o.

percentil de esta distribución con base en una curva normal con $\mu = 39.83$ y $\sigma = 2.05$?

- b. Encuentre la estimación empírica del 95o. percentil y compárela con su respuesta en la parte a. (SUGERENCIA: El 95o. percentil estará en la posición $.95(n + 1) = .95 \times 5739 = 5452.05$ de la cola izquierda de la distribución o en la posición $5738 - 5452.05 = 285.95$ de la cola derecha de la distribución)
- c. Encuentre el 90o. percentil de esta distribución con base en una curva normal con $\mu = 39.83$ y $\sigma = 2.05$. ¿Cuál es el valor del 90o. percentil empírico? ¿Cómo se compara con el valor asumiendo normalidad?

3.179 ¿Distribución normal?, continúa Suponga que las mediciones del pecho en el ejercicio 3.178 están distribuidas normalmente con una media de $\mu = 39.83$ y una desviación estándar de $\sigma = 2.05$.

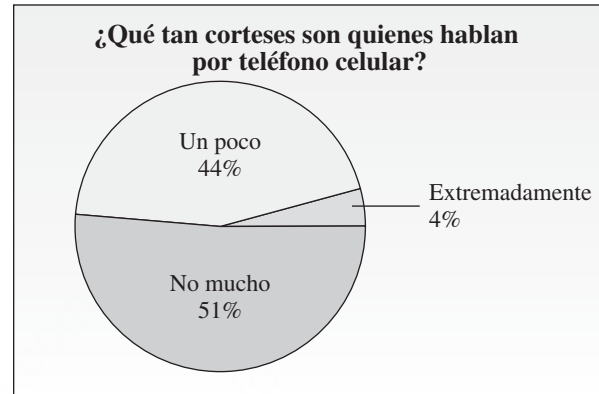
- a. ¿Qué proporción de las observaciones caería entre 36.5 y 43.5 pulgadas?
- b. ¿Entre cuáles dos mediciones caería 95% de las observaciones?
- c. ¿Cuáles son las proporciones reales para las partes a y b usando los datos en forma directa? Comente sobre la precisión de las proporciones encontradas usando la suposición de normalidad de las mediciones del pecho.

3.180 Temperaturas normales Se deduce una distribución de temperaturas del cuerpo humano, que tiene una forma de montículo distintiva. Suponga que asumimos que las temperaturas de personas sanas son aproximadamente normales, con una media de 98.6 grados y una desviación estándar de 0.8 grados.

- a. Si se selecciona al azar una persona sana, ¿cuál es la probabilidad de que la persona tenga una temperatura arriba de 99.0°?
- b. ¿Cuál es el 95o. percentil para las temperaturas corporales de personas sanas?

3.181 Etiqueta de teléfono celular Una encuesta en una universidad indica que 51% de los estudiantes dicen

que la persona promedio no es muy considerada hacia otros cuando habla por un teléfono celular.¹⁹ Suponga que se seleccionan al azar 100 estudiantes. Encuentre la probabilidad aproximada de que 60 estudiantes o más indicarían que la persona promedio no es muy considerada con los demás cuando habla por teléfono celular.



3.182 Calificaciones de examen Las calificaciones en un examen de conocimientos a nivel nacional estuvieron distribuidas normalmente en forma aproximada, con una media de 540 y una desviación estándar de 110.

- a. Si usted recibió una calificación de 680, ¿hasta qué punto, en desviaciones estándar, su calificación se apartó de la media?
- b. ¿Qué porcentaje de quienes tomaron el examen obtuvieron una calificación más alta que usted?

3.183 Trasplante de células Briggs y King inventaron la técnica de trasplante nuclear, en el que el núcleo de una célula de una de las etapas finales del desarrollo de un embrión es trasplantado en un cigoto (un huevo fertilizado unicelular) para ver si el núcleo puede soportar un desarrollo normal. Si la probabilidad de que un solo trasplante de la primera etapa de gástrula sea exitoso es .65, ¿cuál es la probabilidad de que más de 70 trasplantes de entre 100 sean exitosos?

CASO PRÁCTICO

 Un misterio: casos de cáncer cerca de un reactor

¿Qué tan seguro es vivir cerca de un reactor nuclear? Hombres que vivían en una franja costera que se extiende 20 millas al norte de un reactor nuclear en Plymouth, Massachusetts, desarrollaron algunas formas de cáncer a un ritmo 50% mayor que el de todo el estado, de acuerdo con un estudio avalado por el Departamento de Salud Pública de Massachusetts y publicado el 21 de mayo de 1987 en el *New York Times*.

La causa de los casos de cáncer es un misterio, pero se ha sugerido que este padecimiento estaba vinculado al reactor Pilgrim I, que había sido cerrado durante 13 meses debido a problemas de administración. Boston Edison, propietaria del reactor, reconoció escapes de radiación a mediados del decenio de 1970 que estuvieron un poco arriba de los niveles permisibles. Si el reactor era en efecto responsable por el excesivo porcentaje de cáncer, entonces el nivel de radiación reconocido actualmente requerido para causar cáncer debería cambiar. No obstante, el misterio más desconcertante fue el hecho de que las mujeres en esa misma zona aparentemente no fueron afectadas.

En su informe, el doctor Sidney Cobb, epidemiólogo, observó la conexión entre los escapes de radiación en el reactor Pilgrim I y 52 casos de cáncer hematopoyético. El informe indicaba que este número inesperadamente grande podría ser atribuible a escapes radiactivos en el aire salido del Pilgrim I, concentrado a lo largo de la costa por los patrones de viento y no disipados, como supusieron inspectores de reglamentos del gobierno. ¿Qué tan poco común fue este número de casos de cáncer? Esto es, estadísticamente hablando, ¿52 es un número altamente improbable de casos? Si la respuesta es afirmativa, entonces algún factor externo (posiblemente radiación) causó este número anormalmente grande, o bien, ¡hemos observado un evento muy raro!

La distribución de probabilidad de Poisson da una buena aproximación a las distribuciones de variables; por ejemplo, el número de fallecimientos en una región debido a una rara enfermedad, el número de accidentes en una planta manufacturera por mes, o el número de choques de líneas aéreas por mes. Por tanto, es razonable suponer que la distribución de Poisson da un modelo apropiado para el número de casos de cáncer en este ejemplo.

1. Si los 52 casos publicados representaban un porcentaje 50% más alto que el porcentaje a nivel del estado, ¿cuál es una estimación razonable de μ , el número promedio de esos casos de cáncer a nivel de todo el estado?
2. Con base en su estimación respecto a μ , ¿cuál es la desviación estándar estimada del número de casos de cáncer en todo el estado?
3. ¿Cuál es el puntaje z para los $x = 52$ casos observados de cáncer? ¿Cómo interpreta este puntaje z en vista de la preocupación por el elevado porcentaje de cáncer hematopoyético en esta zona?

CASO PRÁCTICO



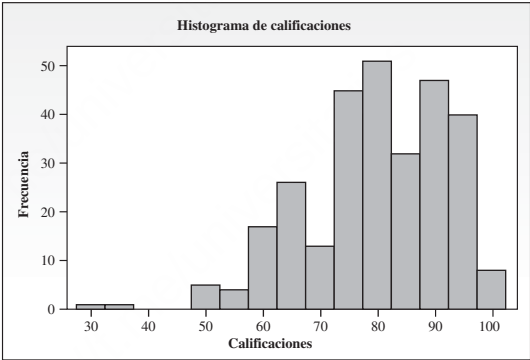
“¿Va a calificar por curva?”

Muy a menudo, al final de un examen que pareció particularmente difícil, los estudiantes preguntan al profesor: “¿Va a calificar por curva?”. Por desgracia, “calificar por curva” no necesariamente significa que recibirá una calificación más alta en un examen, ¡aunque podría gustarle pensar que sí! Calificar por curva en realidad es una técnica por medio de la cual una proporción fija de las calificaciones más altas reciben una A (aun si la calificación más alta es una calificación reprobatoria con base en porcentaje), y una proporción fija de las calificaciones más bajas recibe una F (aun si la puntuación más baja es una calificación aprobatoria con base en porcentaje). Las calificaciones de B, C y D también se asignan de acuerdo con una proporción fija. Una de dichas asignaciones usa las siguientes proporciones.

Calificación por letra	A	B	C	D	F
Proporción de calificaciones	10%	20%	40%	20%	10%

- 1. Si la calificación de C promedio está centrada en la calificación promedio para todos los estudiantes, y si suponemos que las calificaciones están distribuidas normalmente, ¿cuántas desviaciones estándar en cada lado de la media designarían a las calificaciones de C?
- 2. ¿Cuántas desviaciones estándar en cualquier lado de la media serán los puntos límite para las calificaciones de B y D?

A continuación se presenta un histograma de las calificaciones para un curso de introducción a la estadística junto con los estadísticos de resumen.



Estadísticos descriptivos: Calificaciones

Variable	N	N*	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Grades	290	0	79.972	12.271	31.000	73.000	82.000	88.000	100.000

Para facilidad de cálculo, redondee el número de desviaciones estándar para las calificaciones de C a $\pm .5$ desviaciones estándar y para las calificaciones de B y D a ± 1.5 desviaciones estándar.

- 3. Encuentre los puntos límite para las calificaciones de A, B, C, D y F correspondientes a estos valores redondeados.
- 4. Si usted obtuvo una puntuación de 92 en el examen y tuvo la opción de calificar por curva o usar un estándar absoluto de 90–100 para una A, 80–89 para una B, 70–79 para una C, y así en forma sucesiva, ¿cuál sería su elección? Explique su razonamiento. ¿Es un problema la asimetría de la distribución de las calificaciones?

Distribuciones muestrales

OBJETIVOS GENERALES

En los capítulos anteriores estudiamos *poblaciones* y los *parámetros* que las describen. Estas poblaciones eran discretas o continuas y empleamos la *probabilidad* como herramienta para determinar qué tan probables podrían ser ciertos resultados muestrales. En este capítulo, nuestro interés cambia cuando empezamos a estudiar *muestras* y las *estadísticas* que las describen. Estas estadísticas muestrales se usan para hacer inferencias acerca de los correspondientes parámetros de población. Este capítulo comprende muestreo y distribuciones muestrales, que describen el comportamiento de estadísticas muestrales en muestreo repetido.

ÍNDICE DEL CAPÍTULO

- Introducción (4.1)
- Planes muestrales y diseños experimentales (4.2)
- Estadísticas y distribuciones muestrales (4.3)
- El teorema central del límite (4.4)
- La distribución muestral de la media muestral (4.5)
- La distribución muestral de la proporción muestral (4.6)
- Una aplicación muestral: control estadístico de procesos (opcional) (4.7)



NECESITO SABER...

Cuándo el tamaño de la muestra es lo bastante grande para usar el teorema del límite central
 Cómo calcular probabilidades para la media muestral $\bar{x}$
 Cómo calcular probabilidades para la proporción muestral $\hat{p}$

Muestreo de la Ruleta de Monte Carlo

¿Le gustaría probar su mano en un juego sin correr el riesgo de perder? Usted podría hacerlo simulando el proceso de los juegos de azar, hacer apuestas imaginarias y observar los resultados. Esta técnica, llamada método de Monte Carlo, es el tema del caso práctico al final de este capítulo.

4.1

INTRODUCCIÓN

TIP

MI CONSEJO

Parámetro $\Leftrightarrow$ Población
Estadística $\Leftrightarrow$ Muestra

En los capítulos 2 y 3, usted ha aprendido mucho acerca de distribuciones de probabilidad; por ejemplo, las distribuciones binomial y normal. La forma de la distribución normal está determinada por su media μ y su desviación estándar σ , mientras que la forma de la distribución binomial está determinada por p . Estas medidas numéricas descriptivas, llamadas **parámetros**, son necesarias para calcular la probabilidad de observar resultados muestrales.

En situaciones prácticas, usted puede decidir qué *tipo* de distribución de probabilidad usar como modelo, pero los valores de los *parámetros* que especifican su *forma exacta* se desconocen. A continuación veamos dos ejemplos:

- Un entrevistador está seguro de que las respuestas a sus preguntas “de acuerdo/en desacuerdo” seguirán una distribución binomial, pero se desconoce p , la proporción de quienes están “de acuerdo” de la población.
- Un agrónomo cree que la producción por acre de una variedad de trigo está distribuida normalmente en forma aproximada, pero se desconocen la media μ y desviación estándar σ de la producción.

En estos casos, debemos apoyarnos en la *muestra* para saber de estos parámetros. La proporción de quienes están “de acuerdo” en la muestra del entrevistador da información acerca del valor real de p . La media y desviación estándar de la muestra del agrónomo aproximan los valores reales de μ y de σ . Pero, si se desea que la muestra dé *información confiable* acerca de la población, ¡la muestra debe ser seleccionada en cierta forma!

4.2

PLANES MUESTRALES Y DISEÑOS EXPERIMENTALES

La forma en que una muestra se selecciona recibe el nombre de **plan muestral** o **diseño experimental** y determina la cantidad de información de la muestra. Saber el plan muestral empleado en una situación particular permitirá medir la confiabilidad o bondad de la inferencia.

El **muestreo aleatorio simple** es un plan muestral de uso común en el que cada muestra de tamaño n tiene la misma probabilidad de ser seleccionada. Por ejemplo, supongamos que se desea seleccionar una muestra de tamaño $n = 2$ de una población que contiene $N = 4$ objetos. Si los cuatro objetos están identificados por los símbolos x_1, x_2, x_3 y x_4 , hay seis pares distintos que podrían seleccionarse, como se muestra en la tabla 4.1. Si la muestra de $n = 2$ observaciones se selecciona de modo que cada una de estas seis muestras tenga la misma probabilidad de selección, dada por $1/6$, entonces la muestra resultante se denomina **muestra aleatoria simple** o únicamente **muestra aleatoria**.

TABLA 4.1

Formas de seleccionar una muestra de tamaño 2 de 4 objetos

Muestra	Observaciones en la muestra
1	x_1, x_2
2	x_1, x_3
3	x_1, x_4
4	x_2, x_3
5	x_2, x_4
6	x_3, x_4

Definición Si una muestra de n elementos se selecciona de entre una población de N elementos, usando un plan muestral en el que cada una de las posibles muestras tiene la misma probabilidad de selección, entonces se dice que el muestreo es **aleatorio** y la muestra resultante es una **muestra aleatoria simple**.

Un muestreo aleatorio perfecto es difícil de obtener en la práctica. Si el tamaño N de la población es pequeño se podría escribir cada uno de los N números en una ficha, mezclar las fichas y seleccionar una muestra de n fichas. Los números que seleccione corresponden a las n mediciones que aparecen en la muestra. Como este método no siempre es práctico, un método más sencillo y confiable utiliza **números aleatorios**, es decir, dígitos generados de modo que los valores de 0 a 9 se presentan al azar y con igual frecuencia. Estos números pueden ser generados por computadora o incluso aparecer en una calculadora científica. De manera opcional, la tabla 10 del Anexo es una tabla de números aleatorios que se pueden usar para seleccionar una *muestra aleatoria*.

EJEMPLO

4.1

Una base de datos de computadora en una firma de abogados contiene archivos para $N = 1\,000$ clientes. La firma desea seleccionar $n = 5$ archivos para revisión. Seleccione una muestra aleatoria simple de cinco archivos de esta base de datos.

Solución Primero debe marcar cada archivo con un número de 1 a 1 000. Quizá los archivos se guarden alfabéticamente y la computadora ya ha asignado un número a cada uno. A continuación genere una secuencia de 10 números aleatorios de tres dígitos. Si usa la tabla 10 del Anexo, seleccione un punto inicial aleatorio y use una parte de la tabla similar a la de la tabla 4.2. El punto inicial aleatorio asegura que usted no use la misma secuencia una y otra vez.

Los primeros tres dígitos de la tabla 4.2 indican el número del primer archivo a revisarse. El número aleatorio 001 corresponde al archivo #1 y el último archivo, #1000 corresponde al número aleatorio 000. Usando la tabla 4.2, se seleccionarían los cinco archivos numerados 155, 450, 32, 882 y 350 para revisión. De manera opcional, podría elegir leer en sentido horizontal las líneas y seleccionar los archivos 155, 350, 989, 450 y 369 para revisión.

TABLA 4.2

Parte de una tabla de números aleatorios

15574	35026	98924
45045	36933	28630
03225	78812	50856
88292	26053	21121

La situación descrita en el ejemplo 4.1 se denomina **estudio observacional**, porque los datos ya existían antes que usted decidiera *observar* o describir sus características. La mayoría de los estudios muestrales, en los que la información se capta con un cuestionario, caen en esta categoría. Las bases de datos de computadora hacen posible asignar números de identificación a cada elemento aun cuando la población sea grande y seleccionar una muestra aleatoria simple. Tenga cuidado al efectuar un estudio muestral y esté atento a estos problemas que se presentan con frecuencia:

- **No respuesta:** Usted ha seleccionado su muestra aleatoria y enviado sus cuestionarios, pero sólo 50% de los entrevistados devolvió sus cuestionarios. ¿Las respuestas que usted recibió son representativas de toda la población o están **sesgadas** porque sólo quienes eran particularmente tendenciosos en el tema fueron seleccionados para responder?

- **Cobertura demasiado baja:** Usted ha seleccionado su muestra aleatoria usando registros telefónicos como una base de datos. ¿La base de datos que usó sistemáticamente excluye ciertos segmentos de la población, quizá aquellos que usan sólo celulares o que sus números son privados?
- **Sesgo verbal:** Su cuestionario puede tener preguntas que son demasiado complicadas o tienden a confundir al lector. Posiblemente las preguntas son sensibles por naturaleza, por ejemplo, “¿Alguna vez ha consumido usted drogas?” o “¿Alguna vez ha engañado en su declaración de impuestos?”, y quienes responden no contestan con la verdad.

Se han diseñado métodos para resolver algunos de estos problemas, pero sólo si usted sabe que existen. Si su encuesta está *sesgada* por cualquiera de estos problemas, entonces sus conclusiones no serán muy confiables, ¡aunque haya seleccionado una muestra aleatoria!

Alguna investigación realizada comprende la **experimentación**, en la que una condición experimental o *tratamiento* se impone en las *unidades experimentales*. Seleccionar una muestra aleatoria simple es más difícil en esta situación.

EJEMPLO

4.2

Una química investigadora somete a prueba un nuevo método para medir la cantidad de titanio (Ti) en muestras de mineral. Selecciona 10 muestras de mineral del mismo peso para su experimento. Cinco de las muestras se medirán con un método estándar y las otras cinco con el nuevo método. Use números aleatorios para asignar las 10 muestras de mineral a los grupos nuevo y estándar. ¿Estos datos representan una muestra aleatoria simple de entre la población?

Solución Hay realmente dos poblaciones en este experimento. Están formadas por mediciones de titanio, usando ya sea el método nuevo o el estándar, para *todas las posibles* muestras de mineral de este peso. Estas poblaciones no existen en realidad; son **poblaciones hipotéticas**, imaginadas en la mente de la investigadora. Entonces, es imposible seleccionar una muestra aleatoria simple usando los métodos del ejemplo 4.1. En cambio, la investigadora selecciona lo que cree que son 10 muestras *representativas* de mineral y espera que éstas se *comportarán como si* se hubieran seleccionado al azar de las dos poblaciones.

La investigadora puede, no obstante, seleccionar al azar las cinco muestras a medir con cada método. Numere las muestras del 1 al 10. Las cinco muestras seleccionadas para el nuevo método pueden corresponder a cinco números aleatorios de un dígito. Use esa secuencia de dígitos aleatorios generados en una calculadora científica:

948247817184610

Como no se puede seleccionar dos veces la misma muestra de mineral hay que saltarse cualquier dígito que ya haya sido elegido. Las muestras 9, 4, 8, 2 y 7 se medirán usando el nuevo método. Las otras muestras, es decir 1, 3, 5, 6 y 10, se medirán con el método estándar.

Además del *muestreo aleatorio simple*, hay otros planes muestrales con carácter aleatorio y, por tanto, dan una base probabilística para hacer inferencias. Tres de esos planes están basados en *muestreo estratificado*, *conglomerado* y *sistemático*.

Cuando la población está formada por dos o más subpoblaciones, llamadas **estratos**, un plan muestral que asegura que cada subpoblación está representada en la muestra se denomina **muestra aleatoria estratificada**.

Definición Un **muestreo aleatorio estratificado** comprende seleccionar una muestra aleatoria simple de cada uno de un número dado de subpoblaciones o **estratos**.

Las opiniones de la ciudadanía acerca de la construcción de un centro de artes interpretativas podrían ser recolectadas usando una muestra aleatoria estratificada con distritos de votación como estratos. Los sondeos nacionales comprenden alguna forma muestral aleatoria estratificada con los estados como estratos.

Se utiliza otra forma muestral aleatoria cuando las unidades muestrales disponibles son grupos de elementos, llamados **conglomerados**. Por ejemplo, una familia es un *conglomerado* de personas que viven juntas. Una manzana o vecindario de una ciudad podrían ser una unidad muestral conveniente y ser considerada un *conglomerado* para un plan muestral determinado.

Definición Una **muestra de conglomerados** es una muestra aleatoria simple tomada de los conglomerados disponibles en la población.

Cuando un conglomerado particular se incluye en la muestra, se toma un censo de cada uno de los elementos del conglomerado.

En ocasiones la población que será muestreada está ordenada, por ejemplo, una lista alfabética de personas con licencias de manejo, una lista de usuarios de la compañía de luz ordenada por direcciones de servicio o una lista de clientes por números de cuenta. En éstas y otras situaciones, se elige al azar un elemento de los primeros k elementos y, a continuación, cada k -ésimo elemento de ahí en adelante se incluye en la muestra.

Definición Una **muestra aleatoria sistemática 1 en k** involucra la selección aleatoria de uno de los primeros k elementos de una población ordenada y luego la selección sistemática de cada k -ésimo elemento de ahí en adelante.

TIP MI CONSEJO

Todos los planes muestrales empleados para hacer inferencias comprenden el *proceso aleatorio*

Sin embargo, no todos los planes muestrales comprenden una selección aleatoria. Este tipo muestral es una forma de una **muestra de conveniencia**, es decir, una muestra que se puede obtener de manera fácil y sencilla sin selección aleatoria. Hacer publicidad a personas a quienes se les pagará una cuota por participar en un experimento produce una muestra de conveniencia. El **muestreo de juicio** permite que la persona que haga el muestreo decida quién estará o no incluido en la muestra. El **muestreo de cuota**, en el que la composición de la muestra debe reflejar la composición de la población en alguna característica preseleccionada, con frecuencia tiene un componente no aleatorio en el proceso de selección. **Recuerde que las muestras no aleatorias se pueden describir pero no deben usarse para hacer inferencias.**

4.2 EJERCICIOS

TÉCNICAS BÁSICAS

4.1 Una población está formada por $N = 500$ unidades experimentales. Use una tabla numérica aleatoria para seleccionar una muestra aleatoria de $n = 20$ unidades experimentales. (SUGERENCIA: Como es necesario usar números de tres dígitos, asigne dos números de tres dígitos a cada una de las unidades muestrales en la forma ilustrada en la tabla.) ¿Cuál es la probabilidad de que cada unidad experimental sea seleccionada para su inclusión en la muestra?

Unidades experimentales	Números aleatorios
1	001, 501
2	002, 502
3	003, 503
4	004, 504
.	.
.	.
.	.
499	499, 999
500	500, 000

4.2 Un analista político desea seleccionar una muestra de $n = 20$ personas de una población de 2000. Use la tabla numérica aleatoria para identificar a las personas que se incluirán en la muestra.

4.3 Una población contiene 50 000 electores. Use la tabla numérica aleatoria para identificar a los votantes que se incluirán en una muestra aleatoria de $n = 15$.

4.4 Una pequeña ciudad contiene 20 000 electores. Use la tabla numérica aleatoria para identificar a los electores que se incluirán en una muestra aleatoria de $n = 15$.

4.5 Cada décima persona Se obtuvo una muestra aleatoria de opinión pública, en una pequeña ciudad, al seleccionar cada décima persona que pasara por la esquina de mayor movimiento en el centro de la ciudad. ¿Esta muestra tendrá las características de una muestra aleatoria seleccionada de los ciudadanos? Explique.

APLICACIONES

4.6 Ácidos grasos omega 3 Contrario a la opinión actual sobre los ácidos grasos omega 3, una nueva investigación muestra que las grasas benéficas quizá no

ayuden a reducir segundos infartos en los sobrevivientes que han sufrido uno. El estudio incluyó 4 837 hombres y mujeres que recibían tratamiento por enfermedad cardíaca. El grupo experimental recibió 400 mg adicionales diarios de las grasas.¹ Suponga que este experimento se repitió con 50 individuos en el grupo de control y 50 individuos en el grupo experimental. Determine un esquema de aleatorización para asignar los 100 individuos a los dos grupos.

4.7 Tai chi y fibromialgia Un nuevo estudio pequeño muestra que el tai chi, una práctica china antigua de ejercicio y meditación, puede aliviar síntomas de fibromialgia crónica dolorosa. El estudio asignó 66 pacientes con fibromialgia para que tomaran una clase de tai chi durante 12 semanas, o asistieran a una clase de educación para la salud.²

- Proporcione un esquema de aleatorización para asignar los 66 sujetos a los dos grupos.
- ¿Su esquema de aleatorización producirá grupos de tamaño igual? Explique.

ESTADÍSTICAS Y DISTRIBUCIONES MUESTRALES

4.3

Cuando se selecciona una muestra aleatoria de una población las medidas numéricas descriptivas que se calculen de la muestra se denominan **estadísticas**. Las estadísticas varían o cambian para cada muestra aleatoria diferente que se seleccione; esto es, son *variables aleatorias*. Las distribuciones de probabilidad para estadísticas se llaman **distribuciones muestrales** porque, en muestreos repetidos, dan esta información:

- Qué valores de la estadística pueden presentarse.
- Con qué frecuencia se presenta cada valor.

Definición La **distribución muestral de una estadística** es la distribución de probabilidad para los posibles valores de la estadística, que resulta cuando muestras aleatorias de tamaño n se sacan repetidamente de la población.

Hay tres formas de hallar la distribución muestral de una estadística:

- Deducir la distribución *matemáticamente* usando las leyes de probabilidad.
- Usar una *simulación* para aproximar la distribución. Esto es, sacar un gran número de muestras de tamaño n , calculando el valor de la estadística para cada muestra y tabular

los resultados en un histograma de frecuencia relativa. Cuando el número de muestras es grande, el histograma será muy cercano a la distribución muestral teórica.

3. Usar *teoremas estadísticos* para obtener distribuciones muestrales exactas o aproximadas.

El siguiente ejemplo demuestra cómo deducir las distribuciones muestrales de dos estadísticas para una población muy pequeña.

EJEMPLO

4.3

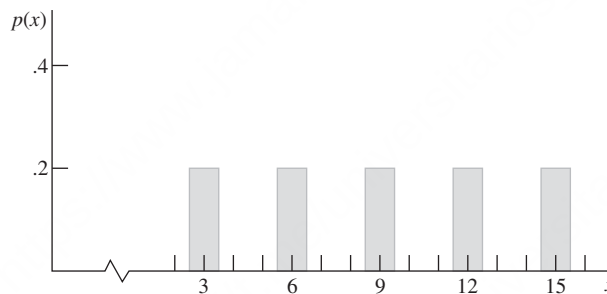
Una población está formada por $N = 5$ números: 3, 6, 9, 12, 15. Si una muestra aleatoria de tamaño $n = 3$ se selecciona sin reemplazo, encuentre las distribuciones muestrales para la media muestral $\bar{x}$ y la mediana m .

Solución Usted está muestreando de la población que aparece en la figura 4.1. Contiene cinco números distintos y cada uno es igualmente probable, con probabilidad $p(x) = 1/5$. Con facilidad puede hallar la media poblacional y mediana como

$$\mu = \frac{3 + 6 + 9 + 12 + 15}{5} = 9 \quad \text{y} \quad M = 9$$

FIGURA 4.1

Histograma de probabilidad para los $N = 5$ valores poblacionales del ejemplo 4.3



TIP MI CONSEJO

Las distribuciones muestrales pueden ser discretas o continuas

Hay 10 posibles muestras aleatorias de tamaño $n = 3$ y cada una es igualmente probable, con probabilidad $1/10$. Estas muestras, junto con los valores calculados de $\bar{x}$ y m para cada una, se muestran en la tabla 4.3. Usted observará que algunos valores de $\bar{x}$ son más probables que otros porque se presentan en más de una muestra. Por ejemplo,

$$P(\bar{x} = 8) = \frac{2}{10} = .2 \quad \text{y} \quad P(m = 6) = \frac{3}{10} = .3$$

Los valores de la tabla 4.3 están tabulados, y las distribuciones muestrales para $\bar{x}$ y m se muestran en la tabla 4.4 y la figura 4.2.

Como la población de $N = 5$ valores es simétrica cerca del valor de $x = 9$, la *media poblacional* y la *mediana* son iguales a 9. Parecería razonable, por tanto, considerar usar ya sea $\bar{x}$ o m como estimador posible de $M = \mu = 9$. ¿Cuál estimador elegiría usted? De la tabla 4.3 se observa que, al usar m como estimador, estaría en un error de $9 - 6 = 3$ con probabilidad .3 o de $9 - 12 = -3$ con probabilidad .3. Esto es, el error en estimación usando m sería 3 con probabilidad .6. Al usar $\bar{x}$, no obstante, un error de 3 ocurriría con probabilidad de sólo .2. Únicamente en estos casos, podría usarse $\bar{x}$ como estimador en preferencia sobre m .

Valores de x y m para muestreo aleatorio simple cuando $n = 3$ y $N = 5$

TABLA 4.3

Muestra	Valores muestrales	$\bar{x}$	m
1	3, 6, 9	6	6
2	3, 6, 12	7	6
3	3, 6, 15	8	6
4	3, 9, 12	8	9
5	3, 9, 15	9	9
6	3, 12, 15	10	12
7	6, 9, 12	9	9
8	6, 9, 15	10	9
9	6, 12, 15	11	12
10	9, 12, 15	12	12

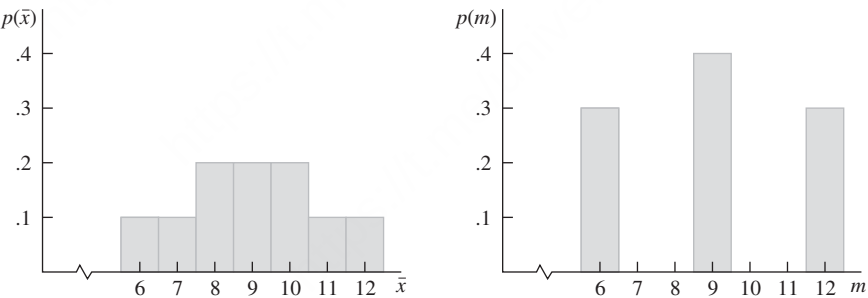
Distribuciones muestrales para a) la media muestral y b) la mediana muestral

TABLA 4.4

(a)	$\bar{x}$	$p(\bar{x})$	(b)	m	$p(m)$
	6	.1		6	.3
	7	.1		9	.4
	8	.2		12	.3
	9	.2			
	10	.2			
	11	.1			
	12	.1			

FIGURA 4.2

Histogramas de probabilidad para las distribuciones muestrales de la media muestral, $\bar{x}$, y la mediana muestral, m , del ejemplo 4.3



TIP

MI CONSEJO

Casi toda estadística tiene una media y una desviación estándar (o *error estándar*) que describen su centro y dispersión

No fue difícil deducir estas distribuciones muestrales en el ejemplo 4.3 porque el número de elementos en la población era muy pequeño. Cuando éste no es el caso, es necesario usar uno de estos métodos:

- Use una simulación para aproximar la distribución muestral de manera empírica.
- Apóyese en teoremas estadísticos y resultados teóricos.

Un teorema estadístico importante, que describe la distribución muestral de estadísticas que son sumas o promedios, se presenta en la siguiente sección.

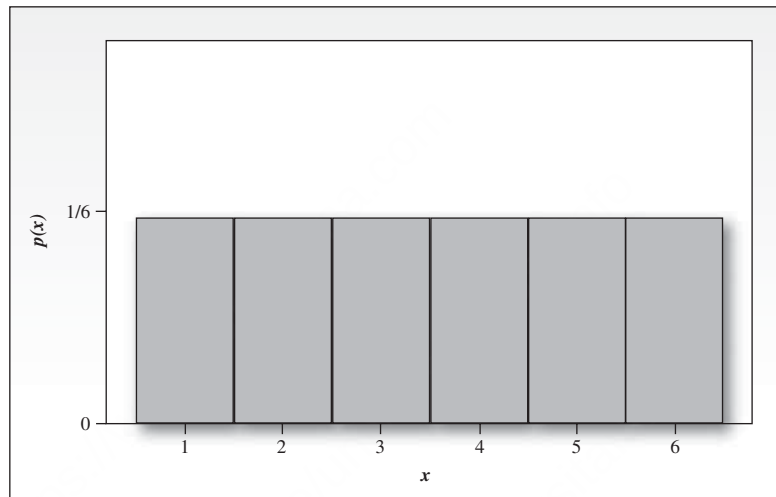
4.4

EL TEOREMA CENTRAL DEL LÍMITE

El **teorema central del límite** establece que, en condiciones más bien generales, las sumas y medias de muestras aleatorias de mediciones tomadas de una población tienden a tener una distribución aproximadamente normal. Suponga que se tira un dado balanceado $n = 1$ vez. La variable aleatoria x es el número observado en la cara superior. Esta variable aleatoria conocida puede tomar seis valores, cada uno con probabilidad $1/6$ y su distribución de probabilidad se muestra en la figura 4.3. La forma de la distribución es *plana* o *uniforme* y simétrica alrededor de la media $\mu = 3.5$, con una desviación estándar $\sigma = 1.71$.

FIGURA 4.3

Distribución de probabilidad para x , el número que aparece en un solo tiro de un dado



Ahora, tome un tamaño muestral de $n = 2$ de esta población; esto es, tire dos veces el dado y registre la suma de los números de las dos caras superiores, $\Sigma x_i = x_1 + x_2$. La tabla 4.5 muestra los 36 posibles resultados, cada uno con probabilidad $1/36$. Las sumas están tabuladas y cada una de las posibles sumas está dividida entre $n = 2$ para obtener un promedio. El resultado es la **distribución muestral** de $\bar{x} = \Sigma x_i/n$, que se ilustra en la figura 4.4. Usted debe observar la considerable diferencia en la forma de la distribución muestral. Ahora tiene casi forma de montículo pero todavía es simétrica alrededor de la media $\mu = 3.5$.

TABLA 4.5

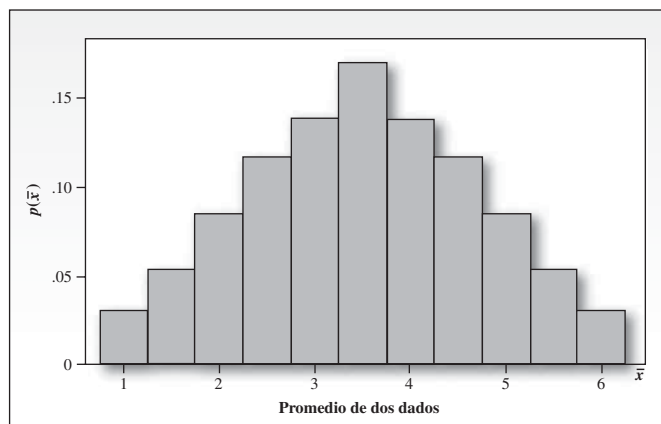
Sumas de las caras superiores de dos dados

Segundo dado	Primer dado					
	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

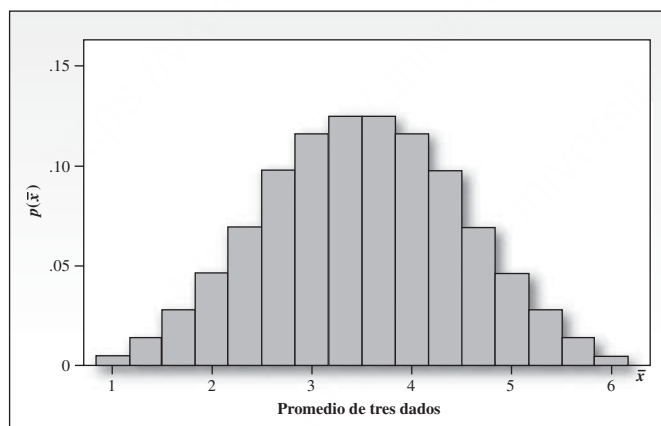
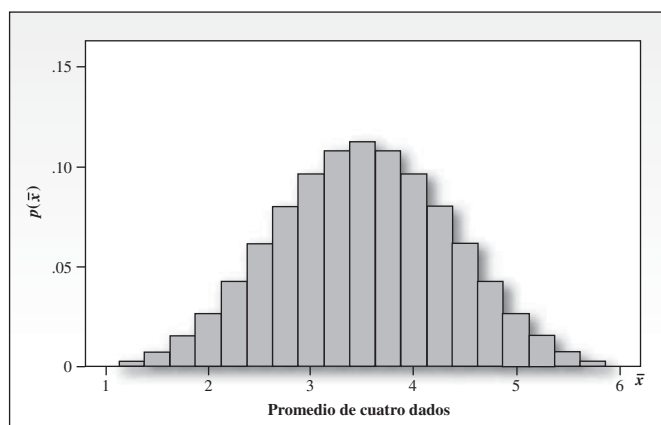
FIGURA 4.4Distribución muestral de $\bar{x}$ para $n = 2$ dados

APPLET EN LÍNEA

Teorema central del límite



Usando un procedimiento similar generamos las distribuciones muestrales de $\bar{x}$ cuando $n = 3$ y $n = 4$. Para $n = 3$, la distribución muestral de la figura 4.5 con toda claridad muestra la forma de montículo de la distribución normal de probabilidad, todavía centrada en $\mu = 3.5$. Observe también que la dispersión de la distribución es lentamente *decreciente* cuando *augmenta* el tamaño muestral n . La figura 4.6 muestra de forma espectacular que la distribución de $\bar{x}$ está distribuida normalmente en forma aproximada con base en una muestra de sólo $n = 4$. Este fenómeno es el resultado de un importante teorema estadístico llamado **teorema central del límite (TCL)**.

FIGURA 4.5Distribución muestral de $\bar{x}$ para $n = 3$ dados**FIGURA 4.6**Distribución muestral de $\bar{x}$ para $n = 4$ dados

Teorema central del límite

Si muestras aleatorias de n observaciones se sacan de una población no normal con media finita μ y desviación estándar σ , entonces, cuando n es grande, la distribución de muestreo de la media muestral $\bar{x}$ está distribuida normalmente en forma aproximada, con media μ y desviación estándar

$$\frac{\sigma}{\sqrt{n}}$$

La aproximación se hace más precisa cuando n se hace grande.

Cualquiera que sea su forma, la distribución muestral de $\bar{x}$ siempre tiene una media idéntica a la media de la población muestreada y una desviación estándar igual a la desviación poblacional estándar σ dividida entre $\sqrt{n}$. En consecuencia, *la dispersión de la distribución de medias muestrales es considerablemente menor que la dispersión de la población muestreada*.

El teorema central del límite se puede expresar de otro modo para aplicar a la **suma de las mediciones muestrales** $\sum x_i$, que, cuando n se hace grande, también tiene una distribución aproximadamente normal con media $n\mu$ y desviación estándar $\sigma\sqrt{n}$.

La aportación importante del teorema central del límite está en la inferencia estadística. Muchos estimadores que se usan para hacer inferencias acerca de parámetros poblacionales son sumas o promedios de las mediciones muestrales. Cuando el tamaño muestral es lo suficientemente grande, se puede esperar que estos estimadores tengan distribuciones muestrales que sean aproximadamente normales. Entonces se puede usar la distribución normal para describir el comportamiento de estos estimadores en muestreo repetido y evaluar la probabilidad de observar ciertos resultados muestrales. Estas probabilidades se calculan usando la variable aleatoria normal estándar

$$z = \frac{\text{Estimador} - \text{Media}}{\text{Desviación estándar}}$$

Cuando usted vuelva a leer el teorema central del límite podrá observar que la aproximación es válida mientras el tamaño muestral n sea “grande”, pero ¿qué tan grande es “grande”? Desafortunadamente, no hay una respuesta clara para esta pregunta. El valor apropiado de n depende de la forma de la población de la cual se muestrea, y cómo se desea usar la aproximación. No obstante, ayudan estas guías:

TIP MI CONSEJO

La distribución muestral de $\bar{x}$ siempre tiene una media μ y desviación estándar $\sigma/\sqrt{n}$. El teorema central del límite ayuda a describir su forma

**NECESITO SABER...**

Cuándo el tamaño de la muestra es lo bastante grande para usar el teorema central del límite

- Si la población muestreada es **normal**, entonces la distribución muestral de $\bar{x}$ también será normal, sin importar cuál sea el tamaño de la muestra que se elija. Este resultado se puede demostrar en forma teórica, pero no debe ser demasiado difícil aceptarla sin prueba.
- Cuando la población muestreada es aproximadamente **simétrica**, la distribución muestral de $\bar{x}$ se hace también aproximadamente normal para valores relativamente pequeños de n . Recuerde la rapidez con la que la distribución “plana” del ejemplo de los dados tomó la forma de montículo ($n = 3$).
- Cuando la población muestreada está **sesgada**, el tamaño muestral n debe ser más grande, con n al menos 30 antes que la distribución muestral de $\bar{x}$ se haga aproximadamente normal.

Estas guías sugieren que, para muchas poblaciones, la distribución muestral de $\bar{x}$ será aproximadamente normal para tamaños muestrales moderados: una excepción a esta regla se presenta al muestrear una población binomial cuando p o $q = (1 - p)$ es muy pequeña. Cuando aparezcan aplicaciones específicas del teorema central del límite, daremos el tamaño muestral apropiado n .

4.5

LA DISTRIBUCIÓN MUESTRAL DE LA MEDIA MUESTRAL

Si la media poblacional μ es desconocida se pueden seleccionar varias *estadísticas* como estimador; la media muestral $\bar{x}$ y la mediana muestral m son dos que con facilidad llegan a nuestra mente. ¿Cuál debe usarse? Considere estos criterios al seleccionar el estimador para μ :

- ¿Es fácil o difícil de calcular?
- ¿Produce estimaciones que sean demasiado altas o demasiado bajas de manera consistente?
- ¿Es más o menos variable que otros estimadores posibles?

Las distribuciones muestrales para $\bar{x}$ y m con $n = 3$ para la pequeña población del ejemplo 4.3 mostraron que, en términos de estos criterios, la media muestral funcionó mejor que la mediana muestral como estimador de μ . En muchas situaciones, la media muestral $\bar{x}$ tiene propiedades deseables como un estimador que no son compartidas por otros estimadores competidores; por tanto, se emplea en forma más amplia.

LA DISTRIBUCIÓN MUESTRAL DE LA MEDIA MUESTRAL, $\bar{x}$

- Si una muestra aleatoria de n mediciones se selecciona de una población con media μ y desviación estándar μ , la distribución muestral de la media muestral $\bar{x}$ tendrá media μ y desviación estándar*

$$\frac{\sigma}{\sqrt{n}}$$

- Si la población tiene una distribución normal, la distribución muestral de $\bar{x}$ estará exactamente *distribuida* en forma normal, *cualquiera que sea el tamaño muestral n* .
- Si la distribución poblacional es *no normal*, la distribución muestral de $\bar{x}$ estará distribuida normalmente en *forma aproximada* para muestras grandes (por el teorema central del límite). De manera conservadora, se requiere $n \geq 30$.

* Cuando muestras repetidas de tamaño n se seleccionan al azar de una población *finita* con N elementos cuya media sea μ y cuya varianza sea σ^2 , la desviación estándar de $\bar{x}$ es

$$\frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

donde σ^2 es la varianza poblacional. Cuando N es grande con respecto al tamaño muestral n , $\sqrt{(N-n)/(N-1)}$ es aproximadamente igual a 1, y la desviación estándar de $\bar{x}$ es

$$\frac{\sigma}{\sqrt{n}}$$

Error estándar

Definición La desviación estándar de una estadística empleada como estimador de un parámetro poblacional también se denomina **error estándar del estimador** (abreviado **SE**) porque se refiere a la precisión del estimador. Por tanto, la desviación estándar de $\bar{x}$, dada por $\sigma/\sqrt{n}$, se conoce como **error estándar de la media** (abreviada $SE(\bar{x})$ o sólo SE).



NECESITO SABER...

Cómo calcular probabilidades para la media muestral $\bar{x}$

Si se sabe que la distribución muestral de $\bar{x}$ es *normal o normalmente aproximada*, se puede describir el comportamiento de la media muestral $\bar{x}$ al calcular la probabilidad de observar ciertos valores de $\bar{x}$ en muestreo repetido.

1. Encuentre μ y calcule $SE(\bar{x}) = \sigma/\sqrt{n}$.
2. Escriba el evento de interés en términos de $\bar{x}$ y localice el área apropiada en la curva normal.
3. Convierta los valores necesarios de $\bar{x}$ en valores z usando

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

4. Use la tabla 3 del Anexo para calcular la probabilidad.

EJEMPLO

4.4

La duración de la enfermedad de Alzheimer desde el principio de los síntomas hasta el fallecimiento varía de 3 a 20 años; el promedio es 8 años con una desviación estándar de 4 años. El administrador de un gran centro médico selecciona al azar los registros médicos de 30 pacientes de Alzheimer ya fallecidos, de la base de datos del centro médico y anota la duración promedio. Encuentre las probabilidades aproximadas para estos eventos:

1. La duración promedio es menor a 7 años.
2. La duración promedio excede de 7 años.
3. La duración promedio está a no más de 1 año de la media poblacional $\mu = 8$.

Solución Como el administrador ha seleccionado una muestra aleatoria de la base de datos en este centro médico, puede sacar conclusiones acerca de pacientes pasados, presentes o futuros con enfermedad de Alzheimer en este centro médico. Si, por el contrario, este centro médico puede ser considerado como representativo de otros centros médicos del país, es posible sacar conclusiones de más largo alcance.

¿Qué se puede decir acerca de la forma de la población muestreada? No es simétrica, porque la media $\mu = 8$ no está a la mitad entre los valores máximo y mínimo. Como la media está más cerca al valor mínimo, la distribución está sesgada a la derecha, con unos pocos pacientes viviendo largo tiempo después que inicia la enfermedad. Cualquiera que sea la forma de la distribución poblacional, no obstante, la distribución muestral de $\bar{x}$ tiene una media de $\mu = 8$ y desviación estándar $\sigma/\sqrt{n} = 4/\sqrt{30} = .73$. Además, como el tamaño

TIP MI CONSEJO

Si x es normal, $\bar{x}$ es normal para cualquier n
Si x no es normal, $\bar{x}$ es aproximadamente normal para n grande



APLET EN LÍNEA

Probabilidades normales
para medias

muestral es $n = 30$, el teorema central del límite asegura la *normalidad aproximada* de la distribución muestral de $\bar{x}$.

1. La probabilidad de que $\bar{x}$ sea menor que 7 está dada por el área sombreada de la figura 4.7. Para hallar esta área, es necesario calcular el valor de z correspondiente a $\bar{x} = 7$:

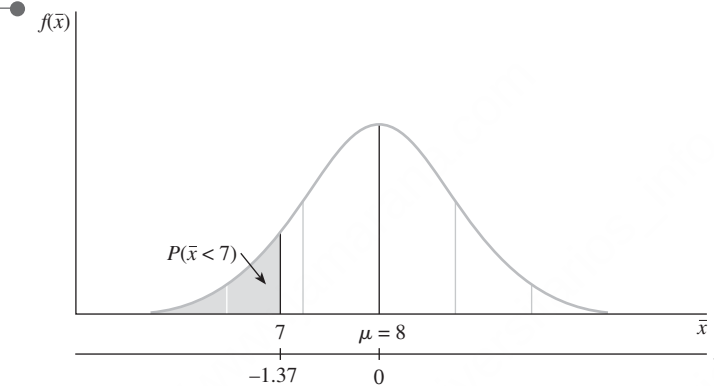
$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{7 - 8}{.73} = -1.37$$

De la tabla 3 del Anexo se puede hallar el área acumulativa correspondiente a $z = -1.37$ y

$$P(\bar{x} < 7) = P(z < -1.37) = .0853$$

FIGURA 4.7

Probabilidad de que $\bar{x}$ sea menor que 7 para el ejemplo 4.4



(NOTA: Se debe usar $\sigma/\sqrt{n}$ (no σ) en la fórmula para z porque estamos buscando un área bajo la distribución muestral para $\bar{x}$, no bajo la distribución de probabilidad para x .)



MI CONSEJO

Recuerde que para variables aleatorias continuas, no hay probabilidad asignada a un punto solo. Por tanto, $P(\bar{x} \leq 7) = P(\bar{x} < 7)$

2. El evento de que $\bar{x}$ exceda de 7 es el complemento del evento que $\bar{x}$ sea menor que 7. Entonces, la probabilidad de que $\bar{x}$ exceda de 7 es

$$\begin{aligned} P(\bar{x} > 7) &= 1 - P(\bar{x} \leq 7) \\ &= 1 - .0853 = .9147 \end{aligned}$$

3. La probabilidad de que $\bar{x}$ se encuentre a no más de 1 año de $\mu = 8$ es el área sombreada en la figura 4.8. El valor z correspondiente a $\bar{x} = 7$ es $z = -1.37$, de la parte 1 y el valor z para $\bar{x} = 9$ es

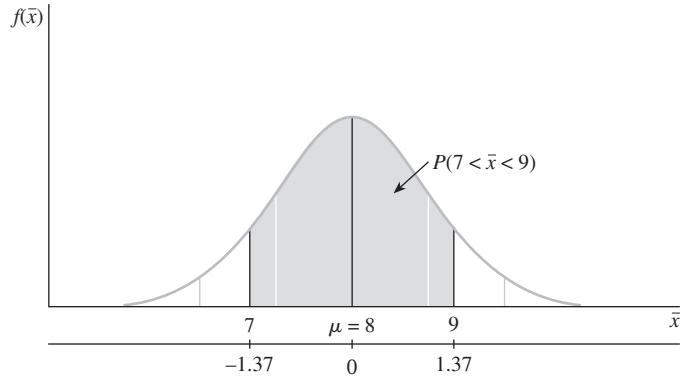
$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{9 - 8}{.73} = 1.37$$

La probabilidad de interés es

$$\begin{aligned} P(7 < \bar{x} < 9) &= P(-1.37 < z < 1.37) \\ &= .9147 - .0853 = .8294 \end{aligned}$$

FIGURA 4.8

La probabilidad de que $\bar{x}$ se encuentre a no más de 1 año de $\mu = 8$ para el ejemplo 4.4

**EJEMPLO****4.5**

Para evitar dificultades con oficinas locales y estatales de protección al consumidor, un embotellador de refrescos debe estar razonablemente seguro de que sus botellas de 12 onzas en realidad contengan 12 onzas de líquido. Para determinar si una máquina embotelladora está funcionando de manera satisfactoria, un embotellador muestrea al azar 10 botellas por hora y mide la cantidad de líquido de cada botella. La media $\bar{x}$ de las 10 medidas de llenado se usa para determinar si se reajusta la cantidad de líquido introducido en la botella por la máquina llenadora.

Si los registros muestran que la cantidad de líquido por botella está normalmente distribuida, con una desviación estándar de .2 onzas y si la máquina embotelladora está ajustada para producir un llenado medio por botella de 12.1 onzas, ¿cuál es la probabilidad aproximada de que la media muestral $\bar{x}$ de las 10 botellas sea menor que 12 onzas?

Solución La media de la distribución muestral de la media muestral $\bar{x}$ es idéntica a la media de la población de llenados de botella, es decir, $\mu = 12.1$ onzas y el error estándar (SE) de $\bar{x}$ es

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{.2}{\sqrt{10}} = .063$$

(NOTA: σ es la desviación estándar de la población de llenados de botella y n es el número de botellas de la muestra.) Como la cantidad de llenado está distribuida normalmente, $\bar{x}$ también está distribuida normalmente, como se muestra en la figura 4.9.

Para hallar la probabilidad de que $\bar{x}$ sea menor que 12 onzas, exprese el valor $\bar{x} = 12$ en unidades de desviación estándar:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{12 - 12.1}{.063} = -1.59$$

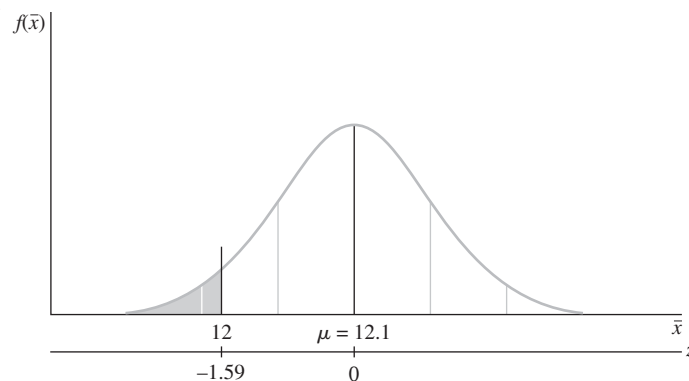
Entonces

$$P(\bar{x} < 12) = P(z < -1.59) = .0559 \approx .056$$

Por tanto, si la máquina está ajustada para producir un llenado promedio de 12.1 onzas, el llenado medio $\bar{x}$ de una muestra de 10 botellas será menor a 12 onzas con una probabilidad igual

FIGURA 4.9

Distribución de probabilidad de $\bar{x}$, la media de los $n = 10$ llenados de botellas, para el ejemplo 4.5



a .056. Cuando se presenta esta señal de riesgo ($\bar{x}$ es menor que 12), el embotellador toma una muestra más grande para volver a verificar el ajuste de la máquina llenadora.

4.5 EJERCICIOS

TÉCNICAS BÁSICAS

4.8 Se seleccionaron muestras aleatorias de tamaño n de poblaciones con las medias y varianzas dadas aquí. Encuentre la media y desviación estándar de la distribución muestral de la media muestral en cada caso:

- a. $n = 36$, $\mu = 10$, $\sigma^2 = 9$
- b. $n = 100$, $\mu = 5$, $\sigma^2 = 4$
- c. $n = 8$, $\mu = 120$, $\sigma^2 = 1$

4.9 Consulte el ejercicio 4.8.

- a. Si las poblaciones muestreadas son normales, ¿cuál es la distribución muestral de $\bar{x}$ para las partes **a**, **b** y **c**?
- b. De acuerdo con el teorema central del límite, si las poblaciones muestreadas *no* son normales, ¿qué se puede decir acerca de la distribución muestral de $\bar{x}$ para las partes **a**, **b** y **c**?

DATOS
EX0410

4.10 Una población está formada por $N = 5$ números: 1, 3, 5, 6 y 7. Se puede demostrar que la media y desviación estándar para esta población son $\mu = 4.4$ y $\sigma = 2.15$, respectivamente.

- a. Construya un histograma de probabilidad para esta población.
- b. Use la tabla de números aleatorios, tabla 10 del Anexo, para seleccionar una muestra aleatoria de tamaño $n = 10$ con reemplazo de la población. Calcule la media muestral, $\bar{x}$. Repita este procedimiento, calculando la media muestral $\bar{x}$ para su segunda muestra. (SUGERENCIA: Asigne los dígitos aleatorios 0 y 1 a la medición $x = 1$;

asigne dígitos 2 y 3 a la medición $x = 3$ y así sucesivamente.)

- c. Para simular la distribución muestral de $\bar{x}$, hemos seleccionado 50 muestras más de tamaño $n = 10$ con restitución y hemos calculado las correspondientes medias muestrales. Construya un histograma de frecuencia relativa para estos 50 valores de $\bar{x}$. ¿Cuál es la forma de esta distribución?

4.8	4.2	4.2	4.5	4.3	4.3	5.0	4.0	3.3	4.7
3.0	5.9	5.7	4.2	4.4	4.8	5.0	5.1	4.8	4.2
4.6	4.1	3.4	4.9	4.1	4.0	3.7	4.3	4.3	4.5
5.0	4.6	4.1	5.1	3.4	5.9	5.0	4.3	4.5	3.9
4.4	4.2	4.2	5.2	5.4	4.8	3.6	5.0	4.5	4.9

4.11 Consulte el ejercicio 4.10.

- a. Use el método de entrada de datos de su calculadora para hallar la media y la desviación estándar de los 50 valores de $\bar{x}$ dados en el ejercicio 4.10, parte **c**.
- b. Compare los valores calculados en la parte **a** contra la media teórica μ y la desviación estándar teórica $\sigma/\sqrt{n}$ para la distribución muestral de $\bar{x}$. ¿Qué tanto se acercan los valores calculados de las 50 mediciones a los valores teóricos?

4.12 Una muestra aleatoria de n observaciones se selecciona de una población con desviación estándar $\sigma = 1$. Calcule el error estándar de la media (SE) para estos valores de n :

- a. $n = 1$
- b. $n = 2$
- c. $n = 4$

- d. $n = 9$ e. $n = 16$ f. $n = 25$
 g. $n = 100$

4.13 Consulte el ejercicio 4.12. Grafique el error estándar de la media (SE) contra el tamaño muestral n y enlace los puntos con una curva suave. ¿Cuál es el efecto de aumentar el tamaño muestral sobre el error estándar?

4.14 Se selecciona una muestra aleatoria de tamaño $n = 49$ de una población con media $\mu = 53$ y desviación estándar $\sigma = 21$.

- a. ¿Cuál será la forma aproximada de la distribución muestral de $\bar{x}$?
 b. ¿Cuáles serán la media y la desviación estándar de la distribución muestral de $\bar{x}$?

4.15 Consulte el ejercicio 4.14. Encuentre la probabilidad de que la media muestral sea mayor que 55.

4.16 Se selecciona una muestra aleatoria de tamaño $n = 40$ de una población con media $\mu = 100$ y desviación estándar $\sigma = 20$.

- a. ¿Cuál será la forma aproximada de la distribución muestral de $\bar{x}$?
 b. ¿Cuáles serán la media y la desviación estándar de la distribución muestral de $\bar{x}$?

4.17 Consulte el ejercicio 4.16. Encuentre la probabilidad de que la media muestral esté entre 105 y 110.

4.18 Suponga que se selecciona una muestra aleatoria de $n = 25$ observaciones de una población que está distribuida normalmente, con media igual a 106 y desviación estándar igual a 12.

- a. Dé la media y desviación estándar de la distribución muestral de la media muestral $\bar{x}$.
 b. Encuentre la probabilidad de que $\bar{x}$ exceda de 110.
 c. Encuentre la probabilidad de que la media muestral se desvíe de la media poblacional $\mu = 106$ en no más de 4.

APLICACIONES

4.19 Error de medición Cuando químicos investigadores realizan experimentos pueden obtener resultados ligeramente diferentes en repeticiones distintas, aunque el experimento se realice de manera idéntica cada vez. Estas diferencias se deben a un fenómeno llamado “error de medición”.

- a. Haga una lista de algunas variables en un experimento de química que podrían causar algunos pequeños cambios en la medición de la respuesta final.
 b. Si se desea verificar que el error de medición es pequeño se puede repetir el experimento y tomar el promedio muestral de todas las mediciones. Para reducir la cantidad de variabilidad en el promedio de

mediciones, ¿debe usarse un número de repeticiones grande o pequeño? Explique.

4.20 Tomates Explique por qué el peso de un empaque de una docena de tomates debe estar distribuido normalmente en forma aproximada si la docena de tomates representa una muestra aleatoria.

4.21 Bacterias en el agua Use el teorema central del límite para explicar por qué una variable aleatoria de Poisson, por ejemplo, el número de un tipo particular de bacterias en un pie cúbico de agua, tiene una distribución que puede ser aproximada por una distribución normal cuando la media μ es grande. (SUGERENCIA: Un pie cúbico de agua contiene 1728 pulgadas cúbicas de líquido.)

4.22 Resistencia del papel Un fabricante de papel exige una resistencia mínima de 20 libras por pulgada cuadrada. Para verificar la calidad del papel, cada hora se selecciona una muestra aleatoria de 10 piezas de papel de la producción de la hora previa, registrándose la medición de su resistencia para cada una. Suponga que las mediciones de resistencia están distribuidas en forma normal con una desviación estándar $\sigma = 2$ libras por pulgada cuadrada.

- a. ¿Cuál es la distribución muestral aproximada de la media muestral de $n = 10$ piezas de papel de prueba?
 b. Si la media de la población de mediciones de resistencia es 21 libras por pulgada cuadrada, ¿cuál es la probabilidad aproximada de que, para una muestra aleatoria de $n = 10$ piezas de papel, $\bar{x} < 20$?
 c. ¿Qué valor se seleccionaría para la resistencia media del papel, μ , para que $P(\bar{x} < 20)$ sea igual a .001?

4.23 Niveles de potasio El requerimiento normal diario de potasio en seres humanos está en el intervalo de 2000 a 6000 miligramos (mg), con cantidades más grandes necesarias durante los meses calurosos de verano. La cantidad de potasio en alimentos varía, pero con frecuencia se asocia a los plátanos con un nivel alto de potasio, con aproximadamente 422 mg en un plátano de tamaño mediano.³ Suponga que la distribución de potasio en un plátano está distribuida normalmente, con media igual a 422 mg y desviación estándar de 13 mg por plátano. Usted come $n = 3$ plátanos al día y T es el número total de miligramos de potasio que recibe de ellos.

- a. Encuentre la media y la desviación estándar de T .
 b. Encuentre la probabilidad de que su ingesta diaria de potasio de los tres plátanos exceda de 1300 mg. (SUGERENCIA: Observe que T es la suma de tres variables aleatorias, x_1 , x_2 y x_3 , donde x_1 es la cantidad de potasio en el plátano 1, etcétera.)

4.24 Temperaturas normales Allen Shoemaker dedujo una distribución de temperaturas del cuerpo humano con una forma distintiva de montículo.⁴ Suponga que

consideramos que las temperaturas de personas sanas es aproximadamente normal, con una media de 98.6° y desviación estándar de 0.8° .

- a. Si se seleccionan al azar 130 personas sanas, ¿cuál es la probabilidad de que la temperatura promedio para ellas sea de 98.25° o menor?
- b. ¿Consideraría usted que una temperatura promedio de 98.25° es un suceso poco común, dado que la verdadera temperatura promedio de personas sanas es de 98.6° ? Explique.

4.25 Deportes y lesiones del tendón de

Aquiles Algunos deportes en los que se practica una cantidad considerable de carreras, saltos de altura o de longitud, ponen a los participantes en riesgo de tendinopatía de Aquiles (AT), que es una inflamación y un engrosamiento del tendón de Aquiles. Especialistas en Medicina del Deporte observaron el diámetro (en mm) de tendones afectados y no afectados de deportistas que

participaron en estos tipos de actividades deportivas. Suponga que los diámetros del tendón de Aquiles en la población general tienen una media de 5.97 milímetros (mm) con una desviación estándar de 1.95 mm.

- a. ¿Cuál es la probabilidad de que una muestra seleccionada al azar de 31 deportistas produzca un diámetro promedio de 6.5 mm o menos para el tendón no afectado?
- b. Cuando los diámetros del tendón afectado se midieron para una muestra de 31 deportistas, el diámetro promedio fue 9.80. Si el promedio del diámetro del tendón en la población de deportistas con AT no es diferente del promedio del diámetro de tendones no afectados (5.97 mm), ¿cuál es la probabilidad de observar un promedio de diámetro de 9.80 o mayor?
- c. ¿Qué conclusiones podrían sacarse de los resultados de la parte b?

LA DISTRIBUCIÓN MUESTRAL DE LA PROPORCIÓN MUESTRAL

4.6

Hay numerosos ejemplos prácticos de la variable aleatoria x binomial. Una aplicación común se relaciona con la preferencia del consumidor o encuestas de opinión, donde usamos una muestra aleatoria de n personas, para estimar la proporción p de personas en la población que tienen una característica especificada. Si x de las personas muestreadas tienen esta característica, entonces la proporción muestral

$$\hat{p} = \frac{x}{n}$$

TIP MI CONSEJO

P: ¿Cómo saber si es o no binomial?

R: Observe si la medición tomada en una sola unidad experimental de la muestra es del tipo “éxito/fracaso”. Si es así, probablemente es binomial

se puede usar para estimar la proporción poblacional p .[†]

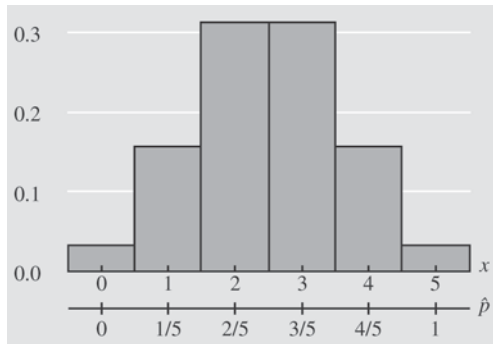
La variable aleatoria x binomial tiene una distribución de probabilidad $p(x)$, descrita en el capítulo 3, con media np y desviación estándar $\sqrt{npq}$. Como $\hat{p}$ es simplemente el valor de

x , expresado como proporción ($\hat{p} = \frac{x}{n}$), la distribución muestral de $\hat{p}$ es idéntica a la distribución de probabilidad de x , excepto que tiene una nueva escala a lo largo del eje horizontal (figura 4.10).

[†] Un “sombbrero” puesto sobre el símbolo de un parámetro poblacional denota una estadística utilizada para estimar el parámetro poblacional. Por ejemplo, el símbolo $\hat{p}$ denota la proporción muestral.

FIGURA 4.10

Distribución muestral de la variable aleatoria x binomial y la proporción muestral $\hat{p}$



Debido a este cambio de escala, la media y la desviación estándar de $\hat{p}$ también tienen cambio de escala, de modo que la media de la distribución muestral de $\hat{p}$ es p , y su error estándar es

$$SE(\hat{p}) = \sqrt{\frac{pq}{n}} \quad \text{donde } q = 1 - p$$

Por último, así como podemos aproximar la distribución de probabilidad de x con una distribución normal cuando el tamaño muestral n es grande, podemos hacer lo mismo con la distribución muestral de $\hat{p}$.

PROPIEDADES DE LA DISTRIBUCIÓN MUESTRAL DE LA PROPORCIÓN MUESTRAL, $\hat{p}$

- Si una muestra aleatoria de n observaciones se selecciona de una población binomial con parámetro p , entonces la distribución muestral de la proporción muestral

$$\hat{p} = \frac{x}{n}$$

tendrá una media

$$p$$

y una desviación estándar

$$SE(\hat{p}) = \sqrt{\frac{pq}{n}} \quad \text{donde } q = 1 - p$$

- Cuando el tamaño muestral n es grande, la distribución muestral de $\hat{p}$ puede ser aproximada por una distribución normal. La aproximación será adecuada si $np > 5$ y $nq > 5$.

EJEMPLO

4.6

En una encuesta se preguntó a 500 padres de familia sobre la importancia del deporte para sus hijos. De los padres entrevistados 60% estuvo de acuerdo en que los géneros son iguales y deben tener iguales oportunidades de participar en deportes. Describa la distribución muestral de la proporción muestral $\hat{p}$ de padres que están de acuerdo en que los géneros son iguales y deben tener iguales oportunidades.

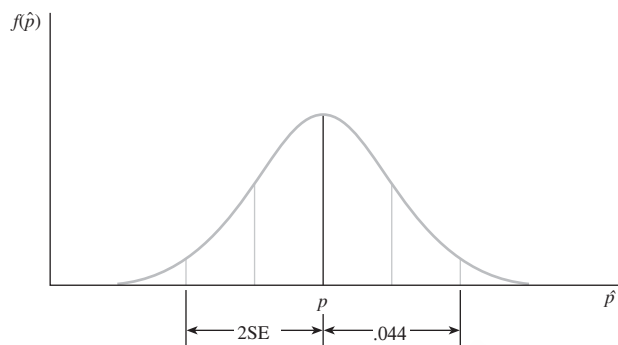
Solución Se puede suponer que los 500 padres representan una muestra aleatoria de los padres de todos los hijos e hijas en un estado y que la verdadera proporción de la población es igual a algún valor desconocido que se puede llamar p . La distribución muestral de $\hat{p}$ puede

ser aproximada por una distribución muestral,^{††} con media igual a p (véase la figura 4.11) y error estándar

$$SE(\hat{p}) = \sqrt{\frac{pq}{n}}$$

FIGURA 4.11

Distribución muestral para $\hat{p}$ basada en una muestra de $n = 500$ padres para el ejemplo 4.6



Se observa de la figura 4.11 que la distribución muestral de $\hat{p}$ está centrada sobre su media p . Aun cuando no se sabe el valor exacto de p (la proporción muestral $\hat{p} = .60$ puede ser mayor o menor que p), puede hallarse un valor aproximado para la desviación estándar de la distribución muestral usando la proporción muestral $\hat{p} = .60$ para aproximar el valor desconocido de p . Entonces,

$$\begin{aligned} SE &= \sqrt{\frac{pq}{n}} \approx \sqrt{\frac{\hat{p}\hat{q}}{n}} \\ &= \sqrt{\frac{(.60)(.40)}{500}} = .022 \end{aligned}$$

Por tanto, aproximadamente 95% de las veces, $\hat{p}$ caerá en no más de $2SE \approx .044$ del valor (desconocido) de p .



NECESITO SABER...

Cómo calcular probabilidades para la proporción muestral $\hat{p}$

1. Encuentre los valores necesarios de n y p .
2. Verifique si la aproximación normal a la distribución binomial es apropiada ($np > 5$ y $nq > 5$).
3. Escriba el evento de interés en términos de $\hat{p}$ y localice el área apropiada en la curva normal.
4. Convierta los valores necesarios de $\hat{p}$ en valores z usando

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

5. Use la tabla 3 del Anexo para calcular la probabilidad.

^{††} Si se verifican las condiciones que permiten la aproximación normal a la distribución de $\hat{p}$, se puede ver que $n = 500$ es adecuado para valores de p cercanos a .60 porque tanto $n\hat{p} = 300$ como $n\hat{q} = 200$ son mayores que 5.

EJEMPLO

4.7

Consulte el ejemplo 4.6. Suponga que la proporción p de padres en la población es en realidad igual a .55. ¿Cuál es la probabilidad de observar una proporción muestral igual de grande o mayor que el valor observado $\hat{p} = .60$?

Solución La figura 4.12 muestra la distribución muestral de $\hat{p}$ cuando $p = .55$, con el valor observado $\hat{p} = .60$ ubicado sobre el eje horizontal. La probabilidad de observar una proporción muestral $\hat{p}$ igual a o mayor que .60 es aproximada por el área sombreada en la cola superior de esta distribución normal con

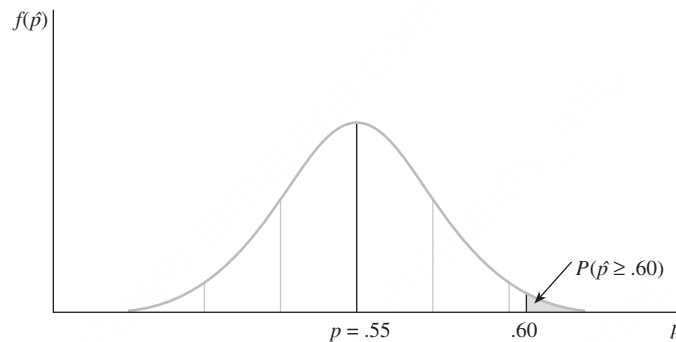
$$p = .55$$

y

$$SE = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(.55)(.45)}{500}} = .0222$$

FIGURA 4.12

Distribución muestral de $\hat{p}$ para $n = 500$ y $p = .55$ para el ejemplo 4.7



Para hallar el área sombreada, primero calcule el valor z correspondiente a $\hat{p} = .60$:

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{.60 - .55}{.0222} = 2.25$$

Usando la tabla 3 del Anexo, encontramos

$$P(\hat{p} > .60) \approx P(z > 2.25) = 1 - .9878 = .0122$$

Esto es, si seleccionáramos una muestra aleatoria de $n = 500$ observaciones de una población con proporción p igual a .55, la probabilidad de que la proporción muestral $\hat{p}$ sea tan grande o mayor que .60 es de sólo .0122.

Cuando la distribución normal se utilizó en el capítulo 3 para aproximar las probabilidades binomiales asociadas con x , una corrección de $\pm .5$ se aplicó para mejorar la aproximación. La corrección equivalente aquí es $\pm (.5/n)$. Por ejemplo, para $\hat{p} = .60$ el valor de z con la corrección es

$$z_1 = \frac{(.60 - .001) - .55}{\sqrt{\frac{(.55)(.45)}{500}}} = 2.20$$

con $P(\hat{p} > .60) \approx .0139$. Con una precisión de dos lugares decimales, este valor concuerda con el resultado anterior. Cuando n es grande, el efecto de usar la corrección es por lo general insignificante. Usted debe resolver problemas de este capítulo y los restantes *sin* el factor de corrección a menos que específicamente se le solicite.

4.6 EJERCICIOS

TÉCNICAS BÁSICAS

4.26 Se seleccionaron muestras aleatorias de tamaño n de poblaciones binomiales con parámetros poblacionales p dados aquí. Encuentre la media y la desviación estándar de la distribución muestral de la proporción muestral $\hat{p}$ en cada caso:

- a. $n = 100, p = .3$
- b. $n = 400, p = .1$
- c. $n = 250, p = .6$

4.27 ¿Es apropiado usar la distribución normal para aproximar la distribución muestral de $\hat{p}$ en las siguientes circunstancias?

- a. $n = 50, p = .05$
- b. $n = 75, p = .1$
- c. $n = 250, p = .99$

4.28 Se seleccionaron muestras aleatorias de tamaño $n = 75$ de una población binomial con $p = .4$. Use la distribución normal para aproximar las siguientes probabilidades:

- a. $P(\hat{p} \leq .43)$
- b. $P(.35 \leq \hat{p} \leq .43)$

4.29 Se seleccionaron muestras aleatorias de tamaño $n = 500$ de una población binomial con $p = .1$.

- a. ¿Es apropiado usar la distribución normal para aproximar la distribución muestral de $\hat{p}$? Verifique para asegurarse que estén satisfechas las condiciones necesarias.

Usando los resultados de la parte a, encuentre estas probabilidades:

- b. $\hat{p} > .12$
- c. $\hat{p} < .10$
- d. $\hat{p}$ está a no más de .02 de p

4.30 Calcule $SE(\hat{p})$ para $n = 100$ y estos valores de p :

- a. $p = .01$ b. $p = .10$ c. $p = .30$
- d. $p = .50$ e. $p = .70$ f. $p = .90$
- g. $p = .99$

- h. Grafique $SE(\hat{p})$ contra p en papel milimétrico y trace una curva suave que pase por los puntos. ¿Para qué valor de p es máxima la desviación estándar de la distribución muestral de $\hat{p}$? ¿Qué ocurre al error estándar cuando p es cercano a 0 o a 1.0?

4.31 Se selecciona una muestra aleatoria de tamaño $n = 50$ de una distribución binomial con una proporción de población $p = .7$.

- a. ¿Cuál será la forma aproximada de la distribución muestral de $\hat{p}$?
- b. ¿Cuáles serán la media y la desviación estándar (o error estándar) de la distribución muestral de $\hat{p}$?
- c. Encuentre la probabilidad de que la proporción muestral $\hat{p}$ sea menor que .8.

4.32 Se selecciona una muestra aleatoria de tamaño $n = 80$ de una distribución binomial con proporción poblacional $p = .25$.

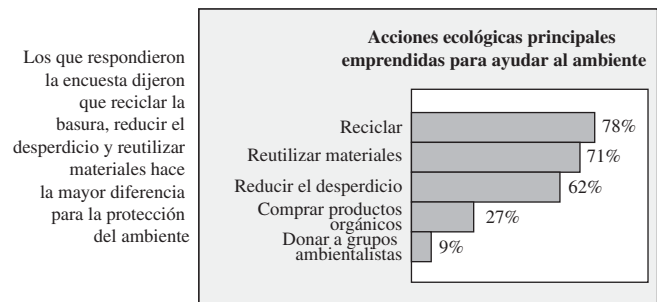
- a. ¿Cuál será la forma aproximada de la distribución muestral de $\hat{p}$?
- b. ¿Cuáles serán la media y la desviación estándar (o error estándar) de la distribución muestral de $\hat{p}$?
- c. Encuentre la probabilidad de que la proporción muestral $\hat{p}$ esté entre .18 y .44.

4.33 a. ¿Es apropiada la aproximación normal a la distribución muestral de $\hat{p}$ cuando $n = 400$ y $p = .8$?

- b. Use los resultados de la parte a para hallar la probabilidad de que $\hat{p}$ es mayor que .83.
- c. Use los resultados de la parte a para hallar la probabilidad de que $\hat{p}$ está entre .76 y .84.

APLICACIONES

4.34 Amigable con la ecología Reciclar basura, reducir los desechos y reutilizar materiales son acciones ecológicas que ayudarán al ambiente. Según una encuesta (ejercicio 3.54), 78% de los que respondieron mencionan el reciclaje como la forma principal de ayudar a nuestro ambiente. Suponga que se selecciona una muestra aleatoria de $n = 100$ adultos y que la cifra de 78% es correcta.



- a. ¿La distribución de $\hat{p}$, la proporción muestral de adultos que listan el reciclaje como la forma principal de ayudar al ambiente, tiene una distribución aproximadamente normal? De ser así, ¿cuál es su media y su desviación estándar?

- b. ¿Cuál es la probabilidad de que la proporción muestral $\hat{p}$ sea menor que 75%?
- c. ¿Cuál es la probabilidad de que $\hat{p}$ caiga en el intervalo de .7 a .75?
- d. ¿Qué podría concluir sobre p si la proporción muestral fuera menor que .65?

4.35 ¡Ah, las nueces! ¿Es usted un “purista” del chocolate o le gusta que tenga otros ingredientes? *American Demographics* informa que casi 75% de los consumidores gustan de ingredientes tradicionales como nueces o caramelos en su chocolate. Son menos entusiastas

hacia el gusto de la menta o el café, que dan sabores más distintivos.⁵ Se selecciona una muestra aleatoria de 200 consumidores y se registra el número de quienes gustan de las nueces o caramelo en su chocolate.

- a. ¿Cuál es la distribución muestral aproximada para la proporción muestral $\hat{p}$? ¿Cuáles son la media y la desviación estándar para esta distribución?
- b. ¿Cuál es la probabilidad de que el porcentaje muestral sea mayor que 80%?
- c. ¿Dentro de qué límites se esperaría que la proporción muestral se encuentre alrededor de 95% de las veces?

UNA APLICACIÓN MUESTRAL: CONTROL ESTADÍSTICO DE PROCESOS (OPCIONAL)

4.7

La metodología del control estadístico de procesos (CEP) se inventó para observar, controlar y mejorar productos y servicios. Por ejemplo, los cojinetes de acero deben apegarse a especificaciones de dimensiones y dureza, los productos químicos industriales tienen un bajo nivel especificado de impurezas y las empresas de contaduría deben reducir al mínimo y finalmente eliminar asientos incorrectos en contabilidades. Con frecuencia se dice que el control estadístico de procesos está formado por 10% de estadística y 90% de ingeniería y sentido común. Desde el punto de vista estadístico podemos observar la media de un proceso y decir cuándo es que la media cae fuera de límites asignados previamente, pero no podemos decir *por qué* está fuera de control. Contestar esta última pregunta requiere conocimiento del proceso y capacidad de resolver problemas, es decir, ¡el otro 90%!

La calidad de un producto suele ser supervisada con el uso de gráficas estadísticas de control. Las mediciones en la variable de un proceso a ser observado cambian con el tiempo. Se dice que la causa de un cambio en la variable es *asignable* si puede ser localizada y corregida. Otra variación, los pequeños cambios fortuitos debidos a la alteración del ambiente de producción, que no sea controlable se considera como *variación aleatoria*. Si la variación en la variable de un proceso es sólo aleatoria, se dice que el proceso está *en control*.

El primer objetivo en el control estadístico de procesos es eliminar causas asignables de variación en la variable del proceso y luego poner el proceso bajo control. El siguiente paso es reducir la variación y poner las mediciones en la variable del proceso dentro de *límites de especificación*, es decir, límites dentro de los cuales deben caer las mediciones sobre objetos o servicios utilizables.

Una vez que un proceso esté en control y produciendo un producto satisfactorio, sus variables son supervisadas con **gráficas de control**. Se toman muestras de n artículos del proceso a intervalos especificados y se calcula una estadística muestral. Estas estadísticas se grafican en la tabla de control, de modo que el proceso se puede verificar respecto a cambios en la variable del proceso que podrían indicar problemas de control.

Una gráfica de control para la media del proceso: la gráfica $\bar{x}$

Suponga que n artículos se seleccionan al azar del proceso de producción a intervalos iguales y que se registran mediciones en la variable del proceso. Si el proceso está en control, las medias muestrales deben variar alrededor de la media poblacional μ en forma aleatoria.

Además, de acuerdo con el teorema central del límite, la distribución muestral de $\bar{x}$ debe ser aproximadamente normal, de modo que casi todos los valores de $\bar{x}$ caen en el intervalo $(\mu \pm 3 \text{ SE}) = \mu \pm 3(\sigma/\sqrt{n})$. Aun cuando los valores exactos de μ y σ sean desconocidos, se pueden obtener estimaciones precisas con el uso de mediciones muestrales.

Toda gráfica de control tiene una *línea de centro* y *límites de control*. La línea del centro para la **gráfica $\bar{x}$** es la estimación de μ , el gran promedio de todas las estadísticas muestrales calculadas de las mediciones en la variable del proceso. Los *límites de control superior e inferior* están colocados a tres desviaciones estándar arriba y debajo de la línea de centro. Si se observa la media del proceso con base en k muestras de tamaño n tomadas a intervalos regulares, la línea de centro es $\bar{\bar{x}}$, el promedio de las medias muestrales y los límites de control están en $\bar{\bar{x}} \pm 3(\sigma/\sqrt{n})$, con σ estimada por s , la desviación estándar de las mediciones nk .

EJEMPLO

4.8

Un sistema estadístico de observación del control del proceso muestrea los diámetros interiores de $n = 4$ cojinetes por hora. En la tabla 4.6 aparecen los datos para $k = 25$ muestras por hora. Construya una gráfica $\bar{x}$ para vigilar la media del proceso.

Solución La media muestral se calculó para cada una de las $k = 25$ muestras. Por ejemplo, la media para la muestra 1 es

$$\bar{x} = \frac{.992 + 1.007 + 1.016 + .991}{4} = 1.0015$$

Las medias muestrales se ven en la última columna de la tabla 4.6. La línea del centro está situada en el promedio de las medias muestrales, o sea

$$\bar{\bar{x}} = \frac{24.9675}{25} = .9987$$

**25 muestras por hora de diámetros de cojinetes,
 $n = 4$ cojinetes por muestra**

TABLA 4.6

Muestra	Mediciones de muestra				Media muestral, $\bar{x}$
1	.992	1.007	1.016	.991	1.00150
2	1.015	.984	.976	1.000	.99375
3	.988	.993	1.011	.981	.99325
4	.996	1.020	1.004	.999	1.00475
5	1.015	1.006	1.002	1.001	1.00600
6	1.000	.982	1.005	.989	.99400
7	.989	1.009	1.019	.994	1.00275
8	.994	1.010	1.009	.990	1.00075
9	1.018	1.016	.990	1.011	1.00875
10	.997	1.005	.989	1.001	.99800
11	1.020	.986	1.002	.989	.99925
12	1.007	.986	.981	.995	.99225
13	1.016	1.002	1.010	.999	1.00675
14	.982	.995	1.011	.987	.99375
15	1.001	1.000	.983	1.002	.99650
16	.992	1.008	1.001	.996	.99925
17	1.020	.988	1.015	.986	1.00225
18	.993	.987	1.006	1.001	.99675
19	.978	1.006	1.002	.982	.99200
20	.984	1.009	.983	.986	.99050
21	.990	1.012	1.010	1.007	1.00475
22	1.015	.983	1.003	.989	.99750
23	.983	.990	.997	1.002	.99300
24	1.011	1.012	.991	1.008	1.00550
25	.987	.987	1.007	.995	.99400

El valor calculado de s , la desviación muestral estándar de todas las $nk = 4(25) = 100$ observaciones, es $s = .011458$ y el error estándar estimado de la media de $n = 4$ observaciones es

$$\frac{s}{\sqrt{n}} = \frac{.011458}{\sqrt{4}} = .005729$$

Los límites de control superior e inferior se encuentran como

$$UCL = \bar{\bar{x}} + 3\frac{s}{\sqrt{n}} = .9987 + 3(.005729) = 1.015887$$

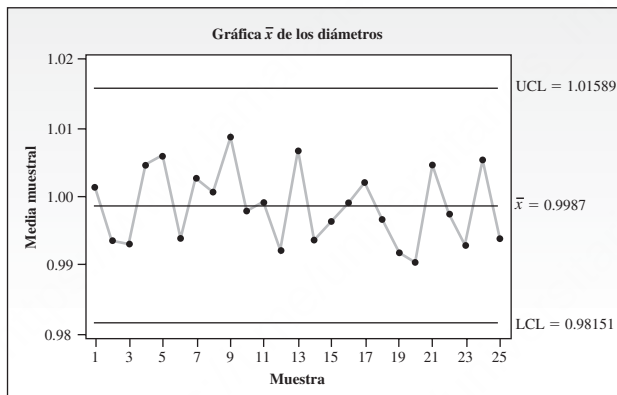
y

$$LCL = \bar{\bar{x}} - 3\frac{s}{\sqrt{n}} = .9987 - 3(.005729) = .981513$$

La figura 4.13 muestra la gráfica $\bar{x}$ construida a partir de los datos. Si se supone que las muestras empleadas para construir la gráfica $\bar{x}$ se recolectaron cuando el proceso estaba en control,

FIGURA 4.13

Gráfica $\bar{x}$ para el ejemplo 4.8



la gráfica se puede usar ahora para detectar cambios en la media del proceso. Las medias muestrales se grafican periódicamente y, si una media muestral cae fuera de los límites de control, el proceso debe verificarse para localizar la causa de la media anormalmente grande o pequeña.

Una gráfica de control para la proporción de piezas defectuosas: la gráfica p

En ocasiones la observación hecha en un artículo es simplemente para saber si satisface o no las especificaciones; entonces, se juzga como defectuosa o no defectuosa. Si la fracción de piezas defectuosas producidas por el proceso es p , entonces x , el número de defectuosas en una muestra de n artículos, tiene una distribución binomial.

Para supervisar un proceso para ver si hay artículos defectuosos, se seleccionan muestras de tamaño n a intervalos periódicos y se calcula la proporción muestral $\hat{p}$. Cuando el proceso está en control, $\hat{p}$ debe caer en el intervalo $p \pm 3SE$, donde p es la proporción de defectuosas en la población (o la fracción defectuosa del proceso) con error estándar

$$SE = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}}$$

La fracción defectuosa del proceso es desconocida pero puede estimarse por el promedio de las proporciones muestrales k :

$$\bar{p} = \frac{\sum \hat{p}_i}{k}$$

y el error estándar es estimado por

$$SE = \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

La línea de centro para la **gráfica p** está ubicada en $\bar{p}$, y los límites de control superior e inferior son

$$UCL = \bar{p} + 3\sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

y

$$LCL = \bar{p} - 3\sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

EJEMPLO

4.9

Un fabricante de bolígrafos muestrea al azar 400 bolígrafos por día y prueba cada uno para ver si el flujo de tinta es aceptable. Las proporciones de bolígrafos que se clasifican como defectuosos por día, en un periodo de 40 días, se presentan en la tabla 4.7. Construya una gráfica de control para la proporción $\hat{p}$ de muestras defectuosas en $n = 400$ bolígrafos seleccionados del proceso.

Solución La estimación de la proporción de piezas defectuosas del proceso es el promedio de las $k = 40$ proporciones muestrales de la tabla 4.7. Por tanto, la línea del centro de la gráfica de control está ubicada en

$$\bar{p} = \frac{\sum \hat{p}_i}{k} = \frac{.0200 + .0125 + \dots + .0225}{40} = \frac{.7600}{40} = .019$$

TABLA 4.7

Proporciones de piezas defectuosas en muestras de $n = 400$ bolígrafos

Día	Proporción	Día	Proporción	Día	Proporción	Día	Proporción
1	.0200	11	.0100	21	.0300	31	.0225
2	.0125	12	.0175	22	.0200	32	.0175
3	.0225	13	.0250	23	.0125	33	.0225
4	.0100	14	.0175	24	.0175	34	.0100
5	.0150	15	.0275	25	.0225	35	.0125
6	.0200	16	.0200	26	.0150	36	.0300
7	.0275	17	.0225	27	.0200	37	.0200
8	.0175	18	.0100	28	.0250	38	.0150
9	.0200	19	.0175	29	.0150	39	.0150
10	.0250	20	.0200	30	.0175	40	.0225

Una estimación del SE, es decir, el error estándar de las proporciones muestrales, es

$$\sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} = \sqrt{\frac{(.019)(.981)}{400}} = .00683$$

y $3SE = (3)(.00683) = .0205$. Por tanto, los límites de control superior e inferior para la gráfica p están ubicados en

$$UCL = \bar{p} + 3SE = .0190 + .0205 = .0395$$

y

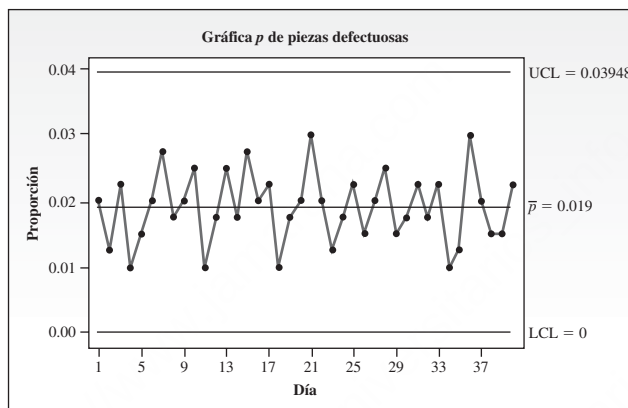
$$LCL = \bar{p} - 3SE = .0190 - .0205 = -.0015$$

O bien, como p no puede ser negativa, el límite inferior de control o $LCL = 0$.

La gráfica de control p se muestra en la figura 4.14. Observe que las 40 proporciones muestrales caen dentro de los límites de control. Si una proporción muestral recolectada en algún tiempo futuro cae fuera de los límites de control, el fabricante debe preocuparse por un aumento en el porcentaje de piezas defectuosas, además de tomar las medidas necesarias para buscar las posibles causas de este aumento.

FIGURA 4.14

Gráfica p para el ejemplo 4.9



Otras gráficas de control que por lo general se utilizan son la *gráfica R*, que se usa para vigilar la variación de la variable del proceso por medio del intervalo muestral y la *gráfica c*, que se usa para vigilar el número de defectos por pieza.

4.7

EJERCICIOS

TÉCNICAS BÁSICAS

4.36 Las medias muestrales se calcularon para 30 muestras de tamaño $n = 10$ para un proceso que se juzgó en control. Las medias de los 30 valores $\bar{x}$ y la desviación estándar de las 300 mediciones combinadas fueron $\bar{\bar{x}} = 20.74$ y $s = .87$, respectivamente.

- Use los datos para determinar los límites superior e inferior de control para una gráfica $\bar{x}$.
- ¿Cuál es el propósito de una gráfica $\bar{x}$?
- Construya una gráfica $\bar{x}$ para el proceso y explique cómo se puede usar.

4.37 Se calcularon las medias muestrales para 40 muestras de tamaño $n = 5$ para un proceso que se juzgó en control. Las medias de los 40 valores y la desviación estándar de las 200 mediciones combinadas fueron $\bar{\bar{x}} = 155.9$ y $s = 4.3$, respectivamente.

- Use los datos para determinar los límites superior e inferior de control para una gráfica $\bar{x}$.
- Construya una gráfica $\bar{x}$ para el proceso y explique cómo se puede usar.

4.38 Explique la diferencia entre una gráfica $\bar{x}$ y una gráfica p .

4.39 Se seleccionaron muestras de $n = 100$ artículos cada hora durante un periodo de 100 horas y la proporción

muestral de defectuosas se calculó cada hora. La media de las 100 proporciones muestrales fue de .035.

- Use los datos para determinar los límites superior e inferior de control para una gráfica p .
- Construya una gráfica p para el proceso y explique cómo se puede usar.

4.40 Se seleccionaron muestras de $n = 200$ artículos cada hora durante un periodo de 100 horas y la proporción muestral de defectuosas se calculó cada hora. La media de las 100 proporciones muestrales fue de .041.

- Use los datos para determinar los límites superior e inferior de control para una gráfica p .
- Construya una gráfica p para el proceso y explique cómo se puede usar.

APLICACIONES

4.41 Veintiuno Un casino de juegos de azar registra y grafica la ganancia o pérdida diaria media, de cinco mesas de veintiuno, en una gráfica $\bar{x}$. La media general de las medias muestrales y la desviación estándar de los datos combinados de 40 semanas fueron $\bar{\bar{x}} = \$10\,752$ y $s = \$1\,605$, respectivamente.

- Construya una gráfica $\bar{x}$ para la ganancia diaria media por mesa de veintiuno.
- ¿Cómo puede serle de valor esta gráfica $\bar{x}$ al gerente del casino?

4.42 Remaches de latón Un fabricante de remaches de latón muestrea al azar 400 remaches cada hora y calcula la proporción de los defectuosos de la muestra. La proporción media muestral calculada de 200 muestras era igual a .021. Construya una gráfica de control para la proporción de defectuosos en muestras de 400 remaches. Explique la forma en que la gráfica de control puede ser de valor para un gerente.

4.43 Especificaciones en madera El gerente de una compañía de materiales de construcción muestrea al azar la madera que recibe para ver si cumple con especificaciones de calidad. De cada embarque, 100 piezas de madera de 2×4 son inspeccionadas y juzgadas de acuerdo con si son de primera (aceptable) o de segunda (defectuosa) clase. Las proporciones de piezas de madera de 2×4 de segunda registradas para 30 embarques fueron como sigue:

.14	.21	.19	.18	.23	.20	.25	.19	.22	.17
.21	.15	.23	.12	.19	.22	.15	.26	.22	.21
.14	.20	.18	.22	.21	.13	.20	.23	.19	.26

Construya una gráfica de control para la proporción de piezas de madera de 2×4 de segunda en muestras de 100 piezas de madera. Explique la forma en que la

gráfica de control puede ser empleada por el gerente de la compañía de materiales de construcción.

4.44 Planta generadora de electricidad a base de carbón Una planta generadora de electricidad a base de carbón prueba y mide tres tipos de carbón al día, para vigilar su porcentaje de ceniza. La media general de 30 medias muestrales diarias y la desviación estándar combinada de todos los datos fueron $\bar{\bar{x}} = 7.24$ y $s = .07$, respectivamente. Construya una gráfica $\bar{x}$ para el proceso y explique la forma en que puede ser de valor para el gerente de la planta generadora de electricidad.

DATOS
EX0445

4.45 Planta nuclear de energía eléctrica Los datos de la tabla son medidas de la radiación en partículas de aire en una planta nuclear de energía eléctrica. Se registraron cuatro mediciones a intervalos semanales en un periodo de 26 semanas. Use los datos para construir una gráfica $\bar{x}$ y grafique los 26 valores de $\bar{x}$. Explique la forma en que se puede usar la gráfica.

Semana	Radiación			
1	.031	.032	.030	.031
2	.025	.026	.025	.025
3	.029	.029	.031	.030
4	.035	.037	.034	.035
5	.022	.024	.022	.023
6	.030	.029	.030	.030
7	.019	.019	.018	.019
8	.027	.028	.028	.028
9	.034	.032	.033	.033
10	.017	.016	.018	.018
11	.022	.020	.020	.021
12	.016	.018	.017	.017
13	.015	.017	.018	.017
14	.029	.028	.029	.029
15	.031	.029	.030	.031
16	.014	.016	.016	.017
17	.019	.019	.021	.020
18	.024	.024	.024	.025
19	.029	.027	.028	.028
20	.032	.030	.031	.030
21	.041	.042	.038	.039
22	.034	.036	.036	.035
23	.021	.022	.024	.022
24	.029	.029	.030	.029
25	.016	.017	.017	.016
26	.020	.021	.020	.022

4.46 Bates de béisbol Una planta fabricante de maderas duras tiene una línea de producción diseñada para elaborar bates que pesan 32 onzas. Durante un tiempo, cuando estaba en control estadístico, se encontró que el peso promedio de un bate era de 31.7 onzas. Los datos observados se reunieron de 50 muestras, cada una de ellas formada por 5 mediciones. Se encontró que la desviación estándar de todas las muestras era $s = .2064$ onzas. Construya una gráfica $\bar{x}$ para vigilar el proceso de producción de bates de 32 onzas.

4.47 Más bates de béisbol Consulte el ejercicio 4.46 y suponga que, durante un día cuando el estado del proceso de producción de bates de 32 onzas era desconocido, se obtuvieron las siguientes mediciones a intervalos de una hora.

Hora	$\bar{x}$	Hora	$\bar{x}$
1	31.6	4	33.1
2	32.5	5	31.6
3	33.4	6	31.8

Cada medición representa una estadística calculada de una muestra de cinco pesos de bates seleccionados del proceso de producción durante cierta hora. Use la gráfica de control construida en el ejercicio 4.46 para supervisar el proceso.

REPASO DEL CAPÍTULO

Conceptos y fórmulas clave

I. Planes muestrales y diseños experimentales

- Muestreo aleatorio simple
 - Cada posible muestra de tamaño n es igualmente probable de ocurrir
 - Use una computadora o tabla de números aleatorios
 - Los problemas son: sin respuesta, baja cobertura y sesgo verbal
- Otros planes muestrales con aleatorización
 - Muestreo aleatorio estratificado
 - Muestreo de conglomerado
 - Muestreo sistemático de 1 en k
- Muestreo no aleatorio
 - Muestreo de conveniencia
 - Muestreo de juicio
 - Muestreo de cuota

II. Estadísticas y distribuciones muestrales

- Las distribuciones muestrales describen los posibles valores de una estadística y con qué frecuencia se presentan en muestreo repetido
- Las distribuciones muestrales se pueden derivar matemáticamente, aproximarse en forma empírica o hallarse usando teoremas estadísticos
- El teorema central del límite dice que las sumas y promedios de mediciones de una población no normal, con media μ finita y desviación estándar σ , tienen distribuciones aproximadamente normales para muestras grandes de tamaño n

III. Distribución muestral de la media muestral

- Cuando muestras de tamaño n se sacan al azar de una población normal con media μ y varianza σ^2 , la media muestral $\bar{x}$ tiene una distribución normal con media μ y desviación estándar $\sigma/\sqrt{n}$
- Cuando muestras de tamaño n se sacan al azar de una población no normal con media μ y varianza σ^2 , el teorema central del límite asegura que la media muestral $\bar{x}$ tendrá una distribución aproximadamente normal con media μ y desviación estándar $\sigma/\sqrt{n}$ cuando n es grande ($n \geq 30$).
- Las probabilidades que contengan la media muestral pueden calcularse al estandarizar el valor de $\bar{x}$ usando z :

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

IV. Distribución muestral de la proporción muestral

- Cuando muestras de tamaño n se toman de una población binomial con parámetro p , la proporción muestral $\hat{p}$ tendrá una distribución aproximadamente normal con media p y desviación estándar $\sqrt{pq/n}$ mientras $np > 5$ y $nq > 5$.
- Las probabilidades que comprendan la proporción muestral se pueden calcular al estandarizar el valor $\hat{p}$ usando z :

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

V. Control estadístico de un proceso

1. Para vigilar un proceso cuantitativo, use una gráfica $\bar{x}$. Seleccione k muestras de tamaño n y calcule la media general $\bar{\bar{x}}$ y la desviación estándar s de todas las nk mediciones. Genere límites de control superiores e inferiores como

$$\bar{\bar{x}} \pm 3 \frac{s}{\sqrt{n}}$$

Si una media muestral excede de estos límites, el proceso está fuera de control.

2. Para vigilar un proceso binomial, use una gráfica p . Seleccione k muestras de tamaño n y calcule el promedio de las proporciones muestrales como

$$\bar{p} = \frac{\sum \hat{p}_i}{k}$$

Genere límites de control superiores e inferiores como

$$\bar{p} \pm 3 \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

Si una proporción muestral excede de estos límites, el proceso está fuera de control.



TECNOLOGÍA ACTUAL

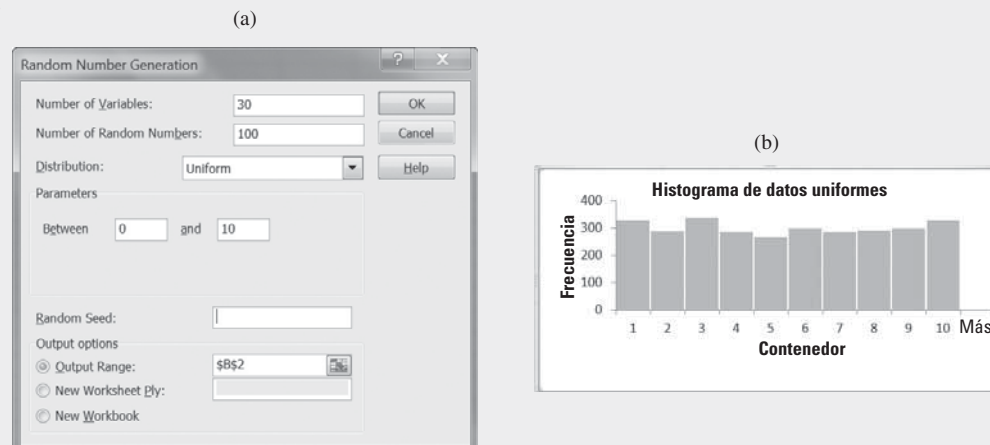
El teorema central del límite en operación: *Microsoft Excel*

Microsoft Excel puede usarse para explorar la forma en que funciona el teorema central del límite en la práctica. Recuerde que, de acuerdo con el teorema central del límite, si se extraen muestras aleatorias de tamaño n de una población no normal con media μ y desviación estándar σ , entonces cuando n es grande, la distribución muestral de la media muestral $\bar{x}$ será aproximadamente normal con la misma media μ y con error estándar $\sigma/\sqrt{n}$. Probemos muestrear de una población no normal usando *Excel*.

En una hoja de cálculo nueva, genere 100 muestras de tamaño $n = 30$ de una distribución uniforme continua del intervalo $(0, 10)$. Etiquete la columna A “Muestra” e introduzca los números 1 a 100 en esa columna. Luego seleccione **Data ► Data Analysis ► Random Number Generation**, para obtener el cuadro de diálogo de la figura 4.15a). Escriba **30** para el número de variables y 100 para el número de números aleatorios. En la lista desplegable “Distribution”, elija “Uniform”, con parámetros entre 0 y 10. Dejaremos vacía la primera fila de la hoja de cálculo, empezando el “Output Range” en la celda B2. Presione **OK** para ver las 100 muestras aleatorias de tamaño $n = 30$. Puede ver la distribución del conjunto entero de datos usando **Data ► Data Analysis ► Histogram**, eligiendo los contenedores 1, 2, ..., 9, 10 y usando los procedimientos descritos en la sección “Tecnología actual” en el capítulo 1. Para nuestros datos, la distribución, mostrada en la figura 4.15b) no tiene forma de montículo, sino que es bastante plana, como se espera para la distribución uniforme.

Para la distribución uniforme que hemos usado, la media y la desviación estándar son $\mu = 5$ y $\sigma = 2.89$, respectivamente. Compruebe las estadísticas descriptivas para las 30×100

FIGURA 4.15



= 3000 mediciones (use las funciones **=AVERAGE(B2:AE101)** e **=STDEV(B2:AE101)**), y encontrará que las 100 observaciones tienen una media muestral y una desviación estándar cercanas pero no exactamente iguales a $\mu = 5$ y $\sigma = 2.89$, respectivamente.

Ahora, genere 100 valores de $\bar{x}$ con base en muestras de tamaño $n = 30$ creando una columna de medias para las 100 filas. Primero, etiquete la columna AF como “x-barra” y coloque el cursor en la celda AF2. Use **Insert Function ► Statistical ► Average** (o escriba **=AVERAGE(B2:AE2)**) para obtener el primer promedio. Luego copie la fórmula en las otras 99 celdas en la columna AF. Ahora puede ver la distribución de estas 100 medias muestrales usando **Data ► Data Analysis ► Histogram** y eligiendo los contenedores 1, 1.5, 2, 2.5, ..., 9, 9.5, 10. La distribución para nuestras 100 medias muestrales se muestra en la figura 4.16.

FIGURA 4.16



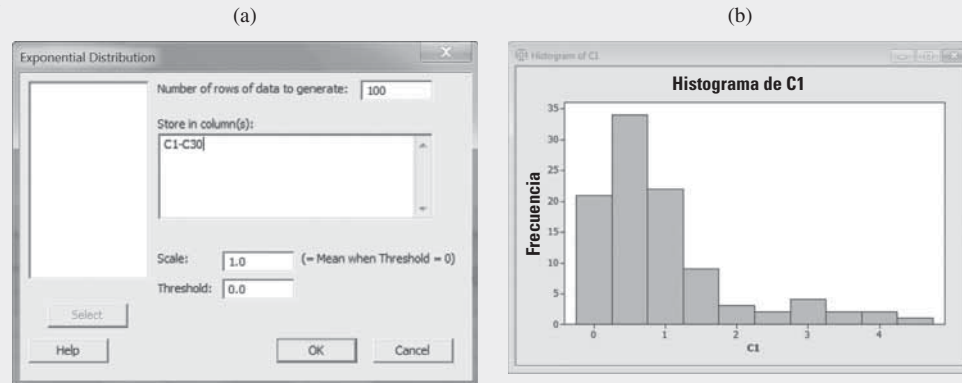
Observe la forma de montículo acentuada de la distribución en la figura 4.16 comparada con la distribución original en la figura 4.15b). Además, si comprueba la media y la desviación estándar para las 100 medias muestrales en la columna AE, encontrará que no son demasiado diferentes de los valores teóricos, $\mu = 5$ y $\sigma/\sqrt{n} = 2.89/\sqrt{30} = .53$. (Para nuestros datos, la media muestral es 4.98 y la desviación estándar es .49.) En vista que sólo tenemos 100 muestras, nuestros resultados no son *exactamente* iguales a los valores teóricos. Si hubiéramos generado un número *infinito* de muestras, habríamos obtenido una correspondencia exacta. ¡Éste es el teorema central del límite en operación!

El teorema central del límite en operación: *MINITAB*

El *MINITAB* es una herramienta perfecta para explorar la forma en que el teorema central del límite funciona en la práctica. Recuerde que, según el teorema central del límite, si muestras aleatorias de tamaño n se sacan de una población no normal con media μ y desviación estándar σ , entonces, cuando n es grande, la distribución muestral de la media muestral $\bar{x}$ será aproximadamente normal con la misma media μ y con error estándar $\sigma/\sqrt{n}$. Intentemos muestrear a partir de una población no normal con la ayuda del *MINITAB*.

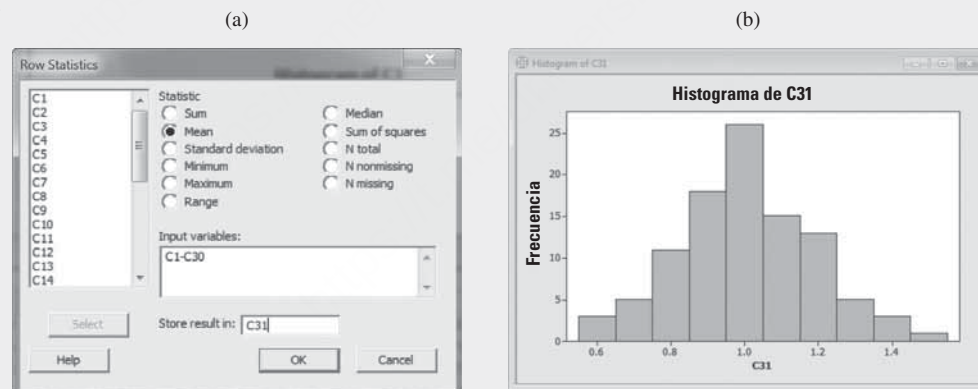
En una hoja de trabajo *MINITAB* nueva, genere 100 muestras de tamaño $n = 30$ de una distribución no normal llamada distribución exponencial. Use **Calc ► Random Data ► Exponential**. Teclee **100** para el número de renglones de datos y guarde los resultados en C1-C30 (véase la figura 4.17a)). Deje la media en el valor predeterminado de 1.0, el umbral en 0.0 y haga clic en **OK**. Los datos se generan y guardan en la hoja de trabajo. Use **Graph ► Histogram ► Simple** para ver la distribución de algunos de los datos, por ejemplo C1 (como en la figura 4.17b)). Observe que la distribución no tiene forma de montículo; está sumamente sesgada a la derecha.

FIGURA 4.17



Para la distribución exponencial que hemos empleado, la media y desviación estándar son $\mu = 1$ y $\sigma = 1$, respectivamente. Verifique las estadísticas descriptivas para una de las columnas (use **Stat ► Basic Statistics ► Display Descriptive Statistics**) y se verá que las 100 observaciones tienen una media muestral y desviación estándar que son *cercanas* pero no exactamente iguales a 1. A continuación, genere 100 valores de $\bar{x}$ con base en muestras de tamaño $n = 30$ al crear una columna de medias para los 100 renglones. Use **Calc ► Row Statistics** y seleccione **Mean**. Para promediar las entradas en todas las 30 columnas, seleccione o teclee **C1-C30** en la caja de variables de Entrada y guarde los resultados en **C31** (véase la figura 4.18a)). Ahora podrá ver la distribución de las medias muestrales usando **Graph ► Histogram ► Simple**, seleccionando **C31** y haciendo clic en **OK**. La distribución de las 100 medias muestrales generadas para nuestro ejemplo se ilustran en la figura 4.18b).

FIGURA 4.18



Observe la forma distintiva de montículo de la distribución de la figura 4.18b) en comparación con la distribución original de la figura 4.17b). También, si verifica las estadísticas descriptivas para C31, encontrará que la media y la desviación estándar de nuestras 100 medias muestrales no son demasiado diferentes de los valores teóricos, $\mu = 1$ y $\sigma/\sqrt{n} = 1/\sqrt{30} = .18$. (Para nuestros datos, la media muestral es 1.0024 y la desviación estándar es .1813.) Como sólo teníamos 100 muestras, nuestros resultados no son *exactamente* iguales a los valores teóricos. Si hubiéramos generado un número *infinito* de muestras, hubiéramos obtenido una coincidencia exacta. ¡Éste es el teorema central del límite en operación!

Ejercicios suplementarios

4.48 Una población finita consta de cuatro elementos:

6, 1, 3, 2.

- ¿Cuántas muestras diferentes de tamaño $n = 2$ se pueden seleccionar de esta población si se muestrea sin reemplazo? (El muestreo se dice que es *sin reemplazo* si un elemento no puede ser seleccionado dos veces para la misma muestra.)
- Haga una lista de posibles muestras de tamaño $n = 2$.
- Calcule la media muestral para cada una de las muestras dadas en la parte **b**.
- Encuentre la distribución muestral de $\bar{x}$. Use un histograma de probabilidad para graficar la distribución muestral de $\bar{x}$.
- Si los cuatro valores poblacionales son igualmente probables, calcule el valor de la media poblacional μ . ¿Alguna de las muestras de la lista de la parte **b** produce un valor de $\bar{x}$ exactamente igual a μ ?

4.49 Consulte el ejercicio 4.48. Encuentre la distribución muestral para $\bar{x}$ si muestras aleatorias de tamaño $n = 3$ se seleccionan *sin reemplazo*. Grafique la distribución muestral de $\bar{x}$.

4.50 Suponga que se selecciona una muestra aleatoria de $n = 5$ observaciones de una población que está distribuida normalmente, con media igual a 1 y desviación estándar igual a .36.

- Dé la media y la desviación estándar de la distribución muestral de $\bar{x}$.
- Encuentre la probabilidad de que $\bar{x}$ exceda 1.3.
- Encuentre la probabilidad de que la media muestral $\bar{x}$ sea menor que .5.
- Encuentre la probabilidad de que la media muestral se desvíe de la media poblacional $\mu = 1$ por más de .4.

4.51 Baterías Se sabe que cierto tipo de batería para automóviles dura un promedio de 1 110 días con una desviación estándar de 80 días. Si se seleccionan 400 de estas baterías, encuentre las siguientes probabilidades para la duración de vida promedio de las baterías seleccionadas:

- El promedio está entre 1 100 y 1 110.
- El promedio es mayor que 1 120.
- El promedio es menor que 900.

4.52 Tubos de plomo Estudios realizados indican que el agua potable, suministrada por algunos viejos sistemas de tuberías con forro interior de plomo (cañerías) en las ciudades, puede contener niveles peligrosos de plomo. Un estudio importante del sistema de abastecimiento de agua de Boston mostró que la distribución de lecturas de contenido de plomo para muestras individuales de agua tenía una media

y desviación estándar de aproximadamente .033 miligramos por litro (mg/l) y .10 mg/l, respectivamente.⁶

- Explique por qué piensa usted que esta distribución está o no normalmente distribuida.
- Debido a que los investigadores están preocupados por la forma de la distribución de la parte **a**, calcularon el promedio de los niveles diarios de plomo en 40 lugares en cada uno de los 23 días seleccionados al azar. ¿Qué se puede decir acerca de la forma de la distribución del promedio de niveles diarios de plomo de los que se tomó la muestra de 23 días?
- ¿Cuáles son la media y la desviación estándar de la distribución del promedio de niveles de plomo en la parte **b**?

4.53 Biomasa Estudios realizados⁷ indican que la masa vegetal de la tierra, o biomasa, para bosques tropicales, estimada en unos 35 kilogramos por metro cuadrado (kg/m^2), puede en realidad ser demasiado alta y que los valores de biomasa tropical varían de una región a otra de alrededor de 5 a 55 kg/m^2 . Supongamos que usted mide la biomasa tropical en 400 lugares seleccionados al azar, de un metro cuadrado cada uno.

- Aproxime σ , la desviación estándar de las mediciones de biomasa.
- ¿Cuál es la probabilidad de que el promedio muestral se encuentre dentro de dos unidades del verdadero promedio de biomasa tropical?
- Si el promedio muestral obtenido es $\bar{x} = 31.75$, ¿qué concluiría usted acerca de la estimación excesiva que preocupa a científicos?

4.54 Cascos de protección Los requisitos de seguridad para cascos de protección utilizados por trabajadores de la construcción y otros, establecidos por el American National Standards Institute (ANSI), especifican que cada uno de tres cascos pase la siguiente prueba. Se coloca un casco sobre una forma de cabeza de aluminio. Se deja caer una esfera de acero de 8 libras de peso sobre el casco desde una altura de 5 pies y se mide la fuerza resultante en la parte inferior de la forma de cabeza. La fuerza ejercida en la forma de cabeza por cada uno de los tres cascos debe ser menor que 1000 libras y el promedio de los tres debe ser menor que 850 libras. (Se desconoce la relación entre esta prueba y el daño real a una cabeza humana.) Suponga que la fuerza ejercida está normalmente distribuida y por tanto la media muestral de tres mediciones de la fuerza está normalmente distribuida. Si una muestra aleatoria de tres cascos se selecciona de un envío con una media igual a 900 y $\sigma = 100$, ¿cuál es la probabilidad de que la media muestral satisfaga la norma del ANSI?

4.55 Imágenes y memoria Un psicólogo investigador está planeando un experimento para determinar si el uso de imágenes, es decir, describir una palabra en la mente, afecta

la capacidad de las personas para aprender de memoria algo. El investigador desea usar dos grupos de sujetos: un grupo que memoriza un conjunto de 20 palabras usando la técnica de imágenes y un grupo control que no usa imágenes.

- Use una técnica de aleatorización para dividir un grupo de 20 individuos en dos grupos de igual tamaño.
- ¿Cómo puede el investigador seleccionar al azar el grupo de 20 individuos?
- Suponga que el investigador ofrece pagar \$50 a cada participante del experimento y utiliza a los primeros 20 estudiantes que lo solicitan. ¿Este grupo se comportaría como si fuera una muestra aleatoria simple de tamaño $n = 20$?

4.56 Rábanos que brotan Se diseñó un experimento de biología para determinar si las semillas nacientes de rábanos inhiben la germinación de semillas de lechuga.⁸ Se usaron tres cajas de Petri de 10 centímetros. La primera contenía 26 semillas de lechuga; la segunda, 26 semillas de rábano y, la tercera, 13 semillas de lechuga y 13 de rábano.

- Suponga que el experimentador tenía un paquete de 50 semillas de rábanos y otro de 50 semillas de lechuga. Diseñe un plan para asignar al azar las semillas de rábanos y lechuga a los tres grupos de tratamiento.
- ¿Qué suposiciones debe hacer el experimentador acerca de los paquetes de 50 semillas para asegurar lo aleatorio del experimento?

4.57 Servicio telefónico Suponga que el ejecutivo de una compañía telefónica desea seleccionar una muestra aleatoria de $n = 20$ entre 7 000 clientes para un estudio de las actitudes de clientes respecto al servicio. Si los clientes se numeran con fines de identificación, indique los clientes a quienes incluiría usted en su muestra. Use la tabla de números aleatorios y explique cómo seleccionaría su muestra.

4.58 Rh positivo La proporción de personas con tipo de sangre Rh positivo es 85%. Usted tiene una muestra aleatoria de $n = 500$ personas.

- ¿Cuáles son la media y la desviación estándar de $\hat{p}$, la proporción muestral con tipo de sangre Rh positivo?
- ¿La distribución de $\hat{p}$ es aproximadamente normal? Justifique su respuesta.
- ¿Cuál es la probabilidad de que la proporción muestral $\hat{p}$ exceda de 82%?
- ¿Cuál es la probabilidad de que la proporción muestral se encuentre entre 83% y 88%?
- 99% de las veces, ¿la proporción muestral estaría entre cuáles dos límites?

4.59 ¿Qué diseño de estudio se usa en cada una de estas situaciones?

- Se selecciona una muestra aleatoria de $n = 50$ manzanas de ciudad y se realiza un censo por cada vivienda unifamiliar en cada manzana.
- Una patrulla de caminos detiene a cada décimo vehículo en una calle determinada, entre las 9:00 a.m. y las 3:00 p.m. para efectuar una revisión rutinaria sobre seguridad de tránsito.
- Cien familias en cada una de cuatro delegaciones citadinas son encuestadas respecto a un referendo pendiente de desgravación de impuestos.
- Se inspecciona cada décimo árbol de una plantación de tala de pinos para comprobar si está infestado del gusano barrenador.
- Una muestra aleatoria de $n = 1\,000$ contribuyentes en una ciudad es seleccionada por el Servicio de impuestos interno y se auditan sus declaraciones de impuestos.

4.60 Cargas de elevadores La carga máxima (con un generoso factor de seguridad) para el elevador de un edificio de oficinas es de 2 000 libras. La distribución de frecuencia relativa de los pesos de todos los hombres y mujeres que usan el elevador tiene forma de montículo (ligeramente sesgada a los pesos pesados), con una media μ igual a 150 libras y desviación estándar σ de 35 libras. ¿Cuál es el número máximo de personas que se pueden permitir en el elevador, si se desea que el peso total de ellas exceda del peso máximo con una pequeña probabilidad (por ejemplo, cercano a .01)? (SUGERENCIA: Use la declaración alternativa del teorema central del límite para $\sum x_i$ dada en la sección 4.4.)

4.61 Paquetes de alambrado El número de paquetes de alambrado que pueden ser ensamblados por los empleados de una compañía tiene una distribución normal, con una media igual a 16.4 por hora y una desviación estándar de 1.3 por hora.

- ¿Cuáles son la media y la desviación estándar del número x de paquetes producidos por trabajador en un día de 8 horas?
- ¿Se espera que la distribución de probabilidad de x sea de forma de montículo y aproximadamente normal? Explique.
- ¿Cuál es la probabilidad de que un trabajador produzca al menos 135 paquetes por día de 8 horas?

4.62 Paquetes de alambrado, continúa Consulte el ejercicio 4.61. Suponga que la compañía emplea 10 ensambladores de paquetes de alambrado.

- Encuentre la media y la desviación estándar de la producción diaria de la compañía (día de 8 horas) de paquetes de alambrado.
- ¿Cuál es la probabilidad de que la producción diaria de la compañía sea menor que 1 280 paquetes de alambrado por día?



4.63 Focos defectuosos La tabla siguiente es una lista del número de focos defectuosos de 60 watts, hallados en muestras de 100

focos seleccionados en 25 días del proceso de un fabricante. Suponga que durante estos 25 días el proceso de manufactura no estuvo produciendo una parte excesivamente grande de focos defectuosos.

Día	1	2	3	4	5	6	7	8	9	10
Defectuosos	4	2	5	8	3	4	4	5	6	1

Día	11	12	13	14	15	16	17	18	19	20
Defectuosos	2	4	3	4	0	2	3	1	4	0

Día	21	22	23	24	25
Defectuosos	2	2	3	5	3

- Construya una gráfica p para vigilar el proceso de manufactura y grafique los datos.
- ¿Qué tan grande debe ser la fracción de artículos defectuosos en una muestra seleccionada del proceso de manufactura, antes que se considere que el proceso está fuera de control?
- Durante un día determinado, suponga que se selecciona una muestra de 100 artículos del proceso de manufactura y se encuentran 15 focos defectuosos. Si se toma una decisión de cerrar el proceso de manufactura, en un intento por localizar la fuente de la variación que se sobreentiende que es controlable, explique la forma en que esta decisión puede llevar a conclusiones erróneas.

4.64 Focos, continúa Una cadena de ferreterías compra grandes cantidades de focos, del fabricante descrito en el ejercicio 4.63 y especifica que cada embarque debe contener no más de 4% de defectuosos. Cuando el proceso de manufactura está en control, ¿cuál es la probabilidad de que sean satisfechas las especificaciones de la cadena de ferreterías?

4.65 Focos, otra vez Consulte el ejercicio 4.63. Durante una semana determinada, el número de focos defectuosos que se encontraron en cada una de cinco muestras de 100 fueron 2, 4, 9, 7 y 11. ¿Hay razón para creer que el proceso de producción ha estado produciendo una proporción excesiva de focos defectuosos en cualquier tiempo durante la semana?



4.66 Tomates enlatados Durante largas series de producción de tomates enlatados, los pesos promedio (en onzas) de muestras de cinco latas de tomates de calidad estándar, en forma de puré, se tomaron en 30 puntos de control durante un periodo de 11 días. Estos resultados se muestran en la tabla.⁹ Cuando la máquina está funcionando normalmente, el peso

promedio por lata es de 21 onzas con una desviación estándar de 1.20 onzas.

- Calcule los límites superior e inferior de control y la línea del centro para la gráfica $\bar{x}$.
- Grafique los datos muestrales en la gráfica $\bar{x}$ y determine si la operación de la máquina está en control.

Número de muestra	Peso promedio	Número de muestra	Peso promedio
1	23.1	16	21.4
2	21.3	17	20.4
3	22.0	18	22.8
4	21.4	19	21.1
5	21.8	20	20.7
6	20.6	21	21.6
7	20.1	22	22.4
8	21.4	23	21.3
9	21.5	24	21.1
10	20.2	25	20.1
11	20.3	26	21.2
12	20.1	27	19.9
13	21.7	28	21.1
14	21.0	29	21.6
15	21.6	30	21.3

Fuente: Adaptado de J. Hackl, *Journal of Quality Technology*, abril de 1991. Utilizado con permiso.

4.67 Fresas Un experimentador desea hallar una temperatura apropiada a la cual almacenar fresas frescas, para reducir al mínimo la pérdida de ácido ascórbico. Hay 20 recipientes de almacenamiento, cada uno con temperatura controlable, en los que se pueden guardar fresas. Si se usan dos temperaturas de almacenamiento, ¿cómo podría el experimentador asignar los 20 contenedores a una de las dos temperaturas?

4.68 Llenado de latas de refrescos Un embotellador de bebidas gaseosas empaca latas en paquetes de seis. Suponga que el líquido por lata tiene una distribución normal aproximada con una media de 12 onzas de líquido y una desviación estándar de 0.2 onzas de líquido.

- ¿Cuál es la distribución del total de líquido para una caja de 24 latas?
- ¿Cuál es la probabilidad de que el total de líquido para una caja sea menor que 286 onzas de líquido?
- Si un paquete de seis latas de refresco se puede considerar como muestra aleatoria de tamaño $n = 6$ de la población, ¿cuál es la probabilidad de que el promedio de líquido por lata para un paquete de seis latas de refresco sea menor que 11.8 onzas de líquido?

4.69 Peso total al envasar Paquetes de alimento cuyo peso promedio es 16 onzas, con desviación estándar de 0.6 onzas, se envían en cajas de 24 paquetes. Si los pesos de los paquetes están normalmente distribuidos en forma aproximada, ¿cuál es la probabilidad de que una caja de 24 paquetes pese más de 392 onzas (24.5 libras)?

4.70 Componentes electrónicos Un proceso de manufactura está diseñado para producir un componente electrónico para uso en pequeños televisores portátiles. Todos los componentes son de tamaño estándar y no necesitan apegarse a ninguna característica mensurable, pero a veces son inoperables cuando emergen del proceso de manufactura. Se seleccionaron 15 muestras en momentos en que se sabía que el proceso estaba en control

estadístico. Se observaron quince componentes dentro de cada muestra y se registró el número de componentes inoperables.

6, 7, 3, 5, 6, 8, 4, 5, 7, 3, 1, 6, 5, 4, 5

Construya una gráfica p para vigilar el proceso de manufactura.

CASO PRÁCTICO



Muestreo de la Ruleta de Monte Carlo

La técnica de simular un proceso que contiene elementos aleatorios y repetir el proceso una y otra vez, para ver cómo se comporta, recibe el nombre de **método de Monte Carlo**. Se usa ampliamente en finanzas y otros campos para investigar las propiedades de una operación que está sujeta a efectos aleatorios, tales como el clima, la conducta humana, etcétera. Por ejemplo, se podría modelar el comportamiento del inventario de una compañía manufacturera al crear, en papel, llegadas y salidas diarias de productos manufacturados desde el almacén de la compañía. Cada día, un número aleatorio de artículos producidos por la compañía sería recibido en inventario. Del mismo modo, cada día un número aleatorio de pedidos de diferentes tamaños aleatorios se enviaría. Con base en la entrada y salida de artículos, se podría calcular el inventario; es decir, el número de artículos disponibles al finalizar cada día. Los valores de las variables aleatorias, el número de artículos producidos, el número de pedidos y el número de artículos por pedido necesarios para la simulación de cada día, se obtendría de distribuciones teóricas de observaciones que modelan muy de cerca las correspondientes distribuciones de las variables que se han observado en el tiempo de la operación de manufactura. Al repetir la simulación del suministro, el envío y el cálculo del inventario diario para un gran número de días (un muestreo de lo que podría realmente ocurrir), se puede observar el comportamiento del inventario diario de la planta. El método de Monte Carlo es particularmente valioso porque hace posible que el fabricante vea cómo se comportaría el inventario diario, cuando se hacen ciertos cambios en el patrón de abastecimiento o en algún otro aspecto de la operación que podría ser controlado.

En un artículo titulado “El Camino a Monte Carlo”, Daniel Seligman comenta sobre el método de Monte Carlo, observando que aun cuando la técnica se usa ampliamente en escuelas de finanzas para estudiar el presupuesto de capital, planeación de inventarios y administración de flujos de efectivo, nadie parece haber usado el procedimiento para estudiar lo bien que la haríamos si fuéramos a jugar en Monte Carlo.¹⁰

Para seguir sobre esta idea, Seligman programó su computadora personal para simular el juego de la ruleta. La ruleta es una rueda con su borde dividido en 38 buchacas. Treinta y seis de éstas están numeradas del 1 al 36 y tienen colores alternados de rojo y negro. Las dos buchacas restantes tienen color verde y están marcadas 0 y 00. Para jugar en la ruleta, se apuesta cierta cantidad de dinero a una o más buchacas, a continuación se hace girar la rueda hasta detenerse. Una pequeña esfera cae en una ranura en la rueda para indicar el número ganador. Si usted tiene dinero en ese número, gana una cantidad especificada.

Por ejemplo, si fuera a jugar el número 20, la paga sería 35 a 1. Si la rueda no se detiene en ese número, usted pierde su apuesta. Seligman decidió ver cómo serían sus ganancias (o pérdidas) nocturnas si apostara cinco dólares en cada giro de la rueda y repitiera el proceso 200 veces por noche. Hizo esto 365 veces, con lo cual simulaba los resultados de 365 noches en el casino. Sin ninguna sorpresa, la “ganancia” media por noche de 1 000 dólares para las 365 noches fue una *pérdida* de 55 dólares, el promedio de las ganancias retenidas por la casa.

La sorpresa, de acuerdo con Seligman, fue la extrema variabilidad de las “ganancias” nocturnas. Siete veces, de entre las 365, el jugador ficticio perdió la apuesta de 1 000 dólares y sólo una vez ganó un máximo de 1 160 dólares. En 141 noches, la pérdida fue más de 250 dólares.

1. Para evaluar los resultados del experimento de Monte Carlo de Seligman, primero encuentre la distribución de probabilidad de la ganancia x en una sola apuesta de cinco dólares.
2. Encuentre el valor y la varianza esperados de la ganancia x del punto 1.
3. Encuentre el valor esperado y la varianza para la ganancia de la noche, la suma de las ganancias o pérdidas para las 200 apuestas de \$5 cada una.
4. Use los resultados del punto 2 para evaluar la probabilidad de que 7 de entre 365 noches resulten en una pérdida de la apuesta total de 1 000 dólares.
5. Use los resultados del punto 3 para evaluar la probabilidad de que las ganancias más grandes de la noche fueran de hasta 1 160 dólares.

Estimación de muestras grandes

OBJETIVO GENERAL

En capítulos previos, usted ya se enteró de las distribuciones de probabilidad de variables aleatorias y las distribuciones muestrales de varias estadísticas que, para tamaños muestrales grandes, pueden ser aproximadas por una distribución normal de acuerdo con el teorema central del límite. Este capítulo presenta un método para estimar parámetros poblacionales e ilustra el concepto con ejemplos prácticos. El teorema central del límite y las distribuciones muestrales presentadas en el capítulo 4 desempeñan un papel clave para evaluar la confiabilidad de estimaciones.

ÍNDICE DEL CAPÍTULO

- Dónde hemos estado (5.1)
- A dónde vamos (5.2)
- Tipos de estimadores (5.3)
- Estimación puntual (5.4)
- Estimación de intervalo (5.5)
- Estimación de la diferencia entre dos medias poblacionales (5.6)
- Estimación de la diferencia entre dos proporciones binomiales (5.7)
- Límites de confianza a una cola (5.8)
- Selección del tamaño muestral (5.9)



NECESITO SABER...

Cómo estimar una media o proporción poblacional
Cómo seleccionar el tamaño muestral

5.1

DÓNDE HEMOS ESTADO

Los primeros cuatro capítulos de este libro han proporcionado el material que usted necesitará para entender la inferencia estadística y cómo se aplica en situaciones prácticas. Los primeros tres capítulos se refieren al uso de estadísticas descriptivas, tanto gráficas como numéricas, para describir e interpretar conjuntos de mediciones.

El capítulo 4 proporcionó el vínculo entre la probabilidad y la inferencia estadística. Numerosas estadísticas son sumas o promedios calculados de mediciones muestrales.

El teorema central del límite dice que, incluso si las poblaciones no son normales, las distribuciones de esas *estadísticas* serán aproximadamente normales cuando el tamaño muestral n es grande. Estas estadísticas son las herramientas que se usarán para las *estadísticas inferenciales*, es decir, hacer inferencias acerca de una población usando información contenida en una muestra.

5.2

A DÓNDE VAMOS: INFERENCIA ESTADÍSTICA

La inferencia, específicamente la toma de decisiones y la predicción, tiene siglos de antigüedad y desempeña un papel muy importante en la vida de casi todas las personas. Veamos a continuación algunas aplicaciones:

- El gobierno necesita predecir las tasas de interés a corto y largo plazos.
- Un corredor financiero desea pronosticar el comportamiento del mercado de acciones.
- Un metalurgista desea determinar si un nuevo tipo de acero es más resistente a altas temperaturas que el actual.
- Una consumidora desea estimar el precio de venta de su casa antes de ponerla en el mercado.

Hay muchas formas de tomar estas decisiones o predicciones, algunas son subjetivas y otras son de naturaleza más objetiva. ¿Qué tan buenas serán las predicciones o decisiones? Aun cuando usted piense que su propia capacidad de tomar decisiones es muy buena, la experiencia sugiere que éste puede no ser el caso. Es la función del estadístico matemático dar métodos para hacer inferencia estadística que son mejores y más confiables que únicamente las suposiciones subjetivas.

La inferencia estadística se ocupa de tomar decisiones o hacer predicciones acerca de **parámetros**, es decir, las medidas numéricas descriptivas que caracterizan a una población. Tres parámetros que encontramos en capítulos anteriores son la media poblacional μ , la desviación poblacional estándar σ y la proporción binomial p . En inferencia estadística, un problema práctico se expone de otra forma en el marco de una población con un parámetro específico de interés. Por ejemplo, el metalurgista podría medir el *promedio* de coeficientes de expansión de ambos tipos de acero y luego comparar sus valores.

Los métodos para hacer inferencias acerca de parámetros poblacionales caen en una de dos categorías:

- **Estimación:** Estimar o predecir el valor del parámetro
- **Prueba de hipótesis:** Tomar una decisión acerca del valor de un parámetro, con base en alguna idea preconcebida acerca de cuál podría ser su valor

TIP MI CONSEJO

Parámetro $\leftrightarrow$ población
Estadística $\leftrightarrow$ muestra

EJEMPLO

5.1

Los circuitos en computadoras y otros equipos electrónicos están formados por una o más tarjetas de circuito impreso (PCB) y es frecuente que las computadoras sean reparadas con solo cambiar una o más tarjetas defectuosas. En un intento por hallar el ajuste apropiado de un proceso de chapa aplicado a uno de los lados de una PCB, un supervisor de producción podría *estimar* el grosor promedio de chapa de cobre en las PCB usando muestras de varios días de operación. Como no sabe del grosor promedio μ antes de observar el proceso de producción, su problema es de *estimación*.

EJEMPLO

5.2

El supervisor del ejemplo 5.1 recibe instrucciones del propietario de la planta de que el grosor de la chapa de cobre no debe ser menor que .001 de pulgada, para que el proceso esté en control. Para decidir si el proceso está o no en control, el supervisor debe formular una prueba. Podría *hacer una hipótesis* de que el proceso está en control, es decir, suponer que el grosor promedio de la chapa de cobre es .001 o mayor, y usar muestras de varios días de operación para decidir si es o no correcta su hipótesis. El método de la toma de decisión del supervisor se denomina *prueba de hipótesis*.

¿Cuál método de inferencia debe usarse? Esto es, ¿debe estimarse el parámetro o probar una hipótesis respecto a su valor? La respuesta está dictada por la pregunta práctica planteada y a veces es determinada por preferencias personales. Como la estimación y las pruebas de hipótesis se usan con frecuencia en literatura científica, incluimos ambos métodos en este capítulo.

Un problema estadístico, que comprende planeación, análisis y formulación de inferencias, está incompleto sin una medida de la **bondad de la inferencia**. Esto es, ¿qué tan preciso o confiable es el método empleado? Si una corredora financiera predice que el precio de una acción será de 80 dólares el próximo lunes, ¿estaría usted dispuesto a comprar o vender su acción sin saber qué tan confiable es la predicción? ¿La predicción estará a no más de 1, 2 o 10 dólares del precio real el próximo lunes? Los procedimientos estadísticos son importantes porque dan dos tipos de información:

- Métodos para hacer la inferencia
- Una medida numérica de la bondad o confiabilidad de la inferencia

5.3

TIPOS DE ESTIMADORES

Para estimar el valor de un parámetro poblacional puede usar información de la muestra en la forma de un **estimador**. Los estimadores se calculan usando información de las observaciones muestrales y, en consecuencia, por definición son también *estadísticas*.

Definición Un **estimador** es una regla, generalmente expresada como fórmula, que nos dice cómo calcular una estimación basada en información de la muestra.

Los estimadores se usan en dos formas diferentes:

- **Estimación puntual:** Con base en datos muestrales, se calcula un solo número para estimar el parámetro poblacional. La regla o fórmula que describe este cálculo se denomina **estimador puntual** y el número resultante recibe el nombre de **estimación puntual**.

- **Estimación de intervalo:** Con base en datos muestrales, dos números se calculan para formar un intervalo dentro del cual se espera esté el parámetro. La regla o fórmula que describe este cálculo se denomina **estimador de intervalo** y el par de números resultantes se llama **estimación de intervalo** o **intervalo de confianza**.

EJEMPLO

5.3

Un veterinario desea estimar el aumento mensual promedio en el peso de cachorros de raza golden retriever, de cuatro meses de edad, que han sido puestos a dieta de carne de cordero y arroz. La *población* está formada por los aumentos mensuales en el peso de todos los cachorros de raza golden retriever a los que se da esta dieta particular. El veterinario desea estimar el parámetro desconocido μ , el aumento mensual promedio en el peso para esta población *hipotética*.

Un posible *estimador* basado en datos muestrales es la media muestral, $\bar{x} = \sum x_i / n$. Podría usarse en la forma de un solo número o *estimación puntual*, por ejemplo, 3.8 libras, o usarse una *estimación de intervalo* y estimar que el aumento promedio en el peso será entre 2.7 y 4.9 libras.

Los procedimientos de estimación tanto puntuales como de intervalo usan información dada por la distribución muestral del estimador específico que se haya elegido para usarse. Empezaremos por exponer la *estimación puntual* y su uso para estimar medias poblacionales y proporciones.

5.4

ESTIMACIÓN PUNTUAL**TIP** MI CONSEJO

Parámetro = diana del blanco

Estimador = bala o flecha

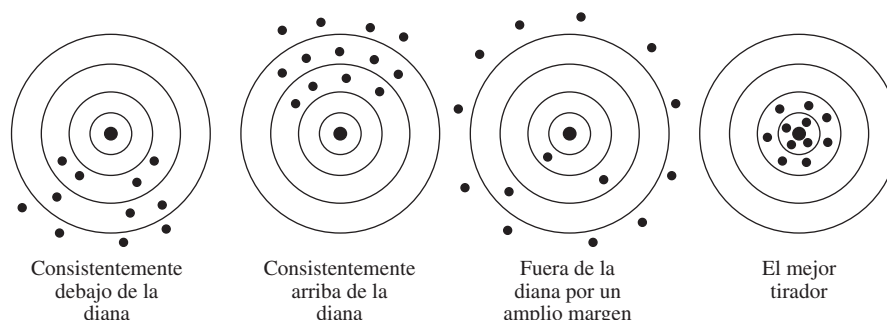
En una situación práctica, existen varias estadísticas que podrían usarse como estimadores puntuales para un parámetro poblacional. Para determinar cuál de las opciones es mejor, usted necesita saber cómo se comporta el estimador en muestreo repetido, descrito por su *distribución muestral*.

Como una analogía, considere disparar con un revólver hacia un blanco. El parámetro de interés es la diana a la cual se disparan balas. Cada bala representa una sola estimación muestral, disparada por el revólver, que representa el estimador.

Suponga que un amigo dispara una sola bala y acierta en la diana. ¿Se puede concluir que él es un excelente tirador? ¿Se pondría usted de pie junto al blanco cuando él dispare una segunda bala? Es probable que no, porque no tiene una medida de lo bien que él dispare en intentos repetidos. ¿Siempre acierta en el blanco o sus tiros son demasiado altos o demasiado bajos en forma consistente? ¿Sus tiros se agrupan alrededor del blanco o fallan en forma consistente en acertar en el blanco por un amplio margen? La figura 5.1 muestra varias configuraciones del blanco. ¿Cuál blanco elegiría usted como perteneciente al mejor tiro?

**FIGURA 5.1**

¿Cuál tirador es el mejor?



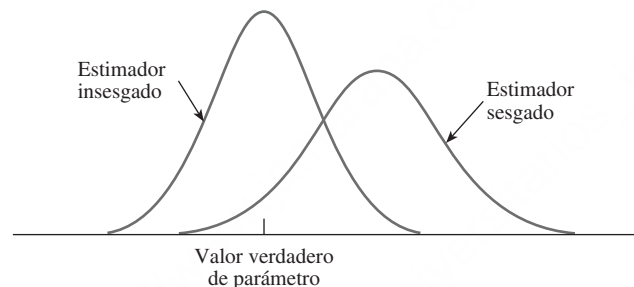
Las distribuciones muestrales dan información que se puede usar para seleccionar el **mejor estimador**. ¿Qué características serían valiosas? Primero, la **distribución muestral del estimador puntual debe estar centrada sobre el valor verdadero del parámetro que se estimará**. Esto es, el estimador no debe subestimar o sobreestimar de manera constante al parámetro de interés. Un estimador como éste se dice que es **insesgado**.

Definición Se dice que un estimador de un parámetro es **insesgado** si la media de su distribución es igual al valor verdadero del parámetro. De otro modo, se dice que el estimador está **sesgado**.

Las distribuciones muestrales para un estimador insesgado y un estimador sesgado se presentan en la figura 5.2. La distribución muestral para el estimador sesgado se encuentra hacia la derecha del valor verdadero del parámetro. Este estimador sesgado es más probable que uno insesgado para sobreestimar el valor del parámetro.

FIGURA 5.2

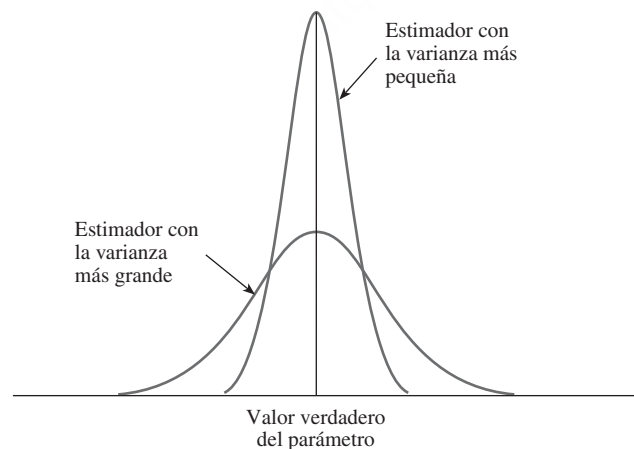
Distribuciones para estimadores sesgados e insesgados



La segunda característica deseable de un estimador es que **la dispersión (medida por la varianza) de la distribución muestral debe ser tan pequeña como sea posible**. Esto asegura que, con una alta probabilidad, una estimación individual quedará cerca del valor verdadero del parámetro. Las distribuciones muestrales para dos estimadores insesgados, una con una varianza pequeña[†] y la otra con una varianza más grande, se presentan en la figura 5.3.

FIGURA 5.3

Comparación de variabilidad de un estimador



[†] En general, los estadísticos usan el término *varianza de un estimador* cuando en realidad es la varianza de la distribución muestral del estimador. Esta expresión contraída se usa casi universalmente.

Por supuesto que sería preferible el estimador con la varianza más pequeña, porque las estimaciones tienden a estar más cerca del valor verdadero del parámetro que en la distribución con la varianza más grande.

En situaciones muestrales prácticas es posible saber que la distribución muestral de un estimador está centrada alrededor del parámetro que se trate de estimar, pero todo lo que se tiene es la estimación calculada de las n mediciones contenidas en la muestra. ¿A qué distancia del valor verdadero del parámetro estará esta estimación? ¿Qué tan cercana está la bala del tirador de la diana o blanco? La distancia entre la estimación y el valor verdadero del parámetro se denomina **error de estimación**.

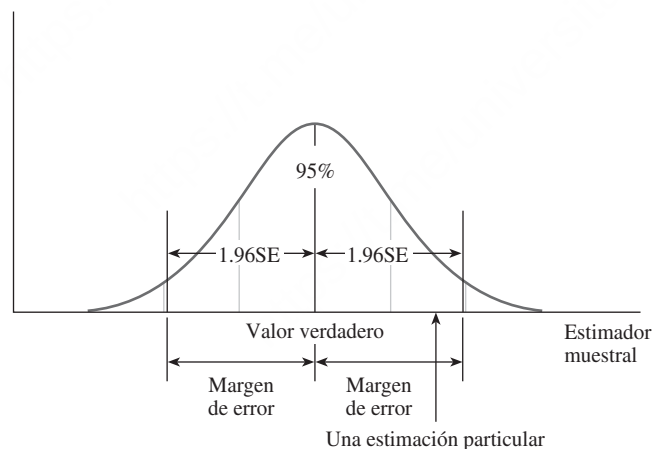
Definición La distancia entre una estimación y el parámetro estimado recibe el nombre de **error de estimación**.

En este capítulo, usted puede suponer que los tamaños muestrales son siempre grandes y, por tanto, que los estimadores *insesgados* que estudiará tienen distribuciones muestrales que pueden ser aproximadas por una distribución normal (por el teorema central del límite). Recuerde que, para cualquier estimador puntual con una distribución normal, la regla empírica dice que aproximadamente 95% de todas las estimaciones puntuales estarán a no más de dos (o más exactamente, 1.96) desviaciones estándar de la media de esa distribución.

Para estimadores *insesgados*, esto implica que la diferencia entre el estimador puntual y el valor verdadero del parámetro será menor que 1.96 desviaciones estándar o 1.96 errores estándar (SE). Esta cantidad, llamada el 95% de **margen de error** (o simplemente “**margen de error**”), da un límite superior práctico para el error de estimación (véase la figura 5.4). Es posible que el error de estimación exceda este margen de error, pero eso es muy poco probable.

FIGURA 5.4

Distribución muestral de un estimador insesgado



TIP MI CONSEJO

95% de margen de error
= $1.96 \times$ error estándar

ESTIMACIÓN PUNTUAL DEL PARÁMETRO DE UNA POBLACIÓN

- Estimador puntual: estadística calculada usando mediciones muestrales
- 95% de margen de error: $1.96 \times$ error estándar del estimador

Las distribuciones muestrales para dos estimadores puntuales insesgados se estudiaron en el capítulo 4. Se puede demostrar que estos dos estimadores puntuales tienen la *mínima variabilidad* de todos los estimadores insesgados y, por lo tanto, son los *mejores estimadores* que se pueden hallar en cada situación.

La variabilidad del estimador se mide usando este error estándar. No obstante, usted habrá observado que el error estándar suele depender de parámetros desconocidos, por ejemplo σ o p . Estos parámetros deben estimarse usando estadísticas muestrales como s y $\hat{p}$. Aun cuando no es exactamente correcto, por lo general los experimentadores se refieren al error estándar estimado como *el error estándar*.



NECESITO SABER...

Cómo estimar una media o proporción poblacional

- Para estimar la media poblacional μ para una población cuantitativa, el estimador puntual $\bar{x}$ es *insesgado* con el error estándar estimado como

$$SE = \frac{s}{\sqrt{n}}^\dagger$$

El 95% de margen de error cuando $n \geq 30$ se estima como

$$\pm 1.96 \left(\frac{s}{\sqrt{n}} \right)$$

- Para estimar la proporción poblacional p para una población binomial, el estimador puntual $\hat{p} = x/n$ es insesgado, con un error estándar estimado como

$$SE = \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

El 95% de margen de error se estima como

$$\pm 1.96 \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Suposiciones: $n\hat{p} > 5$ y $n\hat{q} > 5$.

EJEMPLO

5.4

Un ambientalista realiza un estudio del oso polar, especie que se encuentra en el océano Ártico y sus alrededores. Su zona de distribución está limitada por la existencia de hielo en el mar, que usa como plataforma para cazar focas, principal sostén de su dieta. La destrucción de su hábitat en el hielo del Ártico, que se ha atribuido al calentamiento global, amenaza la supervivencia de los osos como especie; puede extinguirse antes de un siglo.¹ Una muestra aleatoria de $n = 50$ osos polares produjo un peso promedio de 980 libras con una desviación estándar de 105 libras. Use esta información para estimar el peso promedio de todos los osos polares del Ártico.

Solución La variable aleatoria medida es el peso, una variable aleatoria cuantitativa mejor descrita por su media μ . La estimación puntual de μ , el peso promedio de todos los osos polares del Ártico, es $\bar{x} = 980$ libras. El margen de error se estima como

$$1.96 SE = 1.96 \left(\frac{s}{\sqrt{n}} \right) = 1.96 \left(\frac{105}{\sqrt{50}} \right) = 29.10 \approx 29 \text{ libras}$$

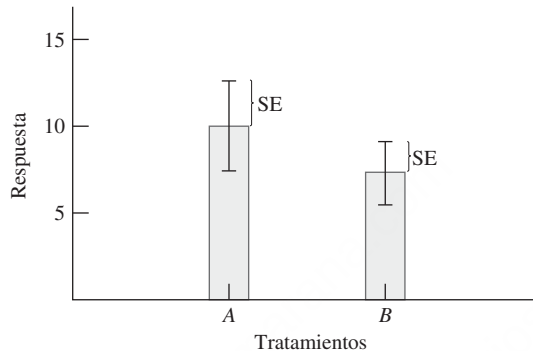
[†] Cuando se muestrea a partir de una distribución normal, el estadístico $(\bar{x} - \mu)/(s/\sqrt{n})$ tiene una distribución t , que se estudiará más adelante. Cuando la muestra es *grande*, este estadístico se encuentra distribuido normalmente en forma aproximada si la población muestreada es normal o no normal.

Se puede tener confianza en que la estimación muestral de 980 libras está a no más de ± 29 libras de la media poblacional.

Al informar los resultados de una investigación es frecuente que los investigadores agreguen ya sea la desviación muestral estándar (a veces llamada SD) o el error estándar $s/\sqrt{n}$ (por lo general llamado SE o SEM) a las estimaciones de medias poblacionales. Siempre se debe buscar una explicación en el texto del informe que diga si el investigador está informando $\bar{x} \pm SD$ o $\bar{x} \pm SE$. Además, las medias muestrales y desviaciones estándar o errores estándar se presentan a menudo como “barras de error” usando el formato gráfico que se ilustra en la figura 5.5.

FIGURA 5.5

Gráfica de medias de tratamiento y sus errores estándar



EJEMPLO

5.5

Además del peso promedio del oso polar del Ártico, el ambientalista del ejemplo 5.4 también está interesado en las opiniones de adultos sobre el tema del calentamiento global. En particular, desea estimar la proporción de personas que piensan que el calentamiento global es un problema muy serio. En una muestra aleatoria de $n = 100$ adultos, 73% indicó que el calentamiento global es un problema muy serio. Estime la verdadera proporción de población de adultos que piensan que el calentamiento global es un problema muy serio y encuentre el margen de error para la estimación.

Solución El parámetro de interés es ahora p , la proporción de personas en la población que piensan que el calentamiento global es un problema muy serio. El mejor estimador de p es la proporción muestral $\hat{p}$, que para esta muestra es $\hat{p} = .73$. Para hallar el margen de error, usted puede aproximar el valor de p con su estimación $\hat{p} = .73$:

$$1.96 \text{ SE} = 1.96 \sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.96 \sqrt{\frac{.73(.27)}{100}} = .09$$

Con este margen de error, se puede estar bastante seguro de que la estimación de .73 está dentro de $\pm .09$ del valor verdadero de p . En consecuencia, se puede concluir que el valor verdadero de p podría ser de sólo .64 o de hasta .82. Este margen de error es bastante grande cuando se compara con la estimación misma y refleja el hecho de que se requiere de muestras grandes para alcanzar un margen de error pequeño cuando se estime p .

TABLA 5.1

Algunos valores calculados de $\sqrt{pq}$

p	pq	$\sqrt{pq}$	p	pq	$\sqrt{pq}$
.1	.09	.30	.6	.24	.49
.2	.16	.40	.7	.21	.46
.3	.21	.46	.8	.16	.40
.4	.24	.49	.9	.09	.30
.5	.25	.50			

La tabla 5.1 muestra la forma en que el numerador del error estándar de $\hat{p}$ cambia para diversos valores de p . Observe que, para casi todos los valores de p , en especial cuando p está entre .3 y .7, hay muy poco cambio en $\sqrt{pq}$, el numerador del SE, que alcanza su máximo valor cuando $p = .5$. Esto significa que el margen de error usando el estimador $\hat{p}$ también será máximo cuando $p = .5$. Cuando estiman p , algunos entrevistadores rutinariamente usan el margen máximo de error, que a veces recibe el nombre de **error muestral**, en cuyo caso calculan

$$1.96 \text{ SE} = 1.96 \sqrt{\frac{.5(.5)}{n}} \quad \text{o a veces} \quad 2 \text{ SE} = 2 \sqrt{\frac{.5(.5)}{n}}$$

Las encuestas Gallup, Harris y Roper generalmente usan tamaños muestrales de alrededor de 1000, de modo que su margen de error es

$$1.96 \sqrt{\frac{.5(.5)}{1000}} = .031 \quad \text{o sea alrededor de 3\%}$$

En este caso, se dice que la estimación está dentro de ± 3 puntos porcentuales de la verdadera proporción de población.

5.4 EJERCICIOS

TÉCNICAS BÁSICAS

5.1 Explique qué significa “margen de error” en estimación puntual.

5.2 ¿Cuáles son dos características del mejor estimador puntual para un parámetro poblacional?

5.3 Calcule el margen de error al estimar una media poblacional μ para estos valores:

- a. $n = 30$, $\sigma^2 = .2$ b. $n = 30$, $\sigma^2 = .9$
c. $n = 30$, $\sigma^2 = 1.5$

5.4 Consulte el ejercicio 5.3. ¿Qué efecto tiene una mayor varianza poblacional sobre el margen de error?

5.5 Calcule el margen de error al estimar una media poblacional μ para estos valores:

- a. $n = 50$, $s^2 = 4$ b. $n = 500$, $s^2 = 4$
c. $n = 5000$, $s^2 = 4$

5.6 Consulte el ejercicio 5.5. ¿Qué efecto tiene un tamaño muestral aumentado sobre el margen de error?

5.7 Calcule el margen de error al estimar una proporción binomial para cada uno de los siguientes valores de n . Use $p = .5$ para calcular el error estándar del estimador.

- a. $n = 30$ b. $n = 100$
c. $n = 400$ d. $n = 1000$

5.8 Consulte el ejercicio 5.7. ¿Qué efecto tiene aumentar el tamaño muestral sobre el margen de error?

5.9 Calcule el margen de error al estimar una proporción binomial p usando muestras de tamaño $n = 100$ y los siguientes valores para p :

a. $p = .1$ b. $p = .3$ c. $p = .5$

d. $p = .7$ e. $p = .9$

f. ¿Cuál de los valores de p produce el máximo margen de error?

5.10 Suponga que está usted escribiendo un cuestionario para una encuesta muestral que comprende $n = 100$ individuos. El cuestionario generará estimaciones para varias proporciones binomiales diferentes. Si desea informar un solo margen de error para la encuesta, ¿qué margen de error del ejercicio 5.9 es el correcto para usar?

5.11 Una muestra aleatoria de $n = 900$ observaciones de una población binomial produjo $x = 655$ éxitos. Estime la proporción binomial p y calcule el margen de error.

5.12 Una muestra aleatoria de $n = 50$ observaciones de una población cuantitativa produjo $\bar{x} = 56.4$ y $s^2 = 2.6$. Dé la mejor estimación puntual para la media poblacional μ , y calcule el margen de error.

5.13 Una muestra aleatoria de $n = 500$ observaciones de una población binomial produjo $x = 450$ éxitos. Estime la proporción binomial p y calcule el margen de error.

5.14 Una muestra aleatoria de $n = 75$ observaciones de una población cuantitativa produjo $\bar{x} = 29.7$ y $s^2 = 10.8$. Dé la mejor estimación puntual para la media poblacional μ y calcule el margen de error.

APLICACIONES

5.15 La falla de San Andrés Una de las fracturas grandes más famosas es la falla de San Andrés, en California. Una geóloga que trata de estudiar el movimiento de la corteza

terrestre, en un lugar en particular, encontró numerosas fracturas en la estructura local de rocas. En un intento por determinar el ángulo medio de las roturas, muestreó $n = 50$ fracturas y encontró que la media muestral y la desviación estándar eran de 39.8° y 17.2° , respectivamente. Estime la dirección angular media de las fracturas y encuentre el margen de error para su estimación.

5.16 Biomasa Las estimaciones de la biomasa de la Tierra, es decir, la cantidad total de vegetación que hay en los bosques del planeta, son importantes para determinar la cantidad de dióxido de carbono no absorbido que se espera permanezca en la atmósfera terrestre.² Suponga que una muestra de 75 terrenos de un metro cuadrado, elegidos al azar en bosques boreales de América del Norte, produjo una biomasa media de 4.2 kilogramos por metro cuadrado (kg/m^2), con una desviación estándar de $1.5 \text{ kg}/\text{m}^2$. Estime el promedio de biomasa para los bosques boreales de América del Norte y encuentre el margen de error para su estimación.

Fuente: Reimpreso con permiso de *Science News*, publicación semanal de Science, copyright 1989 por Science Services, Inc.

5.17 Confianza del consumidor Un aumento en la tasa de ahorros del consumidor está con frecuencia relacionado con la falta de confianza en la economía y se dice que es un indicador de una tendencia a la recesión de la economía. Un muestreo aleatorio de $n = 200$ cuentas de ahorro en una comunidad local mostró un aumento medio en valores de cuentas de ahorro de 7.2% en los últimos 12 meses, con una desviación estándar de 5.6%. Estime el aumento medio en porcentaje en valores de cuentas

de ahorros de los últimos 12 meses para depositantes de esta comunidad. Encuentre el margen de error para su estimación.

5.18 Inmigración ilegal En un sondeo reciente que incluyó preguntas sobre inmigración ilegal en Estados Unidos, y las respuestas federales y estatales al problema, 75% de los $n = 1004$ adultos encuestados pensaban que Estados Unidos no se esforzaba lo suficiente para evitar que los inmigrantes ilegales llegaran a ese país.³

- ¿Cuál es una estimación puntual para la proporción de adultos que creen que Estados Unidos no se esfuerza lo suficiente para evitar que inmigrantes ilegales vayan a ese país? Calcule el margen de error.
- El sondeo refleja un margen de error de $\pm 3.5\%$. ¿Cómo debería calcularse el margen de error mencionado de modo que pueda aplicarse a todas las preguntas en la encuesta? ¿Es correcto el margen de error mencionado?

5.19 Ratas hambrientas En un experimento para evaluar la intensidad del instinto del hambre en ratas, 30 animales previamente entrenados fueron privados de alimento durante 24 horas. Al término de ese periodo, cada rata fue puesta en una jaula donde se les dio alimento si el animal presionaba una palanca. Para cada una, se registró el tiempo en el que continuaba presionando la barra (aun cuando no recibiera alimento). Si los datos dieron una media muestral de 19.3 minutos con una desviación estándar de 5.2 minutos, estime el verdadero tiempo medio y calcule el margen de error.

5.5

ESTIMACIÓN DE INTERVALO

Un *estimador de intervalo* es una regla para calcular dos números, por ejemplo a y b , para crear un intervalo del que usted esté bastante seguro que contiene el parámetro de interés. El concepto de “bastante seguro” significa “con gran probabilidad”. Medimos esta probabilidad usando el **coeficiente de confianza**, designado por $1 - \alpha$.

Definición La probabilidad de que un intervalo de confianza contenga el parámetro estimado se denomina **coeficiente de confianza**.

TIP MI CONSEJO

Cómo lazar:

Parámetro = poste de cerca

Estimación de intervalo = lazo



Por ejemplo, es frecuente que los experimentadores construyan intervalos de confianza de 95%, lo cual significa que el coeficiente de confianza, o la probabilidad de que el intervalo contenga el parámetro estimado, es .95. Puede aumentar o reducir la cantidad de certeza si se cambia el coeficiente de confianza. Algunos valores que por lo general usan experimentadores son .90, .95, .98 y .99.

Considere una analogía, esta vez lanzar un lazo a un poste de una cerca. El poste de la cerca representa el parámetro que se desea estimar y el lazo formado por la cuerda representa el intervalo de confianza. Cada vez que se lance la cuerda, se espera lazar al poste de la cerca; no obstante, a veces falla el lazo. En la misma forma, cada vez que se saque una muestra y

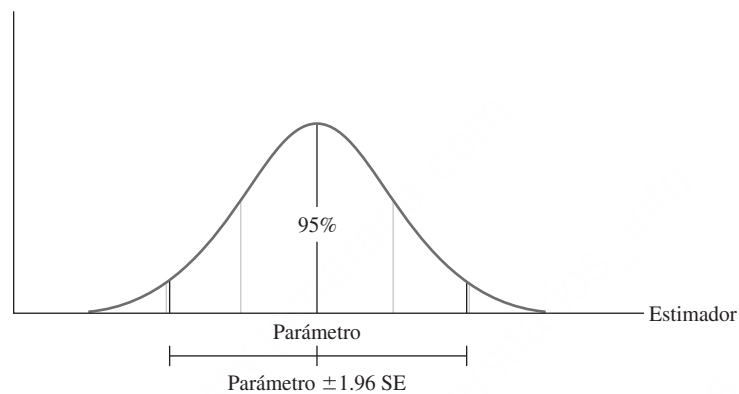
construya un intervalo de confianza para un parámetro, usted espera incluir el parámetro en su intervalo, pero, al igual que el lazo, a veces falla. Su “porcentaje de éxito”, es decir la proporción de intervalos que “lazan al poste” en muestreo repetido, es el coeficiente de confianza.

Construcción de un intervalo de confianza

Cuando la distribución muestral de un estimador puntual es aproximadamente normal, se puede construir un estimador de intervalo o **intervalo de confianza** mediante el siguiente razonamiento. Para mayor sencillez, suponga que el coeficiente de confianza es .95 y consulte la figura 5.6.

FIGURA 5.6

Parámetro ± 1.96 SE



- Sabemos que, de todos los valores posibles del estimador que podríamos seleccionar, 95% de ellos estarán en el intervalo

Parámetro ± 1.96 SE

que se muestra en la figura 5.6.

- Como el valor del parámetro es desconocido, considere construir el intervalo

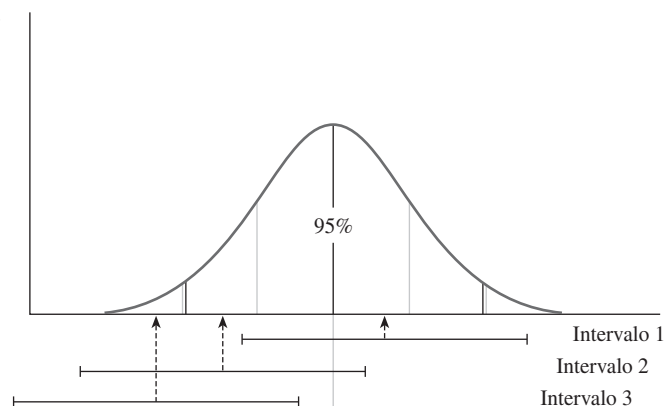
Estimador ± 1.96 SE

que tiene el mismo ancho que el primer intervalo, pero tiene un centro variable.

- ¿Con qué frecuencia funcionará este intervalo en forma correcta y encerrará el parámetro de interés? Consulte la figura 5.7.

FIGURA 5.7

Algunos intervalos de confianza de 95%

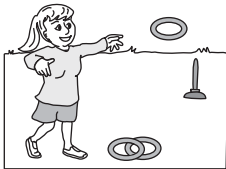


TIP MI CONSEJO

Al igual que en un juego de lanzar un anillo:

Parámetro = estacilla

Estimación de intervalo = anillo



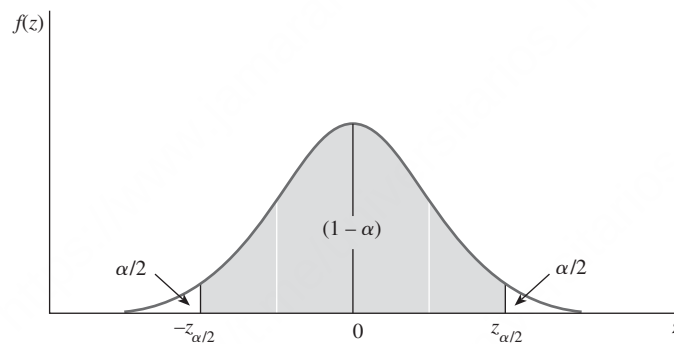
Los primeros dos intervalos funcionan correctamente, es decir, el parámetro (marcado con una línea gris claro) está contenido dentro de ambos intervalos. El tercer intervalo no funciona, porque no encierra al parámetro. Esto ocurrió porque el valor del estimador del centro del intervalo estaba demasiado lejos del parámetro. Por fortuna, valores del estimador sólo caen a esa distancia 5% de las veces y ¡nuestro procedimiento funcionará en forma correcta 95% del tiempo!

Si lo desea puede cambiar el *coeficiente de confianza* de $(1 - \alpha) = .95$ a otro nivel de confianza $(1 - \alpha)$. Para lograr esto, es necesario cambiar el valor $z = 1.96$, que localiza un área de .95 en el centro de la curva normal estándar, a un valor de z que localice el área $(1 - \alpha)$ en el centro de la curva, como se muestra en la figura 5.8.

Como el área total bajo la curva es 1, el área restante en las dos colas es α y cada cola contiene un área $\alpha/2$. El valor de z que tiene “área de cola” $\alpha/2$ a su derecha se denomina $z_{\alpha/2}$, y el área entre $-z_{\alpha/2}$ y $z_{\alpha/2}$ es el coeficiente de confianza $(1 - \alpha)$. Valores de $z_{\alpha/2}$, que por lo general son utilizados por experimentadores, le serán familiares cuando empiece a construir intervalos de confianza para diferentes situaciones prácticas. Algunos de estos valores se dan en la tabla 5.2.

FIGURA 5.8

Ubicación de $z_{\alpha/2}$



INTERVALO DE CONFIANZA DE MUESTRA GRANDE (1 - α)100%

$$(\text{Estimador puntual}) \pm z_{\alpha/2} \times (\text{error estándar del estimador})$$

donde $z_{\alpha/2}$ es el valor z con un área $\alpha/2$ en la cola derecha de una distribución normal estándar. Esta fórmula genera dos valores; el **límite inferior de confianza (LCL)** y el **límite superior de confianza (UCL)**.

TABLA 5.2

Valores de z que comúnmente se usan para intervalos de confianza

Cociente de confianza (1 - α)	α	$\alpha/2$	$z_{\alpha/2}$
.90	.10	.05	1.645
.95	.05	.025	1.96
.98	.02	.01	2.33
.99	.01	.005	2.58

Intervalo de confianza de muestra grande para una media poblacional μ

Es muy frecuente que problemas prácticos lleven a la estimación de μ , la media de una población de mediciones cuantitativas. He aquí algunos ejemplos:

- El promedio de calificaciones de estudiantes universitarios en una universidad particular
- El promedio de resistencia de un nuevo tipo de acero
- El número promedio de fallecimientos por categoría de edad
- El promedio de demanda para un nuevo producto de cosmético

Cuando el tamaño muestral n es grande, la media muestral $\bar{x}$ es el mejor estimador puntual para la media poblacional μ . Como su distribución muestral es aproximadamente normal, puede usarse para construir un intervalo de confianza de acuerdo con el método general dado ya antes.

UN INTERVALO DE CONFIANZA DE MUESTRA GRANDE (1 - α)100% PARA UNA MEDIA POBLACIONAL μ

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

donde $z_{\alpha/2}$ es el valor z correspondiente a un área $\alpha/2$ en la cola superior de una distribución z normal estándar y

n = tamaño muestral

σ = desviación estándar de la población muestreada

Si σ es desconocida, puede ser aproximada por la desviación estándar muestral s cuando el tamaño muestral sea grande ($n \geq 30$) y el intervalo aproximado de confianza es

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

Otra forma de hallar el intervalo de confianza de muestra grande para una media poblacional μ es empezar con la estadística

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

que tiene una distribución normal estándar. Si escribimos $z_{\alpha/2}$ como el valor de z con área $\alpha/2$ a su derecha, entonces se puede escribir

$$P\left(-z_{\alpha/2} < \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1 - \alpha$$

Esta desigualdad puede reescribirse como

$$\begin{aligned} -z_{\alpha/2} \frac{\sigma}{\sqrt{n}} &< \bar{x} - \mu < z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \\ -\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} &< -\mu < -\bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \end{aligned}$$

de modo que

$$P\left(\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Los términos $\bar{x} - z_{\alpha/2}(\sigma/\sqrt{n})$ y $\bar{x} + z_{\alpha/2}(\sigma/\sqrt{n})$, que son los límites inferior y superior de confianza, son en realidad cantidades aleatorias que dependen de la media muestral $\bar{x}$. Por tanto, en muestreo repetido, el intervalo aleatorio $\bar{x} \pm z_{\alpha/2}(\sigma/\sqrt{n})$, contendrá la media poblacional μ con probabilidad $(1 - \alpha)$.

EJEMPLO

5.6

Un nutriólogo seleccionó una muestra aleatoria de $n = 50$ adultos hombres. Se encontró que el promedio de ingesta diaria de productos lácteos fue de $\bar{x} = 756$ gramos por día, con una desviación estándar de $s = 35$ gramos por día. Use esta información muestral para construir un intervalo de confianza de 95% para la ingesta diaria media de productos lácteos para hombres.

Solución Como el tamaño muestral de $n = 50$ es grande, la distribución de la media muestral $\bar{x}$ está distribuida normalmente en forma aproximada, con media μ y error estándar estimado por $s/\sqrt{n}$. El intervalo de confianza aproximado de 95% es

$$\begin{aligned} \bar{x} \pm 1.96\left(\frac{s}{\sqrt{n}}\right) \\ 756 \pm 1.96\left(\frac{35}{\sqrt{50}}\right) \\ 756 \pm 9.70 \end{aligned}$$

TIP MI CONSEJO

Un intervalo de confianza de 95% nos dice que, si construyéramos muchos de estos intervalos (todos los cuales tendrían puntos extremos ligeramente diferentes), 95% de ellos encerrarían la media poblacional

Por tanto, el intervalo de confianza de 95% para μ es de 746.30 a 765.70 gramos por día.

Interpretación del intervalo de confianza

¿Qué significa decir que estamos “95% confiados” que el valor real de la media poblacional μ está dentro de un intervalo determinado? Si construyéramos 20 de esos intervalos, cada uno usando diferente información muestral, nuestros intervalos se verían como los de la figura 5.9a). De los 20 intervalos, podría esperarse que 95% de ellos, o sea 19 de cada 20, funcionarían como se planea y contienen μ dentro de sus límites superior e inferior. Si se construyeran 100 de dichos intervalos (figura 5.9b)), se esperaría que alrededor de 95 de ellos se desem-

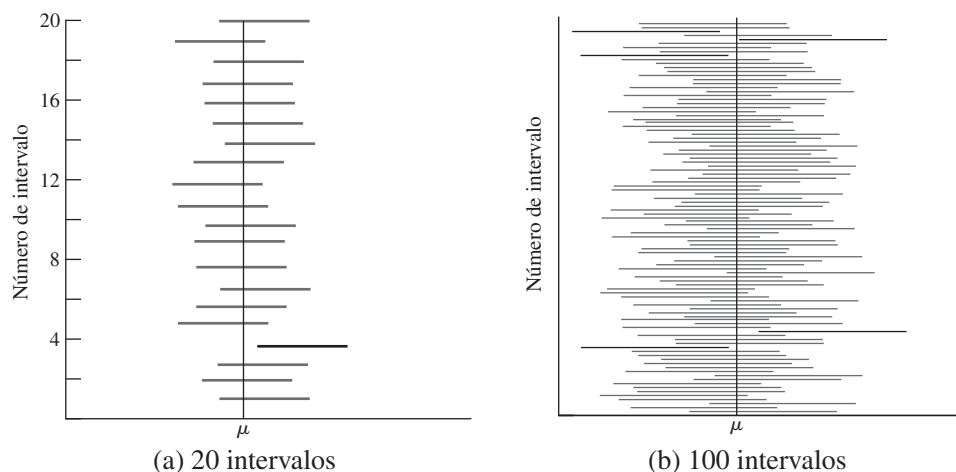
FIGURA 5.9

Interpretación de los intervalos de confianza



APPLET EN LÍNEA

Interpretación de intervalos de confianza



peñaran según lo planeado. Recuerde que no se puede estar absolutamente seguro de que algún intervalo particular contenga la media μ . Nunca se sabrá si ese intervalo particular es uno de los 19 que “funcionaron”, o si es el intervalo que “faltaba”. La confianza en el intervalo estimado proviene del hecho de que cuando se calculan intervalos repetidos, 95% de esos intervalos contendrán μ .

Un buen intervalo de confianza tiene dos características deseables:

- Es tan angosto como es posible. Cuanto más angosto sea el intervalo, más exactamente se habrá localizado el parámetro estimado.
- Tiene un coeficiente de confianza grande, cercano a 1. Cuanto mayor sea el coeficiente de confianza, es más probable que el intervalo contenga el parámetro estimado.

EJEMPLO

5.7

Construya un intervalo de confianza de 99% para la ingesta diaria media de productos lácteos para los hombres adultos del ejemplo 5.6.

Solución Para cambiar el nivel de confianza a .99 se debe hallar el valor apropiado de la z normal estándar que pone el área $(1 - \alpha) = .99$ en el centro de la curva. Este valor, con área de cola $\alpha/2 = .005$ a su derecha, se encuentra de la tabla 5.2 como $z = 2.58$ (véase la figura 5.10). El intervalo de confianza de 99% es entonces

$$\bar{x} \pm 2.58 \left(\frac{s}{\sqrt{n}} \right)$$

$$756 \pm 2.58(4.95)$$

$$756 \pm 12.77$$

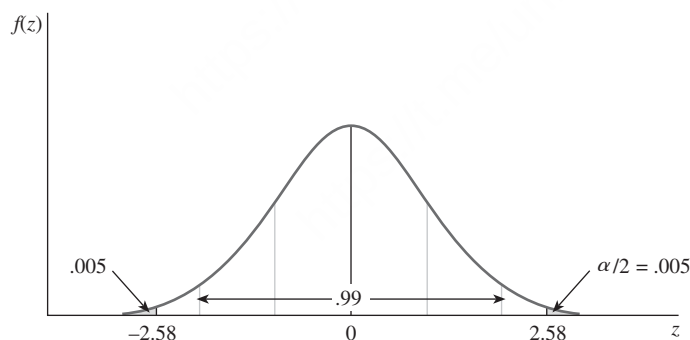
es decir, 743.23 a 768.77 gramos por día. Este intervalo de confianza es *más ancho* que el intervalo de confianza de 95% del ejemplo 5.6.

FIGURA 5.10

Valores estándar normales para un intervalo de confianza de 99%

TIP MI CONSEJO

Área de la cola derecha	Valor z
.05	1.645
.025	1.96
.01	2.33
.005	2.58



El ancho aumentado es necesario para incrementar la confianza, ¡igual que como se desearía un anillo más ancho en su lazo para asegurarse de lazar el poste de una cerca! La única forma de *aumentar la confianza* sin aumentar el ancho del intervalo es *aumentar el tamaño muestral, n* .

El error estándar de $\bar{x}$,

$$SE = \frac{\sigma}{\sqrt{n}}$$



mide la variabilidad o dispersión de los valores de $\bar{x}$. Cuanto más variables sean los datos poblacionales, medidos por σ , más variable será $\bar{x}$ y el error estándar será más grande. Por otra parte, si se aumenta el tamaño muestral n , habrá más información para estimar μ . Las estimaciones deben caer más cerca de μ y el error estándar será más pequeño.

Los intervalos de confianza de los ejemplos 5.6 y 5.7 son aproximados porque se substituyó s como una aproximación para σ . Esto es, en lugar de que el coeficiente de confianza sea .95, el valor especificado en el ejemplo, el valor verdadero del coeficiente puede ser .92, .94 o .97. Pero esta discrepancia es de poco interés desde un punto de vista práctico; en lo que se refiere a la “confianza” del usuario, hay poca diferencia entre estos coeficientes de confianza. Casi todos los estimadores que se emplean en estadística dan intervalos de confianza aproximados, porque las suposiciones sobre las que están basados no se satisfacen exactamente. Habiendo visto este punto, no continuaremos refiriéndonos a intervalos de confianza como “aproximados”. Es de poco interés práctico mientras el coeficiente real de confianza sea cercano al valor especificado.

Intervalo de confianza de muestra grande para una proporción poblacional p

Muchos experimentos de investigación o estudios muestrales tienen como objetivo la estimación de la proporción de personas u objetos de un grupo grande, que posean cierta característica. Veamos algunos ejemplos:

- La proporción de ventas que se puede esperar en un gran número de contactos con clientes.
- La proporción de semillas que germinan.
- La proporción de electores “probables” que planean votar por un candidato político particular.

Cada uno es un ejemplo práctico del experimento binomial y el parámetro a estimarse es la proporción binomial p .

Cuando el tamaño muestral es grande, la proporción muestral,

$$\hat{p} = \frac{x}{n} = \frac{\text{Número total de éxitos}}{\text{Número total de intentos}}$$

es el mejor estimador puntual para la proporción poblacional p . Como su distribución muestral es aproximadamente normal, con media p y error estándar $SE = \sqrt{pq/n}$, $\hat{p}$ puede usarse para construir un intervalo de confianza de acuerdo con el método general dado en esta sección.

TIP

MI CONSEJO

Área de la cola derecha	Valor z
.05	1.645
.025	1.96
.01	2.33
.005	2.58

UN INTERVALO DE CONFIANZA DE MUESTRA GRANDE $(1 - \alpha)100\%$ PARA UNA PROPORCIÓN POBLACIONAL p

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{pq}{n}}$$

donde $z_{\alpha/2}$ es el valor z correspondiente a un área de $\alpha/2$ en la cola derecha de una distribución normal estándar z . Como p y q son incógnitas, se estiman con el uso de los mejores

estimadores puntuales: $\hat{p}$ y $\hat{q}$. El tamaño muestral se considera grande cuando la aproximación normal a la distribución binomial es adecuada, es decir, cuando $n\hat{p} > 5$ y $n\hat{q} > 5$.

EJEMPLO**5.8**

Una muestra aleatoria de 985 “probables” electores, es decir los que probablemente voten en la próxima elección, fueron encuestados durante un maratón telefónico realizado por el Partido Republicano. De ellos, 592 indicaron que tenían la intención de votar por la candidata republicana en la próxima elección. Construya un intervalo de confianza de 90% para p , la proporción de electores probables de la población que tienen la intención de votar por la candidata republicana. Con base en esta información, ¿se puede concluir que la candidata ganará la elección?

Solución La estimación puntual para p es

$$\hat{p} = \frac{x}{n} = \frac{592}{985} = .601$$

y el error estándar estimado es

$$\sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{(.601)(.399)}{985}} = .016$$

El valor z para un intervalo de confianza de 90% es el valor que tiene área de $\alpha/2 = .05$ en la cola superior de la distribución z , o $z_{.05} = 1.645$ de la tabla 5.2. El intervalo de confianza del 90% para p es entonces

$$\hat{p} \pm 1.645 \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$.601 \pm .026$$

Es decir, $.575 < p < .627$. Se estima que el porcentaje de probables electores que tienen intención de votar por la candidata republicana es entre 57.5% y 62.7%. ¿La candidata ganará la elección? Suponiendo que ella necesita más del 50% de los votos para ganar, y como los límites superior e inferior de confianza exceden de este valor mínimo, se puede decir con 90% de confianza que la candidata ganará.

Hay algunos problemas, no obstante, con este tipo de encuesta muestral. ¿Qué pasa si los electores que se consideran a sí mismos “probables para votar” en realidad no van a las casillas? ¿Qué pasa si un elector cambia de idea entre ahora y el día de la elección? ¿Qué pasa si un elector entrevistado no responde honestamente cuando el trabajador de la campaña le hace preguntas? El intervalo de confianza de 90% que ha construido le da 90% de confianza sólo si ha seleccionado una *muestra aleatoria de la población de interés*. Ya no se puede estar seguro de “90% de confianza” si su muestra es sesgada, o ¡si la población de respuestas de votantes cambia antes del día de la elección!

Es posible que usted haya observado que el estimador puntual con su 95% de margen de error se ve muy semejante a un intervalo de confianza de 95% para el mismo parámetro. Esta cercana relación existe para casi todos los parámetros estimados en este libro, pero no es verdadera en general. A veces el mejor estimador puntual para un parámetro *no* cae en la mitad del mejor intervalo de confianza; el mejor intervalo de confianza puede no ser siquiera una función del mejor estimador puntual. Aun cuando ésta es una distinción teórica, debe recordarse que hay una diferencia entre estimación puntual y estimación de intervalo, y que la elección entre las dos depende de la preferencia del experimentador.

5.5

EJERCICIOS

TÉCNICAS BÁSICAS

5.20 Encuentre e interprete un intervalo de confianza de 95% para una media poblacional μ para estos valores:

- a. $n = 36, \bar{x} = 13.1, s^2 = 3.42$
 b. $n = 64, \bar{x} = 2.73, s^2 = .1047$

5.21 Encuentre un intervalo de confianza de 90% para una media poblacional μ para estos valores:

- a. $n = 125, \bar{x} = .84, s^2 = .086$
 b. $n = 50, \bar{x} = 21.9, s^2 = 3.44$
 c. Interprete los intervalos hallados en las partes a y b.

5.22 Encuentre un intervalo de confianza $(1 - \alpha)100\%$ para una media poblacional μ para estos valores:

- a. $\alpha = .01, n = 38, \bar{x} = 34, s^2 = 12$
 b. $\alpha = .10, n = 65, \bar{x} = 1049, s^2 = 51$
 c. $\alpha = .05, n = 89, \bar{x} = 66.3, s^2 = 2.48$

5.23 Una muestra aleatoria de $n = 300$ observaciones de una población binomial produjo $x = 263$ éxitos. Encuentre un intervalo de confianza de 90% para p e interprete el intervalo.

5.24 Suponga que el número de éxitos observado en $n = 500$ intentos de un experimento binomial es 27. Encuentre un intervalo de confianza de 95% para p . ¿Por qué el intervalo de confianza es más angosto que el intervalo de confianza del ejercicio 5.23?

5.25 Se selecciona una muestra aleatoria de n mediciones de una población con media μ desconocida y desviación estándar $\sigma = 10$ conocida. Calcule el ancho de un intervalo de confianza de 95% para μ para estos valores de n :

- a. $n = 100$ b. $n = 200$ c. $n = 400$

5.26 Compare los intervalos de confianza del ejercicio 5.25. ¿Qué efecto tiene cada una de estas acciones sobre el ancho de un intervalo de confianza?

- a. Duplique el tamaño muestral
 b. Cuadruple el tamaño muestral

5.27 Consulte el ejercicio 5.26.

- a. Calcule el ancho de un intervalo de confianza de 90% para μ cuando $n = 100$.
 b. Calcule el ancho de un intervalo de confianza de 99% para μ cuando $n = 100$.
 c. Compare los anchos de intervalos de confianza de 90%, 95% y 99% para μ . ¿Qué efecto tiene un creciente

coeficiente de confianza sobre el ancho del intervalo de confianza?

APLICACIONES

5.28 Un experimento de química En un experimento de electrólisis, un grupo de estudiantes midió la cantidad de cobre precipitado de una solución saturada de sulfato de cobre en un periodo de 30 minutos. Los $n = 30$ estudiantes calcularon una media muestral y desviación estándar igual a .145 y .0051 moles, respectivamente. Encuentre un intervalo de confianza de 90% para la cantidad media de cobre precipitado de la solución en un periodo de 30 minutos.

5.29 Lluvia ácida La lluvia ácida, causada por la reacción de ciertos contaminantes del aire con el agua de lluvia, parece ser un problema creciente en Estados Unidos. La lluvia pura que cae en aire limpio registra un valor pH de 5.7 (el pH es una medida de la acidez: 0 es ácido; 14 es alcalino). Suponga que muestras de agua de 40 lluvias se analizan para el contenido del pH y $\bar{x}$ y s son iguales a 3.7 y .5, respectivamente. Encuentre un intervalo de confianza de 99% para el pH medio en agua de lluvia e interprete el intervalo. ¿Qué suposición debe hacerse para que el intervalo de confianza sea válido?

5.30 Las minivans Una encuesta muestral está diseñada para estimar la proporción de minivans en el estado de California. Se selecciona una muestra aleatoria de 500 registros de una base de datos del Departamento de Vehículos de Motor y 68 se clasifican como minivans.

- a. Use un intervalo de confianza de 95% para estimar la proporción de minivans en California.
 b. ¿Cómo se puede estimar la proporción de minivans en California, con un grado más alto de precisión? (SUGERENCIA: Hay dos respuestas.)

5.31 ¿Qué es normal? ¿Qué es normal, cuando se trata de temperaturas corporales de personas? Una muestra aleatoria de 130 temperaturas corporales humanas tenía una media de 98.25 grados y una desviación estándar de 0.73 grados.

- a. Construya un intervalo de confianza de 99% para la temperatura corporal promedio de personas sanas.
 b. ¿El intervalo de confianza construido en la parte a contiene el valor de 98.6 grados, que es la temperatura promedio usual citada por médicos y otros especialistas? Si no es así, ¿qué conclusiones se pueden sacar?

5.6

ESTIMACIÓN DE LA DIFERENCIA ENTRE DOS MEDIAS POBLACIONALES

Un problema de igual importancia que la estimación de una sola media poblacional μ , para una población cuantitativa, es la comparación de dos medias poblacionales. Usted puede hacer comparaciones como éstas:

- Las calificaciones promedio del examen de admisión para la facultad de medicina (MCAT) para estudiantes cuya especialización fuera bioquímica, y para aquellos cuya especialización fuera biología
- Las producciones promedio en una planta química que usa materias primas suministradas por dos proveedores diferentes
- El promedio de diámetros de tallos de plantas cultivadas con dos tipos diferentes de nutrientes

Para cada uno de estos ejemplos, hay dos poblaciones: la primera con media y varianza μ_1 y σ_1^2 , y la segunda con media y varianza μ_2 y σ_2^2 . Una muestra aleatoria de n_1 mediciones se saca de la población 1 y una segunda muestra aleatoria de tamaño n_2 se saca de manera independiente de la población 2. Por último, las estimaciones de los parámetros poblacionales se calculan a partir de los datos muestrales usando los estimadores $\bar{x}_1$, s_1^2 , $\bar{x}_2$ y s_2^2 , como se presenta en la tabla 5.3.

TABLA 5.3 Muestras de dos poblaciones cuantitativas

	Población 1	Población 2
Media	μ_1	μ_2
Varianza	σ_1^2	σ_2^2

	Muestra 1	Muestra 2
Media	$\bar{x}_1$	$\bar{x}_2$
Varianza	s_1^2	s_2^2
Tamaño de la muestra	n_1	n_2

Intuitivamente, la diferencia entre dos medias muestrales daría la máxima información acerca de la diferencia real entre dos medias poblacionales y éste es de hecho el caso. El mejor estimador puntual de la diferencia ($\mu_1 - \mu_2$) entre las medias poblacionales es $(\bar{x}_1 - \bar{x}_2)$. La distribución muestral de este estimador no es difícil de deducir, pero la expresamos aquí sin demostración.

PROPIEDADES DE LA DISTRIBUCIÓN MUESTRAL DE $(\bar{x}_1 - \bar{x}_2)$, LA DIFERENCIA ENTRE DOS MEDIAS MUESTRALES

Cuando muestras aleatorias independientes de n_1 y n_2 observaciones han sido seleccionadas de poblaciones con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2 , respectivamente, la distribución muestral de la diferencia $(\bar{x}_1 - \bar{x}_2)$ tiene las siguientes propiedades:

1. La media de $(\bar{x}_1 - \bar{x}_2)$ es

y el error estándar es

$$SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

que se puede estimar como

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \text{ cuando los tamaños muestrales son grandes.}$$

2. Si las poblaciones muestreadas están distribuidas normalmente, entonces la distribución muestral de $(\bar{x}_1 - \bar{x}_2)$ está distribuida normalmente **exactamente**, cualquiera que sea el tamaño muestral.
3. Si las poblaciones muestreadas no están distribuidas normalmente, entonces la distribución muestral de $(\bar{x}_1 - \bar{x}_2)$ está distribuida normalmente **aproximadamente** cuando n_1 y n_2 son ambas de 30 o más, debido al teorema central del límite.

Como $(\mu_1 - \mu_2)$ es la media de la distribución muestral, se deduce que $(\bar{x}_1 - \bar{x}_2)$ es un estimador insesgado de $(\mu_1 - \mu_2)$ con una distribución aproximadamente normal cuando n_1 y n_2 son grandes. Esto es, el estadístico

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

tiene una distribución z normal aproximadamente estándar y los procedimientos generales de la sección 5.5 se pueden usar para construir estimaciones puntuales y de intervalo. Aun cuando la elección entre estimación puntual y de intervalo depende de la preferencia personal del usuario, casi todos los experimentadores eligen construir intervalos de confianza para problemas de dos muestras. Las fórmulas apropiadas para ambos métodos se dan a continuación.

TIP MI CONSEJO

Área de la cola derecha	Valor z
.05	1.645
.025	1.96
.01	2.33
.005	2.58

ESTIMACIÓN PUNTUAL DE $(\mu_1 - \mu_2)$ DE MUESTRA GRANDE

Estimador puntual: $(\bar{x}_1 - \bar{x}_2)$

$$95\% \text{ margen de error: } \pm 1.96 SE = \pm 1.96 \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

UN INTERVALO DE CONFIANZA DE MUESTRA GRANDE DE $(1 - \alpha)100\%$ PARA $(\mu_1 - \mu_2)$

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

EJEMPLO

5.9

Las resistencias al desgaste de dos tipos de llantas para automóvil se compararon en muestras de pruebas en camino de $n_1 = n_2 = 100$ llantas para cada tipo y se registró el número de millas recorridas hasta su deterioro, definido como una cantidad específica de desgaste de la llanta. Los resultados de la prueba se muestran en la tabla 5.4. Estime $(\mu_1 - \mu_2)$, la diferencia en la media de millas hasta el completo desgaste, usando un intervalo de confianza de 99%. ¿Hay diferencia en el promedio de calidad de desgaste para los dos tipos de llantas?

TABLA 5.4 ● **Resumen de datos muestrales para dos tipos de llantas**

Llanta 1	Llanta 2
$\bar{x}_1 = 26\,400$ millas	$\bar{x}_2 = 25\,100$ millas
$s_1^2 = 1\,440\,000$	$s_2^2 = 1\,960\,000$

Solución La estimación puntual de $(\mu_1 - \mu_2)$ es

$$(\bar{x}_1 - \bar{x}_2) = 26\,400 - 25\,100 = 1\,300 \text{ millas}$$

y el error estándar de $(\bar{x}_1 - \bar{x}_2)$ se estima como

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{1\,440\,000}{100} + \frac{1\,960\,000}{100}} = 184.4 \text{ millas}$$

El intervalo de confianza de 99% se calcula como

$$\begin{aligned} (\bar{x}_1 - \bar{x}_2) \pm 2.58 \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \\ 1300 \pm 2.58(184.4) \\ 1300 \pm 475.8 \end{aligned}$$

es decir, $824.2 < (\mu_1 - \mu_2) < 1775.8$. La diferencia en el promedio de millas hasta el completo desgaste para los dos tipos de llantas se estima que está entre el límite inferior de confianza LCL = 824.2 y el límite superior de confianza UCL = 1775.8 millas.

Con base en este intervalo de confianza, ¿se puede concluir que hay una diferencia en el promedio de millas hasta el completo desgaste para los dos tipos de llantas? Si no hubiera diferencia en las dos medias poblacionales, entonces μ_1 y μ_2 serían iguales y $(\mu_1 - \mu_2) = 0$. Si observamos el intervalo de confianza construido, se verá que 0 no es uno de los posibles valores para $(\mu_1 - \mu_2)$. Por tanto, no es probable que las medias sean iguales; se puede concluir que hay una diferencia en el promedio de millas hasta el completo desgaste para los dos tipos de llantas. El intervalo de confianza ha permitido *tomar una decisión* acerca de la igualdad de las dos medias poblacionales.

TIP MI CONSEJO

Si 0 no está en el intervalo, *se puede* concluir que hay una diferencia en las medias poblacionales

EJEMPLO

5.10

El científico del ejemplo 5.6 se preguntaba si había diferencia en el promedio de ingesta diaria de productos lácteos entre hombres y mujeres. Tomó una muestra de $n_1 = 50$ hombres adultos y $n_2 = 50$ mujeres adultas y registró sus ingestas diarias de productos lácteos en gramos por día. En la tabla 5.5 se presenta un resumen de sus resultados muestrales. Construya un intervalo de confianza de 95% para la diferencia en el promedio de ingestas diarias de productos lácteos para hombres y mujeres. ¿Se puede concluir que hay una diferencia en el promedio de ingestas diarias para hombres y mujeres?

TABLA 5.5 ● **Valores muestrales para ingestas diarias de productos lácteos**

	Hombres	Mujeres
Tamaño muestral	50	50
Media muestral	756	762
Desviación estándar muestral	35	30

Solución El intervalo de confianza se construye usando un valor de z con área de cola $\alpha/2 = .025$ a su derecha, esto es, $z_{.025} = 1.96$. Usando las desviaciones muestrales estándar para aproximar las desviaciones estándar poblacionales desconocidas, el intervalo de 95% de confianza es

$$\begin{aligned}
 (\bar{x}_1 - \bar{x}_2) \pm 1.96 \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \\
 (756 - 762) \pm 1.96 \sqrt{\frac{35^2}{50} + \frac{30^2}{50}} \\
 -6 \pm 12.78
 \end{aligned}$$

o bien, $-18.78 < (\mu_1 - \mu_2) < 6.78$. Veamos los posibles valores para $(\mu_1 - \mu_2)$ del intervalo de confianza. Es posible que la diferencia $(\mu_1 - \mu_2)$ pudiera ser negativa (lo cual indica que el promedio para mujeres excede del promedio para hombres), ser positiva (lo cual indica que los hombres tienen el promedio más alto) o ser 0 (lo cual indica que no hay diferencia entre los promedios). Con base en esta información, *no estaríamos dispuestos a concluir* que hay una diferencia en el promedio de ingestas diarias de productos lácteos para hombres y mujeres.

Los ejemplos 5.9 y 5.10 merecen más comentarios respecto a usar estimaciones muestrales en lugar de parámetros desconocidos. La distribución muestral de

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

tiene una distribución normal estándar para todos los tamaños muestrales cuando ambas poblaciones muestreadas son normales, y una distribución normal estándar *aproximada* cuando las poblaciones muestreadas no sean normales pero los tamaños muestrales sean grandes (≥ 30). Cuando σ_1^2 y σ_2^2 no se conocen y son estimadas por las estimaciones muestrales s_1^2 y s_2^2 , la estadística resultante todavía tendrá una distribución normal estándar aproximada cuando los tamaños muestrales sean grandes. El comportamiento de esta estadística cuando las varianzas poblacionales son desconocidas, y los tamaños muestrales sean pequeños, se estudiará en el capítulo 6.

5.6

EJERCICIOS

TÉCNICAS BÁSICAS

5.32 Se seleccionaron muestras aleatorias independientes de las poblaciones 1 y 2. Los tamaños muestrales, medias y varianzas son como sigue:

	Población	
	1	2
Tamaño muestral	35	49
Media muestral	12.7	7.4
Varianza muestral	1.38	4.14

- Encuentre un intervalo de confianza de 95% para estimar la diferencia en las medias poblacionales $(\mu_1 - \mu_2)$.
- Con base en el intervalo de confianza de la parte a, ¿se puede concluir que hay una diferencia en las medias para las dos poblaciones? Explique.

5.33 Se seleccionaron muestras aleatorias independientes de las poblaciones 1 y 2. Los tamaños muestrales, medias y varianzas son como sigue:

	Población	
	1	2
Tamaño muestral	64	64
Media muestral	2.9	5.1
Varianza muestral	0.83	1.67

- Encuentre un intervalo de confianza de 90% para la diferencia en las medias poblacionales. ¿Qué significa la frase “90% confiado”?
- Encuentre un intervalo de confianza de 99% para la diferencia en las medias poblacionales. ¿Puede concluir que hay una diferencia en las dos medias poblacionales? Explique.

5.34 Se seleccionaron muestras aleatorias independientes de tamaño $n_1 = n_2 = 100$ de cada una de dos poblaciones. La media y la desviación estándar para las dos muestras fueron $\bar{x}_1 = 125.2$, $\bar{x}_2 = 123.7$, $s_1 = 5.6$ y $s_2 = 6.8$.

- a. Construya un intervalo de confianza de 99% para estimar la diferencia en las dos medias poblacionales.
- b. ¿El intervalo de confianza en la parte a proporciona suficiente evidencia para concluir que hay una diferencia en las dos medias poblacionales? Explique.

5.35 Se seleccionaron muestras aleatorias independientes de tamaño $n_1 = n_2 = 500$ de cada una de dos poblaciones. La media y la desviación estándar para las dos muestras fueron $\bar{x}_1 = 125.2$, $\bar{x}_2 = 123.7$, $s_1 = 5.6$ y $s_2 = 6.8$.

- a. Encuentre una estimación puntual para la diferencia en las dos medias poblacionales. Calcule el margen de error.
- b. Con base en los resultados en la parte a, ¿puede concluir que hay una diferencia en las dos medias poblacionales? Explique.

APLICACIONES

5.36 Selenio Se considera esencial para una buena salud una pequeña cantidad del oligoelemento selenio, 50-200 microgramos (μg) por día. Suponga que se seleccionaron muestras aleatorias de $n_1 = n_2 = 30$ adultos de dos regiones de Estados Unidos y que se registró la ingesta diaria de selenio para cada persona. La media y la desviación estándar de las ingestas diarias de selenio para los 30 adultos de la región 1 fueron $\bar{x}_1 = 167.1$ y $s_1 = 24.3 \mu\text{g}$, respectivamente. Las estadísticas correspondientes para los 30 adultos de la región 2 fueron $\bar{x}_2 = 140.9$ y $s_2 = 17.6 \mu\text{g}$. Encuentre un intervalo de confianza de 95% para la diferencia en las ingestas medias de selenio para las dos regiones. Interprete este intervalo.

5.37 9-1-1 Se realizó un estudio para comparar los números medios de llamadas de emergencia a la policía por turno de 8 horas en dos distritos de una gran ciudad. Se seleccionaron al azar muestras de 100 turnos de 8 horas de los registros policiales para cada una de las dos regiones y el número de llamadas de emergencia se registró para cada turno. Las estadísticas muestrales se indican a continuación:

	Región	
	1	2
Tamaño muestral	100	100
Media muestral	2.4	3.1
Varianza muestral	1.44	2.64

Encuentre un intervalo de confianza de 90% para la diferencia en los números medios de llamadas de emergencia a la policía por turno entre los dos distritos de la ciudad. Interprete el intervalo.

5.38 Enseñando biología Se realizó un experimento para comparar un currículum desarrollado por un maestro, “Biología: un contexto comunitario” (BACC, por sus siglas en inglés) basado en estándares, orientado en actividades y centrado en preguntas con la presentación tradicional usando impartición de clase, vocabulario y hechos memorizados. Los resultados quizá no tan sorprendentes de un examen sobre conceptos de biología, publicados en *The American Biology Teacher*, se muestran en la tabla siguiente.⁴

	Media	Tamaño muestral	Desviación estándar
Examen previo: Todos los grupos de BACC	13.38	372	5.59
Examen previo: Todos los tradicionales	14.06	368	5.45
Examen posterior: Todos los grupos de BACC	18.5	365	8.03
Examen posterior: Todos los tradicionales	16.5	298	6.96

- a. Encuentre un intervalo de confianza de 95% para la calificación media para el examen posterior para todos los grupos de BACC.
- b. Encuentre un intervalo de confianza de 95% para la calificación media del examen posterior para todos los grupos tradicionales.
- c. Encuentre un intervalo de confianza de 95% para la diferencia en calificaciones medias para los grupos BACC y los grupos tradicionales en el examen posterior.
- d. ¿El intervalo de confianza en c da evidencia de que hay una diferencia real en las calificaciones de grupo tradicional y BACC en el examen posterior? Explique.

Fuente: De “Performance Assessment of a Standards-Based High School Biology Curriculum”, por W. Leonard, B. Speziale y J. Pernick en *The American Biology Teacher*, 2001, 63(5), 310-316. Reimpreso con permiso de National Association of Biology Teachers.

5.39 ¿Está usted a dieta? Para comparar dos dietas para bajar de peso, A y B, se seleccionaron al azar 60 personas que seguían dietas. Un grupo de 30 personas se colocó en la dieta A y las otras 30 en la dieta B, y se registraron sus pérdidas de peso durante un periodo de 30 días. Las medias y desviaciones estándar de las mediciones de pérdida de peso para los dos grupos se muestran en la tabla. Encuentre un intervalo de confianza de 95% para la diferencia en pérdida media de peso para las dos dietas. Interprete su intervalo de confianza.

Dieta A	Dieta B
$\bar{x}_A = 21.3$	$\bar{x}_B = 13.4$
$s_A = 2.6$	$s_B = 1.9$

5.40 Salarios iniciales Como grupo, los estudiantes que se especializan en disciplinas de ingeniería tienen las expectativas salariales más altas, seguidos por aquellos que estudian en los campos de las ciencias computacionales, de acuerdo con resultados de la *Encuesta estudiantil 2010* de la NACE.⁵ Se seleccionaron muestras aleatorias de 50 recién graduados universitarios en cada especialidad para comparar los salarios iniciales de graduados en ingeniería

y ciencias computacionales, y se obtuvo la siguiente información:

Especialidad	Media (\$)	Desviación estándar
Ingeniería	56,202	2225
Ciencias computacionales	50,657	2375

- Encuentre una estimación puntual para la diferencia en salarios iniciales de estudiantes universitarios que se especializan en ingeniería y ciencias computacionales. ¿Cuál es el margen de error en su estimación?
- Con base en los resultados de la parte **a**, ¿piensa usted que hay una diferencia importante en el promedio de salarios iniciales para ingenieros y de ciencias computacionales? Explique.

5.41 Conocimientos de biología Consulte el ejercicio 5.38. Además de exámenes que comprenden conceptos de biología, los estudiantes también fueron examinados en conocimientos de procesos. Los resultados de exámenes previos y calificaciones de un examen posterior, publicados en *The American Biology Teacher*, se dan a continuación.⁴

	Media	Tamaño muestral	Desviación estándar
Examen previo: Todos los grupos BACC	10.52	395	4.79
Examen previo: Todos los tradicionales	11.97	379	5.39
Examen posterior: Todos los grupos BACC	14.06	376	5.65
Examen posterior: Todos los tradicionales	12.96	308	5.93

- Encuentre un intervalo de confianza de 95% de la puntuación media en habilidades del proceso para el examen posterior para todos los grupos BACC.
- Encuentre un intervalo de confianza de 95% de la puntuación media en habilidades del proceso para el examen posterior para todos los grupos tradicionales.
- Encuentre un intervalo de confianza de 95% para la diferencia en las puntuaciones medias en habilidades del proceso para el examen posterior para los grupos BACC y los tradicionales.
- ¿El intervalo de confianza de **c** proporciona evidencia de que hay una diferencia real en las puntuaciones medias

de las habilidades del proceso en el examen posterior entre los grupos BACC y los tradicionales? Explique.

Fuente: De "Performance Assessment of a Standards-Based High School Biology Curriculum", por W. Leonard, B. Speziale y J. Pernick en *The American Biology Teacher*, 2001, 63(5), 310-316. Reimpreso con permiso de National Association of Biology Teachers.

5.42 Ruido y estrés Para comparar el efecto del estrés en la forma de ruido sobre la capacidad de realizar un trabajo sencillo, 70 personas fueron divididas en dos grupos. El primer grupo de 30 personas actuó como control, en tanto que el segundo grupo de 40 fue el grupo experimental. Aun cuando cada persona realizó el trabajo, las personas del grupo experimental lo hicieron cuando se reproducía música de *rock* a alto volumen. Se registró el tiempo para terminar el trabajo para cada individuo y se obtuvo el siguiente resumen:

	Control	Experimental
n	30	40
$\bar{x}$	15 minutos	23 minutos
s	4 minutos	10 minutos

- Encuentre un intervalo de confianza de 99% para la diferencia en tiempos medios de terminación para estos dos grupos.
- Con base en el intervalo de confianza de la parte **a**, ¿hay suficiente evidencia para indicar una diferencia en el tiempo promedio de terminación para los dos grupos? Explique.

5.43 ¿Qué es normal?, continúa De las 130 personas del ejercicio 5.31, 65 eran mujeres y 65 eran hombres. Las medias y desviación estándar de sus temperaturas se muestran a continuación.

	Hombres	Mujeres
Media muestral	98.11	98.39
Desviación estándar	0.70	0.74

Encuentre un intervalo de confianza de 95% para la diferencia en el promedio de temperaturas corporales para hombres contra mujeres. Con base en este intervalo, ¿se puede concluir que hay una diferencia en el promedio de temperaturas para hombres contra mujeres? Explique.

ESTIMACIÓN DE LA DIFERENCIA ENTRE DOS PROPORCIONES BINOMIALES

5.7

Una simple extensión de la estimación de una proporción binomial p es la estimación de la diferencia entre dos proporciones binomiales. Se pueden hacer comparaciones como éstas:

- La proporción de artículos defectuosos manufacturados en dos líneas de producción
- La proporción de votantes mujeres y hombres que están a favor de una enmienda por la igualdad de derechos
- Los porcentajes de germinación de semillas no tratadas y semillas tratadas con un fungicida

Estas comparaciones pueden hacerse con la diferencia $(p_1 - p_2)$ entre dos proporciones binomiales, p_1 y p_2 . Muestras aleatorias independientes formadas por n_1 y n_2 intentos se sacan de poblaciones 1 y 2, respectivamente, y se calculan las estimaciones muestrales $\hat{p}_1$ y $\hat{p}_2$. El estimador insesgado de la diferencia $(p_1 - p_2)$ es la diferencia muestral $(\hat{p}_1 - \hat{p}_2)$.

PROPIEDADES DE LA DISTRIBUCIÓN MUESTRAL DE LA DIFERENCIA $(\hat{p}_1 - \hat{p}_2)$ ENTRE DOS PROPORCIONES MUESTRALES

Suponga que las muestras aleatorias independientes de las observaciones n_1 y n_2 han sido seleccionadas de poblaciones binomiales con parámetros p_1 y p_2 , respectivamente. La distribución muestral de la diferencia entre proporciones muestrales

$$(\hat{p}_1 - \hat{p}_2) = \left(\frac{x_1}{n_1} - \frac{x_2}{n_2} \right)$$

tiene estas propiedades:

1. La media de $(p_1 - p_2)$ es

$$p_1 - p_2$$

y el error estándar es

$$SE = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

que se estima como

$$SE = \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

2. La distribución muestral de $(p_1 - p_2)$ puede ser aproximada por una distribución normal cuando n_1 y n_2 son grandes, debido al teorema central del límite.

Aun cuando el margen de una proporción individual es de 0 a 1, la diferencia entre dos proporciones va de -1 a 1 . Para usar una distribución normal para aproximar la distribución de $(\hat{p}_1 - \hat{p}_2)$, tanto $\hat{p}_1$ como $\hat{p}_2$ deben ser aproximadamente normales; esto es, $n_1 \hat{p}_1 > 5$, $n_1 \hat{q}_1 > 5$ y $n_2 \hat{p}_2 > 5$, $n_2 \hat{q}_2 > 5$.

Las fórmulas apropiadas para estimación puntual y de intervalo se dan a continuación.

ESTIMACIÓN PUNTUAL DE MUESTRA GRANDE DE $(p_1 - p_2)$ Estimador puntual: $(\hat{p}_1 - \hat{p}_2)$

$$95\% \text{ de margen de error: } \pm 1.96 \text{ SE} = \pm 1.96 \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

**UN INTERVALO DE CONFIANZA DE MUESTRA GRANDE
(1 - α)100% PARA $(p_1 - p_2)$**

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Suposición: n_1 y n_2 deben ser suficientemente grandes para que la distribución muestral de $(\hat{p}_1 - \hat{p}_2)$ pueda ser aproximada por una distribución normal, es decir, si $n_1 \hat{p}_1$, $n_1 \hat{q}_1$, $n_2 \hat{p}_2$ y $n_2 \hat{q}_2$ son todas mayores que 5.

EJEMPLO

5.11

La propuesta de un bono para la construcción de una escuela será enviada a los votantes en la siguiente elección municipal. Una parte importante del dinero derivado de esta emisión de bonos se empleará en construir escuelas en una zona de rápido desarrollo de la ciudad y lo demás se usará para renovar y actualizar los edificios escolares del resto de ésta. Para evaluar la viabilidad de la propuesta de un bono, a una muestra aleatoria de $n_1 = 50$ residentes de la zona de rápido desarrollo y $n_2 = 100$ de las otras partes de la ciudad, se les preguntó si piensan votar por la propuesta. Los resultados se tabulan en la tabla 5.6.

TABLA 5.6

Valores muestrales para opinión sobre propuesta de bono

	Sección de desarrollo	Resto de la ciudad
Tamaño muestral	50	100
Número a favor de propuesta	38	65
Proporción a favor de propuesta	.76	.65

1. Estime la diferencia en las proporciones verdaderas a favor de la propuesta de bono con un 99% de intervalo de confianza.
2. Si ambas muestras se agrupan en una muestra de tamaño $n = 150$, con 103 a favor de la propuesta, dé una estimación puntual de la proporción de residentes de la ciudad que votarán por la propuesta del bono. ¿Cuál es el margen de error?

Solución

1. La mejor estimación puntual de la diferencia $(p_1 - p_2)$ está dada por

$$(\hat{p}_1 - \hat{p}_2) = .76 - .65 = .11$$

y el error estándar de $(\hat{p}_1 - \hat{p}_2)$ se estima como

$$\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} = \sqrt{\frac{(.76)(.24)}{50} + \frac{(.65)(.35)}{100}} = .0770$$

Para un intervalo de confianza de 99%, $z_{.005} = 2.58$ y el intervalo aproximado de confianza de 99% se encuentra como

$$(\hat{p}_1 - \hat{p}_2) \pm z_{.005} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$.11 \pm (2.58)(.0770)$$

$$.11 \pm .199$$

o bien, $(-.089, .309)$. Como este intervalo contiene el valor $(p_1 - p_2) = 0$, es posible que $p_1 = p_2$, lo cual implica que puede no haber diferencia en las proporciones a favor del asunto del bono en las dos secciones de la ciudad.

2. Si no hay diferencia en las dos proporciones, entonces las dos muestras no son realmente diferentes y podrían combinarse bien para obtener una estimación total de la proporción de los residentes de la ciudad que votarán por el asunto del bono. Si ambas muestras se agrupan, entonces $n = 150$ y

$$\hat{p} = \frac{103}{150} = .69$$

Por tanto, la estimación puntual del valor total de p es .69, con un margen de error dado por

$$\pm 1.96 \sqrt{\frac{(.69)(.31)}{150}} = \pm 1.96(.0378) = \pm .074$$

Observe que $.69 \pm .074$ produce el intervalo .62 a .76, que incluye sólo proporciones mayores que .5. Por tanto, si las actitudes de los votantes no cambian de manera adversa antes de la elección, la propuesta del bono debe aprobarse por una mayoría razonable.

5.7 EJERCICIOS

TÉCNICAS BÁSICAS

5.44 Se seleccionaron muestras aleatorias independientes de $n_1 = 500$ y $n_2 = 500$ observaciones de las poblaciones binomiales 1 y 2, y se observaron $x_1 = 120$ y $x_2 = 147$ éxitos.

- ¿Cuál es el mejor estimador puntual para la diferencia $(p_1 - p_2)$ en las dos proporciones binomiales?
- Calcule el error estándar aproximado para la estadística empleada en la parte a.
- ¿Cuál es el margen de error para esta estimación puntual?

5.45 Se seleccionaron muestras aleatorias independientes de $n_1 = 800$ y $n_2 = 640$ observaciones de las poblaciones binomiales 1 y 2, y se observaron $x_1 = 337$ y $x_2 = 374$ éxitos.

- Encuentre un intervalo de confianza de 90% para la diferencia $(p_1 - p_2)$ de las dos proporciones poblacionales. Interprete el intervalo.
- ¿Qué suposiciones deben hacerse para que el intervalo de confianza sea válido? ¿Se satisfacen estas suposiciones?

5.46 Se seleccionaron muestras aleatorias independientes de $n_1 = 1\,265$ y $n_2 = 1\,688$ observaciones de las poblaciones binomiales 1 y 2, y se observaron $x_1 = 849$ y $x_2 = 910$ éxitos.

- Encuentre un intervalo de confianza de 99% para la diferencia $(p_1 - p_2)$ de las dos proporciones poblacionales. ¿Qué significa “99% de confianza”?
- Con base en el intervalo de confianza de la parte a, ¿puede concluirse que hay una diferencia en las dos proporciones binomiales? Explique.

APLICACIONES

5.47 Prioridades diferentes Conforme se aproximaban las elecciones a mitad de legislatura, en el verano de 2010, demócratas y republicanos estaban divididos acerca de las principales prioridades de la nación.⁶ Se hizo la siguiente pregunta a una muestra de $n = 900$ votantes registrados: “¿Cuál de los siguientes puntos piensa usted que es más importante que atienda el gobierno federal en este preciso momento?” La lista de opciones se muestra en la tabla siguiente. Las opciones se rotaron para reducir el sesgo, y

a los votantes se les permitió indicar “Todos”, “Ninguno” o “No estoy seguro”.

	Todos (%)	Demócratas (%)	Republicanos (%)	Independientes (%)
Economía y empleos	47	55	37	48
Déficit del gasto	15	8	22	16
Terrorismo, Seguridad	8	6	10	10
Irak y Afganistán	7	9	5	4
Inmigración	5	3	7	6
Todos (vol.)	16	18	16	13
Ninguno/Otro (vol.)	1	-	1	2
No estoy seguro	1	-	1	-

Suponga que hay 400 demócratas, 350 republicanos y 150 independientes en la muestra. Use un procedimiento de estimación de muestra grande para comparar las proporciones de republicanos y demócratas que mencionaron economía y empleos como el punto más importante para que sea atendido por el gobierno federal. Compare las proporciones de republicanos e independientes que mencionaron el déficit del gasto como el punto más importante. Explique sus conclusiones.

5.48 Aficionados al beisbol El primer día del beisbol es a fines de marzo y termina en octubre con la Serie Mundial. ¿El apoyo de los aficionados aumenta a medida que avanza la temporada? Dos encuestas de CNN/*USA Today*/Gallup, una realizada en marzo y la otra en noviembre, contenían muestras aleatorias de 1001 adultos de 18 años de edad o mayores. En la muestra de marzo, 45% de los adultos dijeron ser aficionados del beisbol profesional, en tanto que 51% de los adultos de la muestra de noviembre dijo que eran aficionados.⁷

- Construya un intervalo de confianza de 99% para la diferencia en la proporción de adultos que dicen ser aficionados en marzo contra noviembre.
- ¿Los datos indican que la proporción de adultos que dicen ser aficionados aumenta en noviembre, más o menos en el tiempo de la Serie Mundial? Explique.

5.49 Pescar un resfrió ¿Las personas cultas tienen menos resfriados? Un estudio del *Chronicle of Higher Education* fue realizado por científicos de la Carnegie Mellon University, la Universidad de Pittsburgh y la Universidad de Virginia. Encontraron que las personas

que tienen sólo unas pocas reuniones sociales tienen más resfriados que quienes participan en varias actividades sociales.⁸ Suponga que de los 276 hombres y mujeres sanos examinados, $n_1 = 96$ tenían sólo pocas reuniones sociales y $n_2 = 105$ estaban ocupados con seis o más actividades. Cuando estas personas se exponían al virus del resfriado, se observaron los siguientes resultados:

	Pocas reuniones sociales	Muchas reuniones Sociales
Tamaño muestral	96	105
Porcentaje con resfriado	62%	35%

- Construya un intervalo de confianza de 99% para la diferencia en las dos proporciones poblacionales.
- ¿Parece haber una diferencia en las proporciones poblacionales para los dos grupos?
- Podría pensarse que entrar en contacto con más personas llevaría a más resfriados, pero los datos muestran el efecto opuesto. ¿Cómo se puede explicar este hallazgo inesperado?

5.50 Orden de nacimiento y éxito universitario En un estudio de la relación entre el orden de nacimiento y el éxito universitario, un investigador encontró que 126 de una muestra de 180 graduados universitarios eran primogénitos o hijos únicos. En una muestra de 100 no graduados de edad y nivel socioeconómico comparables, el número de primogénitos o hijos únicos fue 54. Estime la diferencia entre las proporciones de primogénitos o hijos únicos en las dos poblaciones de las cuales se tomaron estas muestras. Use un intervalo de confianza de 90% e interprete sus resultados.

5.51 ¿Excedrin o Tylenol? En un estudio para comparar los efectos de dos analgésicos se encontró que, de $n_1 = 200$ personas seleccionadas al azar que usaron el primer analgésico, 93% indicó que alivió su dolor. De $n_2 = 450$ personas seleccionadas al azar que usaron el segundo analgésico, 96% indicó que les alivió el dolor.

- Encuentre un intervalo de confianza de 99% para la diferencia en las proporciones que experimentan alivio por estos dos analgésicos.
- Con base en el intervalo de confianza de la parte a, ¿hay suficiente evidencia para indicar una diferencia en las proporciones que experimentan alivio para los dos analgésicos? Explique.

LÍMITES DE CONFIANZA A UNA COLA

Los intervalos de confianza estudiados en las secciones 5.5 a 5.7 en ocasiones reciben el nombre de **intervalos de confianza a dos colas**, porque producen límites superiores (UCL) e inferiores (LCL) para el parámetro de interés. Sin embargo, a veces un experimentador está interesado en sólo uno de estos límites; esto es, necesita sólo un límite superior (o posiblemente un límite inferior) para el parámetro de interés. En este caso, se puede construir un **límite de confianza de una cola** para el parámetro de interés, por ejemplo μ , p , $\mu_1 - \mu_2$, o $p_1 - p_2$.

Cuando la distribución muestral de un estimador puntual es aproximadamente normal, se puede usar un argumento similar al de la sección 5.5 para mostrar que los límites de confianza de una cola, contruidos usando las siguientes ecuaciones *cuando el tamaño muestral es grande*, contendrán el valor verdadero del parámetro de interés $(1 - \alpha)100\%$ de las veces en muestreo repetido.

UN LÍMITE INFERIOR DE CONFIANZA $(1 - \alpha)100\%$ (LCB)

$$(\text{Estimador puntual}) - z_{\alpha} \times (\text{Error estándar del estimador})$$

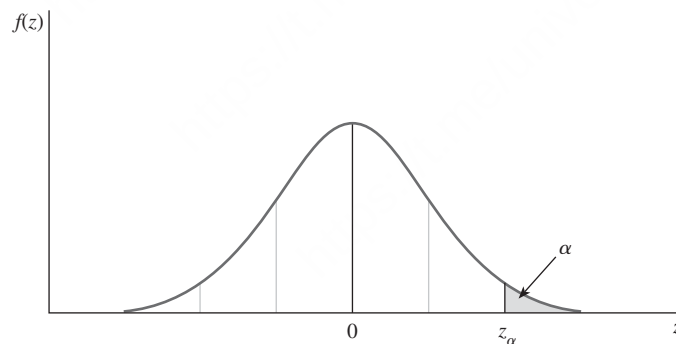
UN LÍMITE SUPERIOR DE CONFIANZA $(1 - \alpha)100\%$ (UCB)

$$(\text{Estimador puntual}) + z_{\alpha} \times (\text{Error estándar del estimador})$$

El valor z para un límite de confianza de una cola $(1 - \alpha)100\%$, z_{α} , localiza un área α en una sola cola de la distribución normal, como se muestra en la figura 5.11.

FIGURA 5.11

Valor z para un límite de confianza de una cola



EJEMPLO

5.12

Una corporación planea emitir algunos documentos a corto plazo y espera que el interés que tendrá para pagar no rebasará 11.5%. Para obtener alguna información acerca de este problema la corporación vendió 40 documentos, uno a través de cada una de las 40 empresas de corretaje de acciones. La media y la desviación estándar para las 40 tasas de interés fueron 10.3% y .31%, respectivamente. Como la corporación está interesada en sólo un límite superior en las tasas de interés, encuentre un límite superior de confianza de 95% para la tasa media de interés que la corporación tendrá que pagar por los documentos.

Solución Como el parámetro de interés es μ , el estimador puntual es $\bar{x}$ con error estándar $SE \approx \frac{s}{\sqrt{n}}$. El coeficiente de confianza es .95, de modo que $\alpha = .05$ y $z_{.05} = 1.645$. Por lo tanto, el límite superior de confianza de 95% es

$$UCB = \bar{x} + 1.645 \left(\frac{s}{\sqrt{n}} \right) = 10.3 + 1.645 \left(\frac{.31}{\sqrt{40}} \right) = 10.3 + .0806 = 10.3806$$

Entonces, se puede estimar que la tasa media de interés que la corporación tendrá que pagar sobre sus documentos será menor que 10.3806%. La corporación no debe preocuparse porque sus tasas de interés rebasen el 11.5%. ¿Qué tan confiado está usted de esta conclusión? Bastante confiado, porque los intervalos construidos en esta forma contienen a μ , 95% de las veces.

5.9

SELECCIÓN DEL TAMAÑO MUESTRAL

Diseñar un experimento es en esencia un plan para comprar cierta cantidad de información. Así como el precio que se paga por un juego de video varía dependiendo de dónde y cuándo se compra, el precio de información estadística varía dependiendo de cómo y dónde se recolecta la información. Al igual que cuando se compra cualquier producto, se debe comprar tanta información estadística como sea posible por el mínimo costo posible.

La cantidad total de información relevante en una muestra es controlada por dos factores:

- El **plan muestral** o **diseño experimental**: el procedimiento para recolectar la información
- El **tamaño muestral n** : la cantidad de información recolectada

Se puede incrementar la cantidad de información recolectada al *aumentar* el tamaño muestral o quizás al *cambiar* el tipo de plan muestral o diseño experimental que se utilice. Trataremos el plan muestral más sencillo, es decir, el muestreo aleatorio de una población relativamente grande y nos concentraremos en las formas para elegir el tamaño muestral n necesario para comprar una cantidad determinada de información.

Un investigador hace poco progreso al planear un experimento antes de encontrar el problema del tamaño muestral. **¿Cuántas mediciones deben incluirse en la muestra?** ¿Cuánta información desea comprar el investigador? Con el fin de responder estas preguntas, el investigador debe especificar primero:

- La confiabilidad que desea lograr, y
- La precisión necesaria para su estimación

En un problema de estimación estadística, la precisión de la estimación se mide por el *margen de error* o el *ancho del intervalo de confianza*, los cuales tienen una confiabilidad especificada. En vista que ambas medidas son una función del tamaño de la muestra, especificar la confiabilidad y la precisión permite determinar el tamaño necesario de la muestra.

Por ejemplo, suponga que se desea estimar el promedio de producción diaria μ de un proceso químico y se necesita que el margen de error sea menor que 4 toneladas. Esto significa que, aproximadamente 95% de las veces en muestreo repetido, la “confiabilidad”, la distancia entre la media muestral $\bar{x}$ y la media poblacional μ será menor que 1.96 SE. Se desea que esta cantidad sea menor que 4 (la “precisión”). Esto es,

$$1.96 SE < 4 \quad \text{o} \quad 1.96 \left(\frac{\sigma}{\sqrt{n}} \right) < 4$$

Despejando n , resulta

$$n > \left(\frac{1.96}{4}\right)^2 \sigma^2 \quad \text{o} \quad n > .24\sigma^2$$

Si se conoce σ , la desviación estándar poblacional, puede sustituirse su valor en la fórmula y despejar n . Si σ es desconocida, que suele ser el caso, puede usarse la mejor aproximación disponible:

- Una estimación s se obtiene de una muestra previa
- Una estimación del rango basada en el conocimiento de las mediciones máximas y mínimas posibles: $\sigma \approx \text{Rango}/4$

Para este ejemplo, suponga que un estudio previo del proceso químico produjo una desviación muestral estándar de $s = 21$ toneladas. Entonces

$$n > .24\sigma^2 = .24(21)^2 = 105.8$$

Usando una muestra de tamaño $n = 106$ o mayor, se puede estar razonablemente seguro (con probabilidad aproximadamente igual a .95) que su estimación del promedio de producción estará a no más de ± 4 toneladas del promedio real de producción.

La solución $n = 106$ es sólo aproximada porque tenía que usarse un valor aproximado de σ para calcular el error estándar de la media. Aun cuando esto puede ser molesto, es el mejor método existente para seleccionar el tamaño muestral ¡y es ciertamente mejor que adivinar!

A veces los investigadores requieren una confiabilidad diferente, o nivel de confianza distinto a 95% de confianza especificado por el margen de error. En este caso, la mitad del ancho del intervalo de confianza da la medida de precisión para su estimación; esto es, el límite B en el error de su estimación es

$$z_{\alpha/2} \left(\frac{\sigma}{\sqrt{n}} \right) < B$$

Este método de seleccionar el tamaño muestral se puede usar para los cuatro procedimientos de estimación presentados en este capítulo. El procedimiento general se describe a continuación.



NECESITO SABER...

Cómo seleccionar el tamaño muestral

Determine el parámetro a estimarse y el error estándar de su estimador puntual. A continuación proceda como sigue:

1. Elija B , el límite en el error de su estimación y un coeficiente de confianza $(1 - \alpha)$.
2. Para un problema de una muestra, de esta ecuación despeje el tamaño muestral n :

$$z_{\alpha/2} \times (\text{error estándar del estimador}) \leq B$$

donde $z_{\alpha/2}$ es el valor de z que tiene un área $\alpha/2$ a su derecha.

3. Para un problema de dos muestras, haga $n_1 = n_2 = n$ y resuelva la ecuación del paso 2.

[NOTA: Para casi todos los estimadores (todos presentados en este texto), el error estándar es una función del tamaño muestral n .]

EJEMPLO

5.13

Los productores de tubo de plástico de polivinilo desean tener un suministro de tubos suficiente para satisfacer las necesidades del mercado. Desean encuestar mayoristas que compren tubos de polivinilo para estimar la proporción que planea aumentar sus compras el año próximo. ¿Qué tamaño muestral se requiere si desean que su estimación se encuentre a no más de .04 de la proporción real con probabilidad igual a .90?

Solución Para este ejemplo particular, el límite B en el error de la estimación es .04. Como el coeficiente de confianza es $(1 - \alpha) = .90$, α debe ser igual a .10 y $\alpha/2$ es .05. El valor z correspondiente a un área igual a .05 en la cola superior de la distribución z es $z_{.05} = 1.645$. Entonces se requiere

$$1.645 \text{ SE} = 1.645 \sqrt{\frac{pq}{n}} \leq .04$$

Para despejar n de esta ecuación, debe sustituirse un valor aproximado de p en la ecuación. Si se desea estar seguro que la muestra es suficientemente grande, debe usarse $p = .5$ (sustituir $p = .5$ dará la máxima solución posible para n porque el máximo valor de pq se presenta cuando $p = q = .5$). Entonces

$$1.645 \sqrt{\frac{(.5)(.5)}{n}} \leq .04$$

o bien

$$\sqrt{n} \geq \frac{(1.645)(.5)}{.04} = 20.56$$

$$n \geq (20.56)^2 = 422.7$$

Por tanto, los productores deben incluir al menos 423 mayoristas en su encuesta si desean estimar la proporción p correcta a no más de .04.

EJEMPLO

5.14

Un director de personal desea comparar la eficacia de dos métodos de capacitar empleados industriales a realizar cierta operación de ensamble. Un número de empleados se divide en dos grupos iguales: el primero recibiendo capacitación del método 1 y, el segundo, el método de capacitación 2. Cada uno efectuará la operación de ensamble y se registrará el tiempo de ensamblado. Se espera que las mediciones para ambos grupos tendrán una variación de alrededor de 8 minutos. Para que la estimación de la diferencia en tiempos medios de ensamble sea correcta a no más de 1 minuto de variación, con una probabilidad igual a .95, ¿cuántos trabajadores deben estar incluidos en cada grupo de capacitación?

Solución En vista que se está estimando la diferencia entre dos medias, el error estándar

de la estimación es $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$. El límite es $B = 1$ minuto, de modo que se requiere

$$1.96 \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq 1$$

Como se desea usar dos grupos iguales, se puede hacer $n_1 = n_2 = n$ y obtener la ecuación

$$1.96 \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}} \leq 1$$

Como ya observamos, la variabilidad (intervalo) de cada método de ensamble es aproximadamente igual y, por lo tanto, $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Como el intervalo, igual a 8 minutos, es aproximadamente igual a 4σ , tenemos

$$4\sigma \approx 8 \quad \text{o bien} \quad \sigma \approx 2$$

Sustituyendo este valor por σ_1 y σ_2 en la ecuación anterior, se obtiene

$$\begin{aligned} 1.96 \sqrt{\frac{(2)^2}{n} + \frac{(2)^2}{n}} &\leq 1 \\ 1.96 \sqrt{\frac{8}{n}} &\leq 1 \\ \sqrt{n} &\geq 1.96\sqrt{8} \end{aligned}$$

Resolviendo, tenemos $n \geq 31$. Por tanto, cada grupo debe contener al menos $n = 31$ trabajadores.

La tabla 5.7 contiene un resumen de las fórmulas empleadas para determinar los tamaños muestrales requeridos para estimación, con un límite dado en el error de la estimación o ancho de intervalo de confianza W ($W = 2B$). Observe que para estimar p , la fórmula del tamaño muestral usa $\sigma^2 = pq$, en tanto que para estimar $(p_1 - p_2)$, la fórmula del tamaño muestral usa $\sigma_1^2 = p_1q_1$ y $\sigma_2^2 = p_2q_2$.

TABLA 5.7 ● **Fórmulas para tamaño muestral**

Parámetro	Estimador	Tamaño muestral	Suposiciones
μ	$\bar{x}$	$n \geq \frac{z_{\alpha/2}^2 \sigma^2}{B^2}$	
$\mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2$	$n \geq \frac{z_{\alpha/2}^2 (\sigma_1^2 + \sigma_2^2)}{B^2}$	$n_1 = n_2 = n$
p	$\hat{p}$	$\begin{cases} n \geq \frac{z_{\alpha/2}^2 pq}{B^2} \\ 0 \\ n \geq \frac{(.25)z_{\alpha/2}^2}{B^2} \end{cases}$	$p = .5$
$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$	$\begin{cases} n \geq \frac{z_{\alpha/2}^2 (p_1q_1 + p_2q_2)}{B^2} \\ 0 \\ n \geq \frac{2(.25)z_{\alpha/2}^2}{B^2} \end{cases}$	$n_1 = n_2 = n$ $p_1 = p_2 = .5$

5.9

EJERCICIOS

TÉCNICAS BÁSICAS

5.52 Encuentre un límite superior de confianza de una cola al 90% para la media poblacional μ para estos valores:

a. $n = 40$, $s^2 = 65$, $\bar{x} = 75$

b. $n = 100$, $s = 2.3$, $\bar{x} = 1.6$

5.53 Encuentre un límite inferior de confianza al 99% para la proporción binomial p cuando una muestra aleatoria de $n = 400$ intentos produjo $x = 196$ éxitos.

5.54 Muestras aleatorias independientes de tamaño 50 son sacadas de dos poblaciones cuantitativas, produciendo la información muestral de la tabla. Encuentre un límite superior de confianza de 95% para la diferencia en las dos medias poblacionales.

	Muestra 1	Muestra 2
Tamaño muestral	50	50
Media muestral	12	10
Desviación estándar muestral	5	7

5.55 Supongamos que usted desea estimar una media poblacional basada en una muestra aleatoria de n observaciones y experiencias anteriores sugieren que $\sigma = 12.7$. Si desea estimar μ correcta a no más de 1.6 de variación, con probabilidad igual a .95, ¿cuántas observaciones deben estar incluidas en su muestra?

5.56 Supongamos que usted desea estimar un parámetro binomial p correcto a no más de .04 de variación, con probabilidad igual a .95. Si sospecha que p es igual a algún valor entre .1 y .3 y desea estar seguro que su muestra es suficientemente grande, ¿qué tan grande debe ser n ? (SUGERENCIA: Cuando calcule el error estándar, use el valor de p en el intervalo $.1 < p < .3$ que dará el tamaño muestral más grande.)

5.57 Muestras aleatorias independientes de $n_1 = n_2 = n$ observaciones son seleccionadas de cada una de las poblaciones 1 y 2. Si desea estimar la diferencia entre las dos medias poblacionales correctas a no más de .17 de variación, con probabilidad igual a .90, ¿qué tan grandes deben ser n_1 y n_2 ? Suponga que se conocen $\sigma_1^2 \approx \sigma_2^2 \approx 27.8$.

5.58 Muestras aleatorias independientes de $n_1 = n_2 = n$ observaciones son seleccionadas de cada una de las poblaciones binomiales 1 y 2. Si desea estimar la diferencia entre las dos proporciones poblacionales correctas a no más de .05 de variación, con probabilidad igual a .98, ¿qué tan grande debe ser n ? Suponga que no hay información anterior sobre los valores de p_1 y p_2 , pero usted desea asegurarse de tener un número adecuado de observaciones en las muestras.

APLICACIONES

5.59 Inmigración ilegal En el ejercicio 5.18 se presentó una encuesta de investigación llevada a cabo para *ABC News* y el *Washington Post* que incluyó preguntas sobre inmigración ilegal en Estados Unidos, y las respuestas federales y estatales al problema.³ Suponga que usted diseña un sondeo de este tipo.

- Explique cómo seleccionaría su muestra. ¿Qué problemas podría encontrar en este proceso?
- Si usted desea estimar el porcentaje de la población que está de acuerdo con una frase particular en su cuestionario de encuesta, correcto a no más de 1% con probabilidad .95, ¿aproximadamente cuántas personas tendrían que ser encuestadas?

5.60 ¡Menos carne roja! Muchas personas están ahora más conscientes de la importancia de una buena nutrición y algunos investigadores creen que podemos estar comiendo menos carne roja. Para probar esta teoría, una investigadora decide seleccionar registros de nutrición en hospitales, para personas encuestadas hace 10 años y comparar el promedio de cantidad de carne consumida por año contra las cantidades consumidas por un número igual de personas a quienes ella entrevistará este año. Ella sabe que la cantidad de carne consumida anualmente por los estadounidenses varía de 0 a alrededor de 104 libras. ¿Cuántas personas debe seleccionar la investigadora de cada grupo si desea estimar la diferencia en el promedio anual de consumo de carne per cápita, correcto a no más de 5 libras con 99% de confianza?

5.61 Carne roja, continúa Consulte el ejercicio 5.60. La investigadora selecciona dos grupos de 400 personas cada uno y recolecta la siguiente información muestral en el consumo anual de carne de res ahora y hace 10 años:

	Hace 10 años	Este año
Media muestral	73	63
Desviación muestral estándar	25	28

- A la investigadora le gustaría demostrar que el consumo per cápita de carne ha disminuido en los últimos 10 años, de modo que necesita demostrar que la diferencia en los promedios es mayor que 0. Encuentre un límite inferior de confianza de 99% para la diferencia en el promedio de consumos per cápita de carne de res para los dos grupos.
- ¿Qué conclusiones puede sacar la investigadora usando el límite de confianza de la parte a?

5.62 Lluvia contaminada Supongamos que usted desea estimar el pH medio de lluvia en una zona que sufre de fuerte contaminación debida a la descarga de humo de una

planta generadora de electricidad. Se sabe que σ está en la cercanía de .5 pH y que desea estimar que se encuentre dentro de .1 de μ , con una probabilidad cercana a .95.

¿Aproximadamente cuántas precipitaciones de lluvia deben incluirse en su muestra (una lectura de pH por lluvia)? ¿Sería válido seleccionar todos sus especímenes de agua de una sola lluvia? Explique.

5.63 pH en lluvia Consulte el ejercicio 5.62. Suponga que desea estimar la diferencia entre la acidez media para lluvias en dos lugares diferentes, uno en una zona relativamente no contaminada y la otra en una zona sujeta a fuerte contaminación del aire. Si usted desea que su estimación sea correcta al .1 de pH más cercano, con probabilidad cercana a .90, ¿aproximadamente cuántas

lluvias (valores de pH) tendrían que incluirse en cada muestra? (Suponga que la varianza de las mediciones de pH es aproximadamente .25 en ambos lugares y que las muestras serán de igual tamaño.)

5.64 Selenio, otra vez Refiérase a la comparación de la ingesta diaria de selenio, que toman adultos en dos regiones diferentes de Estados Unidos, que vimos en el ejercicio 5.36. Supongamos que desea estimar la diferencia en las ingestas diarias medias entre las dos regiones, correcta a no más de 5 microgramos de variación, con probabilidad igual a .90. Si planea seleccionar un número igual de adultos de las dos regiones (es decir, $n_1 = n_2$), ¿qué tan grandes deben ser n_1 y n_2 ?

REPASO DEL CAPÍTULO

Conceptos y fórmulas clave

I. Tipos de estimadores

1. Estimador puntual: un solo número se calcula para estimar el parámetro poblacional
2. Estimador de intervalo: dos números se calculan para formar un intervalo que, con cierta cantidad de confianza, contiene al parámetro

II. Propiedades de buenos estimadores

1. Insesgado: el valor promedio del estimador es igual al parámetro a ser estimado
2. Varianza mínima: de todos los estimadores insesgados, el mejor estimador tiene una distribución muestral con el error estándar más pequeño
3. El margen de error mide la distancia máxima entre el estimador y el valor verdadero del parámetro

III. Estimadores puntuales de muestra grande

Para estimar uno de cuatro parámetros poblacionales cuando los tamaños muestrales sean grandes, use los siguientes estimadores puntuales con los márgenes de error apropiados

Parámetro	Estimador puntual	95% de margen de error
μ	$\bar{x}$	$\pm 1.96 \left(\frac{s}{\sqrt{n}} \right)$
ρ	$\hat{\rho} = \frac{x}{n}$	$\pm 1.96 \sqrt{\frac{\hat{\rho}\hat{q}}{n}}$
$\mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2$	$\pm 1.96 \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
$\rho_1 - \rho_2$	$(\hat{\rho}_1 - \hat{\rho}_2) = \left(\frac{x_1}{n_1} - \frac{x_2}{n_2} \right)$	$\pm 1.96 \sqrt{\frac{\hat{\rho}_1\hat{q}_1}{n_1} + \frac{\hat{\rho}_2\hat{q}_2}{n_2}}$

IV. Estimadores de intervalo de muestra grande

Para estimar uno de cuatro parámetros poblacionales cuando los tamaños muestrales son grandes, use los siguientes estimadores de intervalo

Parámetro 100% Intervalo de confianza $(1 - \alpha)$

$$\mu \quad \bar{x} \pm z_{\alpha/2} \left(\frac{s}{\sqrt{n}} \right)$$

$$\rho \quad \hat{\rho} \pm z_{\alpha/2} \sqrt{\frac{\hat{\rho}\hat{q}}{n}}$$

$$\mu_1 - \mu_2 \quad (\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\rho_1 - \rho_2 \quad (\hat{\rho}_1 - \hat{\rho}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{\rho}_1\hat{q}_1}{n_1} + \frac{\hat{\rho}_2\hat{q}_2}{n_2}}$$

1. Todos los valores del intervalo son posibles valores para el parámetro poblacional desconocido
2. Es improbable que algunos valores fuera del intervalo sean el valor del parámetro desconocido
3. Para comparar dos medias o proporciones poblacionales, busque el valor 0 en el intervalo de confianza. Si 0 está en el intervalo, es posible que las dos medias o proporciones poblacionales sean iguales y no debería declararse una diferencia. Si 0 no está en el intervalo, es improbable que las dos medias o proporciones sean iguales y se puede declarar con seguridad una diferencia

V. Límites de confianza de una cola

Utilice ya sea el límite superior (+) o el inferior (−) de dos colas, con el valor crítico de z cambiado de $z_{\alpha/2}$ a z_{α}

VI. Selección del tamaño muestral

1. Determine el tamaño del margen de error, B , que esté usted dispuesto a tolerar
2. Seleccione el tamaño muestral al despejar n o $n = n_1 = n_2$ en la desigualdad: $z_{\alpha/2} \times SE \leq B$, donde SE (error estándar) es una función del tamaño muestral n
3. Para poblaciones cuantitativas, estime la desviación poblacional estándar usando un valor previamente calculado de s o la aproximación de rango $\sigma \approx \text{Rango}/4$
4. Para poblaciones binomiales, use el método conservativo y aproxime p usando el valor $p = .5$

Ejercicios suplementarios

5.65 Expresé el teorema central del límite. ¿De qué valor es el teorema central del límite en estimación estadística de muestra grande?

5.66 Una muestra aleatoria de $n = 64$ observaciones tiene una media de $\bar{x}$ y una desviación estándar $s = 3.9$.

- a. Dé la estimación puntual de la media poblacional μ y encuentre el margen de error para su estimación.
- b. Encuentre un intervalo de confianza de 90% para μ . ¿Qué significa “90% confiado”?
- c. Encuentre un límite inferior de confianza de 90% para la media poblacional μ . ¿Por qué este límite es diferente del límite inferior de confianza de la parte b?
- d. ¿Cuántas observaciones son necesarias para estimar μ a no más de .5, con probabilidad igual a .95?

5.67 Se seleccionaron muestras aleatorias independientes de $n_1 = 50$ y $n_2 = 60$ observaciones de las poblaciones 1 y 2, respectivamente. Los tamaños muestrales y estadísticas muestrales calculadas se dan en la tabla:

	Población	
	1	2
Tamaño muestral	50	60
Media muestral	100.4	96.2
Desviación muestral estándar	0.8	1.3

Encuentre un intervalo de confianza de 90% para la diferencia en medias poblacionales e interprete el intervalo.

5.68 Consulte el ejercicio 5.67. Supongamos que desea estimar $(\mu_1 - \mu_2)$ correcta a no más de .2, con probabilidad igual a .95. Si planea usar tamaños muestrales iguales, ¿qué tan grande deben ser n_1 y n_2 ?

5.69 Una muestra aleatoria de $n = 500$ observaciones de una población binomial produjo $x = 240$ éxitos.

- a. Encuentre una estimación puntual para p y el margen de error para su estimador.
- b. Encuentre un intervalo de confianza de 90% para p . Interprete este intervalo.

5.70 Consulte el ejercicio 5.69. ¿Qué tan grande debe ser una muestra si se desea estimar p correcta a no más de .025, con probabilidad igual a .90?

5.71 Se seleccionaron muestras aleatorias independientes de $n_1 = 40$ y $n_2 = 80$ observaciones de las poblaciones binomiales 1 y 2, respectivamente. El número de éxitos en las dos muestras fueron $x_1 = 17$ y $x_2 = 23$. Encuentre un intervalo de confianza de 99% para la diferencia entre las dos proporciones poblacionales binomiales. Interprete este intervalo.

5.72 Consulte el ejercicio 5.71. Supongamos que desea estimar $(p_1 - p_2)$ correcta a no más de .06, con probabilidad igual a .99 y planea usar tamaños muestrales iguales, es decir, $n_1 = n_2$. ¿Qué tan grandes deben ser n_1 y n_2 ?

5.73 Fumar y presión sanguínea Se realizó un experimento para estimar el efecto de fumar sobre la presión sanguínea de un grupo de 35 fumadores. Se obtuvo la diferencia para cada participante al tomar la diferencia en las lecturas de presión sanguínea al principio del experimento y otra vez cinco años más adelante. El aumento de la media muestral, medido en milímetros de mercurio, fue de $\bar{x} = 9.7$. La desviación estándar muestral fue $s = 5.8$. Estime el aumento medio en presión sanguínea que se esperaría para fumadores de cigarrillos en el espacio de tiempo indicado por el experimento. Encuentre el margen de error. Describa la población asociada con la media que haya estimado.

5.74 Presión sanguínea, continúa Con el uso de un coeficiente de confianza igual a .90, ponga un intervalo de confianza en el aumento medio de presión sanguínea para el ejercicio 5.73.

5.75 Concentración de yodo Con base en mediciones repetidas de la concentración de yodo en una solución, un químico informa la concentración como 4.614, con un “margen de error de .006”.

- a. ¿Cómo se interpretaría el “margen de error” del químico?

b. Si la concentración informada se basa en una muestra aleatoria de $n = 30$ mediciones, con una desviación muestral estándar $s = .017$, ¿estaría usted de acuerdo en que el “margen de error” del químico es .006?

5.76 Alimento para pollos Un experimentador alimentó con diferentes raciones, A y B, a dos grupos de 100 pollos cada uno. Suponga que todos los factores que no sean raciones son iguales para ambos grupos. De los pollos alimentados con la ración A, 13 murieron y de los pollos alimentados con la ración B, seis murieron.

- a. Construya un intervalo de confianza de 98% para la verdadera diferencia en porcentajes de mortalidad para las dos raciones.
- b. ¿Se puede concluir que hay una diferencia en los porcentajes de mortalidad para las dos raciones?

5.77 Antibióticos Se desea estimar la producción media por hora para un proceso que manufactura un antibiótico. Se observa el proceso durante 100 periodos de una hora elegidos al azar, con los resultados de $\bar{x} = 34$ onzas por hora y $s = 3$. Estime la producción media por hora para el proceso usando un intervalo de confianza de 95%.

5.78 Girasoles En un artículo del *Annals of Botany*, un investigador informó los diámetros basales de tallos de dos grupos de girasoles dicotiledóneos: los que se dejaron balancearse libremente al viento y a los que se les aplicó un soporte artificial.⁹ Un experimento similar se realizó con plantas de maíz monocotiledóneas. Aun cuando los autores midieron otras variables en un diseño experimental más complicado, suponga que cada grupo estuvo formado por 64 plantas (un total de 128 girasoles y 128 plantas de maíz). Los valores que se indican en la tabla siguiente son las medias muestrales más o menos el error estándar.

	Girasol	Maíz
Sin soporte	35.3 ± .72	16.2 ± .41
Con soporte	32.1 ± .72	14.6 ± .40

Use sus conocimientos de estimación estadística para comparar los diámetros basales de plantas sin y con soporte para las dos plantas. Escriba un párrafo que describa sus conclusiones, asegurándose de incluir una medida de la precisión de su inferencia.

5.79 Control de calidad Un ingeniero de control de calidad desea estimar la fracción de defectos en un lote grande de cartuchos de tinta para impresora. De experiencias previas, él sabe que la fracción real de defectos debe estar alrededor de .05. ¿Qué tan grande debe tomar una muestra si él desea estimar la verdadera fracción a no más de .01, usando un intervalo de confianza de 95%?

5.80 Tarjetas de circuitos Se seleccionaron muestras de 400 tarjetas de circuito impreso de cada una de dos líneas

de producción A y B. La línea A produjo 40 defectuosas y la B produjo 80 defectuosas. Estime la diferencia en las fracciones reales de defectuosas para las dos líneas con un coeficiente de confianza de .90.

5.81 Tarjetas de circuitos II Consulte el ejercicio 5.80. Suponga que se probaron 10 muestras de $n = 400$ tarjetas de circuito impreso y se construyó un intervalo de confianza para p para cada una de las 10 muestras. ¿Cuál es la probabilidad de que exactamente uno de los intervalos no contenga el verdadero valor de p ? ¿Y de que al menos un intervalo no contenga el verdadero valor de p ?

5.82 Hockey sobre hielo G. Wayne Marino investigó algunas de las variables relacionadas con la aceleración y rapidez de un jugador de hockey desde una posición en reposo.¹⁰ Sesenta y nueve jugadores, titulares e intramuros, de la Universidad de Illinois se incluyeron en el experimento. A cada jugador se le indicó que se moviera con la mayor rapidez posible desde una posición de reposo, para recorrer una distancia de 6 metros. Las medias y desviaciones estándar de algunas de las variables registradas para cada uno de los 69 patinadores se muestran en la tabla siguiente:

	Media	Desv. est.
Peso (kilogramos)	75.270	9.470
Largo de zancada (metros)	1.110	.205
Rapidez de zancada (zancada/s)	3.310	.390
Promedio de aceleración (m/s ²)	2.962	.529
Velocidad instantánea (m/s)	5.753	.892
Tiempo para patinar (s)	1.953	.131

- a. Dé la fórmula que usaría para construir un intervalo de confianza de 95% para una de las medias poblacionales (por ejemplo, tiempo medio para patinar la distancia de 6 m).
- b. Construya un intervalo de confianza de 95% para el tiempo medio para patinar. Interprete este intervalo.

5.83 Hockey sobre hielo, continúa El ejercicio 5.82 presentó estadísticas de un estudio de arranques rápidos de patinadores de hockey sobre hielo. La media y la desviación estándar de las 69 mediciones individuales de promedio de aceleración, en la distancia de 6 metros, fueron 2.962 y .529 metros por segundo, respectivamente.

- a. Encuentre un intervalo de confianza de 95% para esta media poblacional. Interprete el intervalo.
- b. Suponga que usted no está satisfecho con el ancho de este intervalo de confianza y desea cortar el intervalo a la mitad al aumentar el tamaño muestral. ¿Cuántos patinadores (en total) tendrían que incluirse en el estudio?

5.84 Hockey sobre hielo, continúa La media y la desviación estándar de las magnitudes de rapidez de la muestra de 69 patinadores al final de la distancia de 6 metros del ejercicio 5.82, fueron 5.753 y .892 metros por segundo, respectivamente.

- Encuentre un intervalo de confianza de 95% para la velocidad media en la marca de 6 metros. Interprete el intervalo.
- Supongamos que usted desea repetir el experimento y estimar esta velocidad media a no más de .1 segundos, con probabilidad .99. ¿Cuántos patinadores tendrían que incluirse en su muestra?

5.85 Densidad relativa Si 36 mediciones de la densidad relativa del aluminio tuvieron una media de 2.705 y una desviación estándar de .028, construya un intervalo de confianza del 98% para la densidad relativa real del aluminio.

5.86 Investigación de audiología Se pidió a 70 estudiantes hombres de primer año de universidad que participaran en un estudio para establecer el umbral absoluto de audibilidad. Cada individuo tomó asiento en un cuarto a prueba de sonidos y se le presentó un tono de 150 Hz a un gran número de niveles de estímulo en un orden aleatorio. A la persona se le indicó que presionara un botón si detectaba el sonido; el experimentador registró el nivel más bajo de estímulo al cual se detectó el tono. La

media para el grupo fue de 21.6 dB con $s = 2.1$. Estime el umbral medio absoluto para todos los estudiantes universitarios de primer año y calcule el margen de error.

5.87 Diestro o zurdo Un investigador clasificó a sus sujetos como innatamente diestros o zurdos al comparar la medida del ancho de las uñas de sus pulgares. Tomó una muestra de 400 hombres y encontró que 80 podrían clasificarse como zurdos de acuerdo con este criterio. Estime la proporción de todos los hombres de la población que resultarían zurdos, usando un intervalo de confianza de 95%.

5.88 La garrapata roja de cítricos Una entomóloga desea estimar el tiempo promedio de desarrollo de la garrapata roja de cítricos, a no más de .5 de día. De los experimentos previos, se sabe que σ es cercana a los cuatro días. ¿Qué tan grande debe ser la muestra que tome la entomóloga para tener 95% de confianza de su estimación?

5.89 La garrapata roja de cítricos, continúa Un productor piensa que uno de cada cinco de sus árboles de cítricos está infectado con la garrapata roja de cítricos mencionada en el ejercicio 5.88. ¿Qué tan grande debe ser la muestra si el productor desea estimar, a no más de .08, cuál es la proporción de árboles infectados con la garrapata roja de cítricos?

Inferencia a partir de muestras pequeñas

OBJETIVO GENERAL

En el capítulo 5 se presentaron los conceptos básicos de estimación estadística de muestra grande para situaciones prácticas que involucran medias y proporciones poblacionales. Como todas estas técnicas se apoyan en el teorema central del límite para justificar la normalidad de los estimadores y estadísticos de prueba, sólo se aplican cuando las muestras son grandes. Este capítulo complementa las técnicas de muestra grande al presentar las pruebas de muestra pequeña y los intervalos de confianza para medias y varianzas poblacionales. A diferencia de sus similares de muestras grandes, estas técnicas de muestra pequeña requieren que las poblaciones muestreadas sean normales o aproximadamente normales.

ÍNDICE DEL CAPÍTULO

- Introducción (6.1)
- Distribución t de Student (6.2)
- Inferencias de muestra pequeña respecto a una media poblacional (6.3)
- Inferencias de muestra pequeña para la diferencia entre dos medias poblacionales: muestras aleatorias independientes (6.4)
- Inferencias de muestra pequeña para la diferencia entre dos medias: una prueba de diferencia en pares (6.5)
- Inferencias respecto a la varianza poblacional (6.6)
- Comparación de dos varianzas poblacionales (6.7)
- Repaso de suposiciones de muestra pequeña (6.8)



NECESITO SABER...

Cómo decidir cuál prueba usar

6.1

INTRODUCCIÓN

Suponga que usted necesita correr un experimento para estimar una media poblacional o la diferencia entre dos medias. El proceso de recolectar los datos puede ser muy costoso o lento. Si no es posible recolectar una *muestra grande*, los procedimientos de estimación y prueba de los capítulos 4 y 5 no sirven.

Este capítulo presenta algunos procedimientos estadísticos equivalentes que pueden usarse cuando *el tamaño muestral es pequeño*. Los procedimientos de estimación y prueba comprenden estos parámetros ya conocidos:

- Una sola media poblacional, μ
- La diferencia entre dos medias poblacionales, $(\mu_1 - \mu_2)$
- Una sola varianza poblacional, σ^2
- La comparación de dos varianzas poblacionales, σ_1^2 y σ_2^2

Las pruebas e intervalos de confianza de muestra pequeña para proporciones binomiales se omitirán en nuestro análisis.

6.2

DISTRIBUCIÓN t DE STUDENT

Al efectuar un experimento destinado a evaluar un proceso nuevo pero muy costoso para producir diamantes sintéticos, usted solamente puede estudiar seis diamantes generados por el proceso. ¿Cómo usar estas seis mediciones para hacer inferencias acerca del promedio de peso μ de los diamantes a partir de este proceso?

Al estudiar la distribución muestral de $\bar{x}$ en el capítulo 4, planteamos estos puntos:

- Cuando la población original muestreada es normal, $\bar{x}$ y $z = (\bar{x} - \mu)/(\sigma/\sqrt{n})$ ambas tienen distribuciones normales, *para cualquier tamaño muestral*.
- Cuando la población original muestreada *no* es normal, $\bar{x}$, $z = (\bar{x} - \mu)/(\sigma/\sqrt{n})$, y $z \approx (\bar{x} - \mu)/(s/\sqrt{n})$ todas tienen distribuciones aproximadamente normales, si el tamaño muestral es *grande*.

TIP MI CONSEJO

Cuando $n < 30$, el teorema central del límite no garantiza que

$$\frac{\bar{x} - \mu}{s/\sqrt{n}}$$

sea aproximadamente normal

Desafortunadamente, cuando el tamaño muestral n es pequeño, el estadístico $(\bar{x} - \mu)/(s/\sqrt{n})$ *no tiene* una distribución normal. Por tanto, todos los valores críticos de z que utilizamos en el capítulo 5 ya no son correctos. Por ejemplo, *no se puede decir* que se encontrará a no más de 1.96 errores estándar de μ 95% de las veces.

Este problema no es nuevo; fue estudiado por expertos en estadística y experimentadores a principios del siglo xx. Para hallar la distribución muestral de este estadístico, hay dos formas de proceder:

- Usar un método empírico. Saque repetidas muestras y calcule $(\bar{x} - \mu)/(s/\sqrt{n})$ para cada muestra. La distribución de frecuencia relativa que usted construya usando estos valores aproximarán la forma y ubicación de la distribución muestral.
- Usar un método matemático para deducir la función real de densidad o curva que describa la distribución muestral.

Este segundo método fue utilizado por un inglés llamado W. S. Gosset en 1908. Él dedujo una complicada fórmula para la función de densidad de

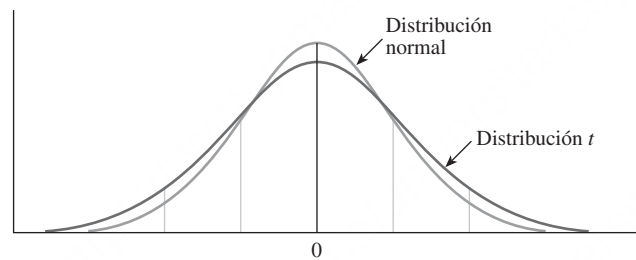
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

para muestras aleatorias de tamaño n desde una población normal y publicó sus resultados con el nombre de “Student”. Desde entonces, la estadística se conoce como ***t* de Student** y tiene las siguientes características:

- Tiene forma de montículo y es simétrica alrededor de $t = 0$, igual que z .
- Es más variable que z , con “colas más pesadas”; esto es, la curva t no se aproxima al eje horizontal con la misma rapidez que z . Esto es porque el estadístico t abarca dos cantidades aleatorias, $\bar{x}$ y s , en tanto que el estadístico z tiene sólo la media muestral, $\bar{x}$. Se observa este fenómeno en la figura 6.1.
- La forma de la distribución t depende del tamaño muestral n . A medida que n aumenta, la variabilidad de t disminuye porque la estimación s de σ está basada en más y más información. En última instancia, cuando n sea infinitamente grande, ¡las distribuciones t y z son idénticas!

FIGURA 6.1

z normal estándar y la distribución t con 5 grados de libertad



TIP MI CONSEJO

Para una t de una muestra, $gl = n - 1$

El divisor $(n - 1)$ en la fórmula para la varianza muestral s^2 se denomina **número de grados de libertad (*gl*) asociados con s^2** . Determina la *forma* de la distribución t . El origen del término *grados de libertad* es teórico y se refiere al número de desviaciones independientes elevadas al cuadrado en s^2 existentes para estimar σ^2 . Estos grados de libertad pueden cambiar para diferentes aplicaciones, y como especifican la distribución t correcta que se debe usar es necesario recordar que hay que calcular los grados de libertad correctos para cada aplicación.

La tabla de probabilidades para la distribución z normal estándar ya no es útil para calcular valores críticos o valores p para el estadístico t . En lugar de ello, se usará la tabla 4 del Anexo que se reproduce parcialmente en la tabla 6.1. Al indizar un número particular de grados de libertad, la tabla registra t_α , un valor de t que tiene área α de cola a su derecha, como se muestra en la figura 6.2.

FIGURA 6.2

Valores tabulados de la t de Student

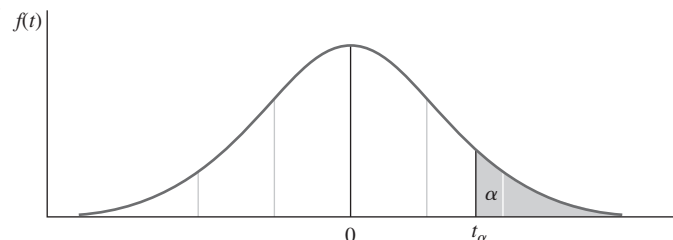


TABLA 6.1

Formato de la tabla t de Student tomado de la tabla 4 del Anexo

APPLET EN LÍNEA

Probabilidades t de
Student comparando t y z

gl	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	gl
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
·	·	·	·	·	·	·
·	·	·	·	·	·	·
·	·	·	·	·	·	·
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
inf.	1.282	1.645	1.960	2.326	2.576	inf.

EJEMPLO

6.1

Para una distribución t con 5 grados de libertad, el valor de t que tiene área .05 a su derecha se encuentra en la fila 5 en la columna marcada $t_{.050}$. Para esta distribución t particular, el área a la derecha de $t = 2.015$ es .05; sólo 5% de todos los valores del estadístico t rebasarán este valor.

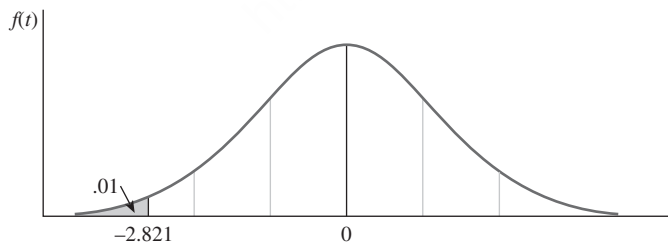
EJEMPLO

6.2

Suponga que usted tiene una muestra de tamaño $n = 10$ de una distribución normal. Encuentre un valor de t tal que sólo 1% de todos los valores de t sea más pequeño.

Solución Los grados de libertad que especifican la distribución t correcta son $gl = n - 1 = 9$, y el valor t necesario debe estar en la parte inferior de la distribución, con área .01 a su izquierda, como se ve en la figura 6.3. Como la distribución t es simétrica alrededor de 0, este valor es simplemente el negativo del valor en el lado derecho con área .01 a su derecha, o $-t_{.01} = -2.821$.

FIGURA 6.3

Distribución t para el
ejemplo 6.2

Podría preguntarse por qué los grados de libertad (gl) en la tabla 6.1 saltan de $gl = 29$ a $gl = \text{inf.}$ (infinito). Los valores críticos de t para varios grados de libertad entre 29 y 300 se dan en la figura 6.4. Notará que el valor de t para la misma área de la cola derecha disminuye conforme aumentan los grados de libertad. Cuando los grados de libertad se vuelven infinita-

mente grandes (*inf.*), el valor de *t* es igual al valor de *z*, el cual se proporciona en la última fila de la figura 6.4.

FIGURA 6.4

Valores críticos de la *t* de Student para grados de libertad entre *gl* = 29 y *gl* = infinito

<i>gl</i>	Área de la cola derecha		
	0.05	0.025	0.01
29	1.699	2.045	2.462
49	1.677	2.010	2.405
69	1.667	1.995	2.382
100	1.660	1.984	2.364
200	1.653	1.972	2.345
300	1.650	1.968	2.339
inf.	1.645	1.96	2.326

Al mismo tiempo, conforme aumentan los grados de libertad, la forma de la distribución *t* se vuelve menos variable hasta que al final se ve como (y es) la distribución normal estándar. Note que cuando los grados de libertad con *t* son *gl* = 300, casi no hay diferencia. Cuando *gl* = 29 y *n* = 30 los valores críticos de *t* son bastante cercanos a sus contrapartes normales; esto puede explicar por qué se usa la línea divisoria arbitraria entre una muestra grande y pequeña como *n* = 30. En lugar de producir una tabla *t* con muchos más valores críticos, los valores críticos de *z* son suficientes cuando *n* llega a 30.

Suposiciones detrás de la distribución *t* de Student

Los valores críticos de *t* le permiten hacer inferencias confiables *sólo si* usted sigue todas las reglas; esto es, su muestra debe cumplir estos requisitos especificados por la distribución *t*:

TIP MI CONSEJO

Suposiciones para *t* de una muestra:

- Muestra aleatoria
- Distribución normal

- La muestra debe ser seleccionada al azar.
- La población de la que haga muestreo debe estar normalmente distribuida.

Estos requisitos parecen bastante restrictivos. ¿Cómo podría conocer la forma de la distribución de probabilidad para toda la población si nada más tiene una muestra? Pero si éste fuera un problema serio el estadístico *t* podría usarse sólo en situaciones muy limitadas. Por fortuna, la forma de la distribución *t* no es afectada en mucho mientras la población muestreada tenga una distribución *aproximadamente en forma de montículo*. Los expertos dicen que el estadístico *t* es **robusto**, lo cual significa que su distribución no cambia de manera significativa cuando se viola la suposición de normalidad.

¿Cómo se puede saber si la muestra es de una población normal? Aun cuando hay procedimientos estadísticos diseñados para este fin, la forma más fácil y rápida de verificar la normalidad es usar las técnicas gráficas: trazar una gráfica de puntos o construir una gráfica de tallo y hoja. Mientras esta gráfica tienda a “hacerse montículo” en el centro, se puede estar razonablemente seguro al usar el estadístico *t* para hacer inferencias.

El requisito de muestreo aleatorio, por otra parte, es bastante crítico si se desea producir inferencias confiables. Si la muestra no es aleatoria, o si no se *comporta al menos como* muestra aleatoria, entonces los resultados de su muestra pueden ser afectados por algún factor desconocido y las conclusiones pueden ser incorrectas. Cuando diseñe un experimento o lea acerca de experimentos realizados por otros investigadores, ¡observe de manera crítica la forma en que los datos han sido recolectados!

6.3

INFERENCIAS DE MUESTRA PEQUEÑA RESPECTO A UNA MEDIA POBLACIONAL

Al igual que con una inferencia de muestra grande, una inferencia de muestra pequeña puede comprender ya sea **estimación** o **prueba de hipótesis**, dependiendo de la preferencia del experimentador. En capítulos previos explicamos los aspectos básicos de estos dos tipos de inferencia y los usamos de nuevo ahora con una estadística muestral diferente, $t = (\bar{x} - \mu)/(s/\sqrt{n})$, y una distribución de muestreo diferente, la t de Student, con $(n - 1)$ grados de libertad.

PRUEBA DE HIPÓTESIS DE MUESTRA PEQUEÑA PARA μ

1. Hipótesis nula: $H_0: \mu = \mu_0$
2. Hipótesis alternativa:

Prueba de una cola

$$H_a: \mu > \mu_0$$

(o, $H_a: \mu < \mu_0$)

Prueba de dos colas

$$H_a: \mu \neq \mu_0$$

3. Estadístico de prueba: $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$

4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola

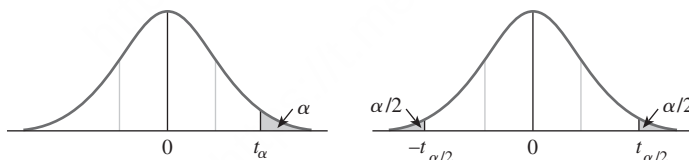
$$t > t_\alpha$$

(o $t < -t_\alpha$ cuando la hipótesis alternativa es $H_a: \mu < \mu_0$)

Prueba de dos colas

$$t > t_{\alpha/2} \quad \text{o} \quad t < -t_{\alpha/2}$$

o cuando el valor $p < \alpha$



Los valores críticos de t , t_α y $t_{\alpha/2}$ están basados en $(n - 1)$ grados de libertad. Estos valores tabulados se encuentran usando la tabla 4 del Anexo.

Suposición: La muestra se selecciona al azar de una población normalmente distribuida.

INTERVALO DE CONFIANZA DE MUESTRA PEQUEÑA ($1 - \alpha$)100% PARA μ

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

donde $s/\sqrt{n}$ es el error estándar estimado de $\bar{x}$, a veces conocido como **error estándar de la media**.

EJEMPLO

6.3

Un nuevo proceso para producir diamantes sintéticos puede ser operado a un nivel rentable sólo si el promedio del peso de éstos es mayor que .5 quilates. Para evaluar la rentabilidad del proceso se generan seis diamantes que registran pesos de .46, .61, .52, .48, .57 y .54 quilates. ¿Estas seis mediciones presentan suficiente evidencia para indicar que el promedio del peso de los diamantes producidos por el proceso es más de .5 quilates?

Solución La población de pesos de diamantes producidos por este nuevo proceso tiene media μ , y usted puede empezar la prueba formal de hipótesis en pasos:

1-2

Hipótesis nula y alternativa:

$$H_0: \mu = .5 \quad \text{contra} \quad H_a: \mu > .5$$

3

Estadístico de prueba: Usted puede usar su calculadora para verificar que la media y desviación estándar para los pesos de los seis diamantes son .53 y .0559, respectivamente. El estadístico de prueba es un estadístico t , calculado como

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{.53 - .5}{.0559/\sqrt{6}} = 1.32$$

Al igual que con las pruebas de muestra grande, el estadístico de prueba proporciona evidencia para rechazar o aceptar H_0 dependiendo de qué tan lejos se encuentre del centro de la distribución t .

4

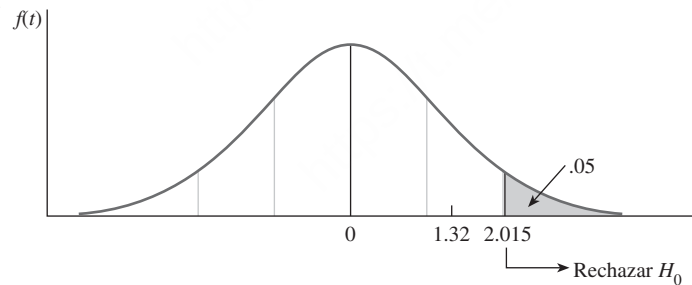
Región de rechazo: Si elige un nivel de significancia de 5% ($\alpha = .05$), la región de rechazo de la cola derecha se encuentra usando los valores críticos de t de la tabla 4 del Anexo. Con $gl = n - 1 = 5$, se puede rechazar H_0 si $t > t_{.05} = 2.015$, como lo muestra la figura 6.5.

5

Conclusión: Como el valor calculado del estadístico de prueba, 1.32, no cae en la región de rechazo, no se puede rechazar H_0 . Los datos no presentan suficiente evidencia para indicar que el peso medio de los diamantes exceda de .5 quilates.

FIGURA 6.5

Región de rechazo para el ejemplo 6.3

**TIP**

MI CONSEJO

Un intervalo de confianza de 95% nos dice que, si construyéramos muchos de estos intervalos (todos los cuales tendrían puntos extremos ligeramente diferentes), 95% de ellos se acercarían a la media poblacional

Igual que en el capítulo 5, la conclusión para *aceptar* H_0 requeriría el difícil cálculo de β , la probabilidad de un error tipo II. Para evitar este problema, elegimos *no rechazar* H_0 . Podemos entonces calcular el límite inferior para μ usando un límite inferior de confianza de un lado y muestra pequeña. Este límite es similar al límite de confianza de un lado y muestra grande, excepto que la z_α crítica es sustituida por una t_α crítica de la tabla 4. Para este ejemplo, un límite inferior de confianza de un lado para μ es:

$$\bar{x} - t_\alpha \frac{s}{\sqrt{n}}$$

$$.53 - 2.015 \frac{.0559}{\sqrt{6}}$$

$$.53 - .046$$

El límite inferior de 95% para μ es $\mu > .484$. El rango de posibles valores incluye la media de los pesos de los diamantes tanto menores como mayores que .5; esto confirma el fracaso de nuestra prueba para demostrar que μ excede de .5.

Hay dos formas de efectuar una prueba de hipótesis:

- **El método del valor crítico:** Proponga una región de rechazo basada en los valores críticos de la distribución muestral del estadístico. Si el estadístico de prueba cae en la región de rechazo, se puede rechazar H_0 .
- **El método del valor p :** Calcule el valor p con base en el valor observado del estadístico de prueba. Si el valor p es menor que el nivel de significancia, α , se puede rechazar H_0 .

Usamos el primer método en la solución del ejemplo 6.3 y el segundo método para resolver el ejemplo 6.4.

EJEMPLO

6.4

Las etiquetas en latas de 1 galón de pintura por lo general indican el tiempo de secado y el área que puede cubrir una capa. Casi todas las marcas de pintura indican que, en una capa, 1 galón cubrirá entre 250 y 500 pies cuadrados, dependiendo de la textura de la superficie a pintarse. Un fabricante, sin embargo, afirma que 1 galón de su pintura cubrirá 400 pies cuadrados de área superficial. Para probar su afirmación, una muestra aleatoria de 10 latas de 1 galón de pintura blanca se empleó para pintar 10 áreas idénticas usando la misma clase de equipo. Las áreas reales (en pies cuadrados) cubiertas por estos 10 galones de pintura se dan a continuación:

310	311	412	368	447
376	303	410	365	350

¿Los datos presentan suficiente evidencia para indicar que el promedio de la cobertura difiere de 400 pies cuadrados? Encuentre el valor p para la prueba y úselo para evaluar la significancia estadística de los resultados.

TIP MI CONSEJO

Recuerde del capítulo 1 cómo calcular $\bar{x}$ y s usando el método de entrada de datos en su calculadora

Solución Para probar lo dicho, las hipótesis que deben probarse son

$$H_0: \mu = 400 \quad \text{contra} \quad H_a: \mu \neq 400$$

La media muestral y desviación estándar para los datos registrados son

$$\bar{x} = 365.2 \quad s = 48.417$$

y el estadístico de prueba es

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{365.2 - 400}{48.417/\sqrt{10}} = -2.27$$

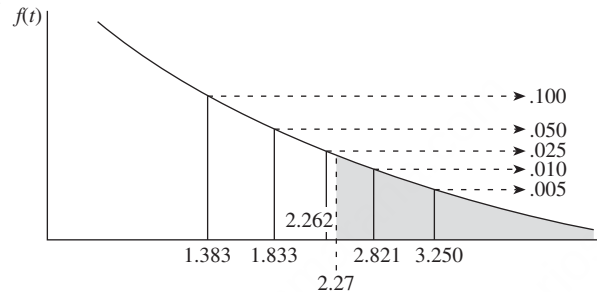
El valor p para esta prueba es la probabilidad de observar un valor de la estadística t como contradictorio a la hipótesis nula como el observado para este conjunto de datos, es decir, $t = -2.27$. Como ésta es una prueba de dos colas, el valor p es la probabilidad de que $t \leq -2.27$ o $t \geq 2.27$.

A diferencia de la tabla z , la tabla para t da los valores de t correspondientes a áreas de cola superior iguales a .100, .050, .025, .010 y .005. En consecuencia, sólo puede aproximarla al área de cola superior que corresponda a la probabilidad de que $t > 2.27$. Como el estadístico t para esta prueba está basado en 9 df , nos referimos a la fila correspondiente a $df = 9$ de la tabla 4. Los cinco valores críticos para varias áreas de cola se muestran en la figura 6.6, un agrandamiento de la cola de la distribución t con 9 grados de libertad. El valor $t = 2.27$ cae entre $t_{.025} = 2.262$ y $t_{.010} = 2.821$. Por tanto, el área de cola derecha correspondiente a la probabilidad de que $t > 2.27$ está entre .01 y .025. Como esta área representa sólo la mitad del valor p , se escribe

$$.01 < \frac{1}{2}(\text{valor } p) < .025 \quad \text{o} \quad .02 < \text{valor } p < .05$$

FIGURA 6.6

Cálculo del valor p para el ejemplo 6.4 (área sombreada $\frac{1}{2} = \text{valor } p$)



¿Qué nos dice esto acerca de la significancia de los resultados estadísticos? Para rechazar H_0 , el valor p debe ser menor que el nivel de significancia especificado, α . En consecuencia, podría rechazarse H_0 al nivel de 5% pero no al nivel de 2% o 1%. Por tanto, el valor p para esta prueba por lo general sería informado por el experimentador como

$$\text{valor } p < .05 \text{ (o a veces } P < .05)$$

Para esta prueba de hipótesis, H_0 es rechazada al nivel de significancia de 5%. Hay suficiente evidencia para indicar que el promedio de la cobertura difiere de 400 pies cuadrados.

¿Dentro de qué límites *realmente* cae este promedio de cobertura? Un intervalo de confianza de 95% da los límites superior e inferior para μ como

$$\bar{x} \pm t_{\alpha/2} \left(\frac{s}{\sqrt{n}} \right)$$

$$365.2 \pm 2.262 \left(\frac{48.417}{\sqrt{10}} \right)$$

$$365.2 \pm 34.63$$

Por tanto, usted puede estimar que el promedio de área cubierta por 1 galón de esta marca de pintura está en el intervalo 330.6 a 399.8. Una estimación más precisa del intervalo (un intervalo más corto) puede obtenerse por lo general al aumentar el tamaño muestral. Observe que el límite superior de este intervalo es muy cercano al valor de 400 pies cuadrados, que es la cobertura marcada en la leyenda de la lata. Esto coincide con el hecho de que el valor observado de $t = -2.27$ es sólo ligeramente menor que el valor crítico de cola izquierda de $t_{.025} = -2.262$, haciendo que el valor p sea sólo ligeramente menor que .05.



APLET EN LÍNEA

Prueba de muestra
pequeña de una media
poblacional

Muchos paquetes de estadística para computadoras contienen programas que ponen en práctica la prueba t de Student o construyen un intervalo de confianza para μ . Aunque *MS Excel* no tiene un comando único que implemente estos procedimientos, usted puede usar la herramienta de función en *Excel* para encontrar por sí mismo el estadístico de prueba, el valor p , y los límites de confianza superior e inferior. *MINITAB*, sin embargo, calcula y reporta todos estos valores con un conjunto de comandos, permitiéndole obtener con rapidez y precisión conclusiones sobre la significancia estadística de los resultados. Los resultados del *MINITAB* de prueba t de una muestra y de procedimientos de intervalo de confianza se dan en la figura 6.7. Además del valor observado de $t = -2.27$ y el intervalo de confianza (330.6, 399.8), la salida da la media muestral, la desviación muestral estándar, el error estándar de la media (media de error estándar = $s/\sqrt{n}$), y el valor p exacto de la prueba ($P = .049$). Esto es consistente con el rango para el valor p que encontramos usando la tabla 4 del Anexo:

$$.02 < \text{valor } p < .05$$

Encontrará instrucciones para generar esta salida *MINITAB* en la sección “Tecnología actual” al final de este capítulo.

FIGURA 6.7

Salida impresa *MINITAB* para el ejemplo 6.4

T de una muestra: área

Test of $\mu = 400$ vs not = 400

Variable	N	Mean	StDev	SE Mean
Area	10	365.2	48.4	15.3
Variable	95% CI		T	P
Area	(330.6, 399.8)		-2.27	0.049

Se puede ver el valor de usar la salida impresa de computadora para evaluar resultados estadísticos:

- El valor p exacto elimina la necesidad de tablas y valores críticos.
- Todos los cálculos numéricos son realizados por el usuario.

El trabajo más importante, que se deja al experimentador, es *interpretar* los resultados en términos de su ¡significancia práctica!

6.3

EJERCICIOS

TÉCNICAS BÁSICAS

6.1 Encuentre los siguientes valores t en la tabla 4 del Anexo:

- a. $t_{.05}$ para 5 gl. b. $t_{.025}$ para 8 gl.
c. $t_{.10}$ para 18 gl. d. $t_{.025}$ para 30 gl.

6.2 Encuentre el (los) valor(es) crítico(s) de t que especifiquen la región de rechazo en estas situaciones:

- a. Una prueba de dos colas con $\alpha = .01$ y 12 gl.
b. Una prueba de cola derecha con $\alpha = .05$ y 16 gl.
c. Una prueba de dos colas con $\alpha = .05$ y 25 gl.
d. Una prueba de cola izquierda con $\alpha = .01$ y 7 gl.

6.3 Use la tabla 4 del Anexo para aproximar el valor p para el estadístico t en cada situación:

- a. Una prueba de dos colas con $t = 2.43$ y 12 gl.
b. Una prueba de cola derecha con $t = 3.21$ y 16 gl.
c. Una prueba de dos colas con $t = -1.19$ y 25 gl.
d. Una prueba de cola izquierda con $t = -8.77$ y 7 gl.

DATOS
EX0604

6.4 Calificaciones de examen Las calificaciones en un examen de 100 puntos, registradas para 20 estudiantes:

71	93	91	86	75
73	86	82	76	57
84	89	67	62	72
77	68	65	75	84

- a. ¿Puede usted suponer razonablemente que estas calificaciones de examen han sido seleccionadas de una población normal? Use una gráfica de tallo y hoja para justificar su respuesta.
- b. Calcule la media y desviación estándar de las calificaciones.
- c. Si estos estudiantes pueden ser considerados como muestra aleatoria de la población de todos los estudiantes, encuentre un intervalo de confianza de 95% para el promedio de calificaciones de examen de la población.

6.5 Las siguientes $n = 10$ observaciones son una muestra de una población normal:

7.4 7.1 6.5 7.5 7.6 6.3 6.9 7.7 6.5 7.0

- a. Encuentre la media y desviación estándar de estos datos.
- b. Encuentre un límite superior de confianza de una cola, de 99%, para la media poblacional μ .
- c. Pruebe $H_0: \mu = 7.5$ contra $H_a: \mu < 7.5$. Use $\alpha = .01$.
- d. ¿Los resultados de la parte b apoyan sus conclusiones de la parte c?

APLICACIONES

DATOS

EX0606

6.6 Atunes ¿Hay diferencia en los precios del atún, dependiendo del método de empaque?

Consumer Reports da el promedio del precio estimado para una lata de 6 onzas o bolsa de 7.06 onzas de atún, con base en los precios pagados a nivel nacional en supermercados.¹ Estos precios están registrados para varias marcas diferentes de atún.

Atún claro en agua	Atún blanco en aceite	Atún blanco en agua	Atún claro en aceite
.99	.53	1.27	1.49
1.92	1.41	1.22	1.29
1.23	1.12	1.19	1.27
.85	.63	1.22	1.35
.65	.67	1.29	1.23
.69	.60	1.00	.67
.60	.66	1.27	
		1.28	

Fuente: Estudio práctico "Pricing of Tuna" Copyright 2001 por Consumers Union of U.S. Inc., Yonkers, NY 10703-1057, organización sin fines de lucro. Reimpreso con permiso de la edición de junio de 2001 de *Consumer Report*® sólo con fines educativos. No se permite uso comercial ni reproducción. www.ConsumerReports.org.

Suponga que las marcas de atún incluidas en este estudio representan una muestra aleatoria de todas las marcas de atún disponibles en Estados Unidos.

- a. Encuentre un intervalo de confianza de 95% para el promedio del precio para el atún claro en agua. Interprete este intervalo. Esto es, ¿a qué se refiere ese "95%"?

- b. Encuentre un intervalo de confianza de 95% para el precio promedio del atún blanco en aceite. ¿Cómo se compara el ancho de este intervalo con el ancho del intervalo de la parte a? ¿Puede explicar por qué?
- c. Encuentre intervalos de confianza de 95% para las otras dos muestras (atún blanco en agua y atún claro en aceite). Grafique las cuatro medias de tratamiento y sus errores estándar en una gráfica bidimensional semejante a la figura 5.5. ¿Qué clase de comparaciones generales pueden hacerse acerca de los cuatro tratamientos? (Analizaremos el procedimiento para comparar más de dos medias poblacionales en el capítulo 7.)

6.7 Contenido de O₂ disuelto Los desechos industriales y residuales descargados en nuestros ríos y arroyos absorben oxígeno y, por tanto, reducen la cantidad de oxígeno disuelto disponible para peces y otras formas de fauna acuática. Una agencia estatal requiere un mínimo de 5 partes por millón (ppm) de oxígeno disuelto para que el contenido de éste sea suficiente para sostener la vida acuática. Seis especímenes de agua tomados de un río en un lugar específico durante la estación de aguas bajas (julio) dio lecturas de 4.9, 5.1, 4.9, 5.0, 5.0 y 4.7 de oxígeno disuelto. ¿Los datos dan suficiente evidencia para indicar que el contenido de oxígeno disuelto es menor que 5 ppm? Pruebe usando $\alpha = .05$.

6.8 Langostas En un estudio de la infestación de la langosta *Thenus orientalis* por dos tipos de lapas, *Octolasmis tridens* y *O. lowei*, se midieron las longitudes del caparazón (en milímetros) de 10 langostas seleccionadas al azar pescadas en los mares cerca de Singapur:²

78 66 65 63 60 60 58 56 52 50

Encuentre un intervalo de confianza de 95% para la media de la longitud del caparazón de langostas *T. orientalis*.

DATOS

EX0609

6.9 Fumar y capacidad pulmonar En un estudio del efecto de fumar sobre la capacidad de difusión de monóxido de carbono (DL) del pulmón, unos investigadores encontraron que los fumadores actuales tenían lecturas de DL considerablemente más bajas que otros que habían sido fumadores o que son no fumadores. Las capacidades de difusión de monóxido de carbono para una muestra aleatoria de $n = 20$ fumadores actuales aparecen a continuación:

103.768 88.602 73.003 123.086 91.052
92.295 61.675 90.677 84.023 76.014
100.615 88.017 71.210 82.115 89.222
102.754 108.579 73.154 106.755 90.479

- a. ¿Estos datos indican que la lectura de la media de DL para fumadores actuales es considerablemente más baja que 100 DL, que es el promedio para no fumadores? Use $\alpha = .01$.

- b. Encuentre un límite superior de confianza de una cola de 99% para la lectura de la media de DL para fumadores actuales. ¿Este límite confirma las conclusiones de la parte a?

6.10 Purificación de un compuesto orgánico Es frecuente que químicos orgánicos purifiquen compuestos orgánicos por un método conocido como cristalización fraccional. Un químico preparó diez cantidades de 4.85 g de anilina y las purificó a acetanilida. Se registraron los siguientes rendimientos en seco:

3.85	3.80	3.88	3.85	3.90
3.36	3.62	4.01	3.72	3.83

Estime la cantidad media en gramos de acetanilido que pueda ser recuperada de una cantidad inicial de 4.85 g de anilina. Use un intervalo de confianza de 95%.

6.11 Compuestos orgánicos, continúa Consulte el ejercicio 6.10. ¿Aproximadamente cuántos especímenes de 4.85 g de anilina se requieren, si se desea estimar el número medio de gramos de acetanilido correcto a no más de .06 g con probabilidad igual a .95?

6.12 Bulimia En un estudio para determinar qué factores pronostican quién se beneficiará del tratamiento para la bulimia nerviosa, un artículo en el *British Journal of Clinical Psychology* indica que la autoestima es uno de estos importantes pronosticadores.³ La tabla siguiente da la media y la desviación estándar de calificaciones de autoestima previas al tratamiento, después del tratamiento y durante un seguimiento:

	Antes del trat.	Después del trat.	Seguimiento
Media muestral $\bar{x}$	20.3	26.6	27.7
Desviación estándar s	5.0	7.4	8.2
Tamaño muestral n	21	21	20

- Use una prueba de hipótesis para determinar si hay suficiente evidencia para concluir que la media verdadera previa al tratamiento es menor que 25.
- Construya un intervalo de confianza de 95% para la media verdadera después del tratamiento.
- En la sección 6.4, presentaremos técnicas de muestra pequeña para hacer inferencias acerca de la diferencia entre dos medias poblacionales. Sin la formalidad de una prueba estadística, ¿qué está usted dispuesto a concluir acerca de las diferencias entre las tres medias poblacionales muestreadas representadas por los resultados de la tabla?

DATOS EX0613 **6.13 Cantidades de RBC** A continuación presentamos las cantidades de células rojas sanguíneas (en 10^6 células por microlitro) de una persona sana, medidas en cada uno de 15 días:

5.4	5.2	5.0	5.2	5.5
5.3	5.4	5.2	5.1	5.3
5.3	4.9	5.4	5.2	5.2

Encuentre una estimación de μ del intervalo de confianza de 95%, la verdadera cantidad media de células rojas sanguíneas para esta persona durante el periodo de prueba.

DATOS EX0614 **6.14 Carne para hamburguesa** Estos datos son los pesos (en libras) de 27 paquetes de carne molida en un exhibidor de carnes de un supermercado:

1.08	.99	.97	1.18	1.41	1.28	.83
1.06	1.14	1.38	.75	.96	1.08	.87
.89	.89	.96	1.12	1.12	.93	1.24
.89	.98	1.14	.92	1.18	1.17	

- a. Interprete las salidas impresas siguientes del MINITAB para los procedimientos de prueba de una muestra y de estimación.

Salidas impresas del MINITAB para el ejercicio 6.14

T de una muestra: peso

Test of $\mu = 1$ vs not = 1

Variable	N	Mean	StDev	SE Mean
Weight	27	1.0522	0.1657	0.0319

Variable	95% CI	T	P
Weight	(0.9867, 1.1178)	1.64	0.113

- b. Verifique los valores calculados de t y los límites de confianza superiores e inferiores.

DATOS EX0615 **6.15 Colesterol** A continuación aparecen los niveles de colesterol seroso de 50 personas seleccionadas al azar de entre los datos del L.A. Heart Data, de un estudio epidemiológico de enfermedades del corazón en empleados del condado de Los Ángeles.⁴

148	304	300	240	368	139	203	249	265	229
303	315	174	209	253	169	170	254	212	255
262	284	275	229	261	239	254	222	273	299
278	227	220	260	221	247	178	204	250	256
305	225	306	184	242	282	311	271	276	248

- Construya un histograma para los datos. ¿Los datos tienen una forma aproximada de montículo?
- Use una distribución t para construir un intervalo de confianza de 95% para el promedio de los niveles de colesterol seroso para empleados del condado de Los Ángeles.

6.16 Colesterol, continúa Consulte el ejercicio 6.15. Como $n > 30$, use los métodos del capítulo 5 para crear un intervalo de confianza de 95% de muestra grande para el promedio del nivel de colesterol seroso para empleados del condado de Los Ángeles. Compare los dos intervalos. (SUGERENCIA: Los dos intervalos deben ser muy semejantes. Ésta es la razón por la que elegimos aproximar la

distribución muestral de $\frac{\bar{x} - \mu}{s/\sqrt{n}}$ con una distribución z cuando $n > 30$.)

INFERENCIAS DE MUESTRA PEQUEÑA PARA LA DIFERENCIA ENTRE DOS MEDIAS POBLACIONALES: MUESTRAS ALEATORIAS INDEPENDIENTES

6.4

El escenario físico para el problema considerado en esta sección es el mismo que el de la sección 5.6, excepto que los tamaños muestrales ya no son grandes. Se extraen muestras aleatorias independientes de n_1 y n_2 mediciones de dos poblaciones, con medias y varianzas μ_1 , σ_1^2 , μ_2 y σ_2^2 , y el objetivo del experimentador es hacer inferencias acerca de $(\mu_1 - \mu_2)$, la diferencia entre las dos medias poblacionales.

Cuando los tamaños muestrales son pequeños ya no se puede confiar en el teorema central del límite para asegurar que las medias muestrales sean normales, pero, si las poblaciones originales *son normales*, entonces la distribución muestral de la diferencia en las medias muestrales, $(\bar{x}_1 - \bar{x}_2)$, será normal (incluso para muestras pequeñas) con media $(\mu_1 - \mu_2)$ y error estándar

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

En los capítulos 4 y 5 usted usó las varianzas muestrales, s_1^2 y s_2^2 , para calcular una *estimación* del error estándar, que entonces se utilizó para formar un intervalo de confianza de muestra grande o una prueba de hipótesis basada en el estadístico z de muestras grandes:

$$z \approx \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Desafortunadamente, cuando los tamaños muestrales son pequeños, este estadístico no tiene una distribución aproximadamente normal ni tiene una distribución t de Student. Para formar una estadística con una distribución de muestreo que pueda deducirse en forma teórica es necesario hacer una suposición más.

Suponga que la variabilidad de la medición en las dos poblaciones normales es la misma y puede ser medida por una varianza común σ^2 . Esto es, *ambas poblaciones tienen exactamente la misma forma*, y $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Entonces el error estándar de la diferencia en las dos medias muestrales es

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

Se puede demostrar matemáticamente que, si se usa la estimación muestral apropiada s^2 para la varianza poblacional σ^2 , entonces el estadístico de prueba resultante,

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

tiene una *distribución t de Student*. El único problema restante es hallar la estimación muestral s^2 y el número apropiado de *grados de libertad* para el estadístico t .

Recuerde que la varianza poblacional σ^2 describe la forma de las distribuciones normales de donde provienen las muestras del experimentador, de modo que s_1^2 o s_2^2 le darían una estimación de σ^2 . Pero, ¿por qué usar sólo una cuando ambas dan información? Un mejor procedimiento es combinar la información en ambas varianzas muestrales usando un *promedio ponderado*, en el que los pesos están determinados por la cantidad relativa de

TIP MI CONSEJO

Suposiciones para la prueba t de dos muestras (independientes):

- Muestras aleatorias independientes
- Distribuciones normales
- $\sigma_1 = \sigma_2$

información (el número de mediciones) en cada muestra. Por ejemplo, si la primera muestra contenía el doble de mediciones que la segunda, se podría considerar dar a la primera varianza muestral el doble de peso. Para obtener este resultado, use esta fórmula:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Recuerde de la sección 6.3 que los grados de libertad para el estadístico t de una muestra son $(n - 1)$, el denominador de la estimación muestral s^2 . Como s_1^2 tiene $(n_1 - 1)$ gl y s_2^2 tiene $(n_2 - 1)gl$, el número total de grados de libertad es la suma

$$(n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$$

mostrada en el denominador de la fórmula para s^2 .

CÁLCULO DE s^2

- Si dispone de una calculadora científica, calcule cada una de las dos desviaciones estándar muestrales s_1 y s_2 por separado, usando el procedimiento de entrada de datos para su calculadora particular. Estos valores son elevados al cuadrado y se usan en esta fórmula:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

TIP MI CONSEJO

Para la prueba t de dos muestras (independientes)
 $gl = n_1 + n_2 - 2$

Se puede demostrar que s^2 es un estimador no sesgado de la varianza poblacional común σ^2 . Si s^2 se usa para estimar σ^2 y si las muestras se han sacado al azar e independientemente de poblaciones normales con una varianza común, entonces el estadístico

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

tiene una distribución t de Student con $(n_1 + n_2 - 2)$ grados de libertad. La estimación de muestra pequeña y procedimientos de prueba para la diferencia entre dos medias se dan a continuación.

PRUEBA DE HIPÓTESIS RESPECTO A LA DIFERENCIA ENTRE DOS MEDIAS: MUESTRAS ALEATORIAS INDEPENDIENTES

1. La hipótesis nula: $H_0: (\mu_1 - \mu_2) = D_0$, donde D_0 es alguna diferencia especificada que el experimentador desea probar. Para numerosas pruebas, el experimentador hará una hipótesis de que no hay diferencia entre μ_1 y μ_2 ; esto es, $D_0 = 0$.
2. Hipótesis alternativa:

Prueba de una cola

$$H_a: (\mu_1 - \mu_2) > D_0$$

$$[o H_a: (\mu_1 - \mu_2) < D_0]$$

Prueba de dos colas

$$H_a: (\mu_1 - \mu_2) \neq D_0$$

$$3. \text{ Estadístico de prueba: } t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad \text{donde}$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

4. Región de rechazo: rechace H_0 cuando

Prueba de una cola

$$t > t_\alpha$$

[o $t < -t_\alpha$ cuando la hipótesis alternativa es

$$H_a: (\mu_1 - \mu_2) < D_0]$$

o cuando el valor $p < \alpha$

Prueba de dos colas

$$t > t_{\alpha/2} \text{ o } t < -t_{\alpha/2}$$

Los valores críticos de t , t_α y $t_{\alpha/2}$ están basados en $(n_1 + n_2 - 2)$ gl. Los valores tabulados se pueden hallar usando la tabla 4 del Anexo.

Suposiciones: Las muestras se seleccionan al azar y en forma independiente de poblaciones distribuidas normalmente. Las varianzas de las poblaciones σ_1^2 y σ_2^2 son iguales.

INTERVALO DE CONFIANZA $(1 - \alpha)100\%$ DE MUESTRA PEQUEÑA PARA $(\mu_1 - \mu_2)$ CON BASE EN MUESTRAS ALEATORIAS INDEPENDIENTES

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

donde s^2 es la estimación agrupada de σ^2 .

EJEMPLO

6.5

Puede tomarse un curso con crédito ya sea asistiendo a clases en horas y días fijos, o en sesiones en línea que el estudiante puede tomar a su propio paso y en los tiempos que elija. El coordinador del curso desea determinar si estas dos formas de tomar el curso resultaron en una diferencia significativa en rendimiento medido por el examen final. La tabla 6.2 da las calificaciones en un examen con 45 puntos posibles para un grupo de $n_1 = 9$ estudiantes que tomaron el curso en línea y un segundo grupo de $n_2 = 9$ estudiantes que lo tomaron en clases convencionales. ¿Estos datos presentan suficiente evidencia para indicar que el promedio de las calificaciones de los estudiantes que tomaron el curso en línea es significativamente más alto que el de quienes asistieron a una clase convencional?

TABLA 6.2

Calificaciones de examen para presentaciones en línea y en salón de clase

En línea	Salón de clases
32	35
37	31
35	29
28	25
41	34
44	40
35	27
31	32
34	31

Solución Sean μ_1 y μ_2 la media de las calificaciones para el grupo en línea y el grupo del salón de clase, respectivamente. Por tanto, como se busca evidencia para apoyar la teoría de que $\mu_1 > \mu_2$, se puede probar la hipótesis nula

$$H_0 : \mu_1 = \mu_2 \quad [\text{o } H_0 : (\mu_1 - \mu_2) = 0]$$

contra la hipótesis alternativa

$$H_a : \mu_1 > \mu_2 \quad [\text{o } H_a : (\mu_1 - \mu_2) > 0]$$

Para efectuar la prueba t para estas dos muestras independientes se debe suponer que las poblaciones muestreadas son normales y tienen la misma varianza σ^2 . ¿Es esto razonable? Las gráficas de tallo y hoja de la figura 6.8 muestran al menos un patrón de “montículo”, de modo que la suposición de normalidad es razonable.

FIGURA 6.8

Gráficas de tallo y hoja para el ejemplo 6.5

En línea		Salón de clases	
2	8	2	579
3	124	3	1124
3	557	3	5
4	14	4	0

TIP MI CONSEJO

Las gráficas de tallo y hoja pueden ayudar a decidir si la suposición de normalidad es razonable

Además, las desviaciones estándar de las dos muestras, calculadas como

$$s_1 = 4.9441 \quad \text{y} \quad s_2 = 4.4752$$

no son tan diferentes para que dudemos de que las dos distribuciones tengan la misma forma. Si se hacen estas dos suposiciones y se calcula (usando precisión total) la estimación agrupada de la varianza común como

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{8(4.9441)^2 + 8(4.4752)^2}{9 + 9 - 2} = 22.2361$$

se puede calcular entonces el estadístico de prueba,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{35.22 - 31.56}{\sqrt{22.2361 \left(\frac{1}{9} + \frac{1}{9} \right)}} = 1.65$$

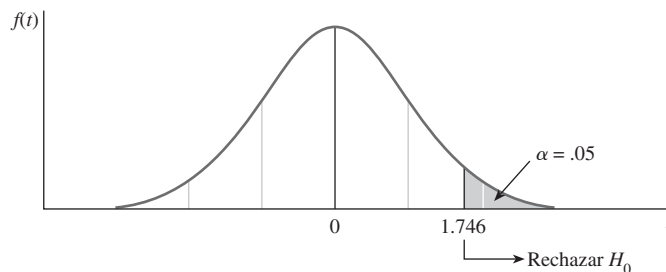
TIP MI CONSEJO

Si usa calculadora, ¡no redondee sino hasta el paso final!

La hipótesis alternativa $H_a : \mu_1 > \mu_2$ o, lo que es equivalente, $H_a : (\mu_1 - \mu_2) > 0$ implica que el experimentador debe usar una prueba de una cola en la cola superior de la distribución t con $(n_1 + n_2 - 2) = 16$ grados de libertad. Se puede hallar el valor crítico apropiado para una región de rechazo con $\alpha = .05$ en la tabla 4 del Anexo, y H_0 será rechazada si $t > 1.746$. Comparando el valor observado del estadístico de prueba $t = 1.65$ con el valor crítico $t_{.05} = 1.746$, no se puede rechazar la hipótesis nula (véase la figura 6.9). Hay insuficiente evidencia para indicar que el promedio de las calificaciones del curso en línea sea más alto que el promedio de las calificaciones del curso convencional al nivel de significancia de 5%.

FIGURA 6.9

Región de rechazo para el ejemplo 6.5



EJEMPLO

6.6

Encuentre el valor p que sería reportado para la prueba estadística del ejemplo 6.5.

Solución El valor observado de t para esta prueba de una cola es $t = 1.65$. Por tanto,

$$\text{valor } p = P(t > 1.65)$$



APLET EN LÍNEA

Prueba t de dos muestras:
muestras independientes

para un estadístico t con 16 grados de libertad. Recuerde que no se puede obtener esta probabilidad directamente de la tabla 4 del Anexo; sólo se puede *limitar* el valor p usando los valores críticos de la tabla. Como el valor observado, $t = 1.65$, está entre $t_{.100} = 1.337$ y $t_{.050} = 1.746$, el área de cola a la derecha de 1.65 está entre .05 y .10. El valor p para esta prueba se informaría como

$$.05 < \text{valor } p < .10$$

Como el valor p es mayor que .05, casi todos los investigadores informarían los resultados como *no significativos*.

EJEMPLO

6.7

Use un límite inferior de confianza de 95% para estimar la diferencia $(\mu_1 - \mu_2)$ en el ejemplo 6.5. ¿El límite de confianza inferior indica que el promedio en línea es significativamente más alto que el promedio de salón de clase?

Solución El límite inferior de confianza toma una forma familiar, es decir, el estimador puntual $(\bar{x}_1 - \bar{x}_2)$ menos una cantidad igual a t_α veces del error estándar del estimador. Sustituyendo en la fórmula, se puede calcular el límite inferior de confianza de 95%:

$$\begin{aligned} (\bar{x}_1 - \bar{x}_2) - t_\alpha \sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \\ (35.22 - 31.56) - 1.746 \sqrt{22.2361 \left(\frac{1}{9} + \frac{1}{9} \right)} \\ 3.66 - 3.88 \end{aligned}$$

o $(\mu_1 - \mu_2) > -22$. Como el valor $(\mu_1 - \mu_2) = 0$ está incluido en el intervalo de confianza, es posible que las dos medias sean iguales. Hay insuficiente evidencia para indicar que el promedio en línea sea más alto que el promedio en salón de clase.



MI CONSEJO

Mayor s^2 /menor $s^2 < 3$



suposición de varianza es
razonable

El procedimiento de dos muestras que usa una estimación agrupada de la varianza común σ^2 se apoya en cuatro importantes suposiciones:

- Las muestras deben ser *seleccionadas al azar*. Las muestras no seleccionadas al azar pueden introducir sesgo en el experimento y así alterar los niveles de significancia que el experimentador informe.
- Las muestras deben ser *independientes*. Si no es así, éste no es el procedimiento estadístico apropiado. En la sección 6.5 analizamos otro procedimiento para muestras dependientes.
- Las poblaciones de las cuales se muestrea deben ser *normales*. No obstante, las desviaciones moderadas desde la normalidad no afectan seriamente la distribución del estadístico de prueba, en especial si los tamaños muestrales son casi iguales.

- Las *varianzas poblacionales deben ser iguales* o casi iguales para asegurar que los procedimientos sean válidos.

Si las varianzas poblacionales están lejos de ser iguales, hay un procedimiento alternativo para estimación y prueba que tiene una distribución *t aproximada* en muestreo repetido. Como regla práctica, se debe usar este procedimiento si la razón entre las dos varianzas muestrales,

$$\frac{\text{Mayor } s^2}{\text{Menor } s^2} > 3$$

Como las varianzas poblacionales no son iguales, el estimador agrupado s^2 ya no es apropiado, y cada varianza poblacional debe ser estimada por su correspondiente varianza muestral. El estadístico de prueba resultante es

$$\frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Cuando los tamaños muestrales sean *pequeños*, los valores críticos para este estadístico se encuentran usando grados de libertad aproximados por la fórmula

$$gl \approx \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}}$$

Los grados de libertad se toman como la parte entera de este resultado.

Se pueden usar paquetes computarizados como el *MINITAB* y el *MS Excel* para poner en práctica este procedimiento, a veces llamado *aproximación de Satterthwaite*, así como el *método agrupado* descrito antes. De hecho, algunos experimentadores eligen analizar sus datos usando *ambos* métodos. Mientras los dos análisis lleven a las mismas conclusiones, no es necesario preocuparse con la igualdad o desigualdad de varianzas.

Las salidas *MINITAB* y *Excel*, resultantes del método agrupado de análisis para los datos del ejemplo 6.5, se muestran en las figuras 6.10a) y b). Observe que la razón entre las dos varianzas muestrales, $(4.94/4.48)^2 = 1.22$ es menor que 3, lo cual hace aproximado el método agrupado. El valor calculado de $t = 1.65$ y el valor p exacto = .059 con 16 grados de libertad se muestran en ambas salidas. El valor p exacto hace muy fácil determinar la significancia o no significancia de los resultados muestrales. Se encontrarán instrucciones para generar esta salida en la sección “Tecnología actual” al final de este capítulo.

FIGURA 6.10(a)

Salida de *MINITAB* para el ejemplo 6.5

● **Prueba t de dos muestras y CI: en línea, salón de clase**

```
Two-sample T for Online vs Classroom
      N      Mean    StDev   SE Mean
Online    9    35.22     4.94     1.6
Classroom  9    31.56     4.48     1.5

Difference = mu (Online) - mu (Classroom)
Estimate for difference: 3.67
95% lower bound for difference: -0.21
T-Test of difference = 0 (vs >): T-Value = 1.65  P-Value = 0.059  DF = 16
Both use Pooled StDev = 4.7155
```

FIGURA 6.10(b)

Salida de Excel para el ejemplo 6.5

D	E	F
t-Test: Two-Sample Assuming Equal Variances		
	Online	Classroom
Mean	35.222	31.556
Variance	24.444	20.028
Observations	9	9
Pooled Variance	22.236	
Hypothesized Mean Difference	0	
df	16	
t Stat	1.649	
P(T<=t) one-tail	0.059	
t Critical one-tail	1.746	
P(T<=t) two-tail	0.119	
t Critical two-tail	2.120	

Si hay razón para pensar que las suposiciones de normalidad han sido violadas se puede probar un cambio en la ubicación de dos distribuciones poblacionales usando la prueba de suma de rango no paramétrica de Wilcoxon del capítulo 11. Este procedimiento de prueba, que requiere menos suposiciones respecto a la naturaleza de las distribuciones de probabilidad poblacionales, es casi igualmente sensible para detectar una diferencia en medias poblacionales cuando las condiciones necesarias para la prueba t se satisfagan. Puede ser más sensible cuando no se satisfaga la suposición de normalidad.

6.4 EJERCICIOS

TÉCNICAS BÁSICAS

6.17 Dé el número de grados de libertad para s^2 , el estimador agrupado de σ^2 , en estos casos:

- a. $n_1 = 16, n_2 = 8$
- b. $n_1 = 10, n_2 = 12$
- c. $n_1 = 15, n_2 = 3$

6.18 Calcule s^2 , el estimador agrupado de σ^2 , en estos casos:

- a. $n_1 = 10, n_2 = 4, s_1^2 = 3.4, s_2^2 = 4.9$
- b. $n_1 = 12, n_2 = 21, s_1^2 = 18, s_2^2 = 23$

6.19 Dos muestras aleatorias independientes de tamaños $n_1 = 4$ y $n_2 = 5$ se seleccionan de cada una de dos poblaciones normales:

Población 1	12	3	8	5	
Población 2	14	7	7	9	6

- a. Calcule s^2 , el estimador agrupado de σ^2 .
- b. Encuentre un intervalo de confianza de 90% para $(\mu_1 - \mu_2)$, la diferencia entre las dos medias poblacionales.
- c. Pruebe $H_0 : (\mu_1 - \mu_2) = 0$ contra $H_a : (\mu_1 - \mu_2) < 0$ para $\alpha = .05$. Exprese sus conclusiones.

6.20 Muestras aleatorias independientes de $n_1 = 16$ y $n_2 = 13$ observaciones fueron seleccionadas de dos poblaciones normales con iguales varianzas:

	Población	
	1	2
Tamaño muestral	16	13
Media muestral	34.6	32.2
Varianza muestral	4.8	5.9

- a. Suponga que usted desea detectar una diferencia entre las medias poblacionales. Exprese las hipótesis nula y alternativa para la prueba.
- b. Encuentre la región de rechazo para la prueba de la parte a para $\alpha = .01$.
- c. Encuentre el valor del estadístico de prueba.
- d. Encuentre el valor p aproximado para la prueba.
- e. Realice la prueba y exprese sus conclusiones.

6.21 Consulte el ejercicio 6.20. Encuentre un intervalo de confianza de 99% para $(\mu_1 - \mu_2)$.

6.22 La salida impresa MINITAB muestra una prueba de la diferencia en dos medias poblacionales.

Salida impresa MINITAB para el ejercicio 6.22

Prueba T de dos muestras y CI: muestra 1, muestra 2

```
Two-sample T for Sample 1 vs Sample 2
      N      Mean    StDev    SE Mean
Sample 1  6      29.00      4.00      1.6
Sample 2  7      28.86      4.67      1.8

Difference = mu (Sample 1) - mu (Sample 2)
Estimate for difference: 0.14
95% CI for difference: (-5.2, 5.5)
T-Test of difference = 0 (vs not =):
T-Value = 0.06 P-Value = 0.95 DF = 11
Both use Pooled StDev = 4.38
```

- ¿Las dos desviaciones muestrales estándar indican que la suposición de una varianza poblacional común es razonable?
- ¿Cuál es el valor observado del estadístico de prueba? ¿Cuál es el valor p asociado con esta prueba?
- ¿Cuál es la estimación agrupada s^2 de la varianza poblacional?
- Use las respuestas a la parte **b** para sacar conclusiones acerca de la diferencia en las dos medias poblacionales.
- Encuentre el intervalo de confianza del 95% para la diferencia en las medias poblacionales. ¿Este intervalo confirma sus conclusiones de la parte **d**?

6.23 La salida impresa de *MS Excel* muestra una prueba para la diferencia en dos medias poblacionales.

Salida de *MS Excel* para el ejercicio 6.23

D	E	F
t-Test: Two-Sample Assuming Equal Variances		
	Sample 1	Sample 2
Mean	28.667	28.286
Variance	5.067	2.238
Observations	6	7
Pooled Variance	3.524	
Hypothesized Mean Difference	0	
df	11	
t Stat	0.365	
P(T<=t) one-tail	0.361	
t Critical one-tail	1.796	
P(T<=t) two-tail	0.722	
t Critical two-tail	2.201	

- ¿Las dos varianzas muestrales indican que es razonable la suposición de una varianza poblacional común?
- ¿Cuál es el valor observado del estadístico de prueba? Si ésta es una prueba de dos colas, ¿cuál es el valor p asociado con la prueba?
- ¿Cuál es la estimación agrupada s^2 de la varianza poblacional?

- Use las respuestas de la parte **b** para sacar conclusiones acerca de la diferencia en las dos medias poblacionales.
- Use la información de la salida impresa para construir un intervalo de confianza del 95% para la diferencia en las medias poblacionales. ¿Este intervalo confirma sus conclusiones de la parte **d**?

APLICACIONES

6.24 Dientes sanos Un odontólogo realizó un estudio sobre el efecto de un enjuague bucal antiplaca sobre la acumulación de placa en dientes.⁵ Catorce personas cuyos dientes estaban muy limpios y pulidos se asignaron al azar a dos grupos de siete personas cada uno. Ambos grupos fueron asignados para usar enjuagues bucales (no cepillado) durante un periodo de dos semanas. El grupo 1 utilizó un enjuague que contenía un agente antiplaca. El grupo 2, el grupo de control, recibió un enjuague similar excepto que, sin que las personas supieran, el enjuague no contenía agente antiplaca. Un índice x de placa, medida de acumulación de placa, fue registrado a los 14 días. Las medias y desviaciones estándar para los dos grupos se muestran en la tabla.

	Grupo de control	Grupo antiplaca
Tamaño muestral	7	7
Media	1.26	.78
Desviación estándar	.32	.32

- Expresar las hipótesis nula y alternativa que deberían usarse para probar la eficacia del enjuague bucal antiplaca.
- ¿Los datos dan suficiente evidencia para indicar que el enjuague bucal antiplaca es eficaz? Pruebe usando $\alpha = .05$.
- Encuentre el valor p aproximado para la prueba.

DATOS
EX0625

6.25 Atún, otra vez En el ejercicio 6.6 hemos presentado datos sobre el precio promedio estimado para una lata de 6 onzas o una bolsa de 7.06 onzas de atún, con base en precios pagados a nivel nacional en supermercados. Una parte de los datos se reproduce en la tabla siguiente. Use la salida impresa MINITAB para contestar las preguntas.

Atún claro en agua	Atún claro en aceite
.99 .53	2.56 .62
1.92 1.41	1.92 .66
1.23 1.12	1.30 .62
.85 .63	1.79 .65
.65 .67	1.23 .60
.69 .60	.67
.60 .66	

Salida impresa MINITAB para el ejercicio 6.25

Prueba T de dos muestras y CI: agua, aceite

```
Two-sample T for Water vs Oil
      N      Mean    StDev   SE Mean
Water  14      0.896    0.400     0.11
Oil    11      1.147    0.679     0.20

Difference = mu (Water) - mu (Oil)
Estimate for difference: -0.251
95% CI for difference: (-0.700, 0.198)
T-Test of difference = 0 (vs not =):
T-Value = -1.16 P-Value = 0.260 DF = 23
Both use Pooled StDev = 0.5389
```

- ¿Los datos de la tabla presentan suficiente evidencia para indicar una diferencia en el promedio de los precios de atún claro en agua contra atún claro en aceite? Pruebe usando $\alpha = .05$.
- ¿Cuál es el valor p para la prueba?
- El análisis del MINITAB usa la estimación agrupada de σ^2 . ¿La suposición de varianzas iguales es razonable? ¿Por qué sí o por qué no?

6.26 Corredores y ciclistas El síndrome del compartimiento anterior crónico es una afección caracterizada por dolor inducido por ejercicio en la parte inferior de las piernas. Hinchazón y función dañada en nervios y músculos también acompañan a este dolor, que se alivia con reposo. Susan Beckham y colegas efectuaron un experimento, donde participaron 10 corredores sanos y 10 ciclistas sanos, para determinar si hay diferencias significativas en mediciones de presión dentro del compartimiento anterior del músculo para corredores y ciclistas.⁶ El resumen de datos, presión del compartimiento en milímetros de mercurio (Hg), es como sigue:

Afección	Corredores		Ciclistas	
	Media	Desviación estándar	Media	Desviación estándar
En reposo	14.5	3.92	11.1	3.98
80% máximo consumo de O_2	12.2	3.49	11.5	4.95
Máximo consumo de O_2	19.1	16.9	12.2	4.47

- Pruebe para una diferencia significativa en el promedio de la presión del compartimiento entre corredores y ciclistas en la condición de reposo. Use $\alpha = .05$.
- Construya una estimación de intervalo de confianza de 95% de la diferencia en las medias para corredores y ciclistas, en la condición de ejercitarse a 80% de consumo máximo de oxígeno.
- Con el fin de probar para una diferencia significativa en presión del compartimiento a máximo consumo de oxígeno, ¿debe usarse la prueba t agrupada o no agrupada? Explique.

6.27 Desinfectantes Un experimento publicado en *The American Biology Teacher* estudió la eficacia de usar 95% de etanol o 20% de blanqueador, como desinfectante para eliminar la contaminación por bacterias y hongos en cultivos de tejidos de plantas. El experimento se repitió 15 veces con cada desinfectante usando berenjenas como tejido de planta cultivada.⁷ Cinco cortes por planta se colocaron en una caja de Petri para cada desinfectante y se guardaron a 25 °C durante 4 semanas. La observación informada fue el número de cortes no contaminados de berenjena después del almacenamiento de 4 semanas.

Desinfectante	95% etanol	20% de blanqueador
Media	3.73	4.80
Varianza	2.78095	.17143
n	15	15
Varianza agrupada 1.47619		

- ¿Está usted dispuesto a suponer que las varianzas subyacentes son iguales?
- Usando la información de la parte a, ¿está usted dispuesto a concluir que hay una diferencia significativa en la media de los números de berenjenas no contaminadas para los dos desinfectantes probados?

DATOS
EX0628

6.28 Titanio Un geólogo recolectó 20 muestras diferentes de mineral, todas del mismo peso y al azar las dividió en dos grupos. Los contenidos de titanio de las muestras, que encontró usando dos métodos diferentes, se detallan en la tabla:

Método 1					Método 2				
.011	.013	.013	.015	.014	.011	.016	.013	.012	.015
.013	.010	.013	.011	.012	.012	.017	.013	.014	.015

- Use un método apropiado con el fin de probar para una diferencia significativa en el promedio de los contenidos de titanio usando los dos métodos diferentes.
- Determine una estimación del intervalo de confianza de 95% para $(\mu_1 - \mu_2)$. ¿La estimación de su intervalo confirma su conclusión de la parte a? Explique.

DATOS
EX0629

6.29 Pasitas Se contó el número de pasitas en cada una de 14 minicajas (tamaño de 1/2 onza), tanto de una marca genérica de pasitas como de la marca Sunmaid®.

Marca genérica				Sunmaid			
25	26	25	28	25	29	24	24
26	28	28	27	28	24	28	22
26	27	24	25	25	28	30	27
26	26			28	24		

- Aun cuando las cantidades tienen una distribución normal, ¿estos datos tienen distribuciones aproximadamente normales? (SUGERENCIA: Use un histograma o gráfica de tallo y hoja.)

- b. ¿Está usted dispuesto a suponer que las varianzas poblacionales subyacentes son iguales? ¿Por qué?
- c. Use el método del valor p para determinar si hay una diferencia significativa en la media de los números de pasitas por minicaja. ¿Cuáles son las implicaciones de su conclusión?

6.30 Contenido de O_2 disuelto, continúa Consulte el ejercicio 6.7, en el que medimos el contenido de oxígeno disuelto en agua de río para determinar si un arroyo tenía suficiente oxígeno para soportar vida acuática. Un inspector de control de contaminación sospechaba que la comunidad de un río estaba vertiendo cantidades de aguas negras poco tratadas al río. Para comprobar su teoría, sacó cinco especímenes de agua de río seleccionados al azar en un lugar aguas arriba del pueblo, y otros cinco de aguas abajo. Las lecturas de oxígeno disuelto (en partes por millón) son como sigue:

Aguas arriba	4.8	5.2	5.0	4.9	5.1
Aguas abajo	5.0	4.7	4.9	4.8	4.9

- a. ¿Los datos dan suficiente evidencia para indicar que la media del contenido de oxígeno aguas abajo del pueblo es menor que la media del contenido de oxígeno aguas arriba? Pruebe usando $\alpha = .05$.
- b. Supongamos que usted prefiere la estimación como método de inferencia. Estime la diferencia en las medias de los contenidos de oxígeno disuelto para lugares aguas arriba y abajo del pueblo. Use un intervalo de confianza de 95%.

DATOS
EX0631 **6.31 Nadadores de estilo libre** En un esfuerzo por comparar el promedio de los tiempos de natación para dos nadadores, a cada nadador se le pidió nadar en estilo libre una distancia de 100 yardas en tiempos seleccionados al azar. Los nadadores descansaron por completo entre vueltas y no corrieron uno contra otro, de modo que cada muestra de tiempos era una muestra

aleatoria independiente. Se muestran los tiempos para cada una de las 10 pruebas para los dos nadadores.

Nadador 1		Nadador 2	
59.62	59.74	59.81	59.41
59.48	59.43	59.32	59.63
59.65	59.72	59.76	59.50
59.50	59.63	59.64	59.83
60.01	59.68	59.86	59.51

Suponga que el nadador 2 fue el ganador del año pasado cuando los dos nadadores compitieron. ¿Le parece que el promedio de tiempo para el nadador 2 es todavía más rápido que el promedio de tiempo para el nadador 1 en el estilo libre de 100 yardas? Encuentre el valor p aproximado para la prueba e interprete los resultados.

6.32 Nadadores de estilo libre, continúa Consulte el ejercicio 6.31. Construya un límite inferior de confianza de una cola y 95% para la diferencia en los promedios del tiempo para los dos nadadores. ¿Este intervalo confirma sus conclusiones en el ejercicio 6.31?

DATOS
EX0633 **6.33 Un hallazgo arqueológico** Un artículo en *Archaeometry* involucraba un análisis de 26 muestras de cerámica romano–inglesa, halladas en hornos de cuatro sitios diferentes en el Reino Unido.⁸ Las muestras se analizaron para determinar su composición química y el porcentaje de óxido de aluminio en cada una de 10 muestras en dos lugares se observa a continuación.

Island Thorns	Ashley Rails
18.3	17.7
15.8	18.3
18.0	16.7
18.0	14.8
20.8	19.1

¿Los datos dan suficiente información para indicar que hay una diferencia en el promedio de porcentaje de óxido de aluminio en los dos lugares? Pruebe al nivel de significancia de 5%.

INFERENCIAS DE MUESTRA PEQUEÑA PARA LA DIFERENCIA ENTRE DOS MEDIAS: UNA PRUEBA DE DIFERENCIA EN PARES

6.5

Para comparar las cualidades de desgaste de dos tipos de llantas de automóvil, A y B, una llanta de tipo A y una de tipo B se asignaron al azar y se montaron en las ruedas traseras a cada uno de cinco automóviles. Éstos se hicieron correr un número especificado de millas y se registró la cantidad de desgaste para cada llanta. Estas mediciones aparecen en la tabla 6.3. ¿Los datos presentan suficiente evidencia para indicar una diferencia en el promedio de desgaste para los dos tipos de llantas?

TABLA 6.3 Promedio de desgaste para dos tipos de llantas

Automóvil	Llanta A	Llanta B
1	10.6	10.2
2	9.8	9.4
3	12.3	11.8
4	9.7	9.1
5	8.8	8.3
$\bar{x}_1 = 10.24$		$\bar{x}_2 = 9.76$
$s_1 = 1.316$		$s_2 = 1.328$

La tabla 6.3 muestra una diferencia de $(\bar{x}_1 - \bar{x}_2) = (10.24 - 9.76) = .48$ entre las dos medias muestrales, en tanto que las desviaciones estándar de ambas muestras son aproximadamente 1.3. Dada la variabilidad de los datos y el pequeño número de mediciones, ésta es una diferencia bastante pequeña y es probable que usted no sospeche una diferencia en el promedio de desgaste para los dos tipos de llantas. Revisemos las sospechas usando los métodos de la sección 6.4.

Vea el análisis MINITAB de la figura 6.11. La prueba t “agrupada” de dos muestras se usa para probar la diferencia en las medias basada en dos muestras aleatorias independientes. El valor calculado de t usado para probar la hipótesis nula $H_0 : \mu_1 = \mu_2$ es $t = .57$ con valor $p = .582$, valor que no es casi suficientemente pequeño para indicar una diferencia significativa en las dos medias poblacionales. El correspondiente intervalo de confianza de 95%, dado como

$$-1.448 < (\mu_1 - \mu_2) < 2.408$$

es bastante ancho y no indica una diferencia significativa en las medias poblacionales.

FIGURA 6.11

Salida impresa MINITAB usando prueba t para muestras independientes para los datos de llantas

Prueba T de dos muestras y CI: llanta A, llanta B

```
Two-sample T for Tire A vs Tire B
      N      Mean    StDev    SE Mean
Tire A  5      10.24     1.32     0.59
Tire B  5       9.76     1.33     0.59

Difference = mu (Tire A) - mu (Tire B)
Estimate for difference: 0.480
95% CI for difference: (-1.448, 2.408)
T-Test of difference = 0 (vs not =): T-Value = 0.57 P-Value = 0.582 DF = 8
Both use Pooled StDev = 1.3221
```

Analice de nuevo los datos y observará que la medición del desgaste para el tipo A es mayor que el correspondiente valor para el tipo B, para *cada uno* de los cinco automóviles. ¿No sería esto improbable, si en realidad no hay diferencia entre los dos tipos de llantas?

Considere una prueba intuitiva sencilla, basada en la distribución binomial del capítulo 3. Si no hay diferencia en el desgaste medio para los dos tipos de llantas, entonces es igualmente probable o no probable que la llanta A muestre más desgaste que la llanta B. Los cinco automóviles entonces corresponden a cinco intentos binomiales con $p = P(\text{llanta A muestra más desgaste que la llanta B}) = .5$.

¿Es poco común el valor observado de $x = 5$ diferencias positivas mostradas en la tabla 6.4? La probabilidad de observar $x = 5$ o el igualmente probable valor de $x = 0$ se puede hallar en la tabla 1 del Anexo que son $2(.031) = .062$, que es bastante pequeño en comparación con la probabilidad de la más potente prueba t , que tenía un valor p de .58. ¿No es peculiar que la prueba t , que usa más información (la muestra real de mediciones) que la prueba binomial, no dé suficiente información para rechazar la hipótesis nula?

Hay una explicación para esta inconsistencia. La prueba t descrita en la sección 6.4 *no* es el estadístico de prueba apropiado para aplicarlo a nuestro ejemplo. El procedimiento de prueba

estadística de la sección 6.4 requiere que las dos muestras sean *independientes* y *aleatorias*. Ciertamente, el requisito de independencia es violado por la forma en la que se realizó el experimento.

El par de mediciones, en las llantas A y B, para un automóvil particular están definitivamente relacionadas. Una mirada a los datos muestra que las lecturas tienen más o menos la misma magnitud para un automóvil particular pero varían en forma marcada de un automóvil a otro. Esto, por supuesto, es exactamente lo que podría esperarse. El desgaste de llantas está determinado en su mayor parte por los hábitos de manejo, el balanceo de las ruedas y la superficie del pavimento. Como cada automóvil tiene un conductor diferente, es de esperarse una gran cantidad de variabilidad en los datos de un automóvil a otro.

Al diseñar el experimento de desgaste de llantas, el experimentador vio que las mediciones variarían en gran medida de un automóvil a otro. Si las llantas (cinco del tipo A y cinco del tipo B) se asignaran al azar a las 10 ruedas, resultando en muestras aleatorias *independientes*, esta variabilidad resultaría en un gran error estándar y haría difícil de detectar una diferencia en las medias. En cambio, el experimentador eligió “parear” las mediciones, comparando el desgaste para llantas tipo A y tipo B en cada uno de los cinco automóviles. Este diseño experimental, a veces llamado diseño de **diferencia en pares** o **pares acoplados**, nos permite eliminar la variabilidad de un auto a otro al observar sólo las cinco mediciones de diferencia mostradas en la tabla 6.4. Estas cinco diferencias forman una sola muestra aleatoria de tamaño $n = 5$.

TABLA 6.4

Diferencias en desgaste de llantas, usando los datos de la tabla 6.3

Automóvil	A	B	$d = A - B$
1	10.6	10.2	.4
2	9.8	9.4	.4
3	12.3	11.8	.5
4	9.7	9.1	.6
5	8.8	8.3	.5
			$\bar{d} = .48$

Observe que en la tabla 6.4 la media muestral de las diferencias, $d = A - B$, se calcula como

$$\bar{d} = \frac{\sum d_i}{n} = .48$$

y es exactamente la misma que la diferencia de las medias muestrales: $(\bar{x}_1 - \bar{x}_2) = (10.24 - 9.76) = .48$. No debe sorprender que esto se pueda demostrar como verdadero en general y también que la misma relación se cumpla para las medias poblacionales. De este modo, el promedio de las diferencias poblacionales es

$$\mu_d = (\mu_1 - \mu_2)$$

Debido a esto, se pueden usar las diferencias muestrales para probar una diferencia significativa en las dos medias poblacionales, $(\mu_1 - \mu_2) = \mu_d$. La prueba es una prueba t de una sola muestra de las mediciones de diferencia para probar la hipótesis nula

$$H_0 : \mu_d = 0 \quad [\text{o } H_0 : (\mu_1 - \mu_2) = 0]$$

contra la hipótesis alternativa

$$H_a : \mu_d \neq 0 \quad [\text{o } H_a : (\mu_1 - \mu_2) \neq 0]$$

Los procedimientos de prueba toman la misma forma que los procedimientos empleados en la sección 6.3 y se describen a continuación.

PRUEBA DE HIPÓTESIS DE DIFERENCIA EN PARES PARA $(\mu_1 - \mu_2) = \mu_d$: MUESTRAS DEPENDIENTES

1. Hipótesis nula: $H_0: \mu_d = 0$

2. Hipótesis alternativa:

Prueba de una cola

$$H_a: \mu_d > 0$$

(o $H_a: \mu_d < 0$)

Prueba de dos colas

$$H_a: \mu_d \neq 0$$

3. Estadístico de prueba: $t = \frac{\bar{d} - 0}{s_d/\sqrt{n}} = \frac{\bar{d}}{s_d/\sqrt{n}}$

donde n = Número de diferencias en pares

$\bar{d}$ = Media de las diferencias muestrales

s_d = Desviación estándar de las diferencias muestrales

$$= \sqrt{\frac{\sum(d_i - \bar{d})^2}{n - 1}} = \sqrt{\frac{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}{n - 1}}$$

4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola

$$t > t_\alpha$$

(o $t < -t_\alpha$ cuando la hipótesis alternativa sea $H_a: \mu_d < 0$)

o cuando valor $p < \alpha$

Prueba de dos colas

$$t > t_{\alpha/2} \quad \text{o} \quad t < -t_{\alpha/2}$$

Los valores críticos de t , t_α y $t_{\alpha/2}$ están basados en $(n - 1)$ gl. Estos valores tabulados se pueden hallar usando la tabla 4 en el Anexo.

INTERVALO DE CONFIANZA DE MUESTRA PEQUEÑA $(1 - \alpha)100\%$ PARA $(\mu_1 - \mu_2) = \mu_d$, CON BASE EN UN EXPERIMENTO DE DIFERENCIA EN PARES

$$\bar{d} \pm t_{\alpha/2} \left(\frac{s_d}{\sqrt{n}} \right)$$

Suposiciones: El experimento está diseñado como una prueba de diferencia en pares de modo que las n diferencias representan una muestra aleatoria tomada de una población normal.

EJEMPLO

6.8

¿Los datos de la tabla 6.3 dan suficiente evidencia para indicar una diferencia en el desgaste medio para llantas tipo A y B? Pruebe usando $\alpha = .05$.

Solución Usted puede verificar, usando su calculadora, que el promedio y desviación estándar de las cinco mediciones de diferencia son

$$\bar{d} = .48 \quad \text{y} \quad s_d = \sqrt{\frac{1.18 - \frac{(2.4)^2}{5}}{4}} = .0837$$

Entonces

$$H_0: \mu_d = 0 \quad \text{y} \quad H_a: \mu_d \neq 0$$

y

$$t = \frac{\bar{d} - 0}{s_d/\sqrt{n}} = \frac{.48}{.0837/\sqrt{5}} = 12.8$$

El valor crítico de t para una prueba estadística de dos colas, $\alpha = .05$ y 4 *gl*, es 2.776. Ciertamente, el valor observado de $t = 12.8$ es muy grande y significativo. En consecuencia, se puede concluir que hay una diferencia en el desgaste medio para llantas tipo A y B.

EJEMPLO

6.9

Encuentre un intervalo de confianza de 95% para $(\mu_1 - \mu_2) = \mu_d$ usando los datos de la tabla 6.3.

Solución Un intervalo de confianza de 95% para la diferencia entre la media de los niveles de desgaste es

$$\begin{aligned} \bar{d} \pm t_{\alpha/2} \left(\frac{s_d}{\sqrt{n}} \right) \\ .48 \pm 2.776 \left(\frac{.0837}{\sqrt{5}} \right) \\ .48 \pm .10 \end{aligned}$$

TIP MI CONSEJO

¡Los intervalos de confianza siempre se interpretan en la misma forma! En el muestreo repetido, los intervalos contruidos de esta manera encierran el valor verdadero del parámetro $100(1 - \alpha)\%$ del tiempo

o sea $.38 < (\mu_1 - \mu_2) < .58$. ¿Cómo se compara el ancho de este intervalo con el ancho de un intervalo que pudiera haberse construido si se hubiera diseñado el experimento en forma no pareada? Es probable que fuera de la misma magnitud que el intervalo calculado en la figura 6.11, donde los datos observados fueron *incorrectamente* analizados usando el análisis no pareado. Este intervalo, $-1.448 < (\mu_1 - \mu_2) < 2.408$, es mucho más ancho que el intervalo en pares, lo cual indica que el diseño de diferencia en pares aumentó la precisión de nuestra estimación y hemos obtenido valiosa información con el uso de este diseño.

La *prueba de diferencia en pares* o *diseño de pares acoplados* empleados en el experimento de desgaste de llantas es un ejemplo sencillo de un diseño experimental denominado **diseño de bloque aleatorizado**. Cuando hay gran cantidad de variabilidad entre las unidades experimentales, aun antes de ponerse en práctica ningún procedimiento experimental, el efecto de esta variabilidad se puede reducir al mínimo al poner en **bloque**, es decir, comparar los diferentes procedimientos dentro de grupos de unidades experimentales relativamente similares llamadas **bloques**. De este modo, el “ruido” causado por la gran variabilidad no oculta las verdaderas diferencias entre los procedimientos. En el capítulo 7 analizaremos con más detalle los diseños aleatorizados de bloque.

Es importante que recuerde que el *pareo* o *poner en bloque* se presenta cuando se planea el experimento y no después de recolectar los datos. Un experimentador tiene la opción de usar pares de gemelos idénticos para comparar dos métodos de aprendizaje. Un médico puede registrar la presión sanguínea de un paciente antes y después de administrar un medicamento en particular. Una vez que usted haya empleado un diseño en pares para un experimento, ya no tiene la opción de usar el análisis no pareado de la sección 6.4. La suposición de independencia ha sido violada a propósito y la única opción es usar el análisis en pares descrito aquí.

Aun cuando hacer un pareado fue benéfico en el experimento de desgaste de llantas, éste puede no ser siempre el caso. En el análisis en pares, los grados de libertad para la prueba t se cortan a la mitad, es decir, de $(n + n - 2) = 2(n - 1)$ a $(n - 1)$. Esta reducción *aumenta* el valor crítico de t para rechazar H_0 y también aumenta el ancho del intervalo de confianza para la diferencia en las dos medias. Si el pareado no es eficiente, este aumento no es compensado por una disminución en la variabilidad y de hecho se puede perder y no ganar información por el

TIP MI CONSEJO

Prueba de diferencia en pares: $gl = n - 1$

pareado. Esto por supuesto que no ocurrió en el experimento de las llantas, ya que la reducción grande en el error estándar compensó la pérdida en grados de libertad.

Excepto por la notación, el análisis de diferencia en pares es igual al análisis de una sola muestra presentado en la sección 6.3. No obstante, *MINITAB* y *MS Excel* dan un solo procedimiento (**Paired *t*** en *MINITAB* y ***t*-Test: Paired Two Sample for Means** en *MS Excel*) para analizar las diferencias. La salida de *MINITAB*, que se muestra en la figura 6.12, presenta el valor *p* para el análisis en pares, .000, que indica una diferencia altamente significativa en las medias. Encontrará instrucciones para generar salidas de *MINITAB* y de *MS Excel* en la sección “Tecnología actual” al final de este capítulo.

FIGURA 6.12

Salida impresa *MINITAB* para análisis de diferencia pareada de datos del desgaste de llantas

Prueba T pareada y CI: llanta A, llanta B

Paired T for Tire A - Tire B

	N	Mean	StDev	SE Mean
Tire A	5	10.240	1.316	0.589
Tire B	5	9.760	1.328	0.594
Difference	5	0.4800	0.0837	0.0374

95% CI for mean difference: (0.3761, 0.5839)

T-Test of mean difference = 0 (vs not = 0): T-Value = 12.83 P-Value = 0.000

6.5 EJERCICIOS

TÉCNICAS BÁSICAS

6.34 Se realizó un experimento de diferencia pareada usando $n = 10$ pares de observaciones.

- Pruebe la hipótesis nula $H_0: (\mu_1 - \mu_2) = 0$ contra $H_a: (\mu_1 - \mu_2) \neq 0$ para $\alpha = .05$, $\bar{d} = .3$ y $s_d^2 = .16$. Dé el valor *p* aproximado para la prueba.
- Encuentre un intervalo de confianza de 95% para $(\mu_1 - \mu_2)$.
- ¿Cuántos pares de observaciones se necesitan si se desea estimar $(\mu_1 - \mu_2)$ correcta a no más de .1 con probabilidad igual a .95?

6.35 Un experimento de diferencia en pares consta de $n = 18$ pares, $\bar{d} = 5.7$, y $s_d^2 = 256$. Suponga que se desea detectar $\mu_d > 0$.

- Dé las hipótesis nula y alternativa para la prueba.
- Efectúe la prueba y exprese sus conclusiones.

6.36 Se realizó un experimento de diferencia en pares para comparar las medias de dos poblaciones:

Población	Pares				
	1	2	3	4	5
1	1.3	1.6	1.1	1.4	1.7
2	1.2	1.5	1.1	1.2	1.8

- ¿Los datos dan suficiente evidencia para indicar que μ_1 difiere de μ_2 ? Pruebe usando $\alpha = .05$.
- Encuentre el valor *p* aproximado para la prueba e interprete su valor.
- Encuentre un intervalo de confianza de 95% para $(\mu_1 - \mu_2)$. Compare su interpretación del intervalo de confianza con sus resultados de la prueba en la parte a.

- ¿Qué suposiciones se deben hacer para lograr que sus inferencias sean válidas?

APLICACIONES

6.37 Corredores y ciclistas II Consulte el ejercicio 6.26. Además de las presiones del compartimiento, el nivel de creatina fosfoquinasa (CPK) en muestras de sangre, una medida del daño muscular, se determinó para cada uno de los 10 corredores y 10 ciclistas antes y después del ejercicio.⁶ El resumen de datos, es decir, valores de CPK en unidades/litro, es como sigue:

Afección	Corredores		Ciclistas	
	Media	Desviación estándar	Media	Desviación estándar
Antes de ejercicio	255.63	115.48	173.8	60.69
Después de ejercicio	284.75	132.64	177.1	64.53
Diferencia	29.13	21.01	3.3	6.85

- Pruebe para una diferencia significativa en la media de los valores de CPK para corredores y ciclistas antes del ejercicio, bajo la suposición de que $\sigma_1^2 \neq \sigma_2^2$; use $\alpha = .05$.
- Pruebe para una diferencia significativa en la media de los valores de CPK para corredores y ciclistas después del ejercicio, bajo la suposición de que $\sigma_1^2 \neq \sigma_2^2$; use $\alpha = .05$. Encuentre una estimación de intervalo de confianza de 95% para la correspondiente diferencia en medias.
- Pruebe para una diferencia significativa en la media de los valores de CPK para corredores antes y después del ejercicio.
- Encuentre un intervalo de confianza de 95% estimado para la diferencia en la media de los valores de CPK

para los ciclistas antes y después del ejercicio. ¿Su estimación indica que no hay diferencia significativa en la media de los niveles de CPK para ciclistas antes y después del ejercicio?

DATOS
EX0638

6.38 Prohibida la vuelta a la izquierda Se realizó un experimento para comparar la media de los tiempos de reacción a dos tipos de señales de tránsito: restrictivos (prohibida la vuelta a la izquierda) y permisivos (sólo vuelta a la izquierda). Diez conductores se incluyeron en el experimento. A cada uno se le presentaron 40 señales de tránsito, 20 restrictivas y 20 permisivas, en orden aleatorio. Se registró el tiempo medio de reacción (en milisegundos) para cada conductor y se muestra aquí.

Conductor	Restrictivas	Permisivas
1	824	702
2	866	725
3	841	744
4	770	663
5	829	792
6	764	708
7	857	747
8	831	685
9	846	742
10	759	610

Salida impresa de *MS Excel* para el ejercicio 6.38

D	E	F
t-Test: Paired Two Sample for Means		
	<i>Prohibitive</i>	<i>Permissive</i>
Mean	818.7	711.8
Variance	1573.344444	2596.4
Observations	10	10
Pearson Correlation	0.693852753	
Hypothesized Mean Difference	0	
df	9	
t Stat	9.14983257	
P(T<=t) one-tail	3.72891E-06	
t Critical one-tail	1.833112933	
P(T<=t) two-tail	7.45782E-06	
t Critical two-tail	2.262157163	

- Explique por qué es un experimento de diferencia en pares; dé razones por las que el pareado debe ser útil para aumentar información sobre la diferencia entre la media de los tiempos de reacción a señales de tránsito restrictivas y permisivas.
- Use la salida impresa de *Excel* para determinar si hay una diferencia significativa en la media de los tiempos de reacción a las señales de tránsito restrictivas y permisivas. Use el enfoque del valor p .

6.39 Dientes sanos II El ejercicio 6.24 describe un experimento dental realizado para investigar la efectividad de un enjuague bucal empleado para inhibir el crecimiento de placa en los dientes. Los sujetos se dividieron en dos grupos: un grupo usó un enjuague con

un ingrediente antiplaca y el grupo de control usó un enjuague que contenía ingredientes inactivos. Suponga que el crecimiento de placa en los dientes de cada persona se midió luego de usar el enjuague después de cuatro horas y además otra vez después de ocho horas. Si usted desea estimar la diferencia en crecimiento de placa de 4 a 8 horas, ¿debe usar un intervalo de confianza basado en un análisis en pares o no pareado? Explique.

6.40 ¿Tierra o aire? La temperatura de la tierra se puede medir usando ya sea sensores al nivel del suelo o aparatos sensibles a rayos infrarrojos instalados en aviones o satélites espaciales. Los sensores instalados en tierra son muy precisos pero tediosos, mientras los sensores de ondas infrarrojas parecen introducir un sesgo en las lecturas de temperatura; es decir, la lectura del promedio de la temperatura puede no ser igual al promedio obtenido por los sensores al nivel del suelo. Para determinar el sesgo, se obtuvieron lecturas en cinco lugares diferentes usando sensores tanto en tierra como en aviones. Las lecturas (en grados Celsius) son como sigue:

Lugar	Tierra	Aire
1	46.9	47.3
2	45.4	48.1
3	36.3	37.9
4	31.0	32.7
5	24.7	26.2

- ¿Los datos presentan suficiente evidencia para indicar un sesgo en las lecturas de temperatura basadas en el aire? Explique.
- Estime la diferencia en temperaturas medias entre sensores basados en tierra y aire usando un intervalo de confianza de 95%.
- ¿Cuántas observaciones en pares se requieren para estimar la diferencia entre temperaturas medias para sensores basados en tierra contra los basados en aire, correcta a no más de $.2^{\circ}\text{C}$, con probabilidad aproximadamente igual a .95?

DATOS
EX0641

6.41 Tintura roja Para probar el color comparativo de dos tinturas rojas se tomaron nueve muestras de tela de una línea de producción y cada muestra se dividió en dos piezas. Una de las dos piezas de cada muestra se eligió al azar y se aplicó la tintura roja 1; la tintura roja 2 se aplicó a la pieza restante. Los datos siguientes representan una “calificación de color” para cada pieza. ¿Hay suficiente evidencia para indicar una diferencia en calificación media de color para las dos tinturas? Use $\alpha = .05$.

Muestra	1	2	3	4	5	6	7	8	9
Tintura 1	10	12	9	8	15	12	9	10	15
Tintura 2	8	11	10	6	12	13	9	8	13

DATOS
EX0642

6.42 Asesores de impuestos En respuesta a una queja de que un asesor de impuestos en particular (A) estaba sesgado, se realizó un experimento para comparar al asesor citado en la queja con otro asesor de impuestos (B) de la misma oficina. Se seleccionaron ocho propiedades y cada una fue evaluada por ambos asesores. Las evaluaciones (en miles de dólares) se muestran en la tabla.

Propiedad	Asesor 1	Asesor 2
1	276.3	275.1
2	288.4	286.8
3	280.2	277.3
4	294.7	290.6
5	268.7	269.1
6	282.8	281.0
7	276.1	275.3
8	279.0	279.1

Use la salida impresa *MINITAB* para contestar las preguntas.

Salida impresa *MINITAB* para el ejercicio 6.42

Prueba T pareada y CI: asesor A, asesor B

```
Paired T for Assessor A - Assessor B
      N      Mean    StDev   SE Mean
Assessor A   8    280.78     7.99     2.83
Assessor B   8    279.29     6.85     2.42
Difference   8     1.487     1.491     0.527

95% lower bound for mean difference: 0.489
T-Test of mean difference = 0 (vs > 0):
T-Value = 2.82 P-value = 0.013
```

- ¿Los datos dan suficiente evidencia para indicar que el asesor A tiende a dar evaluaciones más altas que el asesor B?
- Estime la diferencia en evaluaciones medias para los dos asesores.

- ¿Qué suposiciones deben hacerse para que las inferencias en las partes **a** y **b** sean válidas?
- Suponga que el asesor A había sido comparado con un estándar más estable, es decir, el promedio de las evaluaciones dadas por cuatro asesores seleccionados de la oficina de impuestos. Así, cada propiedad sería evaluada por A y también por cada uno de los otros cuatro asesores y $(x_A - \bar{x})$ se calcularía. Si la prueba de la parte **a** es válida, ¿puede usted usar la prueba *t* de diferencia en pares para probar la hipótesis de que el sesgo, la diferencia media entre las evaluaciones de A y la media de las evaluaciones de los otros cuatro asesores, es igual a 0? Explique.

DATOS
EX0643

6.43 Experimentos de memoria Un grupo de psicología realizó un experimento para comparar si una calificación de recordatorio, en la que se dieron instrucciones para formar imágenes de 25 palabras, es mejor que una calificación inicial de recordatorio para la cual no se dieron instrucciones para formar imágenes. Veinte estudiantes participaron en el experimento con los resultados siguientes:

Estudiante	Con imágenes	Sin imágenes	Estudiante	Con imágenes	Sin imágenes
1	20	5	11	17	8
2	24	9	12	20	16
3	20	5	13	20	10
4	18	9	14	16	12
5	22	6	15	24	7
6	19	11	16	22	9
7	20	8	17	25	21
8	19	11	18	21	14
9	17	7	19	19	12
10	21	9	20	23	13

¿Le parece que el promedio de calificaciones de recordatorio es más alto cuando se usan imágenes?

6.6

INFERENCIAS RESPECTO A LA VARIANZA POBLACIONAL

Ya hemos visto en las secciones precedentes que una estimación de la varianza poblacional σ^2 suele ser necesaria antes de hacer inferencias acerca de medias poblacionales, pero a veces la varianza poblacional σ^2 es el objetivo principal en una investigación experimental. ¡Ésta llega a ser *más* importante para el experimentador que la media poblacional! Considere estos ejemplos:

- Los instrumentos de mediciones científicas deben dar lecturas no sesgadas con un error de medición muy pequeño. Un altímetro de un avión que mida la altitud correcta en *promedio* es bastante inútil si las mediciones tienen un error de hasta 1 000 pies arriba o debajo de la altitud correcta.

- Las piezas maquinadas en un proceso de manufactura deben ser producidas con mínima variabilidad para reducir piezas de dimensiones incorrectas y, por tanto, defectuosas.
- Las pruebas de aptitud deben estar diseñadas de manera que las calificaciones *exhibirán* una cantidad razonable de variabilidad. Por ejemplo, un examen de 800 puntos no es muy discriminatorio si todos los estudiantes obtienen calificaciones entre 601 y 605.

En capítulos previos, hemos usado

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

como estimador no sesgado de la varianza poblacional σ^2 . Esto significa que, en muestreo repetido, el promedio de todas las estimaciones muestrales será igual al parámetro objetivo, σ^2 . Pero, ¿qué tan cercano o lejano del objetivo es probable que esté su estimador s^2 ? Para contestar esta pregunta, usamos la distribución de muestreo de s^2 , que describe su comportamiento en muestreo repetido.

Considere la distribución de s^2 basada en muestreo *aleatorio* repetido de una distribución *normal* con una media y varianza especificadas. Podemos demostrar teóricamente que la distribución empieza en $s^2 = 0$ (porque la varianza no puede ser negativa) con una media igual a σ^2 . Su forma es *no simétrica* y cambia con cada tamaño muestral diferente y cada valor diferente de σ^2 . Hallar valores críticos para la distribución de muestreo de s^2 sería muy difícil y requeriría tablas separadas para cada varianza poblacional. Por fortuna, podemos simplificar el problema por *estandarización*, como hicimos con la distribución z .

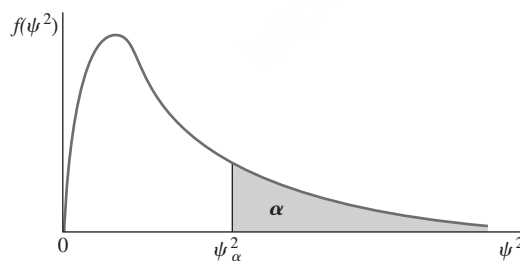
Definición La estadística estandarizada

$$\chi^2 = \frac{(n - 1)s^2}{\sigma^2}$$

recibe el nombre de **variable ji cuadrada** y tiene una distribución de muestreo llamada **distribución de probabilidad ji cuadrada**, con $n - 1$ grados de libertad.

La ecuación de la función de densidad para esta estadística es por demás complicada al verla, pero traza la curva que se muestra en la figura 6.13.

FIGURA 6.13
Distribución ji cuadrada



Ciertos valores críticos del estadístico ji cuadrada, que se usan para hacer inferencias acerca de la varianza poblacional, han sido tabulados por estadísticos y aparecen en la tabla 5 del Anexo. Como la forma de la distribución varía con el tamaño muestral n o, con mayor precisión, los grados de libertad, $n - 1$, asociados con s^2 , tabla 5, parcialmente reproducidos en la tabla 6.5, se construye exactamente en la misma forma que la tabla t , con los grados de libertad en la primera y última columnas. El símbolo χ^2_α indica que el valor χ^2 tabulado tiene un área α a su derecha (véase la figura 6.13).

TABLA 6.5 Formato de la tabla ji cuadrada de la tabla 5 del Anexo

TIP MI CONSEJO

Probando una varianza:
 $gl = n - 1$

MI APPLET EN LÍNEA

Probabilidades ji
cuadrada

gl	$\chi^2_{.995}$	...	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	...	$\chi^2_{.005}$	gl
1	.0000393		.0039321	.0157908	2.70554	3.84146		7.87944	1
2	.0100251		.102587	.210720	4.60517	5.99147		10.5966	2
3	.0717212		.351846	.584375	6.25139	7.81473		12.8381	3
4	.206990		.710721	1.063623	7.77944	9.48773		14.8602	4
5	.411740		1.145476	1.610310	9.23635	11.0705		16.7496	5
6	.0675727		1.63539	2.204130	10.6446	12.5916		18.5476	6
.	.		.	.	.	.		.	.
.	.		.	.	.	.		.	.
.	.		.	.	.	.		.	.
15	4.60094		7.26094	8.54675	22.3072	24.9958		32.8013	15
16	5.14224		7.96164	9.31223	23.5418	26.2962		34.2672	16
17	5.69724		8.67176	10.0852	24.7690	27.5871		35.7185	17
18	6.26481		9.39046	10.8649	25.9894	28.8693		37.1564	18
19	6.84398		10.1170	11.6509	27.2036	30.1435		38.5822	19
.	.		.	.	.	.		.	.
.	.		.	.	.	.		.	.
.	.		.	.	.	.		.	.

Se observa en la tabla 6.5 que, debido a que la distribución no es simétrica y empieza en 0, las áreas de cola superior e inferior deben ser tabuladas para el estadístico ji cuadrada. Por ejemplo, el valor $\chi^2_{.95}$ es el valor que tiene 95% del área bajo la curva a su derecha y 5% del área a su izquierda. Este valor corta un área igual a .05 en la cola inferior de la distribución ji cuadrada.

EJEMPLO 6.10

- Pruebe su capacidad para usar la tabla 5 del Anexo al verificar las siguientes frases:
1. La probabilidad de que χ^2 , basada en $n = 16$ mediciones ($gl = 15$), exceda de 24.9958 es .05.
 2. Para una muestra de $n = 6$ mediciones, 95% del área bajo la distribución X^2 está a la derecha de 1.145476.

Estos valores están sombreados en la tabla 6.5.

La prueba estadística de una hipótesis nula respecto a una varianza poblacional

$$H_0 : \sigma^2 = \sigma_0^2$$

usa el estadístico de prueba

$$\chi^2 = \frac{(n - 1)s^2}{\sigma_0^2}$$

Observe que cuando H_0 es verdadera, s^2/σ_0^2 debe ser cercana a 1, de modo que χ^2 debe ser cercana a $(n - 1)$, los grados de libertad. Si σ^2 es realmente mayor que el valor hipotético σ_0^2 , el estadístico de prueba tenderá a ser mayor a $(n - 1)$ y es probable que caiga hacia la cola superior de la distribución. Si $\sigma^2 < \sigma_0^2$, el estadístico de prueba tenderá a ser menor que $(n - 1)$ y probablemente caerá hacia la cola inferior de la distribución ji cuadrada. Al igual que en otras situaciones de prueba, se puede usar una prueba estadística ya sea de una o de dos colas, dependiendo de la hipótesis alternativa. Esta prueba de hipótesis y el intervalo de confianza $(1 - \alpha)100\%$ para σ^2 están basados ambos en la distribución ji cuadrada y se describen a continuación.

PRUEBA DE HIPÓTESIS RESPECTO A UNA VARIANZA POBLACIONAL

1. Hipótesis nula: $H_0 : \sigma^2 = \sigma_0^2$
2. Hipótesis alternativa:

Prueba de una cola

$$H_a : \sigma^2 > \sigma_0^2$$

$$(\text{o } H_a : \sigma^2 < \sigma_0^2)$$

Prueba de dos colas

$$H_a : \sigma^2 \neq \sigma_0^2$$

$$3. \text{ Estadístico de prueba: } \chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola

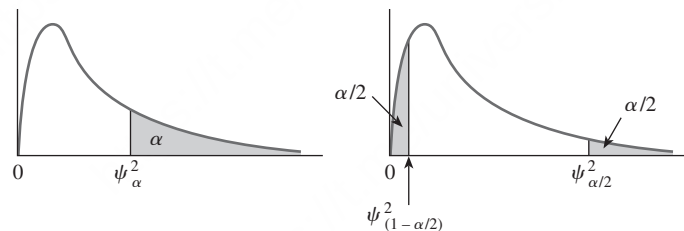
$\chi^2 > \chi_{\alpha}^2$
 (o $\chi^2 < \chi_{(1-\alpha)}^2$ cuando la hipótesis alternativa sea $H_a : \sigma^2 < \sigma_0^2$), donde χ_{α}^2 y $\chi_{(1-\alpha)}^2$ son respectivamente, los valores de cola superior e inferior de χ^2 que pone α en las áreas de cola

Prueba de dos colas

$\chi^2 > \chi_{\alpha/2}^2$ o $\chi^2 < \chi_{(1-\alpha/2)}^2$,
 donde $\chi_{\alpha/2}^2$ y $\chi_{(1-\alpha/2)}^2$ son, respectivamente, los valores de cola superior e inferior de χ^2 que pone $\alpha/2$ en las áreas de cola

o cuando valor $p < \alpha$

Los valores críticos de χ^2 están basados en $(n-1)$ gl. Estos valores tabulados se pueden hallar usando la tabla 5 del Anexo.



INTERVALO DE CONFIANZA DE $(1 - \alpha)100\%$ PARA σ^2

$$\frac{(n-1)s^2}{\chi_{\alpha/2}^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_{(1-\alpha/2)}^2}$$

donde $\chi_{\alpha/2}^2$ y $\chi_{(1-\alpha/2)}^2$ son los valores χ^2 superior e inferior, que localizan la mitad de α en cada cola de la distribución ji cuadrada.

Suposición: La muestra se selecciona al azar de una población normal.

EJEMPLO

6.11

Un fabricante de cemento declara que el concreto preparado con el producto de él tiene resistencia relativamente estable a la compresión y que la resistencia medida en kilogramos por centímetro cuadrado (kg/cm^2) está dentro de un rango de 40 kg/cm^2 . Una muestra de $n = 10$ mediciones produjo una media y varianza igual a, respectivamente,

$$\bar{x} = 312 \quad \text{y} \quad s^2 = 195$$

¿Estos datos presentan suficiente evidencia para rechazar lo dicho por el fabricante?

Solución En la sección 1.5, usted aprendió que el rango de un conjunto de mediciones debe ser aproximadamente cuatro desviaciones estándar. La afirmación del fabricante de que el rango de las mediciones de resistencia está dentro de 40 kg/cm^2 debe significar que la desviación estándar de las mediciones es casi 10 kg/cm^3 o menos. Para probar su declaración, las hipótesis apropiadas son

$$H_0: \sigma^2 = 10^2 = 100 \text{ contra } H_a: \sigma^2 > 100$$

Si la varianza muestral es mucho mayor que el valor hipotético de 100, entonces el estadístico de prueba

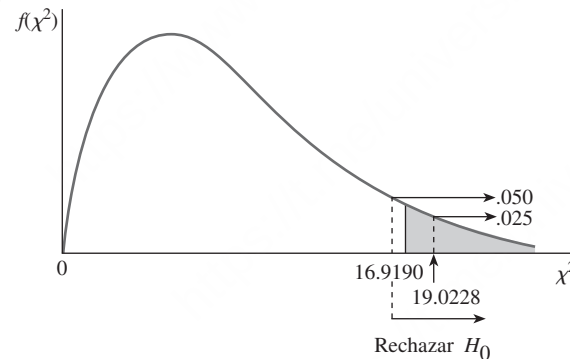
$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{1755}{100} = 17.55$$

será inusualmente grande, favoreciendo el rechazo de H_0 y la aceptación de H_a . Hay dos formas de usar el estadístico de prueba para tomar una decisión para esta prueba.

- **El método del valor crítico:** La prueba apropiada requiere una región de rechazo de una cola en la cola derecha de la distribución χ^2 . El valor crítico para $\alpha = .05$ y $(n-1) = 9 \text{ gl}$ es $\chi^2_{.05} = 16.9190$ de la tabla 5 en el Anexo. La figura 6.14 muestra la región de rechazo; se puede rechazar H_0 si el estadístico de prueba excede de 16.9190. Como el valor observado del estadístico de prueba $\chi^2 = 17.55$, se puede concluir que la hipótesis nula es falsa y que el rango de mediciones de resistencia del concreto excede la declaración del fabricante.

FIGURA 6.14

Región de rechazo y valor p (sombreado) para el ejemplo 6.11



- **El método del valor p :** El valor p para una prueba estadística es el mínimo valor de α para el cual H_0 puede ser rechazado. Se calcula, al igual que en otras pruebas de una cola, como el área en la cola de la distribución χ^2 a la derecha del valor observado, $\chi^2 = 17.55$. Aun cuando algunos paquetes de computadora permiten calcular esta área con toda exactitud, la tabla 5 del Anexo permite sólo acotar el valor p . Como el valor 17.55 está entre $\chi^2_{.050} = 16.9190$ y $\chi^2_{.025} = 19.0228$, el valor p está entre .025 y .05. La mayoría de investigadores rechazarían H_0 y presentan estos resultados como significativos al nivel de 5%, o $P < .05$. De nueva cuenta, el investigador puede rechazar H_0 y concluir que el rango de mediciones excede de lo dicho por el fabricante.

EJEMPLO

6.12

Una experimentadora está convencida de que su instrumento de medición tenía una variabilidad medida por la desviación estándar $\sigma = 2$. Durante un experimento, ella registró las mediciones 4.1, 5.2 y 10.2. ¿Estos datos confirman o desaprueban lo dicho por ella? Pruebe la hipótesis apropiada y construya un intervalo de confianza de 90% para estimar el verdadero valor de la varianza de población.

Solución Como no hay nivel de significancia establecido de antemano, se debe elegir usar el método del valor p para probar estas hipótesis:

$$H_0 : \sigma^2 = 4 \quad \text{contra} \quad H_a : \sigma^2 \neq 4$$

Use su calculadora científica para verificar que la varianza muestral es $s^2 = 10.57$ y el estadístico de prueba es

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{2(10.57)}{4} = 5.285$$

Como ésta es una prueba de dos colas, la región de rechazo se divide en dos partes, la mitad en cada cola de la distribución χ^2 . Si se aproxima el área a la derecha del estadístico de prueba observado, $\chi^2 = 5.285$, se tendrá sólo *la mitad* del valor p para la prueba. Como un valor igualmente improbable de χ^2 podría presentarse en la cola inferior de la distribución, con igual probabilidad, se debe *duplicar* el área superior para obtener el valor p . Con 2 *gl*, el valor observado, 5.29, cae entre $\chi^2_{.10}$ y $\chi^2_{.05}$ de modo que

$$.05 < \frac{1}{2}(\text{valor } p) < .10 \quad \text{o} \quad .10 < \text{valor } p < .20$$

Ya que el valor p es mayor que .10, los resultados no son estadísticamente significativos. Hay insuficiente evidencia para rechazar la hipótesis nula $H_0 : \sigma^2 = 4$.

El correspondiente intervalo de confianza de 90% es

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{(1-\alpha/2)}}$$

Los valores de $\chi^2_{(1-\alpha/2)}$ y $\chi^2_{\alpha/2}$ son

$$\chi^2_{(1-\alpha/2)} = \chi^2_{.95} = .102587$$

$$\chi^2_{\alpha/2} = \chi^2_{.05} = 5.99147$$

Sustituyendo estos valores en la fórmula para la estimación de intervalo, se obtiene

$$\frac{2(10.57)}{5.99147} < \sigma^2 < \frac{2(10.57)}{.102587} \quad \text{o} \quad 3.53 < \sigma^2 < 206.07$$

Así, se puede estimar la varianza poblacional para que caiga en el intervalo 3.53 a 206.07. Este intervalo de confianza muy ancho indica la poca información sobre la varianza poblacional que se obtiene de una muestra de sólo tres mediciones. En consecuencia, no es de sorprenderse que haya insuficiente evidencia para rechazar la hipótesis nula $\sigma^2 = 4$. Para obtener más información sobre σ^2 , el experimentador necesita aumentar el tamaño muestral.

Aunque *MS Excel* no tiene un comando único para implementar estos procedimientos, use la herramienta de función en *Excel* para encontrar usted mismo el estadístico de prueba, el valor p y los límites de confianza superior e inferior. Si usa *MINITAB*, el comando **Stat ► Basic Statistics ► 1 Variance** le permite introducir ya sea datos sin procesar o estadísticas de resumen para ejecutar la prueba ji cuadrada para una sola varianza, y calcular un intervalo de confianza. La parte pertinente de la salida impresa de *MINITAB 16* para el ejemplo 6.12 se muestra en la figura 6.15.

FIGURA 6.15

Salida de MINITAB para el ejemplo 6.12

Prueba y CI para una varianza: mediciones

Null hypothesis			Sigma-squared = 4		
Alternative hypothesis			Sigma-squared not = 4		
Statistics					
Variable	N	StDev	Variance		
Measurements	3	3.25	10.6		
90% Confidence Intervals					
CI for	CI for				
Variable	Method	StDev	Variance		
Measurements	Chi-Square	(1.88, 14.36)	(3.5, 206.1)		
Test					
Variable	Method	Statistic	DF	P-Value	
Measurements	Chi-Square	5.28	2	0.142	

6.6

EJERCICIOS

TÉCNICAS BÁSICAS

6.44 Una muestra aleatoria de $n = 25$ observaciones de una población normal produjo una varianza muestral igual a 21.4. ¿Estos datos dan suficiente evidencia para indicar que $\sigma^2 > 15$? Pruebe usando $\alpha = .05$.

6.45 Una muestra aleatoria de $n = 15$ observaciones fue seleccionada de una población normal. La media muestral y varianza fueron $\bar{x} = 3.91$ y $s^2 = .3214$. Encuentre un intervalo de confianza de 90% para la varianza poblacional σ^2 .

6.46 Una muestra aleatoria de tamaño $n = 7$ de una población normal produjo estas mediciones: 1.4, 3.6, 1.7, 2.0, 3.3, 2.8, 2.9.

- Calcule la varianza muestral, s^2 .
- Construya un intervalo de confianza de 95% para la varianza poblacional, σ^2 .
- Pruebe $H_0 : \sigma^2 = .8$ contra $H_a : \sigma^2 \neq .8$ usando $\alpha = .05$. Exprese sus conclusiones.
- ¿Cuál es el valor p aproximado para la prueba de la parte c?

APLICACIONES

6.47 Precisión de instrumentos Un instrumento de precisión está garantizado para leer con exactitud con variación de no más de 2 unidades. Una muestra de cuatro lecturas de instrumentos en el mismo objeto dio las mediciones 353, 351, 351 y 355.

- Pruebe la hipótesis nula de que $\sigma = .7$ contra la alternativa de $\sigma < .7$. Use $\alpha = .05$.
- Encuentre un intervalo de confianza del 90% para la varianza poblacional.

6.48 Potencia de medicamentos Para tratar a los pacientes en forma apropiada, los medicamentos recetados por los médicos no sólo deben tener un valor de potencia media según se especifica en el envase del medicamento,

sino que también la variación en los valores de potencia debe ser pequeña. De otro modo, los farmacéuticos distribuirían recetas que podrían ser peligrosamente potentes o tendrían una baja potencia y serían ineficaces. Un fabricante de medicinas declara que su medicina tiene una potencia de $5 \pm .1$ miligramos por centímetro cúbico (mg/cc). Una muestra aleatoria de cuatro envases dio lecturas de potencia iguales a 4.94, 5.09, 5.03 y 4.90 mg/cc.

- ¿Los datos presentan suficiente evidencia para indicar que la potencia media difiere de 5 mg/cc?
- ¿Los datos presentan suficiente evidencia para indicar que la variación en potencia difiere de los límites de error especificados por el fabricante? (SUGERENCIA: A veces es difícil determinar con exactitud lo que se entiende por límites de potencia especificados por un fabricante. Como éste implica que los valores de potencia caerán en los intervalos $5 \pm .1$ mg/cc con muy alta probabilidad —la implicación es casi *siempre*— supongamos que el rango .2; o 4.9 a 5.1, representa 6σ , como lo sugiere la Regla Empírica.)

6.49 Potencia de medicamentos, continúa Consulte el ejercicio 6.48. La prueba de 60 envases de medicamento adicionales seleccionados al azar dio una media muestral y varianza igual a 5.04 y .0063 (para el total de $n = 64$ envases). Usando un intervalo de confianza de 95%, estime la varianza de las mediciones de potencia hechas por el fabricante.

6.50 Cascos de seguridad Un fabricante de cascos de seguridad para trabajadores de la construcción, está preocupado por la media y la variación de las fuerzas que transmiten los cascos a los usuarios cuando son sometidos a una fuerza externa estándar. El fabricante desea que la fuerza media transmitida por cascos sea de 800 libras (o menos), muy por debajo del límite legal de 1 000 libras y que σ sea menor que 40. Se probó una muestra aleatoria

de $n = 40$ cascos y se encontró que la media y varianza muestrales eran iguales a 825 libras y 2 350 libras, respectivamente.

- Si $\mu = 800$ y $\sigma = 40$, ¿es probable que cualquier casco, sometido a la fuerza externa estándar, transmita una fuerza de más de 1 000 libras al usuario? Explique.
- ¿Los datos dan suficiente evidencia para indicar que, cuando los casos son sometidos a la fuerza externa estándar, la fuerza media transmitida por los cascos excede de 800 libras?

6.51 Cascos de seguridad, continúa Consulte el ejercicio 6.50. ¿Los datos dan suficiente evidencia para indicar que σ excede de 40?

DATOS
EX0652

6.52 Focos eléctricos A un fabricante de focos eléctricos industriales le gusta que éstos tengan una vida media que sea aceptable para sus clientes, y que su variación de duración de vida sea relativamente pequeña. Una muestra de 20 focos probados produjo las siguientes duraciones de vida útil (en horas):

2100 2302 1951 2067 2415 1883 2101 2146 2278 2019
1924 2183 2077 2392 2286 2501 1946 2161 2253 1827

El fabricante desea controlar la variabilidad en duración de vida útil para que σ sea menor que 150 horas. ¿Los datos dan suficiente evidencia para indicar que el fabricante está alcanzando su objetivo? Pruebe usando $\alpha = .01$.

COMPARACIÓN DE DOS VARIANZAS POBLACIONALES

6.7

Así como una sola varianza poblacional a veces es importante para un experimentador, quizá usted también necesite comparar dos varianzas poblacionales. Es probable que deba comparar la precisión de un aparato de medición con la de otro, la estabilidad de un proceso de manufactura con la de otro, o incluso la variabilidad en el procedimiento de calificaciones de un profesor universitario con el de otro.

Una forma de comparar dos varianzas poblacionales, σ_1^2 y σ_2^2 , es usar la razón entre las varianzas muestrales, s_1^2/s_2^2 . Si s_1^2/s_2^2 es casi igual a 1, se encuentra poca evidencia para indicar que σ_1^2 y σ_2^2 son desiguales. Por otra parte, un valor muy grande o muy pequeño para s_1^2/s_2^2 da evidencia de una diferencia en las varianzas poblacionales.

¿Qué tan grande debe ser s_1^2/s_2^2 para que haya evidencia suficiente para rechazar la siguiente hipótesis nula?

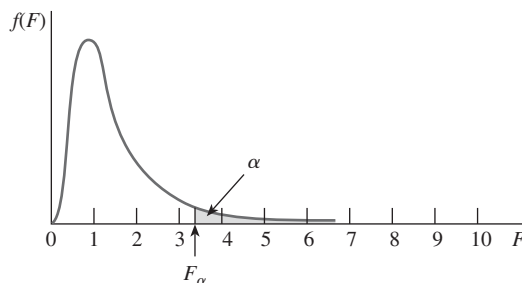
$$H_0 : \sigma_1^2 = \sigma_2^2$$

La respuesta a esta pregunta se halla al estudiar la distribución de s_1^2/s_2^2 en un muestreo repetido.

Cuando las muestras aleatorias independientes se sacan de dos poblaciones *normales* con *varianzas iguales*, es decir, $\sigma_1^2 = \sigma_2^2$; entonces s_1^2/s_2^2 tiene una distribución de probabilidad en un muestreo repetido que los estadísticos conocen como **distribución F**. La ecuación de la función de densidad para este estadístico es bastante complicada para revisarla aquí, pero traza la curva que se muestra en la figura 6.16.

FIGURA 6.16

Una distribución F con $g_l^1 = 10$ y $g_l^2 = 10$



SUPOSICIONES PARA QUE s_1^2/s_2^2 TENGA UNA DISTRIBUCIÓN F

- Las muestras aleatorias e independientes se sacan de cada una de dos poblaciones normales.
- La variabilidad de las mediciones en las dos poblaciones es igual y puede ser medida por una varianza común, σ^2 ; esto es, $\sigma_1^2 = \sigma_2^2 = \sigma^2$.



APPLET EN LÍNEA

Probabilidades F



MI CONSEJO

Probar dos varianzas:

$gl_1 = n_1 - 1$ y

$gl_2 = n_2 - 1$

No es importante que usted conozca la ecuación compleja de la función de densidad para F . Para nuestros fines, sólo se necesita usar los valores críticos bien tabulados de F dados en la tabla 6 en el Anexo.

Al igual que la distribución χ^2 , la forma de la distribución F es no simétrica y depende del número de grados de libertad asociados con s_1^2 y s_2^2 , representados como $gl_1 = (n_1 - 1)$ y $gl_2 = (n_2 - 1)$, respectivamente. Esto complica la tabulación de valores críticos de la distribución F porque se requiere una tabla para cada combinación diferente de gl_1 , gl_2 y a .

En la tabla 6 del Anexo, los valores críticos de F para áreas de cola derecha correspondientes a $a = .100, .050, .025, .010$ y $.005$ están tabulados para varias combinaciones de gl_1 grados de libertad del numerador y gl_2 grados de libertad del denominador. Una parte de la tabla 6 se reproduce en la tabla 6.6. El número de grados de libertad gl_1 del numerador aparece en el margen superior y el número de grados de libertad del denominador gl_2 aparece a lo largo del margen lateral. Los valores de a se encuentran en la segunda columna. Para una combinación fija de gl_1 y gl_2 , los valores críticos apropiados de F se encuentran en la línea indizada por el valor de a requerido.

EJEMPLO

6.13

Compruebe su capacidad para usar la tabla 6 del Anexo al verificar los siguientes enunciados:

1. El valor de F con área .05 a su derecha para $gl_1 = 6$ y $gl_2 = 9$ es 3.37.
2. El valor de F con área .05 a su derecha para $gl_1 = 5$ y $gl_2 = 10$ es 3.33.
3. El valor de F con área .01 a su derecha para $gl_1 = 6$ y $gl_2 = 9$ es 5.80.

Estos valores están sombreados en la tabla 6.6.

TABLA 6.6 Formato para la tabla F de la tabla 6 del Anexo

		gl_1					
gl_2	a	1	2	3	4	5	6
1	.100	39.86	49.50	53.59	55.83	57.24	58.20
	.050	161.4	199.5	215.7	224.6	230.2	234.0
	.025	647.8	799.5	864.2	899.6	921.8	937.1
	.010	4052	4999.5	5403	5625	5764	5859
	.005	16211	20000	21615	22500	23056	23437
2	.100	8.53	9.00	9.16	9.24	9.29	9.33
	.050	18.51	19.00	19.16	19.25	19.30	19.33
	.025	38.51	39.00	39.17	39.25	39.30	39.33
	.010	98.50	99.00	99.17	99.25	99.30	99.33
	.005	198.5	199.0	199.2	199.2	199.3	199.3
3	.100	5.54	5.46	5.39	5.34	5.31	5.28
	.050	10.13	9.55	9.28	9.12	9.01	8.94
	.025	17.44	16.04	15.44	15.10	14.88	14.73
	.010	34.12	30.82	29.46	28.71	28.24	27.91
	.005	55.55	49.80	47.47	46.19	45.39	44.84

9	.100	3.36	3.01	2.81	2.69	2.61	2.55
	.050	5.12	4.26	3.86	3.63	3.48	3.37
	.025	7.21	5.71	5.08	4.72	4.48	4.32
	.010	10.56	8.02	6.99	6.42	6.06	5.80
	.005	13.61	10.11	8.72	7.96	7.47	7.13
10	.100	3.29	2.92	2.73	2.61	2.52	2.46
	.050	4.96	4.10	3.71	3.48	3.33	3.22
	.025	6.94	5.46	4.83	4.47	4.24	4.07
	.010	10.04	7.56	6.55	5.99	5.64	5.39
	.005	12.83	9.43	8.08	7.34	6.87	6.54

La prueba estadística de la hipótesis nula

$$H_0 : \sigma_1^2 = \sigma_2^2$$

utiliza el estadístico de prueba

$$F = \frac{s_1^2}{s_2^2}$$

Cuando la hipótesis alternativa implica una prueba de una cola, esto es,

$$H_a : \sigma_1^2 > \sigma_2^2$$

se puede hallar el valor crítico de cola derecha para rechazar H_0 directamente de la tabla 6 del Anexo. Sin embargo, cuando la hipótesis alternativa requiera una prueba de dos colas, es decir,

$$H_a : \sigma_1^2 \neq \sigma_2^2$$

la región de rechazo se divide entre las colas superior e inferior de la distribución F . Estos valores críticos de cola izquierda *no se dan* en la tabla 6 por la siguiente razón: usted es libre de decidir a cuál de las dos poblaciones llamar “Población 1”. Si siempre elige llamar “Población 1” a la población con la varianza muestral *más grande*, entonces el valor observado de su estadístico de prueba siempre estará en la cola derecha de la distribución F . Aun cuando la mitad de la región de rechazo, el área $\alpha/2$ a su izquierda, estará en la cola inferior de la distribución, nunca será necesario usarla. Pero recuerde estos puntos, para una prueba de dos colas:

- El área de la cola derecha de la región de rechazo es sólo $\alpha/2$.
- El área a la derecha del estadístico de prueba observado es sólo (valor p)/2.

Los procedimientos formales para una prueba de hipótesis y un intervalo de confianza $(1 - \alpha)100\%$ para dos varianzas poblacionales se muestran a continuación.

PRUEBA DE HIPÓTESIS RESPECTO A LA IGUALDAD DE DOS VARIANZAS POBLACIONALES

1. Hipótesis nula: $H_0 : \sigma_1^2 = \sigma_2^2$

2. Hipótesis alternativa:

Prueba de una cola

$$H_a : \sigma_1^2 > \sigma_2^2$$

(o $H_a : \sigma_1^2 < \sigma_2^2$)

Prueba de dos colas

$$H_a : \sigma_1^2 \neq \sigma_2^2$$

(continúa)

PRUEBA DE HIPÓTESIS RESPECTO A LA IGUALDAD DE DOS VARIANZAS POBLACIONALES

(continuación)

3. Estadístico de prueba:

Prueba de una cola

$$F = \frac{s_1^2}{s_2^2}$$

Prueba de dos colas

$$F = \frac{s_1^2}{s_2^2}$$

donde s_1^2 es la varianza muestral más grande.

4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola

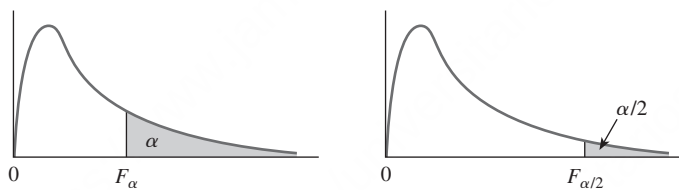
$$F > F_\alpha$$

o cuando valor $p < \alpha$

Prueba de dos colas

$$F > F_{\alpha/2}$$

Los valores críticos de F_α y $F_{\alpha/2}$ están basados en $gl_1 = (n_1 - 1)$ y $gl_2 = (n_2 - 1)$. Estos valores tabulados, para $\alpha = .100, .050, .025, .010$ y $.005$, se pueden hallar usando la tabla 6 del Anexo.



Suposiciones: Las muestras se seleccionan al azar y en forma independiente de las poblaciones normalmente distribuidas.

INTERVALO DE CONFIANZA PARA σ_1^2/σ_2^2

$$\left(\frac{s_1^2}{s_2^2}\right) \frac{1}{F_{gl_1, gl_2}} < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{s_1^2}{s_2^2}\right) F_{gl_2, gl_1}$$

donde $gl_1 = (n_1 - 1)$ y $gl_2 = (n_2 - 1)$. F_{gl_1, gl_2} es el valor crítico tabulado de F correspondiente a gl_1 y gl_2 grados de libertad en el numerador y denominador de F , respectivamente, con área $\alpha/2$ a su derecha.

Suposiciones: Las muestras se seleccionan al azar y en forma independiente de poblaciones normalmente distribuidas.

EJEMPLO

6.14

Un experimentador está preocupado porque la variabilidad de respuestas que usan dos procedimientos experimentales diferentes puede no ser igual. Antes de realizar su investigación, lleva a cabo un estudio previo con muestras aleatorias de 10 y 8 respuestas y obtiene $s_1^2 = 7.14$ y $s_2^2 = 3.21$, respectivamente. ¿Las varianzas muestrales presentan suficiente evidencia para indicar que las varianzas poblacionales son desiguales?

Solución Suponga que las poblaciones tienen distribuciones de probabilidad que son razonablemente en forma de montículo y por tanto satisfacen, para todos los fines prácticos, la suposición de que las poblaciones son normales. Se desea probar estas hipótesis:

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{contra} \quad H_a : \sigma_1^2 \neq \sigma_2^2$$

Usando la tabla 6 del Anexo para $\alpha/2 = .025$, se puede rechazar H_0 cuando $F > 4.82$ con $\alpha = .05$. El valor calculado del estadístico de prueba es

$$F = \frac{s_1^2}{s_2^2} = \frac{7.14}{3.21} = 2.22$$

Debido a que el estadístico de prueba no cae en la región de rechazo, no se puede rechazar $H_0 : \sigma_1^2 = \sigma_2^2$. Por tanto, hay insuficiente evidencia para indicar una diferencia en las varianzas poblacionales.

EJEMPLO

6.15

Consulte el ejemplo 6.14 y encuentre un intervalo de confianza de 90% para σ_1^2/σ_2^2 .

Solución El intervalo de confianza de 90% para σ_1^2/σ_2^2 es

$$\left(\frac{s_1^2}{s_2^2}\right) \frac{1}{F_{gl_1, gl_2}} < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{s_1^2}{s_2^2}\right) F_{gl_2, gl_1}$$

donde

$$\begin{aligned} s_1^2 &= 7.14 & s_2^2 &= 3.21 \\ gl_1 &= (n_1 - 1) = 9 & gl_2 &= (n_2 - 1) = 7 \\ F_{9,7} &= 3.68 & F_{7,9} &= 3.29 \end{aligned}$$

Sustituyendo estos valores en la fórmula para el intervalo de confianza, se obtiene

$$\left(\frac{7.14}{3.21}\right) \frac{1}{3.68} < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{7.14}{3.21}\right) 3.29 \quad \text{o} \quad .60 < \frac{\sigma_1^2}{\sigma_2^2} < 7.32$$

La estimación calculada del intervalo de .60 a 7.32 incluye 1.0, el valor hipotético en H_0 . Esto indica que es muy posible que $\sigma_1^2 = \sigma_2^2$ y por tanto concuerda con las conclusiones de prueba. No rechace $H_0 : \sigma_1^2 = \sigma_2^2$.

La función de *Excel* llamada **F.TEST** (**FTEST** en *Excel 2007* y versiones anteriores) puede usarse para realizar la prueba F para la igualdad de varianzas cuando ha introducido los datos sin procesar en la hoja de cálculo. El comando de *MINITAB Stat* ► **Basic Statistics** ► **2 Variances** es un poco más flexible, porque permite introducir ya sea datos sin procesar o estadísticos de resumen para ejecutar la prueba F . Además, *MINITAB 16* calcula intervalos de confianza para la razón de dos varianzas o dos desviaciones estándar (lo cual todavía no se ha expuesto). La salida impresa de *MINITAB 16* para el ejemplo 6.14, que contiene el estadístico F y su valor p , se muestra en la figura 6.17.

FIGURA 6.17

Salida de MINITAB para el ejemplo 6.14

Prueba y CI para dos varianzas

Null hypothesis Sigma(1) / Sigma(2) = 1
 Alternative hypothesis Sigma(1) / Sigma(2) not = 1
 Significance level Alpha = 0.05

Statistics

Sample	N	StDev	Variance
1	10	2.672	7.140
2	8	1.792	3.210

Ratio of standard deviations = 1.491

Ratio of variances = 2.224

95% Confidence Intervals

CI for

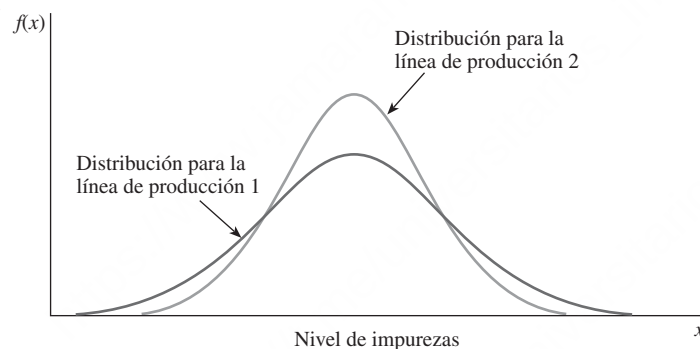
Distribution of Data	CI for StDev Ratio	Variance Ratio
Normal	(0.679, 3.055)	(0.461, 9.335)

Test

Method	DF1	DF2	Statistic	P-Value
F Test (normal)	9	7	2.22	0.304

FIGURA 6.18

Distribuciones de mediciones de impurezas para dos líneas de producción

**EJEMPLO**

6.16

La variabilidad en la cantidad de impurezas, presente en un lote de un producto químico empleado para un proceso particular, depende del tiempo en que el proceso está en operación. Un fabricante que utiliza dos líneas de producción, 1 y 2, ha hecho un ligero ajuste a la línea 2, esperando con ello reducir la variabilidad así como el promedio de la cantidad de impurezas en el producto químico. Muestras de $n_1 = 25$ y $n_2 = 25$ mediciones de los dos lotes dan estas medias y varianzas:

$$\bar{x}_1 = 3.2 \quad s_1^2 = 1.04$$

$$\bar{x}_2 = 3.0 \quad s_2^2 = .51$$

¿Los datos presentan suficiente evidencia para indicar que la variabilidad del proceso es menor para la línea 2?

Solución El experimentador piensa que el promedio de los niveles de impurezas son los mismos para las dos líneas de producción pero que su ajuste puede haber disminuido la variabilidad de los niveles para la línea 2, como se ilustra en la figura 6.18. Este ajuste sería bueno para la compañía porque disminuiría la probabilidad de hacer envíos del producto químico con niveles de impureza inaceptablemente altos.

Para determinar la disminución en variabilidad, la prueba de hipótesis es

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{contra} \quad H_a : \sigma_1^2 > \sigma_2^2$$

y el valor observado para este estadístico de prueba es

$$F = \frac{s_1^2}{s_2^2} = \frac{1.04}{.51} = 2.04$$

Usando el método del valor p , se puede limitar el valor p de una cola usando la tabla 6 del Anexo con $gl_1 = gl_2 = (25 - 1) = 24$. El valor observado de F cae entre $F_{.050} = 1.98$ y $F_{.025} = 2.27$, de modo que $.025 < \text{valor } p < .05$. Los resultados son juzgados como significativos al nivel del 5%, y H_0 es rechazada. Se concluye que la variabilidad de la línea 2 es menor que la de la línea 1.

La prueba F de la diferencia en dos varianzas poblacionales completa la batería de pruebas que usted ha aprendido en este capítulo para hacer inferencias acerca de parámetros poblacionales bajo estas condiciones:

- Los tamaños muestrales son pequeños.
- La muestra o muestras son sacados de poblaciones normales.

Usted encontrará que las distribuciones F y χ^2 , así como la distribución t de Student, son muy importantes en otras aplicaciones en los capítulos que siguen. Se usarán para diferentes estimadores diseñados para contestar distintos tipos de preguntas inferenciales, pero las técnicas básicas para hacer inferencias siguen siendo iguales.

En la siguiente sección, repasamos las suposiciones requeridas para todas estas herramientas de inferencia y analizamos las opciones que existen cuando las suposiciones no parecen ser razonablemente correctas.

6.7 EJERCICIOS

TÉCNICAS BÁSICAS

6.53 Muestras aleatorias independientes de dos poblaciones normales produjeron las varianzas siguientes:

Tamaño muestral	Varianza muestral
16	55.7
20	31.4

- a. ¿Los datos dan suficiente evidencia para indicar que σ_1^2 difiere de σ_2^2 ? Pruebe usando $\alpha = .05$.
- b. Encuentre el valor p aproximado para la prueba e interprete su valor.

6.54 Consulte el ejercicio 6.53 y encuentre un intervalo de confianza de 95% para σ_1^2/σ_2^2 .

6.55 Muestras aleatorias independientes de dos poblaciones normales produjeron las varianzas dadas:

Tamaño muestral	Varianza muestral
13	18.3
13	7.9

- a. ¿Los datos dan suficiente evidencia para indicar que $\sigma_1^2 > \sigma_2^2$? Pruebe usando $\alpha = .05$.

- b. Encuentre el valor p aproximado para la prueba e interprete su valor.

APLICACIONES

6.56 Calificaciones del SAT Resultados de exámenes del SAT en química y física⁹ para dos grupos de 15 estudiantes, habiendo cada uno de éstos seleccionado estos exámenes, se dan a continuación:

Química	Física
$\bar{x} = 644$	$\bar{x} = 658$
$s = 114$	$s = 103$
$n = 15$	$n = 15$

Para usar la prueba t de dos muestras con una estimación agrupada de σ^2 , se debe suponer que las dos varianzas poblacionales son iguales. Pruebe esta suposición usando la prueba F de igualdad de varianzas. ¿Cuál es el valor p aproximado para la prueba?

6.57 Baterías de litio La estabilidad de las mediciones en un producto manufacturado es importante para

mantener su calidad. Un fabricante de baterías de litio, como las que se usan para cámaras digitales, sospecha que una de las líneas de producción producía baterías con una duración que variaba de manera amplia. Para probar esta teoría, seleccionó al azar $n = 50$ baterías de la línea sospechosa y $n = 50$ de una línea que se evaluó como “en control”. Luego midió la duración (en horas) hasta su reducción a 0.85 V con una carga de 5 ohmios para ambas muestras. Las medias y varianzas muestrales para las dos muestras fueron las siguientes:

Línea sospechosa	Línea “en control”
$\bar{x}_1 = 9.40$	$\bar{x}_2 = 9.25$
$s_1 = .25$	$s_2 = .12$

- ¿Los datos proporcionan evidencia suficiente para indicar que las baterías producidas por la “línea sospechosa” tienen una varianza más grande en la duración que aquellas producidas por la línea que se supone está en control? Pruebe usando $\alpha = .05$.
- Encuentre el valor p aproximado para la prueba e interprete su valor.
- Construya un intervalo de confianza del 90% para la razón de la varianza.

DATOS
EX0658

6.58 Roethlisberger y Rodgers Los mariscales de campo no sólo necesitan tener un buen porcentaje de pases, sino que además, deben ser consistentes. Es decir, la variabilidad en el número de pases completados por partido debería ser pequeña. La tabla que sigue da el número de pases completados para Ben Roethlisberger y Aaron Rodgers, mariscales de campo de los Acereros de Pittsburgh y los Empacadores de Green Bay, respectivamente, durante la temporada 2010 de la NFL.¹⁰

Aaron Rodgers			Ben Roethlisberger	
19	21	7	16	20
19	15	25	19	22
34	27	19	17	21
12	22		17	23
27	26		30	22
18	21		18	15

- ¿Los datos indican que hay una diferencia en la variabilidad en el número de pases completados para los dos mariscales de campo? Use $\alpha = .01$.
- Si usted fuera a probar una diferencia en las dos medias poblacionales, ¿sería apropiado usar la prueba t de dos muestras que supone varianzas iguales? Explique.

6.59 Atún III En el ejercicio 6.25 y el conjunto de datos EX0625, se realizó una prueba para detectar una diferencia en el promedio de precios de atún claro en agua contra atún claro en aceite.

- ¿Qué suposición tuvo que hacerse respecto a las varianzas poblacionales para que la prueba fuera válida?

- ¿Los datos presentan suficiente evidencia para indicar que las varianzas violan la suposición en la parte a? Pruebe usando $\alpha = .05$.

6.60 Corredores y ciclistas III Consulte el ejercicio 6.26. Un experimento que comprendía 10 corredores sanos y 10 ciclistas sanos, para determinar si hay diferencias significativas en mediciones de presión dentro del compartimiento del músculo anterior para corredores y ciclistas.⁶ Los datos, es decir, presión del compartimiento, en milímetros de mercurio (Hg), son como sigue:

Afección	Corredores		Ciclistas	
	Media	Desviación estándar	Media	Desviación estándar
En reposo	14.5	3.92	11.1	3.98
80% máximo consumo de O_2	12.2	3.49	11.5	4.95
Máximo consumo de O_2	19.1	16.9	12.2	4.47

Para cada una de las tres variables medidas en este experimento, pruebe para ver si hay una diferencia significativa en las varianzas para corredores contra ciclistas. Encuentre los valores p aproximados para cada una de estas pruebas. ¿Será apropiada una prueba t de dos muestras con una estimación agrupada de σ^2 para estas tres variables? Explique.

6.61 Impurezas Un fabricante farmacéutico compra un material particular a dos proveedores diferentes. El nivel medio de impurezas en la materia prima es aproximadamente igual para ambos proveedores, pero el fabricante está preocupado por la variabilidad de las impurezas de un embarque a otro. Para comparar la variación en el porcentaje de impurezas para los dos proveedores, el fabricante selecciona 10 envíos de cada uno de ellos y mide el porcentaje de impurezas en la materia prima para cada embarque. Las medias y varianzas muestrales se presentan en la tabla.

Proveedor A	Proveedor B
$\bar{x}_1 = 1.89$	$\bar{x}_2 = 1.85$
$s_1^2 = .273$	$s_2^2 = .094$
$n_1 = 10$	$n_2 = 10$

- ¿Los datos dan suficiente evidencia para indicar una diferencia en la variabilidad de los niveles de impureza de envíos para los dos proveedores? Pruebe usando $\alpha = .01$. Con base en los resultados de su prueba, ¿qué recomendación se haría al fabricante de productos farmacéuticos?
- Encuentre un intervalo de confianza de 99% para σ_2^2 e interprete sus resultados.



NECESITO SABER...

Cómo decidir cuál prueba usar

¿Está usted interesado en probar medias? Si el diseño comprende:

- a. Una muestra aleatoria, use el estadístico t de una muestra.
- b. Dos muestras aleatorias independientes, ¿son iguales las varianzas poblacionales?
 - i. Si son iguales, use el estadístico t de dos muestras con s^2 agrupada.
 - ii. Si son desiguales, use la t sin agrupar con gl estimado.
- c. Dos muestras en pares con pares aleatorios, use una t de una muestra para analizar diferencias.

¿Está usted interesado en probar varianzas? Si el diseño comprende:

- a. Una muestra aleatoria, use la prueba χ^2 para una sola varianza.
- b. Dos muestras aleatorias independientes, use la prueba F para comparar dos varianzas.

6.8

REPASO DE SUPOSICIONES DE MUESTRA PEQUEÑA

Todos los procedimientos de prueba y estimación estudiados en este capítulo requieren que los datos satisfagan ciertas condiciones para que las probabilidades de error (para las pruebas) y los coeficientes de confianza (para los intervalos de confianza), sean iguales a los valores que se hayan especificado. Por ejemplo, si usted construye lo que piensa que es un intervalo de confianza de 95%, puede estar seguro que, en muestreo repetido, 95% (y no 85% o 75% o menos) de todos estos intervalos contendrán el parámetro de interés. Estas condiciones se resumen en estas suposiciones.

SUPOSICIONES

1. Para todas las pruebas e intervalos de confianza descritos en este capítulo, se asume que **las muestras se seleccionan al azar de poblaciones normalmente distribuidas**.
2. Cuando se seleccionan dos muestras, se supone que se **seleccionan en forma independiente** excepto en el caso del experimento de diferencia en pares.
3. Para pruebas o intervalos de confianza respecto a la diferencia entre dos medias poblacionales μ_1 y μ_2 con base en muestras aleatorias independientes, se supone que $\sigma_1^2 = \sigma_2^2$.

En realidad, usted nunca sabrá todo acerca de la población muestreada. Si lo supiera no habría necesidad de muestreo o de estadísticos. También es improbable que una población satisfaga *exactamente* las suposiciones dadas en el recuadro. Por fortuna, los procedimientos presentados en este capítulo proporcionan buenas inferencias aun cuando los datos exhiben desviaciones moderadas desde las condiciones necesarias.

Un procedimiento estadístico que no es sensible a desviaciones desde las condiciones en las que está basado se dice que es **robusto**. Las pruebas t de Student son muy robustas para desviaciones moderadas desde normalidad. También, mientras los tamaños muestrales sean casi iguales, no hay mucha diferencia entre los estadísticos t agrupados y no agrupados para la diferencia en dos medias poblacionales. Sin embargo, si los tamaños muestrales no son claramente iguales y si las varianzas poblacionales son desiguales, la estadística t agrupada da conclusiones imprecisas.

Si usted está preocupado porque sus datos no satisfacen las suposiciones, hay otras opciones:

- Si selecciona muestras relativamente grandes, use uno de los procedimientos de muestra grande de los capítulos 4 y 5, que no se apoyan en la normalidad o en suposiciones de varianzas iguales.
- Puede usar una *prueba no paramétrica* para contestar sus preguntas inferenciales. Estas pruebas han sido desarrolladas específicamente para que pocas o ninguna suposición distribucional se requieran para su uso. Las pruebas que se usan para comparar las ubicaciones o la variabilidad de dos poblaciones se presentan en el capítulo 7.

REPASO DEL CAPÍTULO

Conceptos y fórmulas clave

I. Diseños experimentales para muestras pequeñas

1. **Una muestra aleatoria:** La población muestreada debe ser normal
2. **Dos muestras aleatorias independientes:** Ambas poblaciones muestreadas deben ser normales
 - a. Las poblaciones tienen una varianza común σ^2
 - b. Las poblaciones tienen varianzas diferentes: σ_1^2 y σ_2^2
3. **Diseño de una diferencia en pares o pares acoplados:** Las muestras no son independientes

II. Pruebas estadísticas de significancia

1. Con base en las distribuciones t , F y χ^2
2. **Región de rechazo —valores críticos y niveles de significancia:** con base en las distribuciones t , F o χ^2 con grados de libertad apropiados
3. **Pruebas de parámetros poblacionales:** una sola media, la diferencia entre dos medias, una sola varianza y la razón entre dos varianzas

III. Estadístico de prueba de muestra pequeña

Para probar uno de los parámetros poblacionales cuando los tamaños muestrales sean pequeños, use la siguiente estadística de prueba:

Parámetro	Prueba estadística	Grados de libertad
μ	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$	$n - 1$
$\mu_1 - \mu_2$ (Varianzas iguales)	$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	$n_1 + n_2 - 2$
$\mu_1 - \mu_2$ (Varianzas diferentes)	$t \approx \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	Aproximación de Satterthwaite
$\mu_1 - \mu_2$ (Muestras en pares)	$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}}$	$n - 1$
σ^2	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$n - 1$
σ_1^2/σ_2^2	$F = s_1^2/s_2^2$	$n_1 - 1$ y $n_2 - 1$



Prueba y estimación de muestra pequeña: *Microsoft Excel*

Las pruebas de hipótesis para dos medias poblacionales basadas en la distribución t de Student y la prueba F para la razón de dos varianzas pueden encontrarse usando el comando de *Microsoft Excel* **Data ► Data Analysis**. Recuerde que necesita haber cargado los complementos de *Excel* llamados **Analysis ToolPak**. Encontrará tres opciones para las pruebas t de dos muestras y una prueba F en la lista de “Analysis Tools”. Para elegir las pruebas t apropiadas, primero debe decidir si las muestras son independientes o pareadas; para la prueba de muestras independientes, debe decidir si es posible suponer que las varianzas poblacionales son iguales o no.

EJEMPLO

6.17

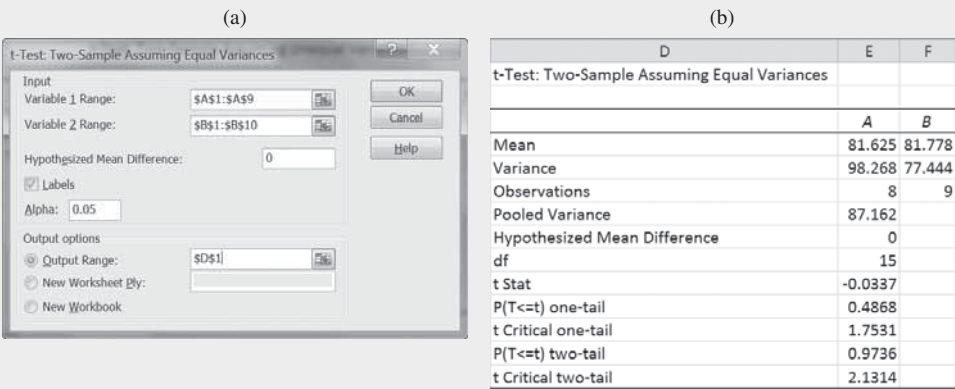
(Prueba t de dos muestras suponiendo varianzas iguales) Se registraron las puntuaciones de examen en la misma prueba de álgebra para nueve estudiantes seleccionados al azar de una clase instruida por el profesor A y ocho estudiantes seleccionados al azar de una clase enseñada por el profesor B. ¿Hay una diferencia en el promedio de las puntuaciones para estudiantes enseñados por estos dos profesores?

Profesor A	65	88	93	95	80	76	79	77	
Profesor B	91	85	70	82	92	68	86	87	75

Introduzca los datos en las columnas A y B de una hoja de cálculo de *Excel*.

1. Use **Data ► Data Analysis ► Descriptive Statistics** o  **Statistical ► STDEV.S (STDEV en Excel 2007 y versiones anteriores)** para hallar las desviaciones estándar para las dos muestras, $s_1 = 9.913$ y $s_2 = 8.800$. En vista que la razón de las dos varianzas es $s_1^2/s_2^2 = 1.27$ (menos de 3), se tiene seguridad al suponer que las varianzas poblacionales son iguales.
2. Seleccione **Data ► Data Analysis ► t-Test: Two-Sample Assuming Equal Variances** para generar el cuadro de diálogo que se muestra en la figura 6.19a). Resalte o teclee **Variable 1 Range** y **Variable 2 Range** (los datos en la primera y segunda columnas) en los primeros dos cuadros. En el cuadro marcado “Hypothesized Mean Difference” teclee **0** (en vista que se está probando $H_0: \mu_1 - \mu_2 = 0$) y seleccione “Labels” si es necesario.
3. En *Excel* el nivel de significancia predeterminado es $\alpha = .05$. Cambie este nivel de significancia si es necesario. Introduzca una ubicación de celda para el **Output Range** y haga clic en **OK**. La salida aparecerá en la ubicación de la celda seleccionada y deberá ajustarse usando **Format ► AutoFit Column Width** en la pestaña **Home** en el grupo **Cells** mientras se encuentra resaltada todavía. Puede disminuir la precisión decimal si lo desea, usando  en la pestaña **Home** en el grupo **Number**.
4. El valor observado del estadístico de prueba $t = -.0337$ se encuentra en la figura 6.19b) en la fila etiquetada “ t Stat” seguido por el valor p de una cola “ $P(T \leq t)$ one-tail” y el valor crítico marcando la región de rechazo para una prueba de una cola con $\alpha = .05$. Las últimas dos filas de la salida proporcionan el valor p y el valor t crítico para una prueba de dos colas.
5. Para este ejemplo, el valor $p = .9736$ indica que no hay una diferencia significativa en el promedio de las puntuaciones para los estudiantes enseñados por los dos profesores.

FIGURA 6.19



6. En la sección 6.7 se presentó una prueba de hipótesis formal para la igualdad de las dos varianzas usando la prueba F . Para efectuar esta prueba usando *Excel*, seleccione **Data ► Data Analysis ► F-Test: Two-Sample for Variances**. Siga las instrucciones para la prueba t de varianzas iguales, pero reemplace el valor “Alpha” con 0.025, y generará la salida que se muestra en la figura 6.20.

FIGURA 6.20

F-Test Two-Sample for Variances		
	A	B
Mean	81.625	81.78
Variance	98.268	77.44
Observations	8	9
df	7	8
F	1.2689	
P(F<=f) one-tail	0.3701	
F Critical one-tail	4.5286	

Note que en la salida sólo se proporcionan el valor p de una cola y el valor crítico, razón por la cual se especifica que la cola única sea 0.025. Por tanto, para nuestra prueba de dos colas, $\alpha = 0.05$ y el valor $p = .7402$. No hay una diferencia significativa en las dos varianzas.

EJEMPLO

6.18

(Prueba t de dos muestras suponiendo varianzas desiguales)

1. Consulte el ejemplo 6.17. Si la razón de las dos varianzas muestrales ha sido tan grande que no podría suponer varianzas iguales (se usa “mayor a 3” como regla general), debería seleccionar **Data ► Data Analysis ► t-Test: Two-Sample Assuming Unequal Variances**.
2. Siga las instrucciones para la prueba t de varianzas iguales, y generará una salida similar. Si se usa esta prueba en los datos del ejemplo 6.17, resulta la siguiente salida (figura 6.21).

FIGURA 6.21

H	I	J
t-Test: Two-Sample Assuming Unequal Variances		
	A	B
Mean	81.625	81.778
Variance	98.268	77.444
Observations	8	9
Hypothesized Mean Difference	0	
df	14	
t Stat	-0.0334	
P(T<=t) one-tail	0.4869	
t Critical one-tail	1.7613	
P(T<=t) two-tail	0.9738	
t Critical two-tail	2.1448	

- Verá ligeras diferencias en el valor observado del estadístico de prueba, los grados de libertad y los valores p para la prueba, pero las conclusiones no cambian.
- NOTA: Cuando calcula los grados de libertad para la *Aproximación de Satterthwaite*, la **Data Analysis Tool** en *Excel* redondea al entero más cercano. Una función alternativa de *Excel* para calcular el valor p para esta prueba ( ► **Statistical** ► **T.TEST**) usa el valor exacto de gl dados por la fórmula de Satterthwaite. Debido a estos enfoques diferentes para determinar los grados de libertad, los resultados de **T.TEST** y la herramienta t -Test diferirán ligeramente en el caso de varianzas desiguales, y también diferirán ligeramente de la salida de *MINITAB*.

EJEMPLO

6.19

(Prueba t en pares) Consulte los datos del desgaste de las llantas de la tabla 6.3.

- Para ejecutar una prueba de diferencia en pares para estas muestras dependientes, introduzca los datos en las primeras dos columnas de una hoja de cálculo de *Excel* y seleccione **Data** ► **Data Analysis** ► **t-Test: Paired Two Sample for Means**.
- Siga las instrucciones para la prueba t de varianzas iguales, y generará una salida similar. Para los datos de la tabla 6.3, obtendrá la salida que se muestra en la figura 6.22. Una vez más, puede disminuir la precisión decimal si lo desea, usando  en la pestaña **Home** en el grupo **Number**.
- Usando el valor observado del estadístico de prueba ($t = 12.83$) con valor p de dos colas = .0002, hay una evidencia fuerte que indica una diferencia en las dos medias poblacionales.

FIGURA 6.22

D	E	F
t-Test: Paired Two Sample for Means		
	<i>Tire A</i>	<i>Tire B</i>
Mean	10.24	9.76
Variance	1.733	1.763
Observations	5	5
Pearson Correlation	0.998	
Hypothesized Mean Difference	0	
df	4	
t Stat	12.8285	
P(T<=t) one-tail	0.0001	
t Critical one-tail	2.1318	
P(T<=t) two-tail	0.0002	
t Critical two-tail	2.7764	

Prueba y estimación de muestra pequeña: *MINITAB*

Las pruebas de hipótesis para dos medias poblacionales basadas en la distribución t de Student y la prueba F para la razón de dos varianzas pueden encontrarse usando el comando de *MINITAB* **Stat** ► **Basic Statistics**. Encontrará opciones para **1-Sample t**, **2-Sample t**, **Paired t** y **2 Variances**, las cuales ejecutarán las pruebas y procedimientos de estimación de las secciones 6.3, 6.4, 6.5 y 6.7. Para elegir las pruebas t de dos muestras apropiadas, primero debe decidir si las muestras son independientes o en pares; para la prueba de muestras independientes, debe decidir si es posible suponer que las varianzas poblacionales son iguales o no.

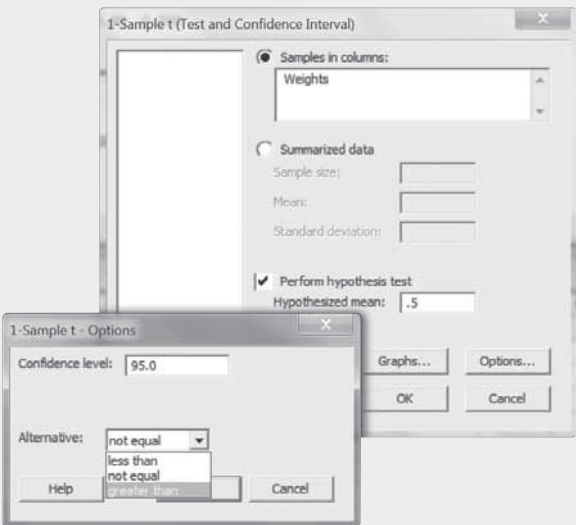
EJEMPLO

6.20

(Prueba *t* de una muestra) Consulte el ejemplo 6.3, en el cual se comparó el promedio del peso de diamantes usando un proceso nuevo con el promedio de un peso de .5 quilates.

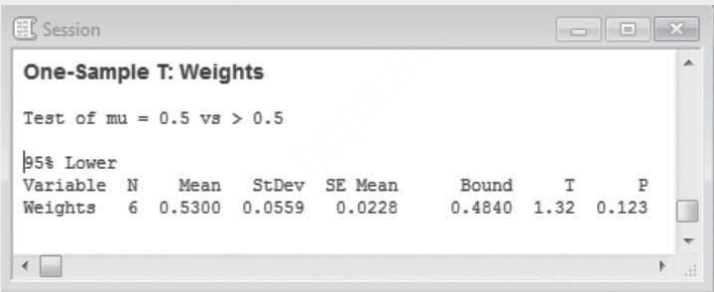
1. Introduzca los seis pesos registrados: .46, .61, .52, .48, .57, .54, en la columna C1 y nómbrala “Pesos”. Use **Stat ► Basic Statistics ► 1-Sample *t*** para generar los cuadros de diálogo que se muestran en la figura 6.23.

FIGURA 6.23



2. Para probar $H_0: \mu = .5$ contra $H_a: \mu > .5$, use la lista de la izquierda para seleccionar “Pesos” para el cuadro marcado “Samples in Columns”. Seleccione el cuadro marcado “Perform hypothesis test”. Luego, coloque el cursor en el cuadro marcado “Hypothesized mean:” e introduzca .5 como el valor de prueba. Por último, use **Options** y el menú desplegable marcado “Alternative” para seleccionar “greater than”. Puede cambiar el coeficiente de confianza por omisión de .95 si lo desea. Haga clic en **OK** dos veces para obtener la salida que se muestra en la figura 6.24.

FIGURA 6.24



3. Note que *MINITAB* produce un intervalo de confianza de una o dos colas para la media poblacional única, consistente con la hipótesis alternativa que haya elegido.

EJEMPLO

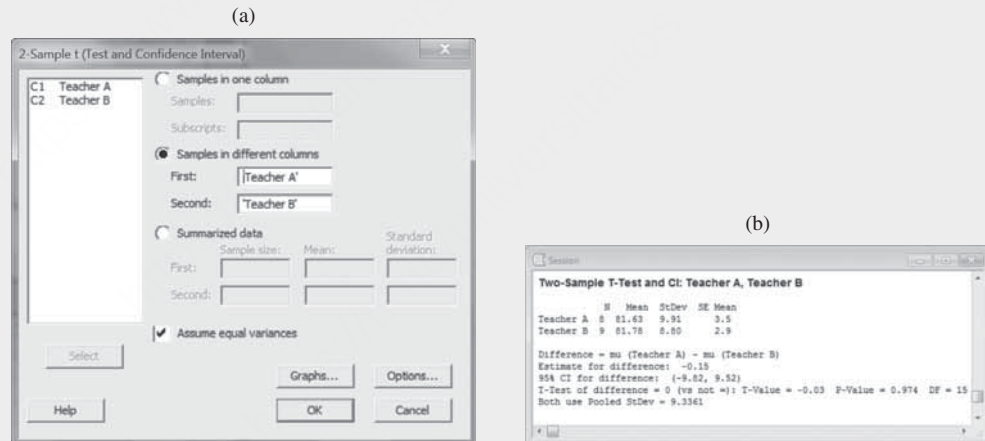
6.21

(Prueba *t* de dos muestras) Se registraron las puntuaciones de examen en la misma prueba de álgebra para ocho estudiantes seleccionados al azar de un salón de clase enseñado por el profesor A y nueve estudiantes seleccionados al azar de un salón de clase enseñado por el profesor B. ¿Hay una diferencia en el promedio de las puntuaciones para estudiantes enseñados por estos dos profesores?

Profesor A	65	88	93	95	80	76	79	77	
Profesor B	91	85	70	82	92	68	86	87	75

1. Los datos pueden introducirse en la hoja de cálculo en una de tres formas:
 - Introduzca las mediciones de ambas muestras en una sola columna e introduzca las letras (A o B) en una segunda columna para identificar la muestra de la cual proviene la medición.
 - Introduzca las muestras en dos columnas separadas.
 - Si no tiene los datos sin procesar, pero en su lugar tiene estadísticas de resumen, *MINITAB* le permitirá usar estos valores al seleccionar “Summarized data” e introducir los valores apropiados en los cuadros.
2. Utilice el segundo método e introduzca los datos en dos columnas de la hoja de cálculo. Use **Stat ► Basic Statistics ► Display Descriptive Statistics** para hallar las desviaciones estándar para las dos muestras, $s_1 = 9.91$ y $s_2 = 8.80$. En vista que la razón de las dos varianzas es $s_1^2/s_2^2 = 1.27$ (menor que 3), se tiene seguridad al suponer que las varianzas poblacionales son iguales.
3. Seleccione **Stat ► Basic Statistics ► 2-Sample t** para generar el cuadro de diálogo que se muestra en la figura 6.25a). Seleccione “samples in different columns”, seleccionando las columnas apropiadas de la lista de la izquierda. Elija el cuadro “Assume equal variances” y seleccione la alternativa apropiada en el cuadro **Options**. La salida de dos muestras aparece cuando haga clic dos veces en **OK** y automáticamente contiene un intervalo de confianza del 95% de una o dos colas al igual que el estadístico de prueba y el valor p (puede cambiar el coeficiente de confianza si lo desea). La salida se muestra en la figura 6.25b).

FIGURA 6.25



4. El valor observado del estadístico de prueba $t = -.03$ es etiquetado “*T-Value*” seguido por el “*P-Value*” de dos colas. Para este ejemplo, el valor $p = .974$ indica que no hay una diferencia significativa en el promedio de las puntuaciones para los estudiantes enseñados por los dos profesores.
5. En la sección 6.7 se presentó una prueba de hipótesis formal para la igualdad de dos varianzas usando la prueba F . Para ejecutar esta prueba usando *MINITAB*, seleccione **Stat ► Basic Statistics ► 2 Variances**. En la lista desplegable, seleccione “Samples in two columns” e introduzca las columnas apropiadas de la lista de la izquierda. La porción pertinente de la salida se muestra en la figura 6.26. Para nuestra prueba de dos colas con $\alpha = 0.05$, el estadístico de prueba es $F = 1.27$ y el valor $p = .740$. No hay una diferencia significativa en las dos varianzas.

FIGURA 6.26



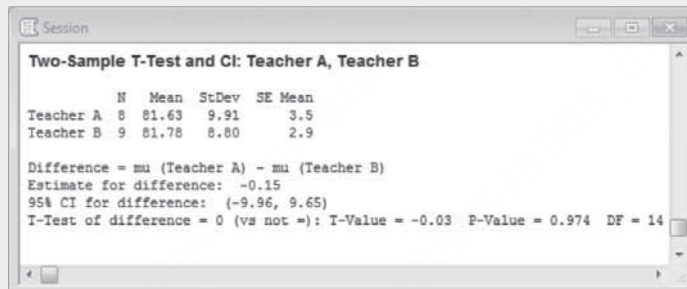
EJEMPLO

6.22

(Prueba t de dos muestras suponiendo varianzas desiguales)

1. Consulte el ejemplo 6.21. Si la razón de las dos varianzas muestrales hubiera sido tan grande que no pudiera suponer varianzas iguales (se usa “mayor que 3” como regla general), debería seleccionar **Stat ► Basic Statistics ► 2-Sample t**, pero NO seleccione el cuadro marcado “Assume Equal Variances”. Si usamos esta prueba en los datos del ejemplo 6.17, resulta la siguiente salida (figura 6.27).

FIGURA 6.27



2. Verá ligeras diferencias en los grados de libertad y no hay lista “Pooled StDev”, pero las conclusiones no cambian.
3. NOTA: Cuando *MINITAB* calcula los grados de libertad para la *Aproximación de Satterthwaite*, *MINITAB* usa la parte entera del valor calculado, el cual es diferente de los procedimientos usados en *MS Excel*. Debido a estos enfoques diferentes para determinar los grados de libertad, los resultados de las salidas de *MS Excel* y *MINITAB* diferirán ligeramente.

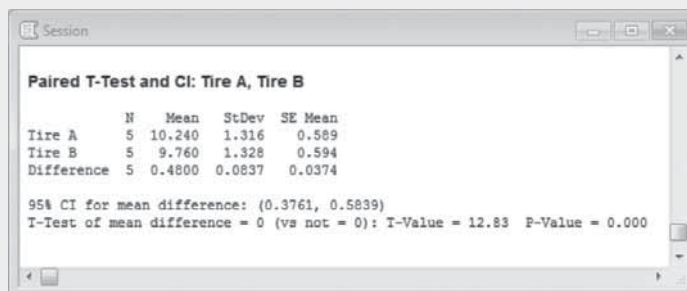
EJEMPLO

6.23

(Prueba de diferencia en pares) Consulte los datos del desgaste de las llantas de la tabla 6.3.

1. Para realizar una prueba de diferencia en pares para estas muestras dependientes, introduzca los datos en las primeras dos columnas de una hoja de cálculo de *MINITAB* y seleccione **Stat ► Basic Statistics ► Paired t**.
2. Siga las instrucciones para la prueba t de muestras independientes, y generará una salida similar. Para los datos en la tabla 6.3, se obtiene la salida en la figura 6.28.
3. Usando el valor observado del estadístico de prueba ($t = 12.83$) con un valor p de dos colas = .000, hay una evidencia fuerte que indica una diferencia en las dos medias poblacionales.

FIGURA 6.28



Ejercicios suplementarios

6.62 ¿Qué suposiciones se hacen cuando se usa la prueba t de Student para probar una hipótesis respecto a una media poblacional?

6.63 ¿Qué suposiciones se hacen acerca de las poblaciones de las cuales se obtienen muestras aleatorias, cuando se usa la distribución t para hacer inferencias de muestra pequeña respecto a la diferencia en medias poblacionales?

6.64 ¿Por qué usar observaciones en pares para estimar la diferencia entre dos medias poblacionales, en lugar de la estimación basada en muestras aleatorias independientes seleccionadas de las dos poblaciones? ¿Un experimento pareado es siempre preferible? Explique.

6.65 Use la tabla 4 en el Anexo para encontrar los siguientes valores críticos:

- Una región de rechazo de una cola superior con $\alpha = .05$ y 11 gl.
- Una región de rechazo de dos colas con $\alpha = .05$ y 7 gl.
- Una región de rechazo de una cola inferior con $\alpha = .01$ y 15 gl.

6.66 Use la tabla 4 en el Anexo para limitar los siguientes valores p :

- $P(t > 1.2)$ con 5 gl
- $P(t > 2) + P(t < -2)$ con 10 gl
- $P(t < -3.3)$ con 8 gl
- $P(t > 0.6)$ con 12 gl

6.67 Una muestra aleatoria de $n = 12$ observaciones de una población normal produjo $\bar{x} = 47.1$ y $s^2 = 4.7$. Pruebe la hipótesis $H_0: \mu = 48$ contra $H_a: \mu \neq 48$ en un nivel de significancia del 5%.

6.68 Impurezas II Un fabricante puede tolerar una pequeña cantidad (.05 miligramos por litro (mg/l)) de impurezas en una materia prima necesaria para manufacturar su producto. Debido a que la prueba de laboratorio para las impurezas está sujeta a error experimental, el fabricante prueba 10 veces cada lote. Suponga que el valor medio del error experimental es 0 y, por tanto, que el valor medio de las 10 lecturas de prueba es una estimación no sesgada de la verdadera cantidad de las impurezas en el lote. Para un lote particular de materia prima, la media de las 10 lecturas de prueba es .058 mg/l, con una desviación estándar de .012 mg/l. ¿Los datos dan suficiente evidencia para indicar que la cantidad de impurezas en el lote excede de .05 mg/l? Encuentre el valor p para la prueba e interprete su valor.

6.69 Pino rojo El crecimiento del tallo principal, medido para una muestra de 17 árboles de pino rojo de cuatro

años de edad, produjo una media y desviación estándar igual a 11.3 y 3.4 pulgadas, respectivamente. Encuentre un intervalo de confianza de 90% para el crecimiento medio de una población de árboles de pino rojo de cuatro años de edad sujetos a condiciones ambientales similares.

6.70 Hidróxido de sodio El objeto de un experimento de química general es determinar la cantidad (en mililitros) de solución de hidróxido de sodio (NaOH) necesaria para neutralizar 1 gramo de un ácido especificado. Ésta será una cantidad exacta, pero cuando el experimento se realice en el laboratorio, ocurrirá variación como resultado de un error experimental. Se hacen tres valoraciones usando fenolftaleína como indicador de la neutralidad de la solución (el pH es 7 para una solución neutra). Los tres volúmenes de NaOH requeridos para obtener un pH de 7 en cada una de las tres valoraciones son como sigue: 82.10, 75.75 y 75.44 mililitros. Use un intervalo de confianza de 99% para estimar el número medio de mililitros necesarios para neutralizar 1 gramo del ácido.

6.71 Cloruro de sodio Mediciones de ingesta de agua, obtenidas de una muestra de 17 ratas que habían sido inyectadas con una solución de cloruro de sodio, produjeron una media y desviación estándar de 31.0 y 6.2 centímetros cúbicos (cm³), respectivamente. Dado que el promedio de ingesta de agua para ratas no inyectadas observado en un periodo comparable es 22.0 cm³, ¿los datos indican que las ratas inyectadas bebieron más agua que las no inyectadas? Pruebe al nivel de significancia de 5%. Encuentre un intervalo de confianza de 90% para la ingesta media de agua para ratas inyectadas.

6.72 Erizos de mar Un experimentador estaba interesado en determinar el grosor medio de la corteza de huevecillos del erizo de mar. El grosor se midió para $n = 10$ huevos de erizo de mar y se obtuvieron estas mediciones.

4.5	6.1	3.2	3.9	4.7
5.2	2.6	3.7	4.6	4.1

Estime el grosor medio de la corteza usando un intervalo de confianza de 95%.

6.73 Sistemas de fabricación Una planta de producción tiene dos sistemas de fabricación complejos, a ambos se les da mantenimiento en intervalos de dos semanas. Sin embargo, un sistema es dos veces más antiguo que el otro. El número de productos terminados fabricados al día por cada uno de los sistemas se registra para 30 días de trabajo. Los resultados se dan en la tabla. ¿Estos datos presentan suficiente evidencia para concluir que la variabilidad en producción diaria justifica incrementar el

mantenimiento del sistema de fabricación antiguo? Use el método del valor p .

Sistema nuevo Sistema viejo

$$\bar{x}_1 = 246 \quad \bar{x}_2 = 240$$

$$s_1 = 15.6 \quad s_2 = 28.2$$

DATOS
EX0674

6.74 Fósiles Los datos de la tabla siguiente son los diámetros y alturas de 10 especímenes fósiles de una especie de molusco pequeño, *Rotularia (Annelida) fallax*, que fueron desenterrados en una expedición de trazado de mapas cerca de la Península Antártica.¹¹ La tabla proporciona un símbolo de identificación para el espécimen de fósil, el diámetro y altura del fósil en milímetros, así como la razón entre diámetro y altura.

Especimen	Diámetro	Altura	D/H
OSU 36651	185	78	2.37
OSU 36652	194	65	2.98
OSU 36653	173	77	2.25
OSU 36654	200	76	2.63
OSU 36655	179	72	2.49
OSU 36656	213	76	2.80
OSU 36657	134	75	1.79
OSU 36658	191	77	2.48
OSU 36659	177	69	2.57
OSU 36660	199	65	3.06
$\bar{x}$:	184.5	73	2.54
s :	21.5	5	.37

- Encuentre un intervalo de confianza de 95% para el diámetro medio de la especie.
- Encuentre un intervalo de confianza de 95% para la altura media de la especie.
- Encuentre un intervalo de confianza de 95% para la razón media entre diámetro y altura.
- Compare los tres intervalos construidos en las partes a, b y c. ¿El promedio de las razones es igual a la razón entre el promedio del diámetro y el promedio de altura?

6.75 Fósiles, continúa Consulte el ejercicio 6.74 y el conjunto de datos EX0674. Suponga que desea estimar el diámetro medio correcto a no más de 5 milímetros, con probabilidad igual a .95 de los especímenes fósiles. ¿Cuántos fósiles se tienen que incluir en su muestra?

DATOS
EX0676

6.76 Alcohol y tiempos de reacción Para probar el efecto del alcohol al aumentar el tiempo de reacción para responder a un estímulo dado, se midieron los tiempos de reacción de siete personas. Después de consumir 3 onzas de alcohol al 40%, el tiempo de reacción para cada una de las siete personas se midió otra vez. ¿Los datos siguientes indican que la media de tiempo de reacción después de consumir alcohol fue mayor que la media de tiempo de reacción antes de consumir alcohol? Use $\alpha = .05$.

Persona	1	2	3	4	5	6	7
Antes	4	5	5	4	3	6	2
Después	7	8	3	5	4	5	5

DATOS
EX0677

6.77 Queso, por favor A continuación aparecen los precios por onza de $n = 13$ marcas diferentes de rebanadas de queso envueltas individualmente:

29.0	24.1	23.7	19.6	27.5
28.7	28.0	23.8	18.9	23.9
21.6	25.9	27.4		

Construya una estimación de intervalo de confianza de 95% del promedio subyacente del precio por onza de rebanadas de queso envueltas individualmente.

6.78 Absorción de medicamento Se realizó un experimento para comparar la media de los tiempos requeridos para la absorción corporal de dos medicamentos, A y B. Se seleccionaron al azar 10 personas y se asignaron para recibir uno de los medicamentos. Se registraron los tiempos (en minutos) para que el medicamento llegara a un nivel especificado en el torrente sanguíneo y el resumen de datos se da en la tabla:

Medicamento A	Medicamento B
$\bar{x}_1 = 27.2$	$\bar{x}_2 = 33.5$
$s_1^2 = 16.36$	$s_2^2 = 18.92$

- ¿Los datos dan suficiente evidencia para indicar una diferencia en la media de los tiempos para la absorción de los dos medicamentos? Pruebe usando $\alpha = .05$.
- Encuentre el valor p aproximado para la prueba. ¿Este valor confirma sus conclusiones?
- Encuentre un intervalo de confianza de 95% para la diferencia en la media de los tiempos de absorción. ¿El intervalo confirma sus conclusiones de la parte a?

6.79 Absorción de medicamento, continúa Consulte el ejercicio 6.78. Supongamos que desea estimar la diferencia en la media de los tiempos para la correcta absorción, a no más de 1 minuto de diferencia, con probabilidad aproximadamente igual a .95.

- ¿Aproximadamente qué tan grande se requiere una muestra para cada medicamento (suponga que los tamaños muestrales son iguales)?
- Si realizar el experimento usando los tamaños muestrales de la parte a requerirá mucho tiempo y dinero, ¿puede hacerse algo para reducir los tamaños muestrales y todavía obtener un margen de error de 1 minuto para la estimación?

DATOS
EX0680

6.80 Faisanes de cuello anillado A continuación se dan los pesos en gramos de 10 machos y 10 hembras jóvenes de faisanes de cuello anillado.

Machos		Hembras	
1384	1672	1073	1058
1286	1370	1053	1123
1503	1659	1038	1089
1627	1725	1018	1034
1450	1394	1146	1281

- Use una prueba estadística para determinar si la varianza poblacional de los pesos de los machos difiere de la de las hembras.
- Pruebe si el promedio del peso de machos jóvenes de faisanes de cuello anillado excede del de las hembras en más de 300 gramos. (SUGERENCIA: El procedimiento que use debe tomar en cuenta los resultados del análisis de la parte a.)

DATOS
EX0681

6.81 Abejas Los insectos que revolotean en vuelo gastan enormes cantidades de energía para su tamaño y peso. Los datos que se muestran aquí se tomaron de un conjunto mucho mayor de datos recolectados por T. M. Casey y colegas.¹² Muestran las frecuencias de aleteo (en hertz) para dos especies diferentes de abejas, $n_1 = 4$ *Euglossa mandibularis* Friese y $n_2 = 6$ *Euglossa imperialis* Cockerell.

<i>E. mandibularis</i> Friese	<i>E. imperialis</i> Cockerell
235	180
225	169
190	180
188	185
	178
	182

- Con base en los rangos observados, ¿piensa usted que hay diferencia entre las dos varianzas poblacionales?
- Use una prueba apropiada para determinar si hay diferencia.
- Explique por qué una prueba t de Student con un estimador s^2 agrupado es inapropiado para comparar las medias de las frecuencias de aleteo para las dos especies de abejas.

DATOS
EX0682

6.82 Calcio El contenido de calcio (Ca) de una sustancia mineral en polvo fue analizada 10 veces, registrándose las siguientes composiciones porcentuales:

.0271	.0282	.0279	.0281	.0268
.0271	.0281	.0269	.0275	.0276

- Encuentre un intervalo de confianza de 99% para el verdadero contenido de calcio de esta sustancia.
- ¿Qué significa la frase “99% de confianza”?
- ¿Qué suposiciones se deben hacer del procedimiento de muestreo para que este intervalo de confianza sea válido? ¿Qué significa esto para el químico que realiza el análisis?

6.83 ¿Sol o sombra? Karl Niklas y T.G. Owens examinaron las diferencias en una planta particular, *Plantago Major L.*, cuando crecen a plena luz del sol contra condiciones de sombra.¹³ En este estudio, las plantas en sombra recibieron luz solar directa durante menos de 2 horas al día, en tanto que las plantas a pleno sol nunca tuvieron sombra. Un resumen parcial de los datos basados en $n_1 = 16$ plantas a pleno sol y $n_2 = 15$ plantas de sombra se muestra a continuación:

	Pleno sol		Sombra	
	$\bar{x}$	s	$\bar{x}$	s
Área de hojas (cm ²)	128.00	43.00	78.70	41.70
Área traslape (cm ²)	46.80	2.21	8.10	1.26
Número de hojas	9.75	2.27	6.93	1.49
Grosor (mm)	.90	.03	.50	.02
Longitud (cm)	8.70	1.64	8.91	1.23
Ancho (cm)	5.24	.98	3.41	.61

- ¿Qué suposiciones se requieren para usar los procedimientos de muestra pequeña dados en este capítulo, para comparar plantas a pleno sol contra las de sombra? A partir del resumen presentado, ¿piensa que cualquiera de estas suposiciones se han violado?
- ¿Los datos presentan suficiente evidencia para indicar una diferencia en la media del área de hojas para plantas de pleno sol contra las de sombra?
- ¿Los datos presentan suficiente evidencia para indicar una diferencia en la media del área de traslape para plantas de pleno sol contra las de sombra?

6.84 Jugo de naranja Se ha de hacer una comparación de las precisiones de dos máquinas desarrolladas para extraer jugo de naranjas, usando los datos siguientes:

Máquina A	Máquina B
$s^2 = 3.1$ onzas ²	$s^2 = 1.4$ onzas ²
$n = 25$	$n = 25$

- ¿Hay suficiente evidencia para indicar que hay una diferencia en la precisión de las dos máquinas al nivel de significancia de 5%?
- Encuentre un intervalo de confianza de 95% para la razón entre las dos varianzas poblacionales. ¿Este intervalo confirma las conclusiones de usted respecto de la parte a? Explique.

DATOS
EX0685

6.85 ¿En casa o en preescolar? Cuatro pares de gemelos idénticos (pares A, B, C y D) se seleccionaron al azar de una base de datos computarizada de gemelos idénticos. Un niño se seleccionó al azar de cada par para formar un “grupo experimental”. Estos cuatro niños fueron enviados a preescolar. Los otros cuatro niños se mantuvieron en casa como grupo de control. Al final del año escolar, se

obtuvieron las siguientes calificaciones de CI (cociente de inteligencia):

Par	Grupo experimental	Grupo de control
A	110	111
B	125	120
C	139	128
D	142	135

¿Esta evidencia justifica la conclusión de que la falta de experiencia preescolar tiene un efecto desalentador en calificaciones del CI? Use el método del valor p .

DATOS
EX0686 **6.86 Dietas** Ocho personas obesas fueron puestas a dieta durante un mes, registrándose sus pesos al principio y a final del mes:

Pesos		
Persona	Inicial	Final
1	310	263
2	295	251
3	287	249
4	305	259
5	270	233
6	323	267
7	277	242
8	299	265

Estime la pérdida media de peso para personas obesas cuando se ponen a dieta durante un periodo de un mes. Use un intervalo de confianza de 95% e interprete sus resultados. ¿Qué suposiciones deben hacerse para que la inferencia sea válida?

DATOS
EX0687 **6.87 Patrones de respiración** Psicólogos investigadores midieron la línea base de patrones de respiración –la ventilación total (en litros de aire por minuto) ajustados para el tamaño del cuerpo–, para cada uno de $n = 30$ pacientes, de modo que pudieran estimar el promedio de ventilación total para pacientes antes de realizar cualquier experimento. Los datos, junto con algunas salidas impresas de MINITAB, se presentan a continuación:

5.23	5.72	5.77	4.99	5.12	4.82
5.54	4.79	5.16	5.84	4.51	5.14
5.92	6.04	5.83	5.32	6.19	5.70
4.72	5.38	5.48	5.37	4.96	5.58
4.67	5.17	6.34	6.58	4.35	5.63

Salida impresa MINITAB para el ejercicio 6.87

Pantalla de tallo y hoja: Lt/min

Stem-and-leaf of Ltrs/min N = 30
Leaf Unit = 0.10

1	4	3
2	4	5
5	4	677
8	4	899
12	5	1111
(4)	5	2333

14	5	455
11	5	6777
7	5	889
4	6	01
2	6	3
1	6	5

Estadística descriptiva: Lt/min

Variable	N	N*	Mean	SE Mean	StDev
Ltrs/min	30	0	5.3953	0.0997	0.5462
Minimum		Q1	Median	Q3	Maximum
4.3500	4.9825	5.3750	5.7850	6.5800	

- ¿Qué información proporciona la gráfica de tallo y hoja acerca de los datos? ¿Por qué es importante esto?
- Use la salida para construir un intervalo de confianza de 99% para el promedio de ventilación total para pacientes.

DATOS
EX0688 **6.88 Tiempos de reacción** Una comparación de tiempos de reacción (en segundos) para dos estímulos diferentes, en un experimento psicológico de asociación de palabras, produjo los siguientes resultados cuando se aplicaron a una muestra aleatoria de 16 personas:

Estímulo 1	1	3	2	1	2	1	3	2
Estímulo 2	4	2	3	3	1	2	3	3

¿Los datos presentan suficiente evidencia para indicar una diferencia en la media de los tiempos de reacción para los dos estímulos? Pruebe usando $\alpha = .05$.

DATOS
EX0689 **6.89 Tiempos de reacción II** Consulte el ejercicio 6.88. Suponga que el experimento de asociación de palabras se realiza usando bloques de ocho personas y haciendo una comparación de tiempos de reacción en cada persona; esto es, cada persona es sometida a ambos estímulos en orden aleatorio. Los tiempos de reacción (en segundos) para el experimento son como sigue:

Persona	Estímulo 1	Estímulo 2
1	3	4
2	1	2
3	1	3
4	2	1
5	1	2
6	2	3
7	3	3
8	2	3

¿Los datos presentan suficiente evidencia para indicar una diferencia en la media de tiempos de reacción para los dos estímulos? Pruebe usando $\alpha = .05$.

6.90 Consulte los ejercicios 6.88 y 6.89. Calcule un intervalo de confianza de 95% para la diferencia en las dos medias poblacionales para cada uno de estos diseños experimentales. ¿Parece que ponerlas en bloque aumentó la cantidad de información disponible en el experimento?

DATOS
EX0691

6.91 Resistencia al impacto Los datos siguientes son lecturas (en pies–libras) de las resistencias al impacto de dos clases de material de empaque:

A	B
1.25	.89
1.16	1.01
1.33	.97
1.15	.95
1.23	.94
1.20	1.02
1.32	.98
1.28	1.06
1.21	.98

Salida de *MS Excel* para el ejercicio 6.91

D	E	F
t-Test: Two-Sample Assuming Equal Variances		
	A	B
Mean	1.2367	0.9778
Variance	0.0042	0.0024
Observations	9	9
Pooled Variance	0.0033	
Hypothesized Mean Difference	0	
df	16	
t Stat	9.5641	
P(T<=t) one-tail	0.0000	
t Critical one-tail	1.7459	
P(T<=t) two-tail	0.0000	
t Critical two-tail	2.1199	

a. Use la salida impresa de *MS Excel* para determinar si hay evidencia de una diferencia en las medias de resistencias para las dos clases de material.

b. ¿Hay implicaciones prácticas para sus resultados?

6.92 ¿De acuerdo con qué suposiciones puede usarse la distribución *F* al hacer inferencias acerca de la razón entre varianzas poblacionales?

6.93 ¿Compró leche? Una lechería busca comprar una nueva máquina llenadora de recipientes y considera dos modelos, manufacturados por la compañía A y la compañía B. La robustez, el costo y la comodidad son comparables en los dos modelos, de modo que el factor decisivo es la variabilidad en el llenado. Se prefiere el modelo que produce llenados con menor varianza. Si se obtienen muestras de llenados para cada uno de los dos modelos, se usaría una prueba *F* para probar la igualdad de varianzas poblacionales. ¿Cuál tipo de región de rechazo sería más favorable para cada una de estas personas?

a. El gerente de la lechería, ¿por qué?

b. Una representante de ventas de la compañía A, ¿por qué?

c. Una representante de ventas de la compañía B, ¿por qué?

DATOS
EX0694

6.94 Producción de alimentos En un tiempo en que la conservación de energía es tan importante,

algunos científicos piensan que debe darse un escrutinio más estrecho al costo (en energía) de producir varias formas de alimentos. Supongamos que usted desea comparar la media de la cantidad de petróleo necesario para producir un acre de maíz contra un acre de coliflor. En la tabla siguiente se muestran las lecturas (en barriles de petróleo por acre), con base en lotes de 20 acres, siete por cada cosecha.

Maíz	Coliflor
5.6	15.9
7.1	13.4
4.5	17.6
6.0	16.8
7.9	15.8
4.8	16.3
5.7	17.1

a. Use estos datos para encontrar un intervalo de confianza de 90% para las diferencias entre la media de las cantidades de petróleo requeridas para producir estos dos cultivos.

b. Con base en el intervalo encontrado en la parte a, ¿hay evidencia de una diferencia en el promedio de las cantidades de petróleo requerida para producir estos dos cultivos? Explique.

6.95 Alcohol y altitud El efecto del consumo de alcohol en el cuerpo parece ser mucho mayor a grandes alturas que al nivel del mar. Para probar esta hipótesis, un científico selecciona al azar 12 personas y las divide aleatoriamente en dos grupos de seis cada uno. Un grupo se pone en una cámara que simula condiciones a una altitud de 12 000 pies y cada persona ingiere una bebida que contiene 100 centímetros cúbicos (cc) de alcohol. El segundo grupo recibe la misma bebida en una cámara que simula condiciones al nivel del mar. Después de 2 horas, se mide la cantidad de alcohol en la sangre (gramos por 100 cc) de cada persona. Los datos se muestran en la tabla. ¿Los datos dan suficiente evidencia para apoyar la hipótesis de que el promedio de la cantidad de alcohol en la sangre después de dos horas es mayor a grandes alturas?

Nivel del mar	12000 pies
.07	.13
.10	.17
.09	.15
.12	.14
.09	.10
.13	.14

6.96 Riesgos accionarios Los precios al cierre de dos acciones comunes se registraron durante un periodo de 15 días. Las medias y varianzas son

$$\bar{x}_1 = 40.33 \quad \bar{x}_2 = 42.54$$

$$s_1^2 = 1.54 \quad s_2^2 = 2.96$$

- a. ¿Estos datos presentan suficiente evidencia, para indicar una diferencia entre las variabilidades de los precios al cierre de las dos acciones para las poblaciones asociadas con las dos muestras? Dé el valor de p para la prueba e interprete este valor.
- b. Construya un intervalo de confianza de 99% para la razón entre las dos varianzas poblacionales.

DATOS
EX0697

6.97 Programas de seguridad Los datos mostrados a continuación se recolectaron de los accidentes incapacitantes (las cifras dadas son las medias de las horas perdidas de trabajo por mes, en un periodo de un año) antes y después de poner en vigor un programa de seguridad industrial. Se registraron datos para seis plantas industriales. ¿Los datos dan suficiente evidencia para indicar si el programa de seguridad fue eficaz para reducir los accidentes incapacitantes? Pruebe usando $\alpha = .01$.

	Planta número					
	1	2	3	4	5	6
Antes del programa	38	64	42	70	58	30
Después del programa	31	58	43	65	52	29

DATOS
EX0698

6.98 Dos platos principales diferentes Para comparar la demanda de dos platos principales diferentes, el gerente de una cafetería registró el número de ventas de cada plato principal en siete días consecutivos. Los datos se muestran en la tabla. ¿Los datos dan suficiente evidencia para indicar una media mayor de demanda para uno de los platos principales? Use la salida impresa de *Excel*.

Día	A	B
Lunes	420	391
Martes	374	343
Miércoles	434	469
Jueves	395	412
Viernes	637	538
Sábado	594	521
Domingo	679	625

Salida de *MS Excel* para el ejercicio 6.98

E	F	G
t-Test: Paired Two Sample for Means		
	A	B
Mean	504.714	471.286
Variance	16191.238	9495.571
Observations	7	7
Pearson Correlation	0.945	
Hypothesized Mean Difference	0	
df	6	
t Stat	1.862	
P(T<=t) one-tail	0.056	
t Critical one-tail	1.943	
P(T<=t) two-tail	0.112	
t Critical two-tail	2.447	

6.99 Control de contaminación El límite de la EPA de descarga permisible de sólidos suspendidos en ríos y arroyos es de 60 miligramos por litro (mg/l) por día. Un estudio de muestras de agua seleccionadas de la descarga en una mina de fosfato muestra que en un tiempo prolongado, la media de la descarga diaria de sólidos suspendidos es 48 mg/l, pero las lecturas de descargas de un día para otro son variables. Unos inspectores del estado midieron los porcentajes de descarga de sólidos suspendidos para $n = 20$ días y hallaron $s^2 = 39$ (mg/l)². Encuentre un intervalo de confianza de 90% para σ^2 . Interprete sus resultados.

DATOS
EX6100

6.100 Enzimas Se emplearon dos métodos para medir la actividad específica (en unidades de actividad enzimática por miligramo de proteína) de una enzima. Una unidad de actividad enzimática es la cantidad que cataliza la formación de 1 micromol de producto por minuto bajo condiciones especificadas. Use una prueba o procedimiento de estimación apropiado para comparar los dos métodos de medición. Comente sobre la validez de cualesquier suposiciones que sean necesarias.

Método 1	125	137	130	151	142
Método 2	137	143	151	156	149

6.101 Barras conectoras Un productor de piezas maquinadas dice que los diámetros de las barras conectoras producidas por su planta tenían una varianza de a lo sumo .03 pulgada². Una muestra aleatoria de 15 barras conectoras de su planta produjo una media y varianza muestrales de .55 pulgadas y .053 pulgadas², respectivamente.

- a. ¿Hay suficiente evidencia para rechazar lo declarado por el productor al nivel de significancia de $\alpha = .05$?
- b. Encuentre un intervalo de confianza de 95% para la varianza de los diámetros de barras conectoras.

6.102 El sueño y el estudiante universitario ¿Cuánto duerme usted en una noche escolar típica? A un grupo de 10 estudiantes universitarios se le pidió informar el número de horas que dormían en la noche previa, con los siguientes resultados:

7, 6, 7.25, 7, 8.5, 5, 8, 7, 6.75, 6

- a. Encuentre un intervalo de confianza de 99% para el promedio del número de horas que duermen los estudiantes universitarios.
- b. ¿Qué suposiciones se requieren para que este intervalo de confianza sea válido?

DATOS
EX6103

6.103 Reacomodo de objetos Los siguientes datos son tiempos de respuesta, en segundos, para $n = 25$ estudiantes de primer año para acomodar tres objetos por tamaño.

5.2	3.8	5.7	3.9	3.7
4.2	4.1	4.3	4.7	4.3
3.1	2.5	3.0	4.4	4.8
3.6	3.9	4.8	5.3	4.2
4.7	3.3	4.2	3.8	5.4

Encuentre un intervalo de confianza de 95% para el promedio de tiempo de respuesta para estudiantes de primer año para acomodar tres objetos por tamaño. Interprete este intervalo.

6.104 Se delatan en el centro comercial Un artículo de *American Demographics* investigó hábitos de consumo en el centro comercial. Tendemos a gastar más dinero de compras en fines de semana y, en particular, los domingos de 4 a 6 p.m. ¡Los miércoles por la mañana los compradores gastan menos!¹⁴ Supongamos que se seleccionó una muestra aleatoria de 20 compradores de fin

de semana y una muestra aleatoria de 20 compradores de día entre semana, y se registró la cantidad gastada por viaje al centro comercial.

	Fines de semana	Día entre semana
Tamaño de la muestra	20	20
Media muestral (\$)	78	67
Desviación muestral estándar (\$)	22	20

- ¿Es razonable suponer que las dos varianzas poblacionales son iguales? Use la prueba F para probar esta hipótesis con $\alpha = .05$.
- Con base en los resultados de la parte **a**, use la prueba apropiada para determinar si hay diferencia en el promedio de la cantidad gastada por viaje en fines de semana contra días laborables. Pruebe usando $\alpha = .05$.

Análisis de varianza

OBJETIVO GENERAL

La cantidad de información contenida en una muestra es afectada por diversos factores que el experimentador puede o no controlar. Este capítulo presenta tres *diseños experimentales* diferentes, dos de los cuales son extensiones directas de los diseños en pares y no pareados del capítulo 6. Una nueva técnica llamada *análisis de varianza* se utiliza para determinar el modo en que diferentes factores experimentales afectan la respuesta promedio.

ÍNDICE DEL CAPÍTULO

- El diseño de un experimento (7.1)
- ¿Qué es un análisis de varianza? (7.2)
- Las suposiciones para un análisis de varianza (7.3)
- El diseño completamente aleatorizado: una clasificación en una dirección (7.4)
- El análisis de varianza para un diseño completamente aleatorizado (7.5)
- Clasificación de medias poblacionales (7.6)
- Diseño de bloque aleatorizado: una clasificación en dos direcciones (7.7)
- El análisis de varianza para un diseño de bloque aleatorizado (7.8)
- El experimento factorial $a \times b$: una clasificación en dos vías (7.9)
- El análisis de varianza para un experimento factorial $a \times b$ (7.10)
- Repaso de las suposiciones del análisis de varianza (7.11)
- Un breve repaso (7.12)



NECESITO SABER...

Cómo determinar si mis cálculos son precisos

¡Cómo ahorrar dinero en comestibles!

Los productos comestibles enlatados o congelados que se compran a granel casi siempre nos ayudan a ahorrar dinero en comparación con los que se compran en los supermercados. Puede ahorrar más de 75% en productos comestibles enlatados y más de 80% en congelados, comprados a granel. El caso práctico al final de este capítulo investiga los costos de comprar a granel enlatados y congelados, usando los procedimientos de análisis de varianza que se presentan en este capítulo.

7.1

EL DISEÑO DE UN EXPERIMENTO

La forma en que se selecciona una muestra se denomina *plan de muestreo* o *diseño experimental* y determina la cantidad de información en la muestra. Algunas investigaciones comprenden un **estudio observacional**, en el que el investigador no produce en realidad los datos sino que sólo *observa* las características de los datos que ya existen. Casi todos los estudios muestrales, en donde se reúne información con un cuestionario, se encuentran en esta categoría. El investigador diseña un plan para recolectar los datos, llamado *plan de muestreo*, y a continuación utiliza los procedimientos estadísticos apropiados para obtener conclusiones acerca de la población o poblaciones de donde proviene la muestra.

Otra investigación comprende la **experimentación**. El investigador puede deliberadamente establecer una o más condiciones en las unidades experimentales para determinar su efecto en la respuesta. Veamos ahora algunos términos nuevos que usaremos para discutir el diseño de un experimento estadístico.

Definición Una **unidad experimental** es el objeto en el que se toma una medición (o mediciones).

Un **factor** es una variable independiente cuyos valores son controlados y variados por el experimentador.

Un **nivel** es el escenario de intensidad de un factor.

Un **tratamiento** es una combinación específica de niveles de factor.

La **respuesta** es la variable que es medida por el experimentador.

EJEMPLO

7.1

Un conjunto de personas se divide al azar en un grupo experimental y uno de control. Al grupo de control se le aplica un examen de aptitud después de haber tomado un desayuno completo; al grupo experimental se le aplica el mismo examen sin haber tomado desayuno. ¿Cuáles son los factores, niveles y tratamientos en este experimento?

Solución Las *unidades experimentales* son las personas en las que la *respuesta* (calificación de examen) se mide. El *factor* de interés podría describirse como “comida” y tiene dos *niveles*: “desayuno” y “sin desayuno”. Como éste es el único factor controlado por el experimentador, los dos niveles de “desayuno” y “sin desayuno” también representan los *tratamientos* de interés en el experimento.

EJEMPLO

7.2

Suponga que el experimentador del ejemplo 7.1 empezó por seleccionar al azar 20 hombres y 20 mujeres para el experimento. Estos dos grupos se dividieron entonces al azar en 10 cada uno para los grupos experimental y de control. ¿Cuáles son los factores, niveles y tratamientos en este experimento?

Solución Ahora hay dos *factores* de interés para el experimentador y cada factor tiene dos *niveles*:

- “Género” a dos niveles: hombres y mujeres
- “Comida” a dos niveles: desayuno y sin desayuno

En este experimento más complejo hay cuatro *tratamientos*, uno para cada combinación específica de niveles de factor: hombres sin desayuno, hombres con desayuno, mujeres sin desayuno y mujeres con desayuno.

En este capítulo nos concentraremos en los experimentos que se han diseñado en tres formas diferentes y usaremos una técnica llamada *análisis de varianza* para juzgar los efectos de varios factores en la respuesta experimental. Dos de estos *diseños experimentales* son extensiones de los diseños en pares y no pareados del capítulo 6.

¿QUÉ ES UN ANÁLISIS DE VARIANZA?

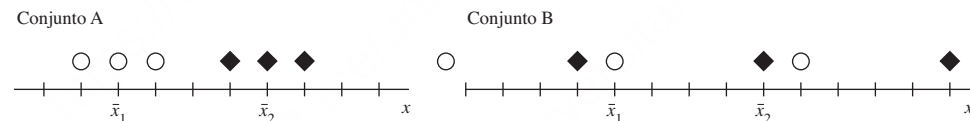
7.2

Las respuestas que se generan en una situación experimental siempre muestran cierta cantidad de *variabilidad*. En un **análisis de varianza** se divide la variación total de las mediciones de respuesta en partes que pueden ser atribuidas a varios *factores* de interés para el experimentador. Si el experimento se ha diseñado en forma apropiada estas partes pueden usarse entonces para contestar preguntas acerca de los efectos de los diversos factores en la respuesta de interés.

Se puede entender mejor la lógica subyacente en un análisis de varianza al analizar un experimento sencillo. Considere dos conjuntos de muestras seleccionadas al azar de las poblaciones 1 (◆) y 2 (○), cada uno con pares idénticos de medias; $\bar{x}_1$ y $\bar{x}_2$. Los dos conjuntos se observan en la figura 7.1. ¿Es más fácil detectar la diferencia en las dos medias cuando se vea el conjunto A o el conjunto B? Es probable que esté de acuerdo en que el conjunto A muestra la diferencia con mayor claridad. En el conjunto A, la variabilidad de las mediciones *dentro* de los grupos (los ◆ y los ○) es mucho menor que la variabilidad *entre* los dos grupos. En el conjunto B, hay más variabilidad *dentro* de los grupos (los ◆ y los ○), causando que los dos grupos se “mezclen” y hagan más difícil ver la diferencia *idéntica* en las medias.

FIGURA 7.1

Dos conjuntos de muestras con las mismas medias



La comparación que usted ha hecho intuitivamente es formalizada por el análisis de varianza. Es más, el análisis de varianza se puede usar no sólo para comparar dos medias sino también para hacer comparaciones de *más de dos* medias poblacionales y para determinar los efectos de varios factores en diseños experimentales más complejos. El análisis de varianza se apoya en estadísticas con distribuciones muestrales que son modeladas por la distribución *F* de la sección 6.7.

LAS SUPOSICIONES PARA UN ANÁLISIS DE VARIANZA

7.3

Las suposiciones necesarias para un análisis de varianza son semejantes a las requeridas para los estadísticos *t* de Student y *F* del capítulo 6. Cualquiera que sea el diseño experimental empleado para generar los datos, se debe suponer que las observaciones dentro de cada grupo de tratamiento están **normalmente distribuidas** con una **varianza común** σ^2 . Al igual que en el capítulo 6, los procedimientos de análisis de varianza son bastante **robustos** cuando los tamaños muestrales son iguales y cuando los datos son de forma de montículo. Violar la suposición de una varianza común es más serio, en especial cuando los tamaños muestrales no son cercanamente iguales.

SUPOSICIONES PARA LA PRUEBA DE ANÁLISIS DE VARIANZA Y LOS PROCEDIMIENTOS DE ESTIMACIÓN

- Las observaciones dentro de cada población están distribuidas normalmente con una varianza común σ^2 .
- Las suposiciones respecto al procedimiento de muestreo se especifican para cada diseño en las secciones que siguen.

Este capítulo describe el análisis de varianza para tres diseños experimentales diferentes. El primer diseño se basa en el muestreo aleatorio independiente de varias poblaciones y es una extensión de la *prueba t independiente* del capítulo 6. El segundo es una extensión del diseño de *diferencia en pares* o *pares acoplados* y comprende una asignación aleatoria de tratamientos dentro de conjuntos de observaciones en pares. El tercero es un diseño que permite juzgar el efecto de dos factores experimentales en la respuesta. Los procedimientos de muestreo necesarios para cada diseño se vuelven a expresar en sus secciones respectivas.

EL DISEÑO COMPLETAMENTE ALEATORIZADO: UNA CLASIFICACIÓN EN UNA DIRECCIÓN

7.4

Uno de los diseños experimentales más sencillos es el **diseño completamente aleatorizado**, en el que muestras aleatorias se seleccionan de manera independiente de cada una de k poblaciones. Este diseño comprende sólo un *factor*, la población de donde proviene la medición, de aquí la designación como una **clasificación en una dirección**. Hay k *niveles* diferentes correspondientes a las k poblaciones, que también son los *tratamientos* para esta clasificación de una dirección. ¿Las k medias poblacionales son todas iguales, o al menos una media es diferente de las otras?

¿Por qué se necesita un nuevo procedimiento, el *análisis de varianza*, para comparar las medias poblacionales cuando ya se tiene disponible la prueba t de Student? Al comparar $k = 3$ medias se podría probar cada uno de los tres pares de hipótesis:

$$H_0 : \mu_1 = \mu_2 \quad H_0 : \mu_1 = \mu_3 \quad H_0 : \mu_2 = \mu_3$$

para averiguar dónde están las diferencias. No obstante, es preciso recordar que cada prueba que se realice está sujeta a la posibilidad de error. Para comparar $k = 4$ medias se necesitarían seis pruebas y se necesitarían 10 pruebas para comparar $k = 5$ medias. Cuantas más pruebas se realicen en un conjunto de mediciones más probable será que al menos una de las conclusiones sea incorrecta. El procedimiento de análisis de varianza provee una prueba general para juzgar la igualdad de las k medias poblacionales. Una vez que haya determinado si hay *en realidad* una diferencia en las medias se puede usar otro procedimiento para averiguar dónde están las diferencias.

¿Cómo se pueden seleccionar estas k muestras aleatorias? A veces las poblaciones existen en realidad y se puede usar un generador computarizado de números aleatorios o una tabla de números aleatorios para seleccionar al azar las muestras. Por ejemplo, en un estudio para comparar los tamaños promedio de reclamaciones de seguro médico en cuatro estados diferentes se podría usar una base de datos computarizada, proporcionada por las compañías de seguros médicos para seleccionar muestras aleatorias de los cuatro estados. En otras situaciones las poblaciones pueden ser *hipotéticas* y se pueden generar respuestas sólo después de que se han aplicado los tratamientos experimentales.

EJEMPLO

7.3

Un investigador está interesado en los efectos de cinco tipos de insecticidas a usar para controlar el gorgojo del algodón en campos algodoneiros. Explique cómo poner en práctica un diseño completamente aleatorizado para investigar los efectos de los cinco insecticidas en la producción de cosechas.

Solución La única forma de generar el equivalente de cinco muestras aleatorias de las poblaciones hipotéticas, correspondientes a los cinco insecticidas, es usar un método llamado **asignación aleatorizada**. Se escoge un número fijo de plantas de algodón para tratamiento y a cada una se le asigna un número aleatorio. Suponga que cada muestra debe tener un número igual de mediciones. Con el uso de un medio de aleatorización, se pueden asignar las primeras n_1 plantas escogidas para recibir el insecticida 1, las segundas n_2 plantas para recibir el insecticida 2 y así sucesivamente, hasta que se hayan asignado los cinco tratamientos.

Ya sea por *selección aleatoria* o *asignación aleatoria*, estos dos ejemplos resultan en un diseño completamente aleatorizado o clasificación en una dirección, para el cual se usa el análisis de varianza.

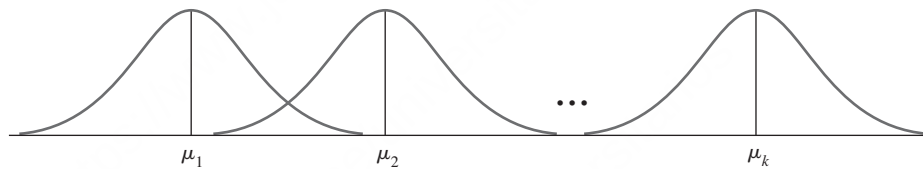
EL ANÁLISIS DE VARIANZA PARA UN DISEÑO COMPLETAMENTE ALEATORIZADO

7.5

Suponga que se desea comparar k medias poblacionales, $\mu_1, \mu_2, \dots, \mu_k$, con base en muestras aleatorias independientes de tamaño $n_1, n_2, \dots, n_k$ de poblaciones normales con una varianza común σ^2 . Esto es, cada una de las poblaciones normales tiene la misma forma, pero sus ubicaciones podrían ser diferentes, como se ve en la figura 7.2.

FIGURA 7.2

Poblaciones normales con una varianza común pero medias diferentes



División de la variación total en un experimento

Sea x_{ij} la j -ésima medición ($j = 1, 2, \dots, n_i$) en la i -ésima muestra. El procedimiento de análisis de varianza empieza por considerar la variación total en el experimento, que es medida por una cantidad llamada **suma total de cuadrados (TSS)**:

$$\text{TSS} = \sum (x_{ij} - \bar{x})^2 = \sum x_{ij}^2 - \frac{(\sum x_{ij})^2}{n}$$

Éste es el conocido numerador de la fórmula para la varianza muestral para todo el conjunto de $n_1 + n_2 + \dots + n_k$ mediciones. La segunda parte de la fórmula de cálculo se denomina a veces **corrección para la media (CM)**. Si con G representamos el *gran total* de todas las n observaciones, entonces

$$\text{CM} = \frac{(\sum x_{ij})^2}{n} = \frac{G^2}{n}$$

Esta suma total de cuadrados (TSS) se divide en dos componentes. El primero, llamado **suma de cuadrados para tratamientos (SST)**, mide la variación entre las k medias muestrales:

$$\text{SST} = \sum n_i (\bar{x}_i - \bar{x})^2 = \sum \frac{T_i^2}{n_i} - \text{CM}$$

donde T_i es el total de observaciones para el tratamiento i . El segundo componente, llamado **suma de cuadrados para el error (SSE)**, se usa para medir la variación agrupada dentro de las k muestras:

$$SSE = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2$$

Esta fórmula es una extensión directa del numerador en la fórmula para la estimación agrupada de σ^2 del capítulo 6. Podemos demostrar algebraicamente que, en el análisis de varianza,

$$TSS = SST + SSE$$

Por tanto, es necesario calcular sólo dos de las tres sumas de cuadrados: TSS, SST y SSE, y la tercera se puede hallar por sustracción.

Cada una de las fuentes de variación, cuando es dividida entre sus apropiados **grados de libertad**, da una estimación de la variación en el experimento. Como TSS involucra n observaciones cuadradas, sus grados de libertad son $gl = (n - 1)$. Del mismo modo, la suma de cuadrados para tratamientos comprende k observaciones cuadradas y sus grados de libertad son $gl = (k - 1)$. Por último, la suma de cuadrados de error, una extensión directa de la estimación agrupada del capítulo 6, tiene

$$gl = (n_1 - 1) + (n_2 - 1) + \cdots + (n_k - 1) = n - k$$

Observe que los grados de libertad para tratamientos y error son aditivos, es decir,

$$gl \text{ (total)} = gl \text{ (tratamientos)} + gl \text{ (error)}$$

Estas dos fuentes de variación y sus respectivos grados de libertad se combinan para formar los **cuadráticos medios** como $MS = SS/gl$. La variación total en el experimento se exhibe entonces en una **tabla de análisis de varianza** (o **ANOVA**).

TIP MI CONSEJO

La columna marcada "SS" satisface:
TSS = SST + SSE

TIP MI CONSEJO

La columna marcada " gl " siempre asciende a $n - 1$

TABLA ANOVA PARA k MUESTRAS ALEATORIAS INDEPENDIENTES: DISEÑO COMPLETAMENTE ALEATORIZADO

Fuente	gl	SS	MS	F
Tratamientos	$k - 1$	SST	$MST = SST/(k - 1)$	MST/MSE
Error	$n - k$	SSE	$MSE = SSE/(n - k)$	
Total	$n - 1$	TSS		

donde

$$\begin{aligned} TSS &= \sum x_{ij}^2 - CM \\ &= (\text{Suma de cuadrados de todos los valores } x) - CM \end{aligned}$$

con

$$CM = \frac{(\sum x_{ij})^2}{n} = \frac{G^2}{n}$$

$$SST = \sum \frac{T_i^2}{n_i} - CM \qquad MST = \frac{SST}{k - 1}$$

$$SSE = TSS - SST \qquad MSE = \frac{SSE}{n - k}$$

y

G = Gran total de todas las n observaciones

T_i = Total de todas las observaciones en la muestra i

n_i = Número de observaciones en la muestra i

$n = n_1 + n_2 + \cdots + n_k$

EJEMPLO

7.4

En un experimento para determinar el efecto de la nutrición sobre los intervalos de atención de los estudiantes de las escuelas básicas, un grupo de 15 estudiantes se asignó al azar a uno de tres planes de comidas: sin desayuno, desayuno ligero y desayuno completo. Sus intervalos de atención (en minutos) se registraron durante un periodo de lectura por la mañana y se muestran en la tabla 7.1. Construya la tabla de análisis de varianza para este experimento.

TABLA 7.1

Intervalos de atención de los estudiantes después de tres planes de comidas

Sin desayuno	Desayuno ligero	Desayuno completo
8	14	10
7	16	12
9	12	16
13	17	15
10	11	12
$T_1 = 47$	$T_2 = 70$	$T_3 = 65$

Solución Para usar las fórmulas de cálculo, se necesitan los $k = 3$ tratamientos totales con $n_1 = n_2 = n_3 = 5$, $n = 15$ y $\sum x_{ij} = 182$. Entonces

$$CM = \frac{(182)^2}{15} = 2208.2667$$

$$TSS = (8^2 + 7^2 + \dots + 12^2) - CM = 2338 - 2208.2667 = 129.7333 \text{ con } (n - 1) = (15 - 1) = 14 \text{ grados de libertad,}$$

$$SST = \frac{47^2 + 70^2 + 65^2}{5} - CM = 2266.8 - 2208.2667 = 58.5333$$

con $(k - 1) = (3 - 1) = 2$ grados de libertad, y por sustracción,

$$SSE = TSS - SST = 129.7333 - 58.5333 = 71.2$$

con $(n - k) = (15 - 3) = 12$ grados de libertad. Estas tres fuentes de variación, sus grados de libertad, sumas de cuadrados y cuadráticos medios se muestran en el área sombreada de las tablas ANOVA generadas por *MS Excel* y *MINITAB* y se dan en la figura 7.3. Se encontrarán instrucciones para generar esta salida impresa en la sección “Tecnología actual” al final de este capítulo.

FIGURA 7.3(a)

Salida MINITAB para el ejemplo 7.4

ANOVA de una dirección: intervalo contra comida

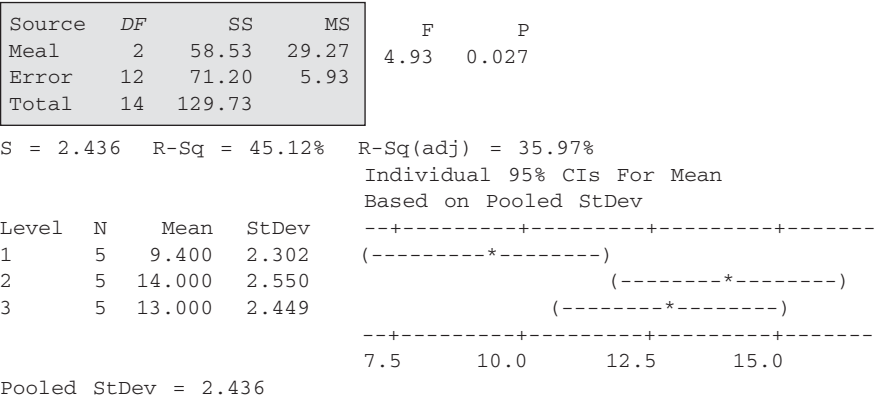


FIGURA 7.3(b)Salida de *MS Excel* para el ejemplo 7.4**Anova: factor único**

RESUMEN

Grupos	Conteo	Suma	Promedio	Varianza
None	5	47	9.4	5.3
Light	5	70	14	6.5
Full	5	65	13	6

ANOVA

Fuente de variación	SS	gl	MS	F	Valor P	F crít.
Entre grupos	58.533	2	29.267	4.933	0.027	3.885
Dentro de grupos	71.2	12	5.933			
Total	129.733	14				

Las salidas de computadora dan alguna información adicional sobre la variación en el experimento. La sección inferior en *MINITAB* y la sección superior en *MS Excel* muestran las medias y desviaciones estándar (o varianzas) para los tres planes de comidas. Más importante aún es que se pueden ver en la sección superior en *MINITAB* y en la sección inferior en *MS Excel* dos columnas marcadas “F” y “P” (“valor F” y “valor P” en *Excel*). Podemos usar estos valores para probar una hipótesis respecto a la igualdad de las tres medias de tratamiento.

Prueba de igualdad de las medias de tratamiento

Los *cuadráticos medios* en la tabla de análisis de varianza se pueden usar para probar la hipótesis nula

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

contra la hipótesis alternativa

$$H_a: \text{Al menos una de las medias es diferente de las otras}$$

usando el siguiente argumento teórico:

- Recuerde que σ^2 es la varianza común para todas las poblaciones k . La cantidad

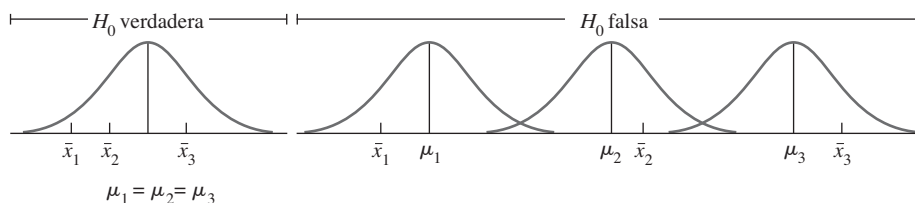
$$MSE = \frac{SSE}{n - k}$$

es una estimación agrupada de σ^2 , un promedio ponderado de todas las k varianzas muestrales, sea H_0 verdadera o no.

- Si H_0 es verdadera, entonces la variación en las medias muestrales, medida por $MST = [SST/(k - 1)]$ también da una estimación insesgada de σ^2 . No obstante, si H_0 es falsa y las medias poblacionales son diferentes, entonces MST , que mide la variación en las medias muestrales, será inusualmente *grande*, como se ve en la figura 7.4.

FIGURA 7.4

Medias muestrales sacadas de poblaciones idénticas contra diferentes



TIP MI CONSEJO

Las pruebas F para tablas ANOVA son **siempre** de cola superior (derecha)

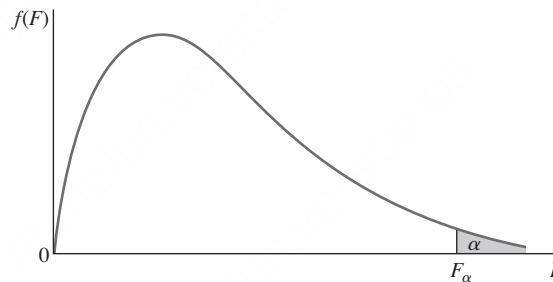
- La prueba estadística

$$F = \frac{MST}{MSE}$$

tiende a ser más grande que lo normal si H_0 es falsa. En consecuencia, se puede rechazar H_0 para valores grandes de F , usando una prueba estadística de *cola derecha*. Cuando H_0 es verdadera, esta prueba estadística tiene una distribución F con $gl_1 = (k - 1)$ y $gl_2 = (n - k)$ grados de libertad y se pueden usar valores críticos de *cola derecha* de la distribución F (de la tabla 6 del Anexo) o valores p generados por computadora, para sacar conclusiones estadísticas acerca de la igualdad de las medias poblacionales.

PRUEBA F PARA COMPARAR K MEDIAS POBLACIONALES

1. Hipótesis nula: $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$
2. Hipótesis alternativa: H_a : Uno o más pares de medias poblacionales difieren
3. Prueba estadística: $F = MST/MSE$, donde F está basada en $gl_1 = (k - 1)$ y $gl_2 = (n - k)$
4. Región de rechazo: rechazar H_0 si $F > F_\alpha$, donde F_α se encuentra en la cola superior de la distribución F (con $gl_1 = k - 1$ y $gl_2 = n - k$) o si el valor $p < \alpha$.



- Las muestras son seleccionadas al azar y en forma independiente de sus respectivas poblaciones.
- Las poblaciones están normalmente distribuidas con $\mu_1, \mu_2, \dots, \mu_k$ y varianzas iguales, $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$.

EJEMPLO

7.5

¿Los datos del ejemplo 7.4 dan suficiente evidencia para indicar una diferencia en el promedio de intervalos de atención, dependiendo del tipo de desayuno tomado por el estudiante?

Solución Para probar $H_0: \mu_1 = \mu_2 = \mu_3$ contra la hipótesis alternativa de que el promedio de intervalo de atención es diferente para al menos uno de los tres tratamientos, se usa el estadístico F de análisis de varianza, calculado como

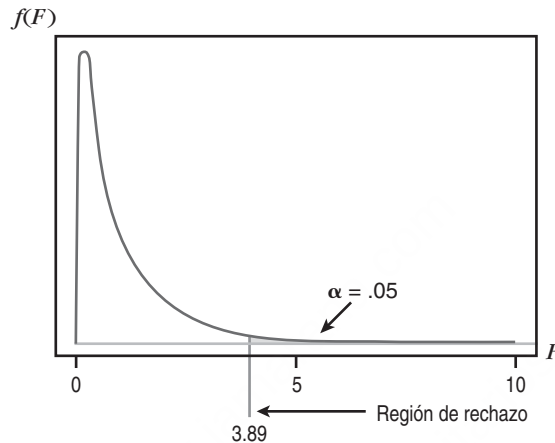
$$F = \frac{MST}{MSE} = \frac{29.2667}{5.9333} = 4.93$$

y se muestra en la columna marcada “F” de la figura 7.3a) y “F” de la figura 7.3b). No es de sorprender saber que el valor en la columna marcada “P” en la figura 7.3a) y “valor P” en la figura 11.3b) sea el valor p exacto para esta prueba estadística.

La prueba estadística MST/MSE calculada líneas antes, tiene una distribución F con $gl_1 = 2$ y $gl_2 = 12$ grados de libertad. Con el uso del método del valor crítico con $\alpha = .05$, se puede rechazar H_0 si $F > F_{.05} = 3.89$ de la tabla 6 del Anexo (véase la figura 7.5). Como el valor observado, $F = 4.93$, excede del valor crítico, se rechaza H_0 . Hay suficiente evidencia para indicar que al menos uno de los tres intervalos de atención promedio es diferente de al menos uno de los otros.

FIGURA 7.5

Región de rechazo para el ejemplo 7.5



TIP MI CONSEJO

Las salidas impresas de computadora dan el valor p exacto; use el valor p para tomar su decisión

Se podría haber llegado a esta misma conclusión usando el valor p exacto, $P = .027$, dado en la figura 7.3. Como el valor p es menor que $\alpha = .05$, los resultados son estadísticamente significativos al nivel de 5%. Todavía se concluye que al menos uno de los tres promedios de intervalos de atención es diferente de al menos uno de los otros.

Estimación de diferencias en las medias de tratamiento

La siguiente pregunta obvia que se podría hacer se refiere a la naturaleza de las diferencias de las medias poblacionales. ¿Cuáles medias son diferentes de las otras? ¿Cómo se puede estimar la diferencia o posiblemente las medias individuales para cada uno de los tres tratamientos? En la sección 7.6, presentaremos un procedimiento que se puede usar para comparar todos los posibles pares de medias de tratamiento simultáneamente. No obstante, si hay un interés especial en una media particular o par de medias, se pueden construir intervalos de confianza usando los procedimientos de muestra pequeña del capítulo 6, con base en la distribución t de Student. Para una sola media poblacional, μ_i , el intervalo de confianza es

$$\bar{x}_i \pm t_{\alpha/2} \left(\frac{s}{\sqrt{n_i}} \right)$$

donde $\bar{x}_i$ es la media muestral para el i -ésimo tratamiento. Del mismo modo, para una comparación de dos medias poblacionales, por ejemplo μ_i y μ_j , el intervalo de confianza es

$$(\bar{x}_i - \bar{x}_j) \pm t_{\alpha/2} \sqrt{s^2 \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

Antes de que se puedan usar estos intervalos de confianza, sin embargo, persisten dos preguntas:

- ¿Cómo se calcula s o s^2 , la mejor estimación de la varianza común σ^2 ?
- ¿Cuántos grados de libertad se usan para el valor crítico de t ?

Para contestar estas preguntas, recuerde que en un análisis de varianza, el cuadrático medio de error, MSE, siempre da un estimador insesgado de σ^2 y usa información de todo el conjunto de mediciones. En consecuencia, es el mejor estimador disponible de σ^2 , cualquiera que sea el procedimiento de prueba o estimación que se use. *Siempre* se debe usar

$$s^2 = \text{MSE} \quad \text{con } gl = (n - k)$$

para estimar σ^2 . Se puede hallar la raíz cuadrada positiva de este estimador, $s = \sqrt{\text{MSE}}$, en el último renglón de la figura 7.3a) marcado “Pooled StDev”.

DISEÑO COMPLETAMENTE ALEATORIZADO: INTERVALOS DE CONFIANZA $(1 - \alpha)100\%$ PARA UNA SOLA MEDIA DE TRATAMIENTO Y LA DIFERENCIA ENTRE DOS MEDIAS DE TRATAMIENTO

Una sola media de tratamiento:

$$\bar{x}_i \pm t_{\alpha/2} \left(\frac{s}{\sqrt{n_i}} \right)$$

Diferencia entre dos medias de tratamiento:

$$(\bar{x}_i - \bar{x}_j) \pm t_{\alpha/2} \sqrt{s^2 \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

con

$$s = \sqrt{s^2} = \sqrt{\text{MSE}} = \sqrt{\frac{\text{SSE}}{n - k}}$$

donde $n = n_1 + n_2 + \dots + n_k$ y $t_{\alpha/2}$ está basada en $(n - k)$ gl.

TIP MI CONSEJO

Los grados de libertad para intervalos de confianza son los gl de error

EJEMPLO

7.6

El investigador del ejemplo 7.4 cree que los estudiantes que no toman desayuno tendrán intervalos de atención significativamente más cortos, pero que puede no haber diferencia entre aquellos que toman un desayuno ligero o un desayuno completo. Encuentre un intervalo de confianza de 95% para el promedio de intervalo de atención para estudiantes que no toman desayuno, así como un intervalo de confianza para la diferencia en los intervalos de atención promedio para quienes toman desayuno ligero contra los de desayuno completo.

Solución Para $s^2 = \text{MSE} = 5.9333$ para que $s = \sqrt{5.9333} = 2.436$ con $gl = (n - k) = 12$, se pueden calcular los dos intervalos de confianza:

- Para sin desayuno:

$$\bar{x}_1 \pm t_{\alpha/2} \left(\frac{s}{\sqrt{n_1}} \right)$$

$$9.4 \pm 2.179 \left(\frac{2.436}{\sqrt{5}} \right)$$

$$9.4 \pm 2.37$$

o sea, entre 7.03 y 11.77 minutos.

- Para los de desayuno ligero contra desayuno completo:

$$(\bar{x}_2 - \bar{x}_3) \pm t_{\alpha/2} \sqrt{s^2 \left(\frac{1}{n_2} + \frac{1}{n_3} \right)}$$

$$(14 - 13) \pm 2.179 \sqrt{5.9333 \left(\frac{1}{5} + \frac{1}{5} \right)}$$

$$1 \pm 3.36$$

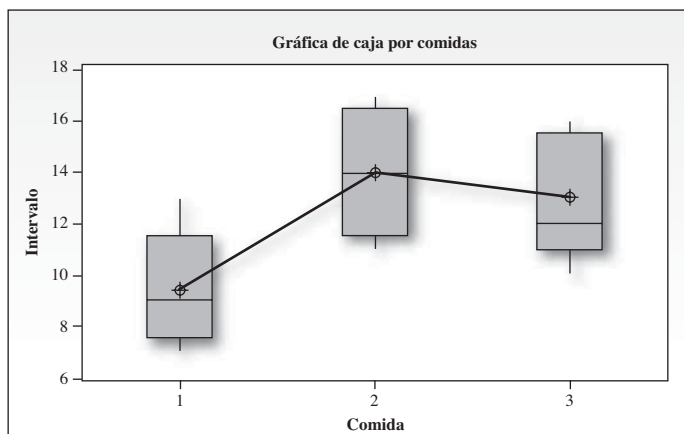
una diferencia de entre -2.36 y 4.36 minutos.

Se puede ver que el segundo intervalo de confianza no indica una diferencia en el promedio de intervalos de atención, para los estudiantes que tomaron un desayuno ligero contra los de desayuno completo, como sospechaba el investigador. Si el investigador, debido a creencias previas, desea probar los otros dos posibles pares de medias, es decir, sin desayuno contra desayuno ligero y sin desayuno contra desayuno completo, los métodos que se proporcionan en la sección 7.6 deben usarse para probar los tres pares.

Algunos programas de computadora tienen opciones de gráficas que dan una excelente descripción visual de datos y las k medias de tratamiento. Una de estas opciones en el programa *MINITAB* se ve en la figura 7.6. Las medias de tratamiento están indicadas por el símbolo $\oplus$ y están conectadas con rectas. Observe que la media “sin desayuno” parece ser un poco diferente de las otras dos medias, como sospechaba el investigador, aun cuando hay un pequeño traslape en las gráficas de caja. En la siguiente sección presentamos un procedimiento formal para probar la significancia de las diferencias entre todos los pares de medias de tratamiento.

FIGURA 7.6

Gráficas de caja para el ejemplo 7.6





NECESITO SABER...

Cómo determinar si mis cálculos son precisos

Las siguientes sugerencias se aplican a todos los análisis de varianza de este capítulo:

1. Cuando calcule sumas de cuadrados, asegúrese de llevar al menos seis cifras significativas antes de hacer restas.
2. Recuerde, las sumas de cuadrados nunca pueden ser negativas. Si obtiene una suma de cuadrados negativa, ha cometido un error aritmético.
3. Siempre verifique su tabla de análisis de varianza para comprobar que los grados de libertad ascienden al total de grados de libertad ($n - 1$) y que las sumas de cuadrados ascienden al TSS.

7.5 EJERCICIOS

TÉCNICAS BÁSICAS

7.1 Suponga que se desea comparar las medias de seis poblaciones basadas en muestras aleatorias independientes, cada una de las cuales contiene 10 observaciones. Inserte, en una tabla ANOVA, las fuentes de variación y sus respectivos grados de libertad.

7.2 Los valores de TSS y SSE para el experimento del ejercicio 7.1 son $TSS = 21.4$ y $SSE = 16.2$.

- a. Complete la tabla ANOVA para el ejercicio 7.1.
- b. ¿Cuántos grados de libertad están asociados con el estadístico F para probar $H_0 : \mu_1 = \mu_2 = \dots = \mu_6$?
- c. Dé la región de rechazo para la prueba de la parte b para $\alpha = .05$.
- d. ¿Los datos dan suficiente evidencia para indicar diferencias entre las medias poblacionales?
- e. Estime el valor p para la prueba. ¿Este valor confirma sus conclusiones de la parte d?

7.3 Las medias muestrales correspondientes a las poblaciones 1 y 2 del ejercicio 7.1 son $\bar{x}_1 = 3.07$ y $\bar{x}_2 = 2.52$.

- a. Encuentre un intervalo de confianza de 95% para μ_1 .
- b. Encuentre un intervalo de confianza de 95% para la diferencia ($\mu_1 - \mu_2$).

7.4 Suponga que se desea comparar las medias de cuatro poblaciones basadas en muestras aleatorias independientes, cada una de las cuales contiene seis observaciones. Inserte, en una tabla ANOVA, las fuentes de variación y sus respectivos grados de libertad.

7.5 Los valores de TSS y SST para el experimento del ejercicio 7.4 son $TSS = 473.2$ y $SST = 339.8$.

- a. Complete la tabla ANOVA para el ejercicio 7.4.
- b. ¿Cuántos grados de libertad están asociados con el estadístico F para probar $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$?
- c. Dé la región de rechazo para la prueba de la parte b para $\alpha = .05$.
- d. ¿Los datos dan suficiente evidencia para indicar diferencias entre las medias poblacionales?
- e. Aproxime el valor p para la prueba. ¿Esto confirma sus conclusiones de la parte d?

7.6 Las medias muestrales correspondientes a las poblaciones 1 y 2 del ejercicio 7.4 son $\bar{x}_1 = 88.0$ y $\bar{x}_2 = 83.9$.

- a. Encuentre un intervalo de confianza de 90% para μ_1 .
- b. Encuentre un intervalo de confianza de 90% para la diferencia ($\mu_1 - \mu_2$).

DATOS

EX0707

7.7 Estos datos son observaciones recolectadas usando un diseño completamente aleatorizado:

Muestra 1	Muestra 2	Muestra 3
3	4	2
2	3	0
4	5	2
3	2	1
2	5	

- a. Calcule CM y TSS.
- b. Calcule SST y MST.
- c. Calcule SSE y MSE.
- d. Construya una tabla ANOVA para los datos.

- e. Exprese las hipótesis nula y alternativa para un análisis de prueba F de varianza.
- f. Use el método del valor p para determinar si hay una diferencia en las tres medias poblacionales.

7.8 Consulte el ejercicio 7.7 y el conjunto de datos EX0707. ¿Los datos dan suficiente evidencia para indicar una diferencia entre μ_2 y μ_3 ? Pruebe usando la prueba t de la sección 6.4 con $\alpha = .05$.

7.9 Consulte el ejercicio 7.7 y el conjunto de datos EX0707.

- a. Encuentre un intervalo de confianza de 90% para μ_1 .
- b. Encuentre un intervalo de confianza de 90% para la diferencia ($\mu_1 - \mu_3$).

APLICACIONES

DATOS
EX0710

7.10 Reducir hostilidad Un psicólogo clínico deseaba comparar tres métodos para reducir niveles de hostilidad en estudiantes universitarios, con el uso de cierto examen psicológico (HLT). Se tomaron las calificaciones altas de este examen como indicio de gran hostilidad. Once estudiantes que obtuvieron calificaciones altas y casi iguales se emplearon en el experimento. Cinco fueron seleccionados al azar de entre los 11 estudiantes y tratados con el método A, tres se tomaron al azar de los seis estudiantes restantes y se trataron con el método B, y los otros tres estudiantes con el método C. Todos los tratamientos continuaron durante todo un semestre, cuando el examen HLT se aplicó de nuevo. Los resultados se muestran en la tabla.

Método	Calificaciones en el examen HLT				
A	73	83	76	68	80
B	54	74	71		
C	79	95	87		

- a. Realice un análisis de varianza para este experimento.
- b. ¿Los datos dan suficiente evidencia para indicar una diferencia en respuesta media de estudiantes a los tres métodos después del tratamiento?

7.11 Hostilidad, continúa Consulte el ejercicio 7.10. Con μ_A y μ_B , respectivamente, denote las calificaciones medias al final del semestre para las poblaciones de estudiantes extremadamente hostiles que fueron tratados en todo ese semestre por el método A y el método B.

- a. Encuentre un intervalo de confianza de 95% para μ_A .
- b. Encuentre un intervalo de confianza de 95% para μ_B .
- c. Encuentre un intervalo de confianza de 95% para ($\mu_A - \mu_B$).
- d. ¿Es correcto decir que los intervalos de confianza hallados en las partes a, b y c son conjuntamente válidos?

DATOS
EX0712

7.12 Ensamblaje de equipo electrónico Se realizó un experimento para comparar la eficacia

de tres programas de capacitación, A, B y C, para capacitar ensambladores de una pieza de equipo electrónico. Quince empleados se asignaron al azar, cinco en cada uno, a los tres programas. Después de terminar los cursos, a cada persona se le pidió ensamblar cuatro piezas de equipo y se registró el promedio de tiempo necesario para completar el ensamble. Varios de los empleados renunciaron durante el curso del programa; el resto fueron evaluados, produciendo los datos que se ven en la tabla siguiente. Use la salida impresa de *Excel* para contestar las preguntas.

Programa de capacitación	Tiempo promedio de ensamblaje (min)				
A	59	64	57	62	
B	52	58	54		
C	58	65	71	63	64

- a. ¿Los datos dan suficiente evidencia para indicar una diferencia en tiempos medios de ensamble para personal capacitado por los tres programas? Dé el valor p para el examen e interprete su valor.
- b. Encuentre un intervalo de confianza de 99% para la diferencia en tiempos medios de ensamble para personas capacitadas en los programas A y B.
- c. Encuentre un intervalo de confianza de 99% para los tiempos medios de ensamble para personas capacitadas en el programa A.
- d. ¿Piensa usted que los datos satisfacen (aproximadamente) la suposición de que han sido seleccionados de poblaciones normales? ¿Por qué?

Salida de *MS Excel* para el ejercicio 7.12

Anova: factor único

RESUMEN						
Grupos	Conteo	Suma	Promedio	Varianza		
A	4	242	60.5	9.667		
B	3	164	54.667	9.333		
C	5	321	64.2	21.7		
ANOVA						
Fuente de variación	SS	df	MS	F	Valor P	F crít.
Entre grupos	170.45	2	85.225	5.704	0.0251	4.256
Dentro de grupos	134.467	9	14.941			
Total	304.917	11				

DATOS
EX0713

7.13 Sitios pantanosos Se realizó un estudio ecológico para comparar los porcentajes de crecimiento de vegetación en cuatro lugares pantanosos sin urbanizar, así como para determinar la causa de cualesquiera diferencias que pudieran observarse. Parte del estudio comprendía medir la longitud de las hojas de una especie particular de plantas en una fecha preseleccionada en mayo. Seis plantas se seleccionaron al azar en cada uno de los cuatro lugares para usarlas en la comparación. Los datos de la tabla son la longitud media de hoja por planta (en centímetros) para una muestra aleatoria de 10 hojas por planta.

También se proporciona la salida de computadora impresa de *MINITAB* del análisis de varianza para estos datos.

Lugar	Longitud media de hoja (cm)					
1	5.7	6.3	6.1	6.0	5.8	6.2
2	6.2	5.3	5.7	6.0	5.2	5.5
3	5.4	5.0	6.0	5.6	4.9	5.2
4	3.7	3.2	3.9	4.0	3.5	3.6

Salida *MINITAB* para el ejercicio 7.13

ANOVA en una dirección: longitud contra lugar

Source	DF	SS	MS	F	P
Location	3	19.740	6.580	57.38	0.000
Error	20	2.293	0.115		
Total	23	22.033			

S = 0.3386 R-Sq = 89.59% R-Sq(adj) = 88.03%

Individual 95% CIs For Mean
Based on Pooled StDev

Level	N	Mean	StDev
1	6	6.0167	0.2317
2	6	5.6500	0.3937
3	6	5.3500	0.4087
4	6	3.6500	0.2881

Pooled StDev = 0.3386

- Usted recordará que los procedimientos de prueba y estimación, para un análisis de varianza requieren que las observaciones se seleccionen de poblaciones normalmente distribuidas (al menos aproximadas). ¿Por qué podría sentir confianza razonable de que sus datos satisfacen esta suposición?
- ¿Los datos dan suficiente evidencia para indicar una diferencia en longitud media de hojas entre los cuatro lugares? ¿Cuál es el valor p para la prueba?
- Suponga, antes de ver los datos, que usted decidió comparar las longitudes medias de hojas de los lugares 1 y 4. Pruebe la hipótesis nula $\mu_1 = \mu_4$ contra la alternativa $\mu_1 \neq \mu_4$.
- Consulte la parte c. Construya un intervalo de confianza de 99% para $(\mu_1 - \mu_4)$.
- En lugar de usar un análisis de varianza de prueba F parecería más sencillo examinar los datos de uno, seleccionar los dos lugares que tengan las longitudes medias muestrales más corta y más larga, y a continuación comparar estas dos medias usando la prueba t de Student. Si hay evidencia para indicar una diferencia en estas medias, hay claramente evidencia de una diferencia entre las cuatro. (Si se usara esta lógica, no habría necesidad del análisis de varianza de prueba F .) Explique por qué este procedimiento es inválido.

DATOS
EX0714

7.14 Contenido de O_2 disuelto Se tomaron muestras de agua de un río en cuatro lugares diferentes para determinar si la cantidad de oxígeno disuelto, una medida de la contaminación del agua, variaba de un lugar a otro. Los lugares 1 y 2 se seleccionaron arriba de una planta industrial, una cerca de la orilla y la otra a mitad del río; el lugar 3 estaba

adyacente a la descarga del agua industrial para la planta; y el lugar 4 estaba ligeramente agua abajo a mitad del río. Cinco especímenes de agua se seleccionaron al azar en cada lugar, pero un espécimen, correspondiente al lugar 4, se perdió en el laboratorio. Los datos y una salida impresa de computadora de *MS Excel* del análisis de varianza se dan a continuación (a mayor contaminación, lecturas más bajas de oxígeno disuelto).

Lugar	Contenido medio de oxígeno disuelto				
1	5.9	6.1	6.3	6.1	6.0
2	6.3	6.6	6.4	6.4	6.5
3	4.8	4.3	5.0	4.7	5.1
4	6.0	6.2	6.1	5.8	

Salida de *MS Excel* para el ejercicio 7.14

Anova: factor único

RESUMEN

Grupos	Conteo	Suma	Promedio	Varianza
1	5	30.4	6.08	0.022
2	5	32.2	6.44	0.013
3	5	23.9	4.78	0.097
4	4	24.1	6.025	0.0292

ANOVA

Fuente de variación	SS	df	MS	F	Valor P	F crt.
Entre grupos	7.8361	3	2.6120	63.656	9E-09	3.287
Dentro de grupos	0.6155	15	0.0410			
Total	8.4516	18				

- ¿Los datos dan suficiente evidencia para indicar una diferencia en el contenido medio de oxígeno disuelto para los cuatro lugares?
- Compare el contenido medio de oxígeno disuelto a mitad del río arriba de la planta, con el contenido medio adyacente a la planta (lugar 2 contra lugar 3). Use un intervalo de confianza de 95%.

DATOS
EX0715

7.15 Calcio El contenido de calcio de una sustancia mineral pulverizada fue analizado cinco veces por cada uno de tres métodos, con desviaciones estándar similares:

Método	Porcentaje de calcio				
1	.0279	.0276	.0270	.0275	.0281
2	.0268	.0274	.0267	.0263	.0267
3	.0280	.0279	.0282	.0278	.0283

Use una prueba adecuada para comparar los tres métodos de medición. Comente sobre la validez de cualesquiera suposiciones que sea necesario hacer.

DATOS
EX0716

7.16 Atún En el ejercicio 6.6, informamos de los precios promedio estimados para una lata de 6 onzas o una bolsa de 7.06 onzas de atún, con base en precios pagados a nivel nacional por una variedad de marcas diferentes de atún.¹

Atún claro en agua	Atún blanco en aceite	Atún blanco en agua	Atún claro en aceite
.99	.53	1.27	1.49
1.92	1.41	1.22	1.29
1.23	1.12	1.19	1.27
.85	.63	1.22	1.35
.65	.67		1.29
.69	.60		1.00
.60	.66		1.27
			1.28

Fuente: De "Pricing of Tuna" Copyright 2001 por Consumers Union of U.S., Inc., Yonkers, NY 10703-1057, organización sin fines de lucro. Reimpreso con permiso de la edición de junio de 2001 de Consumer Reports® sólo para fines educativos. No se permite uso comercial ni reproducción. www.ConsumerReports.org®.

- Use un análisis de varianza para un diseño completamente aleatorizado, para determinar si hay diferencias significativas en los precios de atún empacados en estas cuatro formas. ¿Se puede rechazar la hipótesis de que no hay diferencia en el precio promedio para estos empaques al nivel de significancia de $\alpha = .05$? ¿Y al nivel de significancia de $\alpha = .01$?
- Encuentre una estimación de intervalo de confianza de 95% de la diferencia en precio entre atún claro en agua y atún claro en aceite. ¿Parece haber una diferencia significativa en el precio de estas dos clases de atún empacado?
- Encuentre una estimación de intervalo de confianza de 95% de la diferencia en precio entre atún blanco en agua y atún blanco en aceite. ¿Parece haber una diferencia significativa en el precio de estas dos clases de atún empacado?
- ¿Qué otros intervalos de confianza podrían ser de interés para el investigador que realizó el experimento?

DATOS
EX0717

7.17 El costo de la madera Un constructor de viviendas a nivel nacional desea comparar los precios por 1 000 pies de madera de armazones de abeto Douglas de calidad estándar o mejor. Al azar selecciona cinco proveedores en cada uno de los cuatro estados donde el constructor planea iniciar la construcción. Los precios se dan en la tabla siguiente.

Estado			
1 (\$)	2 (\$)	3 (\$)	4 (\$)
261	236	250	265
255	240	245	270
258	225	255	258
267	233	248	275
270	240	260	275

- ¿Qué tipo de diseño experimental se ha empleado?
- Construya la tabla del análisis de varianza para estos datos.
- ¿Los datos dan suficiente evidencia para indicar que el precio promedio por 1 000 pies de madera de abeto Douglas difiere entre los cuatro estados? Pruebe usando $\alpha = .05$.

DATOS
EX0718

7.18 ¿Bueno para las matemáticas? Veinte alumnos de tercer grado se separaron al azar en cuatro grupos iguales y a cada grupo se le impartieron conceptos matemáticos usando un método diferente de enseñanza, midiéndose el progreso mediante un examen unitario al final del periodo de enseñanza. Las calificaciones se muestran a continuación (un niño del grupo 3 estuvo ausente el día en que se aplicó el examen).

Grupo			
1	2	3	4
112	111	140	101
92	129	121	116
124	102	130	105
89	136	106	126
97	99		119

- ¿Qué tipo de diseño se ha empleado en este experimento?
- Construya una tabla ANOVA para el experimento.
- ¿Los datos presentan suficiente evidencia para indicar una diferencia en el promedio de calificaciones para los cuatro métodos de enseñanza? Pruebe usando $\alpha = .05$.

CLASIFICACIÓN DE MEDIAS POBLACIONALES

7.6

Numerosos experimentos son exploratorios por naturaleza. No hay nociones preconcebidas de los resultados y no se ha decidido (antes de realizar el experimento) hacer comparaciones específicas de tratamiento. En lugar de esto, se desea clasificar las medias de tratamiento, determinar cuáles medias difieren e identificar conjuntos de medias para las cuales no hay evidencia de diferencia.

Una opción podría ser ordenar las medias muestrales de menor a mayor y a continuación efectuar pruebas t para medias adyacentes en el ordenamiento. Si dos medias difieren en más de

$$t_{\alpha/2} \sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

se concluye que el par de medias poblacionales difiere. El problema con este procedimiento es que la probabilidad de hacer un error tipo I, es decir, concluir que dos medias difieren cuando en realidad son iguales, es α para cada prueba. Si se compara un gran número de pares de medias, la probabilidad de detectar al menos una diferencia en medias, cuando en realidad ninguna existe, es bastante grande.

Una forma sencilla de evitar el alto riesgo de declarar diferencias cuando no existen es usar el **rango estudentizado**, que es la diferencia entre la más pequeña y la más grande en un conjunto de k medias muestrales, como la medida para determinar si hay una diferencia en un par de medias poblacionales. Este método, a veces denominado **método de Tukey para comparaciones en pares**, hace que sea igual a α la probabilidad de declarar que existe una diferencia entre al menos un par en un conjunto de k medias de tratamiento, cuando no existe diferencia.

El método de Tukey para hacer comparaciones en pares está basado en el análisis usual de suposiciones de varianza. **Además, supone que las medias muestrales son independientes y están basadas en muestras de igual tamaño.** La medida que determina si existe una diferencia entre un par de medias de tratamiento es la cantidad ω (omega minúscula), que se presenta a continuación.

MEDIDA PARA HACER COMPARACIONES EN PARES

$$\omega = q_{\alpha}(k, gl) \left(\frac{s}{\sqrt{n_t}} \right)$$

donde

k = Número de tratamientos

s^2 = MSE = Estimador de la varianza común σ^2 y $s = \sqrt{s^2}$

gl = Número de grados de libertad para s^2

n_t = Tamaño muestral común, es decir, el número de observaciones en cada una de las k medias de tratamiento

$q_{\alpha}(k, dl)$ = Valor tabulado de las tablas 11a) y 11b) del Anexo, para $\alpha = .05$ y $.01$, respectivamente, y para varias combinaciones de k y gl

Regla: Se juzga que dos medias poblacionales difieren si las medias muestrales correspondientes difieren en ω o más.

Las tablas 11a) y 11b) del Anexo contienen los valores de $q_{\alpha}(k, gl)$ para $\alpha = .05$ y $.01$, respectivamente. Para ilustrar el uso de las tablas, consulte la parte de la tabla 11a) reproducida en la tabla 11.2. Suponga que se desea hacer comparaciones por pares de $k = 5$ medias con $\alpha = .05$ para un análisis de varianza, donde s^2 posee 9 gl . El valor tabulado para $k = 5$, $gl = 9$ y $\alpha = .05$, sombreados en la tabla 11.2, es $q_{.05}(5, 9) = 4.76$.

TABLA 7.2

Reproducción parcial de la tabla 11a) del Anexo; puntos de 5% superiores

<i>gl</i>	2	3	4	5	6	7	8	9	10	11	12
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59	51.96
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39	14.75
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72	9.95
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03	8.21
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65	6.79
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72	5.83
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51	5.61

EJEMPLO

7.7

Consulte el ejemplo 7.4, en el que se comparó el promedio de intervalos de atención para estudiantes a los que se dieron diferentes tratamientos de “comida” en la mañana: sin desayuno, un desayuno ligero o un desayuno completo. La prueba F de ANOVA del ejemplo 11.5 indicó una diferencia significativa en las medias poblacionales. Use el método de Tukey para comparaciones en pares, para determinar cuál de las tres medias poblacionales difiere de las otras.

Solución Para este ejemplo, hay $k = 3$ medias de tratamiento, con $s = \sqrt{\text{MSE}} = 2.436$. El método de Tukey se puede usar, con cada una de las tres muestras conteniendo $n_i = 5$ mediciones y $(n - k) = 12$ grados de libertad. Consulte la tabla 11 en el Anexo para hallar $q_{.05}(k, gl) = q_{.05}(3, 12) = 3.77$ y calcule la “regla” como

$$\omega = q_{.05}(3, 12) \left(\frac{s}{\sqrt{n_i}} \right) = 3.77 \left(\frac{2.436}{\sqrt{5}} \right) = 4.11$$

Las tres medias de tratamiento están dispuestas en orden de la más pequeña, 9.4, a la máxima, 14.0, en la figura 7.7. El siguiente paso es comprobar la diferencia entre cada par de medias. La única diferencia que excede de $\omega = 4.11$ es la diferencia entre no desayuno y un desayuno ligero. Estos dos tratamientos se declaran por tanto, significativamente diferentes. No se puede declarar una diferencia entre los otros dos pares de tratamientos. Para indicar este hecho visualmente, la figura 7.7 muestra una línea bajo los pares de medias que no son significativamente diferentes.

FIGURA 7.7

Medias clasificadas para el ejemplo 7.7

Sin	Completo	Ligero
9.4	13.0	14.0

Estos resultados pueden parecer confusos, pero por lo general ayuda considerar la clasificación de las medias e interpretar diferencias no significativas como nuestra incapacidad para clasificar de manera distintiva las medias subrayadas por la misma línea. Para este ejemplo, el desayuno ligero definitivamente clasificó más alto que sin desayuno, pero el desayuno completo no pudo ser clasificado más alto que sin desayuno, o más abajo que el desayuno ligero. La probabilidad de que cometimos al menos un error entre las tres comparaciones es a lo sumo $\alpha = .05$.

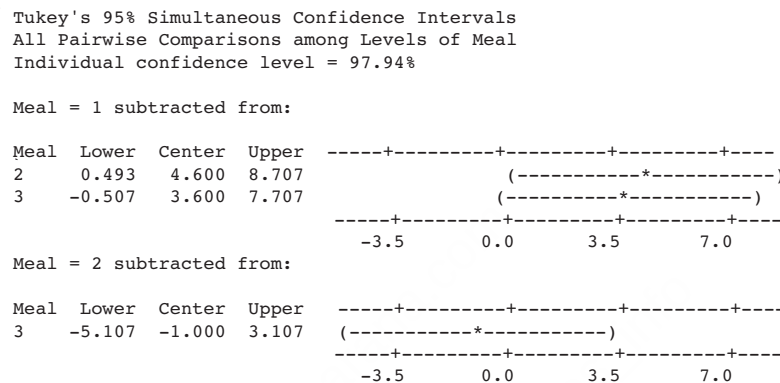
TIP MI CONSEJO

Si cero no está en el intervalo, hay evidencia de una diferencia entre los dos métodos

Muchos programas de computadora dan una opción para realizar **comparaciones en pares**, incluyendo el método de Tukey, aunque *MS Excel* no la tiene. La salida *MINITAB* de la figura 7.8 muestra su forma de la prueba de Tukey, que difiere ligeramente del método que hemos presentado. Los tres intervalos que se ven en la salida impresa marcada “Inferior” y “Superior” representan la diferencia en las dos medias muestrales más o menos la regla ω . Si el intervalo contiene el valor 0, las dos medias se juzgan como no significativamente diferentes. Se puede ver que las únicas medias 1 y 2 (sin desayuno contra desayuno ligero) muestran una diferencia significativa.

FIGURA 7.8

Salida *MINITAB* para el ejemplo 7.7



Cuando usted estudie dos diseños experimentales más en las siguientes secciones de este capítulo recuerde que, una vez que haya encontrado que un factor es significativo, debe usar el método de Tukey u otro método de comparaciones en pares para averiguar exactamente dónde están las diferencias.

7.6 EJERCICIOS**TÉCNICAS BÁSICAS**

7.19 Suponga que usted desea usar el método de Tukey de comparaciones en pares para clasificar un conjunto de medias poblacionales. Además del análisis de suposiciones de varianza, ¿qué otra propiedad deben satisfacer las medias de tratamiento?

7.20 Consulte las tablas 11a) y 11b) del Anexo y encuentre los valores de $q_\alpha(k, gl)$ para estos casos:

- $\alpha = .05, k = 5, gl = 7$
- $\alpha = .05, k = 3, gl = 10$
- $\alpha = .01, k = 4, gl = 8$
- $\alpha = .01, k = 7, gl = 5$

7.21 Si el tamaño muestral para cada tratamiento es n_i y si s^2 está basada en 12 *gl*, encuentre ω en estos casos:

- $\alpha = .05, k = 4, n_i = 5$
- $\alpha = .01, k = 6, n_i = 8$

7.22 Un diseño de muestreo aleatorio independiente se utilizó para comparar las medias de seis tratamientos basados en muestras de cuatro observaciones por

tratamiento. El estimador agrupado de σ^2 es 9.12, y las medias muestrales siguen:

$$\begin{array}{lll} \bar{x}_1 = 101.6 & \bar{x}_2 = 98.4 & \bar{x}_3 = 112.3 \\ \bar{x}_4 = 92.9 & \bar{x}_5 = 104.2 & \bar{x}_6 = 113.8 \end{array}$$

- Dé el valor de ω que usaría para hacer comparaciones por pares de las medias de tratamiento para $\alpha = .05$.
- Ordene las medias de tratamiento usando comparaciones por pares.

APLICACIONES

7.23 Sitios pantanosos, otra vez Consulte el ejercicio 7.13 y el conjunto de datos EX0713. Ordene el crecimiento medio de hojas para los cuatro lugares. Use $\alpha = .01$.

7.24 Calcio Consulte el ejercicio 7.15 y el conjunto de datos EX0715. La opción de comparaciones en pares en *MINITAB* generó la salida impresa siguiente. ¿Qué indican estos resultados acerca de las diferencias en las medias poblacionales? ¿Esto confirma sus conclusiones del ejercicio 7.15?

Salida MINITAB para el ejercicio 7.24

Tukey's 95% Simultaneous Confidence Intervals
All Pairwise Comparisons among Levels of Method
Individual confidence level = 97.94%

Method = 1 subtracted from:

Method	Lower	Center	Upper
2	-0.0014377	-0.0008400	-0.0002423
3	-0.0001777	0.0004200	0.0010177

Method

2	(-----*-----)
3	(-----*-----)

-----+-----+-----+-----+
-0.0010 0.0000 0.0010 0.0020

Method = 2 subtracted from:

Method	Lower	Center	Upper
3	0.0006623	0.0012600	0.0018577

Method

3	(-----*-----)

-----+-----+-----+-----+
-0.0010 0.0000 0.0010 0.0020

DATOS
EX0725

7.25 Tolerancia a la glucosa Los médicos dependen de resultados de exámenes de laboratorio cuando manejan problemas médicos como diabetes o epilepsia. En un examen de uniformidad para tolerancia a la glucosa, a tres laboratorios diferentes se les enviaron $n_i = 5$ muestras de sangre idénticas de una persona que había bebido 50 miligramos (mg) de glucosa disuelta en agua. Los resultados de laboratorio (en mg/dl) son los siguientes:

Lab 1	Lab 2	Lab 3
120.1	98.3	103.0
110.7	112.1	108.5
108.9	107.7	101.1
104.2	107.9	110.0
100.4	99.2	105.4

- ¿Los datos indican una diferencia en el promedio de lecturas para los tres laboratorios?
- Use el método de Tukey para comparaciones en pares para clasificar las tres medias de tratamiento. Use $\alpha = .05$.

7.26 El costo de la madera, continúa El análisis de varianza de la prueba F en el ejercicio 7.17 (y conjunto

de datos EX0717) determinó que, en efecto, había una diferencia en el costo promedio de madera para los cuatro estados. La siguiente información del ejercicio 7.17 se da en la tabla.

Medias muestrales $\bar{x}_1 = 262.2$	MSE	41.25
$\bar{x}_2 = 234.8$	Error gl :	16
$\bar{x}_3 = 251.6$	n_i :	5
$\bar{x}_4 = 268.6$	k :	4

Use el método de Tukey para comparaciones en pares cuando tenga que determinar cuáles medias difieren significativamente respecto de las otras al nivel $\alpha = .01$.

DATOS
EX0727

7.27 Calificaciones del GRE Las calificaciones de razonamiento cuantitativo del Examen de Registro de Graduados (GRE, por sus siglas en inglés),² se registraron para estudiantes admitidos en tres programas diferentes para graduados en una universidad local.

Programa para graduados					
Ciencias de la vida	Ciencias físicas	Ciencias sociales			
630	660	660	760	440	540
640	660	640	670	330	450
470	480	720	700	670	570
600	650	690	710	570	530
580	710	530	450	590	630

- ¿Los datos dan suficiente evidencia para indicar una diferencia en las calificaciones medias del GRE, para solicitantes admitidos a los tres programas?
- Encuentre un intervalo de confianza de 95% para la diferencia en calificaciones medias del GRE para Ciencias de la vida y Ciencias físicas.
- Si encuentra una diferencia significativa en las calificaciones promedio del GRE para los tres programas, use el método de Tukey para comparaciones en pares para determinar cuáles medias difieren significativamente de las otras. Use $\alpha = .05$.

DISEÑO DE BLOQUE ALEATORIZADO: UNA CLASIFICACIÓN EN DOS DIRECCIONES

7.7

El *diseño completamente aleatorizado* que se estudió en la sección 7.4 es una generalización del diseño de las *dos muestras independientes* presentado en la sección 6.4. Está pensado para usarse cuando las unidades experimentales sean bastante similares u *homogéneas* en su conformación y cuando haya sólo un factor, el *tratamiento*, que pueda influir en la respuesta.

Cualquier otra variación en la respuesta se debe a variación aleatoria o a *error experimental*. A veces es claro para el investigador que las unidades experimentales *no sean homogéneas*. Personas o animales con carácter experimental, campos agrícolas, días de la semana y otras unidades experimentales a veces agregan su propia variabilidad a la respuesta. Aun cuando el investigador no esté realmente interesado en esta fuente de variación, sino que más bien en algún *tratamiento* que elija aplicar, puede aumentar la información al aislar esta fuente de variación usando el **diseño de bloque aleatorizado**, que es una extensión directa de los *pares acoplados* o *diseño de diferencia en pares* de la sección 6.5.

En un diseño de bloque aleatorizado, el experimentador está interesado en comparar k medias de tratamiento. El diseño utiliza *bloques* de k unidades experimentales que son relativamente similares, u *homogéneos*, con una unidad dentro de cada bloque asignada *al azar* a cada tratamiento. Si el diseño de bloque aleatorizado contiene k tratamientos dentro de cada uno de los b bloques, entonces el número total de observaciones en el experimento es $n = bk$.

Un supervisor de producción desea comparar los tiempos medios para que operadores de línea de producción ensamblen un artículo usando uno de tres métodos: A, B o C. Esperando variación en tiempos de ensamble de un operador a otro, el supervisor emplea un diseño de bloque aleatorizado para comparar los tres métodos. Cinco operadores de línea de producción se seleccionan para servir como bloques y a cada uno se le asigna ensamblar el artículo tres veces, una vez por cada uno de los tres métodos. Como la secuencia en la que el operador usa los tres métodos puede ser importante (fatiga o mayor destreza pueden ser factores que afecten la respuesta), cada operador debe ser asignado a una secuencia aleatoria de los tres métodos. Por ejemplo, el operador 1 podría ser asignado a efectuar primero el método C, seguido por A y B. El operador 2 podría realizar primero el método A, seguido del C y B, etcétera.

Para comparar cuatro métodos diferentes de enseñanza, un grupo de estudiantes podría ser dividido en bloques de tamaño 4, de modo que los grupos se encuentren casi *acoplados* según su rendimiento académico. Para comparar los costos promedio de tres compañías de telefonía celular diferentes, los costos podrían compararse en cada uno de tres niveles de uso: bajo, medio y alto. Para comparar el promedio de producción de tres especies de árboles frutales, cuando se espera una variación en producción debido al campo en el que se planten los árboles, una investigadora usa cinco campos. Ella divide cada campo en tres *lotes* en los que están plantadas las tres especies de árboles frutales.

El acoplamiento o *bloqueo* puede tener lugar en muchas formas diferentes. Las comparaciones de tratamientos se hacen a veces dentro de bloques de tiempo, dentro de bloques de personas o dentro de ambientes externos similares. El propósito de bloquear es remover o aislar la variabilidad *de un bloque a otro* que pudiera de otro modo ocultar el efecto de los tratamientos. Usted encontrará más ejemplos del uso del diseño de bloque aleatorizado en los ejercicios al final de la siguiente sección.

TIP MI CONSEJO

b = bloques

k = tratamientos

$n = bk$

EL ANÁLISIS DE VARIANZA PARA UN DISEÑO DE BLOQUE ALEATORIZADO

7.8

El diseño de bloque aleatorizado identifica dos factores: **tratamientos** y **bloques**, los cuales afectan la respuesta.

División de la variación total en el experimento

Sea x_{ij} la respuesta cuando el i -ésimo tratamiento ($i = 1, 2, \dots, k$) se aplica en el j -ésimo bloque ($j = 1, 2, \dots, b$). La variación total en las $n = bk$ observaciones es

$$\text{TSS} = \sum (x_{ij} - \bar{x})^2 = \sum x_{ij}^2 - \frac{(\sum x_{ij})^2}{n}$$

Esto se divide en *tres* (en lugar de dos) partes de modo que

$$TSS = SSB + SST + SSE$$

donde

- SSB (suma de cuadrados para bloques) mide la variación entre las medias de bloque.
- SST (suma de cuadrados para tratamientos) mide la variación entre las medias de tratamiento.
- SSE (suma de cuadrados para error) mide la variación de las diferencias entre las observaciones de tratamiento *dentro* de bloques, que mide el error experimental.

Las fórmulas de cálculo para las cuatro sumas de cuadrados son semejantes en forma a las empleadas para el diseño completamente aleatorizado de la sección 7.5. Aun cuando se puede simplificar el trabajo con el uso de un programa computarizado para calcular estas sumas de cuadrados, las fórmulas se dan a continuación.

TIP MI CONSEJO

TSS = SST +
SSB + SSE

CÁLCULO DE LAS SUMAS DE CUADRADOS PARA UN DISEÑO DE BLOQUE ALEATORIZADO, k TRATAMIENTOS EN b BLOQUES

$$CM = \frac{G^2}{n}$$

donde

$$G = \sum x_{ij} = \text{Total de todas las } n = bk \text{ observaciones}$$

$$TSS = \sum x_{ij}^2 - CM$$

$$= (\text{suma de todos los valores de } x) - CM$$

$$SST = \sum \frac{T_i^2}{b} - CM$$

$$SSB = \sum \frac{B_j^2}{k} - CM$$

$$SSE = TSS - SST - SSB$$

con

$$T_i = \text{Total de todas las observaciones que reciben tratamiento } i, i = 1, 2, \dots, k$$

$$B_j = \text{Total de todas las observaciones en bloque } j, j = 1, 2, \dots, b$$

Cada una de las tres **fuentes de variación**, cuando están divididas por los **grados de libertad** apropiados, da una estimación de la variación en el experimento. Como TSS contiene $n = bk$ observaciones al cuadrado, sus grados de libertad son $gl = (n - 1)$. Del mismo modo, SST contiene k totales al cuadrado y sus grados de libertad son $gl = (k - 1)$, en tanto que SSB contiene b totales al cuadrado y $(b - 1)$ grados de libertad. Finalmente, como los grados de libertad son aditivos, los restantes grados de libertad asociados con SSE se puede demostrar algebraicamente que son $gl = (b - 1)(k - 1)$.

Estas tres fuentes de variación y sus respectivos grados de libertad se combinan para formar las **medias cuadráticas** como $MS = SS/gl$, y la variación total en el experimento se exhibe entonces en una **tabla de análisis de varianza** (o ANOVA) como se muestra a continuación:

TIP MI CONSEJO

Los grados de libertad son aditivos

ANOVA PARA UN DISEÑO DE BLOQUE ALEATORIZADO, k TRATAMIENTOS Y b BLOQUES

Fuente	gl	SS	MS	F
Tratamientos	$k - 1$	SST	$MST = SST/(k - 1)$	MST/MSE
Bloques	$b - 1$	SSB	$MSB = SSB/(b - 1)$	MSB/MSE
Error	$(b - 1)(k - 1)$	SSE	$MSE = SSE/(b - 1)(k - 1)$	
Total	$n - 1 = bk - 1$			

EJEMPLO

7.8

La industria de la telefonía celular está inmersa en una feroz batalla por los clientes, con cada compañía ideando su propio y complejo plan de precios para atraer clientes. Como el costo de un minuto por teléfono celular varía en forma drástica dependiendo del número de minutos por mes usados por el cliente, un grupo de vigilancia integrado por consumidores decidió comparar el promedio de costos de cuatro compañías de telefonía celular usando tres diferentes niveles de uso como bloques. Los costos mensuales (en dólares), calculados por las compañías de telefonía celular para usuarios de tiempo pico en niveles de uso bajo (20 minutos por mes), medio (150 minutos por mes) y alto (1 000 minutos por mes), se dan en la tabla 7.3. Construya la tabla de análisis de varianza para este experimento.

TABLA 7.3

Costos mensuales de teléfono

Nivel de uso	Compañía				Totales
	A	B	C	D	
Bajo	27	24	31	23	$B_1 = 105$
Medio	68	76	65	67	$B_2 = 276$
Alto	308	326	312	300	$B_3 = 1246$
Totales	$T_1 = 403$	$T_2 = 426$	$T_3 = 408$	$T_4 = 390$	$G = 1627$

TIP MI CONSEJO

Los bloques contienen unidades experimentales que son relativamente las mismas

Solución El experimento está diseñado como *diseño de bloque aleatorizado* con $b = 3$ niveles de uso (bloques) y $k = 4$ compañías (tratamientos), de modo que hay $n = bk = 12$ observaciones y $G = 1\,627$. Entonces

$$CM = \frac{G^2}{n} = \frac{1627^2}{12} = 220\,594.0833$$

$$TSS = (27^2 + 24^2 + \cdots + 300^2) - CM = 189\,798.9167$$

$$SST = \frac{403^2 + \cdots + 390^2}{3} - CM = 222.25$$

$$SSB = \frac{105^2 + 276^2 + 1246^2}{4} - CM = 189\,335.1667$$

y por sustracción,

$$SSE = TSS - SST - SSB = 241.5$$

Estas cuatro fuentes de variación, sus grados de libertad, sumas de cuadrados y medias cuadráticas se muestran en el área sombreada de la tabla de análisis de varianza, generadas por MINITAB y MS Excel y se dan en las figuras 7.9a) y 7.9b). Encontrará instrucciones para generar esta salida impresa en la sección “Tecnología actual” al final de este capítulo.

FIGURA 7.9(a)

Salida de MINITAB para el ejemplo 7.8

ANOVA de dos vías: dólares contra uso, compañía

Source	DF	SS	MS	F	P
Usage	2	189335	94667.6	2351.99	0.000
Company	3	222	74.1	1.84	0.240
Error	6	242	40.3		
Total	11	189799			

S = 6.344 R-Sq = 99.87% R-Sq(adj) = 99.77%

FIGURA 7.9(b)

Salida de MS Excel para el ejemplo 7.8

ANOVA: dos factores sin réplica

RESUMEN	Conteo	Suma	Promedio	Varianza
Bajo	4	105	26.25	12.917
Medio	4	276	69	23.333
Alto	4	1246	311.5	118.333
A	3	403	134.333	23040.333
B	3	426	142	26068
C	3	408	136	23521
D	3	390	130	22159

ANOVA

Fuente de variación	SS	gl	MS	F	Valor P	F crít.
Uso	189335.167	2	94667.583	2351.990	0.000	5.143
Compañía	222.25	3	74.083	1.841	0.240	4.757
Error	241.5	6	40.25			
Total	189798.917	11				

Observe que ambas tablas ANOVA muestran dos diferentes estadísticos F y valores de p . No debe sorprender el saber que estos estadísticos se usan para probar hipótesis respecto a la igualdad de medias *de tratamiento* y *de bloque*.

Prueba de la igualdad de las medias de tratamiento y de bloque

Las *medias cuadráticas* de la tabla de análisis de varianza se pueden usar para probar las hipótesis nulas

H_0 : No hay diferencia entre las k medias de tratamiento

o bien

H_0 : No hay diferencia entre las b medias de bloque

contra la hipótesis alternativa

H_a : Al menos una de las medias es diferente de al menos otra

con el uso de un argumento teórico similar al empleado para el diseño completamente aleatorizado.

- Recuerde que σ^2 es la varianza común para las observaciones en todas las bk combinaciones de bloque-tratamiento. La cantidad

$$MSE = \frac{SSE}{(b-1)(k-1)}$$

es una estimación insesgada de σ^2 , sea H_0 verdadera o no.

- Las dos medias cuadráticas, MST y MSB, estiman σ^2 sólo si H_0 es verdadera y tienden a ser inusualmente *grandes* si H_0 es falsa y ya sea que las medias de tratamiento o de bloque sean diferentes.
- Los estadísticos de prueba

$$F = \frac{\text{MST}}{\text{MSE}} \quad \text{y} \quad F = \frac{\text{MSB}}{\text{MSE}}$$

se usan para probar la igualdad de medias de tratamiento y de bloque, respectivamente. Ambos estadísticos tienden a ser más grandes de lo normal si H_0 es falsa. En consecuencia, se puede rechazar H_0 para valores grandes de F , usando valores críticos *de cola derecha* de la distribución F con los grados de libertad apropiados (véase la tabla 6 del Anexo) o valores p generados por computadora para sacar conclusiones estadísticas acerca de la igualdad de las medias poblacionales.

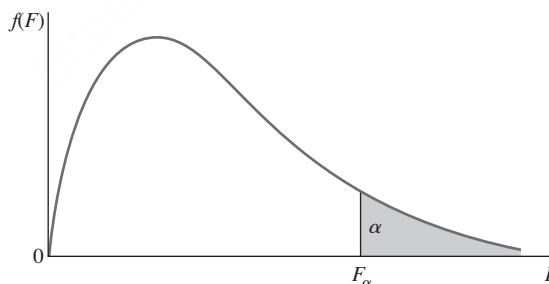
PRUEBAS PARA UN DISEÑO DE BLOQUE ALEATORIZADO

Para comparar medias de tratamiento:

1. Hipótesis nula: H_0 : las medias de tratamiento son iguales
2. Hipótesis alternativa: H_a : al menos dos de las medias de tratamiento difieren
3. Prueba estadística: $F = \text{MST}/\text{MSE}$, donde F está basada en $gl_1 = (k - 1)$ y $gl_2 = (b - 1)(k - 1)$
4. Región de rechazo: rechazar si $F > F_\alpha$, donde F_α se encuentra en la cola superior de la distribución F (véase la figura), o cuando el valor $p < \alpha$

Para comparar medias de bloque:

1. Hipótesis nula: H_0 : las medias de bloque son iguales
2. Hipótesis alternativa: H_a : al menos dos de las medias de bloque difieren
3. Prueba estadística: $F = \text{MSB}/\text{MSE}$, donde F está basada en $gl_1 = (b - 1)$ y $gl_2 = (b - 1)(k - 1)$
4. Región de rechazo: rechazar si $F > F_\alpha$, donde F_α se encuentra en la cola superior de la distribución F (véase la figura), o cuando el valor $p < \alpha$



EJEMPLO

7.9

¿Los datos del ejemplo 7.8 dan suficiente evidencia para indicar una diferencia en el promedio de costo mensual de teléfono celular, dependiendo de la compañía que use el cliente?

Solución Las compañías de telefonía celular representan los *tratamientos* en este diseño de bloque aleatorizado y las diferencias en el promedio de sus costos mensuales son de pri-

mordial importancia para el investigador. Para probar

H_0 : No hay diferencia en el promedio de costo entre compañías

contra la alternativa de que el promedio de costo es diferente para al menos una de las cuatro compañías, se usa el análisis de varianza del estadístico F , calculado como

$$F = \frac{MST}{MSE} = \frac{74.1}{40.3} = 1.84$$

y se muestra en la columna marcada “F” y el renglón marcado “Compañía” en las figuras 7.9a) y 7.9b). El valor p exacto para esta prueba estadística también se da en las figuras 7.9a) y 7.9b) como .240, que es demasiado grande para permitir el rechazo de H_0 . Los resultados no muestran una diferencia significativa en las medias de tratamiento. Esto es, hay insuficiente evidencia para indicar una diferencia en el promedio de costos mensuales para las cuatro compañías.

El investigador del ejemplo 7.9 estaba bastante seguro de que al usar un *diseño de bloque aleatorizado* habría una diferencia significativa en las medias de bloque, es decir, una diferencia significativa en el promedio de costos mensuales dependiendo del nivel de uso. Esta sospecha está justificada al ver la prueba de igualdad de medias de bloque. Observe que la prueba estadística observada es $F = 2351.99$ con valor $p = .000$, mostrando una diferencia altamente significativa, como se esperaba, en las medias de bloque.

Identificación de diferencias en las medias de tratamiento y de bloque

Una vez efectuada la prueba F general para igualdad de las medias de tratamiento o de bloque, ¿qué más se puede hacer para identificar la naturaleza de cualesquiera diferencias que se hayan encontrado? Al igual que en la sección 7.5, se puede usar el método de Tukey de comparaciones en pares para determinar qué pares de medias de tratamiento o de bloque son significativamente diferentes uno de otro. No obstante, si la prueba F no indica una diferencia significativa en las medias, no hay razón para usar el procedimiento de Tukey. Si el investigador tiene especial interés en un *par* particular de medias de tratamiento o de bloque, puede estimar la diferencia usando un intervalo de confianza $(1 - \alpha)100\%$.[†] Las fórmulas para estos procedimientos, que se muestran a continuación, siguen un patrón similar a las fórmulas para el diseño completamente aleatorizado. Recuerde que MSE siempre da un estimador insesgado de σ^2 y utiliza información proveniente de todo el conjunto de mediciones. En consecuencia, es el mejor estimador existente de σ^2 , cualquiera que sea el procedimiento de prueba o estimación que se use. Utilice de nuevo

$$s^2 = \text{MSE} \quad \text{con } gl = (b - 1)(k - 1)$$

para estimar σ^2 al comparar las medias de tratamiento y de bloque.

TIP MI CONSEJO

Los grados de libertad para la prueba de Tukey y para intervalos de confianza son los *gl de error*

COMPARACIÓN DE MEDIAS DE TRATAMIENTO Y DE BLOQUE

La prueba de Tukey para comparar medias de bloque:

$$\omega = q_\alpha(b, gl) \left(\frac{s}{\sqrt{k}} \right)$$

Medidor de Tukey para comparar medias de tratamiento:

$$\omega = q_\alpha(k, gl) \left(\frac{s}{\sqrt{b}} \right)$$

[†] No se puede construir un intervalo de confianza para una media individual, a menos que los bloques hayan sido seleccionados al azar de entre la población de todos los bloques. El procedimiento para construir intervalos para medias individuales está fuera del propósito de este libro.

Intervalo de confianza de $(1 - \alpha)100\%$ para la diferencia en dos medias de bloque:

$$(\bar{B}_i - \bar{B}_j) \pm t_{\alpha/2} \sqrt{s^2 \left(\frac{1}{k} + \frac{1}{k} \right)}$$

donde $\bar{B}_i$ es el promedio de todas las observaciones en el bloque i

Intervalo de confianza de $(1 - \alpha)100\%$ para la diferencia en dos medias de tratamiento:

$$(\bar{T}_i - \bar{T}_j) \pm t_{\alpha/2} \sqrt{s^2 \left(\frac{1}{b} + \frac{1}{b} \right)}$$

donde $\bar{T}_i$ es el promedio de todas las observaciones en el tratamiento i .

Nota: Los valores $q_{\alpha}(*, gl)$ de la tabla 11 en el Anexo, $t_{\alpha/2}$ de la tabla 4 en el Anexo y $s^2 = \text{MSE}$ dependen todos ellos de $gl = (b - 1)(k - 1)$ grados de libertad.

EJEMPLO

7.10

Identifique la naturaleza de cualesquiera diferencias encontradas en el promedio de costos mensuales de teléfono celular del ejemplo 7.8.

Solución Como la prueba F no presentó diferencias significativas en los costos promedio para las cuatro compañías, no hay razón para usar el método de Tukey de comparaciones en pares. Supongamos, sin embargo, que usted es un ejecutivo de la compañía B y su principal competidor es la compañía C. ¿Puede decir que hay una diferencia significativa en los dos costos promedio? Usando un intervalo de confianza de 95%, puede calcular

$$(\bar{T}_2 - \bar{T}_3) \pm t_{0.025} \sqrt{\text{MSE} \left(\frac{2}{b} \right)}$$

$$\left(\frac{426}{3} - \frac{408}{3} \right) \pm 2.447 \sqrt{40.3 \left(\frac{2}{3} \right)}$$

$$6 \pm 12.68$$

de modo que la diferencia entre los dos costos promedio se estima entre $-\$6.68$ y $\$18.68$. Como 0 está contenido en el intervalo, no tiene evidencia para indicar una diferencia significativa en sus costos promedio. ¡Qué pena!

TIP MI CONSEJO

No se puede formar un intervalo de confianza o probar una hipótesis acerca de una media de tratamiento individual en un diseño de bloque aleatorizado

Algunos comentarios de precaución en bloqueo

A continuación veamos algunos puntos importantes que recordar:

- Un diseño de bloque aleatorizado no debe usarse cuando tanto tratamientos como bloques corresponden a factores **experimentales** de interés para el investigador. Al designar un factor como *bloque*, puede suponer que el efecto del tratamiento será el mismo, cualquiera que sea el bloque que utilice. Si éste *no es* el caso, los dos factores, bloques y tratamientos, se dice que **interactúan** y el análisis podría llevar a conclusiones incorrectas respecto a la relación entre los tratamientos y la respuesta. Cuando se sospeche que hay una *interacción* entre dos factores, deben analizarse los datos como **experimento factorial**, que se introduce en la siguiente sección.
- Recuerde que el bloqueo puede no ser siempre benéfico. Cuando el SSB se elimine del SSE, el número de grados de libertad asociado con el SSE se reduce. Para que el bloqueo sea benéfico, la información ganada al aislar la variación de bloque debe

importar más que la pérdida de grados de libertad por error, pero, por lo general, si se sospecha que las unidades experimentales no son homogéneas y se pueden agrupar las unidades en bloques, es bueno usar el *diseño de bloque aleatorizado*.

- Por último, recuerde que no se pueden construir intervalos de confianza para medias de tratamiento individuales a menos que sea razonable suponer que los b bloques se han seleccionado al azar de entre una población de bloques. Si el experimentador construye ese intervalo, la media de tratamiento muestral estará sesgada por los efectos positivos y negativos que los bloques tienen en la respuesta.

7.8

EJERCICIOS

TÉCNICAS BÁSICAS

7.28 Se utilizó un diseño de bloque aleatorizado para comparar las medias de tres tratamientos dentro de seis bloques. Construya una tabla ANOVA que muestre las fuentes de variación y sus respectivos grados de libertad.

7.29 Suponga que los cálculos del análisis de varianza para el ejercicio 7.28 son $SST = 11.4$, $SSB = 17.1$ y $TSS = 42.7$. Complete la tabla ANOVA, mostrando todas las sumas de cuadrados, medias cuadráticas y valores F pertinentes.

7.30 ¿Los datos del ejercicio 7.28 dan suficiente evidencia para indicar diferencias entre las medias de tratamiento? Pruebe usando $\alpha = .05$.

7.31 Consulte el ejercicio 7.28. Encuentre un intervalo de confianza de 95% para la diferencia entre un par de medias de tratamiento A y B si $\bar{x}_A = 21.9$ y $\bar{x}_B = 24.2$.

7.32 ¿Los datos del ejercicio 7.28 dan suficiente evidencia para indicar que el bloque aumentó la cantidad de información en el experimento acerca de las medias de tratamiento? Justifique su respuesta.

DATOS EX0733 **7.33** Los datos que siguen son observaciones recolectadas de un experimento que comparó cuatro tratamientos, A, B, C y D, dentro de cada uno de tres bloques, usando un diseño de bloque aleatorizado.

Bloque	Tratamiento				Total
	A	B	C	D	
1	6	10	8	9	33
2	4	9	5	7	25
3	12	15	14	14	55
Total	22	34	27	30	113

- a. ¿Los datos presentan suficiente evidencia para indicar diferencias entre las medias de tratamiento? Pruebe usando $\alpha = .05$.

- b. ¿Los datos presentan suficiente evidencia para indicar diferencias entre las medias de bloque? Pruebe usando $\alpha = .05$.
- c. Clasifique las cuatro medias de tratamiento usando el método de Tukey de comparaciones en pares con $\alpha = .01$.
- d. Encuentre un intervalo de confianza de 95% para la diferencia en medias para los tratamientos A y B.
- e. ¿Le parece que el uso de un diseño de bloque aleatorizado para este experimento estaba justificado? Explique.

DATOS EX0734

7.34 Los datos que se muestran a continuación son observaciones recolectadas de un experimento que comparó tres tratamientos, A, B y C, dentro de cada uno de cinco bloques, usando un diseño de bloque aleatorizado:

Tratamiento	Bloque					Total
	1	2	3	4	5	
A	2.1	2.6	1.9	3.2	2.7	12.5
B	3.4	3.8	3.6	4.1	3.9	18.8
C	3.0	3.6	3.2	3.9	3.9	17.6
Total	8.5	10.0	8.7	11.2	10.5	48.9

Salida de *MS Excel* para el ejercicio 7.34

ANOVA: dos factores sin réplica

RESUMEN	Conteo	Suma	Promedio	Varianza
A	5	12.5	2.5	0.265
B	5	18.8	3.76	0.073
C	5	17.6	3.52	0.167
1	3	8.5	2.833	0.443
2	3	10	3.333	0.413
3	3	8.7	2.9	0.79
4	3	11.2	3.733	0.223
5	3	10.5	3.5	0.48

ANOVA

Fuente de variación	SS	gl	MS	F	Valor P	F crít.
Filas	4.476	2	2.238	79.929	0.000	4.459
Columnas	1.796	4	0.449	16.036	0.001	3.838
Error	0.224	8	0.028			
Total	6.496	14				

Use la salida de *Excel* para analizar el experimento. Investigue posibles diferencias en las medias de bloque y tratamiento y, si existen algunas diferencias, use un método apropiado para identificar específicamente dónde están las diferencias. ¿El bloqueo ha sido eficaz en este experimento? Presente sus resultados en un informe.

7.35 La tabla ANOVA parcialmente completada para un diseño de bloque aleatorizado se presenta a continuación:

Fuente	gl	SS	MS	F
Tratamientos	4	14.2		
Bloques		18.9		
Error	24			
Total	34	41.9		

- ¿Cuántos bloques intervienen en el diseño?
- ¿Cuántas observaciones hay en cada total de tratamiento?
- ¿Cuántas observaciones hay en cada total de bloque?
- Llene los espacios en blanco de la tabla ANOVA.
- ¿Los datos presentan suficiente evidencia para indicar diferencias entre las medias de tratamiento? Pruebe usando $\alpha = .05$.
- ¿Los datos presentan suficiente evidencia para indicar diferencias entre las medias de bloque? Pruebe usando $\alpha = .05$.

APLICACIONES

DATOS
EX0736

7.36 Rendimiento de gasolina Se realizó un estudio para comparar el rendimiento de gasolina en automóviles para tres fórmulas de gasolina. Se utilizaron cuatro automóviles, todos ellos de la misma marca y modelo, y cada fórmula se probó en cada uno de los autos. El uso de cada fórmula en el mismo auto tiene el efecto de eliminar (bloquear) variabilidad de un auto a otro. Los datos (en millas por galón) se dan a continuación.

	Automóvil			
Fórmula	1	2	3	4
A	25.7	27.0	27.3	26.1
B	27.2	28.1	27.9	27.7
C	26.1	27.5	26.8	27.8

- ¿Los datos presentan suficiente evidencia para indicar diferencias en distancia media en millas por galón para las tres fórmulas de gasolina?
- ¿Hay suficiente evidencia de una diferencia en distancia media en millas para los cuatro automóviles?

- Suponga que *antes de ver los datos*, se ha decidido comparar la distancia media en millas por galón para las fórmulas A y B. Encuentre un intervalo de confianza de 90% para esta diferencia.
- Use un método apropiado para identificar las diferencias por pares, si las hay, en el promedio de las distancias medias para las tres fórmulas.

DATOS
EX0737

7.37 Resistencia al agua en textiles Se realizó un experimento para comparar los efectos de cuatro productos químicos diferentes, A, B, C y D, para producir resistencia al agua en textiles. Una tira de material, seleccionada de un rollo de tela, se cortó en cuatro partes y todas éstas se asignaron al azar para recibir uno de los cuatro productos químicos, A, B, C o D. Este proceso se repitió tres veces, produciendo así un diseño de bloque aleatorizado. Este diseño, con mediciones de resistencia a la humedad, se muestra en la figura siguiente (bajas lecturas indican baja penetración de humedad). Analice el experimento usando un método apropiado para este diseño de bloque aleatorizado. Identifique los bloques y tratamientos e investigue cualesquiera diferencias posibles en las medias de tratamiento. Si existen diferencias, use un método apropiado para identificar específicamente dónde se encuentran las diferencias. ¿Cuáles son las implicaciones prácticas para los productores de las sustancias químicas? ¿El bloqueo ha sido eficaz en este experimento? Presente sus resultados en un informe.

Ilustración para el ejercicio 7.37

Bloques (muestras de rollo)

1	2	3
C 9.9	D 13.4	B 12.7
A 10.1	B 12.9	D 12.9
B 11.4	A 12.2	C 11.4
D 12.1	C 12.3	A 11.9

7.38 Deslumbramiento en espejos retrovisores Se realizó un experimento para comparar las características de deslumbramiento de cuatro tipos de espejos retrovisores en automóviles. Cuarenta conductores se seleccionaron al azar para participar en el experimento; cada uno fue expuesto a deslumbramiento producido por un faro situado a 30 pies detrás de la ventanilla trasera del auto experimental. El conductor clasificó entonces el deslumbramiento producido por el espejo retrovisor en una escala de 1 (bajo) a 10 (alto). Cada uno de los cuatro espejos fue probado por cada conductor; los espejos se asignaron a un conductor en orden aleatorio.

Un análisis de varianza de los datos produjo esta tabla ANOVA:

Fuente	gl	SS	MS	F
Espejos		46.98		
Conductores			8.42	
Error				
Total		638.61		

- Llene los espacios en blanco de la tabla ANOVA.
- ¿Los datos presentan suficiente evidencia para indicar diferencias en las clasificaciones medias de deslumbramiento de los cuatro espejos retrovisores? Calcule el valor p aproximado y úselo para tomar su decisión.
- ¿Los datos presentan suficiente evidencia para indicar que el nivel de deslumbramiento percibido por los conductores varió de un conductor a otro? Use el método del valor p .
- Con base en los resultados de la parte b, ¿cuáles son las implicaciones prácticas de este experimento para los fabricantes de los espejos retrovisores?

DATOS
EX0739

7.39 Plantar pinos ayacahuites Se realizó un experimento para determinar los efectos de tres métodos de preparación del terreno en el primer año de crecimiento de semillas de ayacahuite. Se seleccionaron cuatro lugares (terrenos de bosques estatales) y cada lugar se dividió en tres lotes. Como se tenía la impresión de que la fertilidad del suelo dentro de un lugar era más homogénea que entre éstos, se utilizó un diseño de bloques aleatorizado usando lugares como bloques. Los métodos de preparación del terreno fueron A (sin preparación), B (ligera fertilización) y C (quema). Cada preparación del terreno se aplicó al azar a un lote dentro de cada lugar. En cada lote, el mismo número de semillas se plantó y el promedio de crecimiento del primer año de las semillas se registró en cada lote. Use la salida impresa del MINITAB para contestar las preguntas.

Terreno Preparación	Lugar			
	1	2	3	4
A	11	13	16	10
B	15	17	20	12
C	10	15	13	10

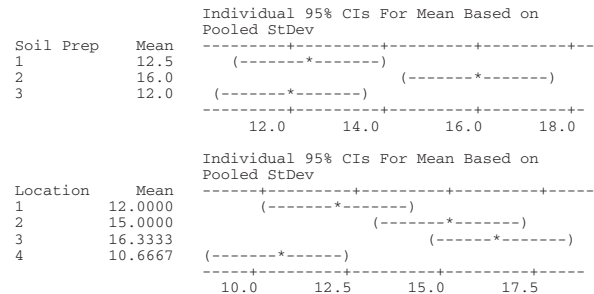
- Efectúe un análisis de varianza. ¿Los datos dan suficiente evidencia para indicar una diferencia en los crecimientos medios para las tres preparaciones del terreno?
- ¿Hay evidencia para indicar una diferencia en porcentajes medios de crecimiento para los cuatro lugares?
- Use el método de Tukey de comparaciones en pares para clasificar los crecimientos medios para las tres preparaciones del terreno. Use $\alpha = .05$.
- Use un intervalo de confianza de 95% para estimar la diferencia en crecimientos medios para los métodos A y B.

Salida MINITAB para el ejercicio 7.39

ANOVA en dos vías: crecimiento contra prep. de terreno, lugar

Source	DF	SS	MS	F	P
Soil Prep	2	38.000	19.0000	10.06	0.012
Location	3	61.667	20.5556	10.88	0.008
Error	6	11.333	1.8889		
Total	11	111.000			

S = 1.374 R-Sq = 89.79% R-Sq(adj) = 81.28%



DATOS
EX0740

7.40 Dedalera e ingesta de calcio Se realizó un estudio para comparar los efectos de tres niveles de dedalera en los niveles de calcio en los músculos cardiacos de perros. Debido a que el nivel general de ingesta de calcio varía de un animal a otro, el tejido de un músculo cardiaco se consideró como un bloque y se hicieron comparaciones de los tres niveles de dedalera (tratamientos) dentro de un animal determinado. Las ingestas de calcio para los tres niveles de dedalera, A, B y C, fueron comparados con base en los músculos cardiacos de cuatro perros y los resultados se dan en la tabla siguiente. Use la salida impresa de Excel para responder las preguntas.

Perros			
1	2	3	4
A	C	B	A
1342	1698	1296	1150
B	B	A	C
1608	1387	1029	1579
C	A	C	B
1881	1140	1549	1319

- ¿Cuántos grados de libertad están asociados con el SSE?
- ¿Los datos presentan suficiente evidencia para indicar una diferencia en las ingestas medias de calcio para los tres niveles de dedalera?
- Use el método de Tukey de comparaciones en pares con $\alpha = .01$ para clasificar las ingestas medias de calcio para los tres niveles de dedalera.
- ¿Los datos indican una diferencia en las ingestas medias de calcio para los cuatro músculos cardiacos?
- Use el método de Tukey de comparaciones en pares con $\alpha = .01$ para clasificar las ingestas medias de calcio para los músculos cardiacos de los cuatro perros empleados en el experimento. ¿Estos resultados son de algún valor práctico para el investigador?

- f. Dé el error estándar de la diferencia entre las ingestas medias de calcio para dos niveles de dedalera.
- g. Encuentre un intervalo de confianza de 95% para la diferencia en respuestas medias entre los tratamientos A y B.

Salida de *MS Excel* para el ejercicio 7.40

ANOVA: dos factores sin réplica

RESUMEN	Conteo	Suma	Promedio	Varianza
A	4	4661	1165.25	16891.583
B	4	5610	1402.5	20261.667
C	4	6707	1676.75	72681.583
1	3	4831	1610.333	72634.333
2	3	4225	1408.333	78182.333
3	3	3874	1291.333	67616.333
4	3	4048	1349.333	46700.333

Fuente de variación	SS	gl	MS	F	Valor P	F. crít.
Dedalera	524177.167	2	262088.583	258.237	0.000	5.143
Perros	173415	3	57805	56.955	0.000	4.757
Error	6089.5	6	1014.917			
Total	703681.667	11				

DATOS

EX0741

7.41 Cotizaciones en trabajos de

construcción Un contratista constructor emplea tres ingenieros de construcción, A, B y C,

para estimar y cotizar trabajos. Para determinar si uno tiende a ser estimador más conservador (o liberal) que los otros, el contratista selecciona cuatro trabajos de construcción proyectados y hace que cada estimador, en forma independiente, estime el costo (en dólares por pie cuadrado) de cada trabajo. Los datos se muestran en la tabla siguiente:

Estimador	Trabajo de construcción				Total
	1	2	3	4	
A	35.10	34.50	29.25	31.60	130.45
B	37.45	34.60	33.10	34.40	139.55
C	36.30	35.10	32.45	32.90	136.75
Total	108.85	104.20	94.80	98.90	406.75

Analice el experimento usando los métodos apropiados. Identifique los bloques y tratamientos e investigue cualesquiera diferencias posibles en medias de tratamiento. Si existe alguna, use un método apropiado para identificar de manera específica en dónde se encuentran las diferencias. ¿El bloqueo ha sido efectivo en este experimento? ¿Cuáles son las implicaciones prácticas del experimento? Presente sus resultados en un informe.

DATOS

EX0742

7.42 ¿Equidad en las primas? El costo del seguro para auto varía por cobertura, ubicación y el historial de manejo del conductor. Las siguientes son estimaciones del costo anual para cobertura estándar al 1 de enero de 2011 para un conductor varón con 6–8 años de experiencia, que conduce un Honda Accord sin accidentes ni infracciones.³ (Éstas son cotizaciones y no primas.)

Ubicación	21st Century	Geico	AAA	Fireman's Fund	State Farm
West Hollywood	3922	4073	3663	4075	3876
Laguna Beach	2378	2512	2478	3056	2508
Redlands	2560	2476	2549	2756	2614
Riverside	2584	2759	2494	2940	2714

Fuente: www.insurance.ca.gov

- ¿Qué tipo de diseño se utilizó para recolectar estos datos?
- ¿Hay suficiente evidencia para indicar que las primas de seguro para el mismo tipo de cobertura difieren de una compañía a otra?
- ¿Hay suficiente evidencia para indicar que las primas de seguro varían de un lugar a otro?
- Use el procedimiento de Tukey para determinar cuáles compañías de seguros citadas aquí difieren de otras en las primas que cobran para este cliente típico. Use $\alpha = .05$.
- Haga un resumen de lo que encuentre.

DATOS

EX0743

7.43 ¿Dónde comprar? ¿Usted compra en la tienda de abarrotes más cercana a su casa o busca la tienda que tenga los mejores precios?

Se compararon los precios normales en cuatro tiendas de abarrotes diferentes para ocho artículos comprados el mismo día.

Artículos	Tiendas			
	Vons	Ralphs	Stater Bros	WinCo
Ensalada mixta, bolsa de 12 oz.	3.99	2.79	1.99	1.78
Salchichas de res ahumadas Hillshire Farm®, 14 oz.	4.29	4.29	3.99	2.50
Raisin Bran® de Kellogg's, 25.5 oz.	4.49	5.49	4.49	3.15
Queso crema Philadelphia® de Kraft®, 8 oz.	2.99	3.19	2.79	1.48
Aderezo Ranch de Kraft®, 16 oz.	3.19	3.49	3.49	1.48
Jugo de manzana Treetop®, 64 oz.	2.99	3.49	3.49	1.58
Jabón Dial®, Gold, 8–4 oz.	5.99	6.49	5.79	5.14
Crema de cacahuete Jif®, cremosa, 28 oz.	5.15	5.49	4.79	4.34

- ¿Cuáles son los bloques y tratamientos en este experimento?
- ¿Los datos dan evidencia para indicar que hay diferencias significativas en precios de una tienda a otra? Apoye su respuesta estadísticamente usando la salida impresa ANOVA que sigue.
- ¿Hay diferencias significativas de un bloque a otro? ¿Fue eficaz el bloqueo?

ANOVA de dos vías: precio contra artículo, tienda

Fuente	GL	SS	MS	F	P
Artículo	7	40.2184	5.74548	29.99	0.000
Tienda	3	14.6695	4.88982	25.53	0.000
Error	21	4.0230	0.19157		
Total	31	58.9108			

$S = 0.4377$ $R-Sq = 93.17\%$ $R-Sq(adj) = 89.92\%$

7.44 ¿Dónde comprar?, continúa Consulte el ejercicio 7.43. La salida impresa que sigue da el promedio de costos de los artículos seleccionados para las $k = 4$ tiendas.

Tienda	Media
Vons	4.1350
Ralphs	4.3400
Stater	3.8525
WinCo	2.5975

- ¿Cuál es el valor apropiado de $q_{.05}(k, gl)$ para probar diferencias entre tiendas?
- ¿Cuál es el valor de $\omega = q_{.05}(k, gl) \sqrt{\frac{MSE}{b}}$?
- Use la prueba de comparación por pares de Tukey entre tiendas usadas para determinar cuáles tiendas difieren significativamente en promedio de precios de los artículos seleccionados.

EL EXPERIMENTO FACTORIAL $a \times b$: UNA CLASIFICACIÓN EN DOS VÍAS

7.9

Suponga que el gerente de una planta manufacturera sospecha que la producción (en número de unidades producidas por turno) de una línea de producción depende de dos factores:

- Cuál de dos supervisores está a cargo de la línea
- Cuál de tres turnos, diurno, vespertino o nocturno, se está midiendo

Esto es, el gerente está interesado en dos *factores*: “supervisor” a dos niveles y “turno” a tres niveles. ¿Puede usted usar un diseño de bloque aleatorizado, diseñando uno de los dos factores como factor de bloque? Para hacer esto necesitaría suponer que el efecto de los dos supervisores es el mismo, cualquiera que sea el turno que considere. Éste puede no ser el caso; quizás el primer supervisor es más eficiente en la mañana y el segundo lo es en el turno de la noche. No se puede generalizar y decir que un supervisor es mejor que el otro o que la producción de un turno particular es mejor. Es necesario investigar no sólo el promedio de producción para los dos supervisores y el promedio de producción para los tres turnos, sino también la **interacción** o relación entre los dos factores. Considere dos ejemplos diferentes que muestran el efecto de *interacción* en las respuestas en esta situación.

EJEMPLO

7.11

Suponga que los dos supervisores son observados en tres días seleccionados al azar para cada uno de los tres turnos diferentes. El promedio de producciones para los tres turnos se muestra en la tabla 7.4 para cada uno de los supervisores. Vea la relación entre los dos factores en la gráfica de líneas para estas medias, mostrada en la figura 7.10. Observe que el supervisor 2 siempre produce más, cualquiera que sea el turno. Los dos factores se comportan *independientemente*; esto es, la producción es siempre de unas 100 piezas más para el supervisor 2, no importa cuál turno se vea.

Ahora considere otro conjunto de datos para la misma situación, mostrado en la tabla 7.5. Hay una diferencia definida en los resultados, dependiendo de cuál turno se vea, y la *interacción* se puede ver en las líneas cruzadas de la gráfica de la figura 7.11.

TABLA 7.4

Promedio de producción para dos supervisores en tres turnos

Supervisor	Turno		
	Diurno	Vespertino	Nocturno
1	487	498	550
2	602	602	637

FIGURA 7.10
Gráfica de interacción para medias en la tabla 7.4

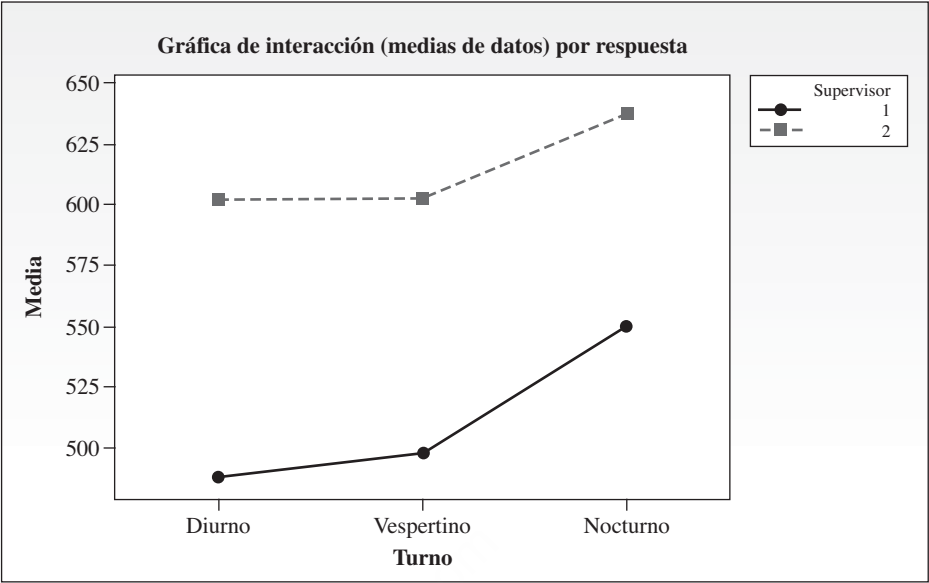
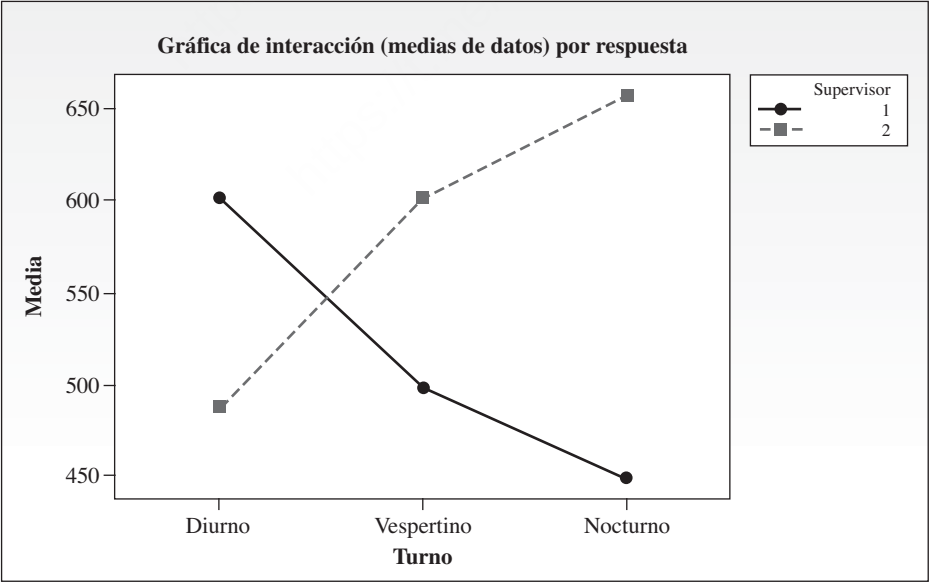


TABLA 7.5 Promedio de producción para dos supervisores en tres turnos

Supervisor	Turno		
	Diurno	Vespertino	Nocturno
1	602	498	450
2	487	602	657

FIGURA 7.11
Gráfica de interacción para medias en la tabla 7.5



TIP MI CONSEJO

Cuando cambia el efecto de un factor sobre la respuesta, dependiendo del nivel al cual se mide el otro factor, se dice que los dos factores **interactúan**.

Esta situación es ejemplo de un **experimento factorial** en el que hay un total de 2×3 posibles combinaciones de los niveles para los dos factores. Estas $2 \times 3 = 6$ combinaciones forman los *tratamientos* y el experimento se denomina **experimento factorial de 2×3** . Este tipo de experimento en realidad se puede usar para investigar los efectos de tres o más factores en una respuesta y explorar las interacciones entre los factores. No obstante, confinamos nuestra discusión a dos factores y su interacción.

Cuando se comparen medias de tratamiento para un experimento factorial (o para cualquier otro experimento), será necesaria más de una observación por tratamiento. Por ejemplo, si el experimentador obtiene dos observaciones para cada una de las combinaciones de factor de un experimento factorial completo, tiene dos **réplicas** del experimento. En la siguiente sección sobre el análisis de varianza para un experimento factorial, se puede suponer que cada tratamiento o combinación de niveles de factor se replica el mismo número de veces r .

EL ANÁLISIS DE VARIANZA PARA UN EXPERIMENTO FACTORIAL $a \times b$

7.10

Un análisis de varianza para un experimento factorial de dos factores replicado r veces sigue el mismo patrón que los diseños previos. Si las letras A y B se usan para identificar los dos factores, la variación total en el experimento

$$\text{TSS} = \sum (x - \bar{x})^2 = \sum x^2 - \text{CM}$$

se divide en *cuatro* partes de modo que

$$\text{TSS} = \text{SSA} + \text{SSB} + \text{SS(AB)} + \text{SSE}$$

donde

- SSA (suma de cuadrados para el factor A) mide la variación entre medias del factor A.
- SSB (suma de cuadrados para el factor B) mide la variación entre medias del factor B.
- SS(AB) (suma de cuadrados para interacción) mide la variación *entre* las diferentes combinaciones de niveles de factor.
- SSE (suma de cuadrados de error) mide la variación de las diferencias entre las observaciones *dentro* de cada combinación de niveles de factor, es decir, el error experimental.

Es frecuente que las sumas de cuadrados SSA y SSB reciban el nombre de sumas de cuadrados de **efecto principal**, para distinguirlas de la suma de cuadrados de **interacción**. Aunque se puede simplificar el trabajo si se usa un programa de computadora para calcular estas sumas de cuadrados, las fórmulas de cálculo se dan a continuación. Se puede suponer que son:

- a niveles del factor A
- b niveles del factor B
- r réplicas de cada una de las ab combinaciones de factor
- Un total de $n = abr$ observaciones

CÁLCULO DE LAS SUMAS DE CUADRADOS PARA UN EXPERIMENTO FACTORIAL DE DOS FACTORES

$$CM = \frac{G^2}{n} \quad TSS = \sum x^2 - CM$$

$$SSA = \sum \frac{A_i^2}{br} - CM \quad SSB = \sum \frac{B_j^2}{ar} - CM$$

$$SS(AB) = \sum \frac{(AB)_{ij}^2}{r} - CM - SSA - SSB$$

donde

G = Suma de todas las $n = abr$ observaciones

A_i = Total de todas las observaciones al i -ésimo nivel del factor A, $i = 1, 2, \dots, a$

B_j = Total de todas las observaciones al j -ésimo nivel del factor B, $j = 1, 2, \dots, b$

$(AB)_{ij}$ = Total de las r observaciones al i -ésimo nivel del factor A y el j -ésimo nivel del factor B

Cada una de las cinco **fuentes de variación**, cuando se dividen entre los **grados de libertad** apropiados, da una estimación de la variación en el experimento. Estas estimaciones se denominan **medias cuadráticas**, $MS = SS/gl$, y se muestran junto con sus respectivas sumas de cuadrados y gl en la **tabla de análisis de varianza** (o ANOVA).

TABLA ANOVA PARA r RÉPLICAS DE UN EXPERIMENTO FACTORIAL DE DOS FACTORES: FACTOR A A NIVELES a Y FACTOR B A NIVELES b

Fuente	gl	SS	MS	F
A	$a - 1$	SSA	$MSA = \frac{SSA}{a - 1}$	$\frac{MSA}{MSE}$
B	$b - 1$	SSB	$MSB = \frac{SSB}{b - 1}$	$\frac{MSB}{MSE}$
AB	$(a - 1)(b - 1)$	SS(AB)	$MS(AB) = \frac{SS(AB)}{(a - 1)(b - 1)}$	$\frac{MS(AB)}{MSE}$
Error	$ab(r - 1)$	SSE	$MSE = \frac{SSE}{ab(r - 1)}$	
Total	$abr - 1$	SS Total		

Finalmente, la igualdad de medias para varios niveles de las combinaciones de factor (el efecto de interacción) y para los niveles de ambos efectos principales, A y B, se puede probar usando las pruebas F de ANOVA, como se ve a continuación.

PRUEBAS PARA UN EXPERIMENTO FACTORIAL

- **Para interacción:**

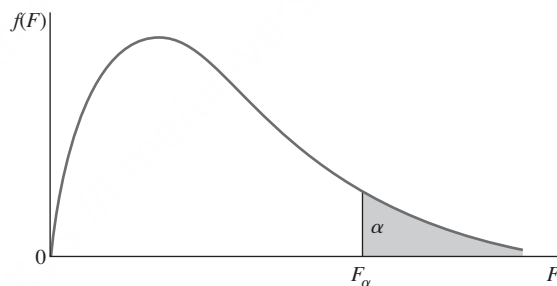
1. Hipótesis nula: H_0 : Los factores A y B no interactúan
2. Hipótesis alternativa: H_a : Los factores A y B interactúan
3. Prueba estadística: $F = MS(AB)/MSE$, donde F está basada en $gl_1 = (a - 1)(b - 1)$ y $gl_2 = ab(r - 1)$
4. Región de rechazo: rechazar H_0 cuando $F > F_\alpha$, donde F_α está en la cola superior de la distribución F (véase la figura), o cuando el valor $p < \alpha$

- **Para efectos principales, factor A:**

1. Hipótesis nula: H_0 : No hay diferencias entre las medias del factor A
2. Hipótesis alternativa: H_a : Al menos dos de las medias del factor A difieren
3. Prueba estadística: $F = MSA/MSE$, donde F está basada en $gl_1 = (a - 1)$ y $gl_2 = ab(r - 1)$
4. Región de rechazo: rechazar H_0 cuando $F > F_\alpha$ (véase la figura), o cuando el valor $p < \alpha$

- **Para efectos principales, factor B:**

1. Hipótesis nula: H_0 : No hay diferencias entre las medias del factor B
2. Hipótesis alternativa: H_a : Al menos dos de las medias del factor B difieren
3. Prueba estadística: $F = MSB/MSE$, donde F está basada en $gl_1 = (b - 1)$ y $gl_2 = ab(r - 1)$
4. Región de rechazo: rechazar H_0 cuando $F > F_\alpha$ (véase la figura), o cuando el valor $p < \alpha$



EJEMPLO

7.12

La tabla 7.6 muestra los datos originales empleados para generar la tabla 7.5 del ejemplo 7.11. Esto es, los dos supervisores fueron observados en tres días seleccionados al azar para cada uno de los tres turnos diferentes, registrándose las salidas de producción. Analice estos datos usando el procedimiento apropiado de análisis de varianza.

TABLA 7.6

Producciones para dos supervisores en tres turnos

Supervisor	Turno		
	Diurno	Vespertino	Nocturno
1	571	480	470
	610	474	430
	625	540	450
2	480	625	630
	516	600	680
	465	581	661

Solución La salida de computadora de la figura 7.12a) fue generada usando el procedimiento de dos vías de análisis de varianza del paquete de software *MINITAB*. Una salida impresa similar generada por *Excel* se muestra en la figura 7.12b). Se pueden verificar las cantidades de la tabla ANOVA usando las fórmulas de cálculo presentadas antes o se puede elegir sólo usar los resultados e interpretar su significado.

FIGURA 7.12(a)

Salida *MINITAB* para el ejemplo 7.12

ANOVA de dos vías: producción contra supervisor, turno

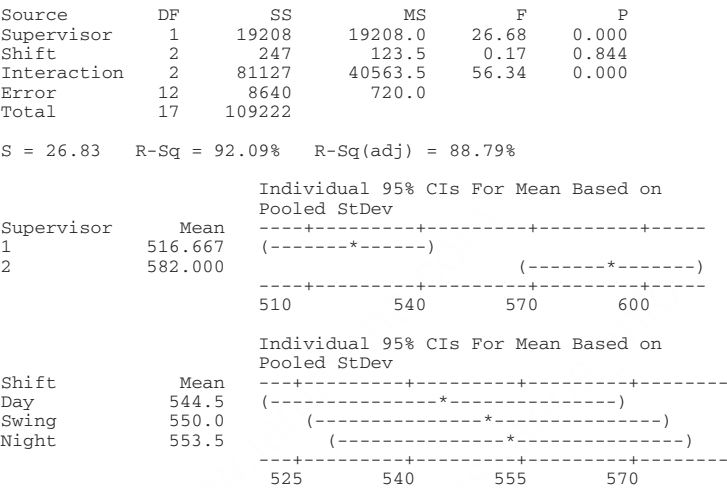


FIGURA 7.12(b)

Salida de *MS Excel* para el ejemplo 7.12

ANOVA: dos factores con réplica

RESUMEN	Diurno	Vespertino	Nocturno	Total
1				
Conteo	3	3	3	9
Suma	1806	1494	1350	4650
Promedio	602	498	450	516.667
Varianza	777	1332	400	5155.25
2				
Conteo	3	3	3	9
Suma	1461	1806	1971	5238
Promedio	487	602	657	582
Varianza	687	487	637	6096.5
Total				
Conteo	6	6	6	
Suma	3267	3300	3321	
Promedio	544.5	550	553.5	
Varianza	4553.1	3972.4	13269.5	

ANOVA	SS	gl	MS	F	Valor P	F crít.
Fuente de variación						
Supervisor	19208	1	19208	26.678	0.000	4.747
Turno	247	2	123.5	0.172	0.844	3.885
Interacción	81127	2	40563.5	56.338	0.000	3.885
Dentro	8640	12	720			
Total	109222	17				

TIP MI CONSEJO

Si la interacción **no es significativa**, pruebe cada uno de los factores individualmente

En este punto es indudable que usted haya descubierto el patrón conocido para probar la significancia de los diversos factores experimentales con el estadístico F y su valor p . El valor p pequeño (.000) en la fila marcada “Supervisor” significa que hay suficiente evidencia para declarar una diferencia en los niveles medios para el factor A, es decir, una diferencia en producciones medias por supervisor. Este hecho es visualmente aparente en los intervalos de confianza que no se traslapan para las medias del supervisor que se ven en la salida impresa. Pero esto es opacado por el hecho de que hay una fuerte evidencia ($P = .000$) de una *interacción* entre los factores A y B. Esto significa que el promedio de producción para un turno determinado depende del supervisor en servicio. Se vio claramente este efecto en la figura 7.10. Las tres producciones medias más grandes ocurren cuando el supervisor 1 está en el turno diurno y cuando el supervisor 2 está ya sea en el turno vespertino o nocturno. Como resultado práctico, el gerente debe programar al supervisor 1 para el turno diurno y al supervisor 2 para el turno nocturno.

Si el efecto de interacción *es* significativo, las diferencias en las medias de tratamiento se pueden estudiar más, *no* comparando las medias para los factores A o B individualmente sino más bien viendo las comparaciones para las combinaciones de factor de nivel de 2×3 (AB). Si el efecto de interacción *no es significativo*, entonces la significancia de las medias de efecto principal debe investigarse, primero con la prueba F general y luego con el método de Tukey para comparaciones en pares o intervalos de confianza específicos. Recuerde que estos procedimientos de análisis de varianza siempre usan $s^2 = \text{MSE}$ como el mejor estimador de σ^2 con grados de libertad iguales a $gl = ab(r - 1)$.

Por ejemplo, usando el medidor de Tukey para comparar el promedio de producciones para los dos supervisores en cada uno de los tres turnos, se podría calcular

$$\omega = q_{.05}(6, 12) \left(\frac{s}{\sqrt{r}} \right) = 4.75 \left(\frac{\sqrt{720}}{\sqrt{3}} \right) = 73.59$$

Como los tres pares de medias (602 y 487 en el turno diurno, 498 y 602 en el turno vespertino y 450 y 657 en el turno nocturno) difieren en más de ω , nuestras conclusiones prácticas han sido confirmadas estadísticamente.

7.10

EJERCICIOS

TÉCNICAS BÁSICAS

7.45 Suponga que se debe realizar un experimento factorial de dos factores, el factor A a cuatro niveles y el factor B a cinco niveles, con tres réplicas por tratamiento.

- ¿Cuántos tratamientos intervienen en el experimento?
- ¿Cuántas observaciones están involucradas?
- Haga una lista de las fuentes de variación y sus respectivos grados de libertad.

7.46 El análisis de la tabla de varianza para un experimento factorial de 3×4 , con el factor A en tres niveles y el factor B en cuatro niveles y con dos observaciones por tratamiento, se muestra a continuación:

Fuente	gl	SS	MS	F
	2	5.3		
	3	9.1		
	6			
	12	24.5		
Total	23	43.7		

- Llene los renglones faltantes de la tabla.
- ¿Los datos dan suficiente evidencia para indicar que los factores A y B interactúan? Pruebe usando $\alpha = .05$. ¿Cuáles son las implicaciones prácticas de su respuesta?

- c. ¿Los datos dan suficiente evidencia para indicar que los factores A y B afectan la variable x de la respuesta? Explique.

7.47 Consulte el ejercicio 7.46. Las medias de dos de las combinaciones a nivel de factor, es decir A_1B_1 y A_2B_1 , son $\bar{x}_1 = 8.3$ y $\bar{x}_2 = 6.3$, respectivamente. Encuentre un intervalo de confianza de 95% para la diferencia entre las dos medias poblacionales correspondientes.

DATOS
EX0748

7.48 La tabla siguiente contiene datos para un experimento factorial de 3×3 , con dos réplicas por tratamiento:

Niveles del factor B	Niveles del factor A		
	1	2	3
1	5, 7	9, 7	4, 6
2	8, 7	12, 13	7, 10
3	14, 11	8, 9	12, 15

- Efectúe un análisis de varianza para los datos y presente los resultados en una tabla de análisis de varianza.
- ¿A qué nos referimos cuando decimos que los factores A y B interactúan?
- ¿Los datos dan suficiente evidencia para indicar interacción entre los factores A y B? Pruebe usando $\alpha = .05$.
- Encuentre el valor p aproximado para la prueba de la parte c.
- ¿Cuáles son las implicaciones prácticas de sus resultados en la parte c? Explique sus resultados usando una gráfica de línea similar a la de la figura 7.11.

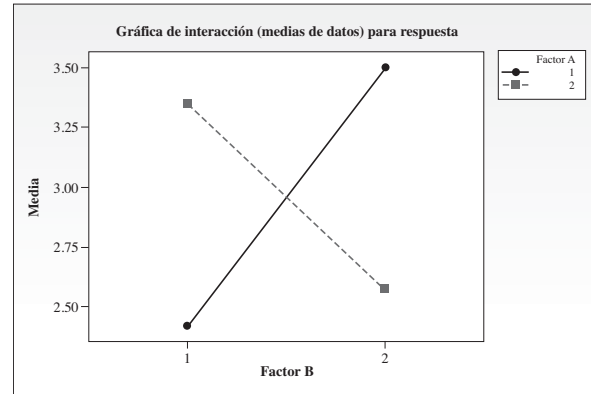
DATOS
EX0749

7.49 Factorial de 2×2 La tabla contiene datos para un experimento factorial de 2×2 , con cuatro réplicas por tratamiento.

Niveles del factor B	Niveles del factor A	
	1	2
1	2.1, 2.7, 2.4, 2.5	3.7, 3.2, 3.0, 3.5
2	3.1, 3.6, 3.4, 3.9	2.9, 2.7, 2.2, 2.5

- La gráfica siguiente fue generada por el MINITAB. Verifique que los cuatro puntos que enlazan las dos rectas sean las medias de las cuatro observaciones dentro de cada combinación del nivel de factor. ¿Qué dice la gráfica acerca de la interacción entre los factores A y B?

Gráfica de interacción MINITAB para el ejercicio 7.49



- Use la salida MINITAB para probar una interacción significativa entre A y B. ¿Esto confirma sus conclusiones de la parte a?

Salida MINITAB para el ejercicio 7.49

ANOVA de dos vías: respuesta contra factor A, factor B

Fuente	DF	SS	MS	F	P
Factor A	1	0.0000	0.00000	0.00	1.000
Factor B	1	0.0900	0.09000	1.00	0.338
Interacción	1	3.4225	3.42250	37.85	0.000
Error	12	1.0850	0.09042		
Total	15	4.5975			

S = 0.3007 R-Sq = 76.40% R-Sq(adj) = 70.50%

- Considerando sus resultados en la parte b, ¿cómo puede explicar el hecho de que ninguno de los efectos principales sea significativo?
- Si se encuentra una interacción significativa, ¿es necesario probar por diferencias significativas del efecto principal? Explique.
- Escriba un párrafo breve que resuma los resultados de este experimento.

APLICACIONES

7.50 Visualización del terreno Se realizó un estudio para determinar el efecto de dos factores sobre el entrenamiento para soldados sobre visualización del terreno.⁴ Durante los programas de entrenamiento los participantes vieron mapas de contorno de varios terrenos y luego se les permitió ver una reconstrucción computarizada del terreno como se vería desde un ángulo específico. Los dos factores que se investigaron en el experimento fueron las capacidades espaciales de los participantes (para visualizar en tres dimensiones) y los procedimientos de visualización (activos o pasivos). La participación activa permitió a los

individuos ver las reconstrucciones del terreno, generadas por computadora, desde cualquiera y de todos los ángulos. La participación pasiva dio a los participantes un conjunto de reconstrucciones del terreno previamente seleccionadas. Los participantes fueron examinados de acuerdo con su capacidad espacial y de las calificaciones del examen 20 fueron clasificados como poseedores de elevada capacidad espacial, 20 regulares y 20 bajos. A continuación, 10 participantes dentro de cada uno de estos grupos fueron asignados a cada uno de los dos modos de entrenamiento, activo o pasivo. Las tablas siguientes son la tabla ANOVA calculada por los investigadores y la tabla de las medias de tratamiento.

Fuente	gl	MS	Error gl	F	p
EFECTOS PRINCIPALES:					
Condición de entrenamiento	1	103.7009	54	3.66	.0610
Capacidad	2	760.5889	54	26.87	.0005
Interacción:					
Condición de entrenamiento × Capacidad	2	124.9905	54	4.42	.0167
Dentro de celdas	54	28.3015			

Condición de entrenamiento		
Capacidad espacial	Activa	Pasiva
Alta	17.895	9.508
Media	5.031	5.648
Baja	1.728	1.610

Nota: Calificación máx = 36.

- Explique cómo llegaron los autores a los grados de libertad indicados en la tabla ANOVA.
- ¿Son correctos los valores F ?
- Interprete los resultados de la prueba. ¿Cuáles son las implicaciones prácticas?
- Use la tabla 6 del Anexo para aproximar los valores p para los estadísticos F indicados en la tabla ANOVA.

Fuente: H. F. Barsam y Z. M. Simutis (1984). "Computer-Based Graphics for Terrain Visualization Training". *Human Factors*, 26. Copyright 1984 por the Human Factors Society, Inc. Reproducido con permiso.

DATOS
EX0751

7.51 ¿La animación ayuda? Para explorar formas de incrementar la experiencia educativa usando animación y no imágenes estáticas en un ambiente de aprendizaje, Cyril Rebetez y sus colegas⁵ realizaron un experimento factorial que midió la retención de información en cuatro condiciones factoriales: con animación o sin animación; y reforzamiento por medio de instantáneas o sin instantáneas de los cuadros importantes en la animación. Se esperaba que la animación conduciría a una mejor retención de la información y que tener disponibles las instantáneas también ayudaría a la

retención de información. Los siguientes datos se basan en los resultados de su experimento:

	Escenario de aprendizaje			
	Estático		Animado	
Instantáneas	Sin	Con	Sin	Con
	58.9	42.0	57.7	64.3
	48.9	53.9	55.9	66.4
	51.8	54.4	57.2	63.1
	53.0	47.6	65.1	55.8
	51.3	50.5	59.3	57.9
	49.8	50.2	65.7	61.5
	61.5	47.0	60.8	61.2
	47.8	52.4	59.3	61.9

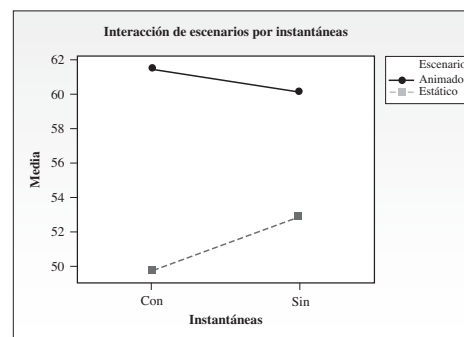
Use la salida del *MINITAB* que sigue para analizar el experimento con el método apropiado. Identifique los dos factores e investigue cualesquier efectos posibles debidos a su interacción o los efectos principales. ¿Cuáles son las implicaciones prácticas de estos resultados? ¿Por qué las gráficas de la interacción parecen contradictorias respecto al análisis?

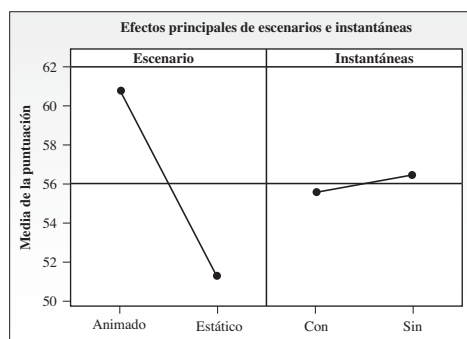
Si el efecto de la interacción es real, ¿qué podría hacer usted como experimentador para mostrar que la interacción es, de hecho, significativa? Explique sus conclusiones en un informe.

Salida del *MINITAB* para el ejercicio 7.51

ANOVA: retención contra escenario, instantáneas

Factor	Tipo	Niveles	Valores		
Escenario	Fijo	2	Animado, Estático		
Instantáneas	Fijo	2	Con, Sin		
Análisis de varianza para la retención					
Fuente	DF	SS	MS		
Escenario	1	722.95	722.95	F	P
Instantáneas	1	6.04	6.04	44.56	0.000
Escenario*				0.37	0.547
Instantáneas	1	40.73	40.73		
Error	28	454.28	16.22	2.51	0.12 4
Total	31	1223.99			
S = 4.02793 R-Sq = 62.89% R-Sq(adj) = 58.91%					





DATOS
EX0752

7.52 Calificaciones de examen de cuarto grado

Una junta directiva local estaba interesada en comparar las calificaciones en un examen estandarizado de lectura para estudiantes de cuarto grado en su distrito. Seleccionaron una muestra aleatoria de cinco hombres y cinco mujeres, estudiantes de cuarto grado en cada una de cuatro escuelas elementales diferentes del distrito y registraron las calificaciones del examen. Los resultados se presentan en la tabla siguiente.

Género	Escuela 1	Escuela 2	Escuela 3	Escuela 4
Hombre	631	642	651	350
	566	710	611	565
	620	649	755	543
	542	596	693	509
	560	660	620	494
Mujer	669	722	709	505
	644	769	545	498
	600	723	657	474
	610	649	722	470
	559	766	711	463

- ¿Qué tipo de diseño experimental es éste? ¿Cuáles son las unidades experimentales? ¿Cuáles son los factores y niveles de interés para la junta directiva?
- Efectúe el análisis de varianza apropiado para este experimento.
- ¿Los datos indican que el efecto del género en el promedio de calificaciones del examen es diferente, dependiendo de la escuela del estudiante? Pruebe la hipótesis apropiada usando $\alpha = .05$.
- Grafique el promedio de calificaciones usando una gráfica de interacción. ¿Cómo describiría usted el efecto del género y escuela en el promedio de las calificaciones del examen?
- ¿Los datos indican que cualquiera de los efectos principales es significativo? Si el efecto principal es significativo, use el método de Tukey de comparaciones en pares para examinar las diferencias en detalle. Use $\alpha = .01$.

7.53 Capacitación gerencial Se realizó un experimento para investigar el efecto de la capacitación gerencial sobre la habilidad de los supervisores para tomar decisiones en una gran compañía. Se seleccionaron 16 supervisores y ocho fueron escogidos al azar para recibir capacitación administrativa. Cuatro supervisores capacitados y cuatro no capacitados se seleccionaron al azar para funcionar en una situación en la que surgió un problema común. A los otros ocho supervisores se les presentó una situación de emergencia en la que los procedimientos comunes no podían usarse. La respuesta fue una clasificación de conducta administrativa para cada supervisor, evaluada por un esquema de calificación diseñado por el experimentador.

- ¿Cuáles son las unidades experimentales de este experimento?
- ¿Cuáles son los dos factores considerados en el experimento?
- ¿Cuáles son los niveles de cada factor?
- ¿Cuántos tratamientos hay en el experimento?
- ¿Qué tipo de diseño experimental se ha empleado?

DATOS
EX0754

7.54 Capacitación gerencial, continúa

Consulte el ejercicio 7.53. Los datos para este experimento se muestran en la tabla.

Situación (B)	Capacitación (A)		Totales
	Capacitado	No capacitado	
Común	85	53	519
	91	49	
	80	38	
	78	45	
Emergencia	76	40	473
	67	52	
	82	46	
	71	39	
Totales	630	362	992

- Construya la tabla ANOVA para este experimento.
- ¿Hay una interacción significativa entre la presencia o ausencia de capacitación y el tipo de situación de toma de decisiones? Pruebe al nivel de 5% de significancia.
- ¿Los datos indican una diferencia significativa en calificaciones de conducta para los dos tipos de situaciones al nivel de significancia de 5%?
- ¿Las calificaciones de conducta difieren significativamente para los dos tipos de categorías de capacitación al nivel de significancia de 5%?
- Grafique el promedio de calificaciones con el uso de una gráfica de interacción. ¿Cómo describiría usted el efecto de la capacitación y la situación de emergencia en las habilidades de toma de decisiones de los supervisores?

REPASO DE LAS SUPOSICIONES DEL ANÁLISIS DE VARIANZA

En la sección 7.3 usted aprendió que las suposiciones y los procedimientos de prueba para el análisis de varianza son semejantes a los requeridos para las pruebas t y F del capítulo 6, es decir, que las observaciones dentro de un grupo de tratamiento deben estar distribuidas normalmente con varianza común σ^2 . También aprendió que los procedimientos del análisis de varianza son bastante robustos cuando los tamaños muestrales son iguales y los datos tienen forma de montículo. Si éste es el caso, un modo de protegerse de conclusiones imprecisas es que dentro de lo posible se trate de seleccionar muestras de tamaños iguales.

Hay algunas formas rápidas y sencillas de comprobar los datos en cuanto a la violación de suposiciones. Primero veamos el tipo de variable de respuesta que se mide. El experimentador podría de inmediato ver un problema con la normalidad o con la suposición de varianza común. Puede que no sea posible medir *cuantitativamente* los datos que haya recolectado. Muchas respuestas, por ejemplo preferencias de productos, se pueden clasificar sólo como “A es mejor que B” o “C es la menos preferible”. Los datos que sean *cualitativos* no pueden tener una distribución normal. Si la variable de respuesta es *discreta* y puede tomar sólo tres valores, es decir 0, 1 o 2, entonces no es razonable suponer que la variable de respuesta esté normalmente distribuida.

Suponga que la variable de respuesta es binomial, es decir, la proporción p de personas a favor de un tipo particular de inversión (véase la sección 4.6). Aun cuando los datos binomiales pueden tener aproximadamente la forma de montículo en ciertas condiciones, violan la suposición de igual varianza. La varianza de una proporción muestral es

$$\sigma^2 = \frac{pq}{n} = \frac{p(1-p)}{n}$$

de modo que la varianza cambia dependiendo del valor de p . Cuando cambian las medias de tratamiento, el valor de p cambia y también cambia la varianza σ^2 . Una situación similar ocurre cuando la variable de respuesta es una variable aleatoria de Poisson, por ejemplo, el número de accidentes industriales por mes en una planta manufacturera (véase la sección 3.5). Como la varianza de una variable aleatoria de Poisson es $\sigma^2 = \mu$, la varianza cambia exactamente como cambia la media de tratamiento.

Si el experimentador no puede ver algunas violaciones flagrantes en el tipo de datos que se miden, vea el rango de los datos dentro de cada grupo de tratamiento. Si estos rangos son casi iguales, entonces es probable que la suposición de varianza común sea razonable. Para ver si hay normalidad se puede hacer una rápida gráfica de puntos o de tallo y hoja para un grupo particular de tratamiento. Sin embargo, es bastante frecuente que no se tengan suficientes mediciones para obtener una gráfica razonable.

Si el experimentador utiliza un programa de cómputo para analizar su experimento, hay algunas valiosas **herramientas de diagnóstico** que puede usar. Estos procedimientos son demasiado complicados para hacer cálculos manualmente, pero son fáciles de usar cuando la computadora hace el trabajo.

Gráficas residuales

En el análisis de varianza, la variación total en los datos se divide en varias partes, dependiendo de los factores identificados como importantes para el investigador. Una vez que los efectos de estas fuentes de variación hayan sido removidos, la variabilidad “sobrante” en cada observación se denomina **residual** para ese punto de datos. Estos residuales representan **error experimental**, la variabilidad básica en el experimento y deben tener una *distribución aproximadamente normal* con una media de 0 y la *misma variación* para cada grupo de tratamiento. Casi todos los paquetes de computadora dan opciones para graficar estos residuales:

- La **gráfica de probabilidad normal de residuales** es aquella que grafica los residuales para cada observación contra el valor esperado de ese residual *si hubiera venido de una distribución normal*. Si los residuales son aproximadamente normales, la gráfica será muy semejante a una *línea recta*, con pendiente hacia arriba a la derecha.
- La **gráfica de residuales contra ajuste o residuales contra variables** es aquella que grafica los residuales contra el valor esperado de esa observación *usando el diseño experimental que hemos empleado*. Si no se han violado suposiciones y no hay fuentes “sobrantes” de variación que no sean el error experimental, esta gráfica debería mostrar una *dispersión aleatoria* de puntos alrededor de la “recta de error cero” horizontal para cada grupo de tratamiento, con aproximadamente la misma dispersión vertical.

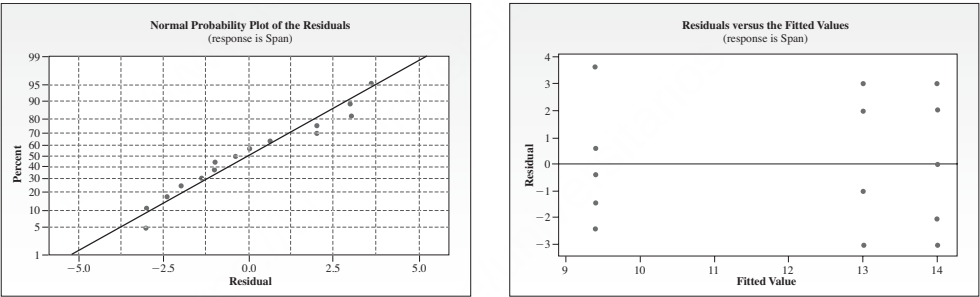
EJEMPLO

7.13

Los datos del ejemplo 7.4, que se refieren a los intervalos de atención de tres grupos de estudiantes de educación básica, se analizaron usando el *MINITAB*. Las gráficas de la figura 7.13, generadas por *MINITAB*, son la gráfica de probabilidad normal y los residuales contra la gráfica de ajuste para este experimento. Observe la forma de línea recta de la gráfica de probabilidad normal, que indica una distribución normal en los residuales. En la otra gráfica, los residuales están graficados contra los valores esperados estimados, que son los promedios muestrales para cada uno de los tres tratamientos en el diseño completamente aleatorizado. La dispersión aleatoria alrededor de la “recta de cero error” horizontal y la dispersión constante indican que *no hay violaciones* en la suposición de varianza constante.

FIGURA 7.13

Gráficas *MINITAB* de diagnóstico para el ejemplo 7.13



EJEMPLO

7.14

Una compañía planea promover un nuevo producto usando una de tres campañas publicitarias. Para investigar el grado de reconocimiento del producto a partir de estas tres campañas, se seleccionaron 15 zonas de mercado y cinco se asignaron al azar a cada plan de publicidad. Al terminar las campañas de publicidad, se seleccionaron muestras aleatorias de 400 adultos en cada zona y se registraron las proporciones que estaban familiarizadas con el nuevo producto, como en la tabla 7.7. ¿Han sido violadas algunas de las suposiciones del análisis de varianza en este experimento?

TABLA 7.7

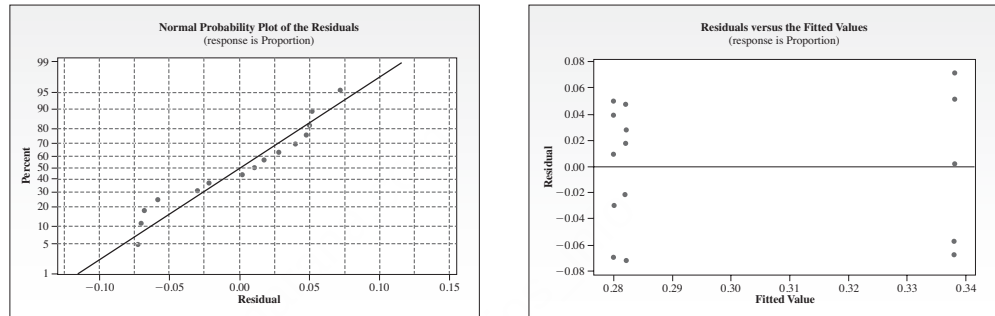
Proporciones de reconocimiento de un producto para tres campañas publicitarias

Campaña 1	Campaña 2	Campaña 3
.33	.28	.21
.29	.41	.30
.21	.34	.26
.32	.39	.33
.25	.27	.31

Solución El experimento está bosquejado como un *diseño completamente aleatorizado*, pero la variable de respuesta es una proporción muestral binomial. Esto indica que la normalidad y las suposiciones de varianza común podrían ser inválidas. Vea la gráfica de probabilidad normal de los residuales y la gráfica de residuales contra ajuste que se ve en la figura 7.14. La forma curva de la gráfica de probabilidad normal indica que los residuales *no tienen una distribución normal*. En la gráfica de residual contra ajuste, se pueden ver tres rectas verticales de residuales, una por cada una de las tres campañas publicitarias. Observe que dos de las rectas (campañas 1 y 3) son cercanas y tienen dispersión similar, pero la tercera recta (campaña 2) está más alejada hacia la derecha, lo cual indica una proporción muestral más grande y en consecuencia una *varianza más grande* en este grupo. Las suposiciones de ambos análisis de varianza son sospechosas en este experimento.

FIGURA 7.14

Gráficas MINITAB de diagnóstico para el ejemplo 7.14



¿Qué se puede hacer cuando las suposiciones ANOVA no se satisfacen? La suposición de *varianza constante* a veces puede corregirse al **transformar** las mediciones de respuesta. Esto es, en lugar de usar las mediciones originales, se pueden usar sus raíces cuadradas, logaritmos o alguna otra función de la respuesta. Las transformaciones que tienden a estabilizar la varianza de la respuesta también hacen sus distribuciones casi normales en mayor medida.

Cuando no se pueda hacer nada para *siquiera aproximadamente* satisfacer las suposiciones ANOVA o si los datos son clasificaciones, se deben usar procedimientos **no paramétricos** de prueba y estimación, mismos que se presentan en el capítulo 11. Hemos mencionado estos procedimientos antes; son casi tan potentes para detectar diferencias de tratamientos como las pruebas presentadas en este capítulo cuando los datos están normalmente distribuidos. Cuando las suposiciones paramétricas ANOVA se violan, las pruebas no paramétricas son por lo general más potentes.

7.12

UN BREVE REPASO

Hemos presentado tres diseños experimentales diferentes en este capítulo, cada uno puede ser analizado usando el procedimiento de análisis de varianza. El objetivo del análisis de varianza es detectar diferencias en las respuestas medias para unidades experimentales que han recibido diferentes tratamientos, es decir, combinaciones diferentes de los niveles de factor experimental. Una vez realizada una prueba general de las diferencias, la naturaleza de estas diferencias (si existe alguna) se pueden explorar usando métodos de comparaciones en pares y/o procedimientos de estimación de intervalo.

Los tres diseños que se han mostrado en este capítulo representan sólo una breve introducción al tema del análisis de experimentos diseñados. Los diseños están disponibles para los experimentos que contienen diversas variables de diseño, así como más de dos factores de tratamiento y otros diseños más complejos. Recuerde que las **variables de diseño** son factores cuyo efecto se desea controlar y por tanto eliminar del error experimental, en tanto que

las **variables de tratamiento** son factores cuyo efecto se desea investigar. Si su experimento está diseñado en forma apropiada, usted podrá analizarlo usando el análisis de varianza. Los experimentos en los que los niveles de una variable son *medidos experimentalmente* en lugar de *controlados* o *preseleccionados* antes de tiempo, pueden analizarse usando **análisis de regresión lineal o múltiple**, que es el tema de los capítulos 8 y 9.

REPASO DEL CAPÍTULO

Conceptos y fórmulas clave

I. Diseños experimentales

1. Unidades experimentales, factores, niveles, tratamientos, variables de respuesta.
2. Suposiciones: las observaciones dentro de cada grupo de tratamiento deben estar normalmente distribuidas con una varianza común σ^2 .
3. Clasificación en una dirección, diseño completamente aleatorizado: las muestras aleatorias independientes se seleccionan de entre cada una de las k poblaciones.
4. Clasificación en dos direcciones; diseño de bloque aleatorizado: k tratamientos se comparan dentro de b grupos relativamente homogéneos de unidades experimentales llamadas *bloques*.
5. Clasificación en dos direcciones, experimento factorial $a \times b$: dos factores, A y B, se comparan a varios niveles. Cada combinación factor-nivel se replica r veces para permitir la investigación de una interacción entre los dos factores.

II. Análisis de varianza

1. La variación total en el experimento está dividida en variación (sumas de cuadrados) explicada por los diversos factores experimentales y variación debida a error experimental (no explicado).
2. Si hay un efecto debido a un factor particular, su media cuadrática ($MS = SS/gl$) por lo general es grande y $F = MS(\text{factor})/MSE$ es grande.
3. Los estadísticos de prueba para los diversos factores experimentales están basados en estadísticos F , con grados de libertad apropiados ($gl_2 = \text{grados de libertad de error}$).

III. Interpretación de un análisis de varianza

1. Para el diseño de bloque completamente aleatorizado y el aleatorizado, en cada factor se prueba su significancia.
2. Para el experimento factorial, primero pruebe para interacción significativa. Si la interacción es significativa, no es necesario probar efectos principales. La naturaleza de las diferencias en las combinaciones factor-nivel deben examinarse más.
3. Si se encuentra una diferencia significativa en las medias poblacionales, el método de Tukey de comparaciones en pares o un método semejante se pueden usar para identificar más la naturaleza de las diferencias.
4. Si el experimentador tiene especial interés en una media poblacional o en la diferencia entre dos medias poblacionales, puede usar una estimación de intervalo de confianza. (Para un diseño de bloque aleatorizado, los intervalos de confianza no dan estimaciones insesgadas para medias poblacionales individuales.)

IV. Verificación del análisis de suposiciones de varianza

1. Para verificar la normalidad, use la gráfica de probabilidad normal para los residuales. Los residuales deben exhibir una forma de línea recta, creciendo hacia arriba a la derecha.
2. Para verificar la igualdad de varianza, use los residuales contra una gráfica de ajuste. La gráfica debe exhibir una dispersión aleatoria, con la misma dispersión vertical alrededor de la "línea de error cero" horizontal.



Procedimientos de análisis de varianza: *Microsoft Excel*

Los procedimientos estadísticos para realizar el análisis de varianza para los tres diseños experimentales que se expusieron en este capítulo pueden encontrarse usando el comando de *Microsoft Excel* **Data ► Data Analysis**. Verá opciones para **Single Factor**, **Two-Factor Without Replication** y **Two-Factor With Replication** que generarán cuadros de diálogo usados para los diseños completamente aleatorizado, de bloque aleatorizado y factorial, respectivamente.

EJEMPLO

7.15

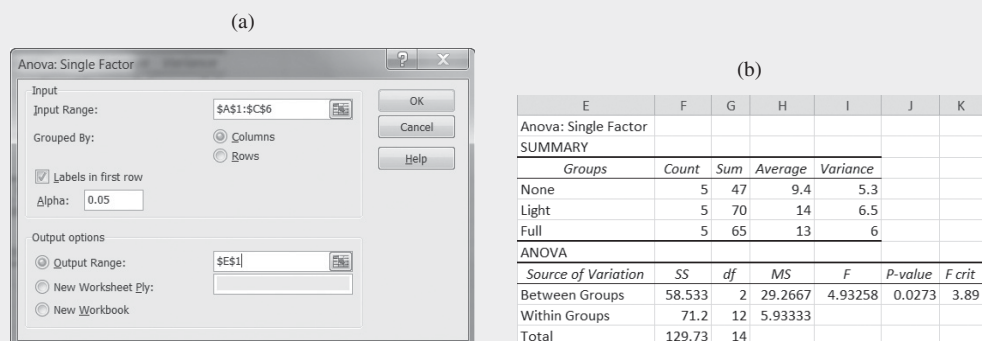
(Diseño completamente aleatorizado) Consulte el estudio del desayuno en el ejemplo 7.4, en el cual se estudió el efecto de la nutrición en el intervalo de atención.

Sin desayuno	Desayuno ligero	Desayuno completo
8	14	10
7	16	12
9	12	16
13	17	15
10	11	12

Introduzca los datos en las columnas A, B y C de una hoja de cálculo de *Excel* con una muestra por columna.

1. Use **Data ► Data Analysis ► Anova: Single Factor** para generar el cuadro de diálogo de la figura 7.15a). Resalte o teclee el **Input Range** (los datos en las primeras tres columnas) en el primer cuadro. En la sección marcada “Grouped by” elija el botón de radio para **Columns** y seleccione “Labels” si es necesario.
2. En *Excel*, el nivel de significancia por omisión es $\alpha = .05$. Cambie este nivel de significancia si es necesario. Introduzca una ubicación de celda para el **Output Range** y haga clic en **OK**. La salida aparecerá en la ubicación de celda seleccionada, y deberá ajustarse usando **Format ► AutoFit Column Width** en la pestaña **Home** en el grupo **Cells** mientras todavía está resaltado. Puede disminuir la precisión decimal si lo desea, usando $\rightarrow .0$ en la pestaña **Home** en el grupo **Number** (véase la figura 7.15b)).
3. El valor observado del estadístico de prueba $F = 4.93$ se encuentra en la fila etiquetada “Between Groups” seguido por el “P-value” y el valor crítico marcando la región de rechazo para una prueba de una cola con $\alpha = .05$. Para este ejemplo, el valor $p = .0273$ indica que hay una diferencia significativa en los intervalos de atención promedio dependiendo del tipo de desayuno.

FIGURA 7.15



EJEMPLO 7.16

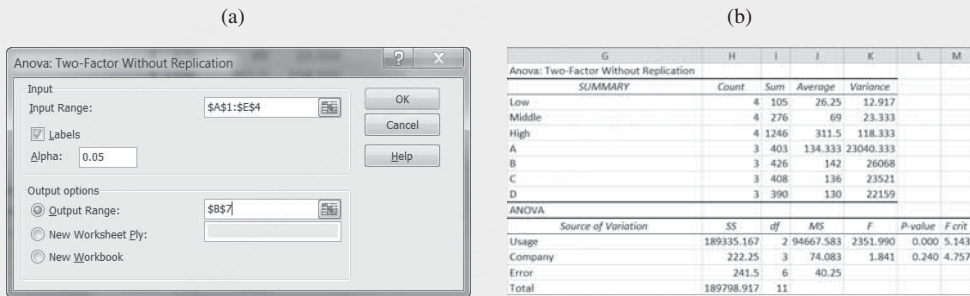
(Diseño de bloque aleatorizado) Consulte el estudio de teléfonos celulares del ejemplo 7.8, en el cual se estudió el efecto del nivel de uso en el costo para cuatro compañías diferentes.

Nivel de uso	Compañía			
	A	B	C	D
Bajo	27	24	31	23
Medio	68	76	65	67
Alto	308	326	312	300

Introduzca los datos en las columnas A–E de una hoja de cálculo de *Excel*, usando la columna A para las etiquetas de uso y la fila 1 para las etiquetas de la compañía, tal como se muestra en la tabla anterior.

1. Use **Data ► Data Analysis ► Anova: Two-Factor Without Replication** para generar el cuadro de diálogo que se muestra en la figura 7.16a). Resalte o teclee el **Input Range** (los datos en las primeras cinco columnas) en el primer cuadro y seleccione “Labels” si es necesario. Cambie el nivel de significancia si es necesario y haga clic en **OK**. Puede ajustar la salida, posiblemente cambiando las etiquetas “Rows” y “Columns” a “Uso” y “Compañía”, como se muestra en la figura 7.16b).
2. El valor observado del estadístico de prueba para tratamientos (compañías) es $F = 1.84$ con valor $p = .240$, lo que indica que no hay una diferencia significativa entre las cuatro compañías. La prueba para bloques (uso) es altamente significativa, con un valor $p = .000$.

FIGURA 7.16



EJEMPLO 7.17

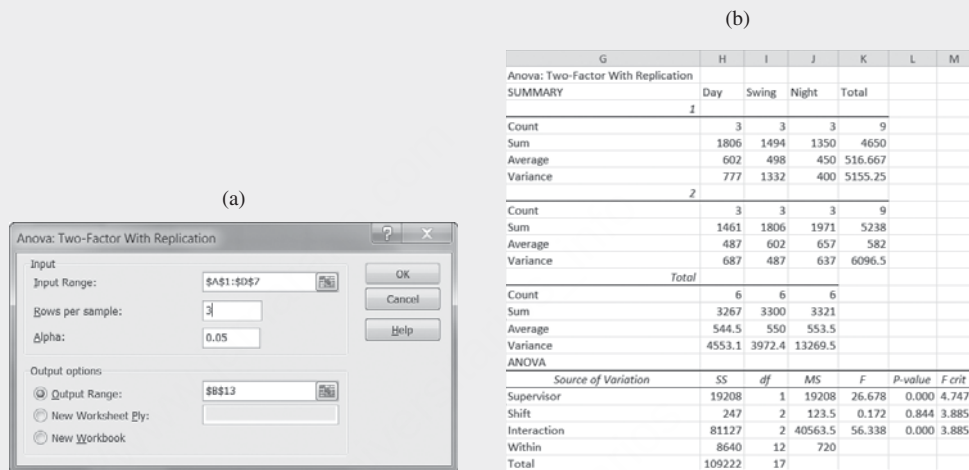
(Experimento factorial) Consulte el estudio de las salidas de producción en el ejemplo 7.12, en el cual se estudió el efecto del supervisor y el turno en la salida de producción.

Supervisor	Turno		
	Diurno	Vespertino	Nocturno
1	571	480	470
	610	474	430
	625	540	450
2	480	625	630
	516	600	680
	465	581	661

Introduzca los datos en las columnas A–D de una hoja de cálculo de *Excel*, usando la columna A para las etiquetas de supervisor y la fila 1 para las etiquetas de turno, como se muestra en la tabla anterior.

1. Use **Data ► Data Analysis ► Anova: Two-Factor With Replication** para generar el cuadro de diálogo que se ve en la figura 7.17a). Resalte o teclee el **Input Range** (los datos en las primeras cuatro columnas) en el primer cuadro. Introduzca el número de réplicas (3) en “Rows per Sample”, cambie el nivel de significancia si es necesario y haga clic en **OK**. Puede ajustar la salida, posiblemente cambiando las etiquetas “Sample” y “Columns” por “Supervisor” y “Turno”, como se muestra en la figura 7.17b).
2. Consulte la tabla ANOVA en la parte inferior de la salida impresa. Hay una interacción significativa entre el turno y el supervisor (valor $p = .000$). Ahora pueden estudiarse las diferencias en las medias de tratamiento observando las comparaciones para las $3 \times 2 = 6$ combinaciones del nivel de factor.

FIGURA 7.17



(NOTA: *MS Excel* no proporciona opciones para ejecutar una prueba de Tukey o para generar gráficas de diagnóstico.

Procedimientos de análisis de varianza: *MINITAB*

Los procedimientos estadísticos empleados para efectuar el análisis de varianza para los tres diseños experimentales diferentes en este capítulo se encuentran usando el comando de *MINITAB* **Stat ► ANOVA**. El usuario verá opciones para **One-way**, **One-way (Unstacked)** y **Two-way** que generará cuadros de diálogo empleados para diseños completamente aleatorizados, de bloque aleatorizado y factoriales, respectivamente.

EJEMPLO

7.18

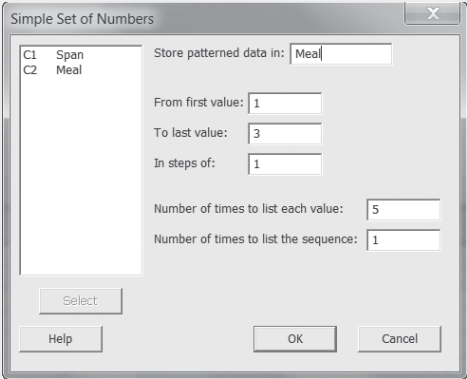
(Diseño completamente aleatorizado) Consulte el estudio del desayuno en el ejemplo 7.4, en el cual se estudió el efecto de la nutrición en la duración de los intervalos de atención.

No desayuno	Desayuno ligero	Desayuno completo
8	14	10
7	16	12
9	12	16
13	17	15
10	11	12

1. Introduzca los 15 intervalos de atención registrados en la columna C1 de una hoja de cálculo de *MINITAB* y nómbrela “Intervalo”. A continuación, introduzca los enteros 1, 2 y 3 en una segunda columna C2 para identificar la asignación de alimento (*tratamiento*)

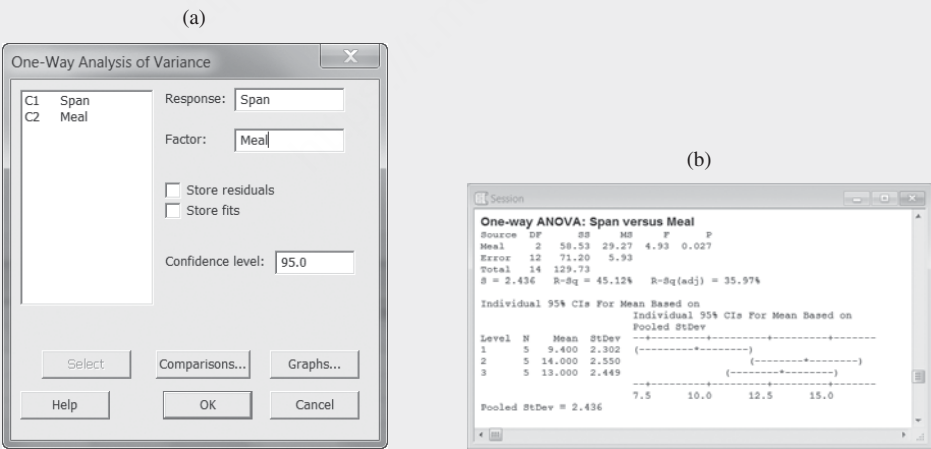
para cada observación. Puede dejar que MINITAB establezca este patrón por usted usando **Calc ► Make Patterned Data ► Simple Set of Numbers** e introduciendo los números apropiados, como se muestra en la figura 7.18.

FIGURA 7.18



1. Use **Stat ► ANOVA ► One-Way** para generar el cuadro de diálogo que se observa en la figura 7.19a).[†] Seleccione la columna de observaciones para el cuadro “Response” y la columna de indicadores de tratamiento para el cuadro “Factor”.
2. Ahora tiene varias opciones. Bajo **Comparisons**, puede seleccionar “Tukey’s family error rate” (que tiene un nivel por omisión de 5%) para obtener una salida de comparaciones en pares. Bajo **Graphs**, puede seleccionar gráficas de valor individual y/o gráficas de caja para comparar las tres asignaciones de alimento, y puede generar gráficas de residuales (use “Normal plot of residuals” y/o “Residuals versus fits”) para verificar la validez de las suposiciones del ANOVA. Haga clic en **OK** desde el cuadro de diálogo principal para obtener la salida de la figura 7.19b).

FIGURA 7.19



[†]Si se introdujo cada una de las tres muestras en columnas separadas, el comando apropiado sería **Stat ► ANOVA ► One-Way (Unstacked)**.

3. El valor observado del estadístico de prueba $F = 4.93$ se encuentra en la fila etiquetada “Alimento” seguida por el valor $p = .0273$. Con $\alpha = .05$, hay una diferencia significativa en los intervalos de atención promedio dependiendo del tipo de desayuno.

El comando **Stat ► ANOVA ► Two-Way** puede usarse tanto para el diseño de bloque aleatorizado como para el diseño factorial. Primero deben introducirse todas las observaciones en una sola columna y luego los enteros o nombres descriptivos para indicar cualquiera de estos casos:

- El *bloque* y *tratamiento* para cada una de las mediciones en un diseño de bloque aleatorizado.
- Los niveles de *los factores A y B* para el experimento factorial.

MINITAB reconocerá un número de réplicas dentro de cada combinación de nivel de factor en el experimento factorial y desglosará la suma de cuadrados para interacción siempre que no se seleccione el cuadro “Fit additive model”. Como estos dos diseños implican la misma secuencia de comandos, se usarán los datos del ejemplo 7.12 para generar el ANOVA.

EJEMPLO

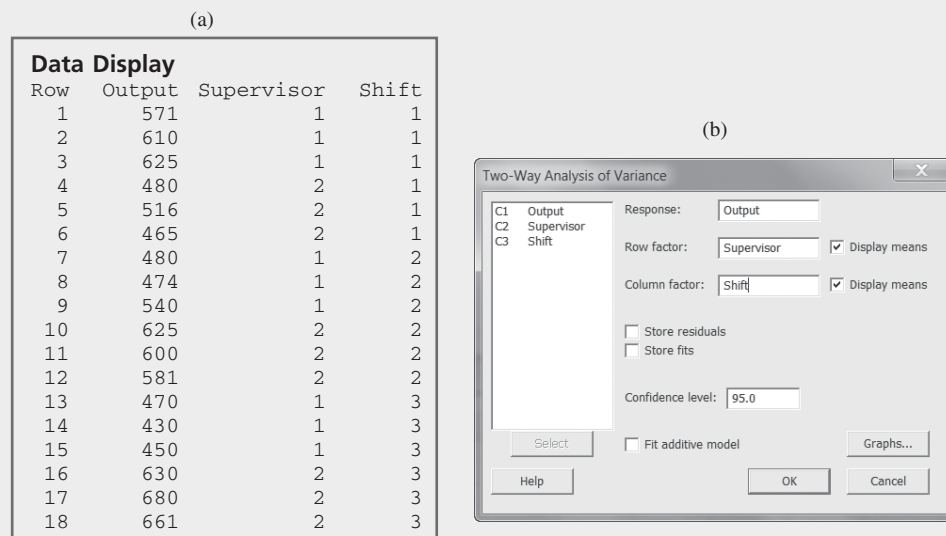
7.19

(Clasificación de dos vías) Consulte el estudio de salidas de producción en el ejemplo 7.12, en el cual se estudió el efecto del supervisor y el turno en las salidas de producción.

Supervisor	Turno		
	Diurno	Vespertino	Nocturno
1	571	480	470
	610	474	430
	625	540	450
2	480	625	630
	516	600	680
	465	581	661

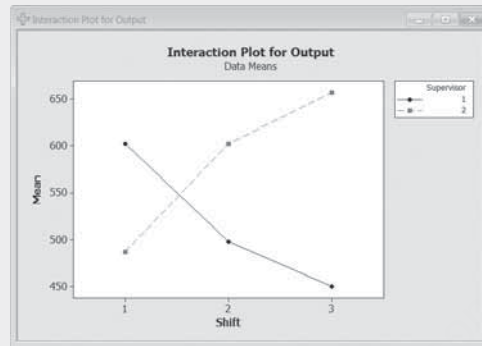
1. Introduzca los datos en la hoja de cálculo como se muestra en la figura 7.20a). Vea si puede usar **Calc ► Make Patterned Data ► Simple Set of Numbers** para introducir los datos en las columnas C2–C3.
2. Use **Stat ► ANOVA ► Two-Way** para generar el cuadro de diálogo que se muestra en la figura 7.20b). Escoja “Output” para el cuadro “Response”, y “Supervisor” y “Turno” para “Row Factor” y “Column Factor”, respectivamente. Puede elegir desplegar las medias del efecto principal junto con intervalos de confianza de 95% seleccionando “Display means”, y puede seleccionar gráficas de residuales si lo desea. Haga clic en **OK** para obtener la salida impresa ANOVA de la figura 7.12a).

FIGURA 7.20



3. En vista que la interacción entre supervisores y turnos es altamente significativa, puede desear explorar la naturaleza de esta interacción al graficar el promedio de producción para cada supervisor en cada uno de los tres turnos. Use **Stat ► ANOVA ► Interaction Plot** y elija la respuesta y las variables de factor apropiadas. La gráfica se muestra en la figura 7.21. Puede ver la diferencia fuerte en los comportamientos de las producciones medias para los dos supervisores, lo que indica una interacción fuerte entre los dos factores.

FIGURA 7.21



Ejercicios suplementarios

DATOS
EX0755

7.55 Tiempos de reacción contra estímulos Veintisiete personas participaron en un experimento para comparar los efectos de cinco diferentes estímulos sobre el tiempo de reacción. El experimento se corrió usando un diseño completamente aleatorizado y, cualquiera que fuera el resultado del análisis de varianza, los experimentadores deseaban comparar los estímulos A y D. Los resultados del experimento se dan aquí. Use la salida impresa *MINITAB* para completar el ejercicio.

Estímulo	Tiempo de reacción (s)					Total	Media
A	.8	.6	.6	.5		2.5	.625
B	.7	.8	.5	.5	.6	4.7	.671
C	1.2	1.0	.9	1.2	1.3	6.4	1.067
D	1.0	.9	.9	1.1	.7	4.6	.920
E	.6	.4	.4	.7	.3	2.4	.480

Salida *MINITAB* para el ejercicio 7.55

Anova en una vía: tiempo contra estímulo

Source	DF	SS	MS	F	P
Stimulus	4	1.2118	0.3030	11.67	0.000
Error	22	0.5711	0.0260		
Total	26	1.7830			

S = 0.1611 R-Sq = 67.97% R-Sq(adj) = 62.14%

Individual 95% CIs For Mean Based on Pooled StDev				
Level	N	Mean	StDev	
A	4	0.6250	0.1258	(-----*-----)
B	7	0.6714	0.1496	(-----*-----)
C	6	1.0667	0.1966	(-----*-----)
D	5	0.9200	0.1483	(-----*-----)
E	5	0.4800	0.1643	(-----*-----)
Pooled StDev = 0.1611				

- Realice un análisis de varianza y pruebe por si hay diferencia en los tiempos medios de reacción debido a los cinco estímulos.
- Compare los estímulos A y D para ver si hay una diferencia en tiempos medios de reacción.

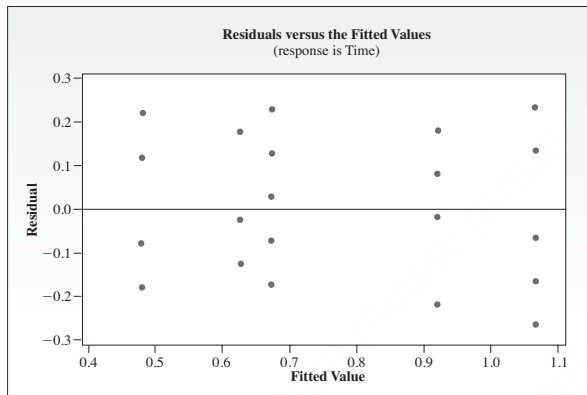
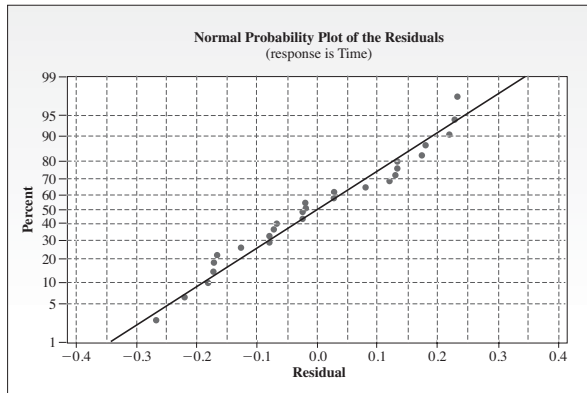
7.56 Consulte el ejercicio 7.55. Use esta salida *MINITAB* para identificar las diferencias en las medias de tratamiento.

Salida *MINITAB* para el ejercicio 7.56

Tukey 95% Simultaneous Confidence Intervals				
All Pairwise Comparisons among Levels of Stimulus				
Individual confidence level = 99.29%				
Stimulus = A subtracted from:				
Stimulus	Lower	Center	Upper	
B	-0.2535	0.0464	0.3463	(-----*-----)
C	0.1328	0.4417	0.7505	(-----*-----)
D	-0.0260	0.2950	0.6160	(-----*-----)
E	-0.4660	-0.1450	0.1760	(-----*-----)
Stimulus = B subtracted from:				
Stimulus	Lower	Center	Upper	
C	0.1290	0.3952	0.6615	(-----*-----)
D	-0.0316	0.2486	0.5288	(-----*-----)
E	-0.4716	-0.1914	0.0888	(-----*-----)
Stimulus = C subtracted from:				
Stimulus	Lower	Center	Upper	
D	-0.4364	-0.1467	0.1431	(-----*-----)
E	-0.8764	-0.5867	-0.2969	(-----*-----)
Stimulus = D subtracted from:				
Stimulus	Lower	Center	Upper	
E	-0.7426	-0.4400	-0.1374	(-----*-----)

7.57 Consulte el ejercicio 7.55. ¿Qué dicen la gráfica de probabilidad normal y la gráfica de residuales contra ajuste acerca de la validez de sus resultados del análisis de varianza?

Gráficas de diagnóstico MINITAB para el ejercicio 7.57



DATOS
EX0758

7.58 Tiempos de reacción II El experimento del ejercicio 7.55 podría haberse realizado en forma más eficiente usando un diseño de bloque aleatorizado con personas como bloques, porque se esperaría que el tiempo medio de reacción variaría de una persona a otra. En consecuencia, cuatro personas se emplearon en un nuevo experimento y cada persona fue sometida a cada uno de los cinco estímulos en orden aleatorio. Los tiempos de reacción (en segundos) se indican a continuación:

Persona	Estímulo				
	A	B	C	D	E
1	.7	.8	1.0	1.0	.5
2	.6	.6	1.1	1.0	.6
3	.9	1.0	1.2	1.1	.6
4	.6	.8	.9	1.0	.4

Salida de Excel para el ejercicio 7.58

ANOVA: dos factores sin réplica

RESUMEN	Conteo	Suma	Promedio	Varianza
A	4	2.8	0.7	0.02
B	4	3.2	0.8	0.0267
C	4	4.2	1.05	0.0167
D	4	4.1	1.025	0.0025
E	4	2.1	0.525	0.009

ANOVA

Fuente de variación	SS	gl	MS	F	Valor P	F crít.
Filas	0.14	3	0.046667	6.588	0.007	3.490
Columnas	0.787	4	0.196750	27.776	0.000	3.259
Error	0.085	12	0.007083			
Total	1.012	19				

- Use la salida impresa de Excel para analizar los datos y probar si hay diferencias en medias de tratamiento.
- Use el método de Tukey de comparaciones en pares para identificar las diferencias significativas por pares en los estímulos.
- ¿Le parece que el bloqueo fue eficaz en este experimento?

DATOS
EX0759

7.59 Frecuencia cardiaca y ejercicio Se realizó un experimento para examinar el efecto de la edad en la frecuencia cardiaca cuando una persona está sometida a una cantidad específica de ejercicio. Diez hombres fueron seleccionados al azar de entre cuatro grupos de edad: 10-19, 20-39, 40-59 y 60-69. Cada uno de ellos se ejercitó en una caminadora a un grado fijo durante 12 minutos, y se registró el aumento en su frecuencia cardiaca (en pulsaciones por minuto), la diferencia antes y después del ejercicio:

	10-19	20-39	40-59	60-69
29	24	37	28	
33	27	25	29	
26	33	22	34	
27	31	33	36	
39	21	28	21	
35	28	26	20	
33	24	30	25	
29	34	34	24	
36	21	27	33	
22	32	33	32	
Total	309	275	295	282

Use un programa apropiado de computadora para contestar estas preguntas:

- ¿Los datos dan suficiente evidencia para indicar una diferencia en el aumento medio en frecuencia cardiaca entre los cuatro grupos de edad? Pruebe usando $\alpha = .05$.
- Encuentre un intervalo de confianza de 90% para la diferencia en el aumento medio de frecuencia cardiaca entre los grupos de 10-19 y 60-69.

- c. Encuentre un intervalo de confianza de 90% para el aumento medio en frecuencia cardiaca para el grupo de edad de 20-39.
- d. ¿Aproximadamente cuántas personas necesitaría en cada grupo, si deseara estimar una media de grupo correcta a no más de dos pulsaciones por minuto con probabilidad igual a .95?

DATOS
EX0760

7.60 Aprendiendo a vender Una compañía deseaba estudiar los efectos de cuatro programas de capacitación en habilidad para vender,

en su personal de ventas. Treinta y dos personas se asignaron al azar a cuatro grupos de igual tamaño y cada grupo participó en uno de los diferentes programas de capacitación en ventas. Debido a que hubo algunas deserciones durante los programas de capacitación debido a enfermedad, vacaciones, etc., el número de estudiantes que completaron los programas varió de un grupo a otro. Al término de éstos, cada vendedor fue asignado al azar a una de un grupo de áreas de ventas cuyos potenciales se consideraron equivalentes. Las ventas efectuadas por cada uno de los cuatro grupos de vendedores durante la primera semana, después de completar el programa de capacitación, aparecen en la siguiente tabla:

Programa de capacitación				
	1	2	3	4
	78	99	74	81
	84	86	87	63
	86	90	80	71
	92	93	83	65
	69	94	78	86
	73	85		79
		97		73
		91		70
Total	482	735	402	588

Analice el experimento usando el método apropiado. Identifique los tratamientos o factores de interés para el experimentador e investigue cualesquier efectos significativos. ¿Cuáles son las implicaciones prácticas de este experimento? Escriba un párrafo que explique los resultados de su análisis.

7.61 Factorial de 4 × 2 Suponga que usted debe realizar un experimento factorial de dos factores, el factor A en cuatro niveles y el factor B en dos niveles, con *r* réplicas por tratamiento.

- a. ¿Cuántos tratamientos intervienen en el experimento?
- b. ¿Cuántas observaciones están involucradas?
- c. Haga una lista de las fuentes de variación y sus grados de libertad respectivos.

7.62 Factorial de 2 × 3 El análisis de varianza para un experimento factorial de 2 × 3, el factor A en dos niveles y el factor B en tres niveles, con cinco observaciones por tratamiento, se muestra en la tabla.

Fuente	gl	SS	MS	F
A		1.14		
B		2.58		
AB		.49		
Error				
Total		8.41		

- a. ¿Los datos dan suficiente evidencia para indicar una interacción entre los factores A y B? Pruebe usando $\alpha = .05$. ¿Cuáles son las implicaciones prácticas de su respuesta?
- b. Dé el valor *p* aproximado para la prueba de la parte a.
- c. ¿Los datos dan suficiente evidencia para indicar que el factor A afecta la respuesta? Pruebe usando $\alpha = .05$.
- d. ¿Los datos dan suficiente evidencia para indicar que el factor B afecta la respuesta? Pruebe usando $\alpha = .05$.

7.63 Consulte el ejercicio 7.62. Las medias de todas las observaciones, a los niveles A₁ y A₂ del factor A son $\bar{x}_1 = 3.7$ y $\bar{x}_2 = 1.4$, respectivamente. Encuentre un intervalo de confianza para la diferencia en respuesta media para los niveles de factor A₁ y A₂.

DATOS
EX0764

7.64 La mosca blanca en California La mosca blanca, que produce la caída de hojas de arbustos y árboles y una reducción en la producción de cosechas negociables, ha emergido como plaga en el sur de California. En un estudio para determinar los factores que afectan el ciclo vital de la mosca blanca, se realizó un experimento en el que las moscas blancas se colocaron en dos tipos diferentes de plantas a tres temperaturas distintas. La observación de interés fue el número total de huevecillos depositados por hembras enjauladas en una de las seis posibles combinaciones de tratamiento. Cada combinación de tratamiento se corrió usando cinco jaulas.

Planta	Temperatura		
	70 °F	77 °F	82 °F
Algodón	37	34	46
	21	54	32
	36	40	41
	43	42	36
	31	16	38
Pepino	50	59	43
	53	53	62
	25	31	71
	37	69	49
	48	51	59

Salida de *Excel* para el ejercicio 7.64

ANOVA						
Fuente de variación	SS	gl	MS	F	Valor P	F crit.
Planta	1512.3	1	1512.3	12.293	0.002	4.260
Temperatura	487.4667	2	243.733	1.981	0.160	3.403
Interacción	111.2	2	55.6	0.452	0.642	3.403
Dentro	2952.4	24	123.017			
Total	5063.367	29				

- ¿Qué tipo de diseño experimental se ha empleado?
- ¿Los datos dan suficiente evidencia para indicar que el efecto de la temperatura en el número de huevos depositados es diferente, dependiendo del tipo de planta? Use la salida impresa para probar la hipótesis apropiada.
- Grafique las medias de tratamiento para algodón como función de la temperatura. Grafique las medias de tratamiento para pepino como función de la temperatura. Comente sobre la similitud o diferencia en estas dos gráficas.
- Encuentre el número medio de huevos depositados en algodón y pepino con base en 15 observaciones cada uno. Calcule un intervalo de confianza de 95% para la diferencia en las medias poblacionales subyacentes.

DATOS
EX0765

7.65 Contaminación proveniente de plantas químicas Cuatro plantas químicas, que producen el mismo producto y son propiedad de la misma compañía, descargan aguas negras en arroyos de la cercanía de sus lugares. Para comprobar el grado de contaminación creada por las aguas negras y para determinar si esto varía de una planta a otra, la compañía recolectó muestras aleatorias de desechos líquidos, cinco especímenes por cada una de las cuatro plantas. Los datos se muestran en la tabla:

Planta	Desechos contaminantes (lb/gal de desechos)				
A	1.65	1.72	1.50	1.37	1.60
B	1.70	1.85	1.46	2.05	1.80
C	1.40	1.75	1.38	1.65	1.55
D	2.10	1.95	1.65	1.88	2.00

- ¿Los datos dan suficiente evidencia para indicar una diferencia en las cantidades medias de aguas negras descargadas por las cuatro plantas?
- Si la descarga media máxima de aguas negras es 1.5 lb/gal, ¿los datos dan suficiente evidencia para indicar que el límite está excedido en la planta A?
- Estime la diferencia en la descarga media de aguas negras entre las plantas A y D, usando un intervalo de confianza de 95%.

DATOS
EX0766

7.66 Artículos básicos en Estados Unidos Un anuncio para una cadena de supermercados popular afirmaba que tenía precios consistentemente más bajos que uno de sus competidores. Como parte de una

encuesta efectuada por una compañía independiente de verificación de precios, se da el total semanal promedio, basado en los precios de aproximadamente 95 artículos, para esta cadena y para su competidor, registrados durante cuatro semanas consecutivas en un mes particular.

Semana	Anunciante (\$)	Competidor (\$)
1	254.26	256.03
2	240.62	255.65
3	231.90	255.12
4	234.13	261.18

- ¿Qué tipo de diseño se ha usado en este experimento?
- Realice un análisis de varianza de los datos.
- ¿Hay suficiente evidencia para indicar que existe diferencia en el promedio total semanal para los dos supermercados? Use $\alpha = .05$.

DATOS
EX0767

7.67 Producción de trigo Las producciones de trigo (en bushels por acre) se compararon para cinco variedades diferentes, A, B, C, D y E, en seis lugares diferentes. Cada variedad fue asignada al azar a un lote en cada lugar. Los resultados del experimento se muestran en la tabla siguiente, junto con una salida impresa *MINITAB* del análisis de varianza. Analice el experimento usando el método apropiado. Identifique los tratamientos o factores de interés para el experimentador e investigue cualquier efecto que exista. Use las gráficas de diagnóstico para comentar sobre la validez de las suposiciones del análisis de varianza. ¿Cuáles son las implicaciones prácticas de este experimento? Redacte un párrafo que explique los resultados de su análisis.

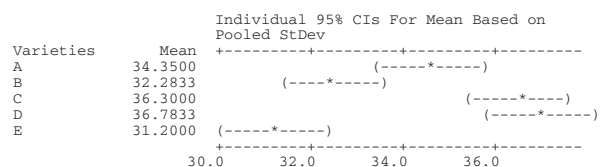
Variedad	Lugar					
	1	2	3	4	5	6
A	35.3	31.0	32.7	36.8	37.2	33.1
B	30.7	32.2	31.4	31.7	35.0	32.7
C	38.2	33.4	33.6	37.1	37.3	38.2
D	34.9	36.1	35.2	38.3	40.2	36.0
E	32.4	28.9	29.2	30.7	33.9	32.1

Salida *MINITAB* para el ejercicio 7.67

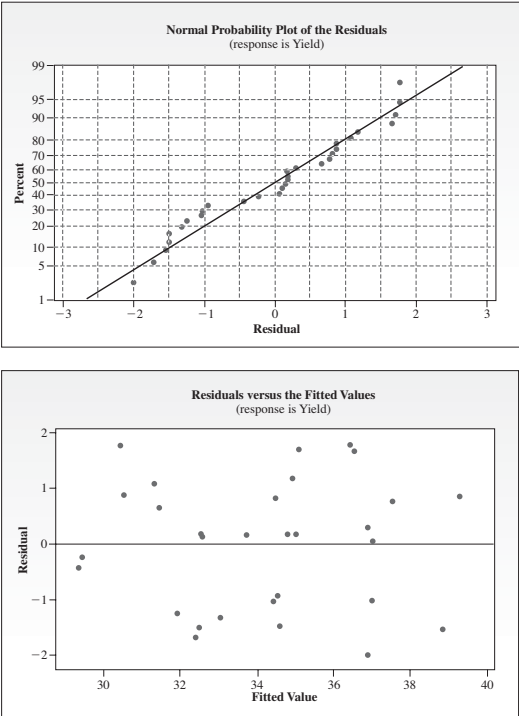
ANOVA de dos vías: producción contra variedades, lugar

Source	DF	SS	MS	F	P
Varieties	4	142.670	35.6675	18.61	0.000
Locations	5	68.142	13.6283	7.11	0.001
Error	20	38.303	1.91565		
Total	29	249.142			

S = 1.384 R-Sq = 84.62% R-Sq(adj) = 77.69%



Gráficas de diagnóstico MINITAB para el ejercicio 7.67



DATOS
EX0768

7.68 Acondicionamiento físico Los investigadores Russell R. Pate y colegas analizaron los resultados de la National Health and Nutrition Examination Survey para evaluar niveles de acondicionamiento cardiorrespiratorio en jóvenes de 12 a 19 años de edad.⁶ La admisión máxima estimada de oxígeno ($VO_{2m\acute{a}x}$) se utilizó para medir el nivel cardiorrespiratorio de una persona. El foco del estudio investiga la relación entre los niveles de actividad física (más que otros, igual a otros o menos que otros) y el género en $VO_{2m\acute{a}x}$. Los datos que siguen se basan en este estudio.

	Actividad física		
	Más	Igual	Menos
Hombres	50.1	45.7	40.9
	47.2	44.2	41.3
	49.7	46.8	39.2
	50.4	44.9	40.9
Mujeres	41.2	37.2	36.5
	39.8	39.4	35.0
	41.5	38.6	37.2
	38.2	37.8	35.4

a. ¿Es éste un experimento factorial o un diseño de bloque aleatorizado? Explique.

- b. ¿Hay interacción significativa entre niveles de actividad física y género? ¿Hay diferencias significativas entre hombres y mujeres? ¿Y entre niveles de actividad física?
- c. Si la interacción es significativa, use el procedimiento de Tukey por pares para investigar diferencias entre las seis medias de celda. Comente sobre los resultados hallados usando este procedimiento. Use $\alpha = .05$.

DATOS
EX0769

7.69 Salarios de profesores En un estudio sobre los salarios iniciales de los profesores auxiliares⁷ cinco profesores varones y cinco mujeres que iniciaban en cada uno de dos tipos de instituciones que otorgan títulos de doctorado participaron en una encuesta y se registraron sus salarios iniciales. Los resultados de la encuesta en unidades de miles de dólares se dan en la tabla siguiente.

Género	Tipo	
	Público (\$)	Sin fines de lucro (\$)
Hombre	64.6	75.2
	63.8	75.7
	63.3	75.0
	64.1	75.1
	64.9	75.8
Mujer	59.7	71.6
	56.9	71.6
	58.4	71.2
	60.1	73.0
	59.8	69.5

- a. ¿Qué tipo de diseño se utilizó para recolectar estos datos?
- b. Use un análisis de varianza para probar si hay diferencias significativas en género y tipo de institución, y para probar si hay interacción significativa de género $\times$ tipo de institución.
- c. Encuentre una estimación de intervalo de confianza de 95% para la diferencia en salarios iniciales para profesores auxiliares hombres y profesoras auxiliares mujeres. Interprete este intervalo en términos de una diferencia de género en salarios iniciales.
- d. Elabore un resumen de los resultados de su análisis.

DATOS
EX0770

7.70 Cerámica en el Reino Unido Un artículo en *Archaeometry* contenía un análisis de 26 muestras de cerámica romano-británica, halladas en los hornos de cuatro lugares diferentes en el Reino Unido.⁸ Como un sitio dio sólo dos muestras, considere las muestras halladas en los otros tres sitios. Las muestras se analizaron para determinar su composición química y a continuación se indica el porcentaje de óxido de hierro.

Llanederyn	Island Thorns	Ashley Rails
7.00	5.78	1.28
7.08	5.49	2.39
7.09	6.92	1.50
6.37	6.13	1.88
7.06	6.64	1.51
6.26	6.69	
4.26	6.44	

- ¿Qué tipo de diseño experimental es éste?
- Use un análisis de varianza para determinar si hay una diferencia en el promedio de porcentaje de óxido de hierro en los tres sitios. Use $\alpha = .01$.
- Si usted tiene acceso a un programa de cómputo, genere las gráficas de diagnóstico para este experimento. ¿Le parece que algunas de las suposiciones del análisis de varianza han sido violadas? Explique.

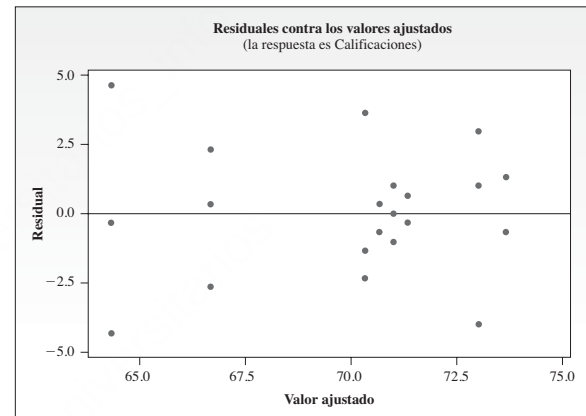
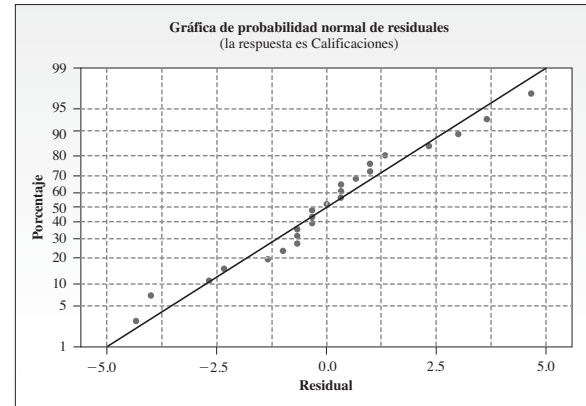
DATOS
EX0771

7.71 Teléfonos inteligentes Un teléfono inteligente es un dispositivo móvil que cuenta con una capacidad de cómputo y una conectividad que son más avanzadas que las de un “teléfono común” básico contemporáneo. Los datos que siguen son las calificaciones para seis teléfonos inteligentes de cuatro proveedores; tres cuestan 150 dólares o más y tres cuestan menos de esa cantidad.⁹ Las calificaciones tienen un valor máximo de 100 y un mínimo de 0.

	Proveedor			
	AT&T	Sprint	T-Mobile	Verizon
Costo $\geq$ \$150	76	74	72	75
	74	69	71	73
	69	68	71	73
Costo $<$ \$150	69	69	71	72
	67	64	71	71
	64	60	70	70

- ¿Qué tipo de experimento se utilizó para evaluar estos teléfonos inteligentes. ¿Cuáles son los factores? ¿Cuántos niveles de cada factor se usaron en el experimento?
- Elabore una tabla de análisis de varianza apropiada para este diseño, especificando las fuentes de variación, grados de libertad, sumas de cuadrados, medias cuadradas y los valores apropiados de F usados en la prueba.
- ¿Hay una interacción significativa entre los dos factores?
- ¿Hay un efecto principal debido a los proveedores? ¿Cuál es su valor p ?
- ¿Hay un efecto principal debido al costo? ¿Cuál es su valor p ?
- Resuma los resultados de las partes c–e.

7.72 Teléfonos inteligentes, continúa Consulte el ejercicio 7.71. Las gráficas de diagnóstico para este experimento se ilustran a continuación. ¿Le parece que alguna de las suposiciones de los análisis de varianza ha sido violada? Explique.



DATOS
EX0773

7.73 Salarios de profesores II El Departamento de Educación de Estados Unidos⁷ informa sobre los salarios de los profesores en las universidades y los colegios de ese país. Los siguientes datos (en miles de dólares) se adaptaron del informe para docentes de tiempo completo con contratos de nueve meses en instituciones sin fines de lucro que ofrecen programas de doctorado. Se tomaron 10 muestras de cada uno de estos tres niveles de catedráticos, tanto para hombres como para mujeres.

Género	Nivel					
	Profesor auxiliar		Profesor adjunto		Profesor de tiempo completo	
Hombre	75.4	76.5	94.8	92.6	133.1	132.0
	75.6	75.1	96.4	95.0	147.7	122.1
	76.6	74.8	89.7	72.6	143.3	132.7
	76.2	74.9	103.1	100.2	158.1	142.0
	76.1	76.3	91.5	88.3	126.3	127.6
Mujer	74.8	71.1	86.3	74.7	161.1	122.9
	70.6	70.3	91.0	77.6	141.4	130.8
	72.4	71.1	88.3	71.8	130.8	145.4
	70.5	71.3	83.0	80.3	145.0	127.9
	71.1	69.3	72.5	80.9	127.5	122.9

- a. Identifique el diseño empleado en este estudio.

b. Use el análisis de varianza apropiado para estos datos.

c. ¿Los datos indican que el salario en los diferentes rangos varía por género?

d. Si no hay interacción, determine si hay diferencias en salarios por rango y si hay diferencias por género. Discuta sus resultados.
- e. Grafique el promedio de salarios usando una gráfica de interacción. Si el principal efecto de rangos es significativo, use el método de comparaciones en pares de Tukey para determinar si hay diferencias significativas entre los rangos. Use $\alpha = .01$.

CASO PRÁCTICO

DATOS Abarrotes

¡Cómo ahorrar dinero en comestibles!

Los productos comestibles enlatados o congelados que se compran a granel casi siempre nos ayudan a ahorrar dinero en comparación con los que se compran en los supermercados. Puede ahorrar más de 75% en productos enlatados, y más de 80% en productos congelados, comprados a granel. Los siguientes precios se encuentran en “Save Money on Groceries”, un artículo de Roberta R. Bailey y Craig Idlebrook en el sitio web www.MotherEarthNews.com.¹⁰

Producto	Costo a granel/lb (\$)	Enlatados		Congelados	
		Casa 16 oz (\$)	Tienda 16 oz (\$)	Casa 16 oz (\$)	Tienda 16 oz (\$)
Ejotes	1.00	1.00	1.31	1.00	1.99*
Maíz dulce	2.00 (docena)	0.83	1.31	0.83	1.99*
Chícharos en vaina	2.00	6.00	1.31	6.00	1.99*
Jitomates enteros	1.00	1.50	1.31*	1.50	N/A
Betabeles	1.00	1.00	1.31	1.00	N/A
Brócoli	1.50	N/A	N/A	1.50	2.29
Espinacas	4.00	N/A	N/A	4.00	4.62*
Peras	1.00	1.00	2.59	1.00	3.59
Arándanos	.50	.50	2.19	.50	3.59*
Duraznos	1.00	1.00	1.99*	1.00	3.59

*Aquí se informa el precio más bajo de un rango.

Aquí hay algunos productos que no están enlatados y otros que no están congelados. Puede eliminar estas entradas para el análisis que sigue.

1. ¿Esta disposición corresponde a cualquiera de los diseños estudiados en este capítulo? De ser así, identifique las filas y/o columnas según se relacionen con ese diseño.
2. En vista que la tabla está incompleta, considere eliminar las filas correspondientes a jitomates enteros, betabeles, brócoli y espinacas antes del análisis. ¿Todavía es válido el diseño dado en el punto 1?
3. Use un análisis apropiado del procedimiento de varianza para analizar estos datos. Si encuentra significancia, use el procedimiento de Tukey para identificar diferencias reales en los precios.
4. Resuma sus resultados en forma de reporte. ¿En realidad puede ahorrar dinero comprando productos a granel? Explique.

Regresión lineal y correlación

OBJETIVOS GENERALES

En este capítulo consideramos la situación en la que el valor medio de una variable aleatoria y está relacionado con otra variable x . Al medir tanto y como x para cada unidad experimental, con lo cual se generan datos bivariados, se puede usar la información dada por x para estimar el valor promedio de y y para predecir valores de y para valores de x asignados previamente.

ÍNDICE DEL CAPÍTULO

- Introducción (8.1)
- Modelo probabilístico lineal simple (8.2)
- El método de mínimos cuadrados (8.3)
- Un análisis de varianza para regresión lineal (8.4)
- Prueba de la utilidad del modelo de regresión lineal (8.5)
- Herramientas de diagnóstico para verificar suposiciones de la regresión (8.6)
- Estimación y predicción usando la recta ajustada (8.7)
- Análisis de correlación (8.8)
- Covarianza (8.9)



NECESITO SABER...

Cómo estar seguro de que mis cálculos son correctos

¿Su automóvil está "Hecho en EUA"?

La frase "Hecho en EUA" se ha convertido en grito de batalla en los últimos años porque los trabajadores estadounidenses tratan de proteger sus empleos contra la competencia extranjera. En el caso práctico del final de este capítulo exploramos las cambiantes actitudes de los consumidores estadounidenses hacia los vehículos fabricados fuera de Estados Unidos usando un sencillo análisis de regresión lineal.

INTRODUCCIÓN

Los estudiantes de último año de preparatoria, los de primer año de universidad, sus padres, así como la administración de una universidad, están preocupados por el rendimiento académico de un estudiante después de haberse inscrito en la universidad. ¿Se puede estimar o predecir el promedio de calificaciones (GPA; siglas de *grade point average*) de un estudiante al terminar su último año de preparatoria, antes de inscribirse en la universidad? A primera vista éste podría ser un problema difícil aunque se espera que los estudiantes altamente motivados, que se hayan graduado con calificaciones altas de una preparatoria, alcancen un GPA alto cuando terminen el primer año. Por otra parte, no es probable que los que carezcan de motivación o que hayan obtenido sólo un éxito moderado en preparatoria salgan bien. Se esperaría que el rendimiento académico de un estudiante sea una función de diversas variables:

- Rango en su grupo de preparatoria
- Nivel general de preparatoria
- Alto promedio GPA
- Calificaciones del SAT

Este problema es de naturaleza más bien general. Usted se interesará por una variable aleatoria y (promedio GPA) relacionada con diversas variables independientes. El objetivo es crear una *ecuación de predicción* que exprese y como función de estas variables independientes. A continuación, si se pueden medir las variables independientes, es posible sustituir estos valores en la ecuación de predicción y obtener la predicción para y , es decir, el promedio GPA del estudiante en nuestro ejemplo. Pero, ¿cuáles variables deben usarse para hacer la predicción? ¿Qué tan fuerte es su relación con y ? ¿Cómo se construye una buena ecuación de predicción para y como función de las variables seleccionadas para la predicción? Contestaremos estas preguntas en los siguientes dos capítulos.

En este capítulo restringimos nuestra atención al sencillo problema de predecir y como función lineal de una sola variable x de pronóstico. Recuerde que usamos la ecuación de una recta para describir la relación entre x y y , y describimos la fuerza de la relación usando el coeficiente de correlación r . Nos apoyaremos en algunos de estos resultados cuando repasemos el tema de regresión lineal y correlación.

MODELO PROBABILÍSTICO LINEAL SIMPLE

Considere el problema de tratar de predecir el valor de una respuesta y basada en el valor de una variable independiente x . La recta de mejor ajuste,

$$y = a + bx$$

está basada en una *muestra* de n observaciones bivariadas tomadas de una *población* más grande de medidas. La recta que describe la relación entre y y x en la *población* es semejante a la recta de mejor ajuste de la *muestra*, pero no es igual. ¿Cómo se puede construir un **modelo de población** para describir la relación entre una variable aleatoria y y una variable independiente x relacionada?

Se empieza por suponer que la variable de interés, y , está *linealmente* relacionada a una variable independiente x . Para describir la relación lineal, se puede usar el **modelo determinista**

$$y = \alpha + \beta x$$

donde α es la intersección y , es decir, el valor de y cuando $x = 0$, y β es la pendiente de la recta, definida como el cambio en y para un cambio unitario en x , como se muestra en la figura 8.1. Este modelo describe una relación determinista entre la variable de interés y , a veces llamada **variable de respuesta**, y la variable independiente x , denominada **variable de pronóstico**. Esto es, la ecuación lineal determina un valor exacto de y cuando se da el valor de x . ¿Este modelo es realista para una situación experimental? Considere el siguiente ejemplo.

FIGURA 8.1

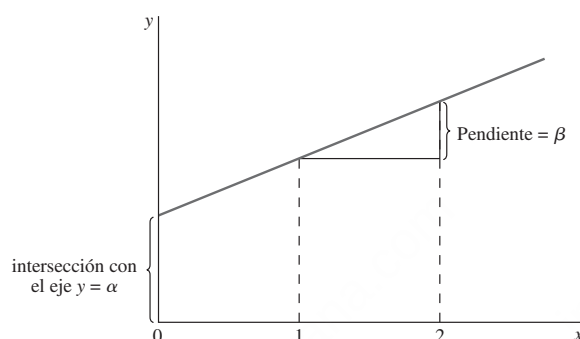
Intersección con el eje y y
pendiente de una recta

TIP

MI CONSEJO

Pendiente = cambio en
 y para un cambio unitario
en x

Intersección con eje y =
valor de y cuando $x = 0$



APPLET EN LÍNEA

Construir una gráfica de
dispersión

La tabla 8.1 muestra las calificaciones del examen de aprovechamiento en matemáticas de una muestra aleatoria de $n = 10$ estudiantes de primer año de universidad, junto con sus calificaciones finales en cálculo. Una gráfica bivariada de estos puntos y calificaciones se da en la figura 8.2. Observe que los puntos *no están exactamente sobre una recta* sino que más bien parecen ser desviaciones alrededor de una recta subyacente. Una forma sencilla de modificar el modelo determinista es agregar un **componente aleatorio de error** para explicar las desviaciones de los puntos alrededor de la recta. Una respuesta particular y se describe usando el **modelo probabilístico**

$$y = \alpha + \beta x + \epsilon$$

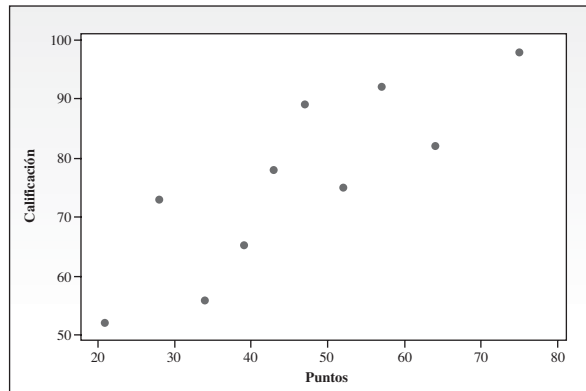
**Calificaciones de examen de aprovechamiento en matemáticas
y puntuaciones finales en cálculo para estudiantes de primer
año de universidad**

TABLA 8.1

Estudiante	Puntuación en el examen de aprovechamiento en matemáticas	Calificación final en cálculo
1	39	65
2	43	78
3	21	52
4	64	82
5	57	92
6	47	89
7	28	73
8	75	98
9	34	56
10	52	75

FIGURA 8.2

Gráfica de dispersión de datos de la tabla 8.1



La primera parte de la ecuación, $\alpha + \beta x$, llamada **recta de medias**, describe el valor promedio de y para un valor determinado de x . El componente de error ϵ permite que cada respuesta individual y se desvíe de la recta de medias en una pequeña cantidad.

Para usar este *modelo probabilístico* para hacer inferencias es necesario ser más específico acerca de esta “pequeña cantidad”, ϵ .

SUPOSICIONES ACERCA DEL ERROR ALEATORIO ϵ

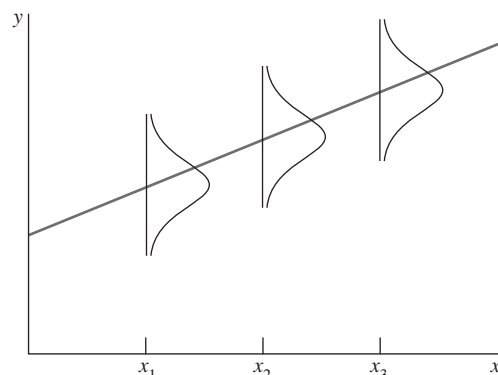
Suponga que los valores de ϵ satisfacen estas condiciones:

- Son independientes en el sentido probabilístico
- Tienen una media de 0 y una varianza común igual a σ^2
- Tienen una distribución normal de probabilidad

Estas suposiciones acerca del error aleatorio ϵ se muestran en la figura 8.3 para tres valores fijos de x , por ejemplo x_1 , x_2 y x_3 . Observe la similitud entre estas suposiciones y las suposiciones necesarias para las pruebas en los capítulos 6 y 7. Repasaremos estas suposiciones más adelante en este capítulo y daremos algunas herramientas de diagnóstico para que usted las use en la verificación de su validez.

FIGURA 8.3

Modelo probabilístico lineal



Recuerde que este modelo se ha creado para una población de mediciones que por lo general es desconocida, pero puede usar información muestral para estimar los valores de α y β , que son los coeficientes de la recta de medias, $E(y) = \alpha + \beta x$. Estas estimaciones se usan para formar la recta de mejor ajuste para un conjunto de datos determinado, llamado **recta de mínimos cuadrados** o **recta de regresión**. En la siguiente sección repasamos la forma de calcular el punto de intersección y la pendiente de esta recta.

8.3

EL MÉTODO DE MÍNIMOS CUADRADOS

TIP MI CONSEJO

Pendiente = coeficiente de x

Intersección con eje y = término constante

$$\hat{y} = a + bx$$

donde a y b son las estimaciones de los parámetros α y β de intersección y pendiente, respectivamente. La recta ajustada para los datos de la tabla 8.1 se muestra en la figura 8.4. Las rectas verticales trazadas desde la recta de predicción $(x_i, \hat{y}_i)$ hasta cada punto (x_i, y_i) representan las desviaciones de los puntos desde la recta; es decir, $(y_i - \hat{y}_i)$.

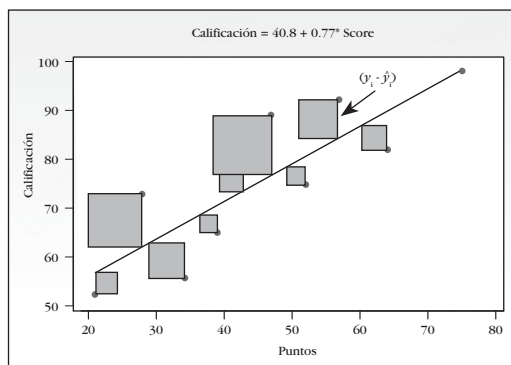
FIGURA 8.4

Método de mínimos cuadrados



APPLET EN LÍNEA

Método de mínimos cuadrados



Observe que algunos puntos están debajo de la recta de predicción, y por tanto $(y_i - \hat{y}_i)$ será negativo. Para evitar que las distancias positivas y negativas se “cancelen entre sí”, elegimos minimizar las distancias desde los puntos hasta la recta ajustada, usando el **principio de mínimos cuadrados**.

PRINCIPIO DE MÍNIMOS CUADRADOS

La recta que reduce al mínimo la suma de cuadrados de las desviaciones de los valores observados de y desde los pronosticados es la **recta de mejor ajuste**. La suma del cuadrado de las desviaciones por lo general se denomina **suma de cuadrados de error** (SSE) y se define como

$$SSE = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - a - bx_i)^2$$

Observe la recta de regresión y los puntos de datos de la figura 8.4. SSE es la suma del cuadrado de las distancias representada por el área de los cuadros azules.

Hallar los valores de a y b , las estimaciones de α y β , usa cálculo diferencial, que está fuera del propósito de este libro. En lugar de derivar sus valores, simplemente presentaremos fórmulas para calcular los valores de a y b , llamados **estimadores de mínimos cuadrados** de α y β . Usaremos una notación que está basada en **las sumas de cuadrados** para las variables del problema de regresión, que es semejante en forma a las sumas de cuadrados empleadas en el capítulo 7.

Usted debe usar el método de entrada de datos para su calculadora científica para introducir los datos muestrales.

- Si su calculadora tiene sólo una función estadística de una variable, todavía puede ahorrar tiempo al hallar las sumas necesarias y sumas de cuadrados.
- Si su calculadora tiene una función estadística de dos variables o si tiene una calculadora graficadora, la calculadora en forma automática guarda todas las sumas y sumas de cuadrados, así como los valores de a , b y el coeficiente de correlación r .
- Asegúrese de consultar el manual de su calculadora para hallar la forma más fácil de obtener los estimadores de mínimos cuadrados.

ESTIMADORES DE MÍNIMOS CUADRADOS DE α Y β

$$b = \frac{S_{xy}}{S_{xx}} \quad \text{y} \quad a = \bar{y} - b\bar{x}$$

donde las cantidades S_{xy} y S_{xx} están definidas como

$$S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}$$

y

$$S_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

EJEMPLO

8.1

Encuentre la recta de predicción de mínimos cuadrados para los datos de la calificación en cálculo de la tabla 8.1.

Solución Use los datos de la tabla 8.2 y el método de introducción de datos en su calculadora científica para hallar las siguientes sumas de cuadrados:

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 23\,634 - \frac{(460)^2}{10} = 2\,474$$

$$S_{xy} = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} = 36\,854 - \frac{(460)(760)}{10} = 1\,894$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{760}{10} = 76 \quad \bar{x} = \frac{\sum x_i}{n} = \frac{460}{10} = 46$$

TABLA 8.2

Cálculos para los datos de la tabla 8.1

y_i	x_i	x_i^2	$x_i y_i$	y_i^2	
65	39	1521	2535	4225	
78	43	1849	3354	6084	
52	21	441	1092	2704	
82	64	4096	5248	6724	
92	57	3249	5244	8464	
89	47	2209	4183	7921	
73	28	784	2044	5329	
98	75	5625	7350	9604	
56	34	1156	1904	3136	
75	52	2704	3900	5625	
Suma	760	460	23634	36854	59816

Entonces

$$b = \frac{S_{xy}}{S_{xx}} = \frac{1894}{2474} = .76556 \quad \text{y} \quad a = \bar{y} - b\bar{x} = 76 - (.76556)(46) = 40.78424$$

TIP MI CONSEJO

Se puede predecir y para un valor determinado de x al sustituir x en la ecuación para hallar $\hat{y}$

La recta de regresión de mínimos cuadrados es entonces

$$\hat{y} = a + bx = 40.78424 + .76556x$$

La gráfica de esta recta se ve en la figura 8.4. Ahora se puede usar para predecir y para un valor determinado de x , ya sea consultando la figura 8.4 o sustituyendo el valor apropiado de x en la ecuación. Por ejemplo, si un alumno de primer año obtuvo $x = 50$ en el examen, la calificación pronosticada de cálculo del estudiante es (usando precisión completa de decimales)

$$\hat{y} = a + b(50) = 40.78424 + (.76556)(50) = 79.06$$



NECESITO SABER...

Cómo estar seguro de que mis cálculos son correctos

- Tenga cuidado con los errores de redondeo. Lleve al menos seis cifras significativas y haga redondeo sólo al informar el resultado final.
- Use una calculadora científica o graficadora para hacer todo el trabajo. Casi todas las calculadoras estimarán los valores de a y b si se le introducen correctamente los datos.
- Use un programa de computadora si tiene acceso a ella.
- Siempre grafique los datos y la recta. Si la recta no se ajusta a los puntos, ¡es probable que el usuario tenga un error!

UN ANÁLISIS DE VARIANZA PARA REGRESIÓN LINEAL

8.4

En el capítulo 7 utilizamos los procedimientos del análisis de varianza para dividir la variación total del experimento en partes atribuidas a diversos factores de interés para el experimentador. En un análisis de regresión, la respuesta y está relacionada con la variable independiente x . En consecuencia, la variación total de la variable de respuesta y , dada por

$$TSS = S_{yy} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$$

está dividida en dos partes:

- La SSR (suma de cuadrados para regresión) mide la cantidad de variación explicada al usar la recta de regresión con una variable independiente x
- La SSE (suma de cuadrados de error) mide la variación “residual” en los datos que no está explicada por la variable independiente x

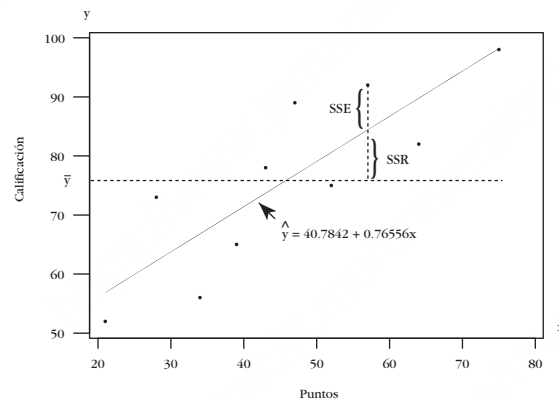
de modo que

$$TSS = SSR + SSE$$

Para un valor particular de la respuesta y_i , se puede visualizar este desglose en la variación usando las distancias verticales ilustradas en la figura 8.5. Se puede ver que la SSR es la suma del cuadrado de desviaciones de las diferencias entre la respuesta estimada sin usar x ($\bar{y}$) y la respuesta estimada usando x (la recta de regresión, $\hat{y}$); la SSE es la suma del cuadrado de diferencias entre la recta de regresión ($\hat{y}$) y el punto y .

FIGURA 8.5

Desviaciones desde la recta ajustada



No es demasiado difícil demostrar algebraicamente que

$$\begin{aligned} SSR &= \sum (\hat{y}_i - \bar{y})^2 = \sum (a + bx_i - \bar{y})^2 = \sum (\bar{y} - b\bar{x} + bx_i - \bar{y})^2 = b^2 \sum (x_i - \bar{x})^2 \\ &= \left(\frac{S_{xy}}{S_{xx}} \right)^2 S_{xx} = \frac{(S_{xy})^2}{S_{xx}} \end{aligned}$$

Como $TSS = SSR + SSE$, se puede completar la partición al calcular

$$SSE = TSS - SSR = S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$$

Recuerde del capítulo 7 que cada una de las diversas fuentes de variación, cuando se dividen entre los **grados de libertad** apropiados, da una estimación de la variación en el experimento. Estas estimaciones se denominan **medias cuadráticas**, $MS = SS/gl$, y se ven en una tabla ANOVA.

Al examinar los grados de libertad asociados con cada una de estas sumas de cuadrados, observe que el total de grados de libertad para n mediciones es $(n - 1)$. Como la estimación de la recta de regresión, $\hat{y} = a + bx_i = \bar{y} - b\bar{x} + bx_i$, abarca la estimación de un parámetro adicional β , hay un grado de libertad asociado con la SSR, dejando $(n - 2)$ grados de libertad con la SSE.

Al igual que con todas las tablas ANOVA que hemos estudiado, el error medio cuadrático

$$\text{MSE} = s^2 = \frac{\text{SSE}}{n - 2}$$

es un estimador insesgado de la varianza subyacente σ^2 . La tabla de análisis de varianza se muestra en la tabla 8.3.

TABLA 8.3

Análisis de varianza para regresión lineal

Fuente	gl	SS	MS
Regresión	1	$\frac{(S_{xy})^2}{S_{xx}}$	MSR
Error	$n - 2$	$S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$	MSE
Total	$n - 1$	S_{yy}	

Para los datos de la tabla 8.1, se puede calcular

$$\text{TSS} = S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = 59\,816 - \frac{(760)^2}{10} = 2\,056$$

$$\text{SSR} = \frac{(S_{xy})^2}{S_{xx}} = \frac{(1\,894)^2}{2\,474} = 1\,449.9741$$

de modo que

$$\text{SSE} = \text{TSS} - \text{SSR} = 2\,056 - 1\,449.9741 = 606.0259$$

y

$$\text{MSE} = \frac{\text{SSE}}{n - 2} = \frac{606.0259}{8} = 75.7532$$

La tabla de análisis de varianza, parte de la *salida de regresión lineal* generada por el MINITAB, es la sección inferior sombreada de la salida de computadora de la figura 8.6a). Las primeras dos rectas dan la ecuación de la recta de mínimos cuadrados, $\hat{y} = 40.8 + .766x$. Las estimaciones de mínimos cuadrados a y b están dadas con mayor precisión en la columna marcada

FIGURA 8.6(a)

Salida MINITAB para los datos de la tabla 8.1

Análisis de regresión: y contra x

The regression equation is
 $y = 40.8 + 0.766 x$

Predictor	Coef	SE Coef	T	P
Constant	40.784	8.507	4.79	0.001
x	0.7656	0.1750	4.38	0.002

S = 8.70363 R-Sq = 70.5% R-Sq(adj) = 66.8%

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	1450.0	1450.0	19.14	0.002
Residual Error	8	606.0	75.8		
Total	9	2056.0			

FIGURA 8.6(b)

Salida de *MS Excel* para los datos de la tabla 8.1

SUMMARY OUTPUT							
Regression Statistics							
Multiple R	0.8398						
R Square	0.7052						
Adjusted R Square	0.6684						
Standard Error	8.7036						
Observations	10						
ANOVA							
	df	SS	MS	F	Significance F		
Regression	1	1449.974	1449.974	19.141	0.002		
Residual	8	606.026	75.753				
Total	9	2056					
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%
Intercept	40.784	8.507	4.794	0.001	21.167	60.401	21.167
Score	0.766	0.175	4.375	0.002	0.362	1.169	0.362

“Coef”. La salida de *MS Excel* para los mismos datos se muestra en la figura 8.6b). La tabla de ANOVA está en medio de la salida sombreada; las estimaciones de mínimos cuadrados se encuentran en la parte inferior de la salida sombreada en la columna etiquetada “Coefficients”. Se pueden hallar instrucciones para generar esta salida impresa en la sección “Tecnología actual” al final de este capítulo.

Las salidas de computadora también dan alguna información acerca de la variación en el experimento. Cada una de las estimaciones de mínimos cuadrados, a y b , tiene un error estándar asociado, marcado “SE Coef” en la figura 8.6a) y “Standard Error” en la figura 8.6b). Hacia la mitad de la salida *MINITAB* se encuentra la mejor estimación insesgada de σ — $s = \sqrt{MSE} = \sqrt{75.7532} = 8.70363$ — que mide el **error residual**, la variación no explicada o “sobrante” del experimento. Esta misma medida se encuentra en la porción superior de la salida de *MS Excel* etiquetada “Standard Error”. No sorprenderá saber que los estadísticos t y F y sus valores p hallados en la salida impresa se usan para probar hipótesis estadísticas. En la siguiente sección explicamos estas entradas.

TIP MI CONSEJO

Busque a y b en la columna llamada “Coef” o “Coefficients”

8.4 EJERCICIOS

TÉCNICAS BÁSICAS

8.1 Grafique la recta correspondiente a la ecuación $y = 2x + 1$ al graficar los puntos correspondientes a $x = 0, 1$ y 2 . Dé el punto de intersección con el eje y y la pendiente para la recta.

8.2 Grafique la recta correspondiente a la ecuación $y = -2x + 1$ al graficar los puntos correspondientes a $x = 0, 1$ y 2 . Dé el punto de intersección con el eje y y la pendiente para la recta. ¿Cómo se relaciona esta recta con la recta $y = 2x + 1$ del ejercicio 8.1?

8.3 Dé la ecuación y la gráfica para una recta con intersección con el eje y igual a 3 y pendiente igual a -1 .

8.4 Dé la ecuación y gráfica para una recta con intersección con el eje y igual a -3 y pendiente igual a 1 .

8.5 ¿Cuál es la diferencia entre modelos matemáticos deterministas y probabilísticos?

8.6 Se le dan cinco puntos con estas coordenadas:

x	-2	-1	0	1	2
y	1	1	3	5	5

- Use el método de entrada de datos en su calculadora científica o graficadora para introducir las $n = 5$ observaciones. Encuentre las sumas de cuadrados y productos cruz, S_{xx} , S_{xy} y S_{yy} .
 - Encuentre la recta de mínimos cuadrados para los datos.
 - Grafique los cinco puntos y grafique la recta de la parte b. ¿La recta parece proporcionar un buen ajuste para los puntos de datos?
 - Construya la tabla ANOVA para la regresión lineal.
- 8.7** Seis puntos tienen estas coordenadas:

x	1	2	3	4	5	6
y	5.6	4.6	4.5	3.7	3.2	2.7

- Encuentre la recta de mínimos cuadrados para los datos.
- Grafique los seis puntos y grafique la recta. ¿La recta parece proporcionar un buen ajuste para los puntos de datos?
- Use la recta de mínimos cuadrados para predecir el valor de y cuando $x = 3.5$.
- Llene los espacios faltantes en la tabla *MINITAB* de análisis de varianza.

Tabla ANOVA MINITAB para el ejercicio 8.7

Analysis of Variance

Source	DF	SS	MS
Regression	*	***	5.4321
Residual Error	*	0.1429	***
Total	*	5.5750	

8.8 Seis puntos tienen estas coordenadas:

x	1	2	3	4	5	6
y	9.7	6.5	6.4	4.1	2.1	1.0

- Encuentre la recta de mínimos cuadrados para los datos.
- Grafique los seis puntos y trace la recta. ¿La recta parece proporcionar un buen ajuste a los puntos de datos?
- Use la recta de mínimos cuadrados para predecir el valor de y cuando $x = 3.5$.
- Llene las entradas faltantes en la tabla de análisis de varianza de *MS Excel*.

ANOVA

	gl	SS	MS
Regresión	*	***	49.7286
Residual	*	1.7848	***
Total	*	51.5133	

APLICACIONES

8.9 Profesor Asimov El profesor Isaac Asimov fue uno de los escritores más prolíficos de todos los tiempos. Escribió casi 500 libros durante una carrera de 40 años. De hecho, conforme su carrera avanzaba fue incluso más productivo en términos del número de libros escritos en un periodo determinado.¹ Los datos siguientes dan el tiempo, en meses, necesario para escribir sus libros en incrementos de 100:

Número de libros, x	100	200	300	400	490
Tiempo en meses, y	237	350	419	465	507

- Suponga que el número de libros x y el tiempo en meses y están relacionados linealmente. Encuentre la recta de mínimos cuadrados que relacione y con x .
- Grafique el tiempo como función del número de libros escritos usando una gráfica de dispersión y grafique la recta de mínimos cuadrados en el mismo papel. ¿Le parece que la recta da un buen ajuste a los puntos de datos?
- Construya una tabla ANOVA para la regresión lineal.

DATOS
EX0810

8.10 Un experimento químico Con el uso de un procedimiento químico llamado *polarografía diferencial de pulsos*, un químico midió la máxima corriente generada (en microamperes) cuando una solución que contenía una cantidad determinada de níquel (en partes por mil millones, ppm) se agregó a un regulador.²

 $x = \text{Ni (ppb)}$ $y = \text{Corriente máxima (mA)}$

19.1	.095
38.2	.174
57.3	.256
76.2	.348
95	.429
114	.500
131	.580
150	.651
170	.722

- Use el método de entrada de datos en su calculadora para calcular las sumas de cuadrados preliminares y productos cruz S_{xx} , S_{yy} y S_{xy} .
- Calcule la recta de regresión de mínimos cuadrados.
- Grafique los puntos y la recta ajustada. ¿Le parece razonable la suposición de una relación lineal?
- Use la recta de regresión para predecir la máxima corriente generada cuando una solución, que contenga 100 ppm de níquel, se agregue al regulador.
- Construya la tabla ANOVA para la regresión lineal.

DATOS
EX0811

8.11 Privación de sueño Se realizó un estudio para determinar los efectos de la privación de sueño en la capacidad de las personas para resolver problemas cuando no duermen. Un total de 10 personas participaron en el estudio, dos en cada uno de cinco niveles de privación de sueño: 8, 12, 16, 20 y 24 horas. Después del periodo de privación de sueño, a cada persona se le aplicó un conjunto de problemas de adición sencillos, registrándose el número de errores. Se obtuvieron estos resultados:

Número de errores, y	8, 6	6, 10	8, 14
Número de horas sin dormir, x	8	12	16
Número de errores, y	14, 12	16, 12	
Número de horas sin dormir, x	20	24	

- ¿Cuántos pares de observaciones hay en el experimento?
- ¿Cuál es el número total de grados de libertad?
- Complete la salida impresa MINITAB.

Salida MINITAB para el ejercicio 8.11

Análisis de regresión: y contra x

The regression equation is
 $y = 3.00 + 0.475 x$

Predictor	Coef	SE Coef	T	P
Constant	3.000	2.127	1.41	0.196
x	***	0.1253	3.79	0.005

S = 2.24165 R-Sq = 64.2% R-Sq(adj) = 59.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	**	72.200	72.200	14.37	0.005
Residual Error	**	***	5.025		
Total	**	***			

d. ¿Cuál es la ecuación de predicción de mínimos cuadrados?

e. Use la ecuación de predicción para predecir el número de errores para una persona que no ha dormido durante 10 horas.

8.12 Privación de sueño II Consulte los datos proporcionados en el experimento de privación de sueño del ejercicio 8.11. Responda las preguntas planteadas en las partes a, b, d y e de dicho ejercicio completando la siguiente salida impresa de *MS Excel*:

ANOVA					
	gl	SS	MS	F	F de significancia
Regresión	**	72.2	72.2	14.36816	0.005308
Residual	**	***	5.025		
Total	**	***			
	Coefficientes	Error estándar	Estad. t	Valor P	
Intersección	3	2.126617	1.410691	0.196016	
x	0.475	0.125312	3.790535	0.005308	

DATOS
EX0813

8.13 Exámenes de aprovechamiento El Índice de Aprovechamiento Académico (API) es una medida de aprovechamiento escolar que se basa en los resultados del examen Stanford 9. Las calificaciones van de 200 a 1000, con 800 considerado como objetivo de largo alcance para escuelas. La tabla siguiente muestra el API para ocho escuelas elementales en el condado de Riverside, California, junto con el porcentaje de estudiantes de esa escuela que son considerados Estudiantes del Idioma Inglés (EL).³

Escuela	1	2	3	4	5	6	7	8
API	745	808	798	791	854	688	801	751
EL	71	18	24	50	17	71	11	57

- a. ¿Cuál de las dos variables es la variable independiente y cuál es la dependiente? Explique su selección.
- b. Use una gráfica de dispersión para graficar los datos. ¿La suposición de una relación lineal entre x y y es razonable?

- c. Suponiendo que x y y estén relacionadas linealmente, calcule la recta de regresión de mínimos cuadrados.
- d. Grafique la recta sobre la gráfica de dispersión de la parte b. ¿La recta ajusta a lo largo de los puntos de datos?

DATOS
EX0814

8.14 ¿Qué tan largo es? ¿Qué tan bueno es usted para hacer estimaciones? Para probar la capacidad de una persona para estimar tamaños, se le mostraron 10 diferentes objetos y se le pidió estimar su longitud o diámetro. A continuación se midió el objeto y los resultados se registraron en la tabla siguiente.

Objeto	Estimado (pulgadas)	Real (pulgadas)
Lápiz	7.00	6.00
Plato de comida	9.50	10.25
Libro 1	7.50	6.75
Teléfono celular	4.00	4.25
Fotografía	14.50	15.75
Juguete	3.75	5.00
Cinturón	42.00	41.50
Pinza para ropa	2.75	3.75
Libro 2	10.00	9.25
Calculadora	3.50	4.75

- a. Encuentre la recta de regresión de mínimos cuadrados para predecir la medida real como función de la medición estimada.
- b. Grafique los puntos y la recta ajustada. ¿Le parece razonable la suposición de una relación lineal?

DATOS
EX0815

8.15 Entrevistas de prueba De dos técnicas existentes para evaluación de personal, la primera requiere una entrevista de prueba de dos horas mientras que la segunda se puede completar en menos de una hora. Las puntuaciones para cada una de las 15 personas que tomaron ambas pruebas se dan en la tabla siguiente.

Solicitante	Prueba 1 (x)	Prueba 2 (y)
1	75	38
2	89	56
3	60	35
4	71	45
5	92	59
6	105	70
7	55	31
8	87	52
9	73	48
10	77	41
11	84	51
12	91	58
13	75	45
14	82	49
15	76	47

- Construya una gráfica de dispersión para los datos. ¿Le parece razonable la suposición de linealidad?
- Encuentre la recta de mínimos cuadrados para los datos.
- Use la recta de regresión para predecir la puntuación en la segunda prueba para un solicitante que obtuvo 85 puntos en la prueba 1.

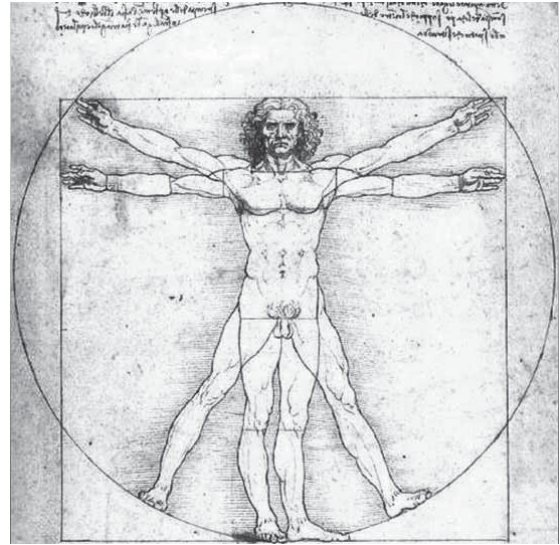
8.16 Entrevistas de prueba, continúa Consulte el ejercicio 8.15. Construya la tabla ANOVA para la regresión lineal que relacione y , la puntuación en la prueba 2, con x , la puntuación en la prueba 1.

DATOS
EX0817

8.17 Distancia entre brazos extendidos y estatura Leonardo da Vinci (1452-1519) dibujó la figura de un hombre, indicando que la distancia entre los brazos extendidos de una persona (midiendo por la espalda con los brazos extendidos para formar una “T”) es casi igual a la estatura de una persona. Para probar lo dicho por él, medimos ocho personas con los siguientes resultados:

Persona	1	2	3	4
Distancia entre los brazos extendidos (pulgadas)	68	62.25	65	69.5
Estatura	69	62	65	70
Persona	5	6	7	8
Distancia entre los brazos extendidos (pulgadas)	68	69	62	60.25
Estatura	67	67	63	62

- Trace la gráfica de dispersión para distancia entre los brazos extendidos y estatura. Use la misma escala en los ejes horizontal y vertical. Describa la relación entre las dos variables.
- Si Da Vinci estaba en lo correcto y la distancia entre los brazos extendidos de una persona es casi igual a la estatura de esa persona, ¿cuál debe ser la pendiente de la recta de regresión?



- Calcule la recta de regresión para predecir la estatura con base en la distancia entre los brazos extendidos de una persona. ¿El valor de la pendiente b confirma sus conclusiones de la parte **b**?
- Si una persona tiene una distancia de 62 pulgadas entre los brazos extendidos, ¿cuál sería el pronóstico de usted respecto a la estatura de la persona?

DATOS
EX0818

8.18 Fresas Los datos siguientes se obtuvieron en un experimento que relacionaba la variable dependiente y (textura de fresas), con x (temperatura codificada de almacenamiento)

x	-2	-2	0	2	2
y	4.0	3.5	2.0	0.5	0.0

- Encuentre la recta de mínimos cuadrados para los datos.
- Grafique los puntos de datos y grafique la recta de mínimos cuadrados como prueba de sus cálculos.
- Construya la tabla ANOVA.

PRUEBA DE LA UTILIDAD DEL MODELO DE REGRESIÓN LINEAL

8.5

Al considerar la regresión lineal, es posible plantear dos preguntas:

- ¿La variable independiente x es útil para predecir la variable de respuesta y ?
- Si es así, ¿qué tan bien funciona?

Esta sección examina varias pruebas estadísticas y medidas que le ayudarán a tener algunas respuestas. Una vez que haya determinado que el modelo está funcionando, puede entonces usarlo para predecir la respuesta y para un valor determinado de x .

Inferencias respecto a β , la pendiente de la recta de medias

¿La recta de regresión de mínimos cuadrados es útil? Es decir, ¿la ecuación de regresión que utiliza información dada por x es considerablemente mejor que la pronosticadora simple $\bar{y}$ que no se apoya en x ? Si la variable independiente x *no es útil* en el modelo de población $y = \alpha + \beta x + \epsilon$, entonces el valor de y no cambia para valores diferentes de x . La única forma en que esto ocurre para todos los valores de x es cuando la pendiente β de la recta de medias es igual a 0. Esto indicaría que la relación entre y y x no es lineal, de modo que la pregunta inicial acerca de la utilidad de la variable independiente x se puede expresar también como: ¿Hay una relación lineal entre x y y ?

Se puede contestar esta pregunta usando ya sea una prueba de hipótesis o un intervalo de confianza para β . Estos procedimientos están basados en la distribución muestral de b , el estimador muestral de la pendiente β . Se puede demostrar que, si las suposiciones acerca del error aleatorio ϵ son válidas, entonces el estimador b tiene una distribución normal en muestreo repetido con media

$$E(b) = \beta$$

y error estándar dado por

$$SE = \sqrt{\frac{\sigma^2}{S_{xx}}}$$

donde σ^2 es la varianza del error aleatorio ϵ . Como el valor de σ^2 se estima con $s^2 = \text{MSE}$, se pueden basar inferencias en la estadística dada por

$$t = \frac{b - \beta}{\sqrt{\text{MSE}/S_{xx}}}$$

que tiene una distribución t con $gl = (n - 2)$, los grados de libertad asociados con MSE.

PRUEBA DE HIPÓTESIS RESPECTO A LA PENDIENTE DE UNA RECTA

1. Hipótesis nula: $H_0: \beta = \beta_0$
2. Hipótesis alternativa

Prueba de una cola

$$H_a: \beta > \beta_0$$

$$(\text{o } \beta < \beta_0)$$

Prueba de dos colas

$$H_a: \beta \neq \beta_0$$

3. Estadístico de prueba $t = \frac{b - \beta_0}{\sqrt{\text{MSE}/S_{xx}}}$

Cuando se satisfacen las suposiciones dadas en la sección 8.2, el estadístico de prueba tendrá una distribución t de Student con $(n - 2)$ grados de libertad.

PRUEBA DE HIPÓTESIS RESPECTO A LA PENDIENTE DE UNA RECTA (continuación)

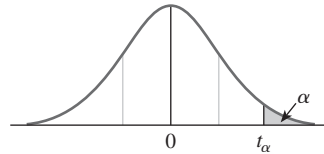
4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola

$$t > t_\alpha$$

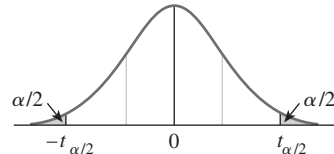
(o $t < t_\alpha$ cuando la hipótesis alternativa sea $H_a: \beta > \beta_0$)

o cuando valor $p < \alpha$



Prueba de dos colas

$$t > t_{\alpha/2} \text{ o } t < -t_{\alpha/2}$$



Los valores de t_α y $t_{\alpha/2}$ correspondientes a $(n - 2)$ grados de libertad pueden encontrarse usando la tabla 4 del Anexo.

EJEMPLO

8.2

Determine si hay una relación lineal significativa entre las calificaciones en cálculo y las puntuaciones de examen de la tabla 8.1. Pruebe al nivel de significancia de 5%.

Solución Las hipótesis a probar son

$$H_0: \beta = 0 \text{ contra } H_a: \beta \neq 0$$

y el valor observado del estadístico de prueba se calcula como

$$t = \frac{b - 0}{\sqrt{\text{MSE}/S_{xx}}} = \frac{.7656 - 0}{\sqrt{75.7532/2474}} = 4.38$$

con $(n - 2) = 8$ grados de libertad. Con $\alpha = .05$, se puede rechazar H_0 cuando $t > 2.306$ o $t < -2.306$. Como el valor observado del estadístico de prueba cae en la región de rechazo, H_0 es rechazada y se puede concluir que hay una relación lineal significativa entre las calificaciones en cálculo y la puntuación de examen para la población de estudiantes de primer año de universidad.

Otra forma de hacer inferencias acerca del valor de β es construir un intervalo de confianza para β y examinar el rango de posibles valores para β .

UN INTERVALO DE CONFIANZA $(1 - \alpha)100\%$ PARA β

$$b \pm t_{\alpha/2}(\text{SE})$$

donde $t_{\alpha/2}$ está basada en $(n - 2)$ grados de libertad, $s^2 = \text{MSE}$, y

$$\text{SE} = \sqrt{\frac{s^2}{S_{xx}}} = \sqrt{\frac{\text{MSE}}{S_{xx}}}$$

EJEMPLO

8.3

Encuentre una estimación de intervalo de confianza de 95% de la pendiente β para los datos de las calificaciones en cálculo de la tabla 8.1.

Solución Sustituyendo valores previamente calculados en

$$b \pm t_{.025} \sqrt{\frac{\text{MSE}}{S_{xx}}}$$



APPLET EN LÍNEA

Prueba t para la pendiente
y probabilidades de
Student

tendremos

$$.766 \pm 2.306 \sqrt{\frac{75.7532}{2474}}$$
$$.766 \pm .404$$

El intervalo de confianza de 95% resultante es .362 a 1.170. Como el intervalo no contiene 0, se puede concluir que el verdadero valor de β no es 0 y se puede rechazar la hipótesis nula $H_0: \beta = 0$ a favor de $H_a: \beta \neq 0$, conclusión que está de acuerdo con los hallazgos del ejemplo 8.2. Además, la estimación del intervalo de confianza indica que hay un aumento desde sólo .4 hasta 1.2 puntos en una puntuación de examen de cálculo por cada aumento de 1 punto en la puntuación del examen de aprovechamiento.

Si usted utiliza un programa de cómputo para hacer un análisis de regresión, encontrará la estadística t y su valor p en la salida impresa. En la segunda sección de la salida del *MINITAB* en la figura 8.7a), encontrará la estimación b de mínimos cuadrados de la recta de la pendiente marcada “x”, junto con su error estándar “SE Coef”, el valor calculado del estadístico de prueba “T” usado para probar la hipótesis $H_0: \beta = 0$ y su valor p “P”. Encontrará la misma información en la última línea de la salida de *MS Excel* en la figura 8.7b), junto con los límites de confianza superior e inferior de un intervalo de confianza de 95% para β . La prueba t para regresión significativa, $H_0: \beta = 0$, tiene un valor p de $P = .002$, y la hipótesis nula se rechaza, como en el ejemplo 8.2. Hay una relación lineal significativa entre x y y .

FIGURA 8.7(a) ● Análisis de regresión: y contra x

Salida *MINITAB* para los datos de calificaciones en cálculo

TIP MI CONSEJO

Busque el error estándar de b en la columna marcada “SE Coef” en la salida de *MINITAB* y “Standard Error” en la salida de *MS Excel*

The regression equation is
 $y = 40.8 + 0.766 x$

Predictor	Coef	SE Coef	T	P
Constant	40.784	8.507	4.79	0.001
x	0.7656	0.1750	4.38	0.002

$S = 8.70363$ $R\text{-Sq} = 70.5\%$ $R\text{-Sq(adj)} = 66.8\%$

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	1450.0	1450.0	19.14	0.002
Residual Error	8	606.0	75.8		
Total	9	2056.0			

FIGURA 8.7(b) ● SUMMARY OUTPUT

Salida de *MS Excel* para los datos de las calificaciones de cálculo

Regression Statistics						
Multiple R	0.8398					
R Square	0.7052					
Adjusted R Square	0.6684					
Standard Error	8.7036					
Observations	10					
ANOVA						
	gl	SS	MS	F	F de significancia	
Regression	1	1449.974	1449.974	19.141	0.002	
Residual	8	606.026	75.753			
Total	9	2056				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	40.784	8.507	4.794	0.001	21.167	60.401
x	0.766	0.175	4.375	0.002	0.362	1.169

TIP MI CONSEJO

Las pruebas F de ANOVA siempre son de una cola (cola superior)

El análisis de varianza de la prueba F

La parte del análisis de varianza de la salida impresa de las figuras 8.7a) y 8.7b) muestra un estadístico F dado por

$$F = \frac{MSR}{MSE} = 19.14$$

con grado de libertad 1 en el numerador y $(n - 2) = 8$ grados de libertad en el denominador. Esto es un *estadístico equivalente de prueba* que también se puede usar para probar la hipótesis $H_0: \beta = 0$. Observe que, dentro del error de redondeo, el valor de F es igual a t^2 con valor p idéntico. En este caso, si se usa una precisión de cinco lugares decimales antes de redondeo, se encuentra que $t^2 = (.76556/.17498)^2 = (4.37513)^2 = 19.14175 \approx 19.14 = F$ como se da en la salida impresa. Esto no es por casualidad y resulta del hecho de que el cuadrado de un estadístico t con gl grados de libertad tiene la misma distribución que un estadístico F con grados de libertad 1 en el numerador y gl en el denominador. La prueba F es una prueba más general de la utilidad del modelo y se puede usar cuando el modelo tenga más de una variable independiente.

Medir la fuerza de la relación: el coeficiente de determinación

¿Qué tan bien se ajusta el modelo de regresión? Para contestar esta pregunta, se puede usar una medida relacionada con el *coeficiente de correlación* r . Recuerde que

$$r = \frac{s_{xy}}{s_x s_y} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} \quad \text{para } -1 \leq r \leq 1$$

donde s_{xy} , s_x y s_y se definieron antes y las diversas sumas de cuadrados se definieron en la sección 8.4.

La suma de cuadrados para regresión, SSR, en el análisis de varianza, mide la parte de la variación total $TSS = S_{yy}$, que puede ser explicada por la regresión de y en x . La parte restante, SSE, es la variación “no explicada” atribuida al error aleatorio. Una forma de medir la fuerza de la relación entre la variable de respuesta y y la variable de predicción x es calcular el **coeficiente de determinación**, la proporción de la variación total que es explicada por la regresión lineal de y en x . Para los datos de calificaciones en cálculo, esta proporción es igual a

$$\frac{SSR}{TSS} = \frac{1450}{2056} = .705 \quad \text{o} \quad 70.5\%$$

Puesto que $TSS = S_{yy}$ y $SSR = \frac{(S_{xy})^2}{S_{xx}}$, se puede escribir

$$\frac{SSR}{TSS} = \frac{(S_{xy})^2}{S_{xx} S_{yy}} = \left(\frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} \right)^2 = r^2$$

Por tanto, el coeficiente de determinación, que se calculó como SSR/TSS , es simplemente el cuadrado del coeficiente de correlación r . Es la entrada marcada “R-Sq” en la figura 8.7a) y “R cuadrada” en la figura 8.7b).

Recuerde que la tabla del análisis de varianza aísla la variación debida a regresión (SSR) de la variación total del experimento. Al hacer esto se reduce la cantidad de *variación aleatoria* del experimento, ahora medida por SSE en lugar de TSS. En este contexto, el **coeficiente de determinación**, r^2 , se puede definir como sigue:

TIP MI CONSEJO

En las salidas impresas de computadora, r^2 a menudo se da como un **porcentaje** más que como una proporción

TIP MI CONSEJO

r^2 se denomina "R-Sq" en la salida impresa de MINITAB y "R Square" en la salida impresa de Excel

Definición El **coeficiente de determinación** r^2 se puede interpretar como el porcentaje de reducción en la variación total en el experimento obtenido al usar la recta de regresión $\hat{y} = a + bx$ en lugar de ignorar x y usar la media muestral $\bar{y}$ para predecir la variable de respuesta y .

Para los datos de calificaciones en cálculo, una reducción de $r^2 = .705$ o sea 70.5% es considerable. El modelo de regresión funciona muy bien.

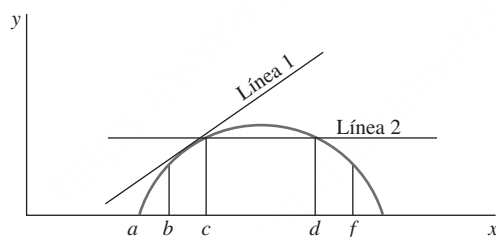
Interpretación de los resultados de una regresión significativa

Una vez que usted haya efectuado la prueba t o la prueba F para determinar la significancia de la regresión lineal, con todo cuidado debe interpretar sus resultados. La pendiente β de la recta de medias se estima con base en datos de sólo una región de observación en particular. Incluso si no rechaza la hipótesis nula de que la pendiente de la recta es igual a 0, no necesariamente significa que y y x no estén relacionadas. Puede ser que haya cometido un error tipo II, declarando falsamente que la pendiente es 0 y que x y y no están relacionadas.

Ajuste del modelo erróneo

Puede ocurrir que y y x estén perfectamente relacionadas en una forma no lineal, como se ve en la figura 8.8. A continuación veamos tres posibilidades:

FIGURA 8.8
Relación curvilínea



Si se tomaron observaciones sólo dentro del intervalo $b < x < c$, la relación aparecería lineal con pendiente positiva.

- Si se tomaron observaciones sólo dentro del intervalo $d < x < f$, la relación aparecería lineal con pendiente negativa.
- Si se tomaron observaciones sobre el intervalo $c < x < d$, la recta estaría ajustada con una pendiente cercana a 0, lo cual indica que no hay relación lineal entre y y x .

Para el ejemplo que se ilustra en la figura 8.8 ninguna recta describe con precisión la verdadera relación entre x y y , que es en realidad una *relación curvilínea*. En este caso, hemos escogido un *modelo erróneo* para describir la relación. A veces este tipo de error se puede detectar usando gráficas residuales, que es el tema de la sección 8.7.

TIP MI CONSEJO

Es peligroso tratar de predecir valores de y fuera del rango de los datos ajustados

Extrapolación

Un problema serio es aplicar los resultados de un análisis de regresión lineal a valores de x que *no estén incluidos* dentro del rango de los datos ajustados. Esto se llama **extrapolación** y puede conducir a errores graves en la predicción, como se ve para la línea 1 de la figura 8.8. Los resultados de una predicción serían buenos en el intervalo $b < x < c$ pero sobreestimarían gravemente los valores de y para $x > c$.

Causalidad

Cuando haya una regresión significativa de y y x , es tentador concluir que x causa a y . No obstante, es posible que una o más variables desconocidas que ni siquiera se hayan medido y que no estén incluidas en el análisis puedan estar causando la relación observada. En general, el estadístico informa los resultados de un análisis pero deja las conclusiones respecto a la causalidad a científicos e investigadores que son expertos en estos campos de actividad. Estos expertos están mejor preparados para tomar esas decisiones.

8.5

EJERCICIOS

TÉCNICAS BÁSICAS

8.19 Consulte el ejercicio 8.6. Los datos se reproducen a continuación.

x	-2	-1	0	1	2
y	1	1	3	5	5

- ¿Los datos presentan suficiente evidencia para indicar que y y x están relacionadas linealmente? Pruebe la hipótesis de que $\beta = 0$ al nivel de significancia de 5%.
- Use la tabla ANOVA del ejercicio 8.6 para calcular $F = MSR/MSE$. Verifique que el cuadrado de la estadística t empleada en la parte **a** es igual a F .
- Compare el valor crítico de dos colas para la prueba t de la parte **a** con el valor crítico para F con $\alpha = .05$. ¿Cuál es la relación entre los valores críticos?

8.20 Consulte el ejercicio 8.19. Encuentre un intervalo de confianza de 95% para la pendiente de la recta. ¿Qué significa la frase “95% de confianza”?

8.21 Consulte el ejercicio 8.7. Los datos, junto con la tabla MINITAB del análisis de varianza se reproducen a continuación.

x	1	2	3	4	5	6
y	5.6	4.6	4.5	3.7	3.2	2.7

Tabla MINITAB ANOVA para el ejercicio 8.21

Análisis de regresión: y contra x

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	5.4321	5.4321	152.10	0.000
Residual Error	4	0.1429	0.0357		
Total	5	5.5750			

- ¿Los datos dan suficiente evidencia para indicar que y y x están relacionados linealmente? Use la información de la salida impresa MINITAB para contestar esta pregunta al nivel de significancia de 1%.
- Calcule el coeficiente de determinación r^2 . ¿Qué información da este valor acerca de la utilidad del modelo lineal?

8.22 Consulte el ejercicio 8.8. Los datos, junto con la tabla del análisis de varianza de MS Excel se reproducen a continuación:

x	1	2	3	4	5	6
y	9.7	6.5	6.4	4.1	2.1	1.0

Tabla ANOVA de MS Excel para el ejercicio 8.22

ANOVA					
	gl	SS	MS	F	Significance F
Regression	1	49.72857	49.72857	111.45	0.000
Residual	4	1.78476	0.4462		
Total	5	51.51333			

- ¿Los datos proporcionan suficiente evidencia para indicar que y y x están relacionados linealmente? Use la información en la salida impresa para responder esta pregunta al nivel de significancia del 5%.
- Calcule el coeficiente de determinación r^2 . ¿Qué información proporciona este valor acerca de la utilidad del modelo lineal?

APLICACIONES



8.23 Chirridos de grillos Se encontró que los grillos machos chirrían frotando sus alas frontales, y su chirrido depende de la temperatura. La tabla siguiente muestra el número de chirridos por segundo para un grillo, registrado a 10 temperaturas diferentes:

Chirridos por segundo	20	16	19	18	18	16	14	17	15	16
Temperatura	88	73	91	85	82	75	69	82	69	83

- Use las fórmulas dadas en este capítulo para encontrar la recta de regresión de mínimos cuadrados relacionando el número de chirridos con la temperatura.
- ¿Los datos proporcionan evidencia suficiente para indicar que hay una relación lineal entre el número de chirridos y la temperatura?
- Calcule r^2 . ¿Qué le dice este valor acerca de la efectividad del análisis de regresión lineal?

DATOS
EX0824

8.24 Tiempos de gestación y longevidad La tabla siguiente muestra el tiempo de gestación en días y la longevidad promedio en años para una variedad de mamíferos en cautiverio.⁴

Animal	Gestación (días)	Longevidad promedio (años)
Babuino	187	20
Bisonte	285	15
Caballo	330	20
Cabra (doméstica)	151	8
Caribú	250	15
Cerdo (doméstico)	112	10
Gato (doméstico)	63	12
Gorila	258	20
Lobo (de crin)	63	5
Mono (rhesus)	166	15
Oso (negro)	219	18
Oveja (doméstica)	154	12
Puma	90	12
Ratón (de campo)	21	3
Zorro (rojo)	52	7

- Si desea estimar la longevidad promedio de un animal con base en su tiempo de gestación, ¿cuál variable es la variable de respuesta y cuál es la variable predictora independiente?
- Suponga que hay una relación lineal entre el tiempo de gestación y la longevidad. Calcule la recta de regresión de mínimos cuadrados que describe la longevidad como una función lineal del tiempo de gestación.
- Grafique los puntos de datos y la recta de regresión. ¿Parece que la recta se ajusta a los datos?
- Use las pruebas estadísticas y medidas apropiadas para explicar la utilidad del modelo de regresión para predecir la longevidad.

8.25 Profesor Asimov, continúa Consulte los datos del ejercicio 8.9, que relacionan x , el número de libros escritos por el profesor Isaac Asimov, con y , el número de meses que le tomó escribir sus libros (en incrementos de 100). Los datos se reproducen a continuación.

Número de libros, x	100	200	300	400	490
Tiempo en meses, y	237	350	419	465	507

- ¿Los datos apoyan la hipótesis de que $\beta = 0$? Use el método del valor p , enlazando el valor p usando la tabla 4 del Anexo. Explique sus conclusiones en términos prácticos.
- Use la tabla ANOVA del ejercicio 8.9, parte c, para calcular el coeficiente de determinación r^2 . ¿Qué

reducción de porcentaje en la variación total se alcanza usando el modelo de regresión lineal?

- Grafique los datos o consulte la gráfica del ejercicio 8.9, parte b. ¿Los resultados de las partes a y b indican que el modelo da un buen ajuste para los datos? ¿Hay algunas suposiciones que pueden haber sido violadas al ajustar el modelo lineal?

8.26 Consulte el experimento de privación de sueño descrito en los ejercicios 8.11 y 8.12 y el conjunto de datos EX0811. Los datos y las salidas impresas de MINITAB y MS Excel se reproducen a continuación.

Número de errores, y	8, 6	6, 10	8, 14
Número de horas sin dormir, x	8	12	16
Número de errores, y	14, 12	16, 12	
Número de horas sin dormir, x	20	24	

Salida MINITAB para el ejercicio 8.26

Análisis de regresión: y contra x

The regression equation is
 $y = 3.00 + 0.475 x$

Predictor	Coef	SE Coef	T	P
Constant	3.000	2.127	1.41	0.196
x	0.4750	0.1253	3.79	0.005

$S = 2.24165$ $R\text{-Sq} = 64.2\%$ $R\text{-Sq(adjusted)} = 59.8\%$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	72.200	72.200	14.37	0.005
Residual Error	8	40.200	5.025		
Total	9	112.400			

Salida de MS Excel para el ejercicio 8.26

ANOVA						
	gl	SS	MS	F	Significance F	
Regression	1	72.2	72.2	14.368	0.005	
Residual	8	40.2	5.025			
Total	9	112.4				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	3	2.1266	1.4107	0.1960	-1.9040	7.9040
x	0.475	0.1253	3.7905	0.0053	0.1860	0.7640

- ¿Los datos presentan suficiente evidencia para indicar que el número de errores está linealmente relacionado con el número de horas sin dormir? Identifique los dos estadísticos de prueba en la salida impresa que puedan usarse para contestar esta pregunta.
- ¿Esperaría usted que la relación entre y y x sea lineal si x varió en un rango más amplio (por ejemplo, $x = 4$ a $x = 48$)?

- c. ¿Cómo describe la fuerza de la relación entre y y x ?
- d. ¿Cuál es la mejor estimación de la varianza poblacional común σ^2 ?
- e. Encuentre un intervalo de confianza de 95% para la pendiente de la recta.

8.27 Fresas II Los datos siguientes (ejercicio 8.18 y conjunto de datos EX0818) se obtuvieron en un experimento que relacionaba la variable dependiente, y (textura de fresas), con x (temperatura de almacenamiento codificada). Use la información del ejercicio 8.18 para contestar las preguntas siguientes:

x	-2	-2	0	2	2
y	4.0	3.5	2.0	0.5	0.0

- a. ¿Cuál es la mejor estimación de σ^2 , la varianza del error aleatorio $\bar{y}$?
- b. ¿Los datos indican que la textura y la temperatura de almacenamiento están relacionadas linealmente? Use $\alpha = .05$.
- c. Calcule el coeficiente de determinación, r^2 .
- d. ¿De qué valor es el modelo *lineal* para aumentar la precisión de la predicción cuando se compara con la variable de predicción $\bar{y}$?

DATOS
EX0828

8.28 Laptops y aprendizaje Se estudiaron dos grupos de primer año de álgebra, uno de los cuales utilizaba computadoras laptop en la escuela y en casa, en tanto que el otro grupo no las utilizaba. En cada grupo, a los estudiantes se les dio una encuesta al principio y al final del semestre, que medía su nivel tecnológico. Se registraron las calificaciones para la encuesta del final de semestre (x) y el examen final (y) para el grupo con laptop.⁵ Los datos y la salida impresa *MINITAB* se muestran aquí.

Estudiante	Encuesta posterior	Examen final	Estudiante	Encuesta posterior	Examen final
1	100	98	11	88	84
2	96	97	12	92	93
3	88	88	13	68	57
4	100	100	14	84	84
5	100	100	15	84	81
6	96	78	16	88	83
7	80	68	17	72	84
8	68	47	18	88	93
9	92	90	19	72	57
10	96	94	20	88	83

Salida *MINITAB* para el ejercicio 8.28

Análisis de regresión: y contra x

The regression equation is
 $y = -26.8 + 1.26 x$

Predictor	Coef	SE Coef	T	P
Constant	-26.82	14.76	-1.82	0.086
x	1.2617	0.1685	7.49	0.000

S = 7.61912 R-Sq = 75.7% R-Sq(adj) = 74.3%

Source	DF	SS	MS	F	P
Regression	1	3254.0	3254.0	56.05	0.000
Residual Error	18	1044.9	58.1		
Total	19	4299.0			

- a. Construya una gráfica de dispersión para los datos. ¿Le parece razonable la suposición de linealidad?
- b. ¿Cuál es la ecuación de la recta de regresión empleada para predecir la calificación del examen final como función de la calificación del examen posterior?
- c. ¿Los datos presentan suficiente evidencia para indicar que la calificación del examen final está linealmente relacionada con la calificación del examen posterior? Use $\alpha = .01$.
- d. Encuentre un intervalo de confianza de 99% para la pendiente de la recta de regresión.

8.29 Laptops y aprendizaje, continúa Consulte el ejercicio 8.28.

- a. Use la salida impresa *MINITAB* para hallar el valor del coeficiente de determinación, r^2 . Demuestre que $r^2 = SSR/TSS$.
- b. ¿Qué reducción de porcentaje en la variación total se obtiene al usar el modelo de regresión lineal?

8.30 Distancia entre brazos extendidos y estatura II

En el ejercicio 8.17 (conjunto de datos EX0817), medimos la distancia entre brazos extendidos y la estatura de ocho personas con los siguientes resultados:

Persona	1	2	3	4
Distancia entre los brazos extendidos (pulgadas)	68	62.25	65	69.5
Estatura (pulgadas)	69	62	65	70
Persona	5	6	7	8
Distancia entre los brazos extendidos (pulgadas)	68	69	62	60.25
Estatura (pulgadas)	67	67	63	62

- a. ¿Los datos dan suficiente evidencia para indicar que hay una relación lineal entre distancia y estatura? Pruebe al nivel de significancia de 5%.
- b. Construya un intervalo de confianza de 95% para la pendiente de la recta de medias, β .
- c. Si Leonardo da Vinci tenía razón y la distancia entre los brazos extendidos de una persona es casi igual a la estatura de esa persona, la pendiente de la recta de regresión es aproximadamente igual a 1. ¿El intervalo de confianza construido en la parte b confirma esta suposición? Explique.

HERRAMIENTAS DE DIAGNÓSTICO PARA VERIFICAR SUPOSICIONES DE LA REGRESIÓN

8.6

Aun cuando ya hemos determinado, con el uso de la prueba t para la pendiente (o la prueba F ANOVA) y el valor de r^2 , que x es útil para predecir el valor de y , los resultados de un análisis de regresión son válidos sólo cuando los datos satisfacen las suposiciones de regresión necesarias.

SUPOSICIONES DE REGRESIÓN

- La relación entre y y x debe ser lineal, dada por el modelo

$$y = \alpha + \beta x + \epsilon$$

- Los valores del término de error aleatorio ϵ : 1) son independientes, 2) tienen una media de 0 y una varianza común σ^2 , independiente de x , y 3) están normalmente distribuidos.

Como estas suposiciones son bastante similares a las presentadas en el capítulo 7 para un análisis de varianza, no debe sorprender hallar que las **herramientas de diagnóstico** para verificar estas suposiciones son las mismas que las que empleamos en ese capítulo. Estas herramientas incluyen el análisis del **error residual**, la variación no explicada en cada observación una vez que la variación explicada por el modelo de regresión se haya eliminado.

Términos de error dependientes

Es frecuente que los términos de error sean dependientes cuando las observaciones se recolectan a intervalos de tiempo regulares. Cuando éste es el caso, las observaciones forman una **serie de tiempo** cuyos términos de error están correlacionados. Esto, a su vez, causa un sesgo en las estimaciones de parámetros del modelo. Los datos de la serie de tiempo deben ser analizados usando métodos de serie de tiempo.

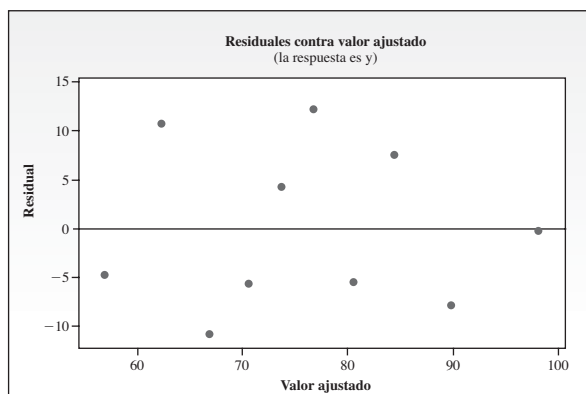
Gráficas residuales

Las otras suposiciones de regresión se pueden verificar con el uso de **gráficas residuales**, que son bastante complicadas de hacer manualmente pero fáciles de usar si se grafican en computadora.

En regresión lineal simple, se puede usar la **gráfica de residuales contra ajuste** para verificar una varianza constante así como asegurarse que el modelo lineal en verdad sea adecuado. Esta gráfica debe estar libre de patrones y aparecer como dispersión aleatoria de puntos alrededor de 0 en el eje vertical, con aproximadamente la misma dispersión vertical para todos los valores de $\hat{y}$. Una propiedad de los residuales es que suman 0 y, por tanto, tienen una media muestral de 0. La gráfica de los residuales contra ajuste para el ejemplo de las calificaciones en cálculo se ve en la figura 8.9. No hay patrones aparentes en esta gráfica residual, lo cual indica que las suposiciones del modelo parecen estar satisfechas para estos datos.

FIGURA 8.9

Gráfica de los residuales contra $\hat{y}$ para el ejemplo 8.1



TIP MI CONSEJO

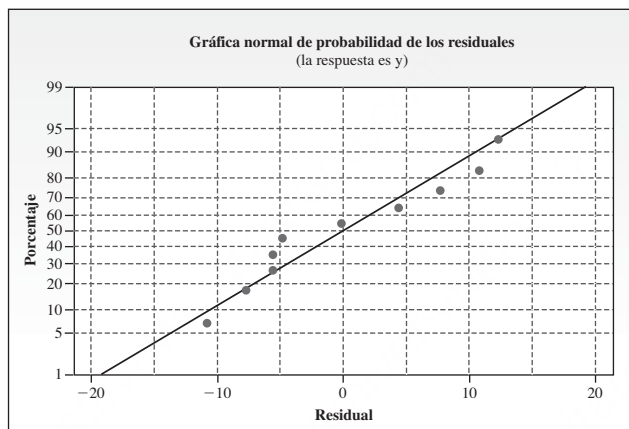
Residuales contra ajustes
 ⇔ dispersión aleatoria

Gráfica normal ⇔
 línea recta, pendiente
 ascendente

Recuerde del capítulo 7 que la **gráfica normal de probabilidad** es una gráfica que traza los residuales contra el valor esperado del residual si hubiera venido de una distribución normal. Cuando los residuales estén distribuidos normalmente o de manera aproximada, la gráfica debe aparecer como una recta con pendiente hacia arriba. La gráfica normal de probabilidad para los residuales del ejemplo 8.1 está en la figura 8.10. Con la excepción de los puntos graficados cuarto y quinto, los puntos restantes parecen estar casi sobre la recta. Esta gráfica no es rara y no indica no normalidad subyacente. Las violaciones más serias de la suposición de normalidad por lo general aparecen en las colas de la distribución porque aquí es donde la distribución normal difiere más de otros tipos de distribuciones con media y medida de dispersión similares. En consecuencia, la curvatura en cualquiera de los extremos o en ambos de la gráfica normal de probabilidad indica no normalidad.

FIGURA 8.10

Gráfica normal de probabilidad de residuales para el ejemplo 8.1

**8.6****EJERCICIOS****TÉCNICAS BÁSICAS**

8.31 ¿Cuál gráfica de diagnóstico se puede usar para determinar si los datos satisfacen la suposición de normalidad? ¿Cómo se vería la gráfica para residuales normales?

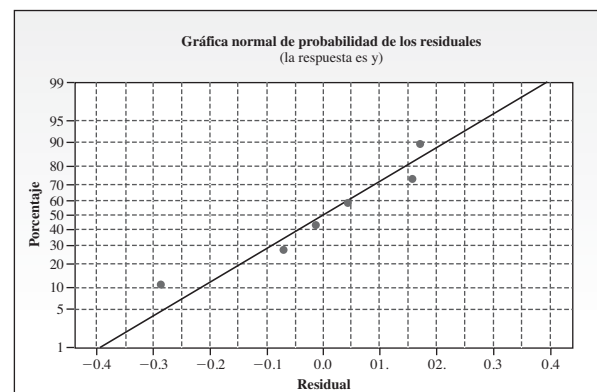
8.32 ¿Cuál gráfica de diagnóstico se puede usar para determinar si se ha usado el modelo incorrecto? ¿Cómo se vería la gráfica si se ha usado el modelo correcto?

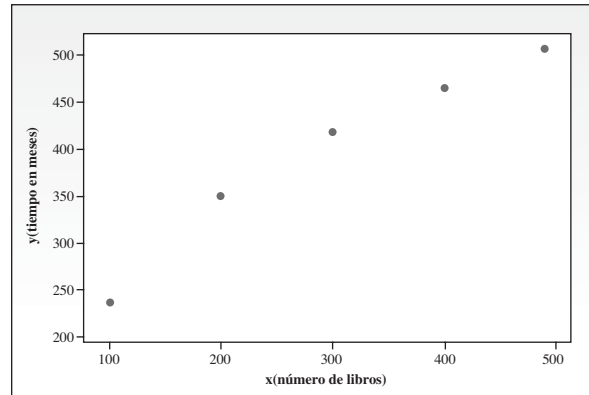
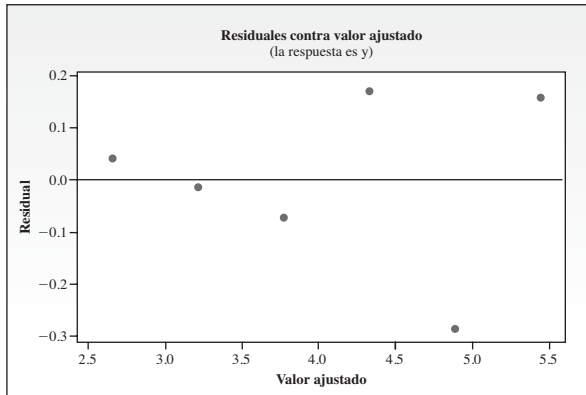
8.33 ¿Cuál gráfica de diagnóstico se puede usar para determinar si se ha violado la suposición de igual varianza? ¿Cómo se vería la gráfica cuando las varianzas son iguales para todos los valores de x ?

8.34 Consulte los datos del ejercicio 8.7. La gráfica normal de probabilidad y las gráficas de residuales contra valores ajustados generadas por *MINITAB* se muestran a

continuación. ¿Le parece que alguna de las suposiciones de regresión ha sido violada? Explique.

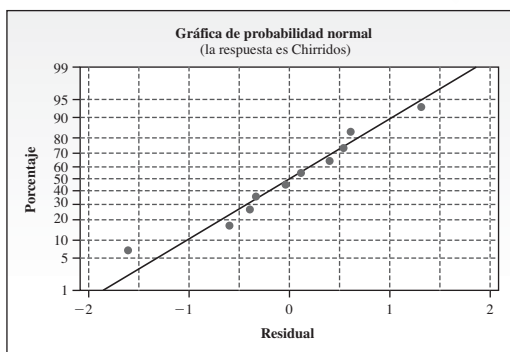
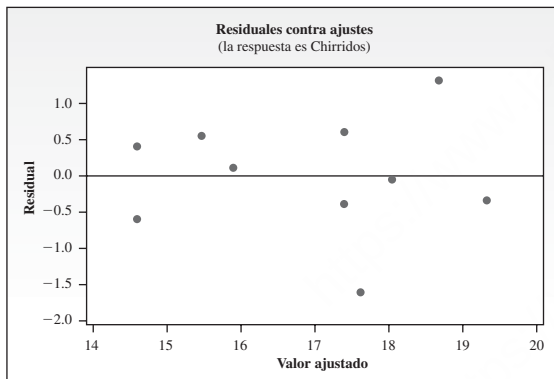
Gráficas de diagnóstico para el ejercicio 8.34





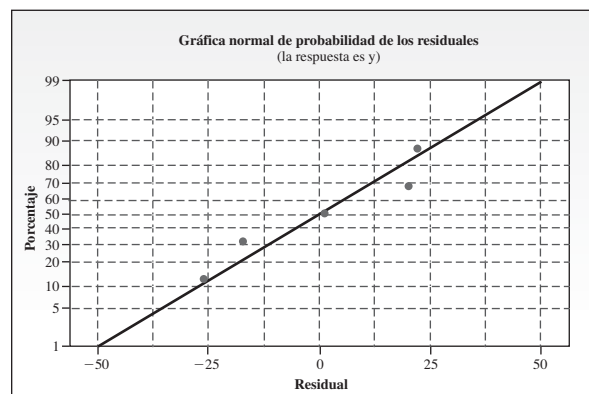
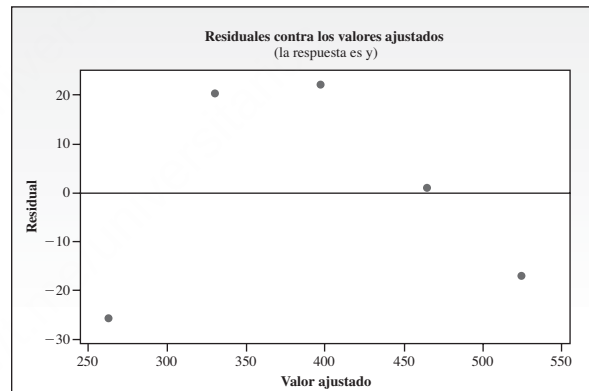
APLICACIONES

8.35 Chirridos de grillos, continúa Consulte el ejercicio 8.23, en el cual se registró el número de chirridos por segundo para un grillo a 10 temperaturas diferentes. Use las gráficas de diagnóstico del *MINITAB* para comentar sobre la validez de las suposiciones de regresión.



8.36 Profesor Asimov, otra vez Consulte el ejercicio 8.9, en el que el número x de libros escritos por Isaac Asimov se relaciona con el número de meses y que tardó en escribirlos. A continuación, una gráfica de los datos.

- ¿Se puede ver cualquier patrón que no sea una relación lineal en la gráfica original?
- El valor de r^2 para estos datos es .959. ¿Qué dice esto acerca del ajuste de la recta de regresión?
- Vea las siguientes gráficas de diagnóstico para estos datos. ¿Se ve algún patrón en los residuales? ¿Sugiere esto que la relación entre el número de meses y el número de libros escritos es algo que no sea lineal?



8.37 Laptops y aprendizaje, otra vez Consulte los datos dados en el ejercicio 8.28. La salida impresa MINITAB se reproduce aquí.

Análisis de regresión: y contra x

The regression equation is
 $y = 6.000 - 0.557 x$

Predictor	Coef	SE Coef	T	P
Constant	6.0000	0.1759	34.10	0.000
x	-0.55714	0.04518	-12.33	0.000

Predicted Values for New Observations

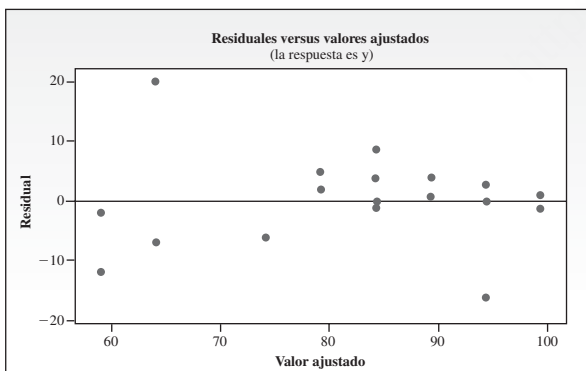
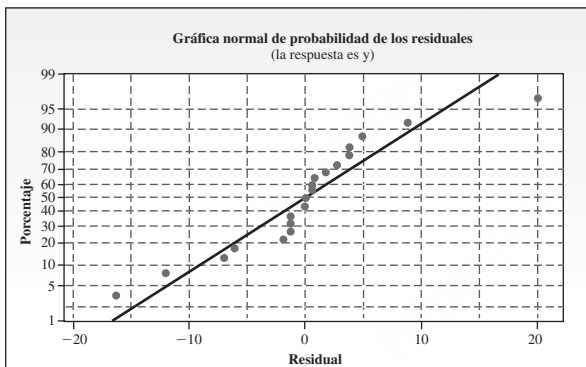
New Obs	Fit	SE Fit	95.0% CI	95.0% PI
1	4.8857	0.1027	(4.6006, 5.1708)	(4.2886, 5.4829)
2	1.5429	0.2174	(0.9392, 2.1466)	(0.7430, 2.3427)

X denotes a point that is an outlier in the predictors.

Values of Predictors for New Observations

New Obs	x
1	2.00
2	8.00

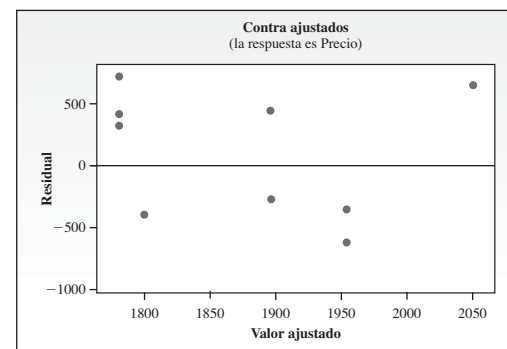
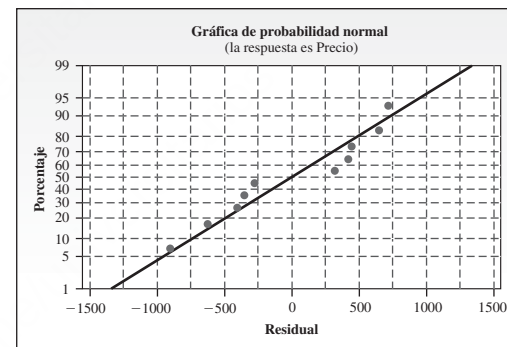
- ¿Qué suposiciones deben hacerse acerca de la distribución del error aleatorio, ε ?
- ¿Cuál es la mejor estimación de σ^2 , la varianza del error aleatorio ε ?
- Use las gráficas de diagnóstico para estos datos para comentar sobre la validez de las suposiciones de regresión.



categorías de 46 pulgadas y más grandes. ¿El precio de un televisor de LCD depende del tamaño de la pantalla?

Marca	Precio (\$)	Tamaño
Sony Bravia KDL-52NX800	2340	52
Samsung LN55C650	1600	55
Vizio VF550M	1330	55
Sony Bravia KDL-60EX700	2700	60
Sharp Aquos LED LC-52LE700UN	1620	52
Sony Bravia KDL-46XBR10	2500	46
Samsung UN46C8000	2200	46
Vizio SV472XVT	1400	47
Samsung UN46C7000	2100	46
LG 47LD450	900	47

- Suponga que intuimos que la relación entre el tamaño y el precio es lineal y realizamos una regresión lineal, que resulta en un valor de $r^2 = .027$. ¿Qué le dice el valor de r^2 acerca de la fuerza de la relación entre precio y tamaño de la pantalla?
- Las gráficas de diagnóstico para estos datos se muestran a continuación. ¿Parece que se han violado las suposiciones de normalidad o de varianza igual?



- Use una gráfica de dispersión para graficar precio contra tamaño de la pantalla para los 10 televisores de LCD. Con base en la información en la parte a, ¿cuál suposición para el modelo de regresión lineal ha sido violada?

DATOS
EX0838

8.38 Cómo elegir un televisor *Consumer Reports*⁶ dio los precios y tamaños de pantalla para los 10 televisores de LCD principales en las

ESTIMACIÓN Y PREDICCIÓN USANDO LA RECTA AJUSTADA

8.7

Ahora que usted ha

- probado y ajustado la recta de regresión, $\hat{y} = a + bx$, para asegurarse que es útil para predicción y
- empleado las herramientas de diagnóstico para asegurarse que ninguna de las suposiciones de regresión han sido violadas

está listo para usar la recta para uno de sus dos propósitos:

- Estimar el valor promedio de y para un valor determinado de x
- Predecir un valor particular de y para un valor determinado de x

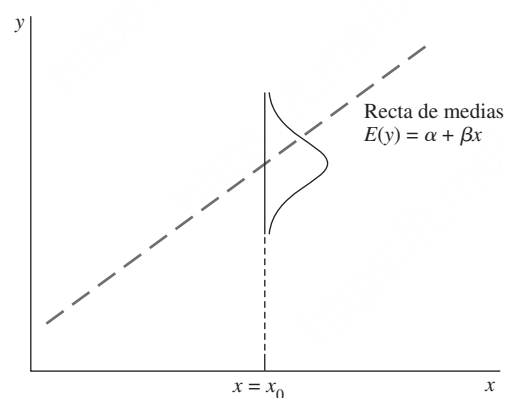
La muestra de n pares de observaciones han sido seleccionados de una población en la que el valor *promedio* de y está relacionado con el valor de la variable de pronóstico x por la **recta de medias**,

$$E(y) = \alpha + \beta x$$

una recta desconocida, que se muestra como recta interrumpida en la figura 8.11. Recuerde que para un valor fijo de x , por ejemplo x_0 , los valores *particulares* de y se desvían desde la recta de medias. Estos valores de y se supone que tienen una distribución normal con media igual a $\alpha + \beta x_0$ y varianza σ^2 , como se ve en la figura 8.11.

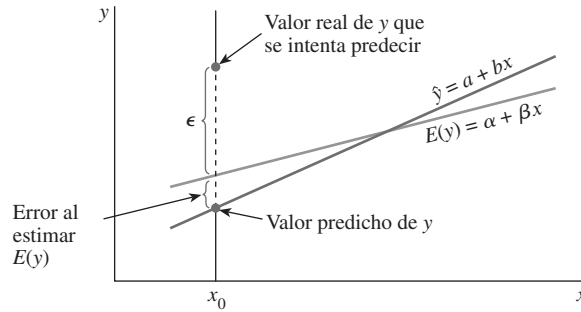
FIGURA 8.11

Distribución de y para
 $x = x_0$



Como los valores calculados de a y b varían de una muestra a otra, cada nueva muestra produce una diferente recta de regresión $\hat{y} = a + bx$, que se puede usar ya sea para estimar la recta de medias o para predecir un valor particular de y . La figura 8.12 muestra una de las posibles configuraciones de la recta ajustada (azul), la recta de medias desconocida (gris), y un valor particular de y (el punto azul).

FIGURA 8.12

Error al estimar $E(y)$ y predecir y 

¿A qué distancia estará nuestro estimador $\hat{y} = a + bx_0$ desde la cantidad a estimar o predecir? Esto depende, como siempre, de la variabilidad de nuestro estimador, medido por su **error estándar**. Se puede demostrar que

$$\hat{y} = a + bx_0$$

el valor estimado de y cuando $x = x_0$, es un estimador insesgado de la recta de medias, $\alpha + \beta x_0$ y que $\hat{y}$ está normalmente distribuida con el error estándar de $\hat{y}$ estimado por

$$SE(\hat{y}) = \sqrt{MSE \left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right)}$$

La estimación y prueba están basadas en el estadístico

$$t = \frac{\hat{y} - E(y)}{SE(\hat{y})}$$

que tiene una distribución t con $(n - 2)$ grados de libertad.

Para formar un intervalo de confianza $(1 - \alpha)100\%$ para el valor promedio de y cuando $x = x_0$, medido por la recta de medias, $\alpha + \beta x_0$, se puede usar la forma usual para un intervalo de confianza basado en la distribución t :

$$\hat{y} \pm t_{\alpha/2} SE(\hat{y})$$

No obstante, si se escoge predecir un valor *particular* de y cuando $x = x_0$, hay algún error adicional en la predicción debido a la desviación de y desde la recta de medias. Si examinamos la figura 8.12, se puede ver que el error en predicción tiene dos componentes:

- El error al usar la recta ajustada para estimar la recta de medias
- El error causado por la desviación de y desde la recta de medias, medida por σ^2

La varianza de la diferencia entre y y $\hat{y}$ es la suma de estas dos varianzas y forma la base para el error estándar de $(y - \hat{y})$ empleado para predicción:

$$SE(y - \hat{y}) = \sqrt{MSE \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]}$$

y el intervalo de predicción $(1 - \alpha)100\%$ se forma como

$$\hat{y} \pm t_{\alpha/2} SE(y - \hat{y})$$

TIP MI CONSEJO

Para un valor determinado de x , el intervalo de predicción es siempre más ancho que el intervalo de confianza

INTERVALOS DE CONFIANZA Y PREDICCIÓN $(1 - \alpha)100\%$

- Para estimar el valor promedio de y cuando $x = x_0$:

$$\hat{y} \pm t_{\alpha/2} \sqrt{\text{MSE} \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]}$$

- Para predecir un valor particular de y cuando $x = x_0$:

$$\hat{y} \pm t_{\alpha/2} \sqrt{\text{MSE} \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]}$$

donde $t_{\alpha/2}$ es el valor de t con $(n - 2)$ grados de libertad y área $\alpha/2$ a su derecha.

EJEMPLO

8.4

Use la información del ejemplo 8.1 para estimar el promedio de calificaciones en cálculo para estudiantes cuya puntuación de aprovechamiento es 50, con un intervalo de confianza de 95%.

Solución La estimación puntual de $E(y|x_0 = 50)$, el promedio de calificación en cálculo para estudiantes cuya puntuación de aprovechamiento es 50, es

$$\hat{y} = 40.78424 + .76556(50) = 79.06$$

El error estándar de $\hat{y}$ es

$$\sqrt{\text{MSE} \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]} = \sqrt{75.7532 \left[\frac{1}{10} + \frac{(50 - 46)^2}{2474} \right]} = 2.840$$

y el intervalo de confianza de 95% es

$$\begin{aligned} 79.06 \pm 2.306(2.840) \\ 79.06 \pm 6.55 \end{aligned}$$

Nuestros resultados indican que el promedio de calificación en cálculo para estudiantes cuya puntuación es de 50 en el examen de aprovechamiento estará entre 72.51 y 85.61.

EJEMPLO

8.5

Un estudiante tomó el examen de aprovechamiento y obtuvo 50 pero todavía no ha tomado el examen de cálculo. Usando la información del ejemplo 8.1, prediga la calificación en cálculo para este estudiante, con un intervalo de predicción de 95%.

Solución El valor predicho de y es $\hat{y} = 79.06$, como en el ejemplo 8.4. No obstante, el error en predicción se mide con $SE(y - \hat{y})$, y el intervalo de predicción de 95% es

$$\begin{aligned} 79.06 \pm 2.306 \sqrt{75.7532 \left[1 + \frac{1}{10} + \frac{(50 - 46)^2}{2474} \right]} \\ 79.06 \pm 2.306(9.155) \\ 79.06 \pm 21.11 \end{aligned}$$

o sea de 57.95 a 100.17. El intervalo de predicción es *más ancho* que el intervalo de confianza del ejemplo 8.4 por la variabilidad extra al predecir el valor real de la respuesta y .

Es frecuente que un punto particular sobre la recta de medias sea de interés para experimentadores, α , la intersección con el eje y , es decir, el valor promedio de y cuando $x_0 = 0$.

EJEMPLO**8.6**

Antes de ajustar una recta a los datos de la calificación en cálculo y las puntuaciones de aprovechamiento, se puede pensar que una puntuación de 0 en el examen de aprovechamiento puede predecir una calificación de 0 en el examen de cálculo. Esto implica que debemos ajustar un modelo con α igual a 0. ¿Los datos apoyan la hipótesis de un punto de intersección en 0?

Solución Se puede contestar esta pregunta al construir un intervalo de confianza de 95% para el punto α de intersección con el eje y , que es el valor promedio de y cuando $x = 0$. La estimación de α es

$$\hat{y} = 40.784 + .76556(0) = 40.784 = a$$

y el intervalo de confianza de 95% es

$$\begin{aligned} \hat{y} \pm t_{\alpha/2} \sqrt{\text{MSE} \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]} \\ 40.784 \pm 2.306 \sqrt{75.7532 \left[\frac{1}{10} + \frac{(0 - 46)^2}{2474} \right]} \\ 40.784 \pm 19.617 \end{aligned}$$

o de 21.167 a 60.401, un intervalo que no contiene el valor $\alpha = 0$. En consecuencia, es improbable que la intersección con el eje y sea 0. Se debe incluir un punto de intersección diferente de cero en el modelo $y = \alpha + \beta x + \epsilon$.

Para esta situación especial en la que estamos interesados en probar o estimar el punto de intersección α con el eje y para la recta de medias, las inferencias comprenden la estimación muestral a . La prueba para un punto de intersección en 0 se da en la figura 8.13 en la recta sombreada marcada como “Constante”. El coeficiente dado como 40.784 es a , con error estándar dado en la columna marcada “SE Coef” como 8.507, que concuerda con el valor calculado en el ejemplo 8.6. El valor de $t = 4.79$ se encuentra al dividir a entre su error estándar con valor $p = .001$.

FIGURA 8.13

Parte de la salida *MINITAB* para el ejemplo 8.6

Predictor	Coef	SE Coef	T	P
Constant	40.784	8.507	4.79	0.001
x	0.7656	0.1750	4.38	0.002

Se puede ver que es bastante lento calcular manualmente estos intervalos de predicción y estimación, además de que es difícil mantener precisión en los cálculos, pero por fortuna hay programas de cómputo que pueden hacer los cálculos. El comando de regresión *MINITAB* tiene una opción para estimación o predicción cuando se especifica el valor (o valores) necesarios de x . La salida impresa de la figura 8.14 da los valores de $\hat{y} = 79.06$ marcados “Fit”, el error estándar de $\hat{y}$, $SE(\hat{y})$, marcado “SE Fit”, el *intervalo de confianza* para el valor promedio de y cuando $x = 50$, marcado “95.0% CI”, y el intervalo de predicción para y cuando $x = 50$, marcado “95.0% PI”.

FIGURA 8.14

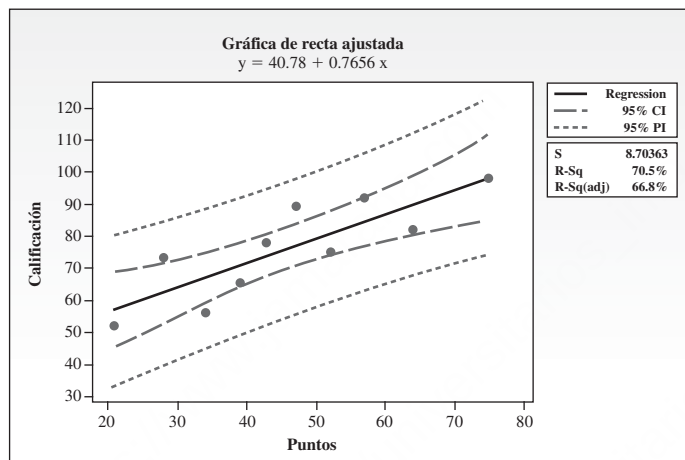
Opción *MINITAB* para estimación y predicción

Predicted Values for New Observations				
New Obs	Fit	SE Fit	95% CI	95% PI
1	79.06	2.84	(72.51, 85.61)	(57.95, 100.17)
Values of Predictors for New Observations				
New Obs	x			
1	50.0			

Las bandas de confianza y bandas de predicción generadas por *MINITAB* para los datos de calificaciones en cálculo se muestran en la figura 8.15. Observe que en general las bandas de confianza son más angostas que las bandas de predicción para todo valor de puntos x de examen de aprovechamiento. Es seguro que usted esperaría predicciones de que un valor individual fuera mucho más variable que las estimaciones del valor promedio. También observe que las bandas parecen hacerse más anchas a medida que el valor de x_0 se aleja de la media $\bar{x}$. Esto es porque los errores estándar empleados en los intervalos de confianza y predicción contienen el término $(x_0 - \bar{x})^2$, que se hace más grande cuando los dos valores divergen. En la práctica, esto significa que estimación y predicción son más precisos cuando x_0 está cerca del centro del rango de los valores de x . Se pueden localizar los intervalos de confianza y predicción calculados cuando $x = 50$ en la figura 8.15.

FIGURA 8.15

Intervalos de confianza y predicción para los datos de la tabla 8.1



8.7 EJERCICIOS

TÉCNICAS BÁSICAS

8.39 Consulte el ejercicio 8.6.

- Estime el valor promedio de y cuando $x = 1$, usando un intervalo de confianza de 90%.
- Encuentre un intervalo de predicción de 90% para algún valor de y a ser observado en el futuro cuando $x = 1$.

8.40 Consulte el ejercicio 8.7. Partes de la salida impresa *MINITAB* se muestran a continuación.

Salida *MINITAB* para el ejercicio 8.40

Análisis de regresión: y contra x

The regression equation is
 $y = 6.00 - 0.557 x$

Predictor	Coef	SE Coef	T	P
Constant	6.0000	0.1759	34.10	0.000
x	-0.55714	0.04518	-12.33	0.000

Predicted Values for New Observations

New Obs	Fit	SE Fit	95.0% CI	95.0% PI
1	4.8857	0.1027	(4.6006, 5.1708)	(4.2886, 5.4829)
2	1.5429	0.2174	(0.9392, 2.1466)	(0.7430, 2.3427)

X denotes a point that is an outlier in the predictors.

Values of Predictors for New Observations

New Obs	x
1	2.00
2	8.00

- Encuentre un intervalo de confianza de 95% para el valor promedio de y cuando $x = 2$.
- Encuentre un intervalo de predicción de 95% para algún valor de y a ser observado en el futuro cuando $x = 2$.
- El último renglón de la tercera sección de la salida impresa indica un problema con uno de los valores ajustados. ¿Qué valor de x corresponde al valor ajustado $\hat{y} = 1.5429$? ¿Qué problema ha detectado el programa *MINITAB*?

APLICACIONES

DATOS
EX0841

8.41 ¿Qué comprar? Se realizó un experimento de investigación de mercado, para estudiar la relación entre el tiempo necesario para que un comprador llegue a una decisión y el número de diseños de empaque alternativos presentados de un producto. Se eliminaron nombres de marca en los empaques y los compradores hicieron su selección usando las descripciones del producto, hechas por los fabricantes en los empaques, como únicas guías para comprar. El tiempo

necesario para tomar una decisión se registró para 15 participantes en el estudio de investigación de mercado.

Tiempo para decisión, y segundos	5, 8, 8, 7, 9	7, 9, 8, 9, 10	10, 11, 10, 12, 9
Número de alternativas, x	2	3	4

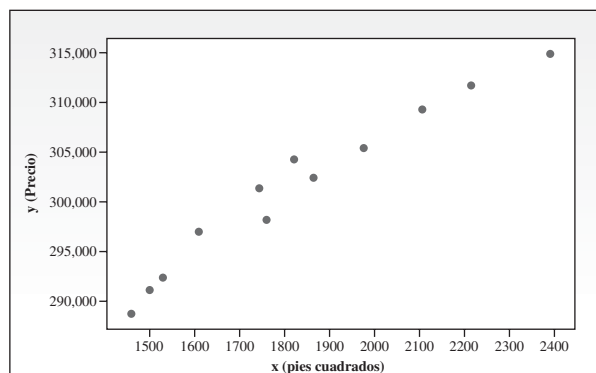
- Encuentre la recta de mínimos cuadrados apropiada para estos datos.
- Grafique los puntos y grafique la recta como prueba en sus cálculos.
- Calcule s^2 .
- ¿Los datos dan suficiente evidencia para indicar que el tiempo para tomar una decisión está linealmente relacionado con el número de diseños alternativos de empaques? (Pruebe al nivel de significancia de $\alpha = .05$.)
- Encuentre el valor p aproximado para la prueba e interprete su valor.
- Si están disponibles, examine las gráficas de diagnóstico para verificar la validez de las suposiciones de regresión.
- Estime el promedio de tiempo necesario para llegar a una decisión cuando se presenten tres alternativas, usando un intervalo de confianza de 95%.

DATOS
EX0842

8.42 Precios de vivienda Los datos de la tabla siguiente dan la superficie en pies cuadrados y los precios de venta de $n = 12$ casas seleccionadas al azar de las ventas en una ciudad pequeña. Use la salida impresa MINITAB para contestar las preguntas.

Pies cuadrados, x	Precio, y	Pies cuadrados, x	Precio, y
1460	\$288 700	1977	\$305 400
2108	309 300	1610	297 000
1743	301 400	1530	292 400
1499	291 100	1759	298 200
1864	302 400	1821	304 300
2391	314 900	2216	311 700

Gráfica de datos para el ejercicio 8.42



Salida MINITAB para el Ejercicio 8.42

Análisis de regresión: y contra x

The regression equation is
 $y = 251206 + 27.4 x$

Predictor	Coef	SE Coef	T	P
Constant	251206	3389	74.13	0.000
x	27.406	1.828	14.99	0.000

S = 1792.72 R-Sq = 95.7% R-Sq(adj) = 95.3%

Predicted Values for New Observations

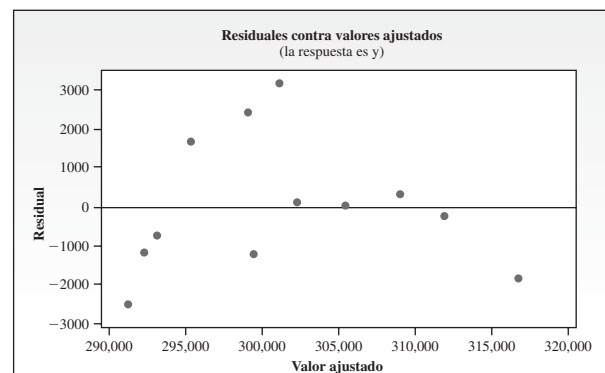
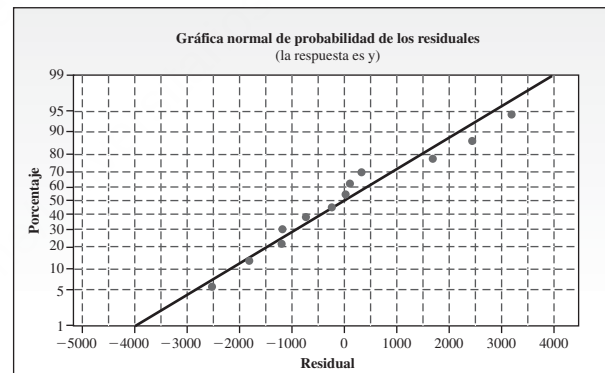
New Obs	Fit	SE Fit	95.0% CI	95.0% PI
1	299989	526	(298817, 301161)	(295826, 304151)
2	306018	602	(304676, 307360)	(301804, 310232)

Values of Predictors for New Observations

New Obs	x
1	1780
2	2000

- ¿Se puede ver algún patrón que no sea una relación lineal en la gráfica original?
- El valor de r^2 para estos datos es .957. ¿Qué nos dice esto acerca del ajuste de la recta de regresión?
- Vea las siguientes gráficas de diagnóstico para estos datos. ¿Se ve algún patrón en los residuales? ¿Esto sugiere que la relación entre precio y pies cuadrados es algo que no sea lineal?

Gráficas de diagnóstico para el ejercicio 8.42



8.43 Precios de vivienda II Consulte el ejercicio 8.42 y el conjunto de datos EX0842.

- Estime el promedio de incremento en el precio para un aumento de 1 pie cuadrado para casas vendidas en la ciudad. Use un intervalo de confianza de 99%. Interprete su estimación.
- Un vendedor de bienes raíces necesita estimar el promedio de precio de venta de casas con un total de 2000 pies cuadrados de espacio con calefacción. Use un intervalo de confianza de 95% e interprete su estimación.
- Calcule el precio por pie cuadrado para cada casa y luego calcule la media muestral. ¿Por qué esta estimación del promedio de costo por pie cuadrado no es igual a la respuesta de la parte a? ¿Debe ser igual? Explique.
- Suponga que una casa con 1780 pies cuadrados de espacio con calefacción se ofrece a la venta. Construya un intervalo de predicción de 95% para el precio al cual se venderá la casa.

8.44 Fresas III Los datos siguientes (ejercicios 8.18 y 8.27) se obtuvieron en un experimento que relacionaba la variable dependiente, y (textura de fresas), con x (temperatura de almacenamiento codificada).

x	-2	-2	0	2	2
y	4.0	3.5	2.0	0.5	0.0

- Estime la textura esperada de fresas para una temperatura de almacenamiento codificada de $x = -1$. Use un intervalo de confianza de 99%.
- Prediga el valor particular de y cuando $x = 1$ con un intervalo de predicción de 99%.
- ¿A qué valor de x permanecerá mínimo el ancho del intervalo de predicción para un valor particular de y , suponiendo que n permanece fija?

DATOS
EX0845
8.45 Drew Brees El número de pases completos y el número total de yardas ganadas por pase para Drew Brees, mariscal de campo de los Santos de Nueva Orleans, se registraron en los 16 juegos regulares de la temporada de fútbol de 2010.⁷ En la semana 10 no se jugó y no hay datos reportados.

Semana	Pases completos	Yardas totales
1	27	237
2	28	254
3	30	365
4	33	275
5	24	279
6	21	263
7	37	356
8	34	305
9	27	253
10	—	—
11	29	382
12	23	352
13	24	313
14	25	221
15	29	267
16	25	302
17	22	196

- ¿Cuál es la recta de mínimos cuadrados que relaciona el total de yardas ganadas por pase con el número de pases completos para Drew Brees?
- ¿Qué proporción de la variación total está explicada por la regresión del total de yardas ganadas por pase (y) en el número de pases completos (x)?
- Si las hay, examine las gráficas de diagnóstico para comprobar la validez de las suposiciones de regresión.

8.46 Drew Brees, continúa Consulte el ejercicio 8.45.

- Estime el número promedio de yardas ganadas por pase para juegos en los que Brees lanza 20 pases completos, usando un intervalo de confianza de 95%.
- Prediga el número real de yardas ganadas por pase para juegos en los que Brees lanza 20 pases completos usando un intervalo de confianza de 95%.
- ¿Sería aconsejable usar la recta de mínimos cuadrados del ejercicio 8.45 para predecir el número total de yardas ganadas por pase de Brees, para un juego en el que lanzó sólo 5 pases completos? Explique.

8.8

ANÁLISIS DE CORRELACIÓN

El coeficiente de correlación, r , formalmente denominado **coeficiente de correlación muestral producto-momento de Pearson**, se define a continuación.

COEFICIENTE DE CORRELACIÓN PRODUCTO-MOMENTO DE PEARSON

$$r = \frac{s_{xy}}{s_x s_y} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} \quad \text{para } -1 \leq r \leq 1$$

TIP MI CONSEJO

r está siempre entre -1 y $+1$

Las varianzas y covarianza se pueden hallar por cálculo directo, con el uso de una calculadora con capacidad de estadística de dos variables o usando un paquete de estadística como el MINITAB o MS Excel. Las varianzas y covarianza se calculan como

$$s_{xy} = \frac{S_{xy}}{n - 1} \quad s_x^2 = \frac{S_{xx}}{n - 1} \quad s_y^2 = \frac{S_{yy}}{n - 1}$$

y use S_{xy} , S_{xx} y S_{yy} , las mismas cantidades empleadas antes en el análisis de regresión en este capítulo. En general, cuando una muestra de n individuos o unidades experimentales se selecciona y dos variables se miden en cada individuo o unidad, de modo que *ambas variables sean aleatorias*, el coeficiente de correlación r es la medida apropiada de linealidad para usar en esta situación.

EJEMPLO

8.7

Las estaturas y los pesos de $n = 10$ jugadores atacantes de fútbol se seleccionan al azar de un equipo de estrellas de un condado. Calcule el coeficiente de correlación para las estaturas (en pulgadas) y pesos (en libras) dado en la tabla 8.4.

TABLA 8.4

Estaturas y pesos de $n = 10$ estrellas profundos

Jugador	Estatura, x	Peso, y
1	73	185
2	71	175
3	75	200
4	72	210
5	72	190
6	75	195
7	67	150
8	69	170
9	71	180
10	69	175

Solución Debe usarse el método apropiado de entrada de datos de su calculadora científica para verificar los cálculos para las sumas de cuadrados y productos cruz,

$$S_{xy} = 328 \quad S_{xx} = 60.4 \quad S_{yy} = 2610$$

usando las fórmulas de cálculo dadas antes en este capítulo. Entonces

$$r = \frac{328}{\sqrt{(60.4)(2610)}} = .8261$$

o sea $r = .83$. Este valor de r es bastante cercano a 1, el máximo valor posible de r , que indica una relación lineal positiva bastante fuerte entre estatura y peso.

TIP MI CONSEJO

El signo de r es siempre igual al signo de la pendiente b

Hay una relación directa entre las fórmulas de cálculo para el coeficiente de correlación r y la pendiente de la recta de regresión b . Como el numerador de ambas cantidades es S_{xy} , tanto r como b tienen el mismo signo. Por tanto, el coeficiente de correlación tiene estas propiedades generales:

- Cuando $r = 0$, la pendiente es $b = 0$ y no hay relación lineal entre x y y .
- Cuando r es positiva, b también es positiva y hay una relación lineal positiva entre x y y .

- Cuando r es negativa, b también es negativa y hay una relación lineal negativa entre x y y .

En la sección 8.5 demostramos que

$$r^2 = \frac{SSR}{TSS} = \frac{TSS - SSE}{TSS}$$

En esta forma se puede ver que r^2 nunca puede ser mayor que 1, de modo que $-1 \leq r \leq 1$. Además, se observa la relación entre la variación aleatoria (medida por SSE) y r^2 .



APPLET EN LÍNEA

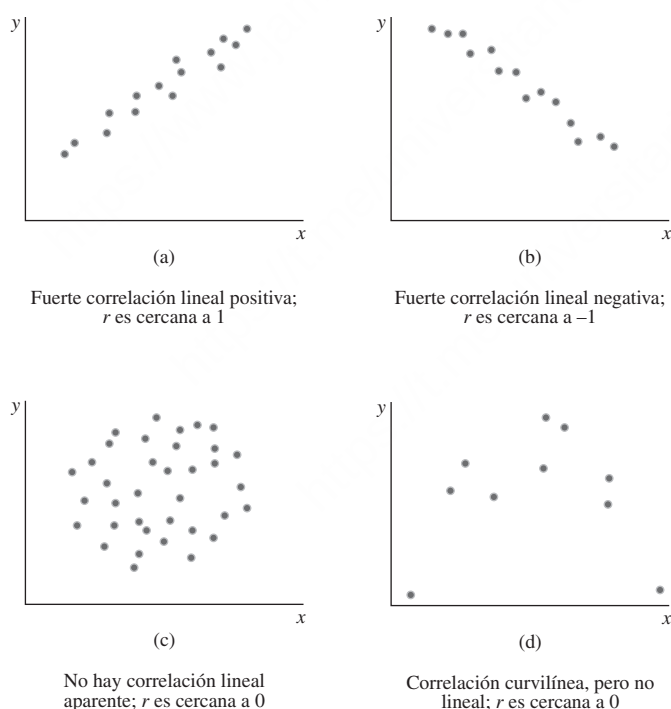
Explorar la correlación

- Si no hay variación aleatoria y todos los puntos caen en la recta de regresión, entonces $SSE = 0$ y $r^2 = 1$.
- Si los puntos están dispersos en forma aleatoria y no hay variación explicada por regresión, entonces $SSR = 0$ y $r^2 = 0$.

La figura 8.16 muestra cuatro gráficas de dispersión típicas y sus coeficientes de correlación asociados. Observe que en la gráfica de dispersión (d) parece haber una relación curvilínea entre x y y , pero r es aproximadamente 0, lo cual refuerza el hecho de que r es una medida de una relación *lineal* (no *curvilínea*) entre dos variables.

FIGURA 8.16

Algunas gráficas de dispersión típicas con valores aproximados de r



Considere una población generada al medir dos variables aleatorias en cada unidad experimental. En esta población *bivariada*, el **coeficiente de correlación poblacional** ρ se calcula e interpreta como está en la muestra. En esta situación, el experimentador puede probar la hipótesis de que no hay correlación entre las variables x y y usando una estadística de prueba que

sea *exactamente equivalente* a la prueba de la pendiente β de la sección 8.5. El procedimiento de prueba se muestra a continuación.

PRUEBA DE HIPÓTESIS RESPECTO AL COEFICIENTE DE CORRELACIÓN ρ

1. Hipótesis nula: $H_0: \rho = 0$
2. Hipótesis alternativa:

Prueba de una cola

$$H_a: \rho > 0 \\ (\text{o } \rho < 0)$$

Prueba de dos colas

$$H_a: \rho \neq 0$$

3. Estadístico de prueba: $t = r \sqrt{\frac{n-2}{1-r^2}}$

Cuando se satisfacen las suposiciones dadas en la sección 8.2, el estadístico de prueba tendrá una distribución t de Student con $(n-2)$ grados de libertad.

4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola

$$t > t_\alpha \\ (\text{o } t < -t_\alpha \text{ cuando la hipótesis} \\ \text{alternativa sea } H_a: \rho < 0)$$

Prueba de dos colas

$$t > t_{\alpha/2} \text{ o } t < -t_{\alpha/2}$$

o cuando valor $p < \alpha$

Los valores de t_α y $t_{\alpha/2}$ correspondientes a $(n-2)$ grados de libertad pueden encontrarse usando la tabla 4 en el Anexo.

TIP MI CONSEJO

Se puede demostrar que

$$t = r \sqrt{\frac{n-2}{1-r^2}} \\ = \frac{b-0}{\sqrt{\text{MSE}/S_{xx}}}$$

EJEMPLO

8.8

Consulte los datos de estatura y peso del ejemplo 8.7. La correlación entre estatura y peso se calculó como $r = .8261$. ¿Esta correlación es significativamente diferente de 0?

Solución Para probar las hipótesis

$$H_0: \rho = 0 \text{ contra } H_a: \rho \neq 0$$

el valor del estadístico de prueba es

$$t = r \sqrt{\frac{n-2}{1-r^2}} = .8261 \sqrt{\frac{10-2}{1-(.8261)^2}} = 4.15$$

que para $n = 10$ tiene una distribución t con 8 grados de libertad. Como este valor es mayor que $t_{.005} = 3.355$, el valor p de dos colas es menor que $2(.005) = .01$ y la correlación se declara significativa al nivel de 1% ($P < .01$). El valor $r^2 = .8261^2 = .6824$ significa que alrededor de 68% de la variación en una de las variables es explicada por la otra. La salida impresa MINITAB de la figura 8.17 muestra la correlación r y el valor p exacto para probar su significancia.

TIP MI CONSEJO

El valor t y valor p para probar $H_0: \rho = 0$ será idéntico al valor t y valor p para probar $H_0: \beta = 0$

FIGURA 8.17

Salida MINITAB para el ejemplo 8.8

Correlaciones: x, y

Pearson correlation of x and y = 0.826
P-Value = 0.003

Si los coeficientes lineales de correlación entre y y cada una de las dos variables x_1 y x_2 se calculan como .4 y .5, respectivamente, no se concluye que un medio de predicción que use ambas variables sea $[(.4)^2 + (.5)^2] = .41$, o sea una reducción de 41% en la suma de cuadrados de desviaciones. En realidad, x_1 y x_2 podrían estar bastante correlacionados y por tanto contribuir prácticamente con la misma información para la predicción de y .

Por último, recuerde que r es una medida de **correlación lineal** y que x y y podrían estar perfectamente relacionadas por alguna función **no lineal** cuando el valor observado de r sea igual a 0. El problema de estimar o predecir y usando información dada por varias variables independientes, $x_1, x_2, \dots, x_k$, es el tema del capítulo 9.

8.8 EJERCICIOS

TÉCNICAS BÁSICAS

8.47 ¿En qué forma el coeficiente de correlación mide la fuerza de la relación lineal entre dos variables y y x ?

8.48 Describa la significancia del signo algebraico y la magnitud de r .

8.49 ¿Qué valor toma r si todos los puntos de datos caen en la misma recta en estos casos?

- La recta tiene pendiente positiva.
- La recta tiene pendiente negativa.

8.50 Nos dan estos datos:

x	-2	-1	0	1	2
y	2	2	3	4	4

- Grafique los puntos. Con base en su gráfica, ¿cuál será el signo del coeficiente de correlación muestral?
- Calcule r y r^2 e interprete sus valores.

8.51 Nos dan estos datos:

x	1	2	3	4	5	6
y	7	5	5	3	2	0

- Grafique los seis puntos en papel para graficar.
- Calcule el coeficiente muestral de correlación r e interprete.
- ¿En qué porcentaje se redujo la suma de cuadrados de desviaciones al usar el pronosticador de mínimos cuadrados $\hat{y} = a + bx$, en lugar de $\bar{y}$ como pronosticador de y ?

8.52 Invierta la pendiente de la recta del ejercicio 8.51 al reordenar las observaciones y , como sigue:

x	1	2	3	4	5	6
y	0	2	3	5	5	7

Repita los pasos del ejercicio 8.51. Observe el cambio en el signo de r y la relación entre los valores de r^2 del ejercicio 8.51 y este ejercicio.

APLICACIONES

DATOS EX0853 **8.53 Langostas** La tabla siguiente da los números de lapas *Octolasmis tridens* y *O. lowei* en cada una de 10 langostas.⁸ ¿Le parece que las lapas compiten por espacio en la superficie de una langosta?

Número de campo de langosta	<i>O. tridens</i>	<i>O. lowei</i>
A061	645	6
A062	320	23
A066	401	40
A070	364	9
A067	327	24
A069	73	5
A064	20	86
A068	221	0
A065	3	109
A063	5	350

- Si compiten, ¿se espera que el número x de lapas *O. tridens* y el número y de *O. lowei* estén correlacionados positiva o negativamente? Explique.
- Si desea probar la teoría de que los dos tipos de lapas compiten por espacio al realizar una prueba de la hipótesis nula “el coeficiente de correlación poblacional r es igual a 0”, ¿cuál es su hipótesis alternativa?
- Realice la prueba de la parte **b** y exprese sus conclusiones.

DATOS EX0854 **8.54 Capacitación de habilidades sociales** Se puso en práctica un programa de capacitación de habilidades sociales con siete estudiantes levemente discapacitados en un estudio para determinar si el programa causó mejora en medidas anteriores y pos-

teriores al estudio y en calificaciones de conducta. Para una de estas pruebas, en la tabla siguiente se dan las puntuaciones de mediciones previas y posteriores para los siete estudiantes.⁹

Sujeto	Anterior	Posterior
Earl	101	113
Ned	89	89
Jasper	112	121
Charlie	105	99
Tom	90	104
Susie	91	94
Lori	89	99

a. ¿Qué tipo de correlación, si la hay, espera ver entre puntuaciones de examen anteriores y posteriores? Grafique los datos. ¿La correlación parece ser positiva o negativa?

b. Calcule el coeficiente de correlación, r . ¿Hay una correlación positiva significativa?

8.55 Hockey G. W. Marino investigó las variables relacionadas con la capacidad de un jugador de hockey para hacer un rápido arranque desde una posición de reposo.¹⁰ En el experimento, cada patinador arrancó desde una posición de reposo y trató de moverse tan rápidamente como le fuera posible en una distancia de 6 metros. El coeficiente de correlación r entre la rapidez de zancada de un patinador (número de zancadas por segundo) y el tiempo para recorrer la distancia de 6 metros para la muestra de 69 patinadores, fue $-.37$.

a. ¿Los datos dan suficiente evidencia para indicar una correlación entre rapidez de zancadas y tiempo para recorrer la distancia? Pruebe usando $\alpha = .05$.

b. Encuentre el valor p aproximado para la prueba.

c. ¿Cuáles son las implicaciones prácticas de la prueba en la parte a?

8.56 Hockey II Consulte el ejercicio 8.55. Marino calculó que el coeficiente de correlación muestral r , para la rapidez de zancadas y el promedio de rapidez de aceleración para los 69 patinadores, era de $.36$. ¿Los datos dan suficiente evidencia para indicar una correlación entre rapidez de zancadas y promedio de aceleración para los patinadores? Use el método del valor p .

DATOS
EX0857 **8.57 Energía geotérmica** La energía geotérmica es una importante fuente de energía. Como la cantidad de energía contenida en 1 libra de agua es función de su temperatura, uno podría preguntarse si el agua obtenida de pozos más profundos contiene más energía por libra. Los datos de la tabla siguiente se

reproducen de un artículo sobre sistemas geotérmicos escrito por A. J. Ellis.¹¹

Ubicación del pozo	Profundidad promedio (máx) de perforación (m)	Temperatura promedio (máx) (°C)
El Tateo, Chile	650	230
Ahuachapan, El Salvador	1000	230
Namafjall, Islandia	1000	250
Larderello (región), Italia	600	200
Matsukawa, Japón	1000	220
Cerro Prieto, México	800	300
Wairakei, Nueva Zelanda	800	230
Kizildere, Turquía	700	190
The Geysers, Estados Unidos	1500	250

¿Hay una correlación positiva significativa entre la profundidad promedio (máx) de perforación y la temperatura promedio (máx)?

DATOS
EX0858 **8.58 ¿Alguien quiere helado?** Por mucho que los estadounidenses traten de evitar los alimentos altos en grasas y altos en calorías, la demanda para un cono de helado frío y cremoso en un día caluroso es difícil de resistir. La popular franquicia de helados *Coldstone Creamery* publicó en su sitio web la información nutricional para sus ofertas de helados en tres tamaños de porciones: “Like it”, “Love it” y “Gotta Have it”.¹² Un fragmento de esa información para el tamaño de porción “Like it” se muestra en la tabla.

Sabor	Calorías	Grasa total (gramos)
Cake Batter	340	19
Cinnamon Bun	370	21
French Toast	330	19
Mocha	320	20
OREO® Crème	440	31
Peanut Butter	370	24
Strawberry Cheesecake	320	21

a. ¿Debería usar los métodos de análisis de regresión lineal o de análisis de correlación para analizar los datos? Explique.

b. Analice los datos para determinar la naturaleza de la relación entre la grasa total y las calorías en el helado *Coldstone Creamery*.

DATOS
EX0859 **8.59 Temperatura corporal y frecuencia cardiaca** ¿Hay alguna relación entre estas dos variables? Para averiguarlo, seleccionamos 12 personas de manera aleatoria de un conjunto de datos construido por Allen Shoemaker (*Journal of Statistics Education*) y registramos sus temperaturas corporales y frecuencias cardíacas.¹³

Persona	1	2	3	4	5	6
Temperatura (grados)	96.3	97.4	98.9	99.0	99.0	96.8
Frecuencia cardiaca (pulsaciones por minuto)	70	68	80	75	79	75

Persona	7	8	9	10	11	12
Temperatura (grados)	98.4	98.4	98.8	98.8	99.2	99.3
Frecuencia cardiaca (pulsaciones por minuto)	74	84	73	84	66	68

- a. Encuentre el coeficiente de correlación r , que relacione temperatura corporal y frecuencia cardiaca.
- b. ¿Hay suficiente evidencia para indicar que hay una correlación entre estas dos variables? Pruebe al nivel de significancia de 5%.

DATOS
EX0860

8.60 Estadísticas de beisbol ¿El promedio de bateo de un equipo depende en alguna forma del número de cuadrangulares conectados por el equipo? Los datos de la tabla siguiente muestran el número de cuadrangulares del equipo y el promedio general de

bateo del equipo para ocho equipos de ligas mayores seleccionados para la temporada de 2010.¹⁴

Equipo	Cuadrangulares totales	Promedio de bateo del equipo
Atlanta Braves	139	.258
Baltimore Orioles	133	.259
Boston Red Sox	211	.268
Chicago White Sox	177	.268
Houston Astros	108	.247
LA Dodgers	120	.252
Philadelphia Phillies	166	.260
Seattle Mariners	101	.236

Fuente: ESPN.com

- a. Grafique los puntos usando una gráfica de dispersión. ¿Le parece que hay alguna relación entre el total de cuadrangulares y el promedio de bateo del equipo?
- b. ¿Hay alguna correlación positiva significativa entre el total de cuadrangulares y el promedio de bateo del equipo? Pruebe al nivel de significancia de 5%.
- c. ¿Piensa usted que la relación entre estas dos variables sería diferente si hubiéramos visto todo el conjunto de franquicias de las ligas mayores?

8.9

COVARIANZA

La covarianza es una medida que se utiliza para describir la forma en que dos variables aleatorias, x y y , varían juntas; medida de asociación entre dichas variables que se da cuando existe una relación lineal entre ellas.

La covarianza se define con la misma expresión, tanto para variables aleatorias discretas como continuas.

$$E(xy) - \mu_x \mu_y$$

donde μ_x y μ_y son los valores esperados de x y y respectivamente, cumpliendo los siguientes supuestos:

- 1) Si las variables aleatorias x y y son independientes (cuando la relación entre ellas es no lineal), la covarianza es cero.
- 2) Si cuando una variable aumenta la otra también aumenta, la covarianza es positiva.
- 3) Si cuando una variable aumenta la otra disminuye, la covarianza es negativa.

Como la covarianza es un “momento producto” o “momento mezclado” que describe la variación de un par de variables aleatorias x , y , la ecuación 1 puede tomar la forma:

$$\text{Cov}(x, y) = E[(x - \mu_x)(y - \mu_y)]$$

La medida que refleja la fuerza y dirección de la relación lineal entre ambas variables, es el coeficiente de correlación, que en términos de la covarianza, y para la población es:

$$\rho = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}$$

donde: σ_x y σ_y son las desviaciones estándar de μ_x y μ_y respectivamente y provienen de las varianzas

$$\sigma_x^2 = E(x) \text{ y } \sigma_y^2 = E(y), \text{ para el caso muestral } r_{xy} = \frac{s_{xy}}{s_x s_y}$$

donde s_{xy} (la covarianza) está definida como:

$$s_{xy} = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y}) \text{ o también}$$

$$s_{xy} = s_x s_y r_{xy}$$

En la función de mínimos cuadrados de una regresión lineal $\hat{\alpha} + \hat{\beta}_x$, despejando $\hat{\alpha}$ y $\hat{\beta}$ se tiene:

$$\beta = \rho \frac{\sigma_y^\dagger}{\sigma_x} \quad \hat{\beta} = \frac{s_{xy}}{s_x^2} = r_{xy} \frac{s_y}{s_x}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}_x$$

El cuadrado medio del error:

$$\hat{\sigma}^2 = s_y^2 (1 - r_{xy}^2)$$

EJEMPLO

8.9

Un especialista en vías respiratorias está interesado en la relación entre x (número de cigarrillos que un paciente fuma diariamente) y y (la ocurrencia de cáncer pulmonar). x , y son variables aleatorias, con probabilidad $P(x)$, $P(y)$. Se considera que $x = 1$ si el paciente es fumador y $x = 0$ si no lo es; $y = 1$ si presenta cáncer pulmonar y $y = 0$ cuando no lo tiene.

Suponga que la probabilidad de que sea fumador y presente cáncer es 0.50, esto es: $P(x = 1, y = 1) = 0.50$; la probabilidad de no ser fumador y no tener cáncer es 0.35 $P(x = 0, y = 0) = 0.35$; que sea fumador y no presentar cáncer es 0.10 $P(x = 1, y = 0) = 0.10$ y no fumador pero tenga cáncer $P(x = 0, y = 1) = 0.05$. Esto se puede mostrar en la siguiente tabla de contingencias:

[†] Entonces $H_0: \beta = 0$ es equivalente a $\rho = 0$.

		y		
		0	1	
x	0	0.35	0.05	0.40
	1	0.10	0.50	0.60
		0.45	0.55	

Donde 0.35, 0.05, 0.10 y 0.50 son conocidas como probabilidades conjuntas y 0.40, 0.60, 0.45 y 0.55 son probabilidades marginales.

Las probabilidades condicionales se obtienen de la siguiente manera:

$$P(x=1 | y=1) = \frac{P(x=1, y=1)}{P(y=1)} = \frac{0.50}{0.55} = 0.91$$

$$P(x=0 | y=0) = \frac{P(x=0, y=0)}{P(y=0)} = \frac{0.35}{0.45} = 0.78$$

$$P(x=0 | y=1) = \frac{P(x=0, y=1)}{P(y=1)} = \frac{0.05}{0.55} = 0.09$$

$$P(x=1 | y=0) = \frac{P(x=1, y=0)}{P(y=0)} = \frac{0.10}{0.45} = 0.22$$

Y como la covarianza refleja la dirección de la relación, se tiene:

$$\begin{aligned} E(xy) &= 0(0)(0.35) + 0(1)(0.05) + 1(0)(0.10) + 1(1)(0.50) = 0.50 \\ &= 0 + 0 + 0 + 0.50 \end{aligned}$$

$E(xy) = 0.50$

Utilizando las probabilidades marginales obtenidas en la tabla anterior, se tiene:

$$E(x) = 0.60$$

$$E(y) = 0.55$$

Sustituyendo estos valores en la definición de la covarianza (2).

$$\text{Cov}(xy) = E(xy) - E(x)E(y)$$

$$\text{Cov}(xy) = 0.50 - (0.60)(0.55) = 0.50 - 0.33 = 0.17$$

Entonces, el valor de la covarianza es

$$\text{Cov}(x, y) = 0.17$$

y como las varianzas σ_x^2 y σ_y^2 se definen como el producto de las probabilidades marginales de x y y , entonces $\sigma_x^2 = E(x) = E[(x=0, x=1)] = 0.40 \times 0.60 = 0.24$

$$\sigma_y^2 = E(y) = E[(y=0, y=1)] = 0.45 \times 0.55 = 0.2475$$

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{0.24} = 0.49$$

$$\sigma_y = \sqrt{\sigma_y^2} = \sqrt{0.2475} = 0.50$$

Sustituyendo ambos resultados en (3)

$$\rho = \frac{0.17}{(0.49)(0.50)} = \frac{0.17}{0.24} = 0.71$$

El coeficiente de correlación es

$$\rho = 0.71$$

REPASO DEL CAPÍTULO

Conceptos y fórmulas clave

I. Un modelo probabilístico lineal

1. Cuando los datos exhiben una relación lineal, el modelo apropiado es $y = \alpha + \beta x + \epsilon$.
2. El error aleatorio ϵ tiene una distribución normal con media de 0 y varianza σ^2 .

II. Método de mínimos cuadrados

1. Las estimaciones a y b , para α y β , se escogen para reducir SSE al mínimo, la suma de desviaciones al cuadrado alrededor de la recta de regresión, $\hat{y} = a + bx$.
2. Las estimaciones de mínimos cuadrados son $b = S_{xy}/S_{xx}$ y $a = \bar{y} - b\bar{x}$.

III. Análisis de varianza

1. $\text{TSS} = \text{SSR} + \text{SSE}$, donde $\text{TSS} = S_{yy}$ y $\text{SSR} = (S_{xy})^2/S_{xx}$.
2. La mejor estimación de σ^2 es $\text{MSE} = \text{SSE}/(n-2)$.

IV. Prueba, estimación y predicción

1. Una prueba para la significancia de la regresión lineal, $H_0: \beta = 0$, se puede implementar usando uno de dos estadísticos de prueba:

$$t = \frac{b}{\sqrt{\text{MSE}/S_{xx}}} \quad \text{o} \quad F = \frac{\text{MSR}}{\text{MSE}}$$

2. La fuerza de la relación entre x y y se puede medir usando

$$r^2 = \frac{\text{SSR}}{\text{TSS}}$$

que se acerca a 1 cuando la relación se hace más fuerte.

3. Use gráficas de residuales para comprobar la no normalidad, desigualdad de varianzas, o un modelo incorrectamente ajustado.
4. Los intervalos de confianza se pueden construir para estimar el punto de intersección α y la pendiente β de la recta de regresión, así como estimar el valor promedio de y , $E(y)$, para un valor determinado de x .

- Se pueden construir intervalos de predicción para predecir una observación particular, y , para un valor determinado de x . Para una x dada, los intervalos de predicción siempre son más anchos que los intervalos de confianza.

V. Análisis de correlación

- Use el coeficiente de correlación para medir la relación entre x y y cuando ambas variables son aleatorias:

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

- El signo de r indica la dirección de la relación; r cerca de 0 indica que no hay relación lineal, y r cerca de 1 o -1 indica una fuerte relación lineal.
- Una prueba de la significancia del coeficiente de correlación usa el estadístico

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

y es idéntico a la prueba de la pendiente β .



TECNOLOGÍA ACTUAL

Procedimientos de regresión lineal: *Microsoft Excel*

Hasta ahora se han usado algunos de los procedimientos de regresión lineal disponibles en *Microsoft Excel* para obtener una gráfica de dispersión de los datos y la recta de regresión de mínimos cuadrados y para calcular el coeficiente de correlación r para un conjunto de datos bivariado. Ahora que ha estudiado las técnicas de prueba y estimación para un análisis de regresión lineal simple, más opciones de *MS Excel* están disponibles para usted.

EJEMPLO

8.10

Consulte la tabla 8.1, en la cual se estudió la relación entre x = puntuación en el examen de aprovechamiento en matemáticas y y = calificación final en cálculo.

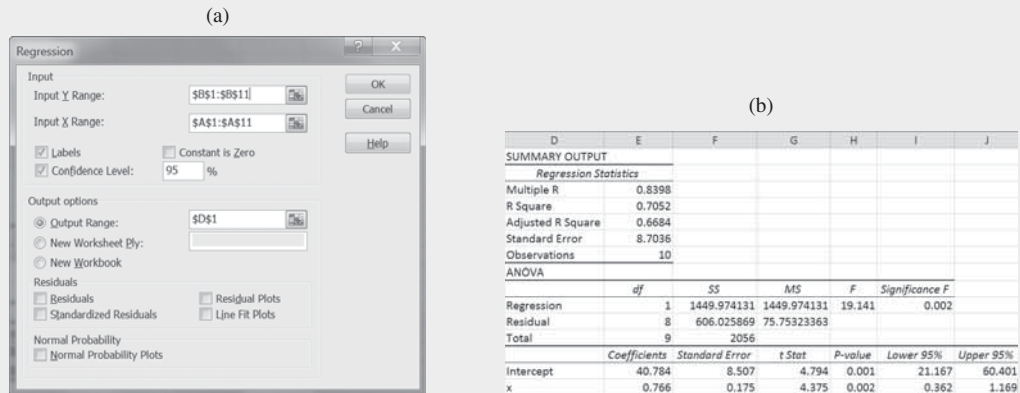
Estudiante	Puntuación en el examen de aprovechamiento en matemáticas (x)	Calificación final en cálculo (y)
1	39	65
2	43	78
3	21	52
4	64	82
5	57	92
6	47	89
7	28	73
8	75	98
9	34	56
10	52	75

Introduzca los valores para x y y en las columnas A y B de una hoja de cálculo de *Excel*.

- Use **Data ► Data Analysis ► Regression** para generar el cuadro de diálogo que se muestra en la figura 8.18a). Resalte o teclee los rangos de celdas para los valores de x y y , y seleccione “Labels” si es necesario.
- Si hace clic en “Confidence Level”, *Excel* calculará intervalos de confianza para las estimaciones de regresión, a y b . Introduzca una ubicación de celda para el **Output Range** y haga clic en **OK**.
- La salida aparecerá en la ubicación de celda seleccionada, y deberá ajustarse usando **Format ► AutoFit Column Width** en la pestaña **Home** en el grupo **Cells** mientras todavía está resaltado. Puede disminuir la precisión decimal si lo desea, usando  en la pestaña **Home** en el grupo **Number** (véase la figura 8.18b)).

4. La salida que se muestra en la figura 8.18b) también puede encontrarse en las figuras 8.6b) y 8.7b), encontrándose su interpretación en las secciones 8.4 y 8.5 del texto.

FIGURA 8.18



NOTA: *MS Excel* no proporciona opciones para la estimación y predicción ni para la prueba de correlación significativa de la sección 8.8. Las gráficas de diagnóstico que pueden generarse en *Excel* no son las mismas gráficas que se expusieron en la sección 8.6 y no se comentarán en esta sección.

Procedimientos de regresión lineal: *MINITAB*

Hasta ahora se han usado algunos de los procedimientos de regresión lineal disponibles en *MINITAB*, para obtener una gráfica de la recta de regresión de mínimos cuadrados de mejor ajuste, así como para calcular el coeficiente de correlación r para un conjunto de datos biva-riados. Ahora que ya hemos estudiado las técnicas de prueba y estimación para un análisis de regresión lineal sencillo, existen más opciones *MINITAB* para usted.

EJEMPLO

8.11

Consulte la tabla 8.1, en la cual se estudió la relación entre x = puntuación en el examen de aprovechamiento en matemáticas y y = calificación final en cálculo.

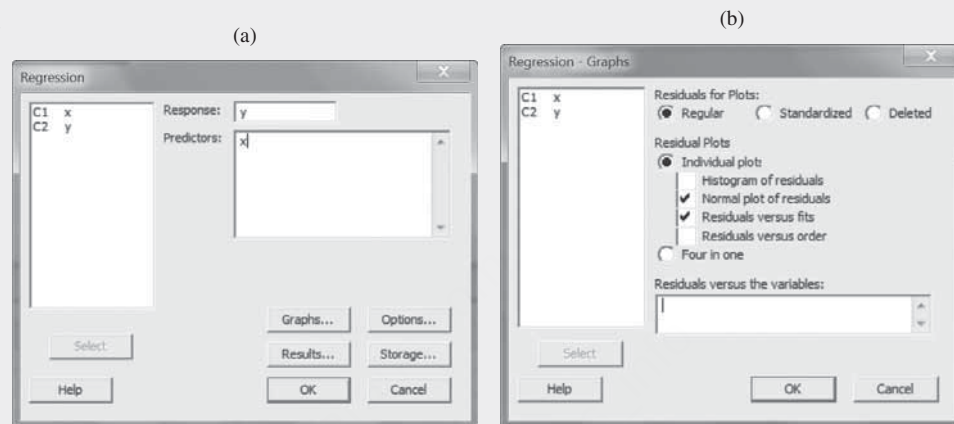
Estudiante	Puntuación en el examen de aprovechamiento en matemáticas (x)	Calificación final en cálculo (y)
1	39	65
2	43	78
3	21	52
4	64	82
5	57	92
6	47	89
7	28	73
8	75	98
9	34	56
10	52	75

Introduzca los valores para x y y en las primeras dos columnas de una hoja de cálculo de *MINITAB*.

- Las herramientas principales para el análisis de regresión lineal se generan usando **Stat ► Regression ► Regression**. (Se usará esta misma secuencia de comandos en el capítulo 9 cuando se estudie el *análisis de regresión múltiple*.) El cuadro de diálogo para el comando Regression se muestra en la figura 8.19a).

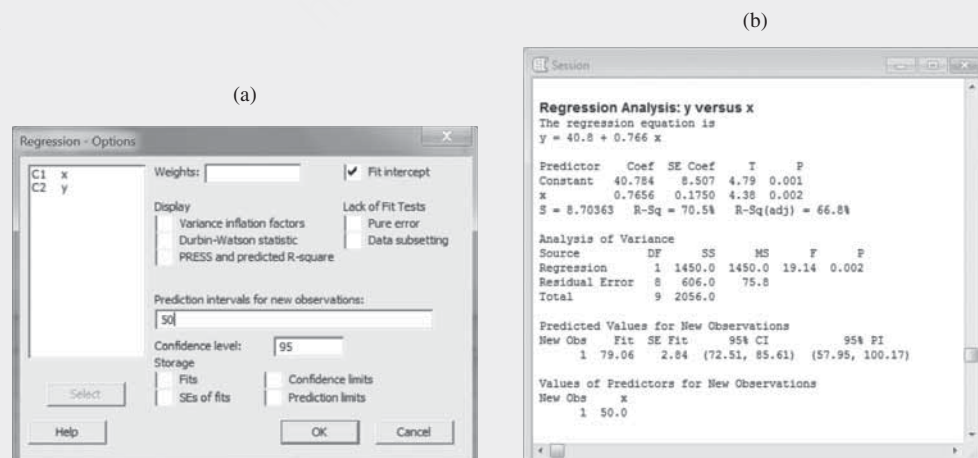
2. Seleccione **y** para la variable “Response” y **x** para la variable “Predictor”. Ahora puede generar algunas gráficas de residuales para comprobar la validez de sus suposiciones de regresión antes de usar el modelo para estimación o predicción. Elija **Graphs** para desplegar el cuadro de diálogo que se muestra en la figura 8.19b). Hemos usado gráficas de residuales **Regular**, seleccionando los cuadros para “Normal plot of residuals” y “Residuals versus fits”. Haga clic en **OK** para regresar al cuadro de diálogo principal.

FIGURA 8.19



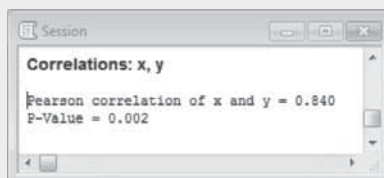
3. Si ahora elige **Options**, puede obtener intervalos de confianza y predicción para cualquiera de estos casos:
 - Un solo valor de x (introducido en el cuadro marcado “Prediction intervals for new observations”).
 - Varios valores de x almacenados en una columna (digamos, C3) de la hoja de cálculo.
4. Introduzca el valor **50** como en la figura 8.20a) para igualar la salida dada en la figura 8.14. Cuando haga clic en **OK** dos veces, se genera la salida de regresión como se muestra en la figura 8.20b). Las dos gráficas de diagnóstico aparecerán en ventanas gráficas separadas.

FIGURA 8.20



5. Si lo desea, ahora puede graficar los puntos de datos, la recta de regresión y los límites de confianza y de predicción superiores e inferiores (véase la figura 8.15) usando **Stat ► Regression ► Fitted Line Plot**. Seleccione y y x para las variables de respuesta y predictor y haga clic en “Display confidence interval” y “Display prediction interval” en el cuadro de diálogo **Options**. Asegúrese de que está seleccionado **Linear** como el “Type of Regression Model”, de modo que obtenga un ajuste lineal para los datos.
6. Se usa el comando **Stat ► Basic Statistics ► Correlation** para obtener el valor del coeficiente de correlación r . Asegúrese de que está seleccionado el cuadro marcado “Display p-values”. La salida para este comando (usando los datos de prueba/calificación) se muestra en la figura 8.21. Note que el valor p para la prueba de $H_0 : \rho = 0$ es idéntico al valor p para la prueba de $H_0 : \beta = 0$, ¡porque las pruebas son exactamente equivalentes!

FIGURA 8.21



Ejercicios suplementarios

DATOS
EX0861

8.61 Potencia de un antibiótico Se realizó un experimento para observar el efecto de un aumento de temperatura en la potencia de un antibiótico.

Tres porciones de una onza del antibiótico se almacenaron durante tiempos iguales a cada una de estas temperaturas: 30°, 50°, 70° y 90°. Las lecturas de potencia observada a cada una de estas temperaturas del periodo experimental se indican a continuación:

Lecturas de potencia, y	38, 43, 29	32, 26, 33	19, 27, 23	14, 19, 21
Temperatura, x	30°	50°	70°	90°

Use un programa de cómputo apropiado para realizar lo siguiente:

- a. Encuentre la recta de mínimos cuadrados apropiada para estos datos.
- b. Grafique los puntos y trace la recta como prueba en sus cálculos.
- c. Construya una tabla ANOVA para regresión lineal.
- d. Si dispone de ellas, examine las gráficas de diagnóstico para comprobar la validez de las suposiciones de regresión.

- e. Estime el cambio en potencia para un cambio de 1 unidad en temperatura. Use un intervalo de confianza de 95%.
- f. Estime el promedio de potencia correspondiente a una temperatura de 50°. Use un intervalo de confianza de 95%.
- g. Suponga que un lote del antibiótico fue almacenado a 50° durante el mismo tiempo que el periodo experimental. Prediga la potencia del lote al término del periodo de almacenamiento. Use un intervalo de predicción de 95%.

DATOS
EX0862

8.62 Ciencia en plantas Se realizó un experimento para determinar el efecto de varios niveles de fósforo, en los niveles de fósforo inorgánico en una planta particular. Los datos de la tabla representan los niveles de fósforo inorgánico en micromoles (μmol) por gramo de peso en seco de raíces de pasto de Sudán, producidas en invernadero durante 28 días, en ausencia de zinc. Use la salida *MINITAB* para contestar las preguntas.

Fosfato aplicado, x	Fósforo en planta, y
.50 μmol	204
	195
	247
	245
.25 μmol	159
	127
	95
	144
.10 μmol	128
	192
	84
	71

- Grafique los datos. ¿Le parece que los datos exhiben una relación lineal?
- Encuentre la recta de mínimos cuadrados que relacione los niveles y de fósforo en planta con la cantidad x de fósforo aplicado al suelo. Grafique la recta de mínimos cuadrados como comprobación de su respuesta.
- ¿Los datos dan suficiente evidencia para indicar que la cantidad de fósforo presente en la planta está linealmente relacionada con la cantidad de fósforo aplicado al suelo?
- Estime la cantidad media de fósforo en la planta si .20 μmol de fósforo se aplica al suelo, en ausencia de zinc. Use un intervalo de confianza de 90%.

Salida MINITAB para el ejercicio 8.62

Análisis de regresión: y contra x

The regression equation is
 $y = 80.9 + 271 x$

Predictor	Coef	SE Coef	T	P
Constant	80.85	22.40	3.61	0.005
x	270.82	68.31	3.96	0.003

S = 39.0419 R-Sq = 61.1% R-Sq(adj) = 57.2%

Predicted Values for New Observations

New Obs	Fit	SE Fit	90.0% CI	90.0% PI
1	135.0	12.6	(112.1, 157.9)	(60.6, 209.4)

Values of Predictors for New Observations

New Obs	x
1	0.200

DATOS
EX0863

8.63 Estadísticas de pista Se realizó un experimento para investigar el efecto de un programa de entrenamiento a lo largo del tiempo para completar la carrera de 100 yardas. Nueve estudiantes se pusieron en el programa. La reducción y en el tiempo para completar la carrera de 100 yardas se midió para tres estudiantes al término de dos semanas, para tres al término de 4 semanas y para tres al término de 6 semanas de entrenamiento. Los datos se dan en la tabla siguiente:

Reducción en tiempo, y (s)	1.6, .8, 1.0	2.1, 1.6, 2.5	3.8, 2.7, 3.1
Tiempo entrenamiento, x (sem)	2	4	6

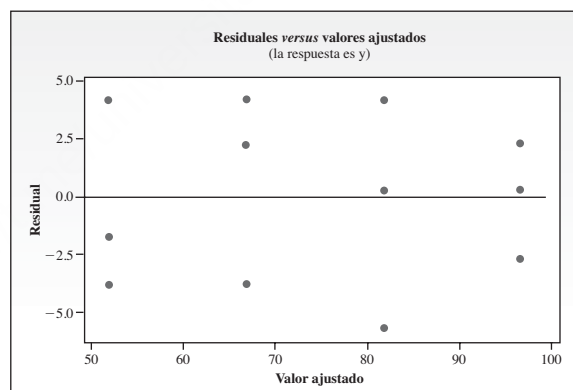
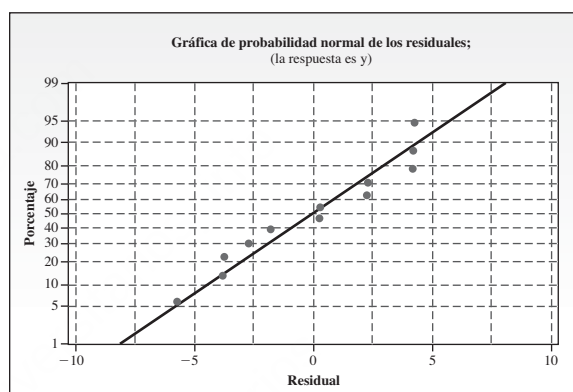
Use un paquete de software apropiado para analizar estos datos. Exprese cualesquiera conclusiones que pueda sacar.

DATOS
EX0864

8.64 Nemátodos Algunas variedades de nemátodos, gusanos que viven en el suelo y se alimentan de raíces de hojas de pasto y de otras plantas, pueden ser tratados con la aplicación de nematicidas. Los datos recolectados sobre el porcentaje de nemátodos muertos por varios rangos de aplicación (dosis dadas en libras por acre de ingrediente activo) son como sigue:

Rango de aplicación, x	2	3	4	5
Porcentaje muertos, y	50, 56, 48	63, 69, 71	86, 82, 76	94, 99, 97

Gráficas MINITAB de diagnóstico para el ejercicio 8.64



Use la salida impresa de computadora apropiada para contestar estas preguntas:

- Calcule el coeficiente de correlación r entre porcentajes de aplicación x y porcentaje de muertes y .
- Calcule el coeficiente de determinación r^2 e interprete.
- Ajuste una recta de mínimos cuadrados a los datos.
- Supongamos que usted desea estimar el porcentaje medio de muertes para una aplicación de 4 libras del nematicida por acre. ¿Qué nos dicen las gráficas de

diagnóstico acerca de la validez de las suposiciones de regresión? ¿Cuáles suposiciones pueden haber sido violadas? ¿Puede explicar por qué?

8.65 Lesiones en rodillas Es frecuente que algunos atletas y otras personas que sufren del mismo tipo de lesiones en las rodillas, requieran reconstrucción del ligamento anterior y posterior. Para determinar la longitud apropiada de injertos entre hueso, tendón rotuliano y otra vez hueso, se realizaron experimentos con el uso de tres técnicas de obtención de imágenes y estos resultados se compararon con la longitud real requerida. Un resumen de los resultados de un análisis de regresión lineal simple, para cada uno de estos tres métodos, se da en la tabla siguiente.¹⁵

Técnica de imagen	Coefficiente de determinación, r^2	Punto de intersección	Pendiente	Valor p
Radiografías	0.80	-3.75	1.031	<0.0001
Resonancia magnética estándar	0.43	20.29	0.497	0.011
Resonancia magnética tridimensional	0.65	1.80	0.977	<0.0001

- ¿Qué se puede decir acerca de la significancia de cada uno de los tres análisis de regresión?
- ¿Cómo se clasificaría la efectividad de los tres análisis de regresión? ¿Cuál es la base de su decisión?
- ¿Cómo se comparan los valores de r^2 y valores p para determinar el mejor pronosticador de longitudes reales de injerto de ligamento requerido?

DATOS EX0866 **8.66 Exámenes de aprovechamiento II** Consulte el ejercicio 8.13 y el conjunto de datos EX0813 respecto a la relación entre el Índice de Aprovechamiento Académico (API), una medida del aprovechamiento escolar basada en los resultados del examen Stanford 9 Achievement y el porcentaje de estudiantes que son considerados Estudiantes del Idioma Inglés (EL). La tabla siguiente muestra el API para ocho escuelas elementales en el condado de Riverside, California, junto con el porcentaje de estudiantes en esa escuela que son considerados Estudiantes del Idioma Inglés.³

Escuela	1	2	3	4	5	6	7	8
API	745	808	798	791	854	688	801	751
EL	71	18	24	50	17	71	11	57

- Use un programa apropiado para analizar la relación entre API y EL.
- Explique todos los detalles pertinentes en su análisis.

8.67 ¿Qué tan largo es? Consulte el ejercicio 8.14 y el conjunto de datos EX0814 respecto a la capacidad de una

persona para estimar tamaños. La tabla siguiente da las longitudes estimadas y reales de los objetos especificados.

Objeto	Estimado (pulgadas)	Real (pulgadas)
Lápiz	7.00	6.00
Plato de comida	9.50	10.25
Libro 1	7.50	6.75
Teléfono celular	4.00	4.25
Fotografía	14.50	15.75
Juguete	3.75	5.00
Cinturón	42.00	41.50
Pinza para ropa	2.75	3.75
Libro 2	10.00	9.25
Calculadora	3.50	4.75

- Use un programa apropiado para analizar la relación entre las longitudes real y estimada de los objetos citados.

- Explique todos los detalles pertinentes de su análisis.

DATOS EX0868 **8.68 Tenis, ¿quiere jugar?** Si usted juega tenis, sabe que las raquetas varían en sus características físicas. Los datos de la tabla siguiente dan medidas de rigidez al doblamiento y rigidez a la torcedura, medidas por pruebas de ingeniería para 12 raquetas de tenis:

Raqueta	Rigidez al doblamiento, x	Rigidez a la torcedura, y
1	419	227
2	407	231
3	363	200
4	360	211
5	257	182
6	622	304
7	424	384
8	359	194
9	346	158
10	556	225
11	474	305
12	441	235

- Si una raqueta tiene rigidez al doblamiento, ¿también es probable que tenga rigidez a la torcedura? ¿Los datos dan evidencia de que x y y están correlacionados positivamente?
- Calcule el coeficiente de determinación r^2 e interprete su valor.

DATOS EX0869 **8.69 Investigación del aguacate** El movimiento de aguacates en Estados Unidos, desde ciertos lugares, está prohibido debido a la posibilidad de introducir moscas de la fruta al país con los embarques de aguacate. No obstante, ciertas variedades de aguacate supuestamente son resistentes a la infesta de moscas de la fruta antes de ablandarse como resultado de su madurez. Los datos de la tabla siguiente resultaron de un experimento en el que aguacates

que iban de 1 a 9 días después de cosecharse fueron expuestos a la mosca de la fruta del Mediterráneo. La penetrabilidad de los aguacates se midió en el día de su exposición, evaluándose el porcentaje de aguacates infestados.

Días después de cosecha	Penetrabilidad	Porcentaje infestado
1	.91	30
2	.81	40
3	.95	45
4	.95	45
5	1.04	57
6	1.22	60
7	1.38	75
9	1.77	100

Use la salida impresa *MINITAB* de la regresión del porcentaje infestado (y) en días después de la cosecha (x), para analizar la relación entre estas dos variables. Explique todas las partes pertinentes de la salida impresa e interprete los resultados de cualesquiera pruebas.

Salida *MINITAB* para el ejercicio 8.69

Análisis de regresión: Porcentaje contra x

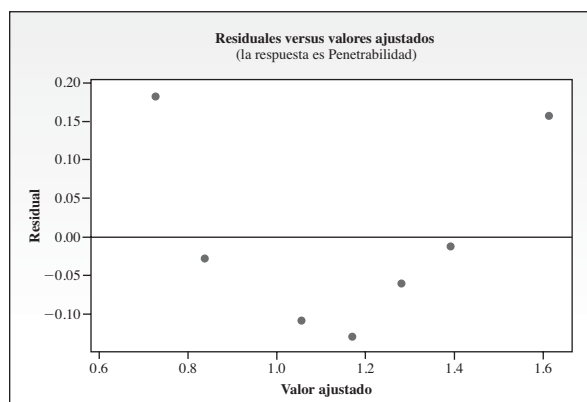
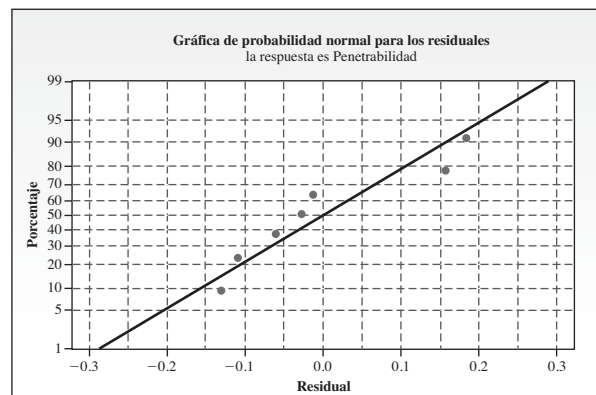
The regression equation is
Percent = 18.4 + 8.18 x

Predictor	Coef	SE Coef	T	P
Constant	18.427	5.110	3.61	0.015
x	8.1768	0.9285	8.81	0.000

S = 6.35552 R-Sq = 93.9% R-Sq(adj) = 92.7%

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	3132.9	3132.9	77.56	0.000
Residual Error	5	202.0	40.4		
Total	6	3334.9			

8.70 Aguacates II Consulte el ejercicio 8.69. Suponga que el experimentador desea examinar la relación entre la penetrabilidad y el número de días después de la cosecha. ¿El método de regresión lineal estudiado en este capítulo es un método apropiado de análisis? Si no es así, ¿qué suposiciones han sido violadas? Use las gráficas de diagnóstico proporcionadas.



DATOS

EX0871

8.71 Metabolismo y aumento de peso ¿Por qué razón una persona puede aumentar de peso, incluso si no come más y se ejercita no menos que un amigo delgado? Estudios recientes sugieren que los factores que controlan el metabolismo pueden depender de su estructura genética. Un estudio comprendió 11 pares de gemelos idénticos alimentados con mil calorías al día más de lo necesario para mantener un peso inicial. Las actividades se mantuvieron constantes y el ejercicio fue mínimo. Al término de 100 días, los cambios en el peso corporal (en kilogramos) se registraron para los 22 gemelos.¹⁶ ¿Hay una correlación positiva significativa entre los cambios en peso corporal para los gemelos? ¿Se puede concluir que esta similitud es causada por semejanzas genéticas? Explique.

Par	Gemelo A	Gemelo B
1	4.2	7.3
2	5.5	6.5
3	7.1	5.7
4	7.0	7.2
5	7.8	7.9
6	8.2	6.4
7	8.2	6.5
8	9.1	8.2
9	11.5	6.0
10	11.2	13.7
11	13.0	11.0

DATOS

EX0872

8.72 Repaso de películas ¿Cuántas semanas puede exhibirse una película y todavía tener utilidades razonables? Los datos que siguen muestran el número de semanas en exhibición (x) y la suma total a la fecha (y) para las mejores 12 películas durante una semana reciente.¹⁷

Película	Suma total a la fecha (en millones)	Semanas de exhibición
1. <i>The Town</i>	33.1	1
2. <i>Easy A</i>	22.1	1
3. <i>Resident Evil: Afterlife 3D</i>	47.1	2
4. <i>Devil</i>	15.3	1
5. <i>Alpha and Omega 3D</i>	10.4	1
6. <i>The American</i>	33.7	3
7. <i>Takers</i>	53.3	4
8. <i>Eat Pray Love</i>	78.3	6
9. <i>Inception</i>	285.8	10
10. <i>Machete</i>	25.1	3
11. <i>The Other Guys</i>	116.0	7
12. <i>Going the Distance</i>	17.3	3

- Grafique los puntos en una gráfica de dispersión. ¿Le parece que la relación entre x y y es lineal? ¿Cómo describiría usted la dirección y fuerza de la relación?
- Calcule el valor de r^2 . ¿Qué porcentaje de la variación general se explica al usar el modelo lineal en lugar de $\bar{y}$ para predecir la variable de respuesta y ?
- ¿Cuál es la ecuación de regresión? ¿Los datos dan evidencia para indicar que x y y están relacionados linealmente? Pruebe usando un nivel de significancia de 5%.
- Dados los resultados de las partes **b** y **c**, ¿es apropiado usar la recta de regresión para estimar y predecir? Explique su respuesta.

8.73 Además de límites cada vez más grandes en el error, ¿por qué un experimentador debe abstenerse de predecir y para valores de x fuera de la región experimental?

8.74 Si el experimentador continúa dentro de la región experimental, ¿cuándo será máximo el error al predecir un valor particular de y ?

DATOS EX0875 **8.75 Avena, ¿alguien quiere?** Un experimentador agrícola, investigando el efecto de la cantidad de nitrógeno x aplicada en 100 libras por acre en la producción de avena y , medida en bushels por acre, recolectó los siguientes datos:

x	1	2	3	4
y	22	38	57	68
	19	41	54	65

- Encuentre la recta de mínimos cuadrados para los datos.
- Construya la tabla ANOVA.
- ¿Hay suficiente evidencia para indicar que la producción de avena está linealmente relacionada con la cantidad de nitrógeno aplicada? Use $\alpha = .05$.
- Prediga la producción esperada de avena con 95% de confianza si se aplican 250 libras de nitrógeno por acre.

- Estime el promedio de aumento en producción para un aumento de 100 libras de nitrógeno por acre con 99% de confianza.
- Calcule r^2 y explique su significancia en términos de predecir y , la producción de avena.

DATOS EX0876

8.76 Rosas frescas Un horticultor inventó una escala para medir la frescura de rosas que fueron empacadas y almacenadas durante periodos variables antes de trasplantarlas. La medición y de frescura y el tiempo x en días que la rosa está empacada y almacenada antes de trasplantarla, se dan a continuación.

x	5	10	15	20	25
y	15.3	13.6	9.8	5.5	1.8
	16.8	13.8	8.7	4.7	1.0

- Ajuste una recta de mínimos cuadrados a los datos.
- Construya la tabla ANOVA.
- ¿Hay suficiente evidencia para indicar que la frescura está linealmente relacionada con el tiempo de almacenaje? Use $\alpha = .05$.
- Estime la rapidez media de cambio en frescura para un aumento de un día en tiempo de almacenaje, usando un intervalo de confianza de 98%.
- Estime la medición de frescura esperada para un tiempo de almacenaje de 14 días con un intervalo de confianza de 95%.
- ¿De qué valor es el modelo lineal cuando se compara con $\bar{y}$ para predecir la frescura?

DATOS EX0877

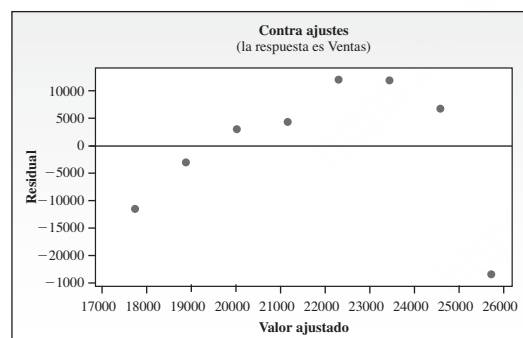
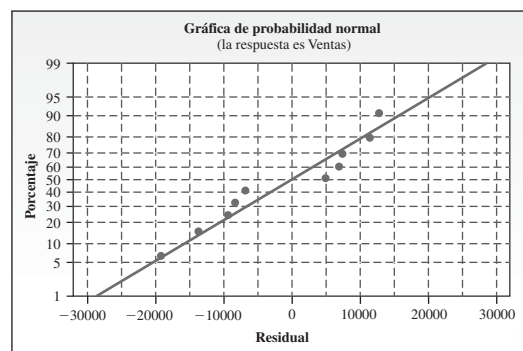
8.77 Lexus, Inc. El Lexus GX es una minivan de tamaño mediano que es vendida en los mercados norteamericano y euroasiático por Lexus. La GX 470 fue introducida en 2002 (como un modelo 2003) y más adelante fue mejorada con un nuevo sistema de suspensión todoterreno. Las ventas de la Lexus GX 470 desde su introducción hasta 2009 se proporcionan en la tabla.¹⁸

Generación	Modelo(s)	Año calendario	Ventas totales (EUA)
UZJ120	GX 470	2002	2,190
		2003	31,376
		2004	35,420
		2005	34,339
		2006	25,454
		2007	23,035
		2008	15,759
		2009	6,235

- Trace los puntos usando una gráfica de dispersión. ¿Cómo describiría usted la relación entre año y ventas de la Lexus GX 470?

- b. Aun cuando la gráfica de dispersión en la parte **a** podría indicarlo de manera diferente, suponga que la relación entre el año y las ventas es lineal. Encuentre la recta de regresión de mínimos cuadrados que relaciona las ventas de la Lexus GX 470 con el año que se está midiendo.
- c. ¿Hay suficiente evidencia para indicar que las ventas están linealmente relacionadas con el año? Use $\alpha = .05$.
- d. Examine las gráficas de diagnóstico que se muestran a continuación. ¿Qué puede concluir acerca de la validez de las suposiciones de la regresión?

Gráficas de diagnóstico para el ejercicio 8.77



- e. Con base en sus conclusiones en la parte **d**, ¿es aconsejable predecir las ventas de 2010 usando la recta de regresión de la parte **b**? Explique.

DATOS
EX0878

8.78 Starbucks Aquí hay algunos datos nutricionales para una muestra de bebidas Espresso de 16 onzas líquidas de Starbucks, elaboradas con 2% de leche. La información nutricional para todos los productos de Starbucks puede encontrarse en el sitio web de la compañía, www.starbucks.com.¹⁹

Producto	Calorías	Grasa (g)	Carb. (g)	Fibra (g)	Proteínas (g)
Caffe Latte	190	7	18	0	12
Caffe Mocha	260	8	41	2	13
Cappuccino	120	4	12	0	8
Caramel Macchiato	240	7	34	0	10
Cinnamon Dolce Latte	260	6	40	0	11
Flavored Latte	250	6	36	0	12
Iced Caffe Latte	130	4.5	13	0	8
Iced Caffe Mocha	200	6	35	2	9
Iced Caramel Macchiato	230	6	33	0	10
Iced Cinnamon Dolce Latte	200	4	34	0	7
Iced Flavored Latte	250	6	36	0	12
Iced Peppermint Mocha	260	6	52	2	8
Iced Peppermint White Chocolate Mocha	400	9	72	0	10
Iced Pumpkin Spice Latte	250	4	44	0	10
Iced Skinny Flavored Latte	110	4	12	0	7
Iced Toffee Mocha	280	3.5	51	2	12
Iced White Chocolate Mocha	340	9	55	0	10
Peppermint Mocha	330	8	57	2	12
Peppermint White Chocolate Mocha	470	12	78	0	14
Pumpkin Spice Latte	310	6	49	0	14
Skinny Cinnamon Dolce Latte	180	6	18	0	12
Skinny Flavored Latte	180	6	18	0	12
Toffee Mocha	350	7	58	2	17
White Chocolate Mocha	400	11	61	0	15

Use los métodos estadísticos apropiados para analizar las relaciones entre algunas de las variables nutricionales dadas en la tabla. Escriba un informe en resumen que explique cualesquiera conclusiones que pueda sacar de su análisis.

CASO PRÁCTICO

DATOS Carros importados

¿Su automóvil está “Hecho en EUA”?

La frase “Hecho en EUA” se ha convertido en un conocido grito de batalla porque los trabajadores de Estados Unidos tratan de proteger sus trabajos de la competencia extranjera. En las últimas décadas, un importante desequilibrio en la balanza comercial en Estados Unidos ha estado causando una inundación de productos importados que entran al país y se venden a menor costo que artículos comparables hechos en él. Una preocupación principal es la industria automotriz, en la que el número de autos importados aumentó continuamente durante las décadas de 1970 y 1980. La industria automotriz de ese país ha estado siendo acosada con quejas por la calidad de sus productos, despidos de trabajadores y altos precios, y ha gastado miles de millones de dólares en publicidad e investigación para producir un auto hecho en Estados Unidos que satisfaga las demandas del consumidor. ¿Han tenido éxito para detener la

inundación de autos importados comprados por consumidores estadounidenses? Los datos de la tabla siguiente representan los números de autos importados y vendidos en Estados Unidos (en millones) durante los años 1969-2009.²⁰ Para simplificar el análisis, hemos codificado el año usando la variable codificada $x = \text{Año} - 1969$.

Año	(Año – 1969), x	Número de autos importados, y	Año	(Año – 1969), x	Número de autos importados, y
1969	0	1.1	1989	20	2.7
1970	1	1.3	1990	21	2.4
1971	2	1.6	1991	22	2.0
1972	3	1.6	1992	23	1.9
1973	4	1.8	1993	24	1.8
1974	5	1.4	1994	25	1.7
1975	6	1.6	1995	26	1.5
1976	7	1.5	1996	27	1.3
1977	8	2.1	1997	28	1.4
1978	9	2.0	1998	29	1.4
1979	10	2.3	1999	30	1.7
1980	11	2.4	2000	31	2.0
1981	12	2.3	2001	32	2.1
1982	13	2.2	2002	33	2.2
1983	14	2.4	2003	34	2.1
1984	15	2.4	2004	35	2.1
1985	16	2.8	2005	36	2.2
1986	17	3.2	2006	37	2.3
1987	18	3.1	2007	38	2.4
1988	19	3.0	2008	39	2.3
			2009	40	1.8

1. Usando una gráfica de dispersión, grafique los datos para los años 1969-1988. ¿Le parece que hay una relación lineal entre el número de autos importados y el año?
2. Use un paquete de software para hallar la recta de mínimos cuadrados para predecir el número de autos importados como función del año para los años 1969-1988.
3. ¿Hay una relación lineal significativa entre el número de autos importados y el año?
4. Use el programa de cómputo para predecir el número de autos que serán importados usando intervalos de predicción de 95% para cada uno de los años 2007, 2008 y 2009.
5. Ahora vea los datos reales para los años 2007-2009. ¿Las predicciones obtenidas en el paso 4 dan estimaciones precisas de los valores reales observados en estos años? Explique.
6. Agregue los datos para 1989-2009 a su base de datos y recalcule la recta de regresión. ¿Qué efecto tienen los nuevos puntos de datos sobre la pendiente? ¿Cuál es el efecto en el SSE?
7. Dada la forma de la gráfica de dispersión para los años 1969-2009, ¿le parece que una recta da un modelo preciso para los datos? ¿Qué otro tipo de modelo podría ser más apropiado? (Use gráficas residuales para ayudar a contestar esta pregunta.)

Análisis de regresión múltiple

OBJETIVO GENERAL

En este capítulo extendemos los conceptos de regresión y correlación lineales a una situación donde el valor promedio de una variable aleatoria y está relacionada con varias variables independientes, $x_1, x_2, \dots, x_k$, en modelos que son más flexibles que el modelo de recta del capítulo 8. Con el análisis de regresión múltiple, podemos usar la información proporcionada por las variables independientes para ajustar varios tipos de modelos a los datos muestrales, para evaluar la utilidad de estos modelos y finalmente para estimar el valor promedio de y o predecir el valor real de y para valores dados de $x_1, x_2, \dots, x_k$.

ÍNDICE DEL CAPÍTULO

- Introducción (9.1)
- El modelo de regresión múltiple (9.2)
- Un análisis de regresión múltiple (9.3)
- Un modelo de regresión polinomial (9.4)
- Uso de variables predictoras cuantitativas y cualitativas en un modelo de regresión (9.5)
- Prueba de conjuntos de coeficientes de regresión (9.6)
- Interpretación de gráficas residuales (9.7)
- Análisis de regresión por pasos (9.8)
- Interpretación errónea de un análisis de regresión (9.9)
- Pasos a seguir al construir un modelo de regresión múltiple (9.10)

9.1

INTRODUCCIÓN

La **regresión lineal múltiple** es una extensión de la regresión lineal simple para tomar en cuenta más de una variable independiente. Esto es, en lugar de usar sólo una variable independiente x para explicar la variación en y , se usan simultáneamente diversas variables independientes (o elementos de predicción). Con el uso de más de una variable independiente se debe hacer un mejor trabajo para explicar la variación en y y en consecuencia hacer predicciones más precisas.

Por ejemplo, las ventas regionales y del producto de una compañía podrían estar relacionadas con tres factores:

- x_1 : la cantidad gastada en publicidad en televisión
- x_2 : la cantidad gastada en publicidad en periódicos
- x_3 : el número de vendedores asignados a la región

Un investigador recolectaría datos para medir las variables y , x_1 , x_2 y x_3 , y luego usaría estos datos muestrales para construir una ecuación de predicción que relacionara y con las tres variables predictoras. Desde luego que surgen varias preguntas, al igual que con la regresión lineal simple:

- ¿Qué tan bien se ajusta el modelo?
- ¿Qué tan fuerte es la relación entre y y las variables predictoras?
- ¿Se han violado suposiciones importantes?
- ¿Qué tan buenas son las estimaciones y predicciones?

Para contestar estas preguntas se pueden usar métodos de **análisis de regresión múltiple**, que casi siempre se hacen con un programa de cómputo. Este capítulo contiene una breve introducción al análisis de regresión múltiple y a la difícil tarea de construcción de modelos, es decir, elegir el modelo correcto para una aplicación práctica.

9.2

EL MODELO DE REGRESIÓN MÚLTIPLE

El **modelo lineal general** para un análisis de regresión múltiple describe una respuesta particular y usando el modelo que damos a continuación.

MODELO LINEAL GENERAL Y SUPOSICIONES

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon$$

donde

- y es la **variable de respuesta** que se desea predecir.
- $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ son constantes desconocidas.
- $x_1, x_2, \dots, x_k$ son **variables predictoras** independientes que se miden sin error.
- ϵ es el error aleatorio, que permite que cada respuesta se desvíe del valor promedio de y en una cantidad ϵ . Se debe suponer que los valores de ϵ : 1) son independientes; 2) tienen una media de 0 y una varianza común σ^2 para cualquier conjunto $x_1, x_2, \dots, x_k$ y 3) están normalmente distribuidos.

Cuando se cumplen estas suposiciones acerca de ϵ , el valor *promedio* de y para un conjunto dado de valores $x_1, x_2, \dots, x_k$ es igual a la parte *determinista* del modelo:

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

Observe que el modelo de regresión simple y las suposiciones son *muy semejantes* al modelo y las suposiciones empleados para la regresión lineal. Es probable que no lo sorprenda el que los procedimientos de prueba y estimación también sean extensiones de los empleados en el capítulo 8.

Los modelos de regresión múltiple son muy flexibles y pueden tomar muchas formas, dependiendo de la manera en que las variables independientes $x_1, x_2, \dots, x_k$ se introduzcan en el modelo. Empezamos con un simple modelo de regresión múltiple, explicando los conceptos y procedimientos básicos con un ejemplo. A medida que nos familiaricemos con los procedimientos de regresión múltiple, aumentamos la complejidad de los ejemplos y veremos que los mismos procedimientos se pueden usar para modelos de formas diferentes, dependiendo de la aplicación particular.

EJEMPLO

9.1

Suponga que se desea relacionar una variable aleatoria y con dos variables independientes x_1 y x_2 . El modelo de regresión múltiple es

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

con el valor medio de y dado como

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

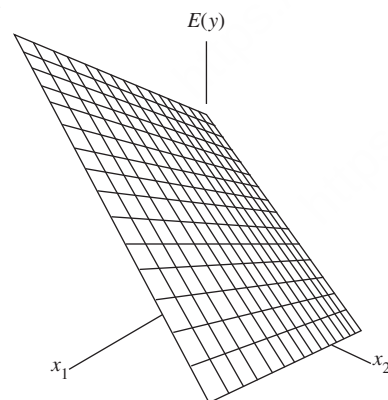
TIP MI CONSEJO

En lugar de x y y graficadas en un espacio en dos dimensiones, y y $x_1, x_2, \dots, x_k$ tienen que graficarse en $(k+1)$ dimensiones

Esta ecuación es una extensión en tres dimensiones de la **recta de medias** del capítulo 8 y traza un **plano** en el espacio tridimensional (véase la figura 9.1). La constante β_0 se denomina **punto de cruce**, que es el valor promedio de y cuando x_1 y x_2 son 0 ambas. Los coeficientes β_1 y β_2 se denominan **pendientes parciales** o **coeficientes de regresión parciales**. La pendiente parcial β_i (para $i = 1$ o 2) mide el cambio en y para un cambio unitario en x_i cuando *todas las otras variables independientes se mantienen constantes*. El valor del coeficiente de regresión parcial, por ejemplo β_1 , con x_1 y x_2 en el modelo generalmente *no* es igual que la pendiente cuando se ajuste una recta sólo con x_1 . Estos coeficientes son las constantes desconocidas, que deben ser estimadas usando datos muestrales para obtener la ecuación de predicción.

FIGURA 9.1

Plano de medias para el ejemplo 9.1



9.3

UN ANÁLISIS DE REGRESIÓN MÚLTIPLE

Un análisis de regresión múltiple comprende procedimientos de estimación, prueba y diagnóstico diseñados para ajustar el modelo de regresión múltiple

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

a un conjunto de datos. Debido a la complejidad de los cálculos involucrados, estos procedimientos casi siempre se ponen en práctica con un programa de regresión de alguno de los varios paquetes de software. Todos dan resultados similares en formas ligeramente diferentes. Seguimos con patrones básicos puestos en regresión lineal simple, empezando con un resumen de los procedimientos generales e ilustrado con un ejemplo.

El método de mínimos cuadrados

La ecuación de predicción

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k$$

es la recta que reduce la SSE, la suma de cuadrados de las desviaciones de los valores y observados en los valores pronosticados $\hat{y}$. Estos valores se calculan usando un programa de regresión.

EJEMPLO

9.2

¿En qué forma los vendedores de bienes raíces determinan el precio de venta para un condominio recién inscrito en lista? La base de datos de una computadora en una pequeña comunidad contiene el precio de venta de lista y (en miles de dólares), la cantidad de área de vivienda x_1 (en cientos de pies cuadrados), así como los números de pisos x_2 , recámaras x_3 y baños x_4 , para $n = 15$ condominios seleccionados al azar actualmente en el mercado. Los datos se muestran en la tabla 9.1.

TABLA 9.1

Datos sobre 15 condominios

Observación	Precio de lista, y	Área de vivienda, x_1	Pisos, x_2	Recámaras, x_3	Baños, x_4
1	169.0	6	1	2	1
2	218.5	10	1	2	2
3	216.5	10	1	3	2
4	225.0	11	1	3	2
5	229.9	13	1	3	1.7
6	235.0	13	2	3	2.5
7	239.9	13	1	3	2
8	247.9	17	2	3	2.5
9	260.0	19	2	3	2
10	269.9	18	1	3	2
11	234.9	13	1	4	2
12	255.0	18	1	4	2
13	269.9	17	2	4	3
14	294.5	20	2	4	3
15	309.9	21	2	4	3

El modelo de regresión múltiple es

$$E(y) = \beta_0 = \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4$$

que se ajusta usando los paquetes de software *MINITAB* o *Microsoft Excel*. En la sección “Tecnología actual”, al final de este capítulo, se encuentran instrucciones para generar esta salida. La primera parte de la salida de regresión de *MINITAB* se muestra en la figura 9.2a). Se encontrará la ecuación de regresión ajustada en los primeros dos renglones de la salida impresa:

$$\hat{y} = 119 + 6.27x_1 - 16.2x_2 - 2.67x_3 + 30.3x_4$$

Los coeficientes de regresión parcial se muestran con ligeramente más precisión en la segunda sección de la salida impresa de *MINITAB*, una salida similar generada por *MS Excel* se muestra en la figura 9.2b). Las columnas son una lista del nombre dado a cada variable independiente de pronóstico, su coeficiente de regresión estimado, su error estándar y los valores t y p que se usan para probar su significancia *en presencia de todas las otras variables predictoras*. En una sección más adelante explicamos estas pruebas con más detalle.

FIGURA 9.2(a)

Una porción de la salida impresa *MINITAB* para el ejemplo 9.2

Análisis de regresión: precio de lista contra pies cuadrados, número de pisos, recámaras, baños

The regression equation is
List Price = 119 + 6.27 Square Feet - 16.2 Number of Floors
- 2.67 Bedrooms + 30.3 Baths

Predictor	Coef	SE Coef	T	P
Constant	118.763	9.207	12.90	0.000
Square Feet	6.2698	0.7252	8.65	0.000
Number of Floors	-16.203	6.212	-2.61	0.026
Bedrooms	-2.673	4.494	-0.59	0.565
Baths	30.271	6.849	4.42	0.001

FIGURA 9.2(b)

Una porción de la salida impresa de *MS Excel* para el ejemplo 9.2

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercepción	118.763	9.207	12.899	0.000	98.248	139.279
Pies cuadrados	6.270	0.725	8.645	0.000	4.654	7.886
Número de pisos	-16.203	6.212	-2.608	0.026	-30.045	-2.362
Recámaras	-2.673	4.494	-0.595	0.565	-12.686	7.340
Baños	30.271	6.849	4.420	0.001	15.011	45.530

El análisis de varianza para regresión múltiple

El análisis de varianza divide la variación total en la variable de respuesta y,

$$SS \text{ Total} = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$$

en dos partes:

- La SSR (suma de cuadrados para regresión) mide la cantidad de variación explicada usando la ecuación de regresión.
- La SSE (suma de cuadrados para error) mide la variación residual en los datos que no está explicada por las variables independientes.

de modo que

$$SS \text{ total} = SSR + SSE$$

Los **grados de libertad** para estas sumas de cuadrados se encuentran usando el siguiente argumento. Hay $(n - 1)$ grados de libertad en total. Estimar la recta de regresión requiere estimar k coeficientes desconocidos: $\beta_1, \beta_2, \dots, \beta_k$; la constante b_0 (la cual estima β_0) es una función de $\bar{y}$ y de las otras estimaciones. En consecuencia, hay k grados de libertad de regresión, dejando $(n - 1) - k$ grados de libertad para error. Al igual que en capítulos anteriores, las medias cuadráticas se calculan como $MS = SS/gl$.

La tabla ANOVA para los datos de bienes raíces de la tabla 9.1 se muestra en la segunda porción de la salida impresa *MINITAB* en la figura 9.3a) y en la sección inferior de la salida impresa de *Excel* en la figura 9.3b). Hay $n = 15$ observaciones y $k = 4$ variables predictoras independientes. Se puede verificar que el total de grados de libertad, $(n - 1) = 14$, se divide en $k = 4$ para regresión y $(n - k - 1) = 10$ para error.

FIGURA 9.3(a)

Una porción de la salida impresa *MINITAB* para el ejemplo 9.2

S = 6.84930 R-Sq = 97.1% R-Sq(adj) = 96.0%

Analysis of Variance						
Source	DF	SS	MS	F	P	
Regression	4	15913.0	3978.3	84.80	0.000	
Residual Error	10	469.1	46.9			
Total	14	16382.2				

Source	DF	Seq SS
Square Feet	1	14829.3
Number of Floors	1	0.9
Bedrooms	1	166.4
Baths	1	916.5

FIGURA 9.3(b)

Una porción de la salida impresa de *MS Excel* para el ejemplo 9.2

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.986
R Square	0.971
Adjusted R Square	0.960
Standard Error	6.849
Observations	15

ANOVA					
	df	SS	MS	F	Significance F
Regression	4	15913.048	3978.262	84.801	0.000
Residual	10	469.129	46.913		
Total	14	16382.177			

La mejor estimación de la variación aleatoria σ^2 en el experimento, es decir la variación que no es explicada por las variables predictoras, como de costumbre está dada por

$$s^2 = \text{MSE} = \frac{\text{SSE}}{n - k - 1} = 46.913$$

de la tabla ANOVA. El primer renglón de la figura 9.3a) y el cuarto renglón en la figura 9.3b) también muestran $s = \sqrt{s^2} = 6.849$. La computadora usa estos valores internamente para producir estadísticos de prueba, intervalos de confianza e intervalos de predicción, que estudiaremos en secciones subsiguientes.

La última sección de la figura 9.3a) muestra una descomposición de $\text{SSR} = 15913.0$ en que la contribución condicional de cada variable de predicción *dadas las variables ya introducidas en el modelo* se muestra para el orden de entrada que se especifique en el programa de regresión. Para el ejemplo de bienes raíces, el programa *MINITAB* introdujo las variables en este orden: pies cuadrados, seguido de números de pisos, recámaras y baños. Cada una de estas **sumas de cuadrados secuenciales** o condicionales explican uno de los $k = 4$ grados de libertad de regresión. Es interesante observar que la variable de predicción x_1 por sí sola da cuenta de $14829.3/15913.0 = .932$ o 93.2% de la variación total explicada por el modelo de regresión, pero, si cambia el orden de entrada, otra variable puede explicar la mayor parte de la suma de cuadrados de regresión.

Prueba de la utilidad del modelo de regresión

Recuerde que en el capítulo 8 probó ver si y y x estaban linealmente relacionadas al probar $H_0 : \beta = 0$ con una prueba t o una prueba F equivalente. En regresión múltiple, hay más de una *pendiente parcial*, que son los *coeficientes de regresión parcial*. Las pruebas t y F ya no son equivalentes.

El análisis de varianza de la prueba F

La ecuación de regresión que usa información dada por las variables predictoras $x_1, x_2, \dots, x_k$ ¿es considerablemente mejor que la predictora simple $\bar{y}$ que no se apoya en ninguno de los valores de x ? Esta pregunta se contesta usando una prueba F general con las hipótesis:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

contra

$$H_a: \text{Al menos una de } \beta_1, \beta_2, \dots, \beta_k \text{ no es } 0$$

El estadístico de prueba se encuentra en la tabla ANOVA (figura 9.3) como

$$F = \frac{\text{MSR}}{\text{MSE}} = \frac{3978.3}{46.9} = 84.80$$

que tiene una distribución F con $gl_1 = k = 4$ y $gl_2 = (n - k - 1) = 10$. Como el valor p exacto, $P = .000$, está dado en la salida impresa, se puede declarar que la regresión es altamente significativa. Esto es, al menos una de las variables predictoras está aportando información significativa para la predicción de la variable de respuesta y .

El coeficiente de determinación, R^2

¿Qué tan bien se ajusta el modelo de regresión? La salida impresa de regresión da una medida estadística de la fuerza del modelo en el **coeficiente de determinación, R^2** ; es decir, la proporción de la variación total que es explicada por la regresión de y en $x_1, x_2, \dots, x_k$, definida como

$$R^2 = \frac{\text{SSR}}{\text{SS Total}} = \frac{15\,913.0}{16\,382.2} = .971 \quad \text{o} \quad 97.1\%$$

El coeficiente de determinación se denomina a veces **R^2 múltiple** y se encuentra en el primer renglón de la figura 9.3a), marcado “R-Sq” y en la segunda línea de la figura 9.3b), etiquetada “R Square”. En consecuencia, para el ejemplo de bienes raíces, 97.1% de la variación total ha sido explicado por el modelo de regresión. El modelo se ajusta muy bien.

Puede ser útil saber que el valor del estadístico F está relacionado con R^2 por la fórmula

$$F = \frac{R^2/k}{(1 - R^2)/(n - k - 1)}$$

de manera que R^2 es grande, F es grande y viceversa.

Interpretación de los resultados de una regresión significativa

Prueba de la significancia de los coeficientes de regresión parcial

Una vez que hayamos determinado que el modelo es útil para predecir y , debemos explorar la naturaleza de la “utilidad” en más detalle. ¿Todas las variables predictoras agregan información importante para la predicción *en presencia de otras variables predictoras que ya están en el modelo*? Las pruebas t individuales de la primera sección de la salida impresa de regresión están diseñadas para probar las hipótesis

TIP MI CONSEJO

La prueba F general (para la significancia del modelo) en regresión múltiple es de una cola

TIP MI CONSEJO

Las salidas impresas MINITAB informan de R^2 como un porcentaje más que una proporción

TIP MI CONSEJO

R^2 es la multivariada equivalente de r^2 , empleada en regresión lineal

TIP MI CONSEJO

Se puede demostrar que

$$F = \frac{\text{MSR}}{\text{MSE}} = \frac{R^2/k}{(1 - R^2)/(n - k - 1)}$$

$$H_0 : \beta_i = 0 \quad \text{contra} \quad H_a : \beta_i \neq 0$$

para cada uno de los coeficientes de regresión parcial, *dado que las otras variables predictoras ya están en el modelo*. Estas pruebas están basadas en el estadístico t de Student dado por

$$t = \frac{b_i - \beta_i}{SE(b_i)}$$

que tiene $gl = (n - k - 1)$ grados de libertad. El procedimiento es idéntico al empleado para probar una hipótesis acerca de la pendiente β del modelo de regresión lineal simple.[†]

La figura 9.4 muestra las pruebas t y los valores p de la parte superior de la salida impresa de *MINITAB* y la sección inferior de la salida impresa de *MS Excel*. Al examinar los valores p de la última columna se puede ver que todas las variables *excepto* x_3 , el número de recámaras, agregan información significativa para predecir y , **aún cuando todas las otras variables independientes ya estén en el modelo**. ¿Podría ser mejor el modelo? Tal vez x_3 sea una variable de predicción innecesaria. Una opción es eliminarla y reajustar el modelo con un nuevo conjunto de datos.

TIP MI CONSEJO

Pruebe la significancia del coeficiente individual β_i , usando pruebas t

FIGURA 9.4(a)

Una porción de la salida impresa *MINITAB* para el ejemplo 9.2

Predictor	Coef	SE Coef	T	P
Constant	118.763	9.207	12.90	0.000
Square Feet	6.2698	0.7252	8.65	0.000
Number of Floors	-16.203	6.212	-2.61	0.026
Bedrooms	-2.673	4.494	-0.59	0.565
Baths	30.271	6.849	4.42	0.001

FIGURA 9.4(b)

Una porción de la salida impresa de *MS Excel* para el ejemplo 9.2

	Coefficients	Standard Error	t Stat	P-value
Intercept	118.763	9.207	12.899	0.000
Square Feet	6.270	0.725	8.645	0.000
Number of Floors	-16.203	6.212	-2.608	0.026
Bedrooms	-2.673	4.494	-0.595	0.565
Baths	30.271	6.849	4.420	0.001

El valor de R^2 ajustado

Observe de la definición de $R^2 = SSR/SS \text{ Total}$ que su valor nunca puede disminuir con la adición de más variables en el modelo de regresión. En consecuencia, R^2 puede estar artificialmente inflada por la inclusión de más y más variables predictoras.

Una medida alternativa de la fuerza del modelo de regresión se ajusta para grados de libertad con el uso de medias cuadráticas en lugar de sumas de cuadrados:

$$R^2(\text{adj}) = \left(1 - \frac{\text{MSE}}{\text{SS Total}/(n - 1)} \right) 100\%$$

TIP MI CONSEJO

Use $R^2(\text{adj})$ para comparar uno o más modelos posibles

Para los datos de bienes raíces de la figura 9.3,

$$R^2(\text{adj}) = \left(1 - \frac{46.9}{16\,382.2/14} \right) 100\% = 96.0\%$$

[†] Algunos paquetes usan el estadístico t que acabamos de describir, mientras que otros usan el estadístico F equivalente ($F = t^2$), puesto que el cuadrado de un estadístico t con ν grados de libertad es igual a un estadístico F con 1 gl en el numerador y ν grados de libertad en el denominador.

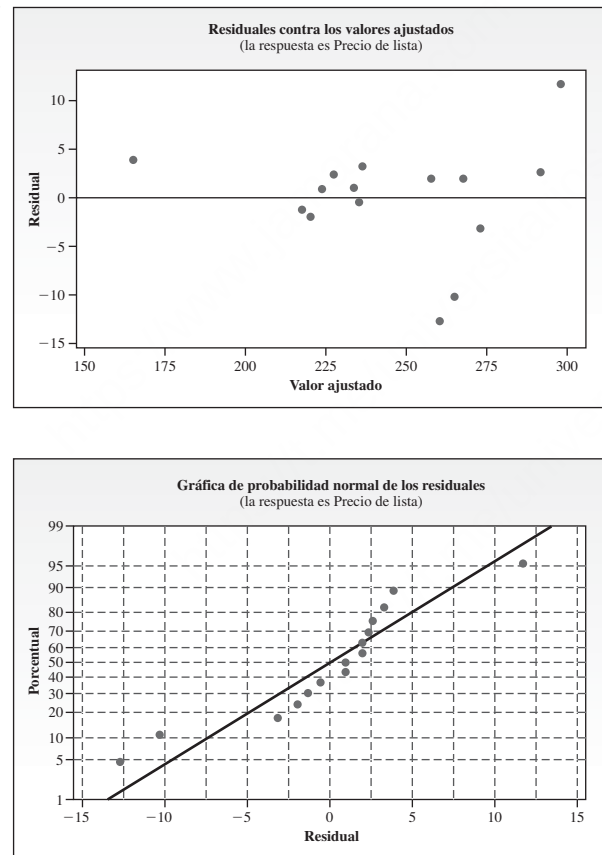
se encuentra en el primer renglón de la salida impresa de *MINITAB* y en el tercer renglón de la salida impresa de *Excel*. El valor “R-Sq(adj) = 96.0%” o “Adjusted R Square = 0.960” representan el porcentaje de variación en la respuesta y explicada por las variables independientes, corregida para grados de libertad. El valor ajustado de R^2 se usa principalmente para comparar dos o más modelos de regresión que usan números diferentes de variables predictoras independientes.

Comprobación de suposiciones de regresión

Antes de usar el modelo de regresión para su propósito principal, que es estimar y predecir y, deben verse **gráficas residuales** generadas por computadora para asegurarse que sean válidas todas las suposiciones de regresión. La *gráfica normal de probabilidad* y la *gráfica de residuales contra ajuste* se presentan en la figura 9.5 para los datos de bienes raíces. Parece haber tres observaciones que no se ajustan al patrón general. Se pueden ver como resultados atípicos en ambas gráficas. Es probable que estas tres observaciones deban investigarse, pero no dan fuerte evidencia de que las suposiciones se han violado.

FIGURA 9.5

Gráficas de diagnóstico



Uso del modelo de regresión para estimación y predicción

Finalmente, una vez que se haya determinado que el modelo es efectivo para describir la relación entre y y las variables predictoras $x_1, x_2, \dots, x_k$, el modelo se puede usar para estos fines:

- Estimar el valor promedio de y , $E(y)$, para valores dados de $x_1, x_2, \dots, x_k$
- Predecir un valor particular de y para valores dados de $x_1, x_2, \dots, x_k$

TIP MI CONSEJO

Para valores dados de $x_1, x_2, \dots, x_k$, el intervalo de predicción **siempre** será más ancho que el intervalo de confianza

Los valores de $x_1, x_2, \dots, x_k$ se introducen en la computadora y ésta genera el valor ajustado $\hat{y}$ junto con su error estándar estimado y los intervalos de confianza y predicción. Recuerde que el intervalo de predicción es *siempre más ancho* que el intervalo de confianza.

Veamos qué tan bien funciona nuestra predicción para los datos de bienes raíces, usando otra casa de la base de datos de la computadora, una casa con 1 000 pies cuadrados de superficie de vivienda, un piso, tres recámaras y dos baños, con precio de lista de 221 500 dólares. La salida impresa de la figura 9.6 muestra los intervalos de confianza y predicción para estos valores. El valor real cae dentro de ambos intervalos, lo que indica que el modelo funciona muy bien.

FIGURA 9.6

Intervalos de confianza y predicción para el ejemplo 9.2

Predicted Values for New Observations					
New Obs	Fit	SE Fit	95% CI	95% PI	
1	217.78	3.11	(210.86, 224.70)	(201.02,	234.54)

Values of Predictors for New Observations					
New Obs	Square Feet	Number of Floors	Bedrooms	Baths	
1	10.0	1.00	3.00	2.00	

9.4

UN MODELO DE REGRESIÓN POLINOMIAL

En la sección 9.3 explicamos en detalle las diversas partes de la salida impresa de regresión múltiple. Cuando se efectúa un análisis de regresión múltiple, se debe usar un método paso a paso:

1. Obtener el modelo de predicción ajustada.
2. Usar el análisis de varianza de la prueba F y R^2 para determinar qué tan bien se ajusta el modelo a los datos.
3. Verificar las pruebas t para los coeficientes de regresión parcial para ver cuáles están aportando información significativa en presencia de los otros.
4. Si se escoge comparar varios modelos diferentes, use $R^2(\text{adj})$ para comparar su efectividad.
5. Usar gráficas residuales generadas por computadora para ver si hay violación de las suposiciones de regresión.

Una vez que se han seguido todos estos pasos estamos listos para usar el modelo en la estimación y predicción.

Las variables predictoras $x_1, x_2, \dots, x_k$ empleadas en el modelo lineal general no tienen que representar variables predictoras *diferentes*. Por ejemplo, si se sospecha que una variable independiente x afecta la respuesta y , pero que la relación es *curvilínea* más que *lineal*, entonces se podría elegir ajustar a un **modelo cuadrático**:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$$

TIP MI CONSEJO

Una ecuación cuadrática es $y = a + bx + cx^2$. La gráfica forma una **parábola**.

El modelo cuadrático es un ejemplo de un **modelo de segundo orden** porque contiene un término cuyos exponentes suman 2 (en este caso, x^2).[†] También es un ejemplo de un **modelo polinomial**, un modelo que toma la forma

$$y = a + bx + cx^2 + dx^3 + \dots$$

Para ajustar este tipo de modelo usando el programa de regresión múltiple, los valores observados de y, x y x^2 se introducen en la computadora y la salida impresa se puede generar como en la sección 9.3.

[†] El *orden* de un término está determinado por la suma de los exponentes de variables que conforman ese término. Los términos que contienen x_1 o x_2 son de primer orden; los que contienen x_1^2, x_2^2 o $x_1 x_2$ son de segundo orden.

EJEMPLO

9.3

En un estudio de variables que afecta la productividad en el comercio de comestibles al menudeo, se utiliza un valor agregado por hora de trabajo para medir la productividad de tiendas de comestibles al menudeo.¹ Se define “valor agregado” como “el excedente [dinero generado por el negocio] disponible para pagar empleados, mobiliario y enseres, y equipo”. Los datos consistentes con la relación entre valor agregado por hora de trabajo y y el tamaño x de una tienda de comestibles descrita en este estudio, se muestran en la tabla 9.2 para 10 tiendas de alimentos ficticias. Elija un modelo para relacionar y con x .

TABLA 9.2

Datos sobre tamaño de tienda y valor agregado

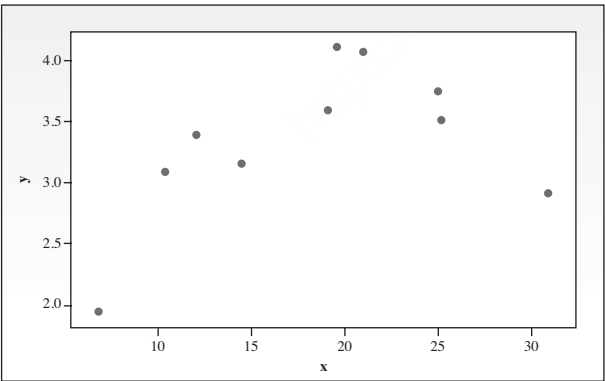
Tienda	Valor agregado por hora de trabajo, y (\$)	Tamaño de tienda (miles de pies cuadrados), x
1	4.08	21.0
2	3.40	12.0
3	3.51	25.2
4	3.09	10.4
5	2.92	30.9
6	1.94	6.8
7	4.11	19.6
8	3.16	14.5
9	3.75	25.0
10	3.60	19.1

Solución Se puede investigar la relación entre y y x al ver la gráfica de los puntos de datos en la figura 9.7. La gráfica sugiere que la productividad, y , aumenta cuando el tamaño de la tienda de comestibles, x , aumenta hasta alcanzar un tamaño óptimo. Arriba de ese tamaño, la productividad tiende a disminuir. La relación parece ser *curvilínea* y un modelo cuadrático,

$$E(y) = \beta_0 + \beta_1x + \beta_2x^2$$

FIGURA 9.7

Gráfica del tamaño de la tienda x y valor agregado y para el ejemplo 9.3



puede ser apropiado. Recuerde que, al elegir usar este modelo, no decimos que la verdadera relación sea cuadrática sino sólo que puede dar estimaciones y predicciones más precisas que, por ejemplo, un modelo lineal.

EJEMPLO

9.4

Consulte los datos sobre productividad y tamaño de una tienda minorista de alimentos del ejemplo 9.3. Se utilizó el *MINITAB* para ajustar un modelo cuadrático a los datos y para graficar la curva cuadrática de predicción, junto con los puntos de datos graficados. Analice lo adecuado del modelo ajustado.

Solución De la salida impresa de la figura 9.8 se puede ver que la ecuación de regresión es

$$\hat{y} = -.159 + .392x - .00949x^2$$

La gráfica de esta ecuación cuadrática junto con los puntos de datos se muestra en la figura 9.9.

FIGURA 9.8

Salida impresa *MINITAB* para el ejemplo 9.4

Análisis de regresión: y contra x, x-sq

The regression equation is
 $y = -0.159 + 0.392x - 0.00949x\text{-sq}$

Predictor	Coef	St Coef	T	P
Constant	-0.1594	0.5006	-0.32	0.760
x	0.39193	0.05801	6.76	0.000
x-sq	-0.009495	0.001535	-6.19	0.000

S = 0.250298 R-Sq = 87.9% R-Sq(adj) = 84.5%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	3.1989	1.5994	25.53	0.001
Residual Error	7	0.4385	0.0626		
Total	9	3.6374			

Source	DF	Seq SS
x	1	0.8003
x-sq	1	2.3986

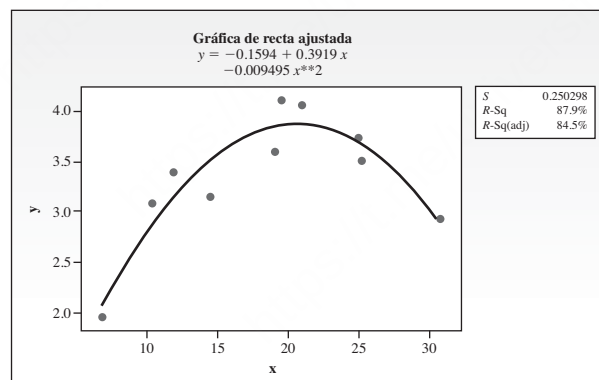
TIP

MI CONSEJO

Vea la salida impresa y encuentre las leyendas para "Predictor". Esto le dirá cuáles variables se han usado en el modelo

FIGURA 9.9

Recta de regresión cuadrática ajustada para el ejemplo 9.4



Para evaluar lo adecuado del modelo cuadrático, la prueba de

$$H_0 : \beta_1 = \beta_2 = 0$$

contra

$$H_a : \text{Ni } \beta_1 \text{ ni } \beta_2 \text{ son } 0$$

se da en la salida impresa como

$$F = \frac{MSR}{MSE} = 25.53$$

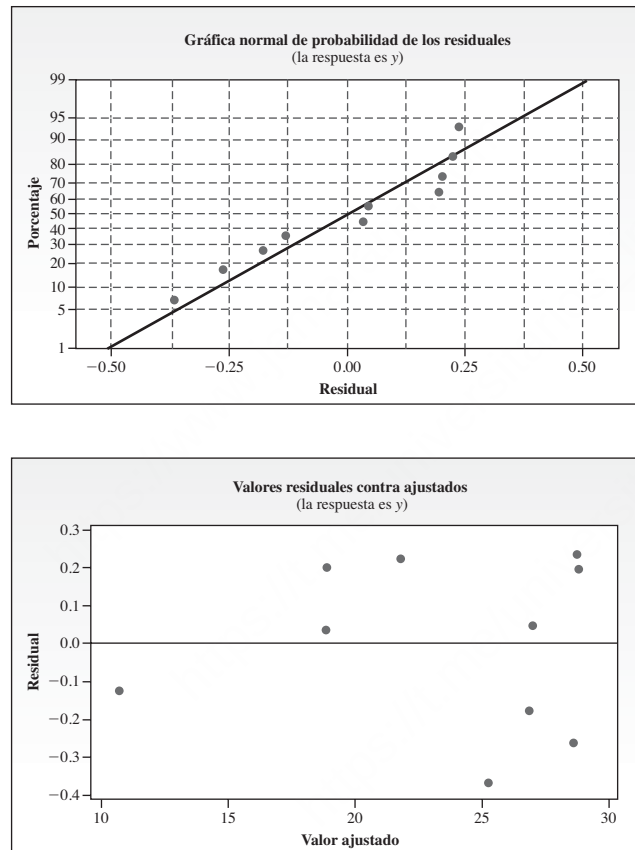
con valor $p = .001$. En consecuencia, el ajuste total del modelo es altamente significativo. La regresión cuadrática explica $R^2 = 87.9\%$ de la variación en y [$R^2(\text{adj}) = 84.5\%$].

De las pruebas t para las variables individuales del modelo se puede ver que b_1 y b_2 son altamente significativas, con valores p iguales a .000. Observe de la sección de suma secuencial de cuadrados que la suma de cuadrados para regresión lineal es .8003, con una suma de cuadrados adicional de 2.3986 cuando se agregue el término cuadrático. Es evidente que el modelo de regresión lineal simple es inadecuado para describir los datos.

Una última mirada a las gráficas residuales en la figura 9.10 asegura que las suposiciones de regresión sean válidas. Observe el aspecto relativamente lineal de la gráfica normal y la relativa dispersión de los residuales contra los ajustes. El modelo cuadrático da predicciones precisas para valores de x que se encuentren *dentro del rango de los valores muestreados de x* .

FIGURA 9.10

Gráficas de diagnóstico para el ejemplo 9.4



9.4 EJERCICIOS

TÉCNICAS BÁSICAS

9.1 Suponga que $E(y)$ está relacionada con dos variables predictoras, x_1 y x_2 , por la ecuación

$$E(y) = 3 + x_1 - 2x_2$$

- a. Grafique la relación entre $E(y)$ y x_1 cuando $x_2 = 2$.
Repita para $x_2 = 1$ y para $x_2 = 0$.

- b. ¿Qué relación tienen entre sí las rectas de la parte a?

9.2 Consulte el ejercicio 9.1.

- a. Grafique la relación entre $E(y)$ y x_2 cuando $x_1 = 0$.
Repita para $x_1 = 1$ y para $x_1 = 2$.

- b. ¿Qué relación tienen entre sí las rectas de la parte a?
- c. Suponga, en una situación práctica, que se desea modelar la relación entre $E(y)$ y dos variables predictoras x_1 y x_2 . ¿Cuál es la implicación de usar el modelo de primer orden $E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$?

9.3 Suponga que se ajusta el modelo

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

a 15 puntos de datos y se encuentra que F es igual a 57.44.

- a. ¿Los datos dan suficiente evidencia para indicar que el modelo aporta información para la predicción de y ? Pruebe usando un nivel de significancia de 5%.

- b. Use el valor de F para calcular R^2 . Interprete su valor.

9.4 La salida impresa de computadora para el análisis de regresión múltiple para el ejercicio 9.3 da esta información:

$$\begin{array}{ll} b_0 = 1.04 & b_1 = 1.29 \\ & SE(b_1) = .42 \\ b_2 = 2.72 & b_3 = .41 \\ SE(b_2) = .65 & SE(b_3) = .17 \end{array}$$

- a. ¿Cuál, si hay alguna, de las variables independientes x_1 , x_2 y x_3 aportan información para la predicción de y ?
- b. Dé la ecuación de predicción de mínimos cuadrados.
- c. En la misma hoja de papel, grafique y contra x_1 cuando $x_2 = 1$ y $x_3 = 0$ y cuando $x_2 = 1$ y $x_3 = .5$.
- d. ¿Qué relación tienen entre sí las dos rectas?
- e. ¿Cuál es la interpretación práctica del parámetro β_1 ?

9.5 Suponga que se ajusta el modelo

$$E(y) = \beta_0 + \beta_1 x + \beta_2 x^2$$

a 20 puntos de datos y se obtiene la salida impresa MINITAB siguiente:

Salida impresa MINITAB para el ejercicio 9.5

Análisis de regresión: y contra x , x -sq

The regression equation is
 $y = 10.6 + 4.44 x - 0.648 x\text{-sq}$

Predictor	Coef	SE Coef	T	P
Constant	10.5638	0.6951	15.20	0.000
x	4.4366	0.5150	8.61	0.000
x-sq	-0.64754	0.07988	-8.11	0.000

S = 1.191 R-Sq = 81.5% R-Sq(adj) = 79.3%

Source	DF	SS	MS	F	P
Regression	2	106.072	53.036	37.37	0.000
Residual Error	17	24.128	1.419		
Total	19	130.200			

- a. ¿Qué tipo de modelo se ha elegido para ajustar los datos?
- b. ¿Qué tan bien ajusta los datos el modelo? Explique.
- c. ¿Los datos dan suficiente evidencia para indicar que el modelo aporta información para la predicción de y ? Use el método del valor p .

9.6 Consulte el ejercicio 9.5.

- a. ¿Cuál es la ecuación de predicción?

- b. Grafique la ecuación de predicción sobre el intervalo $0 \leq x \leq 6$.

9.7 Consulte el ejercicio 9.5.

- a. ¿Cuál es su estimación del valor promedio de y cuando $x = 0$?

- b. ¿Los datos dan suficiente evidencia para indicar que el valor promedio de y difiere de 0 cuando $x = 0$?

9.8 Consulte el ejercicio 9.5.

- a. Suponga que la relación entre $E(y)$ y x es una recta. ¿Qué se sabría acerca del valor de β_2 ?

- b. ¿Los datos dan suficiente evidencia para indicar curvatura en la relación entre y y x ?

9.9 Consulte el ejercicio 9.5. Suponga que y es la utilidad para algún negocio y x es la cantidad de capital invertido, y se sabe que la tasa de aumento en utilidad para un aumento unitario en capital invertido sólo puede disminuir cuando x aumenta. Se desea saber si los datos dan suficiente evidencia para indicar una tasa decreciente de aumento en utilidad cuando la cantidad de capital invertido aumenta.

- a. Las circunstancias descritas implican una prueba estadística de una cola. ¿Por qué?
- b. Realice la prueba al nivel de significancia de 1%. Exprese sus conclusiones.

APLICACIONES



9.10 Libros de texto universitarios Un editor de libros de texto universitarios realizó un estudio para relacionar la utilidad por texto y , con el costo de ventas x , en un periodo de 6 años cuando su fuerza de ventas (y costos de ventas) estaban creciendo rápidamente. Se recolectaron estos datos de inflación ajustada (en miles de dólares):

Utilidad por texto, y	16.5	22.4	24.9	28.8	31.5	35.8
Costo de ventas por texto, x	5.0	5.6	6.1	6.8	7.4	8.6

Esperando que la utilidad por libro subiera y luego se nivelara, el editor ajustó el modelo $E(y) = \beta_0 + \beta_1 x + \beta_2 x^2$ a los datos.

Salida de Excel para el ejercicio 9.10

SUMMARY OUTPUT

Regression Statistics

Multiple R	0.9978
R Square	0.9955
Adjusted R Square	0.9925
Standard Error	0.5944
Observations	6

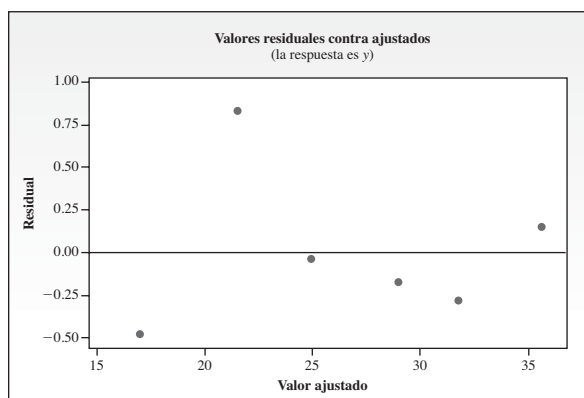
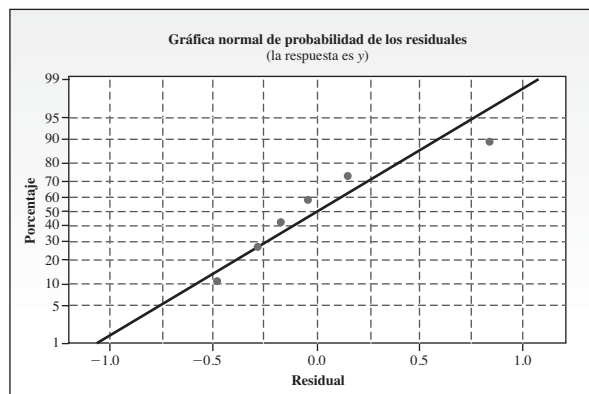
ANOVA

	df	SS	MS	F	Significance F
Regression	2	234.995	117.478	332.528	0.000
Residual	3	1.060	0.353		
Total	5	236.015			

	Coefficients	Standard Error	t Stat	P-value
Intercept	-44.192	8.287	-5.333	0.013
x	16.334	2.490	6.560	0.007
x-sq	-0.820	0.182	-4.494	0.021

- Grafique los puntos de datos. ¿Le parece que el modelo cuadrático es necesario?
- Encuentre s en la salida impresa. Confirme que
$$s = \sqrt{\frac{\text{SSE}}{n - k - 1}}$$
- ¿Los datos dan suficiente evidencia para indicar que el modelo contribuye con información para la predicción de y ? ¿Cuál es el valor p para esta prueba y qué significa?
- ¿Qué signo se esperaría que tenga el valor real de β_2 ? Encuentre el valor de β_2 en la salida impresa. ¿Este valor confirma las expectativas?
- ¿Los datos indican una curvatura significativa en la relación entre y y x ? Pruebe al nivel de significancia de 5%.
- ¿Qué conclusiones se pueden sacar de las gráficas residuales siguientes?

Gráficas de diagnóstico para el ejercicio 9.10



9.11 Libros de texto universitarios II Consulte el ejercicio 9.10.

- Use los valores de SSR y SS Total para calcular R^2 . Compare este valor con el valor dado en la salida impresa.
- Calcule $R^2(\text{adj})$. ¿Cuándo sería apropiado usar este valor en lugar de R^2 para evaluar el ajuste del modelo?
- El valor de $R^2(\text{adj})$ fue de 95.66% cuando un modelo lineal simple se ajustó a los datos. ¿El modelo lineal o el cuadrático ajustan mejor?

DATOS

EX0912

9.12 Elegir una buena cámara Las cámaras vienen con muchas opciones, y al parecer entre más de éstas se desean el costo de la cámara es mayor. *Consumer Reports*² ha calificado $n = 20$ cámaras en cualidades que los consumidores buscan. Las variables que pueden relacionarse con el costo de una cámara se dan en la siguiente tabla, donde y = puntuación general del cliente, x_1 = megapíxeles, x_2 = peso (onzas), x_3 = zoom óptico, x_4 = ángulo más amplio y x_5 = duración de la batería (fotos).

Cámara	Costo	y	x_1	x_2	x_3	x_4	x_5
Nikon Coolpix S8000	300	65	14	6	10.0	30	210
Canon PowerShot SD1400 IS Elph	250	61	14	5	4.0	28	230
Sony Cyber-shot DSC-HX5V	350	60	10	7	10.0	25	310
Sony Cyber-shot DSC-TX7	400	59	10	5	4.0	25	230
Panasonic Lumix DMC-FP1	150	59	12	6	4.0	35	300
Canon PowerShot A3100 IS	180	58	12	6	4.0	35	240
Sony Cyber-shot DSC-W380	140	57	14	4	5.0	24	220
Nikon Coolpix S70	300	57	12	6	5.0	28	200
Pentax Optio I-10	300	56	12	6	5.0	28	250
Olympus Stylus 5010	200	55	14	5	5.0	26	120
Olympus Stylus 7040	250	55	14	5	7.0	28	120
Sony Cyber-shot DSC-W350	200	55	14	4	4.0	26	220
Canon PowerShot A490	110	55	10	7	3.0	37	150
Canon PowerShot A945	130	54	10	7	3.0	37	150
Casio Exilim EX-G1	250	54	12	5	3.0	38	300
Samsung TL210	230	54	12	5	5.0	27	200
Olympus Stylus Tough-3000	230	48	12	6	3.6	28	160
Sony Cyber-shot DSC-W310	150	47	12	5	4.0	28	220
Olympus FE-47	120	47	14	7	5.0	36	120
Sony Cyber-shot DSC-S2100	120	46	12	7	3.0	35	170

La salida impresa de *MINITAB* que se presenta a continuación muestra la regresión de y en las variables predictoras x_1 a x_5 .

Salida de *MINITAB* para el ejercicio 9.12

Análisis de regresión: y contra x1, x2, x3, x4, x5

The regression equation is
 $y = 45.6 - 0.0583 x_1 - 0.83 x_2 + 1.18 x_3 + 1.129 x_4 + 0.0277 x_5$

Predictor	Coef	SE Coef	T	P
Constant	45.584	8.664	5.26	0.000
x1	-0.05832	0.04527	-1.29	0.219
x2	-0.835	1.393	-0.60	0.559
x3	1.1773	0.5994	1.96	0.070
x4	0.1286	0.3040	0.42	0.679
x5	0.02773	0.01664	1.67	0.118

S = 4.09203 R-Sq = 50.3% R-Sq(adj) = 32.6%

Source	DF	SS	MS	F	P
Regression	5	237.37	47.47	2.84	0.057
Residual Error	14	234.43	16.74		
Total	19	471.80			

Source	DF	Seq SS
x1	1	86.68
x2	1	1.69
x3	1	95.43
x4	1	7.05
x5	1	46.51

- Escriba un modelo de regresión múltiple usando cada una de las variables x como variables independientes y y como la variable de respuesta.
- Comente sobre el ajuste del modelo usando la prueba estadística para el ajuste general y el coeficiente de determinación, R^2 .
- Si reajustara el modelo, eliminando una de las variables predictoras, ¿cuál elegiría? ¿Por qué?

9.13 Elegir una buena cámara II Consulte el ejercicio 9.12. Un comando en el menú de regresión de *MINITAB* proporciona una salida en la que se calculan R^2 y $R^2(\text{adj})$ para todos los subconjuntos posibles de las cinco variables independientes. Aquí se proporciona la salida.

Salida de *MINITAB* para el ejercicio 9.13

Regresión de mejores subconjuntos: y contra x1, x2, x3, x4, x5

Vars	R-Sq	R-Sq (adj)	Mallows C-p	S	x	x	x	x	x
1	26.4	22.3	4.7	4.3928					
1	20.7	16.3	6.3	4.5577					
2	41.0	34.1	2.6	4.0449					
2	37.4	30.1	3.6	4.1674					
3	49.0	39.5	2.4	3.8766					
3	43.5	32.9	3.9	4.0807					
4	49.7	36.3	4.2	3.9784					
4	49.0	35.4	4.4	4.0037					

- Si tuviera que comparar estos modelos y elegir el mejor, ¿cuál sería? Explique.
- Comente sobre la utilidad del modelo que eligió en la parte a. ¿Su modelo es valioso para predecir la puntuación general con base en las variables predictoras elegidas?

DATOS
EX0914

9.14 Utilidades corporativas Para estudiar la relación de publicidad e inversión de capital con utilidades corporativas, los datos siguientes, registrados en unidades de 100 000 dólares, se recolectaron para 10 empresas de mediano tamaño en el mismo año. La variable y representa utilidad para el año, x_1 representa inversión de capital y x_2 representa gasto en publicidad.

y	x_1	x_2	y	x_1	x_2
15	25	4	1	20	0
16	1	5	16	12	4
2	6	3	18	15	5
3	30	1	13	6	4
12	29	2	2	16	2

- Usando el modelo

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

y un paquete apropiado de software, encuentre la ecuación de predicción de mínimos cuadrados para estos datos.

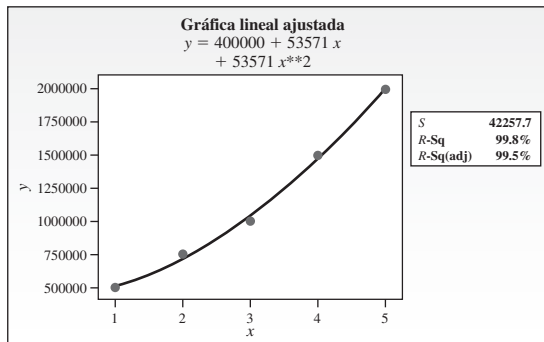
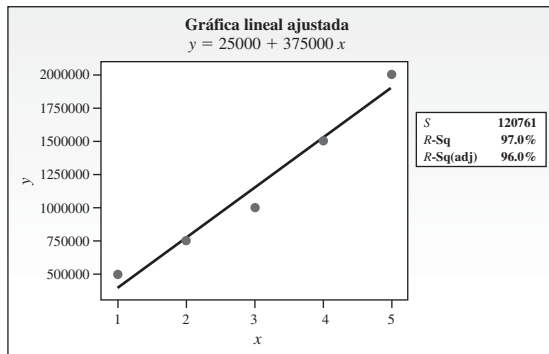
- Use la prueba F general para determinar si el modelo aporta información significativa para la predicción de y . Use $\alpha = .01$.
- ¿El gasto en publicidad x_2 aporta información significativa para la predicción de y , dado que x_1 ya está en el modelo? Use $\alpha = .01$.
- Calcule el coeficiente de determinación, R^2 . ¿Qué porcentaje de la variación general está explicado por el modelo?

DATOS
EX0915

9.15 ¿La nueva Ruta 66? Una de las autopistas nacionales más famosas desde sus inicios en la década de 1920 hasta su cierre alrededor de 1970 fue la Ruta 66. En su trayecto de Chicago a Los Ángeles, una de las últimas paradas era San Bernardino, California. Ahora, un nuevo núcleo de transportación, el Aeropuerto Internacional de San Bernardino, está en las etapas de planeación. Aunque no se han firmado acuerdos con ninguna aerolínea, los funcionarios pronostican 2 millones de pasajeros para 2030, como se muestra en la siguiente tabla.³

Año	2010	2015	2020	2025	2030
Año codificado, x	1	2	3	4	5
Número de viajeros, y	500 000	750 000	1 000 000	1 500 000	2 000 000

A continuación se presentan las gráficas ajustadas lineal y cuadrática para estos datos.



- Con base en las estadísticas de resumen en las gráficas lineales, ¿cuál de los dos modelos se ajusta mejor a los datos?
- Escriba la ecuación teórica para el modelo cuadrático.
- Use la siguiente salida impresa para determinar si el modelo cuadrático contribuye con información significativa para la predicción de y .
- Use la siguiente salida impresa para determinar si el término cuadrático contribuye con información significativa para la predicción de y , en presencia del término lineal.

Análisis de regresión: y contra x , x -sq

The regression equation is
 $y = 400000 + 53571x + 53571x\text{-sq}$

Predictor	Coef	SE Coef	T	P
Constant	400000	90633	4.41	0.048
x	53571	69068	0.78	0.519
x-sq	53571	11294	4.74	0.042

S = 42257.7 R-Sq = 99.8% R-Sq(adj) = 99.5%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	1.44643E+12	7.23214E+11	405.00	0.002
Residual Error	2	3571428571	1785714286		
Total	4	1.45000E+12			

USO DE VARIABLES PREDICTORAS CUANTITATIVAS Y CUALITATIVAS EN UN MODELO DE REGRESIÓN

9.5

Una razón por la que los modelos de regresión múltiple son muy flexibles es que permiten el uso de variables predictoras *cualitativas* y *cuantitativas*. Para los métodos de regresión múltiple que se usan en este capítulo, la variable de respuesta *y* debe ser *cuantitativa*, midiendo una variable aleatoria numérica que tiene una distribución normal (de acuerdo con las suposiciones de la sección 9.2). No obstante, cada variable independiente predictora puede ser *cuantitativa* o *cualitativa*, cuyos niveles representan cualidades o características y sólo pueden ser categorizados.

Las variables cuantitativas y cualitativas introducen el modelo de regresión en formas diferentes. Para complicar las cosas, podemos permitir una combinación de tipos diferentes de variables en el modelo y podemos permitir que las variables *interactúen*, un concepto que puede serle familiar a usted en el *experimento factorial* del capítulo 7. Consideremos estas opciones una por una.

Una **variable cuantitativa** x puede introducirse como término lineal, x , o elevado a alguna potencia de grado superior como x^2 o x^3 , como en el modelo cuadrático del ejemplo 9.3. Cuando sea necesaria más de una variable cuantitativa, la interpretación de los posibles modelos se hace más complicada. Por ejemplo, con dos variables cuantitativas x_1 y x_2 , se puede usar un **modelo de primer orden** como lo es

$$E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2$$

TIP MI CONSEJO

Introduzca variables
cuantitativas como

- una sola x
- una potencia de orden superior, x^2 o x^3
- una interacción con otra variable

que traza un plano en espacio tridimensional (véase la figura 9.1). No obstante, puede ser que una de las variables, por ejemplo x_2 , no esté relacionada con y en la misma forma cuando $x_1 = 1$ como lo está cuando $x_1 = 2$. Para permitir que x_2 se comporte de manera diferente dependiendo del valor de x_1 , sumamos un **término de interacción**, x_1x_2 , y permitimos que el plano bidimensional *se tuerza*. El modelo es ahora un **modelo de segundo orden**:

$$E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2$$

Los modelos se hacen complicados más rápidamente cuando se permiten relaciones curvilíneas e interacción para las dos variables. Una forma de decidir sobre el tipo de modelo que se necesita es graficar algunos de los datos, quizá y contra x_1 , y contra x_2 , y y contra x_2 para varios valores de x_1 .

En contraste con variables predictoras cuantitativas, las **variables predictoras cualitativas** se introducen en un modelo de regresión a través de **variables ficticias** o **indicadoras**.

Por ejemplo, en un modelo que relacione el salario medio de un grupo de empleados con varias variables predictoras, se puede incluir el antecedente étnico del empleado. Si cada empleado incluido en su estudio pertenece a uno de tres grupos étnicos, A, B o C, se puede introducir la variable cualitativa “etnia” en el modelo usando dos *variables ficticias*:

$$x_1 = \begin{cases} 1 & \text{si grupo B} \\ 0 & \text{si no} \end{cases} \quad x_2 = \begin{cases} 1 & \text{si grupo C} \\ 0 & \text{si no} \end{cases}$$

Vea el efecto que estas dos variables tienen en el modelo $E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2$: para empleados del grupo A,

$$E(y) = \beta_0 + \beta_1(0) + \beta_2(0) = \beta_0$$

para empleados del grupo B,

$$E(y) = \beta_0 + \beta_1(1) + \beta_2(0) = \beta_0 + \beta_1$$

y para los del grupo C,

$$E(y) = \beta_0 + \beta_1(0) + \beta_2(1) = \beta_0 + \beta_2$$

TIP MI CONSEJO

Las variables cualitativas se introducen como variables ficticias, una menos que el número de categorías o niveles

El modelo permite una respuesta promedio diferente para cada grupo. β_1 mide la diferencia en las respuestas promedio entre los grupos B y A, en tanto que β_2 mide la diferencia entre los grupos C y A.

Cuando una variable cualitativa contiene k categorías o niveles, $(k - 1)$ variables ficticias deben agregarse al modelo de regresión. Este modelo puede contener otras variables predictoras, cuantitativas o cualitativas, así como productos cruzados (**interacciones**) de las variables ficticias con otras variables que aparecen en el modelo. Como se puede ver, el proceso de construir un modelo, es decir, decidir sobre los términos apropiados a introducir en el modelo de regresión, puede ser bastante complicado. No obstante, usted puede tener más habilidad para construir un modelo, adquiriendo experiencia con los ejercicios del capítulo. El ejemplo siguiente contiene una variable cuantitativa y una cualitativa que interactúan.

EJEMPLO

9.5

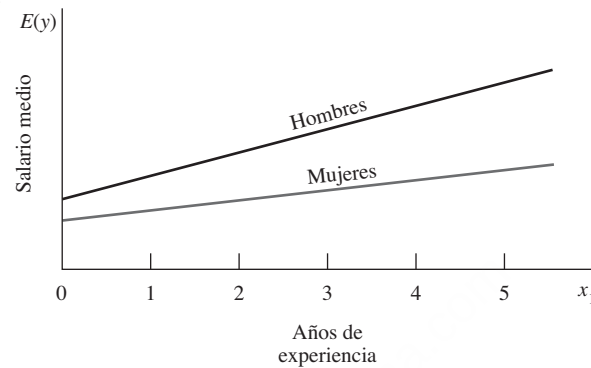
Se realizó un estudio para examinar la relación entre salario en una universidad, y , el número de años de experiencia del miembro del profesorado y el género del miembro del profesorado. Si se espera que haya una relación de línea recta entre salario medio y años de experiencia para caballeros y mujeres, escriba el modelo que relacione salario medio con las dos variables predictoras: años de experiencia (cuantitativa) y género del profesor (cualitativa).

Solución Dado que se puede sospechar que las rectas de salario medio para caballeros y mujeres son diferentes, su modelo para el salario medio $E(y)$ puede aparecer como se muestra en la figura 9.11. Una relación de línea recta entre $E(y)$ y años de experiencia x_1 implica el modelo

$$E(y) = \beta_0 + \beta_1 x_1 \quad (\text{se grafica como línea recta})$$

FIGURA 9.11

Relación hipotética para salario medio $E(y)$, años de experiencia (x_1), y género (x_2) para el ejemplo 9.5



La variable cualitativa “género” contiene $k = 2$ categorías, caballeros y mujeres. Por lo tanto, se necesita $(k - 1) = 1$ variable ficticia, x_2 , definida como

$$x_2 = \begin{cases} 1 & \text{si es hombre} \\ 0 & \text{si es mujer} \end{cases}$$

y el modelo se expande para convertirse en

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad (\text{se grafica como dos rectas paralelas})$$

El hecho de que las pendientes de las dos rectas puedan diferir significa que las dos variables predictoras **interactúan**; esto es, el cambio en $E(y)$ correspondiente a un cambio en x_1 depende de si el profesor es hombre o mujer. Para tomar en cuenta esta interacción (diferencia en pendientes), el término de interacción $x_1 x_2$ se introduce en el modelo. El modelo completo que caracteriza la gráfica de la figura 9.11 es

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$$

Variable ficticia
para género
↓
 ↑ ↑
 Años de Interacción
experiencia

donde

$$x_1 = \text{Años de experiencia}$$

$$x_2 = \begin{cases} 1 & \text{si es hombre} \\ 0 & \text{si es mujer} \end{cases}$$

Se puede ver cómo funciona el modelo al asignar valores a la variable ficticia x_2 . Cuando la profesora es mujer, el modelo es

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2(0) + \beta_3 x_1(0) = \beta_0 + \beta_1 x_1$$

que es una recta con pendiente β_1 y punto de cruce β_0 . Cuando el profesor es hombre, el modelo es

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2(1) + \beta_3 x_1(1) = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) x_1$$

que es una recta con pendiente $(\beta_1 + \beta_3)$ e intersección $(\beta_0 + \beta_2)$. Las dos rectas tienen *pendientes diferentes e intersecciones diferentes*, lo cual permite que la relación entre salario y años de experiencia x_1 se comporte de manera diferente para hombres y mujeres.

EJEMPLO

9.6

Se seleccionaron muestras aleatorias de seis mujeres y seis caballeros profesores auxiliares, de entre los profesores auxiliares de una universidad de artes y ciencias. Los datos sobre el salario y años de experiencia se muestran en la tabla 9.3. Observe que cada una de las dos muestras (hombre y mujer) contenía dos profesores con tres años de experiencia, pero ninguno tenía 2 años de experiencia. Interprete la salida impresa de regresión de *MS Excel* y grafique las rectas de salario predichas.

TABLA 9.3**Salario contra género y años de experiencia**

Años de experiencia, x_1	Salario para caballeros, y (\$)	Salario para mujeres, y (\$)
1	60 710	59 510
2	—	60 440
3	63 160	61 340
3	63 210	61 760
4	64 140	62 750
5	65 760	63 200
5	65 590	—

Solución La salida impresa de regresión de *Excel* para los datos de la tabla 9.3 se muestra en la figura 9.12. Se puede usar un método paso a paso para interpretar este análisis de regresión, empezando con una ecuación de predicción ajustada, $\hat{y} = 58\,593 + 969x_1 + 866.71x_2 + 260.13x_1x_2$. Al sustituir $x_2 = 0$ o 1 en esta ecuación, se obtienen dos rectas, una para mujeres y una para caballeros, para predecir el valor de y para una x_1 determinada. Estas rectas son

$$\text{Mujeres: } \hat{y} = 58\,593 + 969x_1$$

$$\text{Caballeros: } \hat{y} = 59\,459.71 + 1\,229.13x_1$$

y se grafican en la figura 9.13.

FIGURA 9.12

Salida *MS Excel* para el ejemplo 9.6

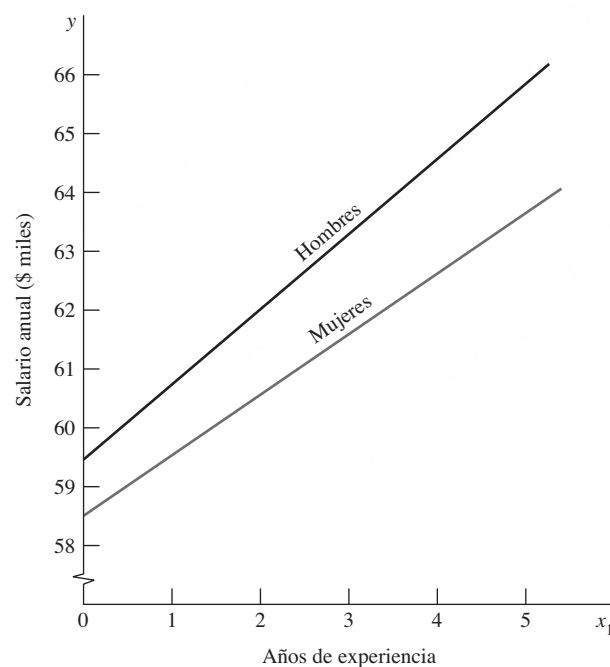
SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.9962
R Square	0.9924
Adjusted R Square	0.9895
Standard Error	201.3438
Observations	12

ANOVA					
	<i>gl</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	3	42108777.03	14036259.01	346.238	0.000
Residual	8	324314.64	40539.330		
Total	11	42433091.67			
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	
Intercept	58593	207.9470	281.7689	0.000	
x1	969	63.6705	15.2190	0.000	
x2	866.710	305.2568	2.8393	0.022	
x1x2	260.130	87.0580	2.9880	0.017	

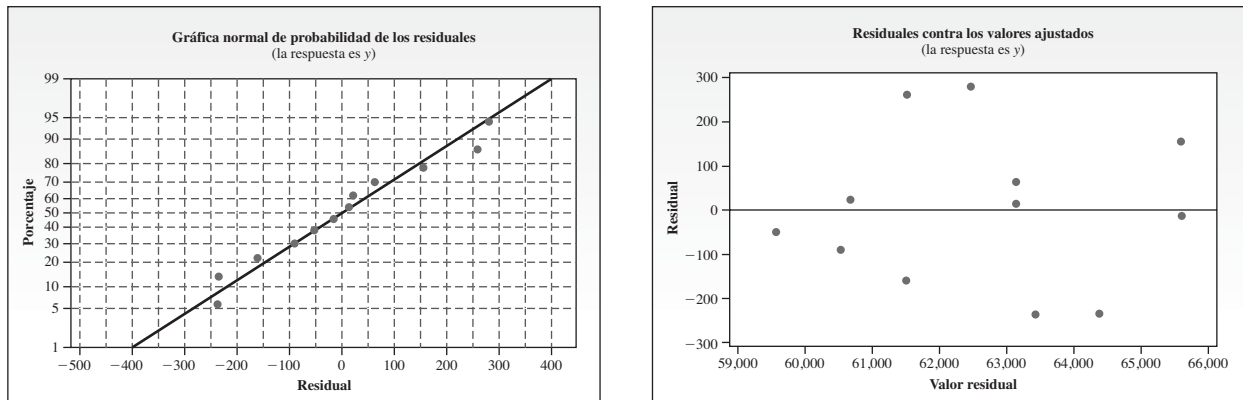
A continuación, considere el ajuste general del modelo usando el análisis de varianza de la prueba F . Como el estadístico de prueba observado en la parte ANOVA de la salida impresa es $F = 346.238$ con valor p (“Significance F”) igual a .000, se puede concluir que al menos una de las variables de pronóstico está aportando información para la predicción de y . La fuerza de este modelo es medida aun más por el coeficiente de determinación, $R^2 = 99.24\%$. Se puede ver que el modelo parece ajustar muy bien.

FIGURA 9.13

Gráfica de las rectas de predicción de salario del profesorado, para el ejemplo 9.6



Para explorar con más detalle el efecto de las variables predictoras, vea las pruebas t individuales para las tres variables predictoras. Los valores p para estas pruebas, es decir, .000, .022 y .017, respectivamente, son todas significativas, lo cual quiere decir que todas las variables predictoras agregan información significativa a la predicción *con las otras dos variables ya en el modelo*. Por último, verifique las gráficas de diagnóstico para asegurarse de que no haya grandes violaciones de las suposiciones de regresión. Estas gráficas, que se comportan como se esperaba para un modelo bien ajustado, se ven en la figura 9.14.

FIGURA 9.14Gráficas de diagnóstico
para el ejemplo 9.6**EJEMPLO****9.7**

Consulte el ejemplo 9.6. ¿Los datos dan suficiente evidencia para indicar que la tasa anual de aumento, en salarios de profesores jóvenes, es mayor que la tasa anual de aumento en salarios de profesoras jóvenes? Es decir, ¿los datos dan suficiente evidencia para indicar que la pendiente de la recta de salarios de profesores es mayor que la pendiente de la recta de salarios de profesoras?

Solución Como β_3 mide la diferencia en pendientes, las pendientes de las dos rectas serán idénticas si $\beta_3 = 0$. Por tanto, se desea probar la hipótesis nula

$$H_0 : \beta_3 = 0$$

esto es, las pendientes de las dos rectas son idénticas, contra la hipótesis alternativa

$$H_a : \beta_3 > 0$$

esto es, la pendiente de la recta de salarios de profesores es mayor que la pendiente de la recta de salarios de profesoras.

El valor calculado de t correspondiente a β_3 , que se ve en el renglón marcado “x1x2” en la figura 9.12, es 2.988. Como la salida de regresión de *Excel* da valores p para dos pruebas de significancia de dos colas, el valor p en la salida impresa, .017, es *el doble* de lo que sería para una prueba de una cola. Para esta prueba de una cola, el valor p es $.017/2 = .0085$ y la hipótesis nula es rechazada. Hay suficiente evidencia para indicar que la tasa anual de aumento en salarios de profesores excede la tasa para mujeres.[†]

[†] Si desea determinar si los datos dan suficiente evidencia para indicar que los profesores empiezan con salarios más altos, probaría $H_0 : \beta_2 = 0$ contra la hipótesis alternativa $H_a : \beta_2 > 0$.

9.5 EJERCICIOS

TÉCNICAS BÁSICAS

9.16 Rendimiento en producción Suponga que se desea predecir el rendimiento en producción y como función de varias variables independientes de predicción. Indique si cada una de las siguientes variables independientes es cualitativa o cuantitativa. Si es cualitativa, defina la(s) variable(s) ficticia(s) apropiada(s).

- La tasa de interés prevaleciente en la región.
- El precio por kilogramo de un artículo empleado en el proceso de producción.
- La planta (A, B o C) en la que se mide el rendimiento en producción.
- El tiempo que la máquina de producción haya estado en operación.
- El turno (de noche o de día) en el que se mide el rendimiento.

9.17 Suponga que $E(y)$ está relacionada con dos variables predictoras x_1 y x_2 por la ecuación

$$E(y) = 3 + x_1 - 2x_2 + x_1x_2$$

- Grafique la relación entre $E(y)$ y x_1 cuando $x_2 = 0$. Repita para $x_2 = 2$ y para $x_2 = -2$.
- Repita las instrucciones de la parte **a** para el modelo $E(y) = 3 + x_1 - 2x_2$
- Observe que la ecuación para la parte **a** es exactamente igual a la ecuación de la parte **b**, excepto que hemos agregado el término x_1x_2 . ¿En qué forma la adición del término x_1x_2 afecta las gráficas de las tres rectas?
- ¿Qué flexibilidad se agrega al modelo de primer orden $E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2$ por la adición del término $\beta_3x_1x_2$, usando el modelo $E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2$?

9.18 Un modelo de regresión lineal múltiple que contiene una variable independiente cualitativa y una cuantitativa produjo esta ecuación de predicción:

$$\hat{y} = 12.6 + .54x_1 - 1.2x_1x_2 + 3.9x_2^2$$

- ¿Cuál de las dos variables es la cuantitativa? Explique.
- Si x_1 puede tomar sólo los valores 0 o 1, encuentre las dos posibles ecuaciones de predicción para este experimento.
- Grafique las dos ecuaciones halladas en la parte **b**. Compare las formas de las dos curvas.

APLICACIONES

DATOS
EX0919

9.19 Menos carne roja Un cambio deseable en la dieta si desea “comer en forma correcta” es reducir la ingesta de carne roja y sustituirla por pollo o pescado. Unos investigadores dieron seguimiento al consumo de carne roja y de pollo, y (en libras anuales por persona) y encontraron que el consumo de carne roja disminuyó y el de pollo aumentó en un periodo de siete años. En la tabla siguiente se ve un resumen de sus datos.

Año	Carne roja	Pollo
1	85	37
2	89	36
3	76	47
4	76	47
5	68	62
6	67	74
7	60	79

Considere ajustar el siguiente modelo, que toma en cuenta ajustar simultáneamente dos rectas de regresión lineal simples:

$$E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2$$

donde y es el consumo anual de carne (roja o de pollo) por persona

$$x_1 = \begin{cases} 1 & \text{si es carne roja} \\ 0 & \text{si es pollo} \end{cases} \quad \text{y} \quad x_2 = \text{Año}$$

Salida MINITAB para el ejercicio 9.19

Análisis de regresión: y contra x1, x2, x1x2

The regression equation is
 $y = 23.6 + 69.0 \, x_1 + 7.75 \, x_2 - 12.3 \, x_1x_2$

Predictor	Coef	SE Coef	T	P
Constant	23.571	3.522	6.69	0.000
x1	69.000	4.981	13.85	0.000
x2	7.7500	0.7875	9.84	0.000
x1x2	-12.286	1.114	-11.03	0.000

S = 4.16705 R-Sq = 95.4% R-Sq(adj) = 94.1%

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	3	3637.9	1212.6	69.83	0.000
Residual Error	10	173.6	17.4		
Total	13	3811.5			

Source	DF	Seq SS
x1	1	1380.1
x2	1	144.6
x1x2	1	2113.1

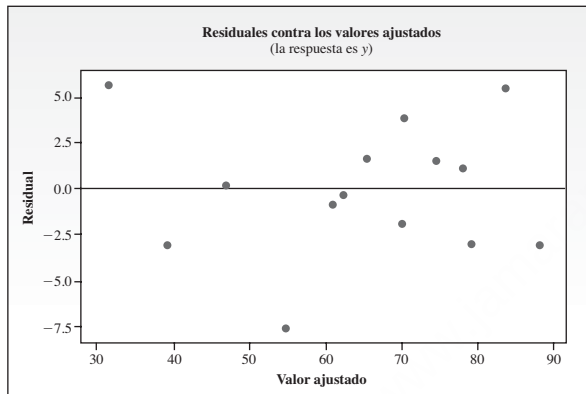
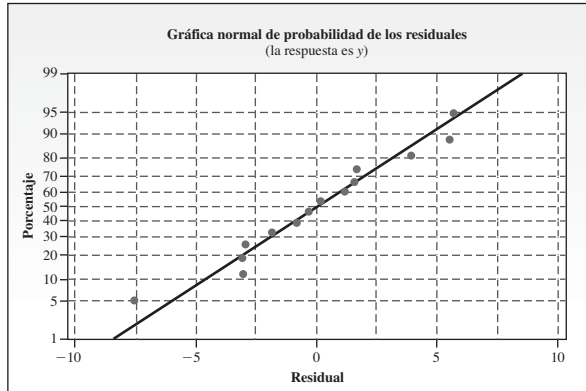
Predicted Values for New Observations

New Obs	Fit	SE Fit	95% CI	95% PI
1	56.29	3.52	(48.44, 64.13)	(44.13, 68.44)

Values of Predictors for New Observations

New Obs	x1	x2	x1x2
1	1.00	8.00	8.00

Gráficas de diagnóstico para el ejercicio 9.19



- ¿Qué tan bien ajusta el modelo? Use cualesquiera estadísticos relevantes y herramientas de diagnóstico de la salida impresa para contestar esta pregunta.
- Escriba las ecuaciones de las dos rectas que describen la tendencia en consumo durante el periodo de 7 años, para carne roja y para pollo.
- Use la ecuación de predicción para hallar una estimación puntual del promedio de consumo de carne roja por persona en el año 8. Compare este valor con el valor marcado “Fit” de la salida impresa.
- Use la salida impresa para hallar un intervalo de confianza de 95% para el promedio de consumo de carne roja por persona en el año 8. ¿Cuál es el intervalo de predicción de 95% para el consumo de carne roja por persona en el año 8? ¿Hay algún problema con la validez del nivel de confianza de 95% para estos intervalos?

DATOS
EX0920

9.20 Algodón contra pepinos En el ejercicio 7.65 utilizamos el procedimiento de análisis de varianza, para analizar un experimento factorial de 2×3 , en el que cada combinación de factor - nivel se

repitió cinco veces. El experimento contenía el número de huevos puestos por moscas blancas enjauladas sobre dos plantas diferentes, a tres niveles diferentes de temperatura. Supongamos que varias de las moscas blancas murieron antes de completar el experimento, de modo que el número de repeticiones ya no fue el mismo para cada tratamiento. El análisis de fórmulas de varianza del capítulo 7 ya no se puede usar, pero el experimento *puede* ser analizado usando un análisis de regresión múltiple. Los resultados de este **experimento factorial de 2×3 con réplicas desiguales** se muestran en la tabla siguiente.

Algodón			Pepino		
70°	77°	82°	70°	77°	82°
37	34	46	50	59	43
21	54	32	53	53	62
36	40	41	25	31	71
43	42		37	69	49
31			48	51	

- Escriba un modelo para analizar este experimento. Asegúrese de incluir un término para la interacción entre planta y temperatura.
- Use un paquete de software para efectuar el análisis de regresión múltiple.
- ¿Los datos dan suficiente evidencia para indicar que el efecto de la temperatura en el número de huevos puestos es *diferente* dependiendo del tipo de planta?
- Con base en los resultados de la parte c, ¿sugiere usted reajustar un modelo diferente? Si es así, vuelva a correr el análisis de regresión usando el nuevo modelo y analice la salida impresa.
- Escriba un párrafo que resuma los resultados de sus análisis.

DATOS
EX0921

9.21 Calificaciones de aprovechamiento III

El Índice de Aprovechamiento Académico (API), descrito en el ejercicio 8.13, es una medida del aprovechamiento escolar con base en los resultados del Examen de Aprovechamiento Stanford 9. Las calificaciones del API para 12 escuelas elementales en el condado de Riverside, California, se muestran a continuación junto con varias variables independientes más.⁴

Escuela	Puntuación API, y	EL (%), x_1	Almuerzo gratuito/reducido (%), x_2	Nivel educativo promedio del padre, x_3	Dotados y talentosos (%), x_4	API del año anterior, x_5
1	745	71	89	1.70	4	705
2	808	18	51	2.91	16	809
3	798	24	79	2.21	10	763
4	791	50	76	2.19	5	786
5	854	17	56	2.84	7	839
6	688	71	27	1.70	6	673

Escuela	Puntuación API, y	EL (%), x_1	Almuerzo gratuito/reducido (%), x_2	Nivel educativo promedio del padre, x_3	Dotados y talentosos (%), x_4	API del año anterior, x_5
7	801	11	39	2.79	7	804
8	751	57	87	1.72	1	750
9	778	34	81	2.14	6	770
10	846	9	31	3.22	22	841
11	690	53	78	2.14	3	706
12	685	77	28	1.46	8	665

Las variables están definidas como

y = Puntuación API en 2010

x_1 = % de estudiantes que son “Estudiantes del Idioma Inglés”

x_2 = % de estudiantes que reciben un almuerzo gratuito o a un costo reducido

x_3 = Nivel educativo promedio del padre (con 1 = No graduado de bachillerato, 2 = Graduado de bachillerato, 3 = Carrera trunca, 4 = Graduado universitario, 5 = Posgrado)

x_4 = % de estudiantes en el Programa de Educación para Dotados y Talentosos

x_5 = Puntuación API en 2009

La salida impresa de MINITAB para un modelo de regresión de primer orden se da a continuación.

Análisis de regresión: y contra x_1 , x_2 , x_3 , x_4 , x_5

The regression equation is
 $y = 15 - 0.306 x_1 + 0.076 x_2 - 48.1 x_3 + 1.93 x_4 + 1.13 x_5$

Predictor	Coef	SE Coef	T	P
Constant	14.9	179.7	0.08	0.936
x_1	-0.3060	0.6483	-0.47	0.654
x_2	0.0763	0.2885	0.26	0.800
x_3	-48.06	34.33	-1.40	0.211
x_4	1.927	1.450	1.33	0.232
x_5	1.1269	0.2503	4.50	0.004

S = 17.0180 R-Sq = 95.4% R-Sq(adj) = 91.6%

Source	DF	SS	MS	F	P
Regression	5	36121.2	7224.2	24.94	0.001
Residual Error	6	1737.7	289.6		
Total	11	37858.9			

Source	DF	Seq SS
x_1	1	28748.4
x_2	1	701.0
x_3	1	591.6
x_4	1	212.3
x_5	1	5868.0

- ¿Cuál es el modelo que se ha ajustado a estos datos?
¿Cuál es la ecuación de predicción de mínimos cuadrados?
- ¿Qué tan bien se ajusta el modelo? Use cualesquier estadístico relevante de la salida impresa para contestar esta pregunta.
- ¿Cuáles de las variables independientes, si las hay, son útiles para predecir el API, dadas las otras variables independientes ya en el modelo? Explique.
- Use los valores de R^2 y $R^2(\text{adj})$ de la siguiente salida impresa para escoger el mejor modelo para predicción.

¿Confiaría usted en usar el modelo elegido, para predecir la calificación API para el siguiente año, con base en un modelo que contenga variables similares? Explique.

Regresión de mejores subconjuntos: y contra x_1 , x_2 , x_3 , x_4 , x_5

Vars	R-Sq	R-Sq (adj)	Mallows Cp	S	x	x	x	x	x
					1	2	3	4	5
1	93.2	92.5	0.9	16.095					X
1	75.9	73.5	23.5	30.184	X				
2	93.8	92.5	2.0	16.090		X			X
2	93.3	91.8	2.8	16.789			X		X
3	95.2	93.3	2.3	15.132			X	X	X
3	94.0	91.7	3.9	16.858	X	X			X
4	95.4	92.7	4.1	15.847	X	X	X		X
4	95.2	92.5	4.2	16.045		X	X	X	X
5	95.4	91.6	6.0	17.018	X	X	X	X	X

9.22 Tabla de partículas Un ingeniero de control de calidad está interesado en predecir la resistencia de una tabla de partículas y , como función del tamaño de las partículas x_1 y dos tipos de compuestos aglutinantes. Si se espera que la respuesta básica sea una función cuadrática del tamaño de una partícula, escriba un modelo lineal que incorpore la variable cualitativa “compuesto aglutinante” en la ecuación de predicción.

DATOS
EX0923

9.23 Proyectos de construcción

En un estudio para examinar la relación entre el tiempo requerido para completar un proyecto de construcción y varias variables independientes pertinentes, un analista compiló una lista de cuatro variables que podrían ser útiles para predecir el tiempo de terminación. Estas cuatro variables eran el tamaño del contrato, x_1 (en unidades de 1000 dólares), el número de días de trabajo adversamente afectados por el clima x_2 , el número de subcontratistas involucrados en el proyecto x_4 y una variable x_3 que midió la presencia ($x_3 = 1$) o ausencia ($x_3 = 0$) de una huelga de trabajadores durante la construcción. Se eligieron al azar 15 proyectos de construcción y se midieron cada una de las cuatro variables, así como el tiempo para terminar el proyecto.

y	x_1	x_2	x_3	x_4
29	60	7	0	7
15	80	10	0	8
60	100	8	1	10
10	50	14	0	5
70	200	12	1	11
15	50	4	0	3
75	500	15	1	12
30	75	5	0	6
45	750	10	0	10
90	1200	20	1	12
7	70	5	0	3
21	80	3	0	6
28	300	8	0	8
50	2600	14	1	13
30	110	7	0	4

Un análisis de estos datos usando un modelo de primer orden en x_1 , x_2 , x_3 y x_4 produjo la siguiente salida impresa. Dé un análisis completo de la salida impresa e interprete sus resultados. ¿Qué se puede decir acerca de la aparente contribución de x_1 y x_2 en la predicción de y ?

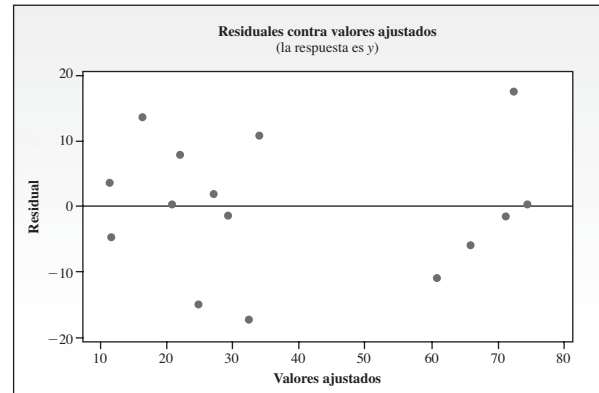
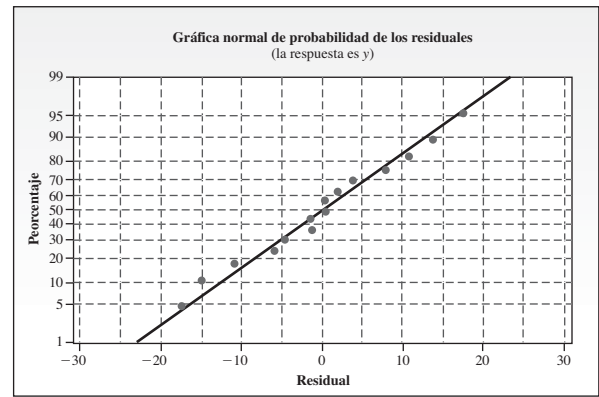
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9204
R Square	0.8471
Adjusted R Square	0.7859
Standard Error	11.8450
Observations	15

ANOVA

	gl	SS	MS	F	Significance F
Regression	4	7770.297	1942.574	13.846	0.000
Residual	10	1403.036	140.304		
Total	14	9173.333			

	Coefficients	Standard Error	t Stat	P-value
Intercept	-1.589	11.656	-0.136	0.894
x1	-0.008	0.006	-1.259	0.237
x2	0.675	1.000	0.675	0.515
x3	28.013	11.371	2.463	0.033
x4	3.489	1.935	1.803	0.102



PRUEBA DE CONJUNTOS DE COEFICIENTES DE REGRESIÓN

9.6

En las secciones precedentes hemos probado el conjunto completo de coeficientes de regresión parcial usando la prueba F para el ajuste general del modelo y los coeficientes de regresión parcial individualmente usando la prueba t de Student. Además de estas dos importantes pruebas se pueden probar hipótesis acerca de algunos subconjuntos de estos coeficientes de regresión.

Por ejemplo, suponga que una compañía sospecha que la demanda y de algún producto podría estar relacionada con hasta cinco variables independientes: x_1 , x_2 , x_3 , x_4 y x_5 . El costo de obtener mediciones de las variables x_3 , x_4 y x_5 es muy alto. Si, en un pequeño estudio piloto, la compañía pudiera demostrar que estas tres variables contribuyen con poca o ninguna información para la predicción de y , pueden ser eliminadas del estudio con grandes ahorros para la compañía.

Si las cinco variables, x_1 , x_2 , x_3 , x_4 y x_5 , se usan para predecir y , el modelo de regresión se escribiría como

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \epsilon$$

No obstante, si x_3 , x_4 y x_5 no aportan información para la predicción de y , entonces no aparecerían en el modelo, es decir, $\beta_3 = \beta_4 = \beta_5 = 0$ y el modelo reducido sería

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

En consecuencia, se desea probar la hipótesis nula

$$H_0 : \beta_3 = \beta_4 = \beta_5 = 0$$

esto es, las variables independientes x_3 , x_4 y x_5 no aportan información para la predicción de y , contra la hipótesis alternativa

$$H_a : \text{Al menos uno de los parámetros } \beta_3, \beta_4 \text{ o } \beta_5 \text{ es diferente de } 0$$

es decir, al menos una de las variables x_3 , x_4 y x_5 aporta información para la predicción de y . Entonces, para decidir si el modelo completo es preferible al modelo reducido para predecir demanda, se lleva a cabo una prueba de hipótesis acerca de un conjunto de tres parámetros, β_3 , β_4 y β_5 .

Una prueba de hipótesis respecto a un conjunto de parámetros de modelo involucra dos modelos:

Modelo 1 (modelo reducido)

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_r x_r$$

Modelo 2 (modelo completo)

$$E(y) = \underbrace{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_r x_r}_{\text{términos en modelo 1}} + \underbrace{\beta_{r+1} x_{r+1} + \beta_{r+2} x_{r+2} + \cdots + \beta_k x_k}_{\text{términos adicionales en modelo 2}}$$

Suponga que se ajustaron ambos modelos al conjunto de datos y se calculó la suma de cuadrados para el error de los dos análisis de regresión. Si el modelo 2 aporta más información para la predicción de y que el modelo 1, entonces los errores de predicción para el modelo 2 deben ser más pequeños que los correspondientes errores para el modelo 1, y la SSE_2 debe ser menor que la SSE_1 . De hecho, cuanto mayor sea la diferencia entre SSE_1 y SSE_2 , mayor es la evidencia para indicar que el modelo 2 aporta más información para la predicción de y que el modelo 1.

La prueba de la hipótesis nula

$$H_0 : \beta_{r+1} = \beta_{r+2} = \cdots = \beta_k = 0$$

contra la hipótesis alternativa

$$H_a : \text{Al menos uno de los parámetros } \beta_{r+1}, \beta_{r+2}, \cdots, \beta_k \text{ difiere de } 0$$

utiliza el estadístico de prueba

$$F = \frac{(SSE_1 - SSE_2)/(k - r)}{MSE_2}$$

donde F está basada en $gl_1 = (k - r)$ y $gl_2 = n - (k + 1)$. Observe que los parámetros $(k - r)$ contenidos en H_0 son los agregados al modelo 1 para obtener el modelo 2. Los grados de libertad gl_1 del numerador siempre son iguales a $(k - r)$, que es el número de parámetros contenidos en H_0 . Los grados de libertad gl_2 del denominador es el número de grados de libertad asociado con la suma de cuadrados para error, SSE_2 , para el modelo completo.

La región de rechazo para la prueba es idéntica a la región de rechazo para todos los análisis de varianza de pruebas F , es decir,

$$F > F_\alpha$$

EJEMPLO

9.8

Consulte los datos de bienes raíces del ejemplo 9.2 que relacionan el precio de venta de lista y con la superficie en pies cuadrados del área de vivienda x_1 , el número de pisos x_2 , el número de recámaras x_3 y el número de baños, x_4 . El agente de bienes raíces sospecha que la superficie en pies cuadrados del área de vivienda es la variable predictora más importante, y que las otras variables podrían ser eliminadas del modelo sin perder mucha información de predicción. Pruebe esta afirmación con $\alpha = .05$.

Solución La hipótesis a probar es

$$H_0: \beta_2 = \beta_3 = \beta_4 = 0$$

contra la hipótesis alternativa que al menos una de β_2 , β_3 o β_4 es diferente de 0. El **modelo 2 completo**, dado como

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$$

fue ajustado en el ejemplo 9.2. Una parte de la salida impresa *MINITAB* de la figura 9.3 se reproduce en la figura 9.15 junto con una parte de la salida impresa *MINITAB* para el análisis de regresión lineal simple del **modelo 1 reducido**, dado como

$$y = \beta_0 + \beta_1 x_1 + \epsilon$$

FIGURA 9.15

Partes de las salidas impresas de regresión *MINITAB* para modelos a) completo y b) reducido para el ejemplo 9.8

● **Análisis de regresión: a) precio de lista contra pies cuadrados, número de pisos, recámaras y baños**

S = 6.84930 R-Sq = 97.1% R-Sq(adj) = 96.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	15913.0	3978.3	84.80	0.000
Residual Error	10	469.1	46.9		
Total	14	16382.2			

Análisis de regresión: b) precio de lista contra pies cuadrados

S = 10.9294 R-Sq = 90.5% R-Sq(adj) = 89.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	14829	14829	124.14	0.000
Residual Error	13	1553	119		
Total	14	16382			

Entonces $SSE_1 = 1553$ de la figura 9.15b) y $SSE_2 = 469.1$ y $MSE_2 = 46.9$ de la figura 9.15a). El estadístico de prueba es

$$\begin{aligned}
 F &= \frac{(SSE_1 - SSE_2)/(k - r)}{MSE_2} \\
 &= \frac{(1553 - 469.1)/(4 - 1)}{46.9} = 7.70
 \end{aligned}$$

El valor crítico de F con $\alpha = .05$, $gl_1 = 3$ y $gl_2 = n - (k + 1) = 15 - (4 + 1) = 10$ es $F_{.05} = 3.71$. Por tanto, H_0 es rechazada. Hay evidencia para indicar que al menos una de las tres variables que son número de pisos, recámaras o baños, está contribuyendo con información significativa para predecir el precio de venta de lista.

INTERPRETACIÓN DE GRÁFICAS RESIDUALES

Una vez más, se pueden usar gráficas residuales para descubrir posibles violaciones en las suposiciones requeridas para un análisis de regresión. Hay varios patrones comunes que se deben reconocer porque se presentan con frecuencia en aplicaciones prácticas.

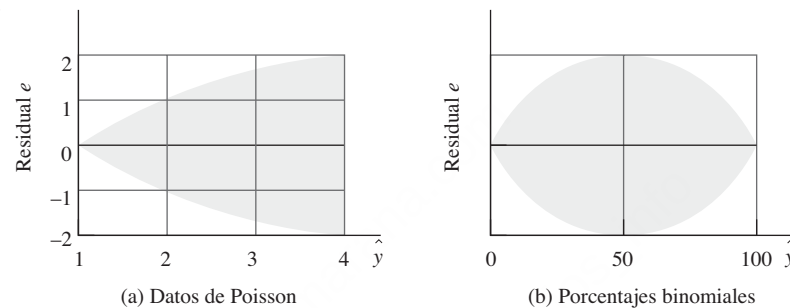
La varianza de algunos tipos de datos cambia conforme cambia la media:

- Los datos de Poisson exhiben variación que *aumenta* con la media.
- Los datos binomiales exhiben variación que *aumenta* para valores de p de .0 a .5 y luego *disminuye* para valores de p de .5 a 1.0.

Las gráficas residuales para estos tipos de datos tienen un patrón semejante al que se ve en la figura 9.16.

FIGURA 9.16

Gráficas de residuales
contra $\hat{y}$



Si el rango de los residuales aumenta cuando $\hat{y}$ aumenta y se sabe que los datos son mediciones sobre variables de Poisson, se puede estabilizar la varianza de la respuesta al correr el análisis de regresión en $y^* = \sqrt{y}$. O bien, si los porcentajes se calculan a partir de datos binomiales, se puede usar la transformación arcsen, $y^* = \sin^{-1}\sqrt{y}$.†

Incluso si no se está seguro de por qué el rango de los residuales aumenta cuando $\hat{y}$ aumenta, todavía se puede usar una transformación de y que afecta valores más grandes de y más que valores pequeños, por ejemplo $yy^* = \sqrt{y}$ o $y^* = \ln y$. Estas transformaciones tienen una tendencia para estabilizar la varianza de y^* y para hacer que la distribución de y^* sea más casi normal cuando la distribución de y sea altamente sesgada.

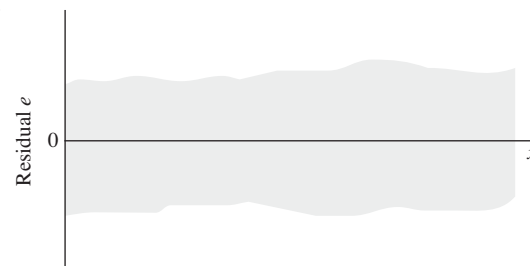
Es frecuente que las gráficas de los residuales contra los ajustes o contra las variables predictoras individuales muestren un patrón que indica que se ha elegido un modelo incorrecto. Por ejemplo, si $E(y)$ y una sola variable independiente x están linealmente relacionadas, es decir,

$$E(y) = \beta_0 + \beta_1 x$$

y se ajusta una recta a los datos, entonces los valores y observados deben variar en una forma aleatoria alrededor de $\hat{y}$, y una gráfica de los residuales contra x aparecerá como se ve en la figura 9.17.

FIGURA 9.17

Gráfica residual cuando
el modelo da una buena
aproximación a la realidad

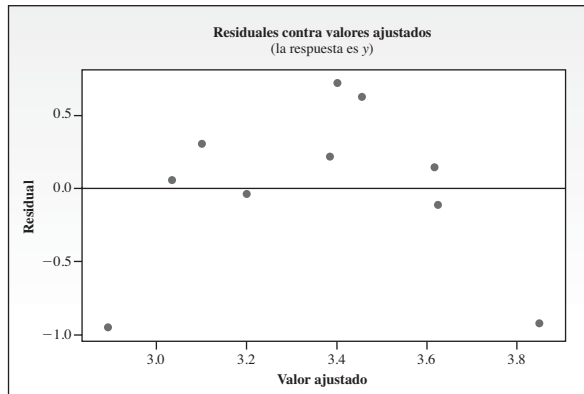


† En el capítulo 7 y anteriores, representamos la respuesta variable con el símbolo x . En los capítulos sobre análisis de regresión, capítulos 8 y 9, la variable de respuesta está representada por el símbolo y .

En el ejemplo 9.3, se ajustó un modelo cuadrático que relacionaba la productividad y con el tamaño de tienda x . Si incorrectamente se hubiera usado un modelo lineal para ajustar estos datos, la gráfica residual de la figura 9.18 mostraría que la variación no explicada exhibe un patrón curvado, que sugiere que hay un efecto cuadrático que no se ha incluido en el modelo.

FIGURA 9.18

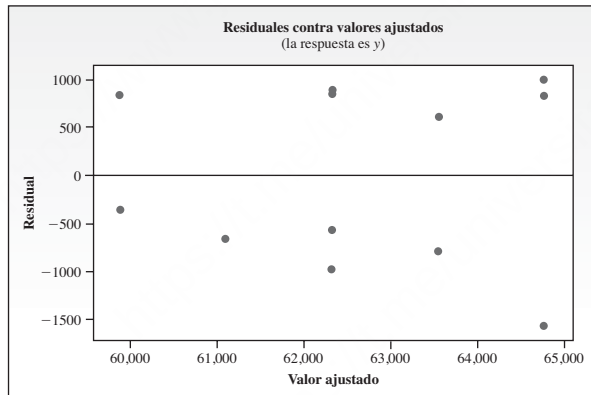
Gráfica residual para ajuste lineal de tamaño de tienda y datos de productividad en el ejemplo 9.3



Para los datos del ejemplo 9.6, los residuales de una regresión lineal de salario con años de experiencia x_1 sin incluir género, x_2 , mostraría un conjunto distinto de residuales positivos correspondientes a los hombres y un conjunto de residuales negativos correspondientes a las mujeres (véase la figura 9.19). Este patrón señala que la variable “género” no estaba incluida en el modelo.

FIGURA 9.19

Gráfica residual para ajuste lineal de datos de salario en el ejemplo 9.6



Desafortunadamente, no todas las gráficas residuales dan una indicación tan clara del problema. Con todo cuidado deben examinarse las gráficas residuales, buscando que no haya aleatoriedad en el patrón de residuales. Si se puede hallar una explicación para el comportamiento de los residuales se puede modificar el modelo para eliminar el problema.

9.8

ANÁLISIS DE REGRESIÓN POR PASOS

A veces hay un gran número de variables predictoras independientes que *podrían* tener un efecto en la variable de respuesta y . Por ejemplo, trate de hacer una lista de todas las variables que podrían afectar el promedio de calificaciones (GPA) de un estudiante de primer año de universidad:

- Calificaciones en cursos de preparatoria, promedio de calificaciones de preparatoria, calificación de examen de aptitud escolar, calificación de examen en universidades
- Especialidad, número de unidades llevadas, número de cursos tomados
- Programa de trabajo, estado civil, se traslada o vive en el plantel

¿Cuáles de este gran número de variables independientes deben incluirse en el modelo? Como el número de términos podría rápidamente hacerse muy difícil de manejar, sería posible elegir un procedimiento llamado **análisis de regresión por pasos**, que se pone en práctica por computadora y está disponible en casi todos los paquetes de estadística.

Un análisis de regresión por pasos ajusta una variedad de modelos a los datos, agregando y eliminando variables cuando la significancia de ellas en presencia de las otras variables es *significativa* o *no significativa*, respectivamente. Una vez que el programa haya hecho un número suficiente de iteraciones y no haya más variables significativas cuando se agreguen al modelo y ninguna de las variables del modelo son no significativas cuando son eliminadas, el procedimiento se detiene.

Un análisis de regresión por pasos es un modo fácil de localizar algunas variables que aportan información para predecir y , pero no es a prueba de errores. Como estos programas siempre ajustan modelos de primer orden de la forma

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

no son útiles para detectar *curvatura* o *interacción* en los datos. El análisis de regresión por pasos se usa mejor como herramienta preliminar para identificar cuál, de un gran número de variables, debe ser considerada en su modelo. Entonces es necesario decidir cómo introducir estas variables en el modelo real que usará para predicción.

INTERPRETACIÓN ERRÓNEA DE UN ANÁLISIS DE REGRESIÓN

9.9

Son comunes varias interpretaciones erróneas de la salida impresa de un análisis de regresión. Ya hemos mencionado la importancia de la selección de un modelo. Si un modelo no se ajusta a un conjunto de datos, no quiere decir que las variables incluidas en el modelo aporten poca o ninguna información para la predicción de y . Las variables pueden ser importantes contribuyentes de información, pero pueden haberse introducido las variables en el modelo en una forma equivocada. Por ejemplo, un modelo de segundo orden en las variables podría dar un muy buen ajuste a los datos cuando un modelo de primer orden parece ser por completo inútil para describir la variable de respuesta y .

Causalidad

Es necesario tener cuidado de no concluir que cambios en x *causan* cambios en y . Este tipo de **relación causal** puede ser detectada sólo con un *experimento cuidadosamente diseñado*. Por ejemplo, si al azar se asignan unidades experimentales a cada uno de dos niveles de una variable x , por ejemplo $x = 5$ y $x = 10$ y los datos muestran que el valor medio de y es mayor cuando $x = 10$, entonces se puede decir que el cambio en el nivel de x causó un cambio en el valor medio de y . Pero en casi todos los análisis de regresión, en los que los experimentos no están diseñados, no hay garantía de que una variable predictora importante, por ejemplo x_1 , cause un cambio en y . Es muy posible que alguna variable que ni siquiera esté en el modelo cause que *tanto y como* x_1 cambien.

Multicolinealidad

Ni el tamaño de un coeficiente de regresión ni su valor t indican la importancia de la variable como contribuyente de información. Por ejemplo, suponga que se desea predecir y , la calificación de cálculo de un estudiante universitario, con base en $x_1 =$ promedio de calificaciones de matemáticas en preparatoria y $x_2 =$ calificación en el examen de aptitud en matemáticas. Como estas dos variables contienen mucho de lo mismo o **información compartida**, no es

de sorprender que una vez que una de las variables se introduzca en el modelo, la otra aporta muy poca información adicional. El valor t individual es pequeño, pero, si las variables se introdujeron en el orden inverso, se vería invertido el tamaño de los valores t .

La situación descrita líneas antes se denomina **multicolinealidad** y se presenta cuando dos o más de las variables predictoras están altamente correlacionadas entre sí. Cuando la multicolinealidad está presente en un problema de regresión, puede tener estos efectos en el análisis:

- Los coeficientes de regresión estimados tendrán errores estándar grandes, causando imprecisión en intervalos de confianza y predicción.
- Agregar o eliminar una variable de predicción puede causar cambios significativos en los valores de los otros coeficientes de regresión.

¿Cómo saber si un análisis de regresión exhibe multicolinealidad? Busque estos indicios:

- El valor de R^2 es grande, lo cual indica un buen ajuste, pero las pruebas t individuales no son significativas.
- Los signos de los coeficientes de regresión son contrarios a lo que intuitivamente se esperaría fueran las contribuciones de esas variables.
- Una matriz de correlaciones, generada por computadora, muestra cuáles variables predictoras están altamente correlacionadas entre sí y con la respuesta y .

La figura 9.20 muestra la matriz de correlaciones generada para los datos de bienes raíces del ejemplo 9.2. La primera columna de la matriz muestra las correlaciones de cada variable de predicción con la variable de respuesta y . Todas son significativamente diferentes de cero, pero la primera variable, x_1 = área de vivienda, es la más altamente correlacionada. Las últimas tres columnas de la matriz muestran correlaciones significativas entre todas las variables predictoras, excepto un par. Éste es un fuerte indicio de multicolinealidad. Si se trata de eliminar una de las variables del modelo, pueden cambiar en forma drástica los efectos de las otras tres. Otro indicio puede hallarse al examinar los coeficientes de la recta de predicción,

$$\text{ListPrice} = 119 + 6.27 \text{ Square Feet} - 16.2 \text{ Number of Floors} \\ - 2.67 \text{ Bedrooms} + 30.3 \text{ Baths}$$

FIGURA 9.20

Matriz de correlación para los datos de bienes raíces del ejemplo 9.2

Correlaciones: precio de lista; pies cuadrados, número de pisos, recámaras, baños

	ListPrice	SqFeet	Numflrs	Bdrms
Square Feet	0.951 0.000			
Number of Fl	0.605 0.017	0.630 0.012		
Bedrooms	0.746 0.001	0.711 0.003	0.375 0.168	
Baths	0.834 0.000	0.720 0.002	0.760 0.001	0.675 0.006

Cell Contents: Pearson Correlation
P-Value

Se podría esperar que más pisos y recámaras aumentaran el precio de lista, pero sus coeficientes son negativos.

Como existe multicolinealidad en alguna medida en todos los problemas de regresión, debemos considerar los términos individuales como *aportadores de información* en lugar de tratar de medir la importancia práctica de cada término. La decisión primaria a tomarse es si un término aporta suficiente información para justificar su inclusión en el modelo.

PASOS A SEGUIR AL CONSTRUIR UN MODELO DE REGRESIÓN MÚLTIPLE

El objetivo final de un análisis de regresión múltiple es desarrollar un modelo que en forma precisa prediga y como función de un conjunto de variables predictoras $x_1, x_2, \dots, x_k$. El procedimiento paso a paso para desarrollar este modelo se presentó en la sección 9.4 y volvemos a expresarlo a continuación con algún detalle adicional. Si se usa este método, lo que puede parecer un problema complicado se puede hacer más sencillo. Al igual que en cualquier procedimiento estadístico, la confianza crecerá a medida que ganemos experiencia con el análisis de regresión múltiple en varias situaciones prácticas.

1. Seleccione las variables predictoras a ser incluidas en el modelo. Como algunas de estas variables pueden contener información compartida, se puede reducir la lista al correr un análisis de regresión por pasos (véase la sección 9.8). Mantenga el número de variables predictoras lo suficientemente pequeño para que sea efectivo pero manejable. Es necesario estar conscientes que el número de observaciones del conjunto de datos debe exceder el número de términos del modelo; cuanto mayor el exceso, mejor.
2. Escriba un modelo usando las variables predictoras seleccionadas. Si las variables son cualitativas, es mejor empezar incluyendo términos de interacción; si las variables son cuantitativas, es mejor empezar con un modelo de segundo orden. Los términos innecesarios pueden eliminarse después. Obtenga el modelo de predicción ajustado.
3. Use el análisis de varianza de la prueba F y R^2 para determinar qué tan bien se ajusta el modelo a los datos.
4. Verifique las pruebas t para los coeficientes de regresión parcial para ver cuáles aporta información significativa en presencia de los otros. Si algunos términos parecen ser no significativos, considere eliminarlos. Si elige comparar varios modelos diferentes, use $R^2(\text{adj})$ para comparar su efectividad.
5. Use gráficas residuales generadas por computadora para ver si hay violación de las suposiciones de regresión.

REPASO DEL CAPÍTULO

Conceptos y fórmulas clave

I. El modelo lineal general

1. $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon$
2. El error aleatorio ϵ tiene una distribución normal con media 0 y varianza 2.

II. Método de mínimos cuadrados

1. Las estimaciones $b_0, b_1, \dots, b_k$, para $\beta_0, \beta_1, \dots, \beta_k$, se eligen para minimizar SSE, la suma del cuadrado de desviaciones alrededor de la recta de regresión, $\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_k x_k$.
2. Las estimaciones de mínimos cuadrados son producidas por computadora.

III. Análisis de varianza

1. $SS \text{ Total} = SSR + SSE$, donde $SS \text{ Total} = S_{yy}$. La tabla ANOVA es producida por computadora.
2. La mejor estimación de σ^2 es

$$MSE = \frac{SSE}{n - k - 1}$$

IV. Prueba, estimación y predicción

1. Una prueba de la significancia de la regresión, $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$, se puede implementar usando el análisis de varianza de prueba F :

$$F = \frac{MSR}{MSE}$$

2. La fuerza de la relación entre x y y se puede medir usando

$$R^2 = \frac{SSR}{SS \text{ Total}}$$

lo cual se acerca a 1 cuando la relación se hace más fuerte.

3. Use gráficas residuales para verificar si hay normalidad, desigualdad de varianzas y un modelo incorrectamente ajustado.
4. Las pruebas de significancia para los coeficientes de regresión parcial se pueden efectuar usando la prueba t de Student:

$$t = \frac{b_i - \beta_i}{SE(b_i)} \quad \text{con error } gl = n - k - 1$$

5. Los intervalos de confianza pueden ser generados por computadora para estimar el valor promedio de y , $E(y)$, para valores dados de $x_1, x_2, \dots, x_k$. Los intervalos de predicción generados por computadora se pueden usar para predecir una observación particular y para valores determinados

de $x_1, x_2, \dots, x_k$. Para $x_1, x_2, \dots, x_k$ dados, los intervalos de predicción son siempre más anchos que los intervalos de confianza.

V. Construcción de un modelo

1. El número de términos en un modelo de regresión no puede exceder al número de observaciones del conjunto de datos y debe ser considerablemente menor.
2. Para explicar un efecto curvilíneo en una variable *cuantitativa*, use un modelo polinomial de segundo orden. Para un efecto cúbico, use un modelo polinomial de tercer orden.
3. Para agregar una variable *cualitativa* con k categorías, use $(k - 1)$ variables ficticias o indicadoras.
4. Puede haber interacciones entre dos variables cuantitativas o entre una variable cuantitativa y una cualitativa. Los términos de interacción se introducen como $\beta_{x_i x_j}$.
5. Compare modelos usando $R^2(\text{adj})$.



TECNOLOGÍA ACTUAL

Procedimientos de regresión múltiple: *Microsoft Excel*

El procedimiento para realizar un análisis de regresión múltiple en *MS Excel* es idéntico al procedimiento de regresión lineal descrito en la sección “Tecnología actual” en el capítulo 8, excepto que el rango de las variables x abarcará más de una columna. Podría desear repasar esa sección antes de continuar.

EJEMPLO

9.9

Suponga que una variable de respuesta y se relaciona con cuatro variables predictoras, x_1, x_2, x_3 y x_4 , de modo que $k = 4$.

1. Introduzca los valores observados de y y cada una de las $k = 4$ variables predictoras en las primeras $(k + 1)$ columnas de una hoja de cálculo de *Excel*. (NOTA: A fin de que el análisis de regresión múltiple funcione de manera apropiada, debe haber una columna de valores para cada variable predictora independiente x_i en su modelo, y las columnas de las x **deben ser adyacentes entre sí**.)
2. Use **Data ► Data Analysis ► Regression** para generar el cuadro de diálogo, resaltando o tecleando los rangos de celdas para los valores x_i y y , y seleccionar “Labels” de ser necesario.
3. Si hace clic en “Confidence Level”, *Excel* calculará intervalos de confianza para las estimaciones de regresión, b_0, b_1, b_2, b_3 y b_4 . Introduzca una ubicación de celda para el **Output Range** y haga clic en **OK** para generar la salida de regresión.

NOTA: *MS Excel* no proporciona opciones para estimación y predicción. Además, las gráficas de diagnóstico que pueden generarse en *Excel* no son las mismas gráficas que se han expuesto en la sección 9.3 y no se comentarán en esta sección.

Procedimientos de regresión múltiple: MINITAB

El procedimiento para ejecutar un análisis de regresión múltiple en *MINITAB* es parecido al procedimiento de regresión lineal descrito en la sección “Tecnología actual” en el capítulo 8, excepto que el rango de las variables x abarcará más de una columna. Podría desear repasar esa sección antes de continuar.

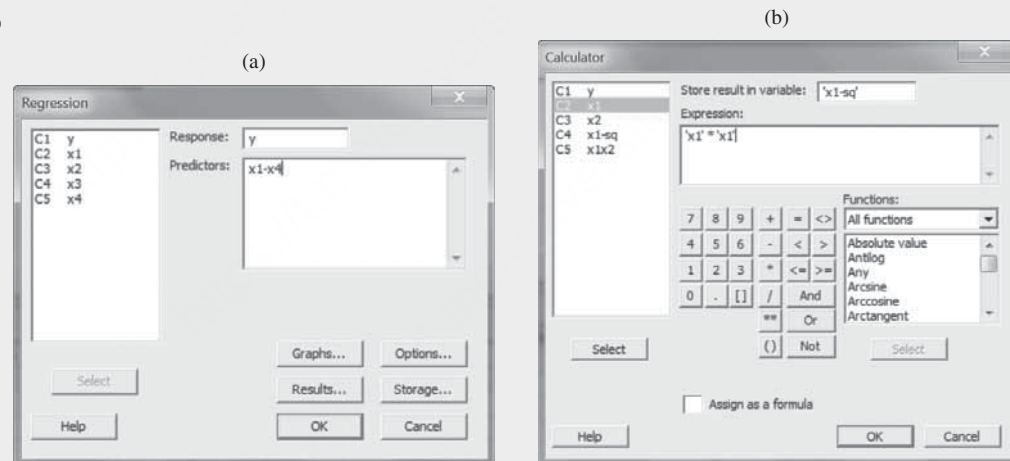
EJEMPLO

9.10

Suponga que una variable de respuesta y se relaciona con cuatro variables predictoras, x_1 , x_2 , x_3 y x_4 , de modo que $k = 4$.

1. Introduzca los valores observados de y y cada una de las $k = 4$ variables predictoras en las primeras $(k + 1)$ columnas de una hoja de cálculo de *MINITAB*. Una vez hecho esto, las principales herramientas inferenciales para el análisis de regresión múltiple se generan usando **Stat ► Regression ► Regression**. El cuadro de diálogo para el comando **Regression** se muestra en la figura 9.21a).
2. Seleccione **y** para la variable Response y **x_1 , x_2 , ..., x_k** para las variables Predictor. Ahora puede elegir generar algunas gráficas residuales para comprobar la validez de sus suposiciones de regresión antes de usar el modelo para estimación o predicción. Elija **Graphs** para mostrar el cuadro de diálogo para gráficas residuales, y elija la gráfica de diagnóstico apropiada.

FIGURA 9.21



3. Una vez que haya verificado la pertinencia de su modelo de regresión múltiple, puede elegir **Options** y obtener intervalos de confianza y predicción para cualquiera de estos casos:
 - Un solo conjunto de valores x_1 , x_2 , ..., x_k (introducidos en el cuadro marcado “Prediction intervals for new observations”)
 - Varios conjuntos de valores x_1 , x_2 , ..., x_k almacenados en k columnas de la hoja de cálculo.

Cuando haga clic en **OK** dos veces se generará la salida de regresión.

4. La única dificultad en la ejecución del análisis de regresión múltiple usando *MINITAB* podría ser introducir de manera apropiada los datos para su modelo particular. Si el modelo implica términos polinomiales o términos de interacción, el comando **Calc ► Calculator** le ayudará. Por ejemplo, suponga que desea ajustar el modelo

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_1 x_2$$

Necesitará introducir los valores observados de y , x_1 y x_2 en las primeras tres columnas de la hoja de cálculo *MINITAB*. Nombre la columna C4 “x1-sq” y nombre C5 “x1x2”. Ahora puede usar el cuadro de diálogo de la calculadora que se muestra en la figura 9.21b) para generar estas dos columnas. En el cuadro **Expression**, seleccione **x1 * x1** o **x1 ** 2** y almacene los resultados en C4 (x1-sq). Haga clic en **OK**. Del mismo modo, para obtener los datos para C5, seleccione **x1 * x2** y almacene los resultados en C5 (x1x2). Haga clic en **OK**. Ahora está listo para ejecutar el análisis de regresión múltiple.

5. Si está ajustando ya sea un modelo cuadrático o un modelo cúbico en una variable x , ahora puede graficar los puntos de datos, la curva de regresión polinomial y los límites de confianza y predicción superior e inferior usando **Stat ► Regression ► Fitted line Plot**. Seleccione y y x para las variables Response y Predictor, y haga clic en “Display confidence interval” y “Display prediction interval” en el cuadro de diálogo **Options**. Asegúrese de que está seleccionado **Quadratic** o **Cubic** como el “Type of Regression Model”, de modo que obtenga el ajuste apropiado para los datos.
6. Recuerde que en el capítulo 8 usó **Stat ► Basic Statistics ► Correlation** para obtener el valor del coeficiente de correlación r . En el análisis de regresión múltiple, el mismo comando generará una matriz de correlaciones, una por cada par de variables en el conjunto $y, x_1, x_2, \dots, x_k$. Asegúrese de que está seleccionado el cuadro marcado “Display p-values”. Los valores p proporcionarán información sobre la correlación significativa entre un par particular, en presencia de todas las otras variables en el modelo, y son idénticas a los valores p para las pruebas t individuales de los coeficientes de regresión.

Ejercicios suplementarios



9.24 Ingesta de biotina en pollos Grupos de pollos de 10 días de edad se asignaron al azar a siete grupos de tratamiento en los que una dieta basal se suplementó con 0, 50, 100, 150, 200, 250 o 300 microgramos/kilogramo ($\mu\text{g/kg}$) de biotina. La tabla siguiente da el promedio de ingesta de biotina (x) en microgramos por día y el promedio de aumento de peso (y) en gramos por día.⁵

Biotina agregada	Ingesta de biotina, x	Aumento de peso, y
0	.14	8.0
50	2.01	17.1
100	6.06	22.3
150	6.34	24.4
200	7.15	26.5
250	9.65	23.4
300	12.50	23.3

En la salida impresa *MINITAB*, el modelo polinomial de segundo orden

$$E(y) = \beta_0 + \beta_1 x + \beta_2 x^2$$

está ajustado a los datos. Use la salida impresa para contestar las preguntas.

- a. ¿Cuál es la recta de mínimos cuadrados ajustada?
- b. Encuentre R^2 e interprete su valor.
- c. ¿Los datos dan suficiente evidencia para concluir que el modelo aporta información significativa para predecir y ?
- d. Encuentre los resultados de la prueba de $H_0: \beta_2 = 0$. ¿Hay suficiente evidencia para indicar que el modelo cuadrático da un mejor ajuste a los datos que un modelo lineal simple?
- e. ¿Las gráficas residuales indican que cualquiera de las suposiciones de regresión han sido violadas? Explique.

Salida impresa MINITAB para el ejercicio 9.24

Análisis de regresión: y contra x, x-sq

The regression equation is
 $y = 8.59 + 3.82 x - 0.217 x\text{-sq}$

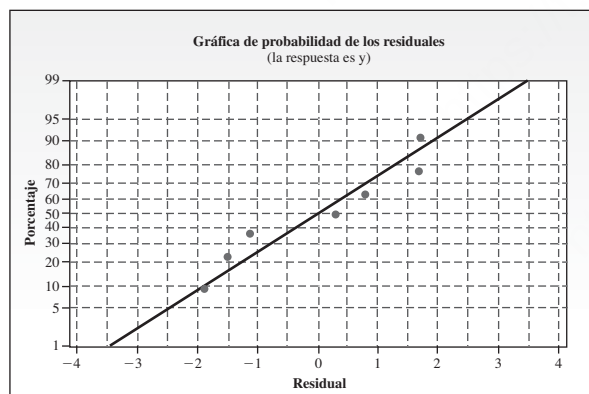
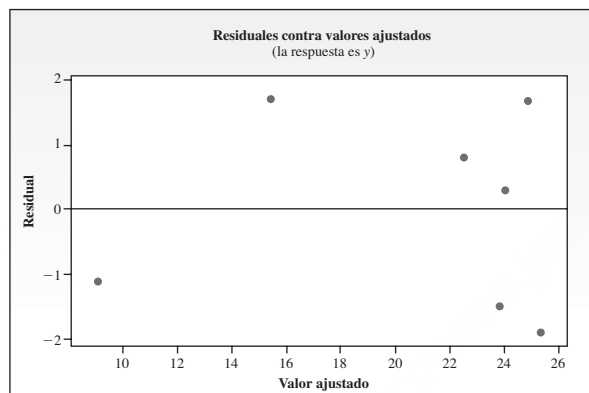
Predictor	Coef	SE Coef	T	P
Constant	8.585	1.641	5.23	0.006
x	3.8208	0.5683	6.72	0.003
x-sq	-0.21663	0.04390	-4.93	0.008

S = 1.83318 R-Sq = 94.4% R-Sq(adj) = 91.5%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	224.75	112.37	33.44	0.003
Residual Error	4	13.44	3.36		
Total	6	238.19			

Source	DF	Seq SS
x	1	142.92
x-sq	1	81.83

**DATOS**

EX0925

9.25 Publicidad y ventas Una tienda de departamentos realizó un experimento con el fin de investigar los efectos de gastos de publicidad en las ventas semanales de sus departamentos de ropa para

caballeros, para niños y para damas. Cinco semanas de observación se seleccionaron al azar de cada departamento y un presupuesto de publicidad x_1 (en cientos de dólares) se asignó a cada uno. Las ventas semanales (en miles de dólares) se muestran en la tabla siguiente para cada uno de los 15 periodos de venta de una semana. Si esperamos que las ventas semanales $E(y)$ estén linealmente relacionadas con el gasto de publicidad x_1 y si esperamos que las pendientes de las rectas correspondientes a los tres departamentos difieran, entonces un modelo apropiado para $E(y)$ es

$$E(y) = \beta_0 + \underbrace{\beta_1 x_1}_{\substack{\text{variable} \\ \text{cuantitativa} \\ \text{"gasto en} \\ \text{publicidad"}}} + \underbrace{\beta_2 x_2 + \beta_3 x_3}_{\substack{\text{variables ficticias} \\ \text{usadas para introducir} \\ \text{la variable cualitativa} \\ \text{"departamento" en el} \\ \text{modelo}}} + \underbrace{\beta_4 x_1 x_2 + \beta_5 x_1 x_3}_{\substack{\text{términos de} \\ \text{interacción} \\ \text{que introducen} \\ \text{diferencias en} \\ \text{pendientes}}}$$

donde

x_1 = gasto en publicidad

$x_2 = \begin{cases} 1 & \text{si el departamento B de ropa para niños} \\ 0 & \text{si no lo es} \end{cases}$

$x_3 = \begin{cases} 1 & \text{si el departamento C de ropa para damas} \\ 0 & \text{si no lo es} \end{cases}$

Departamento	Gasto en publicidad (cientos de dólares)				
	1	2	3	4	5
Ropa para caballeros A	5.2	5.9	7.7	7.9	9.4
Ropa para niños B	8.2	9.0	9.1	10.5	10.5
Ropa para damas C	10.0	10.3	12.1	12.7	13.6

- Encuentre la ecuación de la recta que relacione $E(y)$ con el gasto en publicidad x_1 para el departamento A de ropa para caballeros. [SUGERENCIA: De acuerdo con el código usado para las variables ficticias, el modelo representa ventas medias $E(y)$ para el departamento A de ropa para caballeros cuando $x_2 = x_3 = 0$. Sustituya $x_2 = x_3 = 0$ en la ecuación para $E(y)$ para hallar la ecuación de esta recta.]
- Encuentre la ecuación de la recta que relacione $E(y)$ con x_1 para el departamento B de ropa para niños. [SUGERENCIA: De acuerdo con el código, el modelo representa $E(y)$ para el departamento de ropa para niños cuando $x_2 = 1$ y $x_3 = 0$.]
- Encuentre la ecuación para la recta que relacione $E(y)$ con x_1 para el departamento C de ropa para damas.

- d. Encuentre la diferencia entre los puntos de intersección de las rectas $E(y)$ correspondientes a los departamentos B de ropa para niños y A de ropa para caballeros.
- e. Encuentre la diferencia de pendientes entre las rectas $E(y)$ correspondientes a los departamentos C de ropa para damas y A de ropa para caballeros.
- f. Consulte la parte e. Suponga que se desea probar la hipótesis nula de que las pendientes de las rectas correspondientes a los tres departamentos son iguales. Exprese esto como una prueba de hipótesis acerca de uno o más parámetros del modelo.

9.26 Publicidad y ventas, continúa Consulte el ejercicio 9.25. Use un paquete de software de computadora para efectuar el análisis de regresión múltiple y obtener gráficas de diagnóstico si es posible.

- a. Comente sobre el ajuste del modelo, usando el análisis de varianza de prueba F , R^2 y las gráficas de diagnóstico para comprobar las suposiciones de regresión.
- b. Encuentre la ecuación de predicción y grafique las rectas de ventas de los tres departamentos.
- c. Examine las gráficas de la parte b. ¿Parecen diferir las pendientes de las rectas correspondientes a los departamentos B de ropa para niños y A de ropa para caballeros? Pruebe la hipótesis nula de que las pendientes no difieren ($H_0: \beta_4 = 0$) contra la hipótesis alternativa de que las pendientes son diferentes.
- d. ¿Los términos de interacción del modelo son significativos? Use los métodos descritos en la sección 9.5 para probar $H_0: \beta_4 = \beta_5 = 0$. ¿Los resultados de esta prueba sugieren que el modelo ajustado debe ser modificado?
- e. Escriba una explicación breve de las implicaciones prácticas de este análisis de regresión.

DATOS
EX0927

9.27 Demanda de servicio eléctrico Un estudio a corto plazo se realizó para investigar el efecto de la temperatura diaria mensual x_1 y el costo por kilowatt-hora x_2 en el consumo diario medio de electricidad (en kilowatt-horas, kWh) por familia. Los investigadores esperaban que la demanda de electricidad subiera en tiempo frío (debido a calefacción), bajara cuando el clima fuera moderado y subiera de nuevo cuando la temperatura subiera y hubiera necesidad de poner a funcionar el aire acondicionado. Esperaban que la demanda disminuyera a medida que aumentara el costo por kilowatt-hora, lo que refleja mayor atención a la conservación. Los datos estuvieron disponibles para 2 años, un periodo en el que

el costo por kilowatt-hora x_2 aumentó debido al creciente costo del combustible. La compañía ajustó el modelo

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_2 + \beta_4 x_1 x_2 + \beta_5 x_1^2 x_2$$

a los datos que se ven en la tabla siguiente. La salida impresa de *Excel* para este problema de regresión múltiple también se ilustra.

Precio por kWh, x_2	Temperatura diaria y consumo	Consumo diario medio (kWh) por familia					
8¢	Temperatura diaria media (°F), x_1	31	34	39	42	47	56
		62	66	68	71	75	78
	Consumo diario medio, y	55	49	46	47	40	43
		41	46	44	51	62	73
10¢	Temperatura diaria media, x_1	32	36	39	42	48	56
		62	66	68	72	75	79
	Consumo diario medio, y	50	44	42	42	38	40
		39	44	40	44	50	55

Salida de *Excel* para el ejercicio 9.27

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.948
R Square	0.898
Adjusted R Square	0.870
Standard Error	2.908
Observations	24

ANOVA

	df	SS	MS	F	Significance F
Regression	5	1346.448	269.290	31.852	0.000
Residual	18	152.177	8.454		
Total	23	1498.625			

	Coefficients	Standard Error	t Stat	P-value
Intercept	325.606	83.064	3.920	0.001
x1	-11.383	3.239	-3.515	0.002
x1-sq	0.113	0.029	3.854	0.001
x2	-21.699	9.224	-2.352	0.030
x1x2	0.873	0.359	2.433	0.026
x1sqx2	-0.009	0.003	-2.723	0.014

- a. ¿Los datos dan suficiente evidencia para indicar que el modelo aporta información para la predicción del consumo diario medio en kilowatt-hora por familia? Pruebe al nivel de significancia de 5%.
- b. Grafique la curva que describe $\hat{y}$ como función de la temperatura x_1 cuando el costo por kilowatt-hora es $x_2 = 8¢$. Construya una gráfica similar para el caso cuando $x_2 = 10¢$ por kilowatt-hora. ¿Son diferentes las curvas de consumo?

- c. Si el costo por kilowatt-hora no es importante para predecir el uso, entonces en el modelo no se necesitan los términos que contienen x_2 . Por tanto, la hipótesis nula $H_0: x_2$ no aporta información para la predicción de y y es equivalente a la hipótesis nula $H_0: \beta_3 = \beta_4 = \beta_5 = 0$ (si $\beta_3 = \beta_4 = \beta_5 = 0$, los términos que contienen x_2 desaparecen del modelo). La salida impresa MINITAB, obtenida al ajustar el modelo reducido

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2$$

a los datos, se muestra aquí. Use los métodos de la sección 9.5 para determinar si el precio por kilowatt-hora x_2 aporta información significativa para la predicción de y .

Salida de Excel para el ejercicio 9.27

SUMMARY OUTPUT

Regression Statistics					
Multiple R	0.8304				
R Square	0.6896				
Adjusted R Square	0.6601				
Standard Error	4.7063				
Observations	24				
ANOVA					
	df	SS	MS	F	Significance F
Regression	2	1033.490	516.745	23.330	0.000
Residual	21	465.135	22.149		
Total	23	1498.625			
Coefficients					
		Standard Error	t Stat	P-value	
Intercept	130.009	14.876	8.740	0.000	
x1	-3.502	0.579	-6.049	0.000	
x1-sq	0.033	0.005	6.349	0.000	

- d. Compare los valores de $R^2(\text{adj})$ para el ajuste de los dos modelos en este ejercicio. ¿Cuál de los dos modelos recomendaría usted?

DATOS
EX0928

9.28 Concentración de mercurio en

delfines Debido a que los delfines (y otros grandes mamíferos marinos) son considerados como los principales depredadores en la cadena alimenticia marina, las concentraciones de metales pesados en delfines de franjas se midieron como parte de un estudio de contaminación marina. Se espera que la concentración de mercurio, el metal pesado que se menciona en este estudio, difiera en machos y hembras porque el mercurio en una hembra aparentemente se transmite a su descendencia durante la gestación y la lactancia. Este estudio comprendió 28 machos entre las edades de .21 y 39.5 años y 17 hembras de edades entre .80 y 34.5 años. Para los datos de la tabla,

x_1 = Edad del delfín (en años)

$$x_2 = \begin{cases} 0 & \text{si hembra} \\ 1 & \text{si macho} \end{cases}$$

y = Concentración de mercurio
(en microgramos/gramo) en el hígado

y	x ₁	x ₂	y	x ₁	x ₂
1.70	.21	1	481.00	22.50	1
1.72	.33	1	485.00	24.50	1
8.80	2.00	1	221.00	24.50	1
5.90	2.20	1	406.00	25.50	1
101.00	8.50	1	252.00	26.50	1
85.40	11.50	1	329.00	26.50	1
118.00	11.50	1	316.00	26.50	1
183.00	13.50	1	445.00	26.50	1
168.00	16.50	1	278.00	27.50	1
218.00	16.50	1	286.00	28.50	1
180.00	17.50	1	315.00	29.50	1
264.00	20.50	1			

y	x ₁	x ₂	y	x ₁	x ₂
241.00	31.50	1	142.00	17.50	0
397.00	31.50	1	180.00	17.50	0
209.00	36.50	1	174.00	18.50	0
314.00	37.50	1	247.00	19.50	0
318.00	39.50	1	223.00	21.50	0
2.50	.80	0	167.00	21.50	0
9.35	1.58	0	157.00	25.50	0
4.01	1.75	0	177.00	25.50	0
29.80	5.50	0	475.00	32.50	0
45.30	7.50	0	342.00	34.50	0
101.00	8.05	0			
135.00	11.50	0			

- a. Escriba un modelo de segundo orden que relacione y con x_1 y x_2 . Considere curvatura en la relación entre edad y concentración de mercurio, así como para una interacción entre género y edad.

Use un software de computadora para efectuar el análisis de regresión múltiple. Consulte la salida impresa para contestar estas preguntas.

- b. Comente sobre el ajuste del modelo, usando estadísticos relevantes de la salida impresa.
- c. ¿Cuál es la ecuación de predicción para predecir la concentración de mercurio en un delfín hembra como función de su edad?
- d. ¿Cuál es la ecuación de predicción para predecir la concentración de mercurio en un delfín macho como función de su edad?
- e. ¿El término cuadrático de la ecuación de predicción para hembras hace aportación significativa a la predicción de la concentración de mercurio en un delfín hembra?

- f. ¿Hay otras conclusiones importantes que usted sienta que no se consideraron respecto a la ecuación de predicción ajustada?

DATOS
EX0929

9.29 Puntuaciones en el GRE Se registraron las puntuaciones de razonamiento cuantitativo en el Examen de Registro de Graduados (GRE, siglas de Graduate Record Examination)⁶ para estudiantes admitidos a tres programas diferentes para graduados en una universidad local. Estos datos fueron analizados en el ejercicio 7.27 usando el análisis de varianza para un diseño completamente aleatorizado.

Programa para graduados					
Ciencias de la vida		Ciencias físicas		Ciencias sociales	
630	660	660	760	440	540
640	660	640	670	330	450
470	480	720	700	670	570
600	650	690	710	570	530
580	710	530	450	590	630

- a. Escriba el modelo teórico que relaciona la puntuación en el GRE con la variable cualitativa “programa para graduados” usando dos variables ficticias (indicadoras) para representar los tres programas para graduados.
- b. Use un paquete de cómputo para analizar los datos con un análisis de regresión múltiple. ¿Hay suficiente evidencia para indicar una diferencia en las puntuaciones promedio entre los estudiantes que han sido admitidos a los tres programas para graduados? Use $\alpha = .05$.
- c. Comente sobre el ajuste del modelo y cualesquiera suposiciones de regresión que puedan haber sido violadas. Resuma sus resultados en un informe, incluyendo salidas impresas y gráficas de ser posible.

DATOS
EX0930

9.30 En camino otra vez Hasta fechas recientes, las llantas de alto rendimiento se ajustaban principalmente en vehículos deportivos o de lujo. Ahora se han convertido en estándar en muchos sedanes de uso diario. Los mayores niveles de manejo y agarre al pavimento han aparecido a expensas del desgaste en el dibujo o superficie de rodamiento. Los datos que siguen se han resumido de un informe sobre llantas para todas las temporadas publicado por *Consumer Reports*,⁷ en el que se evaluaron varios aspectos del rendimiento para $n = 26$ llantas diferentes y fueron:

y = puntuación general x_1 = frenado en seco
 x_2 = frenado en húmedo x_3 = manejo
 x_4 = resistencia al rodamiento x_5 = duración del dibujo

Marca y modelo	Precio (\$)	y	x_1	x_2	x_3	x_4	x_5
Michelin HydroEdge	116	84	4	5	4	3	5
Continental ProContact							
ECOPLUS	90	82	4	4	3	5	3
Michelin Energy Saver A/S	120	82	4	4	3	5	3

Marca y modelo	Precio (\$)	y	x_1	x_2	x_3	x_4	x_5
Hankook Optimo H727	96	82	4	4	3	3	3
Pirelli P4 Four Seasons	98	80	4	5	3	2	2
Goodyear Assurance TripleTred	121	80	4	4	3	3	3
Pirelli Cinturato P5	98	78	3	5	3	2	3
Kumho Solus KR21	77	78	4	5	3	3	4
Maxxis Escapade MA-T1	77	76	3	4	4	3	4
Toyo Extensa A/S	99	76	3	4	3	3	3
Cooper GFE	83	76	3	4	3	4	2
Toyo Versado LX	91	76	4	4	3	3	5
BFGoodrich Traction T/A T	90	74	4	4	4	2	3
General Altimax RT	77	74	3	5	3	3	3
Yokohama Avid TRZ	81	74	4	5	3	2	3
Dayton Quadra LE	74	74	3	4	3	3	3
Cooper CS4 Touring	86	72	3	5	3	2	3
Uniroyal Tiger Paw Tour SR	77	72	4	4	2	5	2
Yokohama Avid Touring-S	70	70	4	4	3	4	4
Cooper Lifeline GLS	78	68	3	4	3	4	2
Yokohama Avid T4	85	66	3	4	3	3	2
Bridgestone Turanze EL400	97	66	3	4	2	3	2
Falken Sincera SN828	82	64	3	5	3	3	1
Dunlop SP 60	79	64	3	3	2	2	4
Sumitomo HTR T4	67	64	3	4	3	4	2
Firestone FR710	80	60	3	5	3	3	4

Las variables x_1 a la x_5 están codificadas usando la escala de 5 = excelente, 4 = muy buena, 3 = buena, 2 = regular y 1 = mala.

- a. Use un programa de su elección para hallar la matriz de correlación para las variables bajo estudio, que incluyan costo. ¿El precio está significativamente correlacionado con cualquiera de las variables de estudio? ¿Cuáles variables parecen estar altamente correlacionadas con y , la puntuación total?
- b. Escriba un modelo para describir y , la puntuación total, como función de las variables x_1 = frenado en seco, x_2 = frenado en húmedo, x_3 = manejo, x_4 = resistencia al rodamiento y x_5 = duración del dibujo.
- c. Use un programa de regresión de su elección para ajustar el modelo completo usando todas las variables predictoras. ¿Qué proporción de la variación en y está explicada por regresión? ¿Esto comunica la impresión de que el modelo explica en forma adecuada la variabilidad inherente en y ?
- d. ¿Cuál variable o variables parecen ser buenas variables predictoras de y ? ¿Cómo podría refinarse el modelo en vista de estos resultados? Use estas variables para reajustar el modelo. ¿Qué proporción de la variación es explicada por este modelo reajustado? Comente sobre lo adecuado de este modelo reducido en comparación con el modelo completo.

DATOS
EX0931

9.31 Atunes Los datos sobre atunes del ejercicio 7.16 se analizaron como un diseño completamente aleatorizado con cuatro tratamientos. No obstante,

también podríamos ver el diseño experimental como un experimento factorial de 2×2 con repeticiones desiguales. Los datos se muestran a continuación.⁸

	Aceite		Agua	
Atún claro	2.56	.62	.99	1.12
	1.92	.66	1.92	.63
	1.30	.62	1.23	.67
	1.79	.65	.85	.69
	1.23	.60	.65	.60
		.67	.53	.60
Atún blanco			1.41	.66
	1.27		1.49	1.29
	1.22		1.29	1.00
	1.19		1.27	1.27
	1.22		1.35	1.28

Fuente: Estudio práctico "Tuna Goes Upscale" Copyright 2001 por Consumers Union of U.S. Inc., Yonkers, NY 10703-1057, organización sin fines de lucro. Reimpreso con permiso de la edición de junio de 2001 de *Consumer Reports*® sólo para fines educativos. No se permite uso comercial ni reproducción. www.ConsumerReports.org

Los datos se pueden analizar usando el modelo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon$$

donde

$x_1 = 0$ si en aceite, 1 si en agua

$x_2 = 0$ si atún claro, 1 si atún blanco

- Demuestre cómo introduciría los datos en una hoja de trabajo de computadora, introduciendo los datos en las columnas para y , x_1 , x_2 y $x_1 x_2$.
- La salida impresa generada por *MINITAB* se muestra a continuación. ¿Cuál es la ecuación de predicción de mínimos cuadrados?

Salida impresa *MINITAB* para el ejercicio 9.31

Análisis de regresión: y contra x1, x2, x1x2

The regression equation is
 $y = 1.15 - 0.251 x_1 + 0.078 x_2 + 0.306 x_1 x_2$

Predictor	Coef	SE Coef	T	P
Constant	1.1473	0.1370	8.38	0.000
x1	-0.2508	0.1830	-1.37	0.180
x2	0.0777	0.2652	0.29	0.771
x1x2	0.3058	0.3330	0.92	0.365

S = 0.454287 R-Sq = 11.9% R-Sq(adj) = 3.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	0.9223	0.3074	1.49	0.235
Residual Error	33	6.8104	0.2064		
Total	36	7.7328			

- ¿Hay interacción entre el tipo de atún y el tipo de líquido de empaque?
- ¿Cuál, si lo hay, de los principales efectos (tipo de atún y tipo de líquido de empaque) aporta información significativa para la predicción de y ?
- ¿Qué tan bien se ajusta el modelo a los datos? Explique.

9.32 Atún, continúa Consulte el ejercicio 9.31. La hipótesis probada en el capítulo 7, que el promedio de precios para los cuatro tipos de atún es igual, es equivalente a decir que $E(y)$ no cambiará cuando x_1 y x_2 cambien. Esto puede ocurrir sólo cuando $\beta_1 = \beta_2 = \beta_3 = 0$. Use la salida impresa *MINITAB* para la ANOVA de una vía que se muestra a continuación para efectuar la prueba de igualdad de medias de tratamiento. Verifique que esta prueba sea idéntica a la prueba para regresión significativa del ejercicio 9.31.

Salida impresa *MINITAB* para el ejercicio 9.32

ANOVA de una vía: claro en agua, blanco en aceite, blanco en agua, claro en aceite

Source	DF	SS	MS	F	P
Factor	3	0.922	0.307	1.49	0.235
Error	33	6.810	0.206		
Total	36	7.733			

S = 0.4543 R-Sq = 11.93% R-Sq(adj) = 3.92%

DATOS
EX0933

9.33 Control de calidad Un fabricante registró el número de piezas defectuosas (y) producidas en un día determinado por cada uno de 10 operadores de máquinas y también registró la producción promedio por hora (x_1) por cada operador y el tiempo en semanas desde el último servicio a una máquina (x_2).

y	x_1	x_2
13	20	3.0
1	15	2.0
11	23	1.5
2	10	4.0
20	30	1.0
15	21	3.5
27	38	0
5	18	2.0
26	24	5.0
1	16	1.5

La salida impresa que sigue resultó cuando estos datos se analizaron usando el paquete *MINITAB* utilizando el modelo:

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

Salida de *MINITAB* para el ejercicio 9.33

Análisis de regresión: y contra x1, x2

The regression equation is
 $y = -28.4 + 1.46 x_1 + 3.84 x_2$

Predictor	Coef	SE Coef	T	P
Constant	-28.3906	0.8273	-34.32	0.000
x1	1.46306	0.02699	54.20	0.000
x2	3.8446	0.1426	26.97	0.000

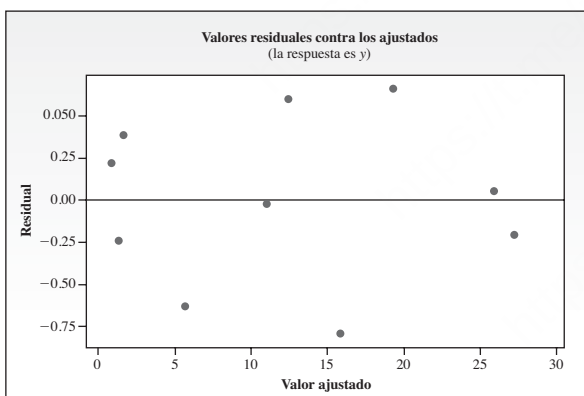
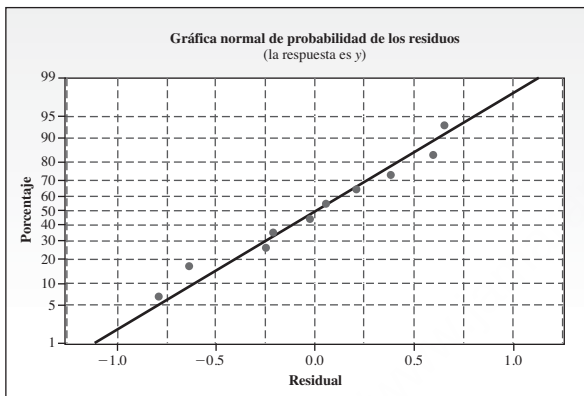
S = 0.548433 R-Sq = 99.8% R-Sq(adj) = 99.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	884.79	442.40	1470.84	0.000
Residual Error	7	2.11	0.30		
Total	9	886.90			

Source	DF	Seq SS
x1	1	666.04
x2	1	218.76

- Interprete R^2 y comente sobre el ajuste del modelo.
- ¿Hay evidencia para indicar que el modelo contribuye de manera significativa a la predicción de y al nivel de significancia de $\alpha = .01$?
- ¿Cuál es la ecuación de predicción que relacione $\hat{y}$ y x_1 cuando $x_2 = 4$?
- Use la ecuación de predicción ajustada para predecir el número de piezas defectuosas producidas por un operador cuya producción promedio por hora es 25 y cuya máquina ha recibido servicio hace tres semanas.
- ¿Qué nos dicen las gráficas residuales acerca de la validez de las suposiciones de regresión?



DATOS
EX0934

9.34 Corrosión de metal y ácidos del suelo En una investigación para determinar la relación entre el grado de corrosión de metal y el tiempo que el metal se exponga a la acción de ácidos del suelo, el porcentaje de corrosión y tiempo de exposición se midieron semanalmente.

y	0.1	0.3	0.5	0.8	1.2	1.8	2.5	3.4
x	1	2	3	4	5	6	7	8

Los datos se ajustaron usando el modelo cuadrático,
 $E(y) = \beta_0 + \beta_1x + \beta_2x^2$, con los siguientes resultados.

Salida de *Excel* para el ejercicio 9.34

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9993
R Square	0.9985
Adjusted R Square	0.9979
Standard Error	0.0530
Observations	8

ANOVA

	gl	SS	MS	F	Significance F
Regression	2	9.421	4.710	1676.610	0.000
Residual	5	0.014	0.003		
Total	7	9.435			

	Coefficients	Standard Error	t Stat	P-value
Intercept	0.196	0.074	2.656	0.045
x	-0.100	0.038	-2.652	0.045
x -sq	0.062	0.004	15.138	0.000

- ¿Qué porcentaje de la variación total es explicada por la regresión cuadrática de y en x ?
- ¿La regresión en x y x^2 es significativa al nivel de significancia de $\alpha = .05$?
- ¿El coeficiente de regresión lineal es significativo cuando x^2 está en el modelo?
- ¿El coeficiente de regresión cuadrática es significativo cuando x está en el modelo?
- Los datos se ajustaron a un modelo lineal sin el término cuadrático con los resultados que siguen. ¿Qué se puede decir acerca de la contribución del término cuadrático cuando está incluido en el modelo?

Salida de *Excel* para el ejercicio 9.34

SUMMARY OUTPUT

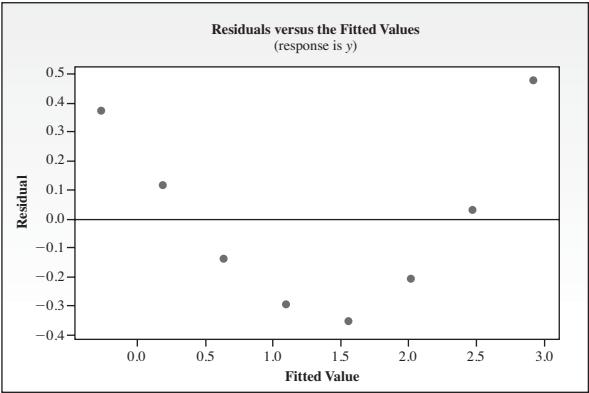
Regression Statistics	
Multiple R	0.9645
R Square	0.9303
Adjusted R Square	0.9187
Standard Error	0.3311
Observations	8

ANOVA

	gl	SS	MS	F	Significance F
Regression	1	8.777	8.777	80.052	0.000
Residual	6	0.658	0.110		
Total	7	9.435			

	Coefficients	Standard Error	t Stat	P-value
Intercept	-0.732	0.258	-2.838	0.030
x	0.457	0.051	8.947	0.000

- La gráfica de los residuales del modelo de regresión lineal en la parte e muestra un patrón específico. ¿Cuál es el término del modelo que parece estar faltando?



DATOS
EX0935

9.35 Manejando su dinero Una corporación particular de ahorros y préstamos está interesada en determinar qué tan bien se puede predecir la cantidad de dinero en ahorros familiares, usando las tres variables independientes de ingreso anual, número de personas en la familia y área en la que viva la familia. Suponga que hay dos áreas específicas de interés para la corporación. Se recolectaron los datos siguientes, donde

y = Cantidad en todas las cuentas de ahorros
 x_1 = Ingreso anual
 x_2 = Número de personas en la familia
 x_3 = 0 si en Área 1; 1 si no

Tanto y como x_1 se registraron en unidades de \$1000.

y	x_1	x_2	x_3
0.5	19.2	3	0
0.3	23.8	6	0
1.3	28.6	5	0
0.2	15.4	4	0
5.4	30.5	3	1
1.3	20.3	2	1
12.8	34.7	2	1
1.5	25.2	4	1
0.5	18.6	3	1
15.2	45.8	2	1

La siguiente salida impresa de computadora resultó cuando se analizaron los datos.

Análisis de regresión: y contra x_1 , x_2 , x_3
The regression equation is
 $y = -3.11 + 0.503 x_1 - 1.61 x_2 - 1.15 x_3$

Predictor	Coef	SE Coef	T	P
Constant	-3.112	3.600	-0.86	0.421
x_1	0.50314	0.07670	6.56	0.001
x_2	-1.6126	0.6579	-2.45	0.050
x_3	-1.155	1.791	-0.64	0.543

$S = 1.89646$ $R\text{-Sq} = 92.2\%$ $R\text{-Sq}(\text{adj}) = 88.4\%$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	256.621	85.540	23.78	0.001
Residual Error	6	21.579	3.597		
Total	9	278.200			

Source **DF** **Seq SS**
 x_1 1 229.113
 x_2 1 26.012
 x_3 1 1.496

- a. Interprete R^2 y comente sobre el ajuste del modelo.
- b. Pruebe para una regresión significativa de y en x_1 , x_2 y x_3 al nivel de significancia de 5%.
- c. Pruebe la hipótesis $H_0 : \beta_3 = 0$ contra $H_a : \beta_3 \neq 0$ usando $\alpha = .05$. Comente sobre los resultados de su prueba.
- d. ¿Qué se puede decir de la utilidad de x_3 como variable de predicción en este problema?

Análisis de datos categóricos

OBJETIVOS GENERALES

Numerosos tipos de estudios y experimentos resultan en variables de respuesta cualitativas y no cuantitativas, de modo que las respuestas pueden ser clasificadas pero no cuantificadas. Los datos de estos experimentos están formados por el conteo o número de observaciones que caen en cada una de las categorías de respuesta incluidas en el experimento. En este capítulo, nos ocupamos de métodos para analizar datos categóricos.

ÍNDICE DEL CAPÍTULO

- Introducción (10.1)
- Una descripción del experimento (10.2)
- Estadístico ji cuadrada de Pearson (10.3)
- Prueba de probabilidades de celda especificada: la prueba de bondad de ajuste (10.4)
- Tablas de contingencia: una clasificación de dos vías (10.5)
- Procedimiento *post hoc* (10.6)
- Prueba exacta de Fisher (10.7)
- Prueba de McNemar (10.8)
- Comparación de varias poblaciones multinomiales: una clasificación de dos vías con totales de fila o columna fijos (10.9)
- La equivalencia de pruebas estadísticas (10.10)
- Otras aplicaciones del estadístico de prueba ji cuadrada (10.11)



NECESITO SABER...

Cómo determinar el número apropiado de grados de libertad

INTRODUCCIÓN

En la actualidad, la evaluación es considerada el medio principal para optimizar los procesos y productos de la enseñanza-aprendizaje; a diferencia de los estudios no evaluativos cuya meta es generalmente comprobar hipótesis. El éxito de los estudios evaluativos radica en la sensibilidad de los instrumentos de medición y su poder para establecer estándares *normativos* o *de criterio* de excelencia. Por ello es conveniente anticipar las limitaciones y potencialidades de uno u otro tipo de estudio a partir del problema o necesidad que pretende solucionarse.

Los modelos estadísticos que se desarrollan en este capítulo se concentran más en los usos de la investigación cualitativa, aunque algunos de ellos también se aplican en la cuantitativa, siempre y cuando la información se pueda representar con datos categóricos.

Estudios evaluativos, un enfoque actual

El movimiento hacia una nueva cultura de la evaluación considera la investigación evaluativa como un elemento indispensable para la emisión de juicios valorativos y toma de decisiones para la mejora de los procesos de enseñanza-aprendizaje.

Los resultados que arrojan los estudios evaluativos permiten el acceso a la información educativa, facilitan la rendición de cuentas para acceder a los resultados y al financiamiento y generan alternativas para optimizar planes y programas curriculares que impacten positivamente en los sistemas educativos. Es por ello que hoy día los resultados evaluativos se reconocen como protagonistas dentro de dicho movimiento, centrando los esfuerzos en la diversificación de sus estrategias para asegurar la confiabilidad validez y en el intercambio de experiencias de evaluación entre instituciones educativas afines.

Definiciones y conceptos

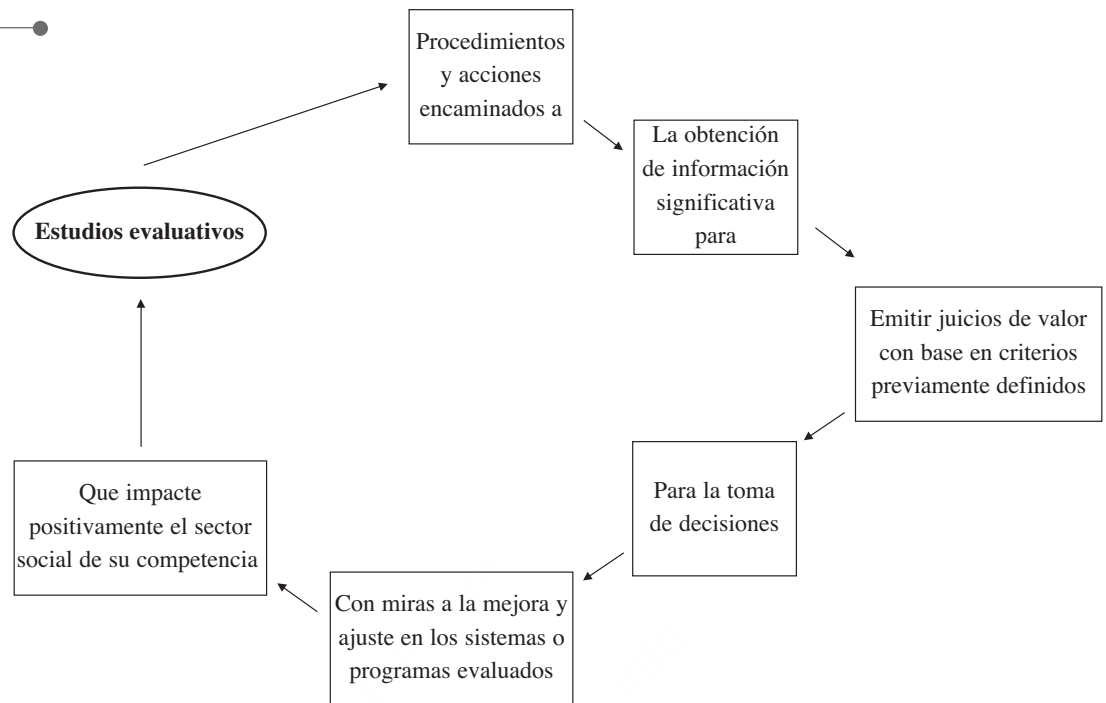
La literatura sobre evaluación es amplia y variada, la evaluación se ha desarrollado bajo múltiples enfoques que han sido determinantes para la evolución del concepto. Por tanto, en el ámbito de la educación existen distintas definiciones, tipos, dimensiones y conceptos relacionados que varían según el marco de referencia y la intención evaluativa de quien la promueve. Sin embargo, para el desarrollo del presente capítulo, el concepto de evaluación es (véase la figura 10.1):

El proceso sistemático que consiste en la recolección de información pertinente, que orienta a la emisión de juicios valorativos respecto de uno o más atributos de los elementos del contexto educativo. La información obtenida es procesada, analizada y contrastada con referentes claramente establecidos y conduce a una toma de decisiones educativas con miras al ajuste y mejora de los planes curriculares y políticas educativas.

Consideraciones para los estudios evaluativos

1. Los principios psicopedagógicos sustentados por las teorías del aprendizaje desde un enfoque constructivista.
2. Los procesos cognoscitivos y las habilidades para el aprendizaje de los individuos, que son destinatarios de un estudio ya en los ámbitos educativo, de la salud y comunitario.

FIGURA 10.1



3. Las finalidades educativas o intenciones del estudio definidas en las políticas sociales, educativas o de salud vigentes, con el fin de buscar congruencia y equilibrio con el modelo de evaluación que sustenta la práctica.
4. Las necesidades de formación, ajustes o mejora en el contexto inmediato, el universo social-profesional-laboral donde el estudio de evaluación se inserta.
5. La caracterización del individuo, considerando edad, nivel de escolaridad, necesidades, expectativas e intereses, entre otros elementos que definen una muestra poblacional.

El trabajo de investigación que se ha de realizar debe considerar los referentes anteriores, como puntos de partida, para formular las preguntas de trabajo o de investigación que van a definir la aplicación de una u otra metodología en el estudio evaluativo.

Los diferentes objetos de la evaluación

Los estudios evaluativos pueden dirigirse a una amplia gama de fenómenos sociales, educativos, comunitarios y del ámbito de la salud, entre otros.

La evaluación del desempeño académico

Es un tipo de evaluación especializada en la obtención de información confiable sobre el nivel de aprendizaje que ha alcanzado un individuo, a través de instrumentos de medición válidos,

estandarizados, objetivos y relevantes que se reportan de manera cuantitativa y pueden interpretarse de manera cualitativa para la emisión de juicios valorativos y la toma de decisiones, ya sea en el contexto educativo, social, laboral o de la salud.

Los estudios evaluativos se realizan para valorar, en el ámbito educativo: a) los requisitos de ingreso a un programa, b) el impacto de un programa en su fase intermedia y reorientar, si es el caso, las acciones curriculares o tomar medidas para favorecer los procesos de aprendizaje de los destinatarios y c) el nivel de conocimientos al egresado, como testimonio del nivel de logro de las intenciones educativas del programa o como coadyuvante para la acreditación-certificación de estudios.

En el contexto laboral, la evaluación puede aplicarse para: a) fines de acreditación de estudios trancos, b) cuando la experiencia laboral ha formado competencias y requieren ser evaluadas a criterio, para definir si el individuo puede ser promovido o certificado a partir de la medición de sus habilidades adquiridas en la práctica con referencia a un estándar.

Los estudios evaluativos que se desarrollan en la actualidad han tenido un fuerte impacto en las prácticas educativas y de salud, también en las políticas sociales, públicas y privadas, pues ha logrado enriquecer las alternativas de medición y evaluación, que hasta hace unas décadas se hacían preferentemente bajo modelos cuantitativos.

La evaluación de programas

Los estudios evaluativos que se dirigen a evaluar programas facilitan los siguientes procesos:

- La comparación, mejoramiento, acreditación o certificación de dichos programas.
- Generalización de su efectividad en diferentes ámbitos y niveles socioeconómicos, estratos sociales, contextos educativos y de salud.
- Participación ciudadana en el equipo de trabajo para un mejor desarrollo y efectividad del programa.
- Toma de decisiones para el ajuste de la mejora.
- Evolución del desempeño de los participantes.

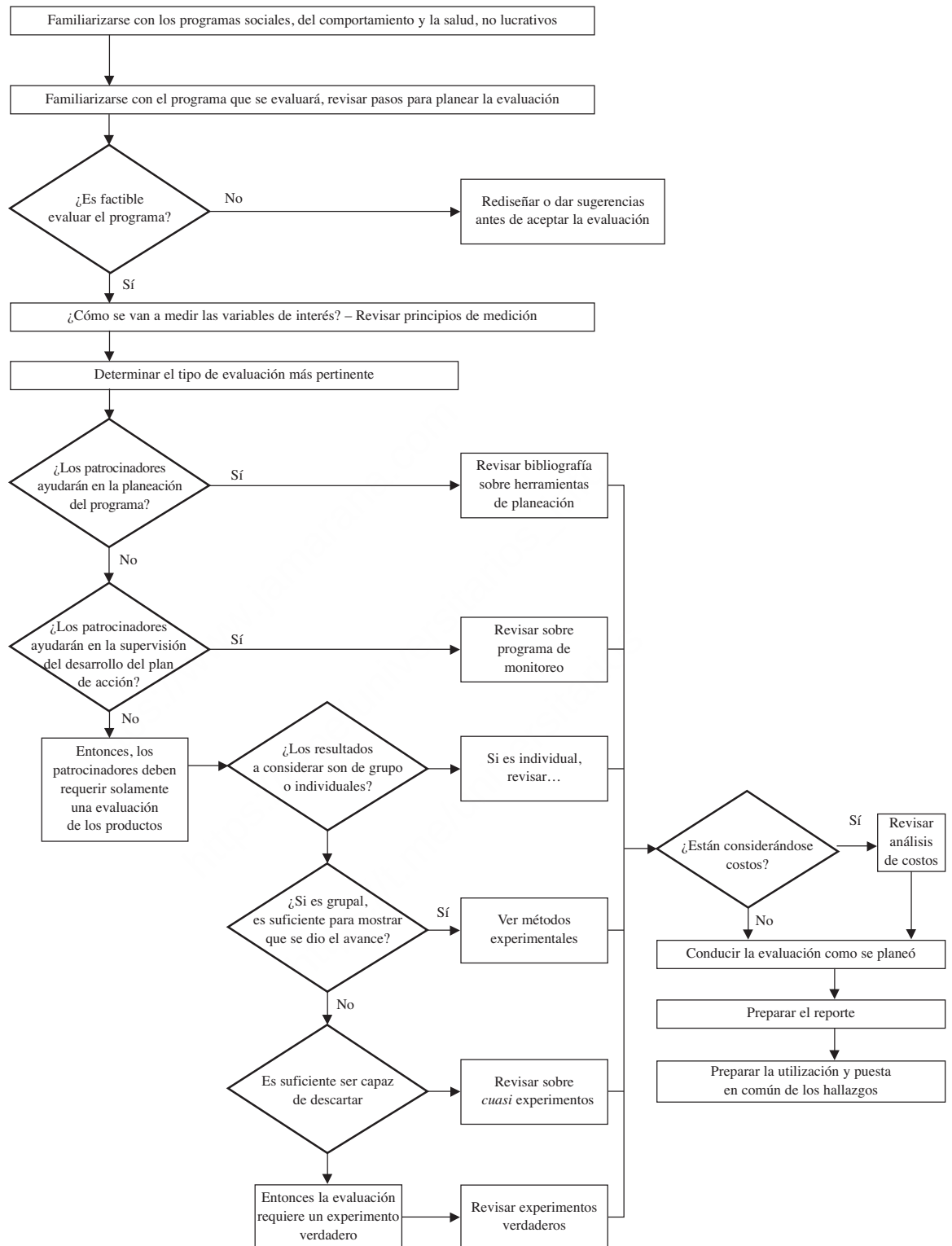
Los estudios que evalúan programas pueden dirigirse hacia el conocimiento de la etapa de desarrollo del programa en un punto de espacio y tiempo de corte, y a partir de los propósitos de quien lo patrocina, considerando:

- Las necesidades detectadas del grupo o comunidad a quien va dirigido el programa.
- Determinar la etapa y proceso en función de los objetivos, costos y beneficios cubiertos.
- El resultado o producto final del programa.

Estudios evaluativos: procedimientos generales

Como ya se mencionó, este tipo de estudios se aplica para evaluar programas previamente diseñados, con la posibilidad de que el evaluador participe desde las primeras etapas del diseño del programa, coordinando los procedimientos de planeación y diseño con los elementos que se van a considerar para su posterior evaluación, de tal manera que la **evaluación** y la **planeación** son procesos similares e interactivos, necesarios para alcanzar el éxito de un programa de beneficio social.

FIGURA 10.2



El evaluador debe:

1. Familiarizarse con los programas sociales no lucrativos, del comportamiento y de la salud.
2. Seguir las etapas y el proceso sugeridos para planear la evaluación.
3. Tener claramente definido el propósito general del proyecto.
4. Reconocer a quiénes va dirigido el programa y el ámbito sociopolítico que subyace al mismo.

Procedimientos:

Paso 1. Identificar al personal y las figuras participantes de importancia para el programa:

- Personas o instituciones que tienen un interés genuino en que el programa tenga éxito y, por consiguiente, consideran la evaluación como un factor necesario e importante.
- Ubicar a los *patrocinadores* del programa.
- Conocer en la medida de lo posible a los destinatarios o beneficiarios del mismo.

Paso 2. Establecer reuniones previas, con el fin de ubicar:

- ¿Quién requiere la evaluación?
- ¿Qué tipo de evaluación desea realizarse?
- ¿Por qué se desea este tipo de evaluación?
- ¿Cuándo y a quiénes deben entregarse los resultados?
- ¿Con qué recursos se dispone?

Paso 3. Decidir el mejor momento para realizar el estudio evaluativo:

- Considerar la sistematización de las actividades propuestas, horarios, calendarización, agendas y eventos de los participantes en dichos procesos, así como actividades y situaciones de oportunidad o de amenaza al proceso de evaluación.

Paso 4. Revisión de la literatura:

- Esto incluye la revisión de literatura, revisión documental institucional y estándares internacionales si es el caso, enriqueciendo estas acciones con trabajos de campo, entrevistas con expertos relacionados o con trabajos afines, etcétera.

Paso 5. Definir la metodología, acorde con el tipo de estudio que se determinó realizar, esto incluye:

- Estrategias y diseño.
- Población objetivo.
- Muestra.
- Control y comparación de grupos.
- Selección de instrumentos de medición.
- Recolección y concentración de la información obtenida.
- Análisis estadístico.
- Conclusiones.
- Sugerencias.
- Reporte final.

Paso 6. Detección de fortalezas-debilidades y oportunidades-amenazas del estudio evaluativo, esta detección se hace desde las reuniones de planeación, con la participación de los involucrados en el proceso evaluativo.

Áreas de interés del estudio evaluativo

Los resultados evaluativos pueden enfocarse a ciertas zonas o áreas del programa que está evaluándose, puede ser a una o varias, dependiendo de los propósitos, de tal manera que se puede realizar:

- a) *Evaluación de las necesidades.* Generalmente implica una detección de necesidades en la comunidad a quien va dirigido el programa.
- b) *Evaluación del proceso.* Sirve para determinar el desarrollo del programa; si está cumpliendo las metas planteadas y si ha considerado las características de los destinatarios.
- c) *Evaluación del resultado o producto final.* Esta evaluación es la que más frecuentemente se realiza, y da cuenta del éxito y el impacto final del programa.

Programas susceptibles de evaluación

Sociales

- Sistema de justicia.
- Seguridad pública.
- Protección ciudadana.
- Prevención del delito.
- Rehabilitación social.
- Asistencia social.

Educativos y del comportamiento

- Calidad de la enseñanza y el aprendizaje.
- Integración de niños, adolescentes y adultos con discapacidad en los sistemas educativos regulares, de todos los niveles.
- Programas curriculares innovadores.
- Efectividad de los métodos educativos.
- Desempeño docente.
- Técnicas de modificación de conductas en el aula.
- Estilos de aprendizaje.

Salud

- Calidad de primero, segundo y tercer nivel de los servicios hospitalarios, clínicas y centros de salud.
- Prevención de enfermedades tanto físicas como mentales.
- Prevención y control de conductas adictivas.
- Control de calidad en los procedimientos de los servicios de salud.

Interpretación de los resultados

Existen dos maneras de interpretar los resultados obtenidos en el estudio educativo, una con referencia a la norma y la otra con referencia al *criterio*. Ambas requieren de una precisión de los requerimientos o parámetros, así como de elementos de juicio previamente establecidos a la intervención evaluativa.

En seguimiento de la tradición, la forma para interpretar los resultados arrojados por un instrumento de medida ha sido la que establece la calificación referida a una norma ya sea grupal o bien poblacional. En el primer caso, la referencia es limitada debido a que está dada por las puntuaciones de donde pertenece el destinatario: una población determinada y a veces reducida para hacer inferencias. La calificación que obtenga un individuo se interpretará en función del desempeño alcanzado por el grupo de referencia, calificándolo según la ubicación de sus resultados en el marco total de puntuaciones del grupo.

La evaluación referida a una norma

Está relacionada con el grupo de referencia en el que se haya basado la comparación, por lo que un cambio en el grupo dará como resultado un cambio en la interpretación de la calificación individual. Siendo ésta relativa, no podría iniciar lo que el individuo puede hacer respecto de sus limitaciones curriculares o en sus dominios para el trabajo, solamente llegaría a saber lo que logra con referencia a otros individuos. Su propósito principal es el de ordenar, calificar y catalogar a los sujetos respecto del atributo por rasgo medido en el mismo grupo de referencia. El referente de comparación es el nivel general de un grupo determinado (otros individuos, instituciones, programas y el promedio de la clase). Aplicada al aprendizaje del alumnado, la evaluación establece la comparación entre el rendimiento de cada sustentante con el rendimiento promedio de la clase en la que se sitúa, aunque esta comparación también puede establecerse con el promedio de otras clases, con la media de toda la institución o con la nacional. En contraposición a las evaluaciones con referencia al criterio.

La evaluación referida al criterio

También llamada evaluación criterial, presenta la característica distintiva de proporcionar la ubicación de los sujetos con relación a dominios de conducta observable previamente definidos, centrándose en lo que hacen o son capaces de hacer los sujetos, en lugar de tomar como referencia la conducta de otros sujetos o de la norma del grupo.

Una de las definiciones más aceptadas de las pruebas referidas al criterio es la de Popham (1978), la cual señala que este tipo de evaluación se utiliza para contrastar el estatus absoluto del sujeto respecto de algún dominio de conductas bien definido. Ésta se hace mediante la determinación precisa y concreta de los rendimientos que pretenden alcanzarse. Ello supone la formulación previa de estándares de evaluación pertinentes que faciliten la tarea de determinar si un individuo ha alcanzado los objetivos previstos; los criterios de evaluación deben estar formulados concreta y claramente. Se pretende que el sustentante alcance un determinado nivel en el ciclo escolar o que aprenda a hacer algo previamente acordado, intentando establecer un equilibrio entre los resultados conseguidos y los objetivos propuestos. La idea es evaluar el resultado del aprendizaje tomando como punto de referencia el criterio marcado o las fases en que éste ha podido desglosarse, independientemente del nivel del grupo.

Ambos tipos de evaluación requieren hacer explícitos los términos en que las conductas o rendimientos que el sujeto exhibe durante la medición van a considerarse adecuadas, deseables o aprobadas; sin embargo, en la referencia a normas, generalmente no se determina un dominio inicial o criterio mínimo, sino la distribución normal, que es el parámetro de referencia. Esto a diferencia de los referidos al criterio, donde, de comenzar con una clara especificación y delimitación de conductas de referencia, se facilitará el diseño y el uso pretendido de la medida, lo cual es un elemento central para la fase de construcción del instrumento. Ambos tipos de evaluación son útiles, dependiendo del tipo de proyecto o programa que va a evaluarse.

10.2

UNA DESCRIPCIÓN DEL EXPERIMENTO

Numerosos experimentos resultan en mediciones que son *cualitativas* o *categorías* en lugar de *cuantitativas*; esto es, una *cualidad* o *característica* (más que un valor numérico) se mide para cada unidad experimental. Se puede resumir este tipo de datos al crear una lista de las categorías o características e informar de una **cantidad** del número de mediciones que caen en cada categoría. A continuación veamos algunos ejemplos:

- Las personas se pueden clasificar en cinco categorías de ingreso.
- Un ratón puede responder en una de tres formas a un estímulo.
- Un proceso industrial manufactura artículos que pueden ser clasificados como “aceptables”, “de segunda categoría” o “defectuosos”.

Éstas son algunas de las muchas situaciones en las que el conjunto de datos tiene características apropiadas para el **experimento multinomial**.

EL EXPERIMENTO MULTINOMIAL

- El experimento consta de n intentos idénticos.
- El resultado de cada intento cae en una de k categorías.
- La probabilidad de que el resultado de un solo intento caiga en una categoría particular, por ejemplo i , es p_i y permanece constante de un intento a otro. Esta probabilidad debe ser entre 0 y 1, para cada una de las k categorías y la suma de todas las k probabilidades es $\sum p_i = 1$.
- Los intentos son independientes.
- El experimentador cuenta el número *observado* de resultados en cada categoría, escrito como $O_1, O_2, \dots, O_k$, con $O_1 + O_2 + \dots + O_k = n$.

Se puede visualizar el experimento multinomial si se considera un número k de cajas o **cel-das** en las que n pelotas se lanzan. Los n tiros son independientes y en cada tiro la probabilidad de hacer blanco en la i -ésima caja es la misma. Pero esta probabilidad puede variar de una caja a otra; podría ser más fácil hacer blanco en la caja 1 que en la caja 3 en cada tiro. Una vez que todas las n pelotas se hayan tirado, se cuenta el número de cada caja o **celda**, $O_1, O_2, \dots, O_k$.

Es probable que haya notado la similitud entre el *experimento multinomial* y el *experimento binomial* introducido en el capítulo 3. De hecho, cuando hay $k = 2$ categorías, los dos experimentos son idénticos, excepto por la notación. En lugar de p y q , escribimos p_1 y p_2 para representar las probabilidades para las dos categorías, “éxito” y “fracaso”. En lugar de x y $(n - x)$, escribimos O_1 y O_2 para representar el número observado de “éxitos” y “fracasos”.

Cuando presentamos la variable aleatoria binomial hicimos inferencias acerca del parámetro binomial p (y de manera predeterminada, $q = 1 - p$) usando métodos de muestra grande basados en el estadístico z . En este capítulo ampliamos esta idea para hacer inferencias acerca de los *parámetros multinomiales*, $p_1, p_2, \dots, p_k$, usando un tipo diferente de estadístico. Este estadístico, cuya distribución de muestreo aproximada fue derivada por un estadístico inglés llamado Karl Pearson en 1900, se llama **estadístico ji cuadrada** (o a veces **ji cuadrada de Pearson**).

TIP MI CONSEJO

El experimento multinomial es una extensión del *experimento binomial*. Para un experimento binomio, $k = 2$

ESTADÍSTICO JI CUADRADA DE PEARSON

Suponga que $n = 100$ pelotas se lanzan a las celdas (cajas) y que sabemos que la probabilidad que una pelota caiga en la primera caja es $p_1 = .1$. ¿Cuántas pelotas *esperaríamos* que caigan en la primera caja? De manera intuitiva se esperaría ver $100(.1) = 10$ pelotas. Esto debe recordarnos el promedio o número esperado de éxitos, $\mu = np$, en el experimento binomial. En general, el número esperado de pelotas que caigan en la celda i , escrito como E_i , se puede calcular usando la fórmula

$$E_i = np_i$$

para cualquiera de las celdas $i = 1, 2, \dots, k$.

Ahora suponga que *hacemos una hipótesis* de valores para cada una de las probabilidades $p_1, p_2, \dots, p_k$ y calculamos el número esperado para cada categoría o celda. Si nuestra hipótesis es correcta, las *cantidades observadas de celda* reales, O_i , no deben ser demasiado diferentes de las *cantidades esperadas de celda*, $E_i = np_i$. Cuanto más grandes sean las diferencias, más probable será que la hipótesis sea incorrecta. El *estadístico ji cuadrada de Pearson* utiliza las diferencias $(O_i - E_i)$ al elevar al cuadrado primeramente estas diferencias para eliminar contribuciones negativas y luego forma un promedio *ponderado* de las diferencias cuadradas.

ESTADÍSTICO DE PRUEBA JI CUADRADA DE PEARSON

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

sumadas en todas las k celdas, con $E_i = np_i$.

Aun cuando la prueba matemática está fuera del propósito de este libro, se puede demostrar que cuando n es grande, χ^2 tiene una **distribución ji cuadrada de probabilidad** aproximada en muestreo repetido. Si los conteos de celda esperados hipotéticos son correctos, las diferencias $(O_i - E_i)$ son pequeñas y χ^2 es cercana a 0. Pero, si las probabilidades hipotéticas son incorrectas, grandes diferencias $(O_i - E_i)$ resultan en un valor *grande* de χ^2 . Debe usarse una **prueba estadística de cola derecha** y buscar un valor inusualmente grande del estadístico de prueba.

La distribución ji cuadrada se utiliza en el capítulo 6 para hacer inferencias acerca de una varianza poblacional σ^2 . Al igual que la distribución F , su forma no es simétrica y depende de un número específico de **grados de libertad**. Una vez especificados estos grados de libertad, se puede usar la tabla 5 del Anexo para hallar valores críticos o para limitar el valor p para un estadístico ji cuadrada particular.

Los grados de libertad apropiados para el estadístico ji cuadrada varían dependiendo de la aplicación particular que se utilice. Aun cuando especificaremos los grados de libertad apropiados para las aplicaciones presentadas en este capítulo, se debe usar la regla general dada a continuación para determinar grados de libertad para el estadístico ji cuadrada.

TIP MI CONSEJO

Las pruebas ji cuadrada de Pearson siempre son pruebas de cola superior

MI APLET EN LÍNEA

Probabilidades ji cuadrada



NECESITO SABER...

Cómo determinar el número apropiado de grados de libertad

1. Empiece con el número de *categorías* o celdas en el experimento.
2. Reste un grado de libertad por cada restricción lineal en las probabilidades de celda. Siempre se pierde un grado de libertad porque $p_1 + p_2 + \dots + p_k = 1$.

3. A veces las cantidades de celda esperadas no se pueden calcular directamente sino que deben estimarse usando los datos muestrales. Reste un grado de libertad por cada parámetro poblacional independiente que deba ser estimado, para obtener los valores estimados de E_i .

Empezamos con las aplicaciones más sencillas del estadístico de prueba ji cuadrada, la prueba de **bondad de ajuste**.

PRUEBA DE PROBABILIDADES DE CELDA ESPECIFICADA: LA PRUEBA DE BONDAD DE AJUSTE

10.4

La hipótesis más sencilla respecto a las probabilidades de celda específica un valor numérico para cada celda. Las cantidades de celda esperadas se calculan fácilmente usando las probabilidades hipotéticas, $E_i = np_i$ y se usan para calcular el valor observado del estadístico de prueba χ^2 . Para un experimento multinomial formado por k categorías o celdas, el estadístico de prueba tiene una distribución χ^2 aproximada con $gl = (k - 1)$.

EJEMPLO

10.1

Un investigador diseña un experimento en el que una rata es atraída al final de una rampa que se divide, llevando a puertas de tres colores diferentes. El investigador hace que la rata baje por la rampa $n = 90$ veces y observa las elecciones citadas en la tabla 10.1. ¿La rata tiene (o ha adquirido) preferencia por una de las tres puertas?

TABLA 10.1

Elecciones de puerta de la rata

	Puerta		
	Verde	Roja	Azul
Cantidad observada (O_i)	20	39	31

Solución Si la rata no tiene preferencia en la elección de una puerta, a la larga debe esperarse que seleccione cada puerta un número igual de veces. Esto es, la hipótesis nula es

$$H_0 : p_1 = p_2 = p_3 = \frac{1}{3}$$

contra la hipótesis alternativa

$$H_a : \text{Al menos una } p_i \text{ es diferente de } \frac{1}{3}$$

donde p_i es la probabilidad de que la rata elija la puerta i , para $i = 1, 2$ y 3 . Las cantidades de celda esperadas son las mismas para cada una de las tres categorías, es decir, $np_i = 90(1/3) = 30$. El estadístico de prueba ji cuadrada puede calcularse ahora como

$$\begin{aligned}\chi^2 &= \sum \frac{(O_i - E_i)^2}{E_i} \\ &= \frac{(20 - 30)^2}{30} + \frac{(39 - 30)^2}{30} + \frac{(31 - 30)^2}{30} = 6.067\end{aligned}$$

Para este ejemplo, el estadístico de prueba tiene $(k - 1) = 2$ grados de libertad porque la única restricción lineal en las probabilidades de celda es que deben sumar 1. En consecuencia, se puede usar la tabla 5 del Anexo para hallar límites para el valor p de cola derecha. Como el

TIP MI CONSEJO

La región de rechazo y valor p están en la cola superior de la distribución ji cuadrada

valor observado, $\chi^2 = 6.607$, se encuentra entre $\chi^2_{.050} = 5.99$ y $\chi^2_{.025} = 7.38$, el valor p está entre .025 y .050. El investigador debe informar los resultados como significativos al nivel de 5% ($P < .05$), lo cual significa que la hipótesis nula de no preferencia es rechazada. Hay suficiente evidencia para indicar que la rata tiene una preferencia por una de las tres puertas.

¿Qué más se puede decir acerca del experimento, una vez que se haya determinado estadísticamente que la rata tiene una preferencia? Veamos los datos para observar dónde están las diferencias.

La puerta azul se eligió sólo un poco más de un tercio de las veces:

$$\frac{31}{90} = .344$$

No obstante, las proporciones muestrales para las otras dos puertas son muy diferentes a un tercio. La rata elige la puerta verde con menos frecuencia, sólo 22% de las veces:

$$\frac{20}{90} = .222$$

La rata elige la puerta roja con más frecuencia, 43% de las veces:

$$\frac{39}{90} = .433$$

Se pueden resumir los resultados del experimento si se dice que la rata tiene preferencia por la puerta roja. ¿Se puede concluir que la preferencia es *causada* por el color de la puerta? La respuesta es negativa; la causa podría ser algún otro factor fisiológico o psicológico que todavía no se haya explorado. ¡Evite declarar una relación *causal* entre color y preferencia!

EJEMPLO

10.2

Las proporciones de fenotipos de sangre A, B, AB y O en la población de todos los caucásicos en Estados Unidos son .41, .10, .04 y .45, respectivamente. Para determinar si las proporciones poblacionales reales se ajustan o no a este conjunto de probabilidades reportadas, se seleccionó una muestra aleatoria de 200 estadounidenses y se registraron sus fenotipos sanguíneos. Las cantidades de celda observadas y esperadas se muestran en la tabla 10.2. Las cantidades de celda esperadas se calculan como $E_i = 200p_i$. Pruebe la bondad de ajuste de estas proporciones de fenotipo sanguíneas.

TABLA 10.2

Cantidades de fenotipos sanguíneos

	A	B	AB	O
Observadas (O_i)	89	18	12	81
Esperadas (E_i)	82	20	8	90

Solución La hipótesis a probar está determinada por las probabilidades del modelo:

$$H_0: p_1 = .41; p_2 = .10; p_3 = .04; p_4 = .45$$

contra

$$H_a: \text{Al menos una de las cuatro probabilidades es diferente del valor especificado}$$

Entonces

$$\begin{aligned}\chi^2 &= \sum \frac{(O_i - E_i)^2}{E_i} \\ &= \frac{(89 - 82)^2}{82} + \dots + \frac{(81 - 90)^2}{90} = 3.70\end{aligned}$$

TIP MI CONSEJO

Grados de libertad para una prueba simple de bondad de ajuste:
 $gl = k - 1$

De la tabla 5 del Anexo, indizando $gl = (k - 1) = 3$, se puede hallar que el valor observado del estadístico de prueba es menor que $\chi^2_{.100} = 6.25$, de modo que el valor p es mayor que .10. No se tiene suficiente evidencia para rechazar H_0 ; esto es, no se puede declarar que los fenotipos sanguíneos de caucásicos estadounidenses sean *diferentes* de los reportados antes. Los resultados son no significativos (NS).

En la sección “Tecnología actual”, al final de este capítulo, se pueden encontrar instrucciones que permiten usar *MINITAB* (versiones 15 o 16) para realizar la prueba ji cuadrada de bondad de ajuste y generar los resultados.

Observe la diferencia en la hipótesis de bondad de ajuste comparada con otras hipótesis que se hayan probado. En la prueba de bondad de ajuste, el investigador usa la hipótesis nula para especificar el modelo que crea ser *verdadero*, más que un modelo que espera demostrar que sea falso. Al no poder rechazar H_0 en el ejemplo del tipo de sangre, los resultados fueron como se esperaba. Es necesario tener cuidado, no obstante, cuando se reporten resultados para pruebas de bondad de ajuste. No se puede declarar con confianza que el modelo sea absolutamente correcto sin reportar el valor de β para algunas alternativas prácticas.

10.4 EJERCICIOS

TÉCNICAS BÁSICAS

10.1 Haga una lista de las características de un experimento multinomial.

10.2 Use la tabla 5 del Anexo para hallar el valor de χ^2 con la siguiente área α a su derecha:

- a. $\alpha = .05$, $gl = 3$ b. $\alpha = .01$, $gl = 8$

10.3 Dé la región de rechazo para una prueba ji cuadrada de probabilidades especificadas si el experimento contiene k categorías en estos casos:

- a. $k = 7$, $\alpha = .05$ b. $k = 10$, $\alpha = .01$

10.4 Use la tabla 5 del Anexo para limitar el valor p para una prueba ji cuadrada:

- a. $\chi^2 = 4.29$, $\alpha = .5$ b. $\chi^2 = 20.62$, $gl = 6$

10.5 Suponga que una respuesta puede caer en una de $k = 5$ categorías con probabilidades $p_1, p_2, \dots, p_5$ y que $n = 300$ respuestas produjeron estas cantidades de categoría:

Categoría	1	2	3	4	5
Cantidad observada	47	63	74	51	65

- a. ¿Es igualmente probable que ocurran las cinco categorías? ¿Cómo se probaría esta hipótesis?
- b. Si se fuera a probar esta hipótesis usando el estadístico ji cuadrada, ¿cuántos grados de libertad tendría la prueba?
- c. Encuentre el valor crítico de χ^2 que defina la región de rechazo con $\alpha = .05$.

d. Calcule el valor observado del estadístico de prueba.

e. Efectúe la prueba y exprese sus conclusiones.

10.6 Suponga que una respuesta puede caer en una de $k = 3$ categorías con probabilidades $p_1 = .4$, $p_2 = .3$ y $p_3 = .3$ y $n = 300$ respuestas produjo estas cantidades de categoría:

Categoría	1	2	3
Cantidad observada	130	98	72

¿Los datos dan suficiente evidencia para indicar que las probabilidades de celda son diferentes respecto de las especificadas para las tres categorías? Encuentre el valor p aproximado y úselo para tomar su decisión.

APLICACIONES

10.7 Su carril preferido Se estudió una autopista de cuatro carriles en cada dirección para ver si los automovilistas prefieren ir en los carriles interiores. Se observaron un total de mil automóviles, durante un intenso tránsito matutino y se registró el número de autos en cada carril:

Carril	1	2	3	4
Cantidad observada	294	276	238	192

¿Los datos presentan suficiente evidencia para indicar que algunos carriles son preferidos sobre otros? Pruebe usando $\alpha = .05$. Si hay algunas diferencias, analice la naturaleza de los diferencias.

10.8 Peonías Una peonia con pétalos rojos fue cruzada con otra planta con pétalos de franjas. Un genetista dice que 75% de los descendientes de esta cruce tendrán flores rojas. Para probar esta afirmación, cien semillas de esta cruce se recolectaron y germinaron, y 58 plantas tenían pétalos rojos. Use la prueba de bondad de ajuste ji cuadrada para determinar si los datos muestrales confirman la predicción del genetista.

10.9 Infartos en lunes Investigadores de Alemania han concluido que el riesgo de un infarto para un trabajador puede ser hasta 50% mayor los lunes que en cualquier otro día.¹ Los investigadores dieron seguimiento a los infartos y paros coronarios en un periodo de 5 años entre 330 000 personas que vivían cerca de Augsburg, Alemania. En un intento por verificar esta afirmación, se hizo un estudio de 200 trabajadores que habían tenido infartos recientemente y se registró el día en que éstos ocurrieron:

Día	Cantidad observada
Domingo	24
Lunes	36
Martes	27
Miércoles	26
Jueves	32
Viernes	26
Sábado	29

¿Los datos presentan suficiente evidencia para indicar que hay diferencia en la incidencia de infartos, dependiendo del día de la semana? Pruebe usando $\alpha = .05$.

10.10 Estadísticas de mortalidad Unas estadísticas médicas muestran que los fallecimientos debidos a cuatro enfermedades principales, llamadas A, B, C y D, constituyen 15%, 21%, 18% y 14%, respectivamente, de todos los fallecimientos no accidentales. Un estudio de las causas de 308 fallecimientos no accidentales en un hospital dieron las siguientes cantidades:

Enfermedad	A	B	C	D	Otra
Fallecimientos	43	76	85	21	83

¿Estos datos dan suficiente evidencia para indicar que las proporciones de personas que fallecen de enfermedades A, B, C y D en este hospital difieren de las proporciones acumuladas para la población en general?

10.11 Esquizofrenia Investigaciones realizadas han sugerido que hay un vínculo entre la prevalencia de esquizofrenia y el nacimiento durante meses particulares del año en que prevalecen infecciones virales. Suponga que trabajamos en un problema similar y sospechamos que hay un vínculo entre una enfermedad observada en la etapa final de la vida y el mes de nacimiento. Tenemos registros de 400 casos de la enfermedad y los clasificamos de acuerdo con el mes de nacimiento. Los datos aparecen en la tabla. ¿Los datos presentan suficiente evidencia para indicar que la proporción de casos de la enfermedad por mes varía de un mes a otro? Pruebe con $\alpha = .05$.

Mes	Enero	Febrero	Marzo	Abril	Mayo	Junio
Nacimientos	38	31	42	46	28	31

Mes	Julio	Agosto	Sept.	Oct.	Nov.	Dic.
Nacimientos	24	29	33	36	27	35

10.12 Chicharos Suponga que usted está interesado en seguir dos características independientes de chicharos: textura de la semilla (S = lisa, s = arrugada) y color de la semilla (Y = amarillo, y = verde) en una cruce de segunda generación de padres heterocigotos. La teoría de Mendel dice que el número de chicharos clasificados como lisos y amarillos, arrugados y amarillos, lisos y verdes, y arrugados y verdes debe estar en la relación 9:3:3:1. Suponga que cien chicharos seleccionados al azar tienen 56, 19, 17 y 8 en estas categorías respectivas. ¿Estos datos indican que el modelo 9:3:3:1 es correcto? Pruebe usando $\alpha = .01$.

10.13 Estándares de inscripción Los registros previos de inscripción en una gran universidad indican que del número total de personas que solicitaron inscripción, 60% son inscritos incondicionalmente, 5% son inscritos a condición de pasar una prueba y el resto son rechazados. De 500 solicitudes a la fecha para el año entrante, 329 solicitantes han sido inscritos incondicionalmente, 43 han sido inscritos a condición de pasar una prueba y el resto han sido rechazados. ¿Estos datos indican una desviación respecto a los porcentajes de inscripción previos? Pruebe usando $\alpha = .05$.

10.5

TABLAS DE CONTINGENCIA: UNA CLASIFICACIÓN DE DOS VÍAS

En algunas situaciones, el investigador clasifica una unidad experimental de acuerdo con *dos variables cualitativas* para generar *datos bivariados*.

- Una pieza defectuosa de mueble se clasifica según el tipo de defecto y el turno de producción durante el que se hizo.
- Una profesora es clasificada por su rango profesional y el tipo de universidad (pública o privada) en la que trabaja.
- Un paciente es clasificado de acuerdo con el tipo de tratamiento preventivo contra la gripe que ha recibido y si ha contraído o no la gripe durante el invierno.

Cuando se registran dos *variables categóricas*, se pueden resumir los datos al contar el número observado de unidades que caen en cada una de las diversas intersecciones de niveles de categoría. Las cantidades resultantes se exhiben en un conjunto ordenado llamado **tabla de contingencia**.

EJEMPLO

10.3

Un total de $n = 309$ defectos en muebles fueron registrados y clasificados en cuatro tipos: A, B, C o D. Al mismo tiempo, cada pieza de mueble fue identificada por el turno de producción en el que se manufacturó. Estas cantidades se presentan en una tabla de contingencia en la tabla 10.3.

TABLA 10.3

Tabla de contingencia

Tipo de defectos	Turno			Total
	1	2	3	
A	15	26	33	74
B	21	31	17	69
C	45	34	49	128
D	13	5	20	38
Total	94	96	119	309

TIP MI CONSEJO

Con clasificaciones de dos vías, no probamos hipótesis acerca de probabilidades específicas. Probamos si los dos métodos de clasificación son independientes

Cuando se estudian datos que contienen dos variables, una consideración importante es la *relación entre las dos variables*. ¿La proporción de mediciones en las diversas categorías para el factor 1 depende de cuál categoría del factor 2 se observe? Para el ejemplo de muebles, ¿las proporciones de los diversos defectos varían de turno a turno, o son iguales estas proporciones, independientemente de cuál turno se observe? Recordemos un fenómeno similar llamado *interacción* en el experimento factorial $a \times b$ del capítulo 7. En el análisis de una tabla de contingencia, el objetivo es determinar si un método de clasificación es o no es **contingente** o **dependiente** del otro método de clasificación. Si no lo es, se dice que los dos métodos de clasificación son **independientes**.

La prueba de independencia ji cuadrada

La cuestión de independencia de los dos métodos de clasificación se puede investigar usando una prueba de hipótesis basada en el estadístico ji cuadrada. Éstas son las hipótesis:

- H_0 : Los dos métodos de clasificación son independientes
 H_a : Los dos métodos de clasificación son dependientes

Suponga que denotamos la cantidad de celdas observada en la fila i y la columna j de la tabla de contingencia como O_{ij} . Si se conocieran las cantidades de celda esperadas ($E_{ij} = np_{ij}$) bajo la hipótesis nula de independencia, entonces se podría usar el estadístico ji cuadrada para comparar las cantidades observadas y esperadas. No obstante, los valores esperados no están especificados en H_0 , como estaban en los ejemplos previos.

Para explicar cómo estimar estas cantidades de celda esperadas, debemos repasar el concepto de *eventos independientes* del capítulo 2. Considere p_{ij} , la probabilidad de que una observación caiga en la fila i y la columna j de la tabla de contingencia. Si las filas y columnas son independientes, entonces

$$\begin{aligned} p_{ij} &= P(\text{observación cae en la fila } i \text{ y columna } j) \\ &= P(\text{observación cae en la fila } i) \times P(\text{observación cae en la columna } j) \\ &= p_i p_j \end{aligned}$$

donde p_i y p_j son las **probabilidades incondicional** o **marginal** de caer en la fila i o columna j , respectivamente. Si se pudieran obtener estimaciones apropiadas de estas probabilidades marginales podrían usarse en lugar de p_{ij} en la fórmula para la cantidad de celda esperada.

Por fortuna, existen estas estimaciones. De hecho, son exactamente lo que en forma intuitiva se elegiría:

- Para estimar la probabilidad de una fila, use

$$\hat{p}_i = \frac{\text{Total de observaciones en fila } i}{\text{número total de observaciones}} = \frac{r_i}{n}$$

- Para estimar la probabilidad de una columna, use

$$\hat{p}_j = \frac{\text{Total de observaciones en la columna } j}{\text{número total de observaciones}} = \frac{c_j}{n}$$

La estimación de la cantidad de celda esperada para la fila i y la columna j se deduce de la suposición de independencia.

TIP MI CONSEJO

Grados de libertad para una tabla de contingencia $r \times c$: $gl = (r - 1)(c - 1)$

CANTIDAD DE CELDA ESPERADA ESTIMADA

$$\hat{E}_{ij} = n \left(\frac{r_i}{n} \right) \left(\frac{c_j}{n} \right) = \frac{r_i c_j}{n}$$

donde r_i es el total para la fila i y c_j es el total para la columna j .

El estadístico de prueba ji cuadrada para una tabla de contingencia con r filas y c columnas se calcula como

$$\chi^2 = \sum \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

y puede mostrarse que tiene una distribución ji cuadrada aproximada con

$$gl = (r - 1)(c - 1)$$

Si el valor observado de χ^2 es demasiado grande, entonces la hipótesis nula de independencia se rechaza.

EJEMPLO

10.4

Consulte el ejemplo 10.3. ¿Los datos presentan suficiente evidencia para indicar que el tipo de defecto de muebles varía con el turno durante el cual la pieza se produjo?

Solución Las cantidades de celda esperadas estimadas se muestran entre paréntesis en la tabla 10.4. Por ejemplo, la cantidad esperada estimada para un defecto tipo C producido durante el segundo turno es

$$\hat{E}_{32} = \frac{r_3 c_2}{n} = \frac{(128)(96)}{309} = 39.77$$

TABLA 10.4

Cantidades de celda esperadas observadas y estimadas

Tipo de defectos	Turno			Total
	1	2	3	
A	15 (22.51)	26 (22.99)	33 (28.50)	74
B	21 (20.99)	31 (21.44)	17 (26.57)	69
C	45 (38.94)	34 (39.77)	49 (49.29)	128
D	13 (11.56)	5 (11.81)	20 (14.63)	38
Total	94	96	119	309

Ahora se pueden usar los valores mostrados en la tabla 10.4 para calcular el estadístico de prueba como

$$\begin{aligned}\chi^2 &= \sum \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} \\ &= \frac{(15 - 22.51)^2}{22.51} + \frac{(26 - 22.99)^2}{22.99} + \dots + \frac{(20 - 14.63)^2}{14.63} \\ &= 19.18\end{aligned}$$

Cuando se indiza la distribución ji cuadrada de la tabla 5 del Anexo con

$$gl = (r - 1)(c - 1) = (4 - 1)(3 - 1) = 6$$

el estadístico de prueba observado es mayor que $\chi^2_{.005} = 18.5476$, lo cual indica que el valor p es menor que .005. Se puede rechazar H_0 y declarar que los resultados son altamente significativos ($P < .005$). Hay suficiente evidencia para indicar que las proporciones de tipos de defecto varían de un turno a otro.



APPLET EN LÍNEA

Prueba de independencia
ji cuadrada

La siguiente pregunta obvia que se debe formular comprende la naturaleza de la relación entre las dos clasificaciones. ¿Cuál turno produce más de qué tipo de defecto? Al igual que con el experimento factorial del capítulo 7, una vez hallada una dependencia (o interacción) se debe buscar dentro de la tabla en las proporciones relativa o *condicional* para cada nivel de clasificación. Por ejemplo, considere el turno 1, que produjo un total de 94 defectos. Estos defectos se pueden dividir en tipos usando las *proporciones condicionales* para esta muestra que aparecen en la primera columna de la tabla 10.5. Si se sigue el mismo procedimiento para los otros dos turnos, se pueden entonces comparar las distribuciones de tipos de defecto para los tres turnos, como se ve en la tabla 10.5.

Ahora comparemos los tres conjuntos de proporciones (cada uno totaliza 1). Se ve que los turnos 1 y 2 producen defectos en el mismo orden general, tipos C, B, A y D de mayor a menor, aunque en proporciones que difieren. El turno 3 presenta un patrón diferente, casi todos los defectos de tipo C pero seguidos por los tipos A, D y B, en ese orden. Dependiendo de cuál tipo de defecto sea el más importante para el fabricante, cada turno debe ser advertido separadamente acerca de las razones para producir demasiados defectos.

TABLA 10.5 Probabilidades condicionales para tipos de defecto dentro de tres turnos

Tipo de defectos	Turno		
	1	2	3
A	$\frac{15}{94} = .16$	$\frac{26}{96} = .27$	$\frac{33}{119} = .28$
B	$\frac{21}{94} = .22$	$\frac{31}{96} = .32$	$\frac{17}{119} = .14$
C	$\frac{45}{94} = .48$	$\frac{34}{96} = .35$	$\frac{49}{119} = .41$
D	$\frac{13}{94} = .14$	$\frac{5}{96} = .05$	$\frac{20}{119} = .17$
Total	1.00	1.00	1.00



NECESITO SABER...

Cómo determinar el número apropiado de grados de libertad

Recuerde el procedimiento general para determinar grados de libertad:

1. Empiece con $k = rc$ categorías o celdas en la tabla de contingencia.
2. Reste un grado de libertad porque todas las probabilidades de celda deben sumar 1.
3. Tuvo que estimar $(r - 1)$ probabilidades de renglón y $(c - 1)$ probabilidades de columna para calcular las cantidades de celda esperadas estimadas. (La última de las probabilidades de fila y de columna está determinada porque las probabilidades *marginales* de fila y columna también totalizan 1.) Reste $(r - 1)$ y $(c - 1)$ grados de libertad (gl).

El total de grados de libertad para la tabla de contingencia $r \times c$ es

$$gl = rc - 1 - (r - 1) - (c - 1) = rc - r - c + 1 = (r - 1)(c - 1)$$

EJEMPLO

10.5

Se realizó un estudio para evaluar la eficacia de una nueva vacuna contra la gripe que había sido administrada en una pequeña comunidad. La vacuna se aplicó sin cargo en una secuencia de dos inyecciones en un periodo de 2 semanas. Algunas personas recibieron la secuencia de dos inyecciones, algunas se presentaron sólo para la primera inyección y otras no recibieron ninguna. Un estudio de mil residentes locales la primavera siguiente dio la información que se ve en la tabla 10.6. ¿Los datos presentan suficiente evidencia para indicar que la vacuna tuvo éxito para reducir el número de casos de gripe en la comunidad?

TABLA 10.6 Tabla de contingencia de 2×3

	Sin vacuna	Una inyección	Dos inyecciones	Total
Gripe	24	9	13	46
Sin gripe	289	100	565	954
Total	313	109	578	1000

Solución El éxito de la vacuna para reducir el número de casos de gripe se puede evaluar en dos partes:

- Si la vacuna tiene éxito, las proporciones de personas que contraen gripe deberían variar, dependiendo de cuál de los tres tratamientos recibieron.
- Esta dependencia no sólo debe existir, sino que la proporción de personas que contraen gripe debería disminuir a medida que aumenta el tratamiento de prevención de la gripe, de cero a una o dos inyecciones.

La primera parte se puede probar usando la prueba ji cuadrada con estas hipótesis:

H_0 : No hay relación entre tratamiento e incidencia de gripe

H_a : La incidencia de gripe depende de la cantidad de tratamiento contra la gripe

TIP MI CONSEJO

Use el valor de χ^2 y el valor p de la salida impresa para probar la hipótesis de independencia

Como de costumbre, paquetes de software pueden eliminar todos los tediosos cálculos y, si los datos se introducen correctamente, dan la salida correcta que contiene el valor observado del estadístico de prueba y su valor p . Esa salida impresa, generada por *MINITAB*, se ilustra en la figura 10.3. Se pueden hallar instrucciones para generar esta salida impresa en la sección “Tecnología actual” al final de este capítulo. El valor observado del estadístico de prueba, $\chi^2 = 17.313$, tiene un valor p de .000 y los resultados son declarados altamente significativos. Esto es, la hipótesis nula es rechazada. Hay suficiente evidencia para indicar una relación entre tratamiento e incidencia de la gripe.

FIGURA 10.3

Salida *MINITAB* para el ejemplo 14.5

Prueba ji cuadrada: sin vacuna, una inyección, dos inyecciones

Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

	No Vaccine	One Shot	Two Shots	Total
1	24 14.40 6.404	9 5.01 3.169	13 26.59 6.944	46
2	289 298.60 0.309	100 103.99 0.153	565 551.41 0.335	954
Total	313	109	578	1000
Chi-Sq = 17.313, DF = 2, P-Value = 0.000				

¿Cuál es la naturaleza de esta relación? Para contestar esta pregunta, véase la tabla 10.7, que da la *incidencia* de gripe en la muestra para cada uno de los tres grupos de tratamiento. La respuesta es obvia. El grupo que recibió dos inyecciones fue menos susceptible a la gripe; una sola inyección no parece reducir la susceptibilidad.

TABLA 10.7

Incidencia de gripe para tres tratamientos

Sin vacuna	Una inyección	Dos inyecciones
$\frac{24}{313} = .08$	$\frac{9}{109} = .08$	$\frac{13}{578} = .02$

10.5 EJERCICIOS

TÉCNICAS BÁSICAS

10.14 Calcule el valor y dé el número de grados de libertad de χ^2 para estas tablas de contingencia:

a.

	Columnas			
Filas	1	2	3	4
1	120	70	55	16
2	79	108	95	43
3	31	49	81	140

b.

	Columnas		
Filas	1	2	3
1	35	16	84
2	120	92	206

10.15 Suponga que un estudio de consumidores compendia las respuestas de $n = 307$ personas en una tabla de contingencia de tres filas y cinco columnas. ¿Cuántos grados de libertad están asociados con el estadístico de prueba ji cuadrada?

10.16 Un estudio de 400 entrevistados produjo estas cantidades de celda en una tabla de contingencia de 2×3 :

	Columnas			
Filas	1	2	3	Total
1	37	34	93	164
2	66	57	113	236
Total	103	91	206	400

- Si se desea probar la hipótesis nula de “independencia”, es decir, que la probabilidad de que una respuesta caiga en cualquier fila, es independiente de la columna en la que caiga y se piensa usar una prueba ji cuadrada, ¿cuántos grados de libertad estarán asociados con el estadístico χ^2 ?
- Encuentre el valor del estadístico de prueba.
- Encuentre la región de rechazo para $\alpha = .01$.
- Efectúe la prueba y exprese sus conclusiones.
- Encuentre el valor p aproximado para la prueba e interprete su valor.

10.17 Diferencias de género Los hombres y mujeres entrevistados en un cuestionario de diferencias de género se clasificaron en tres grupos, según sus respuestas a la primera pregunta:

	Grupo 1	Grupo 2	Grupo 3
Hombres	37	49	72
Mujeres	7	50	31

Use la salida impresa *MINITAB* para determinar si hay una diferencia en las respuestas de acuerdo con el género. Explique la naturaleza de las diferencias, si la hay.

Salida impresa *MINITAB* para el ejercicio 10.17

Prueba ji cuadrada: grupo 1, grupo 2, grupo 3

Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

	Group 1	Group 2	Group 3	Total
1	37 28.26 2.703	49 63.59 3.346	72 66.15 0.517	158
2	7 15.74 4.853	50 35.41 6.007	31 36.85 0.927	88
Total	44	99	103	246

Chi-Sq = 18.352, DF = 2, P-Value = 0.000

APLICACIONES

**10.18 Reforma de la atención a la salud** El

Congreso aprobó la Ley de Reforma de la atención a la salud en 2010, la cual requiere que todos los negocios pequeños tengan un plan de salud establecido para 2014 e incluye la disposición de que a nadie se le podría negar la cobertura debido a condiciones preexistentes. Una encuesta de *Newsweek* basada en $n = 848$ votantes registrados a nivel nacional pedía a quienes respondieron “Por favor diga si piensa que la ley de reforma de la atención a la salud aprobada antes este año es buena o mala para el país en general”.² Los datos que siguen se basan en los resultados de esta encuesta.

Afiliación	Buena	Mala	No está seguro	Total
Republicanos	32	301	17	350
Demócratas	277	60	37	374
Independientes	51	58	19	128

- ¿Hay diferencias significativas en las proporciones de los encuestados que piensan que la ley de reforma de la atención a la salud es buena, mala o no está seguro entre los republicanos, demócratas e independientes? Use $\alpha = .05$.
- Si existen diferencias significativas, describa la naturaleza de las diferencias hallando las proporciones de aquellos que piensan que la ley es buena, mala o no están seguros para cada una de las afiliaciones dadas.

10.9 Infantes ansiosos Joseph Jacobson y Diane Wille realizaron un estudio para determinar el efecto de la atención temprana infantil en los patrones de apego entre madre e hijo.³ En el estudio, 93 infantes fueron clasificados como “seguros” o “ansiosos” usando el paradigma de situación extraña de Ainsworth. Además, los infantes se clasificaron de acuerdo con el número promedio de horas por semana que pasaron recibiendo atención y cuidado. Los datos se presentan en la tabla.

	Baja (0–3 horas)	Moderada (4–19 horas)	Alta (20–54 horas)
Seguro	24	35	5
Ansioso	11	10	8

- a. ¿Los datos dan suficiente evidencia para indicar que hay una diferencia en el patrón de apego para los infantes, dependiendo del tiempo que pasen en atención y cuidado? Pruebe usando $\alpha = .05$.
- b. ¿Cuál es el valor p aproximado para la prueba de la parte a?

DATOS
EX1020

10.20 Color de cabello El color del cabello y de los ojos que sigue se obtuvo por autoinforme de una muestra de estadounidenses caucásicos nacidos entre 1957 y 1965 (en la actualidad de 45–53 años de edad).⁴ Los siguientes datos se adaptaron de ese estudio.

	Rubio claro	Rubio	Castaño claro	Castaño	Negro	Rojo	Total
Hombres	4	46	45	176	23	10	304
Mujeres	5	58	69	164	12	12	320

- a. ¿Hay suficiente evidencia para concluir que la proporción de individuos con estos colores de cabello difieren para hombres y mujeres? Use $\alpha = .05$.
- b. ¿Hay alguna celda con un número esperado menor a cinco? De ser así, combine esas celdas con aquellas que están junto a ellas y vuelva a analizar los datos. ¿Difieren los resultados finales?

10.6

PROCEDIMIENTO POST HOC

Las pruebas de independencia sólo indican la presencia de asociación, pero no su magnitud (valor numérico). La sensibilidad de χ^2 a cualquier grado de asociación se incrementa en función del tamaño de la muestra. Sin embargo, se obtiene un valor significativo de χ^2 con una muestra grande, aunque la relación entre las variables es tan débil que su significado práctico resulta limitado. Cuando se ha rechazado la H_0 de independencia, hay que calcular un índice de asociación para obtener una estimación numérica de la fuerza de la relación. Estos índices incluyen el coeficiente Φ , el de Kramer y el de contingencia.

EJEMPLO

10.6

Un investigador busca establecer si existen diferencias entre los hombres y las mujeres que presentan examen de admisión a la universidad. Admita que las puntuaciones de dicho examen se clasifican como aceptados o rechazados. Los datos se obtuvieron de una muestra aleatoria de 200 aspirantes de la población total que desea ingresar a esa institución; estos datos pueden exponerse en la siguiente tabla de contingencia (2×2).

	Aceptados	Rechazados	Total
Hombres	93 (a)	36 (b)	132
Mujeres	41 (c)	27 (d)	68
Total	134	63	200

Esto indica que la muestra ($n = 200$) la constituían 132 hombres y 68 mujeres; en ella, 93 hombres y 41 mujeres obtuvieron puntuaciones de aceptación. A continuación, se calculan las frecuencias esperadas respecto de si el sexo del aspirante es independiente de la puntuación de aceptación. El procedimiento es el siguiente:

Paso 1. Se calcula el total de puntuaciones aceptadas: $93 + 41 = 134$.

Paso 2. Se encuentra la proporción de 134 (total de aceptados) respecto de n . Así, $\frac{134}{200} = 0.67$.

Paso 3. Se multiplica esta proporción por el número total de hombres en la muestra: $0.67 \times 132 = 88.44$; ésta es la frecuencia esperada para la casilla (a).

Paso 4. Se resta 88.44 de 132 ($132 - 88.44 = 43.56$), y así se encuentra la frecuencia esperada de hombres con puntuaciones de aceptación.

Tome en cuenta que la suma de las frecuencias esperadas para un renglón o columna será igual a la suma de las frecuencias observadas para ese mismo renglón.

Paso 5. Posteriormente, se calculan las frecuencias esperadas para las mujeres con puntuaciones de aceptación: se resta 88.44 de 134, lo que resulta 45.56, y sucesivamente se calculan todas las frecuencias esperadas, donde se aplica la siguiente fórmula en cada casilla:

$$fe = \frac{RC}{n}$$

donde:

fe = Frecuencia esperada de una casilla.

R = Total marginal del renglón de dicha casilla.

C = Total marginal de la columna de dicha casilla.

n = Tamaño de la muestra.

Paso 6. Se calcula χ^2 por medio del siguiente procedimiento para cada casilla:

$$\frac{(|fo - fe| - 0.5)^2}{fe}$$

Las líneas verticales $|fo - fe|$, indican que debe tomarse el valor absoluto de esa diferencia.

Los resultados de las casillas se suman, y el total será el estadístico χ^2 . La corrección de continuidad de Yates se utiliza debido a que el grado de libertad para este diseño es 1 y la distribución real es discontinua.

$$\chi_c^2 = \frac{(4.06)^2}{88.44} + \frac{(4.06)^2}{43.56} + \frac{(4.06)^2}{45.56} + \frac{(4.06)^2}{22.44}$$

$$\chi_c^2 = 0.19 + 0.38 + 0.36 + 0.73 = 1.66$$

$\chi_c^2 = 1.66$

Paso 7. Se aplica la prueba de confianza. El valor del estadístico ji cuadrada teórica (χ^2), al nivel del 5%, con un grado de libertad es de 3.84. como 1.66 es menor que 3.84, se rechaza la hipótesis alterna (H_1), de acuerdo con la regla de decisión del estadístico de prueba ji cuadrada.

Solución El sexo de los aspirantes y las puntuaciones obtenidas en el examen de admisión son independientes entre sí. En otras palabras: no existe diferencia significativa entre los hombres y las mujeres respecto del examen de admisión.

Requisitos para usar el estadístico de prueba ii cuadrada (χ^2)

1. *Aleatoriedad*. Los elementos que constituyen la muestra deben haber sido extraídos de manera aleatoria de la población de estudio.
2. *Independencia de las submuestras*. Las observaciones dentro de las muestras y entre las mismas (submuestras) deben ser mutuamente independientes.
3. *Distribución regular*. Dentro de cada categoría, las observaciones repetidas a través de todos los posibles valores para esa categoría. Así, el valor esperado es un punto de medida poblacional para los datos de frecuencia.
4. *Aleatoriedad*. Sólo se requieren frecuencias de casos o respuestas.
5. *Independencia de las muestras*. Esto requiere la presencia de un cuadro 2×2 como mínimo. La suposición de independencia indica que el estadístico χ^2 no puede aplicarse a una sola muestra (de panel antes-después); deben obtenerse por lo menos dos muestras de entrevistados.
6. Algunos estadísticos sugieren que no es adecuado el uso del estadístico χ^2 cuando las frecuencias esperadas sean menores de cinco, aunque este requisito no constituye una restricción.

Coeficiente ϕ (Φ)

Cuando ambas variables son nominales y dicotómicas, es posible determinar el grado de asociación entre las variables de interés. Dicho coeficiente (Φ) también constituye un caso particular del coeficiente de correlación de Pearson, y se utiliza con cierta frecuencia, aunque no necesariamente con este propósito, en la elaboración y análisis de pruebas. En secciones posteriores se considerará la independencia o dependencia de dos variables en una muestra determinada; cuando se realice este análisis, a partir de las hipótesis establecidas, si la conclusión indica que existe dependencia, el análisis estadístico más pertinente es conocer el grado de asociación que implica la dependencia entre las variables o las muestras. Para determinar esto, se necesita un número que se establecerá a partir del coeficiente de correlación Φ (ϕ); sin embargo, debe recordarse que dicho número está supeditado al diseño 2×2 , al tamaño de la muestra y a la proporción de las variables dicotomizadas. Cuando el número de casos en una variable es igual al de la otra, el coeficiente Φ tendrá el valor máximo de 1; cuando los totales marginales son diferentes, no se alcanzará ese valor.

EJEMPLO

10.7

Se desea establecer una escala de medición de cierto rasgo de agresión en adultos. La interrogante al respecto consiste en determinar si existe relación entre el sexo de la persona y la respuesta (dicotómica) ante una situación que provoque agresividad. Así, se realiza un estudio con 400 personas, 200 del sexo masculino y 200 del femenino. Los datos son los siguientes: 160 personas del sexo masculino están de acuerdo y 40 en desacuerdo; por parte del sexo femenino, 40 están de acuerdo y 160 no lo están. Calcule el coeficiente Φ de correlación y concluya con base en los resultados obtenidos.

Paso 1. Se presentan los datos obtenidos en una tabla de doble entrada de dos renglones y dos columnas (2×2).

	Acuerdo	Desacuerdo	Total
Masculino	160 <i>a</i>	40 <i>b</i>	200 <i>a + b</i>
Femenino	40 <i>c</i>	160 <i>d</i>	200 <i>c + d</i>
Total	200 <i>a + c</i>	200 <i>b + d</i>	400

Paso 2. Se aplica la definición de Φ .

$$\Phi = \frac{ad - bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$$

donde:

a = las puntuaciones del primer renglón y la primera columna (en este caso, masculino y de acuerdo = 160).

b = puntuaciones del primer renglón y la segunda columna (masculino y en desacuerdo = 40).

c = puntuaciones del segundo renglón y la primera columna (femenino y de acuerdo = 40).

d = las del segundo renglón y la segunda columna (femenino y en desacuerdo = 160).

Paso 3. Se sustituyen estos datos en la fórmula anterior y se efectúan las operaciones.

$$\begin{aligned}\Phi &= \frac{(160)(160) - (40)(40)}{\sqrt{(200)(200)(200)(200)}} \\ \Phi &= \frac{25\,600 - 1\,600}{\sqrt{1\,600\,000\,000}} \\ \Phi &= \frac{24\,000}{40\,000} = 0.60 \\ \therefore \Phi &= 0.60\end{aligned}$$

Paso 4. Se concluye teóricamente que el grado de asociación entre el sexo y la agresión es de 60%.

Prueba de significancia de *fi* (Φ)

Para poder comprobar la significancia de dicho coeficiente, se utiliza la siguiente definición:

$$\chi^2 = n \Phi^2$$

donde:

n = número total de casos (suma de los totales marginales, tanto de columnas como de renglones); en el ejemplo, *n* = 400.

Φ = coeficiente *fi* = 0.60.

χ^2 = ji cuadrada, que guarda la siguiente relación:

$$\Phi = \sqrt{\frac{\chi^2}{n}}$$

Al sustituir los valores anteriores, se obtiene:

$$\begin{aligned}\chi_c^2 &= (400) (0.60)^2 \\ &= (400) (0.36) = 144 \\ \therefore \chi^2 &= 144\end{aligned}$$

Este resultado se contrasta con el valor crítico de ji cuadrada, calculado mediante la siguiente regla de decisión RD:

Si $\chi^2 \geq \chi_{crit}^2$, entonces Φ es significativa.

El valor crítico de la χ^2 está en función de los grados de libertad y el nivel de significancia (α), establecido de antemano. Para un diseño 2×2 , los grados de libertad serán siempre igual a 1, pero, en general, para cualquier diseño diferente del mencionado se calcularán mediante las siguientes definiciones:

$$gl = (r - 1) (c - 1)$$

donde:

gl = grados de libertad.

r = número de renglones de la tabla de contingencia.

c = número de columnas de la tabla de frecuencias observadas.

En este caso se tiene $r = 2$ y $c = 2$ (dos renglones y dos columnas), o sea: $gl = (2 - 1) (2 - 1) = 1 \times 1 = 1$.

Se compara el valor calculado de χ^2 con el valor crítico (véase la tabla χ^2 del Anexo), así como con $gl = 1$, y se establecen los niveles de significancia $\alpha = 5\%$ y $\alpha = 1\%$, cuyos valores son $\chi^2 (5\%) = 3.84$ y $\chi^2 = 6.63$, respectivamente. La χ^2 calculada es lo bastante grande para ambos valores, lo cual significa que el coeficiente $\Phi = 60\%$ es significativo.

Requisitos de uso de Φ

Con el fin de utilizar adecuadamente el coeficiente Φ como medida de asociación entre las variables x y y dicotomizadas, deben tomarse en cuenta las siguientes condiciones:

1. *Datos nominales.* Las variables x y y deben ser nominales y dicotomizables, ya que únicamente se requerirán las frecuencias observadas (el número de veces que ocurren en cierta nominación).
2. *Tabla de contingencia 2×2 .* Los datos deben colocarse en un diseño de dos renglones por dos columnas. Es inadecuado aplicar el coeficiente Φ a diseños mayores del indicado, donde se comparan varios grupos o categorías.
3. *Muestreo aleatorio.* Para poder comprobar la significancia y validez de Φ , la muestra en estudio debe extraerse en forma aleatoria (todos los elementos de la población deben tener la misma posibilidad de ser escogidos).
4. *Un caso particular.* Cuando la muestra en estudio es pequeña se utilizará la definición de χ^2 , pero con la corrección de Yates, también llamada de *Pirie-Handem*. Un criterio

para definir la pequeñez, es que la frecuencia Φ observada en dos o más casillas sea menor que 10. La definición de χ^2 consiste en lo siguiente:

$$\chi^2 = \frac{n[|ad - bc| - 0.5]^2}{(a+b)(a+c)(b+d)(c+b)}$$

donde:

$$|ad - bc| = \text{valor absoluto de la diferencia entre } ad \text{ y } bc:$$

y

$$\Phi = \sqrt{\frac{\chi_c^2}{n}}$$

Coeficiente de contingencia (C)

Es un coeficiente de correlación para datos nominales colocados en una tabla de contingencia con un diseño mayor que 2×2 . Con éste se trata de determinar el grado de asociación, al comparar varios grupos o categorías. Puede calcularse mediante la siguiente definición:

$$C = \sqrt{\frac{\chi^2}{n + \chi^2}}$$

donde:

C = coeficiente de contingencia.

χ_c^2 = valor calculado de ji cuadrada.

n = número total de casos (tamaño de la muestra).

EJEMPLO

10.8

Se investiga la posible relación entre cierto tipo de defecto genético en recién nacidos y la edad de la madre al ocurrir el parto. Se necesita encontrar el grado de asociación, si existe dependencia entre ambas variables. El defecto se clasifica en cuatro tipos: A , B , C y D ; la edad de la madre en tres intervalos: menor que 25, entre 25 y 40 y mayor que 40 años. Para llevar a cabo este tipo de estudio, se consideran 309 niños ($n = 309$) en los que se han diagnosticado defectos genéticos.

Paso 1. Los datos obtenidos al realizar este estudio se concentran en una tabla de contingencia:

	Padecimiento genético				Total
	A	B	C	D	
Edad materna					
Menor que 25 años	51	46	25	15	137
Entre 25 y 40 años	33	17	49	20	119
Mayor que 40 años	4	11	35	3	53
Total	88	74	109	38	309

Si las dos clasificaciones son independientes entre sí, la posibilidad (proporción relativa) de una casilla determinada será igual al producto de su total marginal del renglón por el total marginal de su columna, dividido entre n . A esto se le llama *frecuencia esperada* (fe).

Para calcular el estadístico χ^2 puede utilizarse la siguiente fórmula:

$$\chi^2 = \sum \frac{(fo - fe)^2}{fe}$$

donde:

fo = frecuencia observada (se mide directamente).

fe = frecuencia esperada (producto total marginal del renglón por el total marginal de la columna, dividido entre n).

Paso 2. Se calculan las frecuencias esperadas:

	1 A	2 B	3 C	4 D	Total
1 Menor que 25 años	51/39	46/32.8	25/48.3	15/16.8	137
2 Entre 25 y 40 años	33/33.9	17/28.5	49/42	20/14.6	119
3 Mayor que 40 años	4/15.1	11/12.7	35/18.7	3/6.5	53
4 Total	88	74	109	38	309

Por ejemplo:

$$fe_{11} = \frac{137 \times 88}{309} = 39$$

$$fe_{23} = \frac{119 \times 109}{309} = 42$$

Sucesivamente, se obtienen las demás frecuencias, hasta $fe_{34} = \frac{53 \times 38}{309} = 6.5$

Paso 3. Para cada casilla se obtiene $fo - fe$, lo cual se eleva después al cuadrado igualmente en cada casilla con el fin de establecer $(fo - fe)^2$. De este modo, las casillas quedan en particular:

$$(fo_{11} - fe_{11})^2 = (51 - 39)^2 = (12)^2 = 144$$

$$(fo_{23} - fe_{23})^2 = (49 - 42)^2 = 49$$

Paso 4. Se divide cada uno de los resultados anteriores entre su respectiva frecuencia esperada (fe) y se suman los mismos.

$$\frac{144}{39} + \frac{174.24}{32.8} + \frac{542.89}{48.3} + \frac{3.24}{16.8} + \frac{0.81}{33.9} + \frac{132.25}{28.5} +$$

$$\frac{49}{42} + \frac{29.16}{14.6} + \frac{123.21}{15.1} + \frac{2.89}{12.7} + \frac{265.69}{18.7} + \frac{12.25}{6.5}$$

A continuación:

$$3.7 + 5.31 + 11.24 + 0.19 + 0.02 + 4.64 + 1.16 + \\ 3.40 + 8.16 + 0.227 + 14.20 + 1.88 = 54.127$$

$$\therefore \chi_c^2 = 54.127$$

Paso 5. Se calcula el coeficiente C mediante la fórmula:

$$C = \sqrt{\frac{\chi_c^2}{n + \chi_c^2}}$$

$$C = \sqrt{\frac{54.127}{309 + 54.127}}$$

$$C = \sqrt{\frac{54.127}{363.127}}$$

$$C = \sqrt{0.149} = 0.386$$

$$\therefore C = 0.386$$

Paso 6. Teóricamente, el coeficiente de contingencia en la correlación entre el defecto genético del recién nacido y la edad de la madre es de 38.6%.

Prueba de significancia

La significancia estadística del coeficiente de contingencia se consigue a partir de la magnitud del estadístico χ^2 obtenido con la siguiente regla de decisión: si $\chi_c^2 \geq \chi_{crit}^2$, entonces C es significativo. El valor de χ^2 crítico (χ_{crit}^2) se presenta en la tabla χ^2 del Anexo, y se consideran los grados de libertad de la siguiente manera:

$$gl = (r - 1) (c - 1)$$

donde:

r = número de renglones.

c = número de columnas.

En el ejemplo:

$$gl = (3 - 1) (4 - 1) = 2 \times 3 = 6; gl = 6$$

Así, el valor de χ_{crit}^2 al nivel de significancia de 5% es 12.6 y al de 1% es 16.81. Pero el valor de χ^2 calculada es de 54.127, es decir, mayor que cualquiera de los dos valores anteriores. En consecuencia, el coeficiente de contingencia calculado es estadísticamente significativo, por lo que se rechaza la hipótesis nula: los tipos de defectos congénitos dependen de la edad de la madre.

Requisitos para el uso del coeficiente de contingencia:

1. *Datos nominales*. Sólo se requieren datos de las variables dicotomizadas, las que se colocan en una tabla de contingencia $n \times m$.
2. *Muestreo aleatorio*. Con la finalidad de comprobar la significancia estadística del coeficiente, la muestra se obtuvo en forma aleatoria.

Coeficiente V de Kramer

El coeficiente C presenta una desventaja: el número de correspondientes a la tabla de contingencia influirá en el valor obtenido de C . Esto significa que este coeficiente no siempre variará entre 0 y 1 (por supuesto, nunca excederá de 1). En ciertas condiciones, el máximo valor de C podrá ser de 0.94, pero, en general, es de 0.89.

Para resolver este problema se emplea otro coeficiente de correlación, el V de Kramer, que expresa el grado y asociación entre variables nominales, pero en tablas mayores de 2×2 . Dicho coeficiente se define de la siguiente manera:

$$V = \sqrt{\frac{\chi^2}{n(L-1)}}$$

donde:

n = total de casos (tamaño de la muestra).

L = columna o renglón, con casos, más pequeño de los dos.

A diferencia del C , este coeficiente tiene 1 como límite superior.

Como prueba de significancia, cuando el estadístico χ^2 que se utiliza al calcular V es estadísticamente significativo, dicho coeficiente también lo es.

EJEMPLO

10.9

A 350 estudiantes de una escuela preparatoria se les aplica una prueba que mide la habilidad verbal; los resultados se correlacionarán con los correspondientes a Historia de México, lo cual se llevará a cabo al final del curso. Los resultados se muestran en la tabla de contingencias.

		Calificaciones en Historia de México			
		10–8.1	8–6	Inferior a 6	Total
Habilidad verbal	Superior a 600	100	15	5	120
	Entre 500 y 599	50	85	15	150
	Inferior a 500	10	20	50	80
	Total	160	120	70	350

Paso 1. Se calculan las frecuencias esperadas (fe) mediante sus correspondientes totales marginales, que luego se dividen entre el total n : (54.9), (41.1), (24), ..., (16).

100 (54.9)	15 (41.1)	5 (24)	120
50 (68.6)	85 (51.4)	15 (30)	150
10 (36.6)	20 (27.4)	50 (16)	80
160	120	70	350

Paso 2. Se calcula el valor de χ^2 mediante la fórmula conocida:

$$\chi^2 = \sum \frac{(fo - fe)^2}{fe}$$

$$\begin{aligned} \chi_c^2 &= \frac{(100 - 54.9)^2}{54.9} + \frac{(15 - 41.1)^2}{41.1} + \frac{(5 - 24)^2}{24} + \frac{(50 - 68.6)^2}{68.6} + \frac{(85 - 51.4)^2}{51.4} + \\ &\quad \frac{(15 - 30)^2}{30} + \frac{(10 - 36.6)^2}{36.6} + \frac{(20 - 27.4)^2}{27.4} + \frac{(50 - 16)^2}{16} = 196.6 \end{aligned}$$

$$\therefore \chi_c^2 = 196.6$$

Paso 3. Se sustituyen los valores de $n = 350$, $L = 3$, así como de χ_c^2 obtenida anteriormente, en la definición de V .

$$V = \sqrt{\frac{\chi_c^2}{n(L-1)}}$$

$$V = \sqrt{\frac{196.6}{350(3-1)}}$$

$$V = \sqrt{\frac{196.6}{700}} = 0.53$$

$$\therefore V = 0.53$$

Paso 4. Teóricamente, el coeficiente de Kramer de la correlación entre la habilidad verbal y las calificaciones que obtuvieron en Historia de México cada uno de los 350 estudiantes, es de 53%.

Prueba de significancia de V

De la misma manera que C, en el coeficiente V de Kramer se obtiene la prueba de significancia. Es decir, si $\chi^2_c \geq \chi^2_{crit}$, H_0 se rechaza y el coeficiente V es significativo.

Para $\alpha = 5\%$, $\alpha = 1\%$ y $\alpha = 0.005$ con $gl = (3 - 1)(3 - 1) = 4$, se obtienen:

$$\chi^2(4, 0.05) = 9.49$$

$$\chi^2(4, 0.01) = 13.28$$

$$\chi^2(4, 0.005) = 14.86$$

Como $196.6 \gg 14.86$, el coeficiente $V = 0.53$ es significativo.

10.7

PRUEBA EXACTA DE FISHER

Esta prueba se utiliza para comparar proporciones binominales de dos muestras independientes; dichas proporciones se ubican en una tabla de contingencia de dos renglones por dos columnas (2×2).

Cuando las frecuencias esperadas no son menores que 5, en cuando menos una casilla, se aplica el estadístico de prueba χ^2 (ji cuadrada).

Requisitos de uso:

1. Las muestras son independientes.
2. Existe una casilla con una frecuencia esperada menor que 5.
3. Los valores marginales de la tabla son fijos (esto es necesario para calcular probabilidades exactas).
4. Las dos muestras son pequeñas, por lo que no siempre se cumple que las proporciones binominales tengan una aproximación normal.

	Variable 1		
			Total
Grupo 1	a	b	a + b
Grupo 2	c	d	c + d
Total	a + c	b + d	n

$$Pr(a, b, c, d) = \frac{(a + b)! (c + d)! (a + c)! (b + d)!}{n! a! b! c! d!}$$

Hipótesis:

$$H_0: P_1 = P_2$$

$$H_1: P_1 \neq P_2$$

EJEMPLO

10.10

Suponga que 15 personas fallecieron recientemente en un hospital, donde la causa de muerte de ocho de ellos fue no-cardiovascular (P_1), pero en siete de ellos sí lo fue (P_2). De las 15 personas fallecidas, seis tenían alto consumo de sal en su dieta y nueve lo tenían bajo. Los datos se ubican en una tabla de contingencia de la siguiente forma:

		Tipo de dieta		
		Consumo alto de sal	Consumo bajo de sal	Total
Causa de muerte	Cardiovascular	2	5	7
	No cardiovascular	4	4	8
		6	9	15

Calcule la prueba exacta de Fisher

$$P = 0.608$$

$$P(2,5,4,4) = \frac{7! 8! 6! 9!}{15! 2! 5! 4! 4!} = 0.29$$

10.8

PRUEBA DE McNEMAR

Para el caso de una misma muestra, esta prueba mide una situación antes y después de posibles cambios. También se aplica el modelo para dos muestras dependientes o correlacionadas, obtenidas en forma aleatoria de una misma población, midiendo de igual manera los posibles cambios antes y después.

EJEMPLO

10.11

Para un grupo de pacientes que recibe determinado tratamiento (el mismo grupo de pacientes es su control), la variable de interés sería: actitud positiva o negativa, bienestar o malestar, acuerdo o desacuerdo, alivio o no alivio, cumplimiento de la norma establecida o no, entre otros aspectos.

En el caso más simple, se establecería la hipótesis nula (H_0) en términos de igualdad de proporciones correlacionadas; el grupo se mide antes del tratamiento y después del mismo.

La variable es cualitativa (nominal) y dicotómica. La proporción (número de sujetos) que pertenece a una categoría es mutuamente excluyente con la proporción de las otras categorías; dichas categorías pueden representarse en una tabla de contingencia (de doble entrada) como la siguiente:

		Situación 2		
		+	−	Total
Situación 1	+	a	b	$a + b$
	−	c	d	$c + d$
Total		$a + c$	$b + d$	N

En el anterior, a , b , c y d son las probabilidades de intersección (*joint*) y $a + b$, $c + d$, $a + c$ y $b + d$ son las probabilidades marginales. El modelo estadístico por emplear, si el número de casos de $c + b$ es mayor que 20, será una aproximación a la prueba z .

$$z = \frac{c - b \pm 1}{\sqrt{c + b}}, \text{ si:}$$

- 1) $c > b$, el signo es negativo (−).
- 2) $c \leq b$, el signo es positivo (+).

Si $c + b \leq 20$, entonces el valor de b se compara con un valor crítico en la tabla. Las hipótesis se establecen en la siguiente fórmula:

$$H_0 : c = b$$

Para la hipótesis alterna:

- 1) $H_1 : c > b$
- 2) $H_1 : c < b$
- 3) $H_1 : c \neq b$

El valor crítico se obtiene en la tabla z , con la siguiente regla de decisión:
Si $z_{calc} \geq z_{tablas}$, H_0 se rechaza.

EJEMPLO

10.12

Unas flores recibieron un nuevo tipo de mantenimiento, que se valorará mediante un experimento antes-después, en el que será aplicada la prueba de McNemar. Las observaciones proporcionaron los siguientes resultados:

		Después		
		Rosa	Clavel	
Antes	Rosa	20	18	38
		a	b	
	Clavel	6	38	44
		c	d	
		26	56	82

Las hipótesis estadísticas son las siguientes:

H_0 : el mantenimiento de la rosa es igual antes y después.

H_1 : el mantenimiento de la rosa es diferente antes y después.

Se desea realizar la prueba con un $\alpha = 5\%$.

La región de rechazo para H_0 es:

Es decir, si $|z_c| \geq z_{(0.025)} = 1.96$, H_0 se rechaza.

Como $b + c > 20$ y $b > c$, se utiliza la siguiente fórmula:

$$z = \frac{(c - b) + 1}{\sqrt{c + b}}$$

Se sustituye

$$z = \frac{(6 - 18) + 1}{\sqrt{24}}$$

$$z_c = \frac{-12 + 1}{\sqrt{24}} = \frac{-11}{4.9}$$

$$z_c = -2.24$$

Se aplica posteriormente la regla de decisión:

$$z_c = |-2.24| = 2.24$$

$$2.24 > 1.96 = z_{0.025}$$

Entonces, H_0 se rechaza.

Solución El mantenimiento de la rosa es diferente antes y después.

COMPARACIÓN DE VARIAS POBLACIONES MULTINOMIALES: UNA CLASIFICACIÓN DE DOS VÍAS CON TOTALES DE FILA O COLUMNA FIJOS

10.9

Una tabla de contingencia $r \times c$ resulta cuando cada una de las n unidades experimentales se cuenta como si cayera en una de las rc celdas de un experimento multinomial. Cada celda representa un par de niveles de categoría, nivel de fila i y nivel de columna j . A veces, sin embargo, no es aconsejable usar este tipo de diseño experimental, es decir, hacer que n observaciones caigan donde puedan. Por ejemplo, supongamos que se desea estudiar las opiniones de familias estadounidenses acerca de sus niveles de ingreso, es decir, bajos, regulares y altos. Si se seleccionan al azar $n = 1200$ familias para ese estudio, puede que no se encuentre ninguna

que se clasifique a sí misma como de bajos ingresos. Podría ser mejor decidir por anticipado hacer un estudio de 400 familias de cada nivel de ingreso. Los datos resultantes aparecerán todavía como clasificación de dos vías, pero los totales de columna se fijan por anticipado.

EJEMPLO**10.13**

En otro experimento de prevención de gripe como el del ejemplo 10.5, el experimentador decide buscar los registros clínicos de 300 pacientes de cada una de las tres categorías de tratamiento: sin vacuna, una inyección y dos inyecciones. Los $n = 900$ pacientes se encuestarán entonces respecto a su historial de gripe en invierno. El experimento resulta en una tabla de 2×3 con los totales de columna fijos en 300, como se ve en la tabla 10.8. Al fijar los totales de columna, el experimentador ya no tiene un experimento multinomial con $2 \times 3 = 6$ celdas. En cambio, hay tres experimentos binomiales separados, llamémoslos 1, 2 y 3, cada uno con una probabilidad p_j determinada de contraer la gripe y q_j de no contraer la gripe. (Recuerde que para una población binomial, $p_j + q_j = 1$.)

TABLA 10.8**Casos de gripe para tres tratamientos**

	Sin vacuna	Una inyección	Dos inyecciones	Total
Gripe				r_1
Sin gripe				r_2
Total	300	300	300	n

Supongamos que se utilizó el estadístico de prueba ji cuadrada para probar la independencia de clasificaciones de fila y columna. Si un tratamiento particular (nivel de columna) no afecta la incidencia de gripe, entonces cada una de las tres poblaciones binomiales debería tener la misma incidencia de gripe para que $p_1 = p_2 = p_3$ y $q_1 = q_2 = q_3$.

La clasificación de 2×3 del ejemplo 10.13 describe una situación en la que el estadístico de prueba ji cuadrada de independencia es equivalente a una prueba de la igualdad de $c = 3$ proporciones binomiales. Pruebas de este tipo se llaman **pruebas de homogeneidad** y se usan para comparar diversas poblaciones binomiales. Si hay *más de dos* categorías de renglón con totales de columna fijos, entonces la prueba de independencia es equivalente a una prueba de la igualdad de c conjuntos de proporciones multinomiales.

No es necesario preocuparse de la equivalencia teórica de los estadísticos de prueba ji cuadrada para estos dos diseños experimentales. Si las columnas (o filas) son fijas o no, el estadístico de prueba se calcula como

$$\chi^2 = \sum \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} \quad \text{donde } \hat{E}_{ij} = \frac{r_i c_j}{n}$$

que tiene una distribución ji cuadrada aproximada en muestreo repetido con $gl = (r - 1)(c - 1)$.



NECESITO SABER...

Cómo determinar el número apropiado de grados de libertad

Recuerde el procedimiento general para determinar grados de libertad:

- 1. Empiece con las rc celdas en la tabla de dos vías.
- 2. Reste un grado de libertad por cada una de las c poblaciones multinomiales, cuyas probabilidades de columna deben totalizar uno, un total de c *grados de libertad*.
- 3. Tuvo que estimar $(r - 1)$ probabilidades de fila, pero las probabilidades de columna se fijan por anticipado y no necesitaban ser estimadas. Reste $(r - 1)$.

El total de grados de libertad para la tabla $r \times c$ (columna fija) es

$$rc - c - (r - 1) = rc - c - r + 1 = (r - 1)(c - 1)$$

EJEMPLO

10.14

Una encuesta de opiniones de electores fue realizada en cuatro distritos del centro de una ciudad, para comparar las proporciones de electores que están a favor del candidato A. Muestras aleatorias de 200 electores se encuestaron en cada uno de los cuatro distritos con los resultados que se ven en la tabla 10.9. Los valores entre paréntesis de la tabla son las cantidades de celda esperadas. ¿Los datos presentan suficiente evidencia para indicar que las proporciones de electores que están a favor del candidato A difieren en los cuatro distritos?

TABLA 10.9

Opiniones de electores en cuatro distritos

	Distrito				Total
	1	2	3	4	
A favor de A	76 (59)	53 (59)	59 (59)	48 (59)	236
No a favor de A	124 (141)	147 (141)	141 (141)	152 (141)	564
Total	200	200	200	200	800

Solución Como los totales de columna están fijos en 200, el diseño comprende cuatro experimentos binomiales, cada uno de los cuales contiene las respuestas de 200 electores para cada uno de los cuatro distritos. Para probar la igualdad de las proporciones que están a favor del candidato A en los cuatro distritos, la hipótesis nula

$$H_0 : p_1 = p_2 = p_3 = p_4$$

es equivalente a la hipótesis nula

$$H_0 : \text{La proporción a favor del candidato A es independiente del distrito}$$

y será rechazada si el estadístico de prueba χ^2 es demasiado grande. El valor observado del estadístico de prueba, $\chi^2 = 10.722$, y su valor p asociado, .013, se muestran en la figura 10.4. Los resultados son significativos ($P < .025$); esto es, H_0 es rechazada y se puede concluir que hay diferencia en las proporciones de electores que están a favor del candidato A entre los cuatro distritos.

FIGURA 10.4

Salida MINITAB para el ejemplo 10.14

Prueba ji cuadrada: distrito 1, distrito 2, distrito 3, distrito 4Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

	Ward 1	Ward 2	Ward 3	Ward 4	Total
1	76 59.00 4.898	53 59.00 0.610	59 59.00 0.000	48 59.00 2.051	236
2	124 141.00 2.050	147 141.00 0.255	141 141.00 0.000	152 141.00 0.858	564
Total	200	200	200	200	800

Chi-Sq = 10.722 DF = 3, P-Value = 0.013

¿Cuál es la naturaleza de las diferencias descubiertas por la prueba ji cuadrada? Para contestar esta pregunta, véase la tabla 10.10, que muestra las proporciones muestrales que están a favor del candidato A en cada uno de los cuatro distritos. Parece que el candidato A está haciéndolo mejor en el primer distrito y peor en el cuarto distrito. ¿Es esto de alguna *significancia práctica* para el candidato? Posiblemente una observación más importante es que el candidato no tiene una pluralidad de electores en ninguno de los cuatro distritos. Si ésta es una carrera de dos candidatos, el candidato A necesita aumentar su campaña.

TABLA 10.10**Proporciones a favor del candidato A en cuatro distritos**

Distrito 1	Distrito 2	Distrito 3	Distrito 4
76/200 = .38	53/200 = .27	59/200 = .30	48/200 = .24

10.9**EJERCICIOS****TÉCNICAS BÁSICAS**

10.21 Muestras aleatorias de 200 observaciones se seleccionaron de cada una de tres poblaciones y cada observación se clasificó de acuerdo con si cayó en una de tres categorías mutuamente exclusiva:

Población	Categoría			Total
	1	2	3	
1	108	52	40	200
2	87	51	62	200
3	112	39	49	200

Se desea saber si los datos dan suficiente evidencia para indicar que las proporciones de observaciones en las tres categorías dependen de la población de la cual se sacaron.

- Dé el valor de χ^2 para la prueba.
- Dé la región de rechazo para la prueba para $\alpha = .01$.
- Expresé sus conclusiones.
- Encuentre el valor p aproximado para la prueba e interprete su valor.

10.22 Suponga que se desea probar la hipótesis nula de que tres parámetros binomiales p_A , p_B y p_C son iguales contra la hipótesis alternativa de que al menos dos de los parámetros difieren. Muestras aleatorias independientes de cien observaciones se seleccionaron de entre cada una de las poblaciones. Los datos se muestran en la tabla.

	Población			Total
	A	B	C	
Éxitos	24	19	33	76
Fracasos	76	81	67	224
Total	100	100	100	300

- Escriba las hipótesis nula y alternativa para probar la igualdad de las tres proporciones binomiales.
- Calcule el estadístico de prueba y encuentre el valor p aproximado para la prueba en la parte a.
- Use el valor p aproximado para determinar la significancia estadística de sus resultados. Si los resultados son estadísticamente significativos, explore

la naturaleza de las diferencias en las tres proporciones binomiales.

APLICACIONES

10.23 Pollos enfermos Se piensa que una enfermedad particular en pollos no es transmisible. Para probar esta teoría, 30 000 pollos se dividieron al azar en tres grupos de 10 000. Un grupo no tenía contacto con pollos enfermos, otro tenía contacto moderado y el tercero tenía contacto frecuente. Después de seis meses se recabaron datos sobre el número de pollos enfermos en cada grupo de 10 000. ¿Los datos dan suficiente evidencia para indicar una dependencia entre la cantidad de contacto entre aves enfermas y no enfermas y la incidencia de la enfermedad? Use $\alpha = .05$.

	Sin contacto	Contacto moderado	Contacto frecuente
Enfermedad	87	89	124
No enfermedad	9913	9911	9876
Total	10000	10000	10000

DATOS
EX1024 **10.24 Atención a largo plazo** Un estudio realizado en el noroeste de Inglaterra hizo una evaluación de instalaciones de atención a largo plazo que tienen residentes con demencia.⁵ Las casas incluían aquellas que daban servicio especializado a personas ancianas con enfermedad mental o problemas de salud, conocidas como “casas EMI”, así como otras clasificadas como “casas no EMI”. Se esperaba que las casas EMI tuvieran una calificación más alta en varias medidas de calidad de servicio para personas con demencia. Una medida incluía la estructura de la casa y los servicios proporcionados, como se da en la tabla siguiente.

Tipo de atención	Tipo de casa		
	EMI	No EMI	Total
Enfermería	54	22	76
Atención residencial	59	77	136
Doble registro	49	26	75
Total	162	125	287

- a. Describa los experimentos binomiales cuyas proporciones se han comparado en este experimento.
- b. ¿Estos datos indican que el tipo de atención proporcionado varía por los tres tipos de casa? Pruebe al nivel $\alpha = .01$.

- c. Con base en los resultados del inciso b), explique la naturaleza práctica de la relación entre el tipo de casa y el tipo de atención.

10.25 Investigación de mares profundos W.W.

Menard ha realizado investigaciones respecto a nódulos de manganeso, una mezcla rica en minerales hallada en abundancia en el lecho de mares profundos.⁶ En una parte de su informe, Menard proporciona datos que relacionan la edad magnética de la corteza terrestre con la “probabilidad de hallar nódulos de manganeso”. La tabla siguiente da el número de muestras del núcleo de la Tierra y el porcentaje de las que contienen nódulos de manganeso para cada una de un conjunto de edades de la corteza magnética. ¿Estos datos dan suficiente evidencia para indicar que la probabilidad de hallar nódulos de manganeso en la corteza de mares profundos de la Tierra depende de la clasificación de edad magnética?

Edad	Número de muestras	Porcentaje con nódulos
Mioceno; reciente	389	5.9
Oligoceno	140	17.9
Eoceno	214	16.4
Paleoceno	84	21.4
Cretácico tardío	247	21.1
Cretácico temprano o medio	1120	14.2
Jurásico	99	11.0

DATOS
EX1026 **10.26 ¿Qué tan grande es la familia?** Una cámara de comercio local encuestó a 120 familias en su ciudad, 40 en cada uno de tres tipos de residencia (departamentos, dúplex o casas solas) y registró el número de miembros en cada una de las familias. Los datos se muestran en la tabla.

Miembros de familia	Tipo de residencia		
	Departamento	Dúplex	Casa sola
1	8	20	1
2	16	8	9
3	16	10	14
4 o más	6	2	16

¿Hay diferencia significativa en las distribuciones del tamaño de familia para los tres tipos de residencia? Pruebe usando $\alpha = .01$. Si hay diferencias significativas, describa su naturaleza.

LA EQUIVALENCIA DE PRUEBAS ESTADÍSTICAS

10.10

Recuerde que cuando hay sólo $k = 2$ categorías en un experimento multinomial, el experimento se reduce a un *experimento binomial* donde se registra el número de éxitos x (o O_1) en n (o $O_1 + O_2$) intentos. Del mismo modo, los datos que resultan de *dos experimentos binomiales* se pueden exhibir en una clasificación de dos vías con $r = 2$ y $c = 2$, de modo que el estadístico de prueba ji cuadrada de *homogeneidad* se puede usar para comparar las dos proporciones binomiales, p_1 y p_2 . Para estas dos situaciones, hemos presentado pruebas estadísticas para las proporciones binomiales basadas en el estadístico z :

• **Una muestra:**
$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

 $k = 2$

Éxitos	Fracasos

• **Dos muestras:**
$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

 $r = c = 2$

Muestra 1	Muestra 2
Éxitos	Éxitos
Fracasos	Fracasos

TIP MI CONSEJO

Las pruebas binomiales de una y de dos muestras del capítulo 9 son equivalentes a pruebas ji cuadrada, $z^2 = \chi^2$

¿Por qué hay dos pruebas diferentes para la misma hipótesis estadística? ¿Cuál debería usarse? Para estas dos situaciones, se puede usar *ya sea* la prueba z o *bien* la prueba ji cuadrada, y se obtendrán resultados idénticos. Para la prueba de una o de dos muestras, podemos demostrar algebraicamente que

$$z^2 = \chi^2$$

de modo que el estadístico de prueba será la raíz cuadrada (ya sea positiva o negativa, dependiendo de los datos) del estadístico ji cuadrada. Además, podemos demostrar teóricamente que la misma relación se cumple para los valores críticos de las tablas z y χ^2 del Anexo, que produce *valores p idénticos* para las dos pruebas equivalentes. Para probar una hipótesis alternativa de una cola como $H_0: p_1 > p_2$, primero se determina si $\hat{p}_1 - \hat{p}_2 > 0$, es decir, si la diferencia en proporciones muestrales tiene el signo apropiado. Si es así, el valor crítico apropiado de χ^2 de la tabla 5 en el Anexo tendrá un grado de libertad y un área de cola derecha de 2α . Por ejemplo, el valor crítico χ^2 con 1 gl y $\alpha = .05$ será $\chi^2_{.10} = 2.70554 = 1.645^2$.

En resumen, usted es libre de elegir la prueba (z o χ^2) que sea más cómoda. Como casi todos los paquetes de computadora incluyen la prueba ji cuadrada y la mayoría de ellos no incluyen las pruebas de muestra grande, la prueba ji cuadrada puede ser preferible.

OTRAS APLICACIONES DEL ESTADÍSTICO DE PRUEBA JI CUADRADA

10.11

La aplicación de la prueba ji cuadrada para analizar datos de cantidades es sólo uno de muchos problemas de clasificación que resultan en datos multinomiales. Algunas de estas aplicaciones son bastante complejas, requiriendo procedimientos complicados o difíciles desde el punto de

vista de cálculos para estimar las cantidades de celda esperadas. No obstante, varias aplicaciones se utilizan con suficiente frecuencia para hacerlas dignas de mención:

- **Pruebas de bondad de ajuste:** Se puede diseñar una prueba de bondad de ajuste para determinar si los datos son consistentes con datos tomados de una distribución particular de probabilidad, posiblemente normal, binomial, de Poisson u otras distribuciones. Las celdas de un histograma de frecuencia muestral corresponden a las k celdas de un experimento multinomial. Las cantidades de celda esperadas se calculan usando las probabilidades asociadas con la distribución hipotética de probabilidad.
- **Multinomiales dependientes del tiempo:** Se puede usar el estadístico ji cuadrada para investigar la rapidez de cambio de proporciones multinomiales (o binomiales) en el tiempo. Por ejemplo, suponga que la proporción de respuestas correctas en un examen de 100 preguntas se registra para un estudiante, que entonces repite el examen en cada una de las siguientes cuatro semanas. ¿La proporción de respuestas correctas aumenta con el tiempo? ¿Tiene lugar un aprendizaje? En un proceso monitoreado por un plan de control de calidad, ¿hay una tendencia positiva en la proporción de artículos defectuosos como función del tiempo?
- **Tablas de contingencia multidimensionales:** En lugar de sólo dos métodos de clasificación, se puede investigar una dependencia entre tres o más clasificaciones. La tabla de contingencia de dos vías se extiende a una tabla en más de dos dimensiones. La metodología es similar a la que se emplea para la tabla de contingencia de $r \times c$, pero el análisis es un poco más complejo.
- **Modelos log-lineales:** Modelos complejos se pueden crear en donde el logaritmo de la probabilidad de celda ($\ln p_{ij}$) es alguna función lineal de las probabilidades de renglón y columna.

Casi todas estas aplicaciones son más bien complejas y podrían requerir el consejo de un estadístico profesional antes de realizar un experimento.

En todas las aplicaciones estadísticas que usen *estadístico ji cuadrada de Pearson*, las suposiciones deben estar satisfechas para que el estadístico de prueba tenga una distribución de probabilidad ji cuadrada aproximada.

SUPOSICIONES

- Las cantidades de celda $O_1, O_2, \dots, O_k$ deben satisfacer las condiciones de un experimento multinomial, o un conjunto de experimentos multinomiales creados al fijar ya sea los totales de renglón o de columna.
- Las cantidades de celda esperadas $E_1, E_2, \dots, E_k$ deben ser iguales a o mayores que 5.

Por lo general se puede estar razonablemente seguro de haber satisfecho la primera suposición si con todo cuidado se prepara o diseña un experimento o encuesta muestral. Cuando calcule las cantidades de celda esperadas, si encuentra que una o más es menor que 5, existen estas opciones:

- Elija un tamaño n muestral más grande. Cuanto más grande sea el tamaño muestral, más cerca se aproximará la distribución ji cuadrada a la distribución de su estadístico de prueba χ^2 .
- Puede ser posible combinar una o más de las celdas con pequeñas cantidades de celdas esperadas, con lo cual se satisface la suposición.

Por último, asegúrese de estar calculando los *grados de libertad* correctamente y que con todo cuidado se evalúan las conclusiones estadísticas y prácticas que se pueden sacar de la prueba.

REPASO DEL CAPÍTULO

Conceptos y fórmulas clave

I. El experimento multinomial

1. Hay n intentos idénticos y cada resultado cae en una de k categorías
2. La probabilidad de caer en la categoría i es p_i y permanece constante de un intento a otro
3. Los intentos son independientes, $\sum p_i = 1$, y medimos O_i , el número de observaciones que caen en cada una de k categorías

II. Estadístico ji cuadrada de Pearson

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \quad \text{donde } E_i = np_i$$

que es una distribución ji cuadrada aproximada con *grados de libertad* determinados por la aplicación

III. La prueba de bondad de ajuste

1. Ésta es una clasificación de una vía con probabilidades de celda especificadas en H_0
2. Use el estadístico ji cuadrada con $E_i = np_i$ calculada con las probabilidades hipotéticas
3. $GL = k - 1$ (Número de parámetros estimados para hallar E_i)
4. Si H_0 es rechazada, investigue la naturaleza de las diferencias usando las proporciones muestrales

IV. Tablas de contingencia

1. Una clasificación de dos vías con n observaciones en categorías de $r \times c$ celdas de una tabla de dos vías, que usa dos métodos diferentes de clasificación, se denomina *tabla de contingencia*

2. La prueba de independencia de métodos de clasificación usa el estadístico ji cuadrada

$$\chi^2 = \sum \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

$$\text{con } \hat{E}_{ij} = \frac{r_i c_j}{n} \quad \text{y} \quad gl = (r - 1)(c - 1)$$

3. Si la hipótesis nula de independencia de clasificaciones se rechaza, investigue la naturaleza de la dependencia usando proporciones condicionales dentro de ya sea filas o columnas de la tabla de contingencia

V. Fijar totales de fila o columna

1. Cuando sea que los totales de fila o los de columna sean fijos, la prueba de independencia de clasificaciones se convierte en una prueba de la homogeneidad de probabilidades de celda para varios experimentos multinomiales
2. Use el mismo estadístico ji cuadrada para tablas de contingencia
3. Las pruebas z de muestra grande para una y dos proporciones binomiales son casos especiales del estadístico ji cuadrada

VI. Suposiciones

1. Las cantidades de celda satisfacen las condiciones de un experimento multinomial o un conjunto de experimentos multinomiales con tamaños muestrales fijos
2. Todas las cantidades de celda esperadas deben ser iguales a o mayores que cinco para que la aproximación ji cuadrada sea válida



TECNOLOGÍA ACTUAL

La prueba ji cuadrada: *Microsoft Excel*

El procedimiento para realizar una prueba de independencia ji cuadrada en *MS Excel* requiere que introduzca tanto los conteos de celda observados como los esperados en una hoja de cálculo de *Excel*. Si se han almacenado en la hoja de cálculo los datos categóricos sin procesar en lugar de los conteos de celda observados, puede ser necesario contar los datos para obtener los conteos de celda antes de continuar.

EJEMPLO

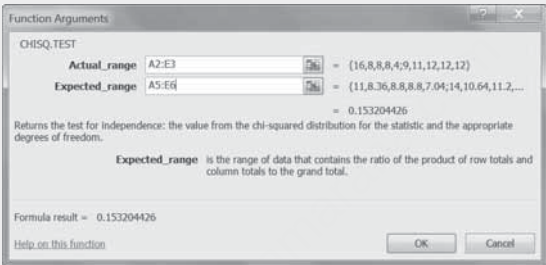
10.15

Suponga que ha registrado el género (M o F) y el nivel universitario (1o, 2o, 3o, 4o, Graduado) para 100 estudiantes de estadística, como se muestra en la siguiente tabla.

Género	Nivel				
	1o	2o	3o	4o	Graduado
F	16	8	8	8	4
M	9	11	12	12	12

- 1. Introduzca los valores observados en las primeras cinco columnas de una hoja de cálculo de *Excel*.
- 2. Calcule (en forma manual) los 10 conteos de celda esperados estimados e introdúzcalos en otro rango en la hoja de cálculo.
- 3. Coloque el cursor en una celda vacía y use **Formulas ► More Functions ► Statistical ► CHISQ.TEST (CHITEST en versiones anteriores de Excel)** para generar el cuadro de diálogo que se muestra en la figura 10.5. Resalte o teclee los rangos de celda para los conteos de celda observados y esperados.

FIGURA 10.5



- 4. Cuando haga clic en **OK**, *MS Excel* calculará el valor *p* asociado con la prueba de independencia ji cuadrada. Para estos datos, el valor *p* grande (.153) indica un resultado no significativo. Hay evidencia insuficiente para indicar que el género de un estudiante depende de su nivel en clase.

NOTA: *MS Excel* no proporciona un comando **único que permita** realizar la prueba de bondad de ajuste ji cuadrada; sin embargo, podría crear en forma manual fórmulas en *MS Excel* para ejecutar esta prueba y obtener el valor *p* apropiado.

La prueba ji cuadrada: MINITAB

Existen varios procedimientos en el paquete *MINITAB* para analizar datos categóricos. El procedimiento apropiado depende de si los datos representan una clasificación de una vía (un solo experimento multinomial) o una clasificación de dos vías o tabla de contingencia. Si los *datos categóricos sin elaborar* se han guardado en la hoja de cálculo *MINITAB* más que las *cantidades de celda observadas*, puede ser necesario totalizar o clasificar en forma cruzada los datos para obtener las cantidades de celda antes de continuar.

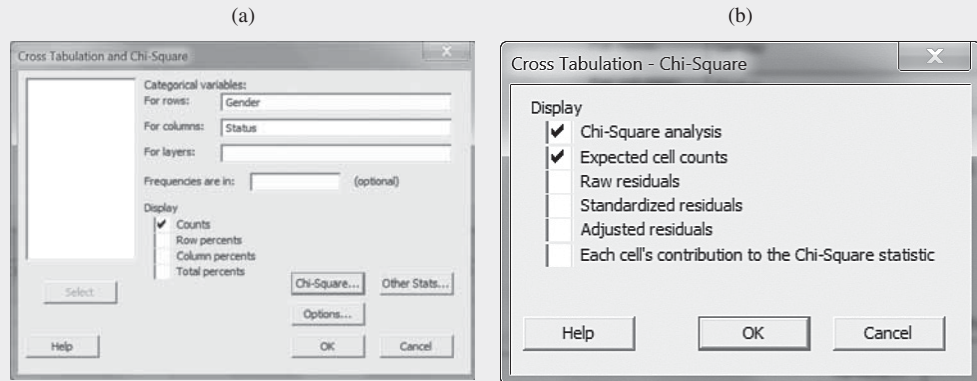
EJEMPLO

10.16

Suponga que ha registrado el género (M o F) y el nivel universitario (1o, 2o, 3o, 4o, graduado) para 100 estudiantes de estadística. La hoja de cálculo de *MINITAB* contendría dos columnas de 100 observaciones cada una. Cada fila contendría un género del individuo en la columna 1 y el nivel universitario en la columna 2.

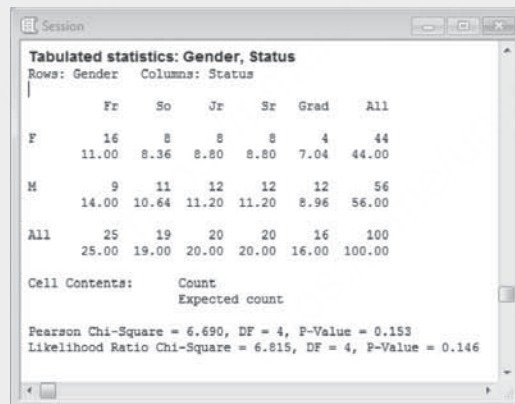
- 1. Para obtener los conteos de celda observados (O_{ij}) para la tabla de contingencia de 2×5 , use **Stat ► Tables ► Cross Tabulation and Chi-Square** para generar el cuadro de diálogo que se muestra en la figura 10.6a).

FIGURA 10.6



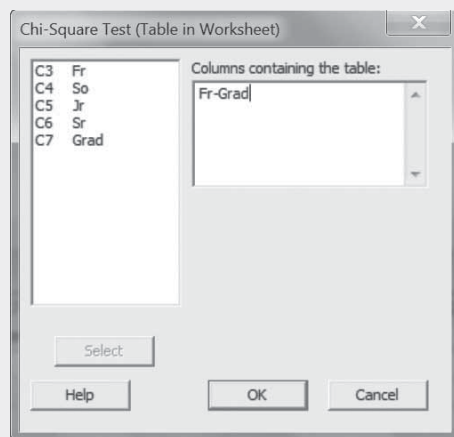
2. Bajo “Categorical Variables”, seleccione “Género” para la variable de fila y “Nivel” para la variable de columna. Deje en blanco los cuadros marcados “For Layers” y “Frequencies are in:”. Asegúrese que está seleccionado el cuadro etiquetado “Display Counts”.
3. Haga clic en el botón **Chi-Square...** para mostrar el cuadro de diálogo que aparece en la figura 10.6b). Seleccione los cuadros para “Chi-Square Analysis” y “Expected Cell Counts”. Haga clic en **OK** dos veces. Esta secuencia de comandos no sólo tabula la tabla de contingencia sino también realiza la prueba de independencia ji cuadrada y muestra los resultados en la ventana Session mostrada en la figura 10.7. Para los datos de género/nivel universitario, el valor p grande ($P = .153$) indica un resultado no significativo. Hay evidencia insuficiente para indicar que el género de un estudiante depende de su nivel en clase.

FIGURA 10.7



4. Si los conteos de celda observados en la tabla de contingencia ya han sido tabulados, simplemente introduzca los conteos en c columnas de la hoja de cálculo de MINITAB, use **Stat ► Tables ► Chi-Square Test (Two-Way Table in Worksheet)** y seleccione las columnas apropiadas antes de hacer clic en **OK**. Para los datos de género/nivel universitario, puede introducir los conteos en las columnas C3–C7 como se muestra en la figura 10.8. La salida resultante será etiquetada de manera diferente pero se verá exactamente como la salida en la figura 10.7.

FIGURA 10.8



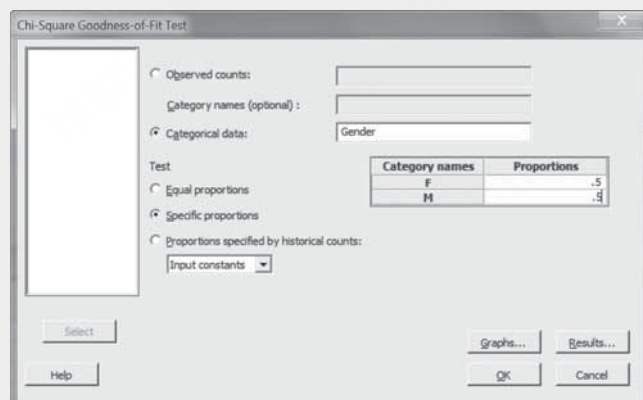
EJEMPLO

10.17

Se puede establecer una prueba simple de un experimento multinomial único al considerar si las proporciones de hombres y mujeres estudiantes de estadística son iguales; es decir, $p_1 = .5$ y $p_2 = .5$.

1. En *MINITAB 15* o *16*, use **Stat ► Tables ► Chi-Square Goodness-of-Fit Test (One Variable)** para desplegar el cuadro de diálogo que se muestra en la figura 10.9. Si tiene datos categóricos sin procesar en una columna, haga clic en el botón “Categorical data:” e introduzca la columna “Género” en la celda. Si tiene valores de resumen de conteos observados para cada categoría, elija “Observed counts”. Luego introduzca la columna que contiene los conteos observados o teclee los conteos observados para cada categoría.

FIGURA 10.9



2. Para esta prueba, se puede elegir “Equal proportions” para probar $H_0: p_1 = p_2 = .5$. Cuando tiene diferentes proporciones para cada categoría, use “Specific proportions”. Puede almacenar las proporciones para cada categoría en una columna, elija “Input column” e introduzca la columna. Si desea teclear la proporción para cada categoría, seleccione “Input constants” y teclee las proporciones para las categorías correspondientes. Haga clic en **OK**.
3. La salida resultante incluirá varias gráficas junto con los valores para O_i y E_i para cada categoría, el valor observado del estadístico de prueba, $\chi^2 = 1.44$ y su valor $p = 0.230$, el cual no es significativo. Hay evidencia insuficiente para indicar una diferencia en la proporción de hombres y mujeres estudiantes de estadística.

NOTA: Si usa una versión previa de *MINITAB* tendrá que determinar los conteos de celda observados y esperados, e introducirlos en columnas separadas en la hoja de cálculo. Luego use **Calc ► Calculator** y la expresión $\text{SUM}((\text{'O'} - \text{'E'})**2/\text{'E'})$ para calcular el valor observado del estadístico de prueba.

Ejercicios suplementarios

10.27 Cera para pisos Un fabricante de cera para pisos realizó un experimento de preferencia del consumidor para ver si una nueva cera para pisos A era mejor que las producidas por cuatro competidores, B, C, D y E. Una muestra de cien amas de casa vieron cinco parches de piso que habían recibido las cinco ceras y cada una indicó el parche que consideraba mejor en apariencia. La iluminación, el fondo y otros factores eran aproximadamente iguales para los cinco parches. Los resultados del estudio se ven a continuación:

Cera	A	B	C	D	E
Frecuencia	27	17	15	22	19

¿Los datos presentan suficiente evidencia para indicar una preferencia por uno o más de los parches pulidos de piso sobre los otros? Si uno fuera a rechazar la hipótesis de no preferencia para este experimento, ¿implicaría esto que la cera A es mejor que las otras? ¿Se puede sugerir una mejor forma de realizar el experimento?

10.28 Buena condición física en adultos Se realizó una encuesta para determinar si la participación de adultos en programas de acondicionamiento físico varía de una región a otra. Se entrevistó a una muestra aleatoria de personas en cada uno de cuatro estados y se registraron estos datos:

	A	B	C	D
Participan	46	63	108	121
No participan	149	178	192	179

¿Los datos indican una diferencia en participación de adultos en programas de acondicionamiento físico de un estado a otro? Si es así, describa la naturaleza de las diferencias.

10.29 Accidentes fatales Se analizaron datos de accidentes para determinar los números de accidentes fatales para automóviles de tres tamaños. Los datos de 346 accidentes son como sigue:

	Pequeño	Mediano	Grande
Fatal	67	26	16
No fatal	128	63	46

¿Los datos indican que la frecuencia de accidentes fatales depende del tamaño de los automóviles? Escriba un breve párrafo que describa sus resultados estadísticos y las implicaciones prácticas de éstos.

10.30 Médicos y pacientes de asistencia médica gratuita Se realizó un experimento para investigar el efecto de la experiencia general en hospital, en las actitudes de médicos hacia pacientes de asistencia médica

gratuita. Una muestra aleatoria de 50 médicos que acababan de terminar 4 semanas de servicio en un hospital general fueron clasificados de acuerdo con su interés por pacientes de asistencia médica gratuita, antes y después de su experiencia en el hospital general. Los datos se ven en la tabla. ¿Estos datos dan suficiente evidencia para indicar un cambio en “interés” después de la experiencia en el hospital general? Si es así, describa la naturaleza del cambio.

Interés antes	Interés después		Total
	Alto	Bajo	
Bajo	27	5	32
Alto	9	9	18

Salida impresa parcial MINITAB para el ejercicio 10.30

Prueba ji cuadrada: alto, bajo

Chi-Sq = 6.752, DF = 1, P-Value = 0.009



10.31 Enseñanza basada en

descubrimientos Dos profesores de biología se propusieron evaluar los efectos de la enseñanza basada en descubrimientos, en comparación con el método de enseñanza estándar basado en clases en el laboratorio.⁷ Este último dio una lista de instrucciones a seguir en cada paso del ejercicio de laboratorio, mientras que el método basado en descubrimientos hizo preguntas en lugar de dar instrucciones y utilizó informes de grupo pequeño para decidir la mejor forma de continuar para llegar al objetivo de laboratorio. Una evaluación de las técnicas comprendía evaluaciones por escrito de ambos procedimientos por estudiantes al final del curso. La comparación del número de respuestas positivas y negativas para ambas técnicas se da en la tabla siguiente.

Grupo	Evaluaciones positivas	Evaluaciones negativas	Total
Descubrimiento	37	11	48
Control	31	17	48

a. ¿Hay una diferencia significativa en la proporción de respuestas positivas para cada uno de los métodos de enseñanza? Use $\alpha = .05$. Si es así, ¿cómo describiría esta diferencia?

b. ¿Cuál es el valor p aproximado para la prueba de la parte a?

10.32 Posición de un bebé dormido ¿La posición de un bebé dormido afecta el desarrollo de habilidades motrices? En un estudio, 343 niños nacidos a término fueron examinados en su revisión de cuarto mes en busca de varios aspectos importantes de desarrollo, por ejemplo rodar, tomar una sonaja o alcanzar un objeto.⁸ La posición predominante

del bebé al dormir, ya sea boca abajo, de espaldas o de costado, fue determinada por una entrevista telefónica con los padres. Los resultados de la muestra de 320 de los 343 bebés de quienes se recibió información se ven en la tabla siguiente. El investigador informó que los bebés que dormían de costado o boca arriba eran menos susceptibles de rodarse en su revisión de cuarto mes que los que dormían predominantemente boca abajo ($P < .001$).

	Boca abajo	Boca arriba o de costado
Número de bebés	121	199
Número que se rodaban	93	119

- a. Use una prueba z de muestra grande para confirmar o refutar la conclusión del investigador.
- b. Reescriba los datos de la muestra como una tabla de contingencia de 2×2 . Use la prueba ji cuadrada para homogeneidad para confirmar o refutar la conclusión del investigador.
- c. Compare los resultados de las partes a y b. Confirme que los dos estadísticos de prueba están relacionados como $z^2 = \chi^2$ y que los valores críticos para rechazar H_0 tienen la misma relación.

10.33 Consulte el ejercicio 10.32. Encuentre el valor p para la prueba z de muestra grande del inciso a). Compare este valor p con el valor p para la prueba ji cuadrada, mostrada en la salida impresa parcial MINITAB.

Salida impresa parcial MINITAB para el ejercicio 10.33

Prueba ji cuadrada: boca abajo, de costado

Chi-Sq = 9.795, DF = 1, P-Value = 0.002

10.34 Posición de un bebé dormido II Los investigadores en el ejercicio 10.32 también midieron otros varios aspectos del desarrollo y sus relaciones con la posición predominante del bebé al dormir.⁸ Los resultados de su investigación se presentan en la tabla para los 320 bebés y su revisión de cuarto mes.

Aspecto	Puntuación	Boca abajo	Boca arriba o de lado	P
Se jala para sentarse sin doblar la cabeza	Aprueba	79	144	< .21
	No aprueba	6	20	
Toma una sonaja	Aprueba	102	167	< .13
	No aprueba	3	1	
Alcanza un objeto	Aprueba	107	183	< .97
	No aprueba	3	5	

Use su conocimiento del análisis de datos categóricos para explicar el diseño experimental empleado por los investigadores. ¿Qué hipótesis fueron de interés para los investigadores y qué prueba estadística hubieran usado los investigadores? Explique las conclusiones que se puedan sacar de los tres valores p en la última columna de la tabla, así como las implicaciones prácticas que puedan

sacarse de los resultados estadísticos. ¿Se han violado algunas suposiciones estadísticas?

10.35 Color y forma de una flor Un botánico realiza un cruce secundario de petunias con factores independientes que controlan la forma de la hoja y el color de la flor, donde el factor A representa el color rojo, a representa color blanco, B representa hojas redondas y b representa hojas largas. De acuerdo con el modelo de Mendel, las plantas deben exhibir las características AB , Ab , aB y ab en la proporción 9:3:3:1. De 160 plantas experimentales, se observaron los números siguientes:

AB	Ab	aB	ab
95	30	28	7

¿Hay suficiente evidencia para refutar el modelo de Mendel al nivel $\alpha = .01$?



10.36 Oportunidades de éxito Una encuesta de CBS News⁹ planteó la pregunta “Comparado con la generación de sus padres, ¿piensa que en general sus oportunidades de tener éxito en la vida son mejores que las tuyas, más o menos las mismas que las tuyas o peores que las tuyas?” en tres fechas separadas durante un periodo de 10 años. Todas las encuestas incluyeron $n = 1\,048$ individuos.

Fecha de la encuesta	Mejores	Iguales	Peores	No estoy seguro	Totales
Diciembre de 2009	493	252	283	20	1048
Junio de 2007	650	189	189	20	1048
Febrero de 2000	755	231	52	10	1048

- a. ¿Hay una diferencia significativa entre las respuestas a esta pregunta a lo largo del tiempo? Use $\alpha = .05$
- b. Si se encontraron diferencias significativas en la parte a, describa la naturaleza de estas diferencias.

10.37 Medicina para la artritis Un estudio para determinar la eficacia de un medicamento (suero) para la artritis resultó en la comparación de dos grupos, cada uno formado por 200 pacientes de artritis. Un grupo fue inoculado con el suero; el otro recibió un placebo (inoculación que parece contener suero pero en realidad no es activo). Después de un tiempo, a cada persona del estudio se le pidió que dijera si su afección de artritis había mejorado. Éstos son los resultados:

	Tratado	No tratado
Mejóro	117	74
No mejoró	83	126

Se desea saber si estos datos presentan suficiente evidencia para indicar que el suero fue eficaz para mejorar la condición de pacientes de artritis.

- a. Use la prueba ji cuadrada de homogeneidad para comparar las proporciones mejoradas de las poblaciones de personas tratadas y no tratadas. Pruebe al nivel de significancia de 5%.
- b. Pruebe la igualdad de las dos proporciones binomiales usando la prueba z de dos muestras. Verifique que el valor elevado al cuadrado del estadístico de prueba $z^2 = \chi^2$ de la parte a. ¿Sus conclusiones son iguales a las de la parte a?

10.38 La prueba ji cuadrada empleada en el ejercicio 10.37 es equivalente a la prueba z de dos colas siempre que α sea igual para las dos pruebas. Demuestre algebraicamente que el estadístico χ^2 de la prueba ji cuadrada es el cuadrado de la prueba estadística z para la prueba equivalente.

10.39 Ajuste de una distribución binomial Se puede usar una prueba de bondad de ajuste para determinar si en realidad todos los criterios para un experimento binomial se han satisfecho en una aplicación determinada. Suponga que un experimento consistente en cuatro intentos se repitió cien veces. El número de repeticiones en las que se obtuvo un número dado de éxitos se registró en la tabla:

Resultados posibles (número de éxitos)	Número de veces obtenido
0	11
1	17
2	42
3	21
4	9

Estime p (suponiendo que el experimento fue binomial), obtenga estimaciones de las frecuencias de celda esperadas y pruebe la bondad del ajuste. Para determinar el número apropiado de grados de libertad para χ^2 , observe que p fue estimado por una combinación lineal de las frecuencias observadas.

10.40 Antibióticos e infección A veces ocurren infecciones cuando se aplican transfusiones de sangre durante operaciones quirúrgicas. Se realizó un experimento para determinar si la inyección de anticuerpos redujo la probabilidad de infección. Un examen de los registros de 138 pacientes produjo los datos que se ven en la tabla siguiente. ¿Los datos dan suficiente evidencia para indicar que las inyecciones de anticuerpos afectan la probabilidad de infecciones en transfusiones? Pruebe usando $\alpha = .05$.

	Infección	No infección
Anticuerpos	4	78
Sin anticuerpos	11	45

10.41 Manufactura alemana Los sindicatos de Estados Unidos han estado tradicionalmente contentos con dejar la administración de compañías a gerentes y ejecutivos corporativos. Pero, en Europa, la participación de los

trabajadores en la toma de decisiones administrativas es una idea aceptada que se está extendiendo en forma continua. Para estudiar el efecto de la participación de trabajadores en la toma de decisiones administrativas, cien trabajadores fueron entrevistados en cada una de dos plantas manufactureras alemanas. Una de ellas tenía participación activa de trabajadores en la toma de decisiones administrativas; la otra, no. A cada trabajador seleccionado se le preguntó si en general aprobaba las decisiones administrativas tomadas dentro de la firma. Los resultados de las entrevistas se muestran en la tabla siguiente:

	Participación	No participación
Generalmente aprueban	73	51
No aprueban	27	49

- a. ¿Los datos dan evidencia suficiente para indicar que la aprobación o desaprobación de decisiones administrativas depende de si los trabajadores participan en la toma de decisiones? Pruebe usando el estadístico de prueba χ^2 . Use $\alpha = .05$.
- b. ¿Estos datos apoyan la hipótesis de que trabajadores de una firma, con toma de decisiones participativa, aprueban de manera más general las decisiones administrativas de la firma que los empleados sin toma de decisiones participativas? Pruebe usando la prueba z . Este problema requiere una prueba de una cola. ¿Por qué?

10.42 Tres entradas Un estudio de tránsito de ocupantes se realizó para ayudar en la remodelación de un edificio de oficinas que contiene tres entradas. La selección de entrada se registró para una muestra de 200 personas que entraron al edificio. ¿Los datos de la tabla indican que hay una diferencia en preferencia para las tres entradas? Encuentre un intervalo de confianza de 95% para la proporción de personas que están a favor de la entrada 1.

Entrada	1	2	3
Número que entran	83	51	56



10.43 Profesores en prácticas graduados Las responsabilidades de los estudiantes graduados con frecuencia se relacionan con sus funciones como profesores en prácticas o asistentes de investigación. Como parte de un estudio más grande, K.M. McGoldrick y sus colegas investigaron el nivel de preparación de estudiantes graduados de economía para sus deberes relacionados con la enseñanza para estudiantes en escuelas “de excelencia” y aquellos en escuelas “de segundo nivel”.¹⁰ A continuación se encuentran las respuestas a la pregunta “¿Está usted satisfecho con el nivel de preparación que tuvo para sus deberes relacionados con su año de enseñanza?”.

	De excelencia	De segundo nivel
Estoy muy satisfecho	85	197
Estoy algo satisfecho	102	171
Estoy insatisfecho	22	29
Total	209	397

- a. ¿Hay una diferencia significativa en las respuestas a la pregunta entre estudiantes de escuelas “de excelencia” en comparación con aquellos en escuelas “de segundo nivel”?
- b. Si es significativo, describa la naturaleza de las diferencias en la respuesta para estudiantes graduados en escuelas “de excelencia” contra escuelas “de segundo nivel”.

DATOS
EX1044 **10.44 ¿Sus alimentos son seguros?** ¿Cuánta confianza tiene de que es seguro comer los alimentos que compra? Esta pregunta se hizo en una encuesta de *CBS News*.¹¹ Los datos que siguen reflejan los resultados de las respuestas a esta encuesta.

	Mucha confianza	Alguna confianza	No demasiada confianza	Ninguna confianza en absoluto	Total
Hombres	210	241	68	5	524
Mujeres	129	306	73	16	524
Total	329	547	141	21	1048

- a. ¿Hay suficiente evidencia para concluir que hay diferencias significativas en las respuestas entre hombres y mujeres en el nivel de significancia de $\alpha = .05$?
- b. Encuentre el valor p aproximado para la prueba.

10.45 Colores de vehículos Cada modelo parece introducir nuevos colores y tonos diferentes para una amplia variedad de vehículos, desde automóviles de lujo, autos grandes o modelos intermedios, hasta compactos y autos deportivos, así como para camiones ligeros. No obstante, el color blanco y el plateado/gris siguen estando entre los cinco o seis colores principales en todas estas categorías de automóviles. Los seis colores principales y el porcentaje de su participación en el mercado para autos compactos/deportivos se muestran en la tabla siguiente.¹²

Color	Plateado	Negro	Gris	Azul	Rojo	Blanco
Porcentaje	19	17	17	15	12	12

Para verificar las cifras se tomó una muestra aleatoria formada por 250 autos compactos/deportivos y se registró el color de los vehículos. La muestra dio las siguientes cantidades para las categorías dadas antes: 52, 43, 48, 41, 32 y 19, respectivamente.

- a. ¿Falta alguna categoría en la clasificación? ¿Cuántos vehículos pertenecen a esa categoría?
- b. ¿Hay suficiente evidencia para indicar que nuestros porcentajes de los colores de autos compactos/

deportivos difieren de los dados? Encuentre el valor p aproximado para la prueba.

DATOS
EX1046 **10.46 Colores de vehículos, de nuevo** Consulte el ejercicio 10.45. El investigador desea ver si hay una diferencia en las distribuciones de los colores para autos compactos/deportivos contra autos grandes/medianos.¹² Se tomó otra muestra aleatoria de 250 autos grandes/medianos y se registraron los colores de los vehículos. La tabla siguiente muestra los resultados tanto para autos compactos/deportivos como para grandes/medianos.

Color	Plateado	Negro	Gris	Azul	Rojo	Blanco
Compacto/deportivo	52	43	48	41	32	19
Grande/mediano	50	33	37	32	27	38

¿Los datos indican que hay una diferencia en las distribuciones de los colores dependiendo del tipo de vehículo? Use $\alpha = .05$. (SUGERENCIA: Recuerde incluir una columna llamada “Otro” para autos que no caen en ninguna de las seis categorías mostradas en la tabla.)

DATOS
EX1047 **10.47 Medicina que sabe bien** Pfizer Canada Inc. es una empresa farmacéutica que hace acitromicina, antibiótico en suspensión con sabor a cereza, que se usa para tratar infecciones bacterianas en niños. Para comparar el sabor de su producto con tres medicamentos de la competencia, Pfizer probó 50 niños sanos y 20 adultos sanos. Entre otras medidas para probar el sabor, registraron el número de probadores que calificaron cada una de las cuatro suspensiones de antibiótico como el de mejor sabor.¹³ Los resultados se muestran en la tabla siguiente. ¿Hay diferencia en la percepción del mejor sabor entre adultos y niños? Si es así, ¿cuál es la naturaleza de la diferencia y por qué es de importancia práctica para la compañía farmacéutica?

	Sabor de antibiótico			
	Plátano	Cereza*	Fruta silvestre	Fresa-Plátano
Niños	14	20	7	9
Adultos	4	14	0	2

* Acitromicina producida por Pfizer Canada Inc.

DATOS
EX1048 **10.48 Lesiones en rugby** La prevalencia y patrones de lesiones en rodillas entre mujeres universitarias jugadoras de rugby se investigaron usando un cuestionario de muestra, al que respondieron 42 clubes de rugby.¹⁴ Un total de 76 lesiones de rodilla fueron clasificadas por tipo y posición de la jugadora (delantera o defensa).

	Tipo de lesión de rodilla				
Posición	Desgarre de menisco	Desgarre MCL	Desgarre ACL	Dislocación de rótula	Desgarre PCL
Delantera	13	14	7	3	1
Defensa	12	9	14	2	1

Salida MINITAB para el ejercicio 10.48

Prueba ji cuadrada: desgarrar de menisco, desgarrar MCL, desgarrar ACL, dislocación de rótula, desgarrar PCL

Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

	Men	Tear	MCL	Tear	ACL	Tear	Patella	PCL	Tear	Total
1	13	14	7	3	1					38
	12.50	11.50	10.50	2.50	1.00					
	0.020	0.543	1.167	0.100	0.000					
2	12	9	14	2	1					38
	12.50	11.50	10.50	2.50	1.00					
	0.020	0.543	1.167	0.100	0.000					
Total	25	23	21	5	2					76

Chi-Sq = 3.660, DF = 4, P-Value = 0.454
4 cells with expected counts less than 5.0

- Use la salida impresa MINITAB para determinar si hay diferencia en la distribución de tipos de lesiones para delanteras y defensas de rugby. ¿Ha sido violada alguna suposición necesaria para la prueba ji cuadrada? ¿Qué efecto tendrá esto en la magnitud del estadístico de prueba?
- Los investigadores informan de una diferencia significativa en la proporción de los desgarrar del MCL para las dos posiciones ($P < .05$) y una diferencia significativa en la proporción de desgarrar del ACL ($P < .05$), pero indican que todas las otras lesiones ocurren con igual frecuencia para las dos posiciones. ¿Está de acuerdo con esas conclusiones? Explique.

10.49 Pescar un resfriado ¿La probabilidad de contraer un resfriado está influida por el número de contactos sociales que usted tenga? Un estudio de Sheldon Cohen, profesor de psicología de la Universidad Carnegie Mellon, parece mostrar que cuantas más relaciones sociales tenga alguien *menos susceptible* es a los resfriados.¹⁵ Un grupo de 276 hombres y mujeres sanos se agruparon de acuerdo con su número de relaciones (padres, amigos, miembros de la iglesia, vecinos). Se les expone entonces a un virus que causa los resfriados. Una adaptación de los resultados se muestra en la tabla.

	Número de relaciones		
	Tres o menos	Cuatro o cinco	Seis o más
Resfriado	49	57	34
No resfriado	31	43	62
Total	80	100	96

- ¿Los datos dan suficiente evidencia para indicar que la susceptibilidad a los resfriados es afectada por el número de relaciones que se tenga? Pruebe al nivel de significancia de 5%.
- Con base en los resultados del inciso a), describa la naturaleza de la relación entre las dos variables

categorías: incidencia de resfriado y número de relaciones sociales. ¿Sus observaciones concuerdan con las conclusiones del autor?

10.50 Se realiza una encuesta con un grupo de 100 pacientes con cáncer y 100 personas sanas, para conocer su conducta frente a la bebida alcohólica.

Ingestión de alcohol	Con cáncer	Sin cáncer
Abstemio	37	39
Moderada	34	30
Excesiva	15	21
Adicción	14	10
Total	100	100

Aplique la prueba de χ^2 y concluya.

10.51 En un estudio donde participaron 8 726 personas, entre las que había 6 037 sanas, 883 con un cáncer incipiente en el estómago y 1 796 con úlcera duodenal, se aplicó una clasificación a partir del tipo de sangre: O, A y B.

Tipo de sangre	GRUPOS			
	Úlcera	Cáncer	Sanos	Total
O	983	393	2 842	4 758
A	679	416	2 625	3 720
B	134	84	570	788
Total	1 796	893	6 037	8 726

Aplique la prueba de independencia.

10.52 Calcule la ji cuadrada si se tienen 2 789 personas, de las cuales 1 058 fuman y 1 731 no fuman al analizar sus historias clínicas 472 fallecieron y 2 317 siguen vivos.

	Fallecieron	Vivos	Total
No fumadores	264	1 467	1 731
Fumadores	208	850	1 058
Total	472	2 317	2 789

- Utilizando la corrección de Yates.
- Sin utilizar la corrección de Yates.

10.53 En un estudio realizado a 3 000 mujeres, de las cuales 300 presentan cáncer de mama se requiere conocer si está relacionado con la ingesta de alcohol.

		Consumo de alcohol			
		Leve	Moderado	Severo	Total
Abstemio	Sí	75	40	60	125
Cáncer de mama	No	2 000	280	120	300
		2 075	320	180	425

10.54 Se obtienen datos de los expedientes de un hospital de 300 personas, donde 100 forman el grupo testigo y 200 el de estudio (tratamiento), dos años después del tratamiento.

	Muertos	Vivos	Total
Tratamiento	10	190	200
Testigo	15	85	100
	25	275	300

Calcule χ^2 .

Estadística no paramétrica

OBJETIVO GENERAL

En los capítulos 5 a 7 presentamos técnicas estadísticas para comparar dos poblaciones contrastando sus respectivos parámetros poblacionales (por lo general, sus medias poblacionales). Las técnicas del capítulo 5 son aplicables a los datos que son al menos cuantitativos y las del capítulo 6 son aplicables a los datos que tienen distribuciones normales. El propósito de este capítulo es presentar varias pruebas estadísticas útiles en la comparación de poblaciones para los numerosos tipos de datos que no satisfagan las suposiciones especificadas en los capítulos 5 a 7.

ÍNDICE DEL CAPÍTULO

- Introducción (11.1)
- La prueba de suma de rango de Wilcoxon: muestras aleatorias independientes (11.2)
- La prueba del signo para un experimento de dos poblaciones (11.3)
- Una comparación de pruebas estadísticas (11.4)
- La prueba de rango con signo de Wilcoxon para un experimento de dos poblaciones (11.5)
- La prueba H de Kruskal-Wallis para diseños completamente aleatorizados (11.6)
- La prueba F_r de Friedman para diseños de bloque aleatorizados (11.7)
- Prueba de Nemenyi (11.8)
- Prueba de la mediana (11.9)
- Coeficiente de correlación de rango (11.10)
- Prueba de significancia de r_s (11.11)
- Coeficiente tau (τ) de Kendall (11.12)
- Coeficiente de concordancia (ω) de Kendall (11.13)
- Coeficiente de correlación (r_{bp}) biserial de punto (11.14)
- Prueba de Kappa (11.15)
- Resumen (11.16)

¿Cómo está su nivel de colesterol?

¿Cuál es su nivel de colesterol? En los últimos años, muchos de nosotros nos hemos hecho más conscientes de nuestra salud cuando leemos las etiquetas de información nutrimental en productos alimenticios que compramos y elegimos alimentos que sean bajos en grasas y colesterol y altos en fibras. El caso práctico del final de este capítulo contiene un experimento de degustación, para comparar tres tipos de sustitutos de huevo, usando técnicas no paramétricas.

11.1

INTRODUCCIÓN

Algunos experimentos generan respuestas que pueden ser ordenadas o clasificadas, pero no es posible medir numéricamente el valor real de la respuesta excepto con una escala arbitraria que usted cree. Es probable que usted sólo pueda decir si una observación es mayor que otra. Quizá sea capaz de clasificar todo un conjunto de observaciones sin saber en realidad los valores numéricos exactos de las mediciones. A continuación revisemos algunos ejemplos:

- Las habilidades de ventas de cuatro vendedores son clasificadas de la mejor a la peor.
- Las características de comestible y sabor de cinco marcas de fibra de pasitas se clasifican en una escala arbitraria de 1 a 5.
- Cinco diseños de automóviles se clasifican del más atractivo al menos atractivo.

¿Cómo se analizan estos tipos de datos? Los métodos estadísticos de muestra pequeña presentados en los capítulos 8 a 10 son válidos sólo cuando la(s) población(es) muestreada(s) es (son) normal(es) o aproximadamente lo es (son). Los datos formados por rangos o escalas arbitrarias de 1 a 5 *no satisfacen la suposición de normalidad*, incluso a un grado razonable. En algunas aplicaciones, las técnicas son válidas si las muestras se toman al azar de poblaciones cuyas varianzas son iguales.

Cuando los datos no parecen satisfacer éstas y suposiciones similares, es posible usar un método alternativo, **métodos estadísticos no paramétricos**. Los métodos no paramétricos por lo general especifican las hipótesis en términos de distribuciones poblacionales más que parámetros como medias y desviaciones estándar. Es frecuente que las suposiciones paramétricas sean sustituidas por suposiciones más generales acerca de las distribuciones poblacionales y las clasificaciones de las observaciones se usen a veces en lugar de las mediciones reales.

Investigaciones realizadas demuestran que las pruebas estadísticas no paramétricas son tan capaces de detectar diferencias entre poblaciones como los métodos paramétricos de capítulos anteriores, cuando se satisfacen la normalidad y otras suposiciones. Pueden ser, y con frecuencia son, más potentes para detectar diferencias poblacionales cuando estas suposiciones no se satisfacen. Por esta razón, algunos estadísticos están a favor de usar procedimientos no paramétricos en lugar de sus similares paramétricos.

Presentaremos métodos no paramétricos apropiados para comparar dos o más poblaciones usando ya sea muestras independientes o en pares. También presentaremos una medida de asociación que es útil para determinar si una variable aumenta cuando la otra aumenta o si una variable disminuye cuando la otra aumenta.

TIP MI CONSEJO

Cuando los tamaños muestrales sean pequeños y las poblaciones originales no sean normales, use técnicas no paramétricas

LA PRUEBA DE SUMA DE RANGO DE WILCOXON: MUESTRAS ALEATORIAS INDEPENDIENTES

11.2

Al comparar las medias de dos poblaciones basadas en muestras independientes, el estadístico pivote era la diferencia en las medias muestrales. Si no está seguro de que las suposiciones requeridas para una prueba t de dos muestras sea satisfecha, una alternativa es sustituir los valores de las observaciones por sus rangos y proceder como si éstos fueran las observaciones reales. Dos pruebas no paramétricas diferentes usan un estadístico de prueba basado en estos rangos de muestra:

- Prueba de suma de rango de Wilcoxon
- Prueba U de Mann-Whitney

Son *equivalentes* porque usan la misma información muestral. El procedimiento que presentaremos es la prueba de suma de rango de Wilcoxon, propuesta por Frank Wilcoxon, que está basada en la suma de los rangos de la muestra que tiene el tamaño muestral más pequeño.

Suponga que tenemos n_1 observaciones de la población 1 y n_2 observaciones de la población 2. La hipótesis nula a probar es que las dos distribuciones poblacionales son idénticas, contra la hipótesis alternativa de que las distribuciones poblacionales son diferentes. Éstas son las posibilidades para las dos poblaciones:

- Si H_0 es verdadera y las observaciones han provenido de las mismas o idénticas poblaciones, entonces las observaciones de ambas muestras deben mezclarse al azar cuando conjuntamente sean de rango de pequeño a grande. La suma de los rangos de las observaciones de la muestra 1 debe ser similar a la suma de los rangos de la muestra 2.
- Si, por el contrario, las observaciones de la población 1 tienden a ser más pequeñas que las de la población 2, entonces estas observaciones tendrían los rangos más pequeños porque la mayoría de éstas serían más pequeñas que las de la población 2. La suma de los rangos de ellas sería “pequeña”.
- Si las observaciones de la población 1 tienden a ser más grandes que las de la población 2, se les asignarían rangos más grandes. La suma de los rangos de estas últimas tendería a ser “grande”.

Por ejemplo, suponga que tenemos $n_1 = 3$ observaciones de la población 1, es decir, 2, 4 y 6, y $n_2 = 4$ observaciones de la población 2, o sea 3, 5, 8 y 9. La tabla 11.1 muestra siete observaciones ordenadas de pequeñas a grandes.

TABLA 11.1

Siete observaciones en orden

Observación	x_1	y_1	x_2	y_2	x_3	y_3	y_4
Datos	2	3	4	5	6	8	9
Rango	1	2	3	4	5	6	7

A la observación más pequeña, $x_1 = 2$, se le asigna el rango 1; a la siguiente observación más pequeña, $y_1 = 3$, se le asigna el rango 2; y así sucesivamente. La *suma de los rangos* de las observaciones de la muestra 1 es $1 + 3 + 5 = 9$ y la **suma de rangos** de la muestra 2 es $2 + 4 + 6 + 7 = 19$. ¿Cómo se determina si la suma de rangos de las observaciones de la muestra 1 es significativamente pequeña o significativamente grande? Esto depende de la distribución de probabilidad de la suma de rangos de una de las muestras. Como los rangos para $n_1 + n_2 = N$ observaciones son los primeros N enteros, se puede demostrar que la suma de estos rangos es $N(N + 1)/2$. En este sencillo ejemplo, la suma de $N = 7$ rangos es $1 + 2 + 3 + 4 + 5 + 6 + 7 = 7(8)/2$ o sea 28. En consecuencia, si el experimentador conoce la suma de rangos para una de las muestras, puede hallar la otra por sustracción. En nuestro ejemplo, observe que la suma de rangos para la muestra 1 es 9, en tanto que la segunda suma de rangos es $(28 - 9) = 19$. Esto significa que sólo una de las dos sumas de rangos es necesaria para la prueba. Para simplificar la tabulación de valores críticos para esta prueba debe usarse la suma de rangos de la muestra más pequeña como estadístico de prueba. ¿Qué ocurre si dos o más observaciones son iguales? A observaciones empatadas se les asigna el promedio de los rangos que tendrían las observaciones si hubieran sido ligeramente diferentes en valor.

Para poner en práctica la prueba de la suma de rangos de Wilcoxon, supongamos que muestras aleatorias independientes de tamaño n_1 y n_2 se seleccionan de las poblaciones 1 y 2, respectivamente. Representemos con n_1 al menor de los dos tamaños muestrales y

con T_1 la suma de los rangos de las observaciones de la muestra 1. Si la población 1 está a la izquierda de la población 2, T_1 será “pequeña”. T_1 será “grande” si la población 1 está a la derecha de la 2.

FÓRMULAS PARA EL ESTADÍSTICO DE LA SUMA DE RANGOS DE WILCOXON (PARA MUESTRAS INDEPENDIENTES)

Sean

$$T_1 = \text{Suma de rangos para la primera muestra}$$
$$T_1^* = n_1(n_1 + n_2 + 1) - T_1$$

T_1^* es el valor de la suma de rangos para n_1 si las observaciones se hubieran ordenado de *grande a pequeña*. (No es la suma de rangos para la segunda muestra.) Dependiendo de la naturaleza de la hipótesis alternativa, uno de estos dos valores se elegirá como estadístico de prueba, T .

La tabla 7 del Anexo se puede usar con el fin de localizar *valores críticos* para el estadístico de prueba para cuatro valores diferentes de pruebas de una cola con $\alpha = .05, .025, .01$ y $.005$. Para usar la tabla 7 del Anexo para una prueba de dos colas, los valores de α se duplican, es decir, $\alpha = .10, .05, .02$ y $.01$. La entrada de la tabla da el valor de a tal que $P(T \leq a) \leq \alpha$. Con la intención de ver cómo localizar un valor crítico para la prueba de suma de rango de Wilcoxon, suponga que $n_1 = 8$ y $n_2 = 10$ para una prueba de una cola con $\alpha = .05$. Se puede usar la tabla 7a) del Anexo, una parte de la cual se reproduce en la tabla 11.2. Observe que la tabla está construida suponiendo que $n_1 \leq n_2$. Es por esta razón que designamos la población con el tamaño muestral más pequeño como población 1. Los valores de n_1 se muestran en sentido horizontal en la parte superior de la tabla y los de n_2 se muestran en sentido vertical al lado izquierdo. La entrada $a = 56$, sombreada, es el valor crítico para rechazar H_0 . La hipótesis nula de igualdad de las dos distribuciones debe ser rechazada si el valor observado del estadístico de prueba T es menor o igual a 56.

TABLA 11.2 Parte de los valores críticos de cola izquierda al 5%, tabla 7 del Anexo

	n_1						
n_2	2	3	4	5	6	7	8
3	—	6					
4	—	6	11				
5	3	7	12	19			
6	3	8	13	20	28		
7	3	8	14	21	29	39	
8	4	9	15	23	31	41	51
9	4	10	16	24	33	43	54
10	4	10	17	26	35	45	56

PRUEBA DE LA SUMA DE RANGO DE WILCOXON

Con n_1 denotemos la más pequeña de las dos muestras. Esta muestra proviene de la población 1. Las hipótesis a probar son

H_0 : Las distribuciones para las poblaciones 1 y 2 son idénticas
contra una de tres hipótesis alternativas:

- H_a : Las distribuciones para las poblaciones 1 y 2 son diferentes (una prueba de dos colas)
 H_a : La distribución para la población 1 está a la izquierda de aquella de la población 2 (una prueba de cola izquierda)
 H_a : La distribución para la población 1 está a la derecha de aquella de la población 2 (una prueba de cola derecha)

Procedimiento

1. Ordene todas las $n_1 + n_2$ observaciones de pequeña a grande.
2. Encuentre T_1 , la suma de rangos para las observaciones de la muestra 1. Éste es el estadístico de prueba para la prueba de cola izquierda.
3. Encuentre $T_1^* = n_1(n_1 + n_2 + 1) - T_1$, la suma de los rangos de las observaciones de la población 1 si los rangos asignados se hubieran invertido de grandes a pequeños. (El valor de T_1^* no es la suma de los rangos de las observaciones de la muestra 2.) Éste es el estadístico de prueba para una prueba de cola derecha.
4. El estadístico de prueba para una prueba de dos colas es T , la mínima de T_1 y T_1^* .
5. H_0 es rechazada si el estadístico de prueba observado es menor o igual al valor crítico hallado usando la tabla 7 del Anexo.

Ilustramos el uso de la tabla 7 con el siguiente ejemplo.

EJEMPLO

11.1

Las frecuencias de aleteo de dos especies de abejas Euglossine fueron registradas para una muestra de $n_1 = 4$ *Euglossa mandibularis* Friese (especie 1) y $n_2 = 6$ *Euglossa imperialis* Cockerell (especie 2).¹ Las frecuencias se detallan en la tabla 11.3. ¿Puede usted concluir que las distribuciones de aleteo difieren para estas dos especies? Pruebe usando $\alpha = .05$.

TABLA 11.3

Frecuencias de aleteo para dos especies de abejas

Especie 1	Especie 2
235	180
225	169
190	180
188	185
	178
	182

Solución Primero es necesario ordenar las observaciones de pequeña a grande, como se muestra en la tabla 11.4.

TABLA 11.4

Frecuencias de aleteo ordenadas de pequeña a grande

Datos	Especie	Orden
169	2	1
178	2	2
180	2	3
180	2	4
182	2	5
185	2	6
188	1	7
190	1	8
225	1	9
235	1	10

Las hipótesis a probar son

H_0 : Las distribuciones de las frecuencias de aleteo son las mismas para las dos especies
contra

H_a : Las distribuciones de las frecuencias de aleteo difieren para las dos especies

Como el tamaño muestral para individuos de la especie 1, $n_1 = 4$, es el más pequeño de los dos tamaños muestrales, tenemos

$$T_1 = 7 + 8 + 9 + 10 = 34$$

y

$$T_1^* = n_1(n_1 + n_2 + 1) - T_1 = 4(4 + 6 + 1) - 34 = 10$$

Para una prueba de dos colas, el estadístico de prueba es $T = 10$, la menor de $T_1 = 34$ y $T_1^* = 10$.

Para esta prueba de dos colas con $\alpha = .05$, puede usar la tabla 7b) del Anexo con $n_1 = 4$ y $n_2 = 6$. El valor crítico de T tal que $P(T \leq a) \leq \alpha/2 = .025$ es 12, y se debe rechazar la hipótesis nula si el valor observado de T es 12 o menos. Como el valor observado del estadístico de prueba, $T = 10$, es menor que 12, se puede rechazar la hipótesis de distribuciones de frecuencias de aleteo iguales al nivel de significancia de 5%.

Una salida impresa *MINITAB* de la prueba de suma de rango de Wilcoxon (llamada Mann-Whitney por *MINITAB*) para estos datos se da en la figura 11.1. Al final de este capítulo se encuentran instrucciones para generar esta salida impresa en la sección “Tecnología actual”. Observe que la suma de rango de la primera muestra se da como $W = 34.0$, que concuerda con nuestros cálculos. Con un valor p reportado de .0142 calculado por *MINITAB*, se puede rechazar la hipótesis nula al nivel de 5%.

FIGURA 11.1

Salida impresa para el ejemplo 11.1

● Prueba Mann-Whitney y CI: especie 1, especie 2

	N	Median
Species 1	4	207.50
Species 2	6	180.00

Point estimate for ETA1-ETA2 is 30.50
95.7 Percent CI for ETA1-ETA2 is (5.99,56.01)
W = 34.0
Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.0142
The test is significant at 0.0139 (adjusted for ties)

Aproximación normal para la prueba de suma de rango de Wilcoxon

La tabla 7 del Anexo contiene valores críticos para tamaños muestrales de $n_1 \leq n_2 = 3, 4, \dots, 15$. Siempre que n_1 no sea demasiado pequeña,[†] las aproximaciones a las probabilidades para el estadístico T de la suma de rango de Wilcoxon se pueden hallar usando una aproximación normal a la distribución de T . Se puede demostrar que la media y la varianza de T son

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2} \quad \text{y} \quad \sigma_T^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

[†] Algunos investigadores indican que la aproximación normal es adecuada para muestras de hasta sólo $n_1 = n_2 = 4$.

La distribución de

$$z = \frac{T - \mu_T}{\sigma_T}$$

es aproximadamente normal con media 0 y desviación estándar 1 para valores de n_1 y n_2 de sólo 10.

Si se intenta esta aproximación para el ejemplo 11.1, se obtiene

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2} = \frac{4(4 + 6 + 1)}{2} = 22$$

y

$$\sigma_T^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12} = \frac{4(6)(4 + 6 + 1)}{12} = 22$$

El valor p para esta prueba es $2P(T \geq 34)$. Si se usa una corrección de .5 para continuidad al calcular el valor de z porque n_1 y n_2 son pequeñas,[†] tenemos

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{(34 - .5) - 22}{\sqrt{22}} = 2.45$$

El valor p para esta prueba es

$$2P(T \geq 34) \approx 2P(z \geq 2.45) = 2(.0071) = .0142$$

el valor reportado en la salida impresa *MINITAB* de la figura 11.1.

PRUEBA DE LA SUMA DE RANGO DE WILCOXON PARA MUESTRAS GRANDES: $n_1 \geq 10$ Y $n_2 \geq 10$

1. Hipótesis nula: H_0 : Las distribuciones de las poblaciones son idénticas.
2. Hipótesis alternativa: H_a : Las dos distribuciones poblacionales no son idénticas (una prueba de dos colas); o H_a : La distribución de la población 1 se corre a la derecha (o a la izquierda) de la distribución de la población 2 (una prueba de una cola).
3. Estadístico de prueba: $z = \frac{T - n_1(n_1 + n_2 + 1)/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}}$
4. Región de rechazo:
 - a. Para una prueba de dos colas, rechace H_0 si $z > z_{\alpha/2}$ o $z < -z_{\alpha/2}$.
 - b. Para una prueba de una cola en la cola derecha, rechace H_0 si $z > z_{\alpha}$.
 - c. Para una prueba de una cola en la cola izquierda, rechace H_0 si $z < -z_{\alpha}$.

O rechace H_0 si el valor p es $< \alpha$.

Los valores tabulados de z se encuentran en la tabla 3 del Anexo.

EJEMPLO

11.2

Se realizó un experimento para comparar las resistencias de dos tipos de papel de estraza: uno es un papel de estraza estándar de un peso especificado y el otro es el mismo papel tratado con una sustancia química. Diez piezas de cada tipo de papel, seleccionadas al azar de la producción, produjeron las mediciones de resistencia que se muestran en la tabla 11.5. Pruebe la hipótesis nula de no diferencia en las distribuciones de resistencias para los dos tipos de papel

[†] Como el valor de $T = 34$ está a la derecha de la media 22, la resta de .5 al usar la aproximación normal toma en cuenta el límite inferior de la barra arriba del valor 34 en la distribución de probabilidad de T .

contra la hipótesis alternativa de que el papel tratado tiende a ser más fuerte (es decir, su distribución de mediciones de resistencia se corre a la derecha de la distribución correspondiente para el papel no tratado).

TABLA 11.5 Mediciones de resistencia (y sus rangos) para dos tipos de papel

Estándar 1	Tratado 2
1.21 (2)	1.49 (15)
1.43 (12)	1.37 (7.5)
1.35 (6)	1.67 (20)
1.51 (17)	1.50 (16)
1.39 (9)	1.31 (5)
1.17 (1)	1.29 (3.5)
1.48 (14)	1.52 (18)
1.42 (11)	1.37 (7.5)
1.29 (3.5)	1.44 (13)
1.40 (10)	1.53 (19)
Suma de rango	$T_1 = 85.5$
	$T_1^* = n_1(n_1 + n_2 + 1) - T_1 = 210 - 85.5 = 124.5$

Solución Como los tamaños muestrales son iguales, estamos en libertad de decidir cuál de las dos muestras debe ser la muestra 1. Si se elige el tratamiento estándar como la primera muestra, se pueden clasificar 20 mediciones y los valores de T_1 y T_1^* se ven en la parte inferior de la tabla. Como se desea detectar un corrimiento en las mediciones estándar (1) a la izquierda de las mediciones tratadas (2), se realiza una prueba de cola izquierda:

H_0 : No hay diferencia en las distribuciones de resistencia

H_a : La distribución estándar se encuentra a la izquierda de la distribución tratada

y se usa $T = T_1$ como el estadístico de prueba, buscando un valor inusualmente pequeño de T .

Para hallar el valor crítico para una prueba de una cola con $\alpha = .05$, indicemos de la tabla 7a) del Anexo con $n_1 = n_2 = 10$. Usando la entrada de la tabla, se puede rechazar H_0 cuando $T \leq 82$. Como el valor observado del estadístico de prueba es $T = 85.5$, no se puede rechazar H_0 . Hay insuficiente evidencia para concluir que el papel de estraza tratado es más fuerte que el papel estándar.

Para usar la aproximación normal a la distribución de T , se calcula

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2} = \frac{10(21)}{2} = 105$$

y

$$\sigma_T^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12} = \frac{10(10)(21)}{12} = 175$$

con $\sigma_T = \sqrt{175} = 13.23$. Entonces

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{85.5 - 105}{13.23} = -1.47$$

El valor p de una cola correspondiente a $z = -1.47$ es

$$\text{valor } p = P(z \leq -1.47) = .5 - .4292 = .0708$$

que es mayor que $\alpha = .05$. La conclusión es la misma. No se puede concluir que el papel de estraza tratado sea más fuerte que el papel estándar.

¿Cuándo se puede usar la prueba de la suma de rango de Wilcoxon en lugar de la prueba t no pareada de dos muestras? La prueba t de dos muestras funciona bien si los datos están normalmente distribuidos con varianzas iguales. Si hay duda respecto a estas suposiciones, se usa una gráfica de probabilidad normal para evaluar el grado de no normalidad y también una prueba F de dos muestras de varianzas muestrales para verificar la igualdad de varianzas. Si estos procedimientos indican ya sea no normalidad o desigualdad de varianza, entonces es apropiada la prueba de la suma de rango de Wilcoxon.

11.2 EJERCICIOS

TÉCNICAS BÁSICAS

11.1 Suponga que se desea usar la prueba de la suma de rangos de Wilcoxon, para detectar un corrimiento en la distribución 1 a la derecha de la distribución 2, con base en muestras de tamaño $n_1 = 6$ y $n_2 = 8$.

- ¿Se debe usar T_1 o T_1^* como el estadístico de prueba?
- ¿Cuál es la región de rechazo para la prueba si $\alpha = .05$?
- ¿Cuál es la región de rechazo para la prueba si $\alpha = .01$?

11.2 Consulte el ejercicio 11.1. Suponga que la hipótesis alternativa es que la distribución 1 se corre ya sea a la izquierda o a la derecha de la distribución 2.

- ¿Se debe usar T_1 o T_1^* como el estadístico de prueba?
- ¿Cuál es la región de rechazo para la prueba si $\alpha = .05$?
- ¿Cuál es la región de rechazo para la prueba si $\alpha = .01$?

11.3 Observaciones de dos muestras aleatorias e independientes, tomadas de las poblaciones 1 y 2, se dan aquí. Use la prueba de la suma de rango de Wilcoxon para determinar si la población 1 está corrida a la izquierda de la población 2.

Muestra 1	1	3	2	3	5
Muestra 2	4	7	6	8	6

- Expresar las hipótesis nula y alternativa a probar.
- Ordene la muestra combinada de menor a mayor. Calcule T_1 y T_1^* .
- ¿Cuál es la región de rechazo para $\alpha = .05$?
- ¿Los datos dan suficiente evidencia para indicar que la población 1 está corrida a la izquierda de la población 2?

11.4 Muestras aleatorias independientes de tamaño $n_1 = 20$ y $n_2 = 25$ se toman de las poblaciones 1 y 2 no normales. La muestra combinada está ordenada y $T_1 = 252$. Use la aproximación de muestra grande a la prueba de la suma de rango de Wilcoxon para determinar si hay una diferencia en las dos distribuciones de población. Calcule el valor p para la prueba.

11.5 Suponga que desea detectar un corrimiento en la distribución 1 a la derecha de la distribución 2, con base en

tamaños muestrales $n_1 = 12$ y $n_2 = 14$. Si $T_1 = 193$, ¿qué se concluye? Use $\alpha = .05$.

APLICACIONES

11.6 Enfermedad de Alzheimer En algunas pruebas de ancianos sanos un nuevo medicamento ha restaurado su memoria casi como la de los jóvenes. Pronto se probará en pacientes con enfermedad de Alzheimer, esa fatal enfermedad del cerebro que destruye la mente. Según el doctor Gary Lynch, de la Universidad de California en Irvine, el medicamento, llamado ampakina CX-516, acelera las señales entre las células cerebrales que parecen agudizar significativamente la memoria.² En una prueba preliminar en estudiantes de poco más de 20 años y en hombres de entre 65 y 70 años de edad, los resultados fueron particularmente sorprendentes. Después de recibir dosis moderadas de este medicamento, las personas de entre 65 y 70 años de edad calificaron casi tan alto como los jóvenes. Los datos siguientes son los números de sílabas sin sentido recordadas después de 5 minutos, para 10 hombres de poco más de 20 años de edad y 10 señores de entre 65 y 70 años. Use la prueba de la suma de rango de Wilcoxon para determinar si las distribuciones para el número de sílabas sin sentido recordadas son iguales para estos dos grupos.

20	3	6	4	8	7	1	1	2	7	8
65-70	1	0	4	1	2	5	0	2	2	3

11.7 Alzheimer, continúa Consulte el ejercicio 11.6. Suponga que dos grupos más de 10 hombres cada uno son probados sobre el número de sílabas sin sentido que podían recordar después de 5 minutos. No obstante, esta vez a los señores de entre 65 y 70 se les da una dosis moderada de ampakina CX-516. ¿Los datos dan suficiente evidencia para concluir que este medicamento mejora la memoria en los pacientes entre 65 y 70 años de edad en comparación con los de 20 años? Use un nivel apropiado de α .

20	11	7	6	8	6	9	2	10	3	6
65-70	1	9	6	8	7	8	5	7	10	3

11.8 Contenido de oxígeno disuelto Las observaciones de la tabla son el contenido de oxígeno disuelto en agua. Cuanto más alto el contenido de oxígeno disuelto, mayor es la capacidad de un río, lago o arroyo para sostener fauna acuática. En este experimento, un inspector de control de contaminación sospechaba que una comunidad ribereña descargaba aguas negras tratadas parcialmente hacia un río. Para comprobar esta teoría, cinco especímenes de agua de río, tomados al azar, se seleccionaron en un lugar aguas arriba del pueblo y, otras cinco, aguas abajo. Éstas son las lecturas de oxígeno disuelto (en partes por millón):

Aguas arriba	4.8	5.2	5.0	4.9	5.1
Aguas abajo	5.0	4.7	4.9	4.8	4.9

- a. Use una prueba de suma de rango de Wilcoxon de una cola con $\alpha = .05$ para confirmar o refutar la teoría.
- b. Use una prueba t de Student (con $\alpha = .05$) para analizar los datos. Compare la conclusión alcanzada en la parte a.

DATOS
EX1109

11.9 Movimiento de ojos En una investigación del comportamiento de exploración visual de niños sordos se tomaron medidas del movimiento de ojos en nueve niños sordos y nueve que sí escuchaban. La tabla siguiente da la rapidez de movimiento de ojos y sus rangos (entre paréntesis). ¿Le parece que difieren las distribuciones de la rapidez de movimiento de ojos de niños sordos y de niños que sí escuchan?

	Niños sordos	Niños que sí escuchan
	2.75 (15)	.89 (1)
	2.14 (11)	1.43 (7)
	3.23 (18)	1.06 (4)
	2.07 (10)	1.01 (3)
	2.49 (14)	.94 (2)
	2.18 (12)	1.79 (8)
	3.16 (17)	1.12 (5.5)
	2.93 (16)	2.01 (9)
	2.20 (13)	1.12 (5.5)
Suma de rango	126	45

DATOS
EX1110

11.10 Pesos de tortugas Los pesos de tortugas capturadas en dos lagos diferentes se midieron para comparar los efectos de los ambientes de los dos lagos en el crecimiento de las tortugas. Todas las tortugas eran de la misma edad y fueron marcadas antes de soltarlas en los lagos. A continuación veamos los pesos para $n_1 =$

10 tortugas marcadas y capturadas en el lago 1 y $n_2 = 8$ capturadas en el lago 2:

Lago	Pesos (onzas)									
1	14.1	15.2	13.9	14.5	14.7	13.8	14.0	16.1	12.7	15.3
2	12.2	13.0	14.1	13.6	12.4	11.9	12.5	13.8		

¿Los datos dan suficiente evidencia para indicar una diferencia en las distribuciones de peso para las tortugas marcadas, expuestas a los ambientes de los dos lagos? Use la prueba de la suma de rango de Wilcoxon con $\alpha = .05$ para contestar la pregunta.

DATOS
EX1111

11.11 Quimioterapia El tratamiento de cáncer por medios químicos, llamado quimioterapia, mata células cancerosas y células normales. En algunos casos, la toxicidad del medicamento para el cáncer, es decir, su efecto sobre células normales, puede reducirse con la inyección simultánea de un segundo medicamento. Se realizó un estudio para determinar si la inyección de un medicamento en particular reducía los efectos dañinos de un tratamiento de quimioterapia en el tiempo de sobrevivencia de ratas. Dos grupos de 12 ratas seleccionados al azar se emplearon en un experimento en el que ambos grupos, llamémoslos A y B, recibieron la droga tóxica en una dosis lo suficientemente grande para causarles la muerte, pero, además, el grupo B recibió la antitoxina que iba a reducir el efecto tóxico de la quimioterapia en células normales. La prueba finalizó al término de 20 días, o sea, 480 horas. Los tiempos de sobrevivencia para los dos grupos de ratas, a las 4 horas más cercanas, se muestran en la tabla siguiente. ¿Los datos dan suficiente evidencia para indicar que las ratas que recibieron la antitoxina tienden a sobrevivir más después de la quimioterapia que las que no recibieron la antitoxina? Use la prueba de la suma de rango de Wilcoxon con $\alpha = .05$.

Sólo quimioterapia	Quimioterapia más droga
A	B
84	140
128	184
168	368
92	96
184	480
92	188
76	480
104	244
72	440
180	380
144	480
120	196

11.3

LA PRUEBA DEL SIGNO PARA UN EXPERIMENTO DE DOS POBLACIONES

La prueba del signo es un procedimiento bastante sencillo que se usa para comparar dos poblaciones cuando las muestras consisten en observaciones en pares. Este tipo de diseño experimental se denomina diseño de **diferencia en pares** o **pares comparados**, que en la sección 6.5 se usó para comparar el promedio de desgaste para dos tipos de llantas. En general, para cada par se mide si la primera respuesta, por ejemplo A, excede a la segunda respuesta, por ejemplo a B. El estadístico de prueba es x , el número de veces que A excede a B en los n pares de observaciones.

Cuando las dos distribuciones poblacionales son idénticas, la probabilidad de que A exceda a B es igual a $p = .5$, y x , el número de veces que A excede a B, tiene una distribución binomial. Sólo pares sin empates se incluyen en la prueba. En consecuencia, se puede probar la hipótesis de distribuciones poblacionales idénticas al probar $H_0 : p = .5$ contra una alternativa ya sea de una o dos colas. Los valores críticos para la región de rechazo o valores p exactos se hallan usando las tablas binomiales acumulativas del Anexo.

LA PRUEBA DEL SIGNO PARA COMPARAR DOS POBLACIONES

1. Hipótesis nula: H_0 : Las dos distribuciones poblacionales son idénticas y $P(\text{A excede a B}) = p = .5$
2. Hipótesis alternativa:
 - a. H_a : Las distribuciones poblacionales no son idénticas y $p \neq .5$
 - b. H_a : La población de A mediciones se corre a la derecha de la población de B mediciones y $p > .5$
 - c. H_a : La población de A mediciones se corre a la izquierda de la población de B mediciones y $p < .5$
3. Estadístico de prueba: Para n , el número de pares sin empates, use x , el número de veces que $(A - B)$ es positivo.
4. Región de rechazo:
 - a. Para la prueba de dos colas $H_a : p \neq .5$, rechazar H_0 si $x \leq x_L$ o $x \geq x_U$, donde $P(x \leq x_L) \leq \alpha/2$ y $P(x \geq x_U) \leq \alpha/2$ para x que tenga una distribución binomial con $p = .5$.
 - b. Para $H_a : p > .5$, rechazar H_0 si $x \geq x_U$ con $P(x \geq x_U) \leq \alpha$.
 - c. Para $H_a : p < .5$, rechazar H_0 si $x \leq x_L$ con $P(x \leq x_L) < \alpha$.

O calcule el valor p y rechace H_0 si el valor $p < \alpha$.

Un problema que probablemente ocurra cuando se realiza una prueba del signo es que las mediciones asociadas con uno o más pares pueden ser iguales y, por tanto, resultan en **observaciones empatadas**. Cuando esto ocurra, elimine los pares empatados y reduzca n , el número total de pares. El ejemplo siguiente le ayudará a entender cómo se construye y utiliza la prueba del signo.

EJEMPLO

11.3

Los números de fusibles eléctricos defectuosos producidos por dos líneas de producción, A y B, se registraron a diario durante un periodo de 10 días, los resultados aparecen en la tabla 11.6. La variable de respuesta, el número de fusibles defectuosos, tiene una distribución binomial exacta con un gran número de fusibles producidos por día. Aun cuando esta variable

tendrá aproximadamente una distribución normal, el supervisor de planta preferiría una prueba estadística rápida y fácil para determinar si una línea de producción tiende a producir más fusibles defectuosos que la otra. Use la prueba del signo para probar la hipótesis apropiada.

TABLA 11.6

Fusibles defectuosos de dos líneas de producción

Día	Línea A	Línea B	Signo de diferencia
1	170	201	—
2	164	179	—
3	140	159	—
4	184	195	—
5	174	177	—
6	142	170	—
7	191	183	+
8	169	179	—
9	161	170	—
10	200	212	—

Solución Para este experimento de *diferencia en pares*, x es el número de veces que la observación de la línea A excede la de la línea B en un día determinado. Si no hay diferencia en las distribuciones de fusibles defectuosos para las dos líneas de producción, entonces p , la proporción de días en los que A excede a B, es .5, que es el valor hipotético en una prueba del parámetro binomial p . Valores muy pequeños o muy grandes de x , el número de veces que A excede de B, son contrarios a la hipótesis nula.

Como $n = 10$ y el valor hipotético de p es .5, la tabla 1 del Anexo se puede usar para hallar el valor p exacto para la prueba de

$$H_0 : p = .5 \text{ contra } H_a : p \neq .5$$

El valor observado del estadístico de prueba, que es el número de signos “más” en la tabla, es $x = 1$, y el valor p se calcula como

$$\text{valor } p = 2P(x \leq 1) = 2(.011) = .022$$

El valor p bastante pequeño = .022 permite rechazar H_0 al nivel de 5%. Hay evidencia significativa para indicar que el número de fusibles defectuosos no es el mismo para las dos líneas de producción; de hecho, la línea B produce más fusibles defectuosos que la A. En este ejemplo, la prueba del signo es una herramienta aproximada, fácil de calcular, para detectar líneas de producción con falla y funciona perfectamente bien para detectar una diferencia significativa usando sólo una cantidad mínima de información.

Aproximación normal para la prueba del signo

Cuando el número de pares n es grande, los valores críticos para el rechazo de H_0 y los valores p aproximados se pueden hallar usando una aproximación normal a la distribución de x . Debido a que la distribución binomial es perfectamente simétrica cuando $p = .5$, esta aproximación funciona muy bien, incluso para n de sólo 10.

Para $n \geq 25$, se puede efectuar la prueba del signo usando el estadístico z ,

$$z = \frac{x - np}{\sqrt{npq}} = \frac{x - .5n}{.5\sqrt{n}}$$

como el estadístico de prueba. Al usar z , se prueba la hipótesis nula $p = .5$ contra la alternativa $p \neq .5$ para una prueba de dos colas o contra la alternativa $p > .5$ (o $p < .5$) para una prueba de una cola.

PRUEBA DEL SIGNO PARA MUESTRAS GRANDES: $n \geq 25$

1. La hipótesis nula: $H_0 : p = .5$ (un tratamiento no se prefiere a un segundo tratamiento).
2. Hipótesis alternativa: $H_a : p \neq .5$, para una prueba de dos colas (NOTA: Usamos la prueba de dos colas como ejemplo. Muchos análisis podrían requerir una prueba de una cola.)
3. Estadístico de prueba: $z = \frac{x - .5n}{.5\sqrt{n}}$
4. Región de rechazo: Rechace H_0 si $z \geq z_{\alpha/2}$ o $z \leq -z_{\alpha/2}$, donde $z_{\alpha/2}$ es el valor z de la tabla 3 del Anexo correspondiente a un área de $\alpha/2$ en la cola superior de la distribución normal.

EJEMPLO

11.4

Un superintendente de producción dice que no hay diferencia entre los porcentajes de accidentes de empleados en turnos de día o de noche en una gran planta manufacturera. Se registra el número diario de accidentes para los turnos de día y de noche durante $n = 100$ días. Se encuentra que el número diario de accidentes en el turno de noche x_E excedió al número correspondiente de éstos en el turno de día x_D en 63 de los 100 días. ¿Estos resultados dan suficiente evidencia para indicar que más accidentes tienden a ocurrir en un turno que en el otro, o bien, es equivalente, que $P(x_E > x_D) \neq 1/2$?

Solución Este estudio es un experimento de diferencia en pares, con $n = 100$ pares de observaciones correspondientes a los 100 días. Para probar la hipótesis nula de que las dos distribuciones de accidentes son idénticas, se puede usar el estadístico de prueba

$$z = \frac{x - .5n}{.5\sqrt{n}}$$

donde x es la cantidad de días en los cuales el número de accidentes en el turno de noche excedió al de accidentes en el turno de día. Entonces, para $\alpha = .05$, se puede rechazar la hipótesis nula si $z \geq 1.96$ o $z \leq -1.96$. Sustituyendo en la fórmula para z , se obtiene

$$z = \frac{x - .5n}{.5\sqrt{n}} = \frac{63 - (.5)(100)}{.5\sqrt{100}} = \frac{13}{5} = 2.60$$

Como el valor calculado de z excede de $z_{\alpha/2} = 1.96$, se puede rechazar la hipótesis nula. Los datos dan suficiente evidencia para indicar una diferencia en las distribuciones del porcentaje de accidentes para el turno de día contra el de noche.

¿Cuándo debe usarse la prueba del signo en vez de la prueba t ? Cuando se da sólo la *dirección* de la diferencia en la medición, sólo se puede usar la prueba del signo. Por el contrario, cuando los datos son cuantitativos y satisfacen las suposiciones de normalidad y varianza constante, debe usarse la prueba t en pares. Se utiliza una gráfica de probabilidad normal para evaluar normalidad, en tanto que una gráfica de los residuales ($d_i - d$) puede revelar grandes desviaciones que indicarían una varianza diferente de un par a otro. Cuando haya dudas acerca de la validez de las suposiciones, los estadísticos recomiendan con frecuencia que se realicen ambas pruebas. Si las dos llegan a las mismas conclusiones, entonces los resultados de la prueba paramétrica se consideran válidos.

11.3 EJERCICIOS

TÉCNICAS BÁSICAS

11.12 Suponga que desea usar la prueba del signo para probar $H_a: p > .5$ para un experimento de diferencia en pares con $n = 25$ pares.

- Expresar la situación práctica que dicta la hipótesis alternativa dada.
- Usar la tabla 1 del Anexo para hallar los valores de α ($\alpha < .15$) disponible para la prueba.

11.13 Repita las instrucciones del ejercicio 11.12 para $H_a: p \neq .5$.

11.14 Repita las instrucciones de los ejercicios 11.12 y 11.13 para $n = 10, 15$ y 20 .

DATOS
EX1115 **11.15** Se realizó un experimento de diferencia en pares para comparar dos poblaciones. Los datos se muestran en la tabla siguiente. Use la prueba del signo para determinar si las distribuciones poblacionales son diferentes.

Población	Pares						
	1	2	3	4	5	6	7
1	8.9	8.1	9.3	7.7	10.4	8.3	7.4
2	8.8	7.4	9.0	7.8	9.9	8.1	6.9

- Expresar las hipótesis nula y alternativa para la prueba.
- Determinar una región de rechazo apropiada con $\alpha \approx .01$.
- Calcular el valor observado del estadístico de prueba.
- ¿Los datos presentan suficiente evidencia para indicar que las poblaciones 1 y 2 son diferentes?

APLICACIONES

DATOS
EX1116 **11.16 Valores de propiedades** En el ejercicio 6.42 comparamos las evaluaciones de propiedades de dos asesores de impuestos, A y B. Sus evaluaciones para ocho propiedades se muestran en la tabla:

Propiedad	Asesor A	Asesor B
1	276.3	275.1
2	288.4	286.8
3	280.2	277.3
4	294.7	290.6
5	268.7	269.1
6	282.8	281.0
7	276.1	275.3
8	279.0	279.1

- Usar la prueba del signo para determinar si los datos presentan evidencia suficiente para indicar que uno de los asesores tiende a ser consistentemente más conservador que el otro; es decir, $P(x_A > x_B) \neq 1/2$.

Pruebe usando un valor de α cercano a .05. Encuentre el valor p para la prueba e interprete su valor.

- El ejercicio 6.42 usa el estadístico t para probar la hipótesis nula de que no hay diferencia en las evaluaciones medias de propiedades entre los asesores A y B. Compruebe la respuesta para el ejercicio 6.42 y compárela con su respuesta de la parte a. ¿Concuerdan los resultados de la prueba? Explique por qué las respuestas son (o no son) consistentes.

DATOS
EX1117 **11.17 Cocina de gourmet** Dos gourmets, A y B, calificaron 22 comidas en una escala del 1 al 10. Los datos se muestran en la tabla. ¿Los datos dan suficiente evidencia para indicar que uno de los gourmets tiende a dar calificaciones más altas que el otro? Pruebe usando la prueba del signo con un valor de α cercano a .05.

Comida	A	B	Comida	A	B
1	6	8	12	8	5
2	4	5	13	4	2
3	7	4	14	3	3
4	8	7	15	6	8
5	2	3	16	9	10
6	7	4	17	9	8
7	9	9	18	4	6
8	7	8	19	4	3
9	2	5	20	5	4
10	4	3	21	3	2
11	6	9	22	5	3

- Usar las tablas binomiales del Anexo para hallar la región de rechazo exacta para la prueba.
- Usar el estadístico z de muestra grande. (NOTA: Aun cuando la aproximación de muestra grande se sugiere para $n \geq 25$, funciona bastante bien para los valores de n tan pequeños como 15.)
- Comparar los resultados de las partes a y b.

11.18 Niveles de plomo en la sangre Un estudio publicado en la *American Journal of Public Health* (*Science News*), primera publicación en seguir los niveles de plomo en la sangre de los aficionados al tiro al blanco con pistola en polígonos de tiro bajo techo y que cumplen con la ley, documenta un riesgo significativo de envenenamiento por plomo.³ Se tomaron mediciones de exposición al plomo a 17 miembros del grupo de entrenamiento del cuerpo de policía antes, durante y después de un periodo de tres meses de instrucción de disparo en un polígono de tiro bajo techo propiedad del Estado. Ningún recluta tenía niveles elevados de plomo en la sangre antes del entrenamiento, pero 15 de los 17 terminaron su entrenamiento con niveles de plomo considerados “elevados” por la Agencia para la Seguridad y la Salud en el Trabajo (OSHA). Si el uso de un polígono de tiro no causa un aumento en niveles de plomo

en la sangre, entonces p , la probabilidad de que aumente el nivel de plomo en la sangre de una persona es menor o igual a .5. Pero si el uso del polígono de tiro bajo techo causa un aumento en los niveles de plomo en la sangre de una persona, entonces $p > .5$. Use la prueba del signo para determinar si el uso de un polígono de tiro bajo techo tiene el efecto de aumentar el nivel de plomo en la sangre con $\alpha = .05$. (SUGERENCIA: La aproximación normal a probabilidades binomiales es bastante precisa para $n = 17$.)

DATOS
EX1119

11.19 Porcentajes de recuperación De 10 hospitales se recolectaron datos clínicos respecto a la eficacia de dos medicamentos para tratar una

enfermedad particular. La cantidad de pacientes tratados con los medicamentos varió de un hospital a otro. Usted desea saber si los datos presentan evidencia suficiente con el fin de indicar un porcentaje más alto de recuperación para uno de los dos medicamentos.

- Pruebe usando la prueba del signo. Escoja su región de rechazo de modo que α sea cercana a .05.
- ¿Por qué podría ser inapropiado usar la prueba t de Student al analizar los datos?

Hospital	Medicamento A		
	Número en grupo	Número recuperado	Porcentaje recuperado
1	84	63	75.0
2	63	44	69.8
3	56	48	85.7
4	77	57	74.0
5	29	20	69.0
6	48	40	83.3
7	61	42	68.9
8	45	35	77.8
9	79	57	72.2
10	62	48	77.4

Hospital	Medicamento B		
	Número en grupo	Número recuperado	Porcentaje recuperado
1	96	82	85.4
2	83	69	83.1
3	91	73	80.2
4	47	35	74.5
5	60	42	70.0
6	27	22	81.5
7	69	52	75.4
8	72	57	79.2
9	89	76	85.4
10	46	37	80.4

UNA COMPARACIÓN DE PRUEBAS ESTADÍSTICAS

11.4

El experimento del ejemplo 11.3 está diseñado como experimento de diferencia en pares. Si se satisfacen las suposiciones de normalidad y varianza constante, σ_d^2 , para las diferencias, ¿la prueba del signo detectaría un cambio en la ubicación para las dos poblaciones tan eficientemente como la prueba t en pares? Es probable que no, porque la prueba t usa mucha más información que la prueba del signo. Usa no sólo el signo de la diferencia, sino también los valores reales de las diferencias. En este caso, diríamos que la prueba del signo no es tan *eficiente* como la prueba t en pares. No obstante, la prueba del signo podría ser más eficiente si no se satisfacen las suposiciones acostumbradas.

Cuando dos pruebas estadísticas diferentes se pueden usar para probar una hipótesis basada en los mismos datos, es natural preguntar: ¿cuál es mejor? Una forma de contestar esta pregunta sería mantener constante el tamaño muestral n y α también constante para ambos procedimientos y comparar β , la probabilidad de un error tipo II. Los expertos en estadística, sin embargo, prefieren examinar la **potencia** de una prueba.

Definición $\text{Potencia} = 1 - \beta = P(\text{rechazar } H_0 \text{ cuando } H_a \text{ es verdadera})$

Como β es la probabilidad de no rechazar la hipótesis nula cuando es falsa, la **potencia** de la prueba es la probabilidad de rechazar la hipótesis nula cuando es falsa y alguna alternativa especificada es verdadera. Es la probabilidad de que la prueba haga aquello para lo que está diseñada, es decir, detectar una desviación de la hipótesis nula cuando exista una desviación.

Probablemente el método más común de comparar dos procedimientos de prueba es en términos de la eficiencia relativa de un par de pruebas. La **eficiencia relativa** es la razón entre los tamaños muestrales, para los dos procedimientos de prueba requeridos para alcanzar la misma α y β para una alternativa determinada a la hipótesis nula.

En algunas situaciones, es posible que usted no esté demasiado preocupado si está usando la prueba más potente. Por ejemplo, podría elegir usar la prueba del signo por encima de una competidora más potente por su facilidad de aplicación. Entonces, se podrían ver las pruebas como microscopios que se usan para detectar desviaciones de una teoría hipotética. Uno no tiene que saber la potencia exacta de un microscopio para usarlo en una investigación biológica, y lo mismo se aplica a las pruebas estadísticas. Si el procedimiento de prueba detecta una desviación de la hipótesis nula, estaría encantado; si no es así, volvería a analizar los datos usando un microscopio (prueba) más potente, o aumentar la potencia del microscopio (prueba) incrementando el tamaño muestral.

LA PRUEBA DE RANGO CON SIGNO DE WILCOXON PARA UN EXPERIMENTO DE DOS POBLACIONES

11.5

Se puede usar una prueba de rango con signo, propuesta por Frank Wilcoxon para analizar el experimento de diferencia en pares de la sección 6.5 al considerar las diferencias en pares de dos tratamientos, 1 y 2. Bajo la hipótesis nula de que no hay diferencias en las distribuciones para 1 y 2 se esperaría que (en promedio) la mitad de las diferencias en pares sean negativas y la mitad sean positivas; esto es, el número esperado de diferencias negativas entre pares sería $n/2$ (donde n es el número de pares). Además, se deduce que las diferencias positivas y negativas de igual magnitud absoluta deben presentarse con igual probabilidad. Si fuéramos a ordenar las diferencias de acuerdo con sus valores absolutos y de menor a mayor, las sumas de rango esperadas para las diferencias negativas y positivas sería igual. Las diferencias grandes en las sumas de los rangos, asignadas a las diferencias positivas y negativas, darían evidencia para indicar un cambio en lugar entre las distribuciones de respuestas para los dos tratamientos, 1 y 2.

Si la distribución 1 se corre a la derecha de la distribución 2, entonces se espera que más de las diferencias sean positivas y esto resulta en un número pequeño de diferencias negativas. Por tanto, para detectar esta alternativa de una cola, use la suma de rango T^- (la suma de los rangos de las diferencias negativas) y rechace la hipótesis nula para valores significativamente pequeños de T^- . Del mismo modo, si la distribución 1 se corre a la izquierda de la distribución 2, entonces se espera que más de las diferencias sean negativas y que el número de diferencias positivas sea pequeño. En consecuencia, para detectar esta alternativa de una cola, use T^+ (la suma de los rangos de las diferencias positivas) y rechace la hipótesis nula si T^+ es significativamente pequeña.

CÁLCULO DEL ESTADÍSTICO DE PRUEBA PARA LA PRUEBA DE RANGO CON SIGNO DE WILCOXON

1. Calcule las diferencias $(x_1 - x_2)$ para cada uno de los n pares. Las diferencias iguales a 0 se eliminan y el número de pares, n , se reduce de conformidad.
2. Ordene los **valores absolutos** de las diferencias asignando 1 a la más pequeña, 2 a la segunda más pequeña, y así sucesivamente. A las observaciones empatadas se les asigna el promedio de los rangos que se hubieran asignado sin empates.
3. Calcule la **suma de rango** para las diferencias **negativas** y marque este valor T^- . Del mismo modo, calcule T^+ , la **suma de rango** para las diferencias **positivas**.

Para una **prueba de dos colas** use la **menor de estas dos cantidades T como un estadístico de prueba** para probar la hipótesis nula de que los dos histogramas de frecuencia relativa poblacional son idénticos. Cuanto menor sea el valor de T , mayor es el peso de evidencia a favor de rechazar la hipótesis nula. **Por tanto, se rechazará la hipótesis nula si T es menor o igual a algún valor, por ejemplo T_0 .**

Para detectar la **alternativa de una cola**, esa **distribución 1 se corre a la derecha de la distribución 2, use la suma de rango T^-** de las diferencias negativas y rechace la hipótesis nula para valores pequeños de T^- , por ejemplo $T^+ \leq T_0$. Si desea detectar un **corrimiento de la distribución 2 a la derecha de la distribución 1, use la suma de rango T^+** de las diferencias positivas como estadístico de prueba y rechace la hipótesis nula para valores pequeños de T^+ , por ejemplo, $T^+ \leq T_0$.

La probabilidad de que T sea menor o igual a algún valor T_0 se ha calculado para una combinación de tamaños muestrales y valores de T_0 . Estas probabilidades, dadas en la tabla 8 del Anexo, se pueden usar para hallar la región de rechazo para la prueba T .

Una versión abreviada de la tabla 8 se muestra en la tabla 11.7. En sentido horizontal en la parte superior de la tabla se ve el número de diferencias (el número de pares) n . Los valores de α para una prueba de una cola aparecen en la primera columna de la tabla. La segunda columna da valores de α para una prueba de dos colas. Las entradas de la tabla son los valores críticos de T . Usted recordará que el valor crítico de un estadístico de prueba es el valor que localiza la frontera de la región de rechazo.

Por ejemplo, suponga que tenemos $n = 7$ pares y se realiza una prueba de dos colas de la hipótesis nula de que las dos distribuciones de frecuencia relativa poblacional son idénticas. Al comprobar la columna $n = 7$ de la tabla 11.7 y usar el segundo renglón (correspondiente a $\alpha = .05$ para una prueba de dos colas), se ve la entrada 2 (sombreada). Este valor es T_0 , el valor crítico de T . Como ya se vio antes, cuanto menor sea el valor de T , mayor es la evidencia para rechazar la hipótesis nula. Por tanto, se rechazará la hipótesis nula para todos los valores de T menores o iguales a 2. La región de rechazo para la prueba de rango con signo de Wilcoxon para un experimento en pares es siempre de la forma: rechazar H_0 si $T \leq T_0$, donde T_0 es el valor crítico de T . La región de rechazo se muestra simbólicamente en la figura 11.2.

TABLA 11.7

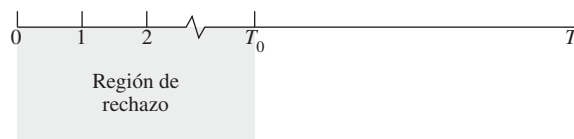
Versión abreviada de la tabla 8 del Anexo; valores críticos de T

Una cola	Dos colas	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$	$n = 11$
$\alpha = .050$	$\alpha = .10$	1	2	4	6	8	11	14
$\alpha = .025$	$\alpha = .05$		1	2	4	6	8	11
$\alpha = .010$	$\alpha = .02$			0	2	3	5	7
$\alpha = .005$	$\alpha = .01$				0	2	3	5

Una cola	Dos colas	$n = 12$	$n = 13$	$n = 14$	$n = 15$	$n = 16$	$n = 17$
$\alpha = .050$	$\alpha = .10$	17	21	26	30	36	41
$\alpha = .025$	$\alpha = .05$	14	17	21	25	30	35
$\alpha = .010$	$\alpha = .02$	10	13	16	20	24	28
$\alpha = .005$	$\alpha = .01$	7	10	13	16	19	23

FIGURA 11.2

Región de rechazo para la prueba de rango con signo de Wilcoxon para un experimento pareado (rechazar H_0 si $T \leq T_0$)



PRUEBA DE RANGO CON SIGNO DE WILCOXON PARA UN EXPERIMENTO DE DOS POBLACIONES

1. Hipótesis nula: H_0 : Las dos distribuciones de frecuencia relativa poblacional son idénticas.
2. Hipótesis alternativa: H_a : Las dos distribuciones de frecuencia relativa poblacional difieren en ubicación (una prueba de dos colas); o H_a : La distribución de frecuencia relativa de la población 1 se corre a la derecha de la distribución de frecuencia relativa para la población 2 (una prueba de una cola).
3. Estadístico de prueba
 - a. Para una prueba de dos colas, use T , la menor de la suma de rango para diferencias positivas y la suma de rango para diferencias negativas.
 - b. Para una prueba de una cola (detectar la hipótesis alternativa descrita líneas antes), use la suma de rango T^- de las diferencias negativas.
4. Región de rechazo
 - a. Para una prueba de dos colas, rechace H_0 si $T \leq T_0$, donde T_0 es el valor crítico dado en la tabla 8 del Anexo.
 - b. Para una prueba de una cola (detectar la hipótesis alternativa descrita líneas antes), use la suma de rango T^- de las diferencias negativas. Rechace H_0 si $T^- \leq T_0$.[†]

$$\left[\text{NOTA: se puede demostrar que } T^+ + T^- = \frac{n(n+1)}{2} \right]$$

EJEMPLO

11.5

Se realizó un experimento para comparar las densidades de pasteles elaborados de dos clases diferentes de harinas para pastel, A y B. Seis moldes para hornear pastel recibieron la masa A y seis la masa B. Esperando una variación en la temperatura del horno, el experimentador colocó un pastel A y uno B juntos en seis lugares diferentes del horno. Pruebe la hipótesis de que no hay diferencia en las distribuciones poblacionales de densidades de pastel para dos masas para pastel diferentes.

Solución Los datos (densidad en onzas por pulgada cúbica) y diferencias en densidad para seis pares de pasteles se dan en la tabla 11.8. La gráfica de caja de las diferencias de la figura 11.3 muestra un sesgo bastante fuerte y una diferencia muy grande en la cola derecha, lo cual indica que los datos pueden no satisfacer la suposición de normalidad. La muestra de dos diferencias es demasiado pequeña para tomar decisiones válidas acerca de normalidad y varianza constante. En esta situación, la prueba de rango con signo de Wilcoxon es la prueba más prudente a usar.

Al igual que con otras pruebas no paramétricas, la hipótesis nula a probar es que las dos distribuciones de frecuencia poblacionales de densidades de pastel son idénticas. La hipótesis alternativa, que implica una prueba de dos colas, es que las distribuciones son diferentes. Como la cantidad de datos es pequeña, se puede efectuar la prueba usando $\alpha = .10$. De la tabla 8 del Anexo, el valor crítico de T para una prueba de dos colas, $\alpha = .10$, es $T_0 = 2$. Por tanto, se puede rechazar H_0 si $T \leq 2$.

[†] Para detectar un cambio de la distribución 2 a la derecha de la distribución 1, use la suma de rango T^+ de las diferencias positivas como estadístico de prueba y rechace H_0 si $T^+ \leq T_0$.

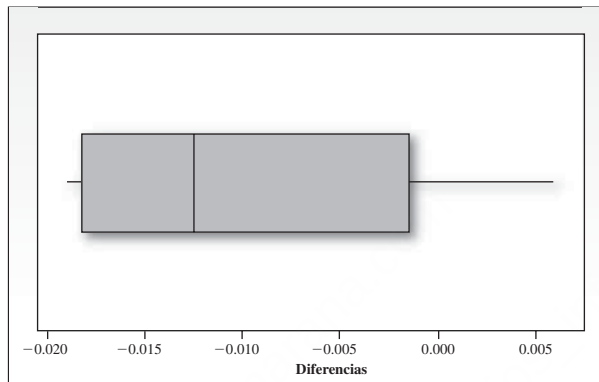
TABLA 11.8

Densidades de seis pares de pasteles

x_A	x_B	Diferencia ($x_A - x_B$)	Rango
.135	.129	.006	2
.102	.120	-.018	5
.098	.112	-.014	4
.141	.152	-.011	3
.131	.135	-.004	1
.144	.163	-.019	6

FIGURA 11.3

Gráfica de caja de diferencias para el ejemplo 11.5



Las diferencias ($x_1 - x_2$) están calculadas y ordenadas de acuerdo con sus valores absolutos en la tabla 11.8. La suma de rangos positivos es $T^+ = 2$ y la suma de rangos negativos es $T^- = 19$. El estadístico de prueba es la más pequeña de estas dos sumas de rango, o $T = 2$. Como $T = 2$ cae en la región de rechazo, se puede rechazar H_0 y concluir que las dos distribuciones de frecuencia poblacional de densidades de pastel difieren.

Una salida impresa *MINITAB* de la prueba de rango con signo de Wilcoxon se da en la figura 11.4. En la sección “Tecnología actual”, al final de este capítulo, encontrará instrucciones para generar esta salida impresa. Observará que el valor del estadístico de prueba concuerda con los otros cálculos y el valor p indica que se puede rechazar H_0 al nivel de significancia de 10%.

FIGURA 11.4

Salida impresa *MINITAB* para el ejemplo 11.5

Prueba de rango con signo de Wilcoxon: diferencia

Test of median = 0.000000 versus median not = 0.000000

	N	N for Test	Wilcoxon Statistic	P	Estimated Median
Difference	6	6	2.0	0.093	-0.01100

Aproximación normal para la prueba de rango con signo de Wilcoxon

Aun cuando la tabla 8 del Anexo tiene valores críticos para n de hasta 50, T^+ , al igual que la prueba de rango con signo de Wilcoxon, estará distribuida en forma aproximadamente normal

cuando la hipótesis nula sea verdadera y sea grande, por ejemplo 25 o más. Esto hace posible construir una prueba z de muestra grande, donde

$$E(T) = \frac{n(n+1)}{4}$$

$$\sigma_T^2 = \frac{n(n+1)(2n+1)}{24}$$

Entonces el estadístico z

$$z = \frac{T^+ - E(T)}{\sigma_T} = \frac{T^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

se usa como estadístico de prueba. Por tanto, para una prueba de dos colas y $\alpha = .05$, usted puede rechazar la hipótesis de distribuciones poblacionales idénticas cuando $|z| \geq 1.96$.

PRUEBA DE RANGO CON SIGNO DE WILCOXON DE MUESTRA GRANDE PARA UN EXPERIMENTO EN PARES: $n \geq 25$

1. Hipótesis nula: H_0 : las distribuciones 1 y 2 de frecuencia relativa poblacional son idénticas.
2. Hipótesis alternativa: H_a : las dos distribuciones de frecuencia relativa poblacional difieren en ubicación (una prueba de dos colas); o bien, H_a : la distribución de frecuencia relativa poblacional 1 está corrida a la derecha (o izquierda) de la distribución de frecuencia relativa para la población 2 (una prueba de una cola).
3. Estadístico de prueba: $z = \frac{T^+ - [n(n+1)/4]}{\sqrt{[n(n+1)(2n+1)]/24}}$
4. Región de rechazo: rechazar H_0 si $z > z_{\alpha/2}$ o $z < -z_{\alpha/2}$ para una prueba de dos colas. Para una prueba de una cola, pónganse todas las a en una cola de la distribución z . Para detectar un cambio en la distribución 1 a la derecha de la distribución 2, rechazar H_0 cuando $z > z_{\alpha}$. Para detectar un cambio en la dirección opuesta, rechace H_0 si $z < -z_{\alpha}$.

Los valores tabulados de z se dan en la tabla 3, Anexo.

11.5 EJERCICIOS

TÉCNICAS BÁSICAS

11.20 Suponga que se desea detectar una diferencia en las ubicaciones de dos distribuciones poblacionales, basadas en un experimento de diferencia en pares de $n = 30$ pares.

- a. Dé las hipótesis nula y alternativa para la prueba de rango con signo de Wilcoxon.
- b. Dé el estadístico de prueba.
- c. Dé la región de rechazo para la prueba para $\alpha = .05$.
- d. Si $T^+ = 249$, ¿cuáles son sus conclusiones? [NOTA: $T^+ + T^- = n(n+1)/2$.]

11.21 Consulte el ejercicio 11.20. Suponga que desea detectar sólo un corrimiento en la distribución 1 a la derecha de la distribución 2.

- a. Dé las hipótesis nula y alternativa para la prueba de rango con signo de Wilcoxon.
- b. Dé el estadístico de prueba.
- c. Dé la región de rechazo para la prueba para $\alpha = .05$.
- d. Si $T^+ = 249$, ¿cuáles son sus conclusiones? [NOTA: $T^+ + T^- = n(n+1)/2$.]

11.22 Consulte el ejercicio 11.20. Realice la prueba usando la prueba z de muestra grande. Compare sus resultados con los resultados de prueba no paramétrica del ejercicio 11.21, parte **d**.

11.23 Consulte el ejercicio 11.21. Realice la prueba usando la prueba z de muestra grande. Compare sus resultados con los resultados de la prueba no paramétrica del ejercicio 11.20, parte **d**.

11.24 Consulte el ejercicio 11.15 y el conjunto de datos EX1115. Los datos de esta tabla son de un experimento de diferencia en pares con $n = 7$ pares de observaciones.

Población	Pares						
	1	2	3	4	5	6	7
1	8.9	8.1	9.3	7.7	10.4	8.3	7.4
2	8.8	7.4	9.0	7.8	9.9	8.1	6.9

- Use la prueba de rango con signo de Wilcoxon para determinar si hay una diferencia significativa entre las dos poblaciones.
- Compare los resultados de la parte **a** con el resultado que obtuvo en el ejercicio 11.15. ¿Son iguales? Explique.

APLICACIONES

11.25 Valores de propiedades II En el ejercicio 11.16 se usó la prueba del signo para determinar si los datos daban evidencia suficiente con el fin de indicar una diferencia, en las distribuciones de evaluaciones de propiedad, para los asesores A y B.

- Use la prueba del rango con signo de Wilcoxon para un experimento en pares, con el fin de probar la hipótesis nula de que no hay diferencia en las distribuciones de evaluaciones de propiedades entre los asesores A y B. Pruebe usando un valor de α cercano a .05.
- Compare las conclusiones de la prueba de la parte **a** con las conclusiones derivadas de la prueba t del ejercicio 6.42, y la prueba del signo del ejercicio 11.16. Explique por qué estas conclusiones de prueba son (o no son) consistentes.

DATOS EX1126 **11.26 Descomposturas de máquinas** El número de descomposturas mensuales en máquinas se registró durante nueve meses en dos máquinas idénticas, A y B, empleadas para hacer cables:

Mes	A	B
1	3	7
2	14	12
3	7	9
4	10	15
5	9	12
6	6	6
7	13	12
8	6	5
9	7	13

- ¿Los datos dan suficiente evidencia para indicar una diferencia en los porcentajes de descomposturas mensuales para las dos máquinas? Pruebe usando un valor de α cercano a .05.
- ¿Piensa en alguna razón por la que los porcentajes de descomposturas para las dos máquinas varían de un mes a otro?

11.27 Cocina de gourmet II Consulte la comparación de calificaciones de comidas de gourmet del ejercicio 11.17 y use la prueba de rango con signo de Wilcoxon, para determinar si los datos dan evidencia suficiente con el fin de indicar una diferencia en las calificaciones de los dos gourmets. Pruebe usando un valor de α cercano a .05. Compare los resultados de esta prueba con los de la prueba con signo del ejercicio 11.17. ¿Las conclusiones de la prueba son consistentes?

DATOS EX1128

11.28 Control de tránsito Dos métodos para controlar el tránsito, A y B, se usaron en cada uno de $n = 12$ cruceros durante una semana y los números de accidentes que ocurrieron durante ese tiempo se registraron. El orden de uso (cuál método se emplearía para la primera semana) se seleccionó de una manera aleatoria. Se desea saber si los datos dan suficiente evidencia para indicar una diferencia en las distribuciones de porcentajes de accidentes para los métodos A y B de control de tránsito.

Crucero	Método		Crucero	Método	
	A	B		A	B
1	5	4	7	2	3
2	6	4	8	4	1
3	8	9	9	7	9
4	3	2	10	5	2
5	6	3	11	6	5
6	1	0	12	1	1

- Analice usando una prueba de signo.
- Analice usando la prueba de rango con signo de Wilcoxon para un experimento en pares.

DATOS EX1129

11.29 Rompecabezas Se le pidió a 8 personas una tarea sencilla de armar un rompecabezas bajo condiciones normales y bajo condiciones de estrés. Durante el tiempo de estrés, se presentó un estímulo a las personas tres minutos después del inicio del experimento y cada 30 segundos de ahí en adelante, hasta que terminaran la tarea. Se tomaron lecturas de la presión sanguínea en ambas condiciones. Los datos de la tabla son las lecturas más altas durante el experimento. ¿Los datos presentan suficiente evidencia para indicar lecturas de presión sanguínea más altas en condiciones de estrés? Analícelos usando la prueba de rango con signo de Wilcoxon para un experimento en pares.

Persona	Normal	Con estrés
1	126	130
2	117	118
3	115	125
4	118	120
5	118	121
6	128	125
7	125	130
8	120	120

DATOS
EX1130

11.30 Imágenes y recordar palabras Un grupo de psicología realizó un experimento para determinar si una puntuación de recuerdo, en la que se dieron instrucciones para formar imágenes de 25 palabras, difiere de una puntuación de recuerdo inicial para la que no se dieron instrucciones de formar imágenes. Veinte estudiantes participaron en el experimento, con los resultados que se indican en la tabla siguiente.

Estudiante	Con imágenes	Sin imágenes	Estudiante	Con imágenes	Sin imágenes
1	20	5	11	17	8
2	24	9	12	20	16
3	20	5	13	20	10
4	18	9	14	16	12
5	22	6	15	24	7
6	19	11	16	22	9
7	20	8	17	25	21
8	19	11	18	21	14
9	17	7	19	19	12
10	21	9	20	23	13

- ¿Cuáles tres procedimientos de prueba se pueden usar para probar si hay diferencias en la distribución de las puntuaciones de recuerdo con y sin imágenes? ¿Qué suposiciones se requieren para el procedimiento paramétrico? ¿Estos datos satisfacen esas suposiciones?
- Use la prueba del signo y la prueba de rango con signo de Wilcoxon para probar si hay diferencias en las distribuciones de las puntuaciones de recuerdo bajo estas dos condiciones.
- Compare los resultados de las pruebas de la parte **b**. ¿Son iguales las conclusiones? Si no es así, ¿por qué no?

LA PRUEBA H DE KRUSKAL-WALLIS PARA DISEÑOS COMPLETAMENTE ALEATORIZADOS

11.6

Así como la prueba de la suma de rango de Wilcoxon es la alternativa no paramétrica a la prueba t de Student para una comparación de medias poblacionales, la prueba H de Kruskal-Wallis es la alternativa no paramétrica a la prueba F del análisis de varianza para un diseño completamente aleatorizado. Se usa para detectar diferencias en ubicaciones entre más de dos distribuciones poblacionales basadas en el muestreo aleatorio independiente.

El procedimiento para realizar la prueba H de Kruskal-Wallis es semejante al empleado para la prueba de suma de rango de Wilcoxon. Suponga que se comparan k poblaciones basadas en muestras aleatorias independientes n_1 de la población 1, n_2 de la población 2, ..., n_k de la población k , donde

$$n_1 + n_2 + \cdots + n_k = n$$

El primer paso es ordenar las n observaciones de la más pequeña (rango 1) a la mayor (rango n). A las observaciones con empate se les asigna un rango igual al promedio de los rangos que hubieran recibido de haber sido casi iguales pero no empatadas. A continuación se calculan las sumas de rangos $T_1, T_2, \dots, T_k$ para las k muestras y también el estadístico de prueba

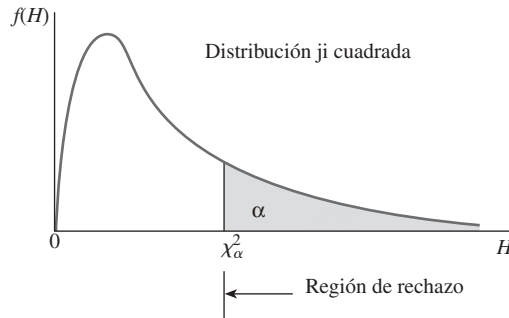
$$H = \frac{12}{n(n+1)} \sum \frac{T_i^2}{n_i} - 3(n+1)$$

que es proporcional a $\sum n_i(\bar{T}_i - \bar{T})^2$, la suma de desviaciones cuadradas de las medias de rango alrededor de la gran media $\bar{T} = n(n+1)/2n = (n+1)/2$. Entre más grandes sean las diferencias en ubicaciones entre las k distribuciones poblacionales, mayor es el valor del estadístico H . Por tanto, se puede rechazar la hipótesis nula de que las k distribuciones poblacionales son idénticas para valores grandes de H .

¿Qué tan grande es grande? Se puede mostrar (se omite la demostración) que cuando los tamaños muestrales son de moderados a grandes —por ejemplo, cada tamaño muestral es igual a cinco o mayor— y H_0 es verdadera, el estadístico H tendrá aproximadamente una distribución ji cuadrada con $(k-1)$ grados de libertad. Por tanto, para un valor determinado de α , se puede rechazar H_0 cuando el estadístico H exceda de χ^2_α (véase la figura 11.5).

FIGURA 11.5

Distribución aproximada del estadístico H cuando H_0 es verdadera



EJEMPLO

11.6

Los datos de la tabla 11.9 se recolectaron usando un diseño completamente aleatorizado. Son las calificaciones del examen de aprovechamiento para cuatro diferentes grupos de estudiantes, donde cada grupo recibió enseñanza mediante una técnica diferente. El objetivo del experimento es probar la hipótesis de que no hay diferencia, en las distribuciones poblacionales de las calificaciones del examen de aprovechamiento, contra la alternativa de que difieren en ubicación; esto es, al menos una de las distribuciones se corre arriba de las otras. Realice la prueba usando la prueba H de Kruskal-Wallis con $\alpha = .05$.

TABLA 11.9

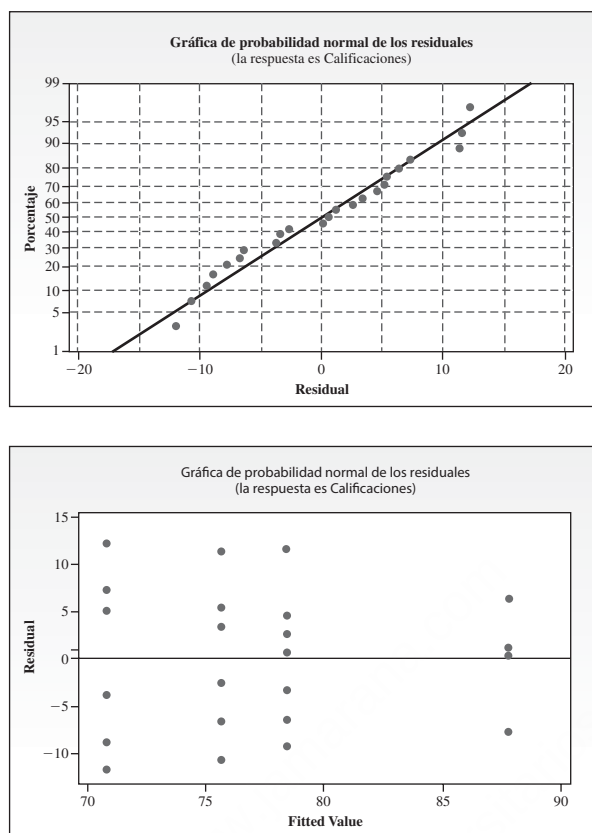
Calificaciones de examen (y rangos) para cuatro técnicas de enseñanza

	1	2	3	4
	65 (3)	75 (9)	59 (1)	94 (23)
	87 (19)	69 (5.5)	78 (11)	89 (21)
	73 (8)	83 (17.5)	67 (4)	80 (14)
	79 (12.5)	81 (15.5)	62 (2)	88 (20)
	81 (15.5)	72 (7)	83 (17.5)	
	69 (5.5)	79 (12.5)	76 (10)	
		90 (22)		
Suma de rango	$T_1 = 63.5$	$T_2 = 89$	$T_3 = 45.5$	$T_4 = 78$

Solución Antes de realizar un análisis no paramétrico de estos datos, se puede usar un análisis de varianza de una vía para producir las dos gráficas de la figura 11.6. Parece que la técnica 4 tiene una varianza más pequeña que las otras tres y que hay una marcada desviación en la cola derecha de la gráfica de probabilidad normal. Estas desviaciones se considerarían menores y entonces usarse un análisis ya sea paramétrico o no paramétrico.

FIGURA 11.6

Gráfica de probabilidad normal y una gráfica residual resultantes de un análisis de varianza de una vía para el ejemplo 11.6



En el procedimiento de la prueba H de Kruskal-Wallis, el primer paso es ordenar las $n = 23$ observaciones de menor (rango 1) a mayor (rango 23). Estos rangos se muestran entre paréntesis en la tabla 11.9. Observe cómo se manejan los empates. Por ejemplo, dos observaciones en 69 están empatadas para el rango 5. Por tanto, se les asigna el promedio 5.5 de los dos rangos (5 y 6) que hubieran ocupado si hubieran sido ligeramente diferentes. Las sumas de rango T_1 , T_2 , T_3 y T_4 para las cuatro muestras se ven en el renglón inferior de la tabla. Sustituyendo sumas de rango y tamaños muestrales en la fórmula para el estadístico H , obtenemos

$$\begin{aligned}
 H &= \frac{12}{n(n+1)} \sum \frac{T_i^2}{n_i} - 3(n+1) \\
 &= \frac{12}{23(24)} \left[\frac{(63.5)^2}{6} + \frac{(89)^2}{7} + \frac{(45.5)^2}{6} + \frac{(78)^2}{4} \right] - 3(24) \\
 &= 79.775102 - 72 = 7.775102
 \end{aligned}$$

La región de rechazo para el estadístico H para $\alpha = .05$ incluye valores de $H \geq \chi_{.05}^2$, donde $\chi_{.05}^2$ está basada en $(k-1) = (4-1) = 3$ grados de libertad. El valor de χ^2 dado en la tabla 5 en el Anexo es $\chi_{.05}^2 = 7.81473$. El valor observado del estadístico H , $H = 7.775102$, no cae en la región de rechazo para la prueba. Por tanto, hay evidencia insuficiente para indicar diferencias en las distribuciones de calificaciones de examen de aprovechamiento para las cuatro técnicas de enseñanza.

En la figura 11.7 se da una salida impresa *MINITAB* de la prueba H de Kruskal-Wallis para estos datos. Observe que el valor p , .051, es sólo ligeramente mayor que el nivel de 5% necesario para declarar significancia estadística.

FIGURA 11.7

Salida impresa *MINITAB*
para la prueba Kruskal-
Wallis para el ejemplo 11.6

Prueba Kruskal-Wallis: calificaciones contra técnicas

Kruskal-Wallis Test on Scores

Techniques	N	Median	Ave Rank	Z
1	6	76.00	10.6	-0.60
2	7	79.00	12.7	0.33
3	6	71.50	7.6	-1.86
4	4	88.50	19.5	2.43
Overall	23		12.0	

H = 7.78 DF = 3 P = 0.051

H = 7.79 DF = 3 P = 0.051 (adjusted for ties)

* NOTE * One or more small samples

EJEMPLO**11.7**

Compare los resultados de la prueba F del análisis de varianza y la prueba H de Kruskal-Wallis para probar si hay diferencias en las distribuciones de calificaciones del examen de aprovechamiento para las cuatro técnicas de enseñanza del ejemplo 11.6.

Solución La salida impresa *MINITAB* para un análisis de varianza de una vía para los datos de la tabla 11.9 se dan en la figura 11.8. El análisis de varianza muestra que la prueba F para probar si hay diferencias entre las medias para las cuatro técnicas es significativa al nivel de .028. La prueba H de Kruskal-Wallis no detectó un corrimiento en las distribuciones poblacionales al nivel de significancia de .05. Aun cuando estas conclusiones parecen estar muy separadas, los resultados de la prueba no difieren en mucho. El valor $p = .028$ correspondiente a $F = 3.77$, con $gl_1 = 3$ y $gl_2 = 19$, es ligeramente menor a .05, en contraste con el valor $p = .051$ para $H = 7.78$, $gl = 3$ que es ligeramente mayor a .05. Alguien que vea los valores p para las dos pruebas notaría poca diferencia en los resultados de las pruebas F y H . No obstante, si nos apegamos a la elección de $\alpha = .05$, no se puede rechazar H_0 usando la prueba H .

FIGURA 11.8

Salida impresa *MINITAB*
para el ejemplo 11.7

Anova de una vía: calificaciones contra técnicas

Source	DF	SS	MS	F	P
Techniques	3	712.6	237.5	3.77	0.028
Error	19	1196.6	63.0		
Total	22	1909.2			

LA PRUEBA H DE KRUSKAL-WALLIS PARA COMPARAR MÁS DE DOS POBLACIONES: DISEÑO COMPLETAMENTE ALEATORIZADO (MUESTRAS ALEATORIAS INDEPENDIENTES)

1. Hipótesis nula: H_0 : las k distribuciones poblacionales son idénticas.
2. Hipótesis alternativa: H_a : al menos dos de las k distribuciones poblacionales difieren en ubicación.

3. Estadístico de prueba:
$$H = \frac{12}{n(n+1)} \sum \frac{T_i^2}{n_i} - 3(n+1)$$

LA PRUEBA H DE KRUSKAL-WALLIS PARA COMPARAR MÁS DE DOS POBLACIONES: DISEÑO COMPLETAMENTE ALEATORIZADO (MUESTRAS ALEATORIAS INDEPENDIENTES) (continuación)

donde

- n_i = Tamaño muestral para la población i
- T_i = Suma de rango para la población i
- n = Número total de observaciones = $n_1 + n_2 + \dots + n_k$

4. Región de rechazo para una α determinada: $H > \chi^2_{\alpha}$ con $(k - 1) gl$

Suposiciones

- Todos los tamaños muestrales son mayores o iguales a 5.
- Los empates toman el promedio de los rangos que hubieran ocupado de no haber estado empatados.

La prueba H de Kruskal-Wallis es una valiosa alternativa para un análisis de varianza de una vía cuando son violadas las suposiciones de normalidad e igualdad de varianza. De nuevo, gráficas normales de probabilidad de residuales y gráficas de residuales por grupo de tratamiento son útiles para determinar si estas suposiciones han sido violadas. Recuerde que una gráfica de probabilidad normal debe aparecer como una recta con pendiente positiva; las gráficas de residuales por grupos de tratamiento deben exhibir la misma dispersión arriba y abajo de la línea 0.

11.6 EJERCICIOS

TÉCNICAS BÁSICAS

DATOS
EX1131 **11.31** Se compararon tres tratamientos usando un diseño completamente aleatorizado. Los datos se muestran en la tabla siguiente.

Tratamiento		
1	2	3
26	27	25
29	31	24
23	30	27
24	28	22
28	29	24
26	32	20
	30	21
	33	

¿Los datos dan suficiente evidencia para indicar una diferencia en ubicación para al menos dos de las distribuciones poblacionales? Pruebe usando el estadístico H de Kruskal-Wallis con $\alpha = .05$.

DATOS
EX1132 **11.32** Se compararon cuatro tratamientos usando un diseño completamente aleatorizado. Los datos se muestran en la tabla siguiente:

Tratamiento			
1	2	3	4
124	147	141	117
167	121	144	128
135	136	139	102
160	114	162	119
159	129	155	128
144	117	150	123
133	109		

¿Los datos aportan suficiente evidencia para indicar una diferencia en ubicación para al menos dos de las distribuciones poblacionales? Pruebe usando el estadístico H de Kruskal-Wallis con $\alpha = .05$.

APLICACIONES

11.33 Sitios pantanosos II El ejercicio 7.13 presenta datos (véase el conjunto de datos EX0713) sobre la tasa de crecimiento de vegetación en cuatro lugares pantanosos no urbanizados. Se seleccionaron al azar seis plantas en cada uno de los cuatro sitios para usarlas en la comparación. Los datos son la media de longitud de hojas por planta (en centímetros) para una muestra aleatoria de 10 hojas por planta.

Lugar	Media de longitud (cm)					
1	5.7	6.3	6.1	6.0	5.8	6.2
2	6.2	5.3	5.7	6.0	5.2	5.5
3	5.4	5.0	6.0	5.6	4.9	5.2
4	3.7	3.2	3.9	4.0	3.5	3.6

- ¿Los datos aportan suficiente evidencia para indicar diferencias de ubicación para al menos dos de las distribuciones de la media de longitud de hoja correspondientes a los cuatro lugares? Pruebe usando la prueba H de Kruskal-Wallis con $\alpha = .05$
- Encuentre el valor p aproximado para la prueba.
- Analizamos este mismo conjunto de datos en el ejercicio 7.13 usando un análisis de varianza. Encuentre el valor p para la prueba F usada para comparar las cuatro medias de lugar en el ejercicio 7.13.
- Compare los valores p en las partes **b** y **c** y explique las implicaciones de la comparación.

11.34 Frecuencia cardiaca y ejercicio El ejercicio 7.60 presentó datos (conjunto de datos EX0760) sobre las frecuencias cardiacas para muestras de 10 hombres seleccionados al azar de cada uno de los cuatro grupos de edades. Cada hombre se ejercitó en una caminadora a un ritmo fijo durante 12 minutos, registrándose el aumento de frecuencia (la diferencia antes y después del ejercicio, en pulsaciones por minuto). Los datos se presentan en la tabla siguiente:

	10–19	20–39	40–59	60–69
	29	24	37	28
	33	27	25	29
	26	33	22	34
	27	31	33	36
	39	21	28	21
	35	28	26	20
	33	24	30	25
	29	34	34	24
	36	21	27	33
	22	32	33	32
Total	309	275	295	282

- ¿Los datos presentan suficiente evidencia para indicar diferencias de ubicación para al menos dos de los cuatro grupos de edad? Pruebe usando la prueba H de Kruskal-Wallis con $\alpha = .01$.

- Encuentre el valor p aproximado para la prueba de la parte **a**.
- Como la prueba F del ejercicio 7.60 y la prueba H de la parte **a** son pruebas para detectar diferencias en la ubicación de las cuatro poblaciones de frecuencia cardiaca, ¿cómo se comparan los resultados de las pruebas? Compare los valores p para las dos pruebas y explique las implicaciones de la comparación.

DATOS
EX1135

11.35 Niveles de pH en el agua Un muestreo de acidez de agua de lluvia para 10 aguaceros seleccionados al azar se registró en tres lugares diferentes en Estados Unidos: el noreste, la región media del Atlántico y el sureste. Las lecturas de pH para estas 30 lluvias se muestran en la tabla. (NOTA: Las lecturas de pH van de 0 a 14; 0 es ácida, 14 es alcalina. El agua pura que cae en aire limpio tiene una lectura de pH de 5.7.)

Noreste	Atlántico medio	Sureste
4.45	4.60	4.55
4.02	4.27	4.31
4.13	4.31	4.84
3.51	3.88	4.67
4.42	4.49	4.28
3.89	4.22	4.95
4.18	4.54	4.72
3.95	4.76	4.63
4.07	4.36	4.36
4.29	4.21	4.47

- ¿Los datos presentan suficiente evidencia para indicar diferencias en los niveles de acidez en lluvias en los tres diferentes lugares? Pruebe usando la prueba H de Kruskal-Wallis.
- Encuentre el valor p aproximado para la prueba de la parte **a** e interprétela.

DATOS
EX1136

11.36 Campañas publicitarias Los resultados de un experimento para investigar el reconocimiento de productos, durante tres campañas publicitarias, se informaron en el ejemplo 7.14. Las respuestas fueron el porcentaje de 400 adultos que estaban familiarizados con el producto recién anunciado. La gráfica de probabilidad normal indicó que los datos no eran aproximadamente normales y debía usarse otro método de análisis. ¿Hay una diferencia significativa entre las tres distribuciones poblacionales de donde vinieron estas muestras? Use un método no paramétrico apropiado para contestar esta pregunta.

Campaña		
1	2	3
.33	.28	.21
.29	.41	.30
.21	.34	.26
.32	.39	.33
.25	.27	.31

LA PRUEBA F_r DE FRIEDMAN PARA DISEÑOS DE BLOQUE ALEATORIZADOS

11.7

La prueba F_r de Friedman, propuesta por Milton Friedman, economista ganador del Premio Nobel, es una prueba no paramétrica para comparar las distribuciones de mediciones para k tratamientos diseñados en b bloques usando un diseño aleatorizado de bloques. El procedimiento para realizar la prueba es muy semejante al empleado para la prueba H de Kruskal-Wallis. El primer paso en el procedimiento es ordenar las k observaciones de tratamiento dentro de cada bloque. Los empates se tratan en la forma usual; es decir, reciben un promedio de los rangos ocupados por las observaciones empatadas. Por tanto, las sumas de rango $T_1, T_2, \dots, T_k$ se obtienen y el estadístico de prueba

$$F_r = \frac{12}{bk(k+1)} \sum T_i^2 - 3b(k+1)$$

se calcula. El valor del estadístico F_r está en un mínimo cuando las sumas de rango son iguales, esto es, $T_1 = T_2 = \dots = T_k$ y aumenta en valor cuando aumentan las diferencias entre las sumas de rango. Cuando el número k de tratamientos o el número b de bloques sea mayor a cinco, la distribución muestral de F_r puede ser aproximada por una distribución ji cuadrada con $(k-1)$ gl. Por tanto, al igual que para la prueba H de Kruskal-Wallis, la región de rechazo para la prueba F_r está formada por valores de F_r para los cuales

$$F_r > \chi^2_\alpha$$

EJEMPLO

11.8

Suponga que se desea comparar los tiempos de reacción de las personas expuestas a seis estímulos diferentes. Una medición del tiempo de reacción se obtiene al someter a una persona a un estímulo y luego medir el tiempo hasta que la persona presente alguna reacción especificada. El objetivo del experimento es determinar si existen diferencias en los tiempos de reacción para los estímulos empleados en el experimento. Para eliminar la variación de una persona a otra en el tiempo de reacción, cuatro personas participaron en el experimento y se registró el tiempo de reacción (en segundos) de cada una ante cada estímulo. Los datos se dan en la tabla 11.10 (los rangos de las observaciones se muestran entre paréntesis). Use la prueba F_r de Friedman para determinar si los datos presentan suficiente evidencia para indicar diferencias en las distribuciones de tiempos de reacción para los seis estímulos. Pruebe usando $\alpha = .05$.

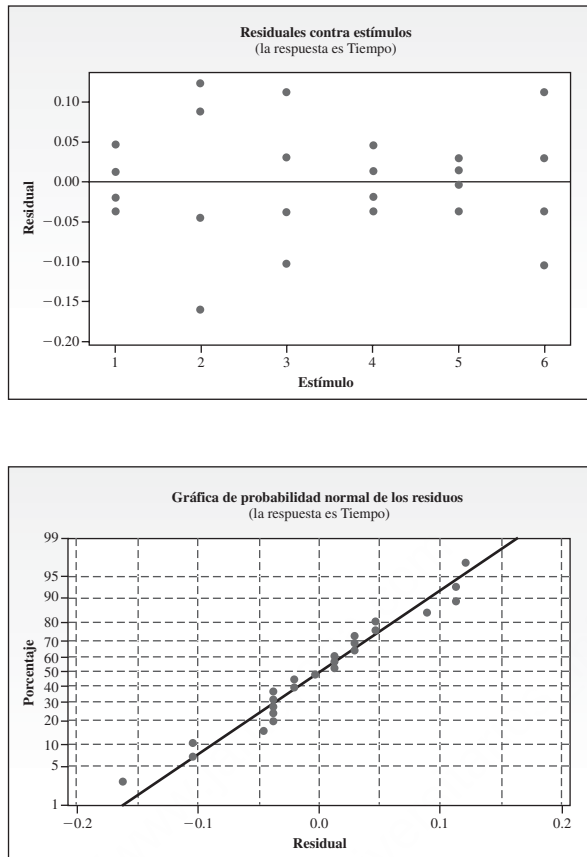
TABLA 11.10 **Tiempos de reacción a seis estímulos**

Persona	Estímulos					
	1	2	3	4	5	6
1	.6 (2.5)	.9 (6)	.8 (5)	.7 (4)	.5 (1)	.6 (2.5)
2	.7 (3.5)	1.1 (6)	.7 (3.5)	.8 (5)	.5 (1.5)	.5 (1.5)
3	.9 (3)	1.3 (6)	1.0 (4.5)	1.0 (4.5)	.7 (1)	.8 (2)
4	.5 (2)	.7 (5)	.8 (6)	.6 (3.5)	.4 (1)	.6 (3.5)
Suma de rango	$T_1 = 11$	$T_2 = 23$	$T_3 = 19$	$T_4 = 17$	$T_5 = 4.5$	$T_6 = 9.5$

Solución En la figura 11.9, la gráfica de los residuales para cada uno de los seis estímulos deja ver que los estímulos 1, 4 y 5 tienen varianzas un poco menores que los otros. Además, la gráfica de probabilidad normal de los residuales revela un cambio en la pendiente de la recta que sigue a los primeros tres residuales, así como una curvatura en la parte superior de la gráfica. Parece que un análisis no paramétrico es apropiado para estos datos.

FIGURA 11.9

Una gráfica de tratamientos contra residuales y una gráfica de probabilidad normal de residuales para el ejemplo 11.8



Se desea probar

H_0 : las distribuciones de tiempos de reacción para los seis estímulos son idénticas

contra la hipótesis alternativa

H_a : Al menos dos de las distribuciones de tiempos de reacción para los seis estímulos difieren en ubicación

La tabla 11.10 muestra los rangos (entre paréntesis) de las observaciones dentro de cada bloque y las sumas de rango para cada uno de los seis estímulos (los tratamientos). El valor del estadístico F_r para estos datos es

$$\begin{aligned}
 F_r &= \frac{12}{bk(k+1)} \sum T_i^2 - 3b(k+1) \\
 &= \frac{12}{(4)(6)(7)} [(11)^2 + (23)^2 + (19)^2 + \cdots + (9.5)^2] - 3(4)(7) \\
 &= 100.75 - 84 = 16.75
 \end{aligned}$$

Como el número $k = 6$ de tratamientos excede de cinco, la distribución de muestreo de F_r puede ser aproximada por una distribución ji cuadrada con $(k-1) = (6-1) = 5$ gl. Por tanto, para $\alpha = .05$, se puede rechazar H_0 si

$$F_r > \chi_{.05}^2 \quad \text{donde} \quad \chi_{.05}^2 = 11.0705$$

Esta región de rechazo se muestra en la figura 11.10. Como el valor observado $F_r = 16.75$ excede de $\chi^2_{.05} = 11.0705$, cae en la región de rechazo. Por tanto, se puede desechar H_0 y concluir que las distribuciones de tiempos de reacción difieren en ubicación para al menos dos estímulos. La salida impresa MINITAB de la prueba F_r de Friedman para los datos se da en la figura 11.11.

FIGURA 11.10
Región de rechazo para el ejemplo 11.8

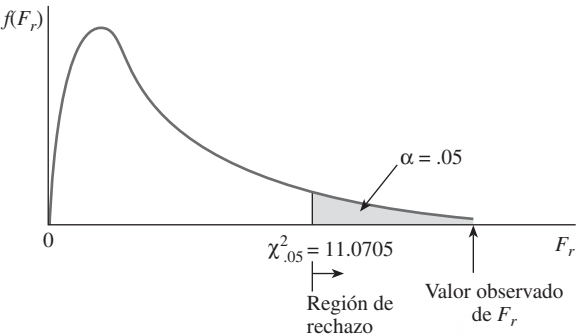


FIGURA 11.11
Salida impresa MINITAB para el ejemplo 11.8

Prueba de Friedman: tiempo contra estímulo en bloques por persona

S = 16.75 DF = 5 P = 0.005
S = 17.37 DF = 5 P = 0.004 (adjusted for ties)

Stimulus	N	Est	Sum of
		Median	Ranks
1	4	0.6500	11.0
2	4	1.0000	23.0
3	4	0.8000	19.0
4	4	0.7500	17.0
5	4	0.5000	4.5
6	4	0.6000	9.5
Grand median		=	0.7167

EJEMPLO

11.9

Encuentre el valor p aproximado para la prueba del ejemplo 11.8.

Solución Si se consulta la tabla 5 del Anexo con 5 gl, se encuentra que el valor observado de $F_r = 16.75$ excede del valor de la tabla $\chi^2_{.005} = 16.7496$. Por tanto, el valor p es muy cercano a .005, pero ligeramente menor a éste.

PRUEBA F_r DE FRIEDMAN PARA UN DISEÑO ALEATORIZADO DE BLOQUES

- 1. Hipótesis nula: H_0 : las k distribuciones poblacionales son idénticas
- 2. Hipótesis alternativa: H_a : al menos dos de las k distribuciones poblacionales difieren en ubicación
- 3. Estadístico de prueba: $F_r = \frac{12}{bk(k + 1)} \sum T_i^2 - 3b(k + 1)$

donde

b = Número de bloques

k = Número de tratamientos

T_i = Suma de rango para el tratamiento i , $i = 1, 2, \dots, k$

4. Región de rechazo: $F_r > \chi_{\alpha}^2$, donde χ_{α}^2 está basada en $(k-1) gl$

Suposición: o bien el número k de tratamientos o el número b de bloques es mayor a cinco.

11.8

PRUEBA DE NEMENYI

Cuando la prueba de Friedman (χ_r^2) resulta significativa, es necesario efectuar comparaciones por pares de los k niveles de la variable independiente (tratamiento) para conocer cuál es el más efectivo.

Existen ${}_k C_2$ comparaciones posibles (pares) entre la medida de los rangos de los tratamientos.

$$\bar{R}_1 - \bar{R}_2, \bar{R}_1 - \bar{R}_3, \dots$$

A continuación se obtiene la diferencia mínima (DM) por medio de:

$$DM = \sqrt{\frac{\chi_{(gl, \alpha)}^2 [k(k+1)]}{6n}}$$

donde $\chi_{gl, \alpha}^2$ es el valor crítico de tablas cuando χ_r^2 es significativa, k es el número de tratamientos, n de sujetos y el 6 es una constante. Cuando la diferencia entre dos tratamientos (representada por la diferencia de sus rangos promediados) es mayor que la DM, entonces el tratamiento que hace la diferencia es el más efectivo.

En el ejemplo anterior se tiene:

Paso 1. Calcule todas las parejas posibles de los promedios de tratamiento, así como sus valores.

$${}_k C_2 = {}_3 C_2 = \frac{3!}{2!1!} = 3$$

$$\bar{R}_B - \bar{R}_A, \bar{R}_M - \bar{R}_A \text{ y } \bar{R}_B - \bar{R}_M$$

$$\bar{R}_A = \frac{\sum R_A}{n} = \frac{14}{10} = 1.40$$

$$\bar{R}_B - \bar{R}_A = 2.80 - 1.40 = 1.40$$

$$\bar{R}_B = \frac{\sum R_B}{n} = \frac{28}{10} = 2.80$$

$$\bar{R}_M - \bar{R}_A = 1.80 - 1.40 = 0.40$$

$$\bar{R}_M = \frac{\sum R_M}{n} = \frac{18}{10} = 1.80$$

$$\bar{R}_B - \bar{R}_M = 2.80 - 1.80 = 1.0$$

Paso 2. Se obtiene la DM, para $\chi^2_{(2, 0.05)} = 5.991$

$$\begin{aligned} DM &= \sqrt{\frac{X^2_{(gl, \alpha)}[k(k + 1)]}{6n}} \\ DM &= \sqrt{\frac{5.991[3 \times 4]}{6 \times 10}} \\ DM &= \sqrt{\frac{5.991 \times 2}{10}} = \sqrt{1.1982} \end{aligned}$$

DM = 1.09462

Paso 3. Como $1.4 > 1.09462$, que corresponde a la pareja $\bar{R}_B - \bar{R}_A$; por consiguiente, el nivel $\bar{R}_B$ es el que hace la diferencia estadísticamente significativa. Para $\chi^2_{(2, 0.01)} = 9.210$, la diferencia mínima (DM) = 1.3572 es menor que 1.4, por lo que RB también lo cumple al 1%.

11.8 EJERCICIOS

TÉCNICAS BÁSICAS

DATOS **11.37** Se usa un diseño aleatorizado de bloques para comparar tres tratamientos en seis bloques.

Bloque	Tratamiento		
	1	2	3
1	3.2	3.1	2.4
2	2.8	3.0	1.7
3	4.5	5.0	3.9
4	2.5	2.7	2.6
5	3.7	4.1	3.5
6	2.4	2.4	2.0

- a. Use la prueba F_r de Friedman para detectar diferencias en ubicación entre las tres distribuciones de tratamiento. Pruebe usando $\alpha = .05$.
- b. Encuentre el valor p aproximado para la prueba de la parte a.
- c. Efectúe un análisis de varianza y dé la tabla ANOVA para el análisis.
- d. Dé el valor del estadístico F para probar la igualdad de las tres medias de tratamiento.
- e. Dé el valor p aproximado para el estadístico F de la parte d.
- f. Compare los valores p para las pruebas de las partes a y d, y explique las implicaciones prácticas de la comparación.

DATOS **11.38** Un diseño aleatorizado de bloques se usa para comparar cuatro tratamientos en ocho bloques.

Bloque	Tratamiento			
	1	2	3	4
1	89	81	84	85
2	93	86	86	88
3	91	85	87	86
4	85	79	80	82
5	90	84	85	85
6	86	78	83	84
7	87	80	83	82
8	93	86	88	90

- a. Use la prueba F_r de Friedman para detectar diferencias en ubicación entre las cuatro distribuciones de tratamiento. Pruebe usando $\alpha = .05$.
- b. Encuentre el valor p aproximado para la prueba de la parte a.
- c. Efectúe un análisis de varianza y dé la tabla ANOVA para el análisis.
- d. Dé el valor del estadístico F para probar la igualdad de las cuatro medias de tratamiento.
- e. Dé el valor p aproximado para el estadístico F de la parte d.

- f. Compare los valores p para las pruebas de las partes **a** y **d**, y explique las implicaciones prácticas de la comparación.

APLICACIONES

DATOS
EX1139 **11.39 Precios de supermercado** En el ejercicio 7.43 (y en el conjunto de datos EX0743) se compararon los precios regulares en cuatro tiendas de abarrotes diferentes para ocho artículos comprados en el mismo día. Los precios se detallan en la tabla.

Artículos	Tienda			
	Vons	Ralphs	Stater Bros	WinCo
Mezcla para ensalada, bolsa de 12 oz.	3.99	2.79	1.99	1.78
Salchichas ahumadas de res Hillshire Farm®, 14 oz.	4.29	4.29	3.99	2.50
Raisin Bran® de Kellogg, 25.5 oz.	4.49	5.49	4.49	3.15
Queso crema Philadelphia® de Kraft®, 8 oz.	2.99	3.19	2.79	1.48
Aderezo Ranch de Kraft®, 16 oz.	3.19	3.49	3.49	1.48
Jugo de manzana TreeTop®, 64 oz.	2.99	3.49	3.49	1.58
Jabón de tocador Dial® Gold, 8-4 oz.	5.99	6.49	5.79	5.14
Crema de cacahuete Jif®, cremosa, 28 oz.	5.15	5.49	4.79	4.34

- a. ¿La distribución de los precios difiere de un supermercado a otro? Pruebe usando la prueba F_r de Friedman con $\alpha = .05$.
- b. Encuentre el valor p aproximado para la prueba e interprételo.

DATOS
EX1140 **11.40 Productos químicos tóxicos** Se realizó un experimento para comparar los efectos de tres productos químicos tóxicos, A, B y C, en la piel de ratas. Cuadros de una pulgada de piel fueron tratados con los productos y luego calificados del 0 al 10, dependiendo del grado de irritación. Tres cuadros adyacentes de 1 pulgada fueron marcados en los lomos de ocho ratas y cada uno de los tres productos se aplicó a cada rata. Así, el experimento se realizó en bloques sobre ratas para eliminar la variación en sensibilidad de la piel de una rata a otra.

Ratas							
1	2	3	4	5	6	7	8
B	A	A	C	B	C	C	B
5	9	6	6	8	5	5	7
A	C	B	B	C	A	B	A
6	4	9	8	8	5	7	6
C	B	C	A	A	B	A	C
3	9	3	5	7	7	6	7

- a. ¿Los datos dan suficiente evidencia para indicar una diferencia en los efectos tóxicos de los tres productos químicos? Pruebe usando la prueba F_r de Friedman con $\alpha = .05$.

- b. Encuentre el valor p aproximado para la prueba e interprételo.

DATOS
EX1141 **11.41 Medicina que sabe bien** En un estudio del sabor de antibióticos en niños, la doctora Doreen Matsui y colegas emplearon una muestra de niños sanos voluntarios para evaluar sus reacciones al sabor de cuatro antibióticos.⁴ Las respuestas de los niños se midieron en una escala analógica visual de 10 centímetros (cm) que incorporaba el uso de rostros, de tristes (baja calificación) a alegres (alta calificación). La calificación mínima era 0 y la máxima era 10. Para los datos siguientes (simulados de los resultados del informe de Matsui), a cada uno de cinco niños se le pidió que probara cada uno de los cuatro antibióticos y los calificara usando la escala analógica visual (rostros) de 0 a 10 cm.

Niño	Antibiótico			
	1	2	3	4
1	4.8	2.2	6.8	6.2
2	8.1	9.2	6.6	9.6
3	5.0	2.6	3.6	6.5
4	7.9	9.4	5.3	8.5
5	3.9	7.4	2.1	2.0

- a. ¿Qué diseño se usa para recolectar estos datos?
- b. Usando el paquete estadístico apropiado para una clasificación de dos vías, elabore una gráfica de probabilidad normal de los residuales así como una gráfica de residuales contra antibióticos. ¿Las suposiciones usuales del análisis de varianza parecen quedar satisfechas?
- c. Use la prueba no paramétrica apropiada para probar si hay diferencias en las distribuciones de respuestas a los sabores de los cuatro antibióticos.
- d. Comente sobre los resultados del análisis de varianza de la parte **b** comparados con la prueba no paramétrica de la parte **c**.

PRUEBA DE LA MEDIANA

Cuando es necesario comparar dos o más poblaciones independientes y dado que no hay normalidad no es posible utilizar las medias aritméticas, existe la opción de emplear la mediana, otra medida de tendencia central sobre todo con distribuciones asimétricas.

A fin de aplicar esta prueba, que contrastaría la hipótesis nula de que la mediana poblacional es la misma para todas las poblaciones que intervienen, resulta indispensable que dichas poblaciones tengan una distribución continua y que las muestras que las representan se hayan seleccionado de manera aleatoria.

Al aplicar los ocho pasos, resulta:

Paso 1. Se agrupan los datos de los k grupos en uno solo y se calcula la mediana (Me).

Paso 2. Se proponen las siguientes hipótesis:

$$H_0: Me_1 = Me_2 = \dots = Me_k = Me_0$$

H_1 : Al menos dos poblaciones tienen mediana diferente.

Paso 3. Se establece el valor de α .

Paso 4. Una vez obtenida la mediana, la frecuencia de los datos se expone en una tabla de contingencia de la siguiente manera:

	Muestras (k)				Total
Datos menores o iguales a la Me	f_{o11}	f_{o11}	...	f_{o1k}	a
Datos mayores que la Me	f_{o21}	f_{o22}	...	f_{o2k}	b
Total	n_1	n_2	...	n_k	N

Se aplica el modelo estadístico χ^2

$$\chi^2 = \sum \frac{(f_{oi} - f_{ei})^2}{f_{ei}}$$

Donde $f_{ei} = \frac{n_i}{2}$

Paso 5. Se obtienen los grados de libertad, considerando el número de renglones y de columnas de la tabla de contingencia, y se lee el valor crítico de χ^2 , según α y gl .

Paso 6. Se considera la siguiente regla de decisión: si $\chi^2_{calc} \geq \chi^2_{(gl, \alpha)}$, entonces H_0 se rechaza.

Paso 7. Se calcula el valor de χ^2 .

Paso 8. Si no se rechaza H_0 , entonces las poblaciones anteriores tienen la misma mediana. Cuando son dos poblaciones, en el paso 4 puede utilizarse el siguiente modelo:

$$\chi^2 = \frac{N \left[\left| ad - bc \right| - \frac{N}{2} \right]^2}{klmn}^\dagger$$

[†] Donde k, l, m, n son los totales marginales.

El correspondiente cuadro de contingencia sería el siguiente:

	Población		Total
	1	2	
Mayor o igual a la Me	a	b	k
Menor que la Me	c	d	l
Total	m	n	N

EJEMPLO

11.10

A dos grupos de estudiantes de la misma escuela y grado (referidos como $n_1 = 20$ y $n_2 = 17$), se les enseña una técnica odontológica con dos métodos diferentes (I y II). Ambos grupos se evalúan y se obtienen las siguientes puntuaciones:

n_1	n_2
Método I	Método II
48 27	54 36
47 27	32
46 26	50 30
46 25	48 25
44 24	47 25
40 18	46 22
38 18	45 20
32 16	45 14
30 12	43 42
28 10	$n_2 = 17$
$n_1 = 20$	

Paso 1. Se combinan todas las puntuaciones en una sola distribución y se ordenan en forma creciente o decreciente.

$$n = n_1 + n_2 = 20 + 17 = 37$$

Paso 2. Se calcula la mediana a través de la fórmula de datos no agrupados. En el ejemplo: $Me = 32$.

Paso 3. Para cada distribución (métodos I y II), se cuentan los casos que se sitúan en el rango superior e inferior de la mediana, y se presentan los valores en un cuadro de contingencia.

	Método I	Método II	Total
Arriba de la mediana	8 (a)	11 (b)	19
Debajo de la mediana	12 (c)	6 (d)	18
	20	17	37

Paso 4. Se calcula χ_c^2 mediante la siguiente fórmula:

$$\chi_c^2 = \frac{n [|ad - bc| - 0.5]^2}{(a + b)(a + c)(b + d)(c + d)}$$

$$\begin{aligned} &= \frac{37 \left[|(8)(6) - (11)(12)| - 0.5 \right]^2}{(19)(20)(17)(18)} \\ &= \frac{37 \left[|48 - 132| - 0.5 \right]^2}{116\ 280} = \frac{257\ 973.25}{116\ 280} = 2.22 \end{aligned}$$

$\therefore \chi_c^2 = 2.22$

Paso 5. Se compara este valor con el observado en la tabla de χ^2 con $gl = 1$, y a un nivel de significación de 5%. Como $2.22 < 3.841$, se rechaza la hipótesis alternativa en el sentido de que las dos medianas provienen de poblaciones diferentes.

Solución Como la χ^2 no es significativa (ni siquiera en el nivel de 10%, donde $\chi_{i.}^2 = 2.706$, no se rechaza la hipótesis de que las medianas proceden de la misma población. En otros términos, ambos métodos de enseñanza son iguales. Esta prueba puede aplicarse a más de dos grupos al utilizar tablas de contingencia de tipo 2×2 .

Cuando el número de frecuencias esperadas es pequeño (menor que 2), no es conveniente utilizar la χ^2 , sino la prueba exacta de Fisher. Es necesario considerar que la potencia de la prueba de la mediana es baja, por lo que se desarrolla una prueba más potente.

11.10

COEFICIENTE DE CORRELACIÓN DE RANGO

En las secciones precedentes usamos rangos para indicar la magnitud relativa de observaciones en pruebas no paramétricas para la comparación de tratamientos. A continuación usaremos la misma técnica para probar una relación entre dos variables ordenadas. Dos coeficientes comunes de correlación de rango son **r_s de Spearman** y **τ de Kendall**. Presentaremos el r_s de Spearman porque su cálculo es idéntico al del coeficiente de correlación muestral r del capítulo 8.

Suponga que ocho profesores de ciencias de escuela básica han sido clasificados por un juez de acuerdo con su capacidad de enseñanza y todos han presentado un “examen nacional para maestros”. Los datos se dan en la tabla 11.11. ¿Los datos sugieren un acuerdo entre la clasificación del juez y la calificación del examen? Esto es, ¿hay una correlación entre rangos y calificaciones de examen?

TABLA 11.11 Clasificaciones y calificaciones de examen para ocho profesores

Profesor	Clasificación del juez	Calificación de examen
1	7	44
2	4	72
3	2	69
4	6	70
5	1	93
6	3	82
7	8	67
8	5	80

Las dos variables de interés son clasificación del juez y calificación de examen. La primera ya está en forma de rango y las calificaciones se pueden ordenar de manera similar, como se ve en la tabla 11.12. Los rangos para las observaciones empatadas se obtienen al promediar los rangos que habrían tenido las observaciones empatadas si no se hubieran observado empates. El coeficiente de correlación de rango de Spearman, r_s , se calcula usando los rangos de las mediciones en pares en las dos variables x y y en la fórmula para r (véase el capítulo 8).

TABLA 11.12 Rangos de datos de la tabla 11.11

Profesor	Clasificación del juez, x_i	Calificación de examen, y_i
1	7	1
2	4	5
3	2	3
4	6	4
5	1	8
6	3	7
7	8	2
8	5	6

COEFICIENTE DE CORRELACIÓN DE RANGO DE SPEARMAN

$$r_s = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

donde x_i y y_i representan los rangos del i -ésimo par de observaciones y

$$S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}$$

$$S_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

$$S_{yy} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$$

Cuando no haya empates en las x observaciones o las y observaciones, la expresión para r_s algebraicamente se reduce a la expresión más sencilla

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad \text{donde } d_i = (x_i - y_i)$$

Si el número de empates es pequeño en comparación con el número de pares de datos, se produce poco error al usar esta fórmula breve.

EJEMPLO

11.11

Calcule r_s para los datos de la tabla 11.12.

Solución Las diferencias y cuadrados de diferencias entre los dos rangos se dan en la tabla 11.13. Sustituyendo valores en la fórmula para r_s , tenemos

$$\begin{aligned} r_s &= 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \\ &= 1 - \frac{6(144)}{8(64 - 1)} = -.714 \end{aligned}$$

TABLA 11.13 Diferencias y cuadrados de diferencias para los rangos del profesor

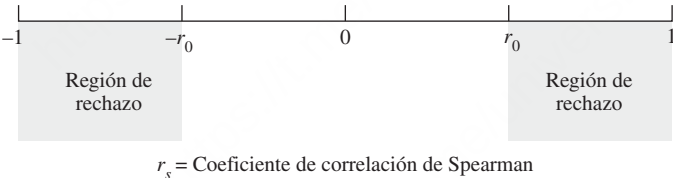
Profesor	x_i	y_i	d_i	d_i^2
1	7	1	6	36
2	4	5	-1	1
3	2	3	-1	1
4	6	4	2	4
5	1	8	-7	49
6	3	7	-4	16
7	8	2	6	36
8	5	6	-1	1
Total				144

El coeficiente de correlación de rango de Spearman se usa como estadístico de prueba para verificar la hipótesis de que no hay asociación entre dos poblaciones. Se puede suponer que los n pares de observaciones (x_i, y_i) se han seleccionado al azar y, por tanto, ninguna asociación entre las poblaciones implica una asignación aleatoria de los n rangos dentro de cada muestra. Cada asignación aleatoria (para las dos muestras) representa un evento sencillo asociado con el experimento, y un valor de r_s se calcula para cada una. Entonces, es posible calcular la probabilidad de que r_s tome un valor absoluto grande debido sólo a una casualidad y, por tanto, sugiere una asociación entre poblaciones cuando no existe ninguna.

La región de rechazo para una prueba de dos colas se muestra en la figura 11.12. Si la hipótesis alternativa es que la correlación entre x y y sea negativa, se rechazaría H_0 para valores negativos de r_s cercanos a -1 (en la cola inferior de la figura 11.12). Del mismo modo, si la hipótesis alternativa es que la correlación entre x y y es positiva, se rechazaría H_0 para valores positivos grandes de r_s (en la cola superior de la figura 11.12).

FIGURA 11.12

Región de rechazo para una prueba de dos colas de la hipótesis nula de no asociación, usando la prueba de correlación de rango de Spearman



Los valores críticos de r_s se dan en la tabla 9 del Anexo. Una versión abreviada se muestra en la tabla 11.14. En sentido horizontal en la tabla 11.14 (y la tabla 9 del Anexo) están los valores registrados de α que podrían usarse para una prueba de una cola de la hipótesis nula de no asociación entre x y y . El número de pares de rango n aparece en el lado izquierdo de la tabla. Las entradas de la tabla dan el valor crítico r_0 para una prueba de una cola. Entonces, $P(r_s \geq r_0) = \alpha$.

Por ejemplo, suponga que tenemos $n = 8$ pares de rango y la hipótesis alternativa es que la correlación entre los rangos es positiva. Se desearía rechazar la hipótesis nula de no asociación sólo para los valores positivos grandes de r_s y se usaría una prueba de una cola. Si se consulta la tabla 11.14 y se usa el renglón correspondiente a $n = 8$ y la columna para $\alpha = .05$, se lee $r_0 = .643$. Por tanto, se puede rechazar H_0 para todos los valores de r_s mayores o iguales a $.643$.

La prueba se realiza exactamente en la misma forma si se desea probar sólo la hipótesis alternativa de que los rangos tienen correlación negativa. La única diferencia es que se rechazaría la hipótesis nula si $r_s \leq -.643$. Esto es, se usa el negativo del valor tabulado de r_0 para obtener el valor crítico de cola inferior.

TABLA 11.14

Versión abreviada de la tabla 9 del Anexo; para la prueba de correlación de rango de Spearman

n	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$
5	.900	—	—	—
6	.829	.886	.943	—
7	.714	.786	.893	—
8	.643	.738	.833	.881
9	.600	.683	.783	.833
10	.564	.648	.745	.794
11	.523	.623	.736	.818
12	.497	.591	.703	.780
13	.475	.566	.673	.745
14	.457	.545		
15	.441	.525		
16	.425			
17	.412			
18	.399			
19	.388			
20	.377			

Para efectuar una prueba de dos colas se rechaza la hipótesis nula si $r_s \geq r_0$ o $r_s \leq -r_0$. El valor de α para la prueba es el doble del valor que se ve en la parte superior de la tabla. Por ejemplo, si $n = 8$ y se elige la columna .025, se rechaza H_0 si $r_s \geq .738$ o $r_s \leq -.738$. El valor α para la prueba es $2(.025) = .05$.

PRUEBA DE CORRELACIÓN DE RANGO DE SPEARMAN

1. Hipótesis nula: H_0 : no hay asociación entre los pares de rangos.
2. Hipótesis alternativa: H_a : hay asociación entre los pares de rangos (una prueba de dos colas); o bien, H_a : la correlación entre los pares de rangos es positiva o negativa (una prueba de una cola).

3. Estadístico de prueba:
$$r_s = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

donde x_i y y_i representan las filas del i -ésimo par de observaciones.

4. Región de rechazo: para una prueba de dos colas, rechace H_0 si $r_s \geq r_0$ o $r_s \leq -r_0$, donde r_0 se da en la tabla 9 del Anexo. Duplique la probabilidad tabulada para obtener el valor de α para la prueba de dos colas. Para una prueba de una cola, rechace H_0 si $r_s \geq r_0$ (para una prueba de cola superior) o $r_s \leq -r_0$ (para una prueba de cola inferior). El valor α para una prueba de una cola es el valor que se muestra en la tabla 9 del Anexo.

EJEMPLO

11.12

Pruebe la hipótesis de que no hay asociación entre las poblaciones para el ejemplo 11.11.

Solución El valor crítico de r_s para una prueba de una cola con $\alpha = .05$ y $n = 8$ es .643. Se puede suponer que una correlación entre la clasificación del juez y las calificaciones de examen de profesores no podrían ser posiblemente positivas. (Una clasificación baja significa buena enseñanza y debe estar asociada con una alta calificación de examen si el juez y la prueba miden la capacidad de enseñanza.) La hipótesis alternativa es que el **coeficiente de correlación de rango poblacional** ρ_s es menor que 0 y nos interesa una prueba estadística de una cola. Entonces, α para la prueba es el valor tabulado para .05 y se puede rechazar la hipótesis nula si $r_s \leq -.643$.

El valor calculado del estadístico de prueba, $r_s = -.714$, es menor que el valor crítico para $\alpha = .05$. En consecuencia, la hipótesis nula se rechaza al nivel de significancia $\alpha = .05$. Es evidente que existe algún acuerdo entre las clasificaciones del juez y las calificaciones del examen. No obstante, debe observarse que este acuerdo podría existir cuando *ninguna* da una medida adecuada para medir la capacidad de enseñanza. Por ejemplo, la asociación podría existir si el juez y quienes formularon el examen de profesores tuvieran un concepto completamente erróneo, pero semejante, de las características de la buena enseñanza.

¿Qué es exactamente lo que mide r_s ? El coeficiente de correlación de Spearman detecta no sólo una relación lineal entre dos variables, sino que también mide cualquier otra relación monotónica (y aumenta cuando x aumenta o y disminuye cuando x aumenta). Por ejemplo, si se calculó r_s para los dos conjuntos de datos de la tabla 11.15, ambos producirían un valor de $r_s = 1$ porque los rangos asignados para x y y en ambos casos concuerdan para todos los pares (x, y) . Es importante recordar que un valor significativo de r_s indica una relación entre x y y que es creciente o decreciente, pero no necesariamente lineal.

TABLA 11.15

Conjuntos de datos gemelos con $r_s = 1$

x	$y = x^2$	x	$y = \log_{10}(x)$
1	1	10	1
2	4	100	2
3	9	1000	3
4	16	10,000	4
5	25	100,000	5
6	36	1,000,000	6

11.11

PRUEBA DE SIGNIFICANCIA DE r_s

Cuando el tamaño de la muestra es menor que 10, se consultará la tabla de Spearman del Anexo (al final del libro), ya que si n es pequeña, el valor de r_s debe ser muy grande para que sea significativo. Si $n > 10$, se podrá utilizarse la siguiente fórmula (al igual que en el caso del coeficiente de correlación de Pearson):

$$t = \frac{r_s \sqrt{n-2}}{\sqrt{1-r_s^2}}$$

Al sustituir los valores anteriores se obtiene:

$$\begin{aligned} t &= \frac{0.74 \sqrt{10-2}}{\sqrt{1-(0.74)^2}} \\ &= \frac{0.74 \sqrt{8}}{\sqrt{1-0.5476}} \\ &= \frac{2.093}{0.6726} = 3.11 \\ \therefore t &= 3.11 \end{aligned}$$

Este resultado se contrasta con la tabla del Anexo de la razón t de Student, y se concluye lo siguiente:

$$gl = n - 2 = 10 - 2 = 8; \alpha = 0.05$$

Para una prueba bilateral, la t crítica será:

$$t_{(g,0.05)} = 2.36$$

Por consiguiente, r_s es significativo, ya que la regla de decisión es: si $t \geq t_{crít}$, H_0 se rechaza, y como $3.11 > 2.36$, la hipótesis nula (H_0) se descarta.

Se concluye que

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

donde ρ significa la correlación no paramétrica en la población.

11.12

COEFICIENTE TAU (τ) DE KENDALL

Esta medida de correlación se basa en intervalos jerarquizados de las observaciones, más que en los números mismos; lo anterior tiene la ventaja de que la distribución de dicho coeficiente no depende de la que presentan x y y , siempre que las observaciones representadas por x y y sean independientes y continuas. Algunos investigadores prefieren este coeficiente, desarrollado por Kendall (1938), al de Spearman, aunque es ligeramente más difícil de calcular que r_s , en especial cuando existen empates. No obstante esto, el coeficiente de Kendall cuenta con una gran ventaja: su distribución tiende a la normal más rápidamente que el de Spearman. Este “ajuste” a la distribución normal es mejor para el τ de Kendall, siempre y cuando la hipótesis nula (H_0) de independencia entre x y y sea cierta.

EJEMPLO

11.13

Se cataloga a un grupo de nueve niños como los más agresivos en el rango de 5 años de edad de la guardería donde acuden con regularidad. Para confirmar el grado de agresividad, se aplicó a los infantes una prueba: durante una semana se efectuaron registros observacionales, diarios y de acuerdo con ciertas condiciones. Los resultados promedio fueron los siguientes: registros observacionales en la guardería (RO_y), registros observacionales en los hogares (RO_x); rangos o intervalos en el hogar (R_x), así como rangos o intervalos en la guardería (R_y). Dichos registros se presentan en la tabla siguiente:

Niños	RO_x	RO_y	R_x	R_y
A	84	60	1	4
B	80	64	2	2
C	78	71	3	1
D	76	61	4	3
E	70	58	5	5
F	64	57	6	6
G	62	54	7	8
H	50	55	8	7
I	47	52	9	9

A continuación se calcula el coeficiente de correlación de Kendall (τ) entre lo detectado por los padres de los niños y lo que se advierte en la guardería.

- Paso 1.** Cada distribución de puntuaciones, que representa a la variable x o a la y , se jerarquiza casi de la misma manera que cuando se calcula el coeficiente de Spearman para obtener R_x y R_y ; la modificación consiste en que un conjunto de rangos (x o y) debe estar ordenado en una secuencia natural y creciente. El objetivo de este paso es tener una referencia que se utilizará posteriormente.
- Paso 2.** Se obtiene la columna de rangos más altos (P) y la de rangos menores (Q) que tengan como referencia la columna R_y . Con ese fin se procede así: se considera el valor numérico correspondiente al primer niño (4, en la columna R_y del ejemplo) y se cuentan hacia abajo los valores numéricos menores que él (en este caso son 2, 1 y 3); a continuación, se cuentan los valores que son mayores que él (5, 6, 8, 7 y 9).

R_x	R_y		Cantidad de rangos más altos	Cantidad de rangos más bajos
1	4	Primer sujeto	5	3
2	2			
3	1	Tres rangos		
4	3	más bajos que el primer sujeto		
5	5			
6	6			
7	8	Cinco rangos		
8	7	más altos que el primer sujeto		
9	9			

El procedimiento se aplica de nuevo para obtener el segundo sujeto (2 en el ejemplo); así, hay un sujeto más bajo que el 2 (1) y seis más altos (3, 5, 6, 8, 7 y 9).

R_x	R_y		Cantidad de rangos más altos	Cantidad de rangos más bajos
1	4		5	3
2	2	Segundo sujeto	6	1
3	1	Un rango		
4	3	más bajo que el primer sujeto		
5	5			
6	6			
7	8	Seis rangos		
8	7	más altos que el segundo sujeto		
9	9			

Con el fin de determinar el tercer sujeto (1), se excluyen los sujetos anteriores a él y se cuenta hacia abajo cuántos son menores y cuántos mayores. En este caso, no existe valor numérico menor; los más altos son seis (3, 5, 6, 8, 7 y 9).

R_x	R_y		Cantidad de rangos más altos	Cantidad de rangos más bajos
1	4		5	3
2	2		6	1
3	1	Tercer sujeto (cero rangos más bajos que él)	6	0
4	3			
5	5			
6	6			
7	8			
8	7	Seis rangos más altos que el tercer sujeto		
9	9			

El método se aplicará sucesivamente hasta llegar al último sujeto, que siempre tendrá cero rangos más altos y cero más bajos.

Paso 3. Una vez determinadas las columnas anteriores, se obtienen la sumatoria correspondiente a la columna de rangos más altos (P), así como la relativa a la columna de rangos más bajos (Q).

Sujetos	RO_x	RO_y	R_x	R_y	Rangos más altos P	Rangos más bajos Q
A	84	60	1	4	5	3
B	80	64	2	2	6	1
C	78	71	3	1	6	0
D	76	61	4	3	5	0
E	70	58	5	5	4	0
F	64	57	6	6	3	0
G	62	54	7	8	1	1
H	50	55	8	7	1	0
I	47	52	9	9	0	0
					31	5

Paso 4. El resultado se sustituye en la fórmula *tau* de Kendall.

$$\tau = \frac{P - Q}{\frac{n(n - 1)}{2}}$$

donde:

n = número de casos o sujetos; en el ejemplo, $n = 9$

P = suma de rangos más altos ($P = 31$)

Q = suma de rangos más bajos ($Q = 5$)

Así:

$$\tau = \frac{P - Q}{\frac{n(n - 1)}{2}}$$

$$\begin{aligned} &= \frac{31-5}{\frac{72}{2}} \\ &= \frac{26}{36} = 0.72 \end{aligned}$$

$\therefore \tau = 0.72$

COEFICIENTE DE CONCORDANCIA (ω) DE KENDALL

11.13

Es posible considerar este coeficiente como el promedio de un grupo de coeficientes de Spearman; o sea, ω es una medida del grado de acuerdo o concordancia entre m conjuntos de n rangos. En otros términos, para un grupo de n objetos que m jueces jerarquizan o evalúan, ω provee información sobre el grado de acuerdo existente entre m conjuntos de rangos o jerarquizaciones que otorgan los jueces.

Una diferencia entre r_s (Spearman) y ω (Kendall) es que éste siempre será positivo, por lo que su intervalo de valores fluctúa de 0 a 1. De esa manera, si la evaluación que otorga cada juez a los n objetos es la misma, entonces $\omega = 1.0$; en cambio, si existe un total desacuerdo entre ellos, $\omega = 0$.

Sin demérito de lo anterior, hay que destacar que lo medido por ω es el acuerdo (concordancia) entre los jueces con respecto a la materia evaluada; por tanto, aunque ω sea grande (muy cercana a 1.0), no siempre significa que los jueces valoren correctamente, ya que pueden coincidir por completo en una evaluación incorrecta en términos de un criterio externo. Tampoco implica que los jueces utilizaron los mismos criterios de evaluación o los mismos estándares, independientemente de si son los más adecuados.

Por otra parte, si los jueces no se ponen de acuerdo ($\omega = 0$), quizá se deba a que los atributos por evaluar son ambiguos o están pobremente definidos, o también a que dichos objetos no difieren significativamente en la característica o atributo medido. En consecuencia, la discriminación no es posible, ni puede ser concordante la opinión de los jueces.

EJEMPLO 11.14

A cuatro catadores especializados ($m = 4$) se les presenta un grupo de cinco vinos diferentes ($n = 5$). La información obtenida se expresa de la siguiente manera:

Vino	Opinión de los catadores	Sumas de los rangos (R)
	1 2 3 4	
1	5 4 5 5	19
2	3 3 2 3	11
3	1 2 1 2	6
4	2 1 3 1	7
5	4 5 4 4	17
		$\Sigma R = 60$

Calcule la concordancia entre las opiniones otorgadas por los cuatro catadores, que son los jueces. La mayor puntuación (5) significa que el vino reúne varias características de calidad, y éstas descienden sucesivamente hasta llegar a 1, donde el vino es de baja calidad.

Paso 1. Las jerarquizaciones otorgadas por los cuatro jueces a cada vino se suman y se colocan en una columna, como se ve en la tabla anterior.

Paso 2. En caso de no existir relación entre la jerarquización otorgada por los jueces ($\omega = 0$), la suma de dichas jerarquizaciones para cada uno de los n objetos (vino) será la misma. Si esto hubiese ocurrido en el ejemplo, esta suma sería 12; o sea, la total entre el número de casos $\left(\frac{60}{5} = 12\right)$.

Paso 3. Se obtiene la diferencia (D) entre la suma de cada grupo de jerarquizaciones y el promedio de 12, luego se colocan estos resultados en la columna 2.

n Vino	m Opinión de los catadores	(1) Sumas de los rangos (R)	(2) Diferencia de rangos
	1 2 3 4	R	$D = \Sigma R - 12$
1	5 4 5 5	19	7
2	3 3 2 3	11	1
3	1 2 1 2	6	6
4	2 1 3 1	7	5
5	4 5 4 4	17	5
		$\Sigma R = 60$	

Paso 4. Se elevan al cuadrado estas diferencias (D), se colocan los resultados en la columna 3 y se obtiene su suma.

n Vino	m Opinión de los catadores	(1) Sumas de los rangos (R)	(2) Diferencia de rangos	(3)
	1 2 3 4	R	D	D^2
1	5 4 5 5	19	7	49
2	3 3 2 3	11	1	1
3	1 2 1 2	6	6	36
4	2 1 3 1	7	5	25
5	4 5 4 4	17	5	25
		$\Sigma R = 60$		$\Sigma D^2 = 136$

Paso 5. Se calcula ω mediante la siguiente fórmula:

$$\omega = \frac{12 \Sigma D^2}{m^2 n (n^2 - 1)}$$

$$\omega = \frac{12(136)}{(4^2)(5)(5^2 - 1)}$$

$$\omega = \frac{1\,632}{(16)(5)(24)} = \frac{1\,632}{1\,920} \therefore \omega = 0.85$$

Con los resultados obtenidos anteriormente, se comprueba que la suma de las jerarquizaciones (ΣR) es igual al producto del número de jueces ($m = 4$) por el de objetos evaluados ($n = 5$) y por $(n + 1)$. Todo lo anterior se divide entre 2; o sea que debe obtenerse $60 = 60$.

Esto es:

$$\Sigma R = \frac{mn(n + 1)}{2}$$

Como $\Sigma R = 60$:

$$60 = \frac{4 \cdot 5 \cdot 6}{2}$$

$$60 = \frac{120}{2}$$

$$60 = 60$$

Prueba de significancia de ω

La significancia estadística de ω se evalúa por medio de la tabla del Anexo. Los datos para consultar esta tabla son $m = 4$ (jueces) y $n = 5$ (vinos) que constituyen los objetos jerarquizados o calificados. Se observa que con un nivel de significancia de 1%, el valor de ω_0 es de 0.67 y, dado que la ω calculada asciende a 0.85 y supera el valor de 0.67 (el de la tabla), puede concluirse que este coeficiente de concordancia es significativo a partir de 1% y, por supuesto, α de 5% ($\omega_0 = 0.54$). Por consiguiente, los catadores realizaron un trabajo correcto al evaluar estos cinco vinos.

En general, la regla de decisión para el coeficiente de Kendall es la siguiente:

Si $\omega \geq \omega_0$, entonces ω es significativo.

Solución $\omega = 0.85$ indica que existe un alto grado de acuerdo (concordancia) entre los cuatro jueces en la evaluación de las características de calidad de los cinco vinos.

11.14

COEFICIENTE DE CORRELACIÓN (r_{bp}) BISERIAL DE PUNTO

La correlación biserial, en general, es una medida de asociación entre dos variables continuas que se distribuyen de esa misma forma, en donde, para facilitar el proceso, una de ellas se ha dicotomizado (sí-no; falso-verdadero; masculino-femenino; casado-soltero; bueno-malo; entre otras). La variable dicotomizada se supone discreta o discontinua cuando trata de relacionarse con la que permanece continua; el coeficiente más apropiado para este propósito es el de correlación biserial puntual (r_{bp}), que se define en la siguiente forma:

$$r_{bp} = \frac{n(\sum fcx) - n_c(\sum fx)}{\sqrt{(n_c)(n_i) [n \sum fx^2 - (\sum fx)^2]}}$$

donde:

n = número de sujetos

f = frecuencia de ocurrencia de las puntuaciones obtenidas por los sujetos

x = puntuaciones obtenidas por los sujetos

fx = producto de la frecuencia por las puntuaciones

fc = número de sujetos que obtuvieron exactamente las puntuaciones x

$n_c = \Sigma fc = (\dots)$ = número de sujetos que no obtuvieron exactamente las puntuaciones x

- f_i = número de sujetos que no obtuvieron exactamente las puntuaciones x
 n_i = Σf_i = número total de sujetos que no obtuvieron las puntuaciones x
 fx = producto del número de sujetos que obtuvieron exactamente las puntuaciones x por la puntuación x de cada sujeto representativo
 Σfx^2 = suma de los productos de la frecuencia absoluta por las puntuaciones x al cuadrado
 $(\Sigma fx)^2$ = cuadrado de la suma del producto de la frecuencia por las puntuaciones x

EJEMPLO**11.15**

A 100 estudiantes de una escuela preparatoria se les aplica una prueba de conocimiento x , que consta de 40 reactivos, y se determina la frecuencia con que se distribuyen las puntuaciones de los 100 estudiantes. Se obtienen los siguientes resultados:

x	f
40	2
38	4
37	6
36	12
32	12
31	10
30	12
28	10
27	10
25	4
24	4
22	3
20	3
18	3
16	2
12	2
10	1

 $n = 100$

Cuando se elaboró la prueba se hizo especial hincapié en el reactivo número 33; el criterio para evaluarla será incorrecto o correcto. Por consiguiente, se calculará el coeficiente r_{bp} mediante las puntuaciones obtenidas por los 100 estudiantes en la prueba x y la respuesta al reactivo 33, que se dicotomizó artificialmente como “incorrecto-correcto”.

Paso 1. Se concentra en un cuadro la información sobre los resultados de la prueba y sobre la respuesta de los 100 estudiantes al reactivo 33. En la columna 1 se presentan las puntuaciones; en la columna 2, la frecuencia con que ocurren éstos; en la 3, el número de estudiantes que respondieron de manera correcta el reactivo. Por último, en la columna 4, al restar las puntuaciones de la columna 3 a la 2, se obtiene la frecuencia de los estudiantes que respondieron incorrectamente al reactivo 33.

(1)	(2)	(3)	(4)
x	f	f_c	f_i
40	2	2	0
38	4	4	0
37	6	5	1
36	12	10	2
32	12	9	3
31	10	8	2
30	12	7	5
28	10	6	4
27	10	7	3
25	4	1	3
24	4	1	3
22	3	1	2
20	3	1	2
18	3	0	3
16	2	1	1
12	2	0	2
10	1	0	1

Paso 2. Se obtienen los productos de los valores de la columna (1) multiplicada por la columna (2). O sea fx , lo cual forma la columna (5). Por último, se suman estos resultados para obtener Σfx .

(5)
(fx)
80
152
222
432
384
310
360
280
270
100
96
66
60
54
32
24
10
$\Sigma fx = 2\,932$

Paso 3. Se elevan al cuadrado las puntuaciones x de la columna 1 y este resultado se multiplica por cada valor f de la columna 2; esta operación origina la columna (6). Por último, se suman todos los valores y se obtiene Σfx^2 .

Paso 4. Se obtiene la columna (7) al multiplicar las puntuaciones x de la columna 1 con los valores f_c de la columna (3); después se suman todos estos resultados y se obtiene $\Sigma f_c x$.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
x	f	fc	f_i	fx	fx^2	fcx
40	2	2	0	80	3200	80
38	4	4	0	152	5776	152
37	6	5	1	222	8214	185
36	12	10	2	432	15552	360
32	12	9	3	384	12288	288
31	10	8	2	310	9610	248
30	12	7	5	360	10800	210
28	10	6	4	280	7840	168
27	10	7	3	270	7290	189
25	4	1	3	100	2500	25
24	4	1	3	96	2304	24
22	3	1	2	66	1452	22
20	3	1	2	60	1200	20
18	3	0	3	54	972	0
16	2	1	1	32	512	16
12	2	0	2	24	288	0
10	1	0	1	10	100	0
	100	63	37	2932	89898	1987
	Σf	nc	n_i	Σfx	Σfx^2	Σfcx

Paso 5. Los resultados obtenidos en la tabla anterior se sustituyen en la definición del coeficiente de corrección biserial-puntual.

$$r_{bp} = \frac{n(\Sigma fcx) - nc(\Sigma fx)}{(nc \ n_i) \left[n \Sigma fx^2 - (\Sigma fx)^2 \right]}$$

$$r_{bp} = \frac{(100)(1987) - 63(2932)}{\sqrt{(63)(37) \left[100(89898) - (2932)^2 \right]}}$$

$$r_{bp} = \frac{13984}{\sqrt{(2331)(393176)}}$$

$$r_{bp} = \frac{13984}{30274} = 0.46$$

$r_{bp} = 0.46$

Prueba de significancia de r_{bp}

Debido a que el coeficiente de correlación biserial de punto o puntual, es un caso particular del coeficiente de correlación lineal de Pearson, su significancia y relación en la población se llevan a cabo de la siguiente manera: en primer término, se establecen las hipótesis acerca de la existencia y no existencia de correlación en la población. La notación estadística es:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

También se establece el nivel de significancia (α) con una probabilidad de error de 0.05, 0.01, 0.001, etc. Los grados de libertad se determinan de la siguiente forma: $gl = n - 2$; con éstos se compara el coeficiente r_{bp} , calculado con el coeficiente crítico que se obtiene por medio de la tabla del Anexo, con la siguiente regla de decisión:

Si $r_{bp} \geq r_{bp(crit)}$, entonces H_0 se rechaza.

Para el ejemplo desarrollado, resulta lo siguiente:

$$gl = 98, \text{ debido a que } n = 100$$

$$\alpha = 0.001$$

$r_{bp(crit)} = r_{bp(gl, \alpha)} = r_{bp(98, 0.001)} = 0.3211$. Este valor es para una prueba bilateral (de dos colas).

Como $0.46 > 0.3211$, entonces H_0 se rechaza, lo que significa que sí existe correlación biserial puntual en la población y que el valor de $r_{bp} = 0.46$ es significativo. Debido a que las correlaciones biserial puntual se usan frecuentemente como índices de asociación entre una pregunta en particular y el puntaje total de la prueba, son muy útiles para seleccionar preguntas (rubros) durante el proceso de diseño y construcción de pruebas, con el fin de incluir en las mismas ciertas características específicas que se necesita medir.

Por último, se toma en consideración que r_{bp} presenta una limitación cuando la variable continua se distribuye normalmente y la variable dicotómica tiene una probabilidad de ocurrencia de 0.5. Dicha limitación proviene de que cada vector de la variable dicotómica (falso-verdadero, correcto-incorrecto) tiene la misma posibilidad de ser escogido, por lo que el máximo valor que r_{bp} adquiriría es 0.80; pero si la variable no dicotomizada es bimodal, este valor máximo asciende a 0.90.

Finalmente, al interpretar el resultado de r_{bp} en el ejemplo de los 100 estudiantes a los que se aplica la prueba con un reactivo especial, se concluye que el coeficiente no es alto (0.46), pero existe correlación significativa (dadas las limitaciones del biserial-puntual de 0.80). Además, sobresale que la correlación es positiva; esto significa que los estudiantes con las mejores puntuaciones en la prueba tienden a responder correctamente al reactivo 33 y que, por lo contrario, quienes obtuvieron puntuaciones bajas tienden a contestarlo incorrectamente. Entonces, la pregunta 33 realiza lo que se supone constituye el propósito de la prueba: separar a los estudiantes buenos de los mediocres.

11.15

PRUEBA DE KAPPA

Este estadístico es útil para determinar por un mismo experto la confiabilidad de dos mediciones (antes-después), o mediciones realizadas de dos especialistas en forma independiente cuando la variable es categórica y lo que se mide es hasta cierto punto subjetivo (dolor, inflamación, efectividad de un tratamiento, etc.) por lo que es necesario establecer una escala; por ejemplo: 0 = no hay, 1 = leve, 2 = regular y 3 = severo.

EJEMPLO

11.16

La inflamación de 90 encías en pacientes con problemas periodontales, fueron evaluadas por dos odontólogos certificados (A y B), considerando la siguiente escala:

- 0 = sin inflamación
- 1 = inflamación leve
- 2 = inflamación regular
- 3 = inflamación severa

Para conocer la confiabilidad de las mediciones de los odontólogos, sus evaluaciones se ubican en una tabla de doble entrada, llamada también de contingencia, de 2×2 (2 renglones por 2 columnas).

		Odontólogo B				
		0	1	2	3	Total
Odontólogo A	0	25	2	1	0	28
	1	4	17	3	2	26
	2	1	3	15	1	20
	3	0	1	2	13	16
	Total	30	23	21	16	90

Las proporciones obtenidas por los dos odontólogos son independientes.

El estadístico Kappa (k) se define como:

$$k = \frac{P_o - P_e}{1 - P_e}$$

donde:

- P_o = proporción observada de respuestas similares para ambos observadores A y B
- P_e = proporción esperada de respuestas similares para ambos observadores A y B

$$P_e = \sum_{i=1}^c Q_i b_i$$

donde:

- Q_i = proporción de respuestas en la categoría i , para el observador A
- b_i = proporción de respuestas en la categoría i , para el observador B
- c = número de categorías

La prueba Kappa varía entre 0 y 1, donde una perfecta confiabilidad se denota por 1, mientras que 0 indica que no existe confiabilidad de las dos mediciones, o sea $P_o = P_e$; si Kappa > 0.75 se considera excelente. Entre 0.40 y 0.75 buena y < 0.4 pobre. Para el ejemplo anterior se tiene:

$$P_e = \frac{25 + 17 + 15 + 13}{90} = 0.78$$

$$Q_1 = \frac{28}{90} = 0.31 \quad b_1 = \frac{30}{90} = 0.33$$

$$Q_2 = \frac{26}{90} = 0.29 \quad b_2 = \frac{23}{90} = 0.26$$

$$Q_3 = \frac{20}{90} = 0.22 \quad b_3 = \frac{21}{90} = 0.23$$

$$Q_4 = \frac{16}{90} = 0.18 \quad b_4 = \frac{16}{90} = 0.18$$

$$P_e = (0.31)(0.33) + (0.29)(0.26) + (0.22)(0.23) + (0.18)(0.18)$$

$$P_e = 0.10 + 0.08 + 0.05 + 0.03 = 0.26$$

$$\text{Kappa} = \frac{0.78 - 0.26}{1 - 0.26} = \frac{0.52}{0.74} = 0.70$$

$$k = 0.70$$

Esto indica una buena confiabilidad del sistema de evaluación.

11.15

EJERCICIOS

TÉCNICAS BÁSICAS

11.42 Dé la región de rechazo en una prueba para detectar correlación de rango positiva si el número de pares de rangos es 16 y se tienen estos valores α :

- a. $\alpha = .05$ b. $\alpha = .01$

11.43 Dé la región de rechazo en una prueba para detectar correlación de rango negativa si el número de pares de rangos es 12 y se tienen estos valores α :

- a. $\alpha = .05$ b. $\alpha = .01$

11.44 Dé la región de rechazo en una prueba para detectar correlación de rango si el número de pares de rangos es 25 y se tienen estos valores α :

- a. $\alpha = .05$ b. $\alpha = .01$

11.45 Las siguientes observaciones en pares se obtuvieron en dos variables x y y :

x	1.2	.8	2.1	3.5	2.7	1.5
y	1.0	1.3	.1	−.8	−.2	.6

- a. Calcule el coeficiente de correlación de rango de Spearman r_s .
- b. ¿Los datos presentan suficiente evidencia para indicar una correlación entre x y y ? Pruebe usando $\alpha = .05$.

APLICACIONES



EX1146

11.46 Calificación de candidatos políticos Un

estudiante de Ciencias Políticas desea examinar la relación entre la imagen que tiene un elector, respecto de un candidato político y la distancia (en millas) entre las residencias del elector y el candidato. Cada uno de 12 electores calificó al candidato en una escala de 1 a 20.

Elector	Calificación	Distancia
1	12	75
2	7	165
3	5	300
4	19	15
5	17	180
6	12	240
7	9	120
8	18	60
9	3	230
10	8	200
11	15	130
12	4	130

- a. Calcule el coeficiente de correlación de rango de Spearman r_s .
- b. ¿Estos datos dan suficiente evidencia para indicar una correlación negativa entre calificación y distancia?

DATOS
EX1147

11.47 Carreras de competencia ¿El número de años de experiencia en carreras de competencia está relacionado con el rendimiento en carreras de distancia de un corredor? Los datos de nueve corredores, obtenidos de un estudio hecho por Scott Powers y colegas, se muestran en la tabla siguiente:⁵

Corredor	Años de correr en competencias	Tiempo de llegada en 10 kilómetros (min)
1	9	33.15
2	13	33.33
3	5	33.50
4	7	33.55
5	12	33.73
6	6	33.86
7	4	33.90
8	5	34.15
9	3	34.90

- Calcule el coeficiente de correlación de rango entre años de carreras en competencias x y el tiempo y de llegada a la meta en una carrera de 10 kilómetros.
- ¿Los datos dan suficiente evidencia para indicar una correlación de rango significativa entre y y x ? Pruebe usando $\alpha = .05$.

DATOS
EX1148

11.48 Raquetas de tenis Los datos mostrados en la tabla siguiente dan medidas de rigidez al doblamiento y de rigidez al torcimiento, determinadas por pruebas de ingeniería en 12 raquetas de tenis.

Raqueta	Rigidez al doblamiento	Rigidez al torcimiento
1	419	227
2	407	231
3	363	200
4	360	211
5	257	182
6	622	304
7	424	384
8	359	194
9	346	158
10	556	225
11	474	305
12	441	235

- Calcule el coeficiente de correlación de rango r_s entre rigidez al doblamiento y rigidez al torcimiento.
- Si una raqueta tiene rigidez al doblamiento, ¿también es probable que tenga rigidez al torcimiento? Use el coeficiente de correlación de rango para determinar si hay una relación positiva significativa entre rigidez al doblamiento y rigidez al torcimiento. Use $\alpha = .05$.

11.49 Clasificaciones de estudiante El director de una escuela sospechaba que la actitud de un profesor hacia un alumno de primer año dependía de su juicio original sobre la capacidad del niño. El director también sospechaba que

gran parte de ese juicio estaba basado en la puntuación del coeficiente de inteligencia (CI) del alumno, que por lo general era conocida por el profesor. Después de tres semanas de enseñanza, se le pidió a un profesor que clasificara a los nueve niños de su grupo de 1 (la más alta) a 9 (la más baja) de acuerdo con su opinión de la capacidad de los niños. Calcule r_s para estos rangos de CI-profesor:

Rango	1	2	3	4	5	6	7	8	9
CI	3	1	2	4	5	7	9	6	8

11.50 Clasificaciones de estudiante,

continúa Consulte el ejercicio 11.49. ¿Los datos dan suficiente evidencia para indicar una correlación positiva entre las clasificaciones del profesor y los rangos de los CI? Use $\alpha = .05$.

DATOS
EX1151

11.51 Críticos de arte Dos críticos de arte clasificaron 10 pinturas de artistas contemporáneos (pero anónimos), de acuerdo con su atractivo para los críticos respectivos. Las clasificaciones se ilustran en la tabla siguiente. ¿Los críticos parecen estar de acuerdo en sus clasificaciones de arte contemporáneo? Es decir, ¿los datos dan suficiente evidencia para indicar una correlación positiva entre los críticos A y B? Pruebe usando un valor α cercano a .05.

Pintura	Crítico A	Crítico B
1	6	5
2	4	6
3	9	10
4	1	2
5	2	3
6	7	8
7	3	1
8	8	7
9	5	4
10	10	9

11.52 Calificación de hojas de tabaco Se realizó un experimento para estudiar la relación entre las calificaciones de un experto en clasificar hojas de tabaco y el contenido de humedad de las hojas. Doce de éstas fueron calificadas por el experto en una escala de 1 a 10 y se tomaron las lecturas correspondientes del contenido de humedad.

Hoja	Calificación del experto	Contenido de humedad
1	9	.22
2	6	.16
3	7	.17
4	7	.14
5	5	.12
6	8	.19
7	2	.10
8	6	.12
9	1	.05
10	10	.20
11	9	.16
12	3	.09

Calcule r_s . ¿Los datos dan suficiente evidencia para indicar una asociación entre las calificaciones del experto y el contenido de humedad de las hojas?

DATOS
EX1153

11.53 Educación en habilidades sociales Se implementó un programa de educación en habilidades sociales para siete estudiantes con discapacidad leve, en un estudio para determinar si el programa causaba mejoras en medidas de antes y después, así como en calificaciones de conducta. Las calificaciones de antes y después de una de las pruebas de los siete estudiantes se muestran en la tabla siguiente:

Estudiante	Antes de la prueba	Después de la prueba
Earl	101	113
Ned	89	89
Jasper	112	121
Charlie	105	99
Tom	90	104
Susie	91	94
Lori	89	99

- a. Use una prueba no paramétrica para determinar si hay una relación positiva significativa entre las calificaciones de exámenes anteriores y posteriores.
- b. ¿Estos resultados concuerdan con los resultados de la prueba paramétrica del ejercicio 8.54?

11.54 Se aplican a 12 personas tres tratamientos nuevos, que se comparan con el tratamiento tradicional, obteniendo los siguientes datos:

Sujetos	Tratamiento			
	(Control) A	B	C	D
1	8	21	23	15
2	10	29	30	21
3	20	16	19	18
4	15	20	19	18
5	12	13	10	14
6	10	15	12	17
7	9	5	12	6
8	15	18	18	12
9	17	26	32	21
10	12	17	20	9
11	10	4	10	2
12	6	12	15	4

Calcule la prueba de Friedman.

11.55 Aplique la prueba Kappa. Xerostomía en pacientes sometidos a radioterapia en cabeza y cuello.

x = xerostomía

0 = flujo salival normal

1 = flujo salival moderado

2 = flujo salival medio

3 = flujo salival escaso

OBSERVACIÓN B

OBSERVACIÓN A	OBSERVACIÓN B				
	0	1	2	3	Total
0	5	0	3	1	9
1	1	4	4	2	11
2	1	4	5	4	14
3	0	2	4	10	16
Total	7	13	13	17	50

11.56 Calidad ósea en reborde alveolar anterior mandibular en pacientes desdentados.

x = hueso alveolar

0 = normal

1 = buena

2 = regular

3 = pobre

A	B				
	0	1	2	3	Total
0	13	3	2	1	19
1	2	5	4	2	13
2	6	5	4	1	16
3	1	1	0	0	2
Total	22	14	10	4	50

11.57 Calcular la prueba (k) Kappa

x = problema inflamación de papila gustativa

0 = normal

1 = leve

2 = moderada

3 = severa

$n = 74$

Inflamación (Observación "A")	Inflamación (Observación "B")				
	0	1	2	3	Total
0	0	4	8	3	15
1	2	18	0	1	21
2	7	5	9	3	24
3	7	1	0	5	14
Total	16	28	17	12	74

11.58 Calcular la prueba Kappa

x = problema periodontal

0 = normal

1 = leve

2 = moderado

3 = severo

$n = 90$

		Problema periodontal Observación "B"				
		0	1	2	3	Total
Problema periodontal "A"	0	16	8	4	2	30
	1	7	9	1	8	25
	2	4	0	5	7	16
	3	0	9	4	6	19
	Total	27	26	14	23	90

11.16

RESUMEN

Las pruebas no paramétricas expuestas en este capítulo son sólo algunas de las numerosas pruebas de este tipo disponibles para los experimentadores. Para las pruebas que se presentan aquí hay tablas de valores críticos fácilmente disponibles.

Los métodos estadísticos no paramétricos son especialmente útiles cuando las observaciones se pueden ordenar pero no colocar exactamente en una escala de medición. Además, los métodos no paramétricos son los únicos que son usados cuando se hayan apegado correctamente a los diseños de muestreo, pero los datos no se supone que sigan o no es posible asumir que sigan una o más suposiciones de distribución prescritas.

Hemos presentado una amplia variedad de técnicas no paramétricas, que se pueden usar cuando los datos no están normalmente distribuidos, o las otras suposiciones requeridas no se satisfacen. En la literatura existen procedimientos de una muestra; sin embargo, nos hemos concentrado en analizar dos o más muestras que han sido correctamente seleccionadas usando muestreo aleatorio e independiente, como lo requiere el diseño de que se trate. Los análogos no paramétricos de los procedimientos paramétricos presentados en los capítulos 6 a 10 son sencillos y muy fáciles de poner en práctica:

- La prueba de suma de rango de Wilcoxon es el análogo no paramétrico de la prueba t de dos muestras.
- La prueba del signo y las pruebas de rango con signo de Wilcoxon son los análogos no paramétricos de la prueba t de muestra en pares.
- La prueba H de Kruskal-Wallis es el equivalente de rango del análisis de varianza de una vía de la prueba F .
- La prueba F_r de Friedman es el equivalente de rango del análisis de varianza de dos vías del diseño aleatorizado de bloques de la prueba F .
- La correlación de rango r_s de Spearman es el equivalente de rango del coeficiente de correlación de Pearson.

Estos procedimientos no paramétricos y muchos más están disponibles como alternativas a las pruebas paramétricas presentadas anteriormente. Es importante tener en cuenta que cuando los supuestos requeridos de las poblaciones muestreadas se relajan, nuestra capacidad para detectar diferencias significativas en una o más características de la población disminuye.

REPASO DEL CAPÍTULO

Conceptos y fórmulas clave

I. Métodos no paramétricos

- Estos métodos se pueden utilizar cuando no es posible medir los datos en una escala cuantitativa, o cuando
- La escala numérica de medición sea fijada arbitrariamente por el investigador, o bien, cuando
- Las suposiciones paramétricas tales como normalidad o varianza constante sean violadas gravemente.

II. Prueba de la suma de rango de Wilcoxon: muestras aleatorias independientes

- Conjuntamente ordene las dos muestras. Designe la muestra más pequeña como muestra 1. Entonces

$$T_1 = \text{Suma de rango de la muestra 1}$$

$$T_1^* = n_1(n_1 + n_2 + 1) - T_1$$
- Use T_1 para probar si la población 1 está a la izquierda de la población 2. Use T_1^* para probar si la población 1 está a la derecha de la población 2. Use la menor de T_1 y T_1^* para probar si hay diferencia en las ubicaciones de las dos poblaciones.
- La tabla 7 del Anexo tiene valores críticos para el rechazo de H_0 .
- Cuando los tamaños muestrales sean grandes, use la aproximación normal:

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\sigma_T^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

$$z = \frac{T - \mu_T}{\sigma_T}$$

III. Prueba del signo para un experimento en pares

- Encuentre x , el número de veces que la observación A exceda de la observación B para un par determinado.
- Para probar si hay diferencia en dos poblaciones, pruebe $H_0: p = .5$ contra una alternativa de una o de dos colas.
- Use la tabla 1 del Anexo para calcular el valor p para la prueba.

- Cuando los tamaños muestrales sean grandes, use la aproximación normal:

$$z = \frac{x - .5n}{.5\sqrt{n}}$$

IV. Prueba de rango con signo de Wilcoxon: experimento en pares

- Calcule las diferencias en las observaciones en pares. Ordene los valores absolutos de las diferencias. Calcule las sumas de rango T^+ y T^- para las diferencias positivas y negativas, respectivamente. El estadístico de prueba T es la menor de las dos sumas de rango.
- La tabla 8 del Anexo tiene valores críticos para el rechazo de H_0 para pruebas de una y de dos colas.
- Cuando los tamaños muestrales sean grandes, use la aproximación normal:

$$z = \frac{T^+ - [n(n+1)/4]}{\sqrt{[n(n+1)(2n+1)]/24}}$$

V. Prueba H de Kruskal-Wallis: diseño completamente aleatorizado

- Conjuntamente ordene las n observaciones de las k muestras. Calcule las sumas de rango, T_i = suma de rango de la muestra i y el estadístico de prueba

$$H = \frac{12}{n(n+1)} \sum \frac{T_i^2}{n_i} - 3(n+1)$$

- Si la hipótesis nula de igualdad de distribuciones es falsa, H será inusualmente grande, resultando en una prueba de una cola.
- Para tamaños muestrales de cinco o mayores, la región de rechazo para H está basada en la distribución ji cuadrada con $(k-1)$ grados de libertad.

VI. Prueba F_r de Friedman: diseño aleatorizado de bloques

- Ordene las respuestas dentro de cada bloque de 1 a k . Calcule las sumas de rango, $T_1, T_2, \dots, T_k$, y el estadístico de prueba

$$F_r = \frac{12}{bk(k+1)} \sum T_i^2 - 3b(k+1)$$

2. Si la hipótesis nula de igualdad de distribuciones de tratamiento es falsa, F_r será inusualmente grande, resultando en una prueba de una cola.
3. Para tamaños de bloques de cinco o mayores, la región de rechazo para F_r está basada en la distribución ji cuadrada con $(k - 1)$ grados de libertad.

VII. Coeficiente de correlación de rango de Spearman

1. Ordene las respuestas para las dos variables de menor a mayor.

2. Calcule el coeficiente de correlación para las observaciones ordenadas:

$$r_s = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \quad \text{o} \quad r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

3. La tabla 9 del Anexo da valores críticos para correlaciones de rango significativamente diferentes de 0.
4. El coeficiente de correlación de rango detecta no sólo correlación lineal significativa sino también cualquier otra relación monótonica entre las dos variables.



TECNOLOGÍA ACTUAL

Procedimientos no paramétricos: MINITAB

Aunque no hay opciones para procedimientos no paramétricos en *MS Excel*, muchos de estos procedimientos se encuentran disponibles en el paquete *MINITAB*, incluyendo la mayor parte de las pruebas estudiadas en este capítulo. Los cuadros de diálogo son conocidos para el usuario por ahora y veremos las pruebas en el orden presentado en el capítulo.

Para poner en práctica la prueba de la suma de rango de Wilcoxon para dos muestras aleatorias independientes introduzca los dos conjuntos de datos muestrales en dos columnas (por ejemplo, C1 y C2) de la hoja de trabajo *MINITAB*. El cuadro de diálogo de la figura 11.13 se genera usando **Stat ► Nonparametrics ► Mann-Whitney**. Seleccione C1 y C2 para la **First** y **Second Samples** e indique el coeficiente de confianza apropiado (para un intervalo de confianza) e hipótesis alternativa. Al hacer clic en **OK** se genera la salida de la figura 11.1.

La prueba del signo y la prueba de rango con signo de Wilcoxon para muestras en pares se efectúan exactamente en la misma forma, con un cambio sólo en el último comando de la secuencia. Incluso los cuadros de diálogo son idénticos. Introduzca los datos en dos columnas de la hoja de trabajo *MINITAB* (usamos los datos de harina para pastel en la sección 11.5). Antes de que se pueda implementar cada prueba es preciso generar una columna de diferencias usando **Calc ► Calculator**, como se ve en la figura 11.14. Use **Stat ► Nonparametrics ► 1-Sample Sign** o **Stat ► Nonparametrics ► 1-Sample Wilcoxon** para generar el cuadro de diálogo apropiado que se muestra en la figura 11.15.

FIGURA 11.13

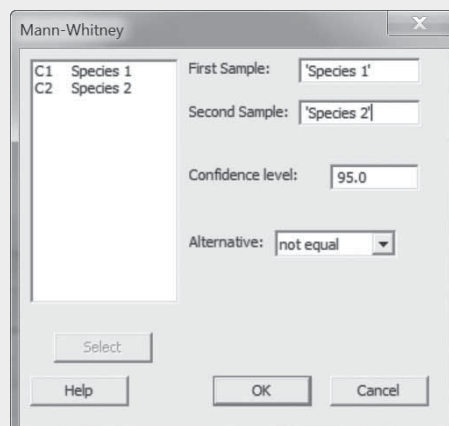
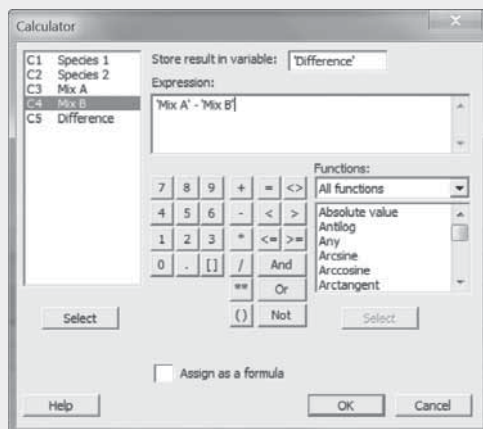


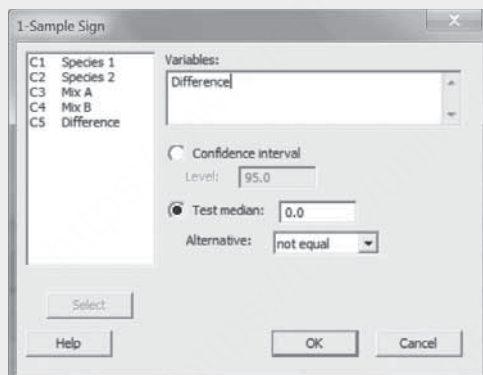
FIGURA 11.14



Recuerde que la mediana es el valor de una variable tal que 50% de los valores son más pequeños y 50% son más grandes. En consecuencia, si las dos distribuciones poblacionales son iguales, la mediana de las diferencias será 0. Esto es equivalente a la hipótesis nula

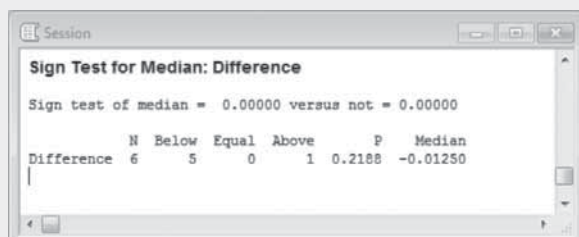
$$H_0 : P(\text{diferencia positiva}) = P(\text{diferencia negativa}) = .5$$

FIGURA 11.15



empleada para la prueba del signo. Seleccione la columna de diferencias para el cuadro Variables y seleccione la prueba de la mediana igual a 0 con la alternativa apropiada. Haga clic en **OK** para obtener la salida impresa de cualquiera de las dos pruebas. La salida impresa de la ventana Session para la prueba del signo, ilustrada en la figura 11.16, indica una diferencia no significativa en las distribuciones de densidades para las dos harinas para pastel. Observe que el valor p (.2188) no es igual que el valor p para la prueba de rango con signo de Wilcoxon (.093 de la figura 11.4). No obstante, si se prueba al nivel de 5%, ambas pruebas producen diferencias no significativas.

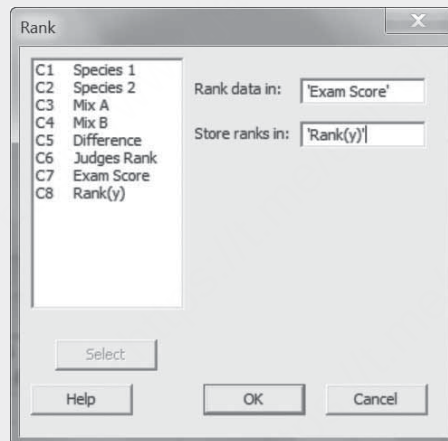
FIGURA 11.16



Los procedimientos para implementar la prueba H de Kruskal-Wallis para k muestras independientes y la prueba F_r de Friedman para un diseño aleatorizado de bloques, son idénticos a los procedimientos empleados para sus equivalentes paramétricos. Repase los métodos descritos en la sección “Tecnología actual” del capítulo 7. Una vez que haya introducido los datos como se explica en esa sección, los comandos **Stat ► Nonparametrics ► Kruskal-Wallis** o **Stat ► Nonparametrics ► Friedman** van a generar un cuadro de diálogo en el que el usuario especifica la columna Response y la columna Factor o la columna de respuesta, la columna del tratamiento y la columna de bloque, respectivamente. Haga clic en **OK** para obtener las salidas impresas para estas pruebas no paramétricas.

Finalmente, se puede generar el coeficiente no paramétrico de correlación de rango r_s si se introducen los datos en dos columnas y se ordenan los datos usando **Data ► Rank**. Por ejemplo, los datos sobre la clasificación del juez y las calificaciones del examen se introdujeron en las columnas C6 y C7 de nuestra hoja de cálculo *MINITAB*. Como las clasificaciones del juez ya están ordenadas, sólo necesitamos ordenar C7 al seleccionar “Exam Score” y guardar los rangos en C8 [llamada “Rank (y)” en la figura 11.17]. Los comandos **Stat ► Basic Statistics ► Correlation** producirán ahora el coeficiente de correlación de rango cuando C6 y C8 se seleccionen. No obstante, el valor p que se ve en la salida *no produce* exactamente la misma prueba que los valores críticos de la tabla 11.14. Debe compararse el valor de r_s obtenido por el usuario con el valor tabulado para comprobar si hay una asociación significativa entre las dos variables.

FIGURA 11.17



Ejercicios suplementarios

DATOS

EX1159

11.59 Tiempos de respuesta Se realizó un experimento para comparar los tiempos de respuesta para dos estímulos diferentes. Para eliminar la variabilidad natural de una persona a otra en las respuestas, ambos estímulos se presentaron a cada una de nueve personas, permitiendo así un análisis de las diferencias entre estímulos *dentro* de cada persona. La tabla es una lista de tiempos de respuesta (en segundos).

Persona	Estímulo 1	Estímulo 2
1	9.4	10.3
2	7.8	8.9
3	5.6	4.1
4	12.1	14.7
5	6.9	8.7
6	4.2	7.1
7	8.8	11.3
8	7.7	5.2
9	6.4	7.8

- Use la prueba del signo para determinar si existe suficiente evidencia para indicar una diferencia en los tiempos medios de respuesta para los dos estímulos. Use una región de rechazo para la cual $\alpha \leq .05$.
- Pruebe la hipótesis de que no hay diferencia en tiempos medios de respuesta usando la prueba t de Student.

11.60 Tiempos de respuesta, continúa Consulte el ejercicio 11.59. Pruebe la hipótesis de que no hay diferencia en las distribuciones de tiempos de respuesta para los dos estímulos, usando la prueba del rango con signo de Wilcoxon. Use una región de rechazo para la cual α sea tan cercana como sea posible a la α obtenida en la parte **a** del ejercicio 11.59.

DATOS

EX1161

11.61 Gemelos idénticos Para comparar dos escuelas secundarias, A y B, en eficiencia académica, se diseñó un experimento que requería el uso de 10 pares de gemelos idénticos, y en el que cada gemelo acabara de terminar el sexto grado. En cada caso, los gemelos del mismo par obtuvieron su enseñanza en los mismos salones de clase en cada nivel de grado. Un niño fue seleccionado al azar de cada par de gemelos y asignado a la escuela A. Los demás niños fueron enviados a la escuela B. Cerca del final del noveno grado, se aplicó cierto examen de aprovechamiento a cada niño del experimento. Las calificaciones del examen se muestran en la tabla siguiente.

Par de gemelos	Escuela A	Escuela B
1	67	39
2	80	75
3	65	69
4	70	55
5	86	74
6	50	52
7	63	56
8	81	72
9	86	89
10	60	47

- Pruebe (usando la prueba del signo) la hipótesis de que las dos escuelas son iguales en efectividad académica, medida por calificaciones en el examen de aprovechamiento, contra la alternativa de que las escuelas no son igualmente eficaces.
- Suponga que se sabe que la escuela secundaria A tenía un mejor profesorado y mejores instalaciones de enseñanza. Pruebe la hipótesis de igual efectividad académica contra la alternativa de que la escuela A es superior.

11.62 Gemelos idénticos II Consulte el ejercicio 11.61. ¿Qué respuestas se obtienen si se usa la prueba de rango con signo de Wilcoxon para analizar los datos? Compare con sus respuestas anteriores.

DATOS

EX1163

11.63 Brilantez de papel Los valores codificados para una medida de la brillantez del papel (reflectividad de la luz), preparados por dos procesos diferentes, aparecen en la tabla para muestras de nueve observaciones tomadas al azar de cada uno de los dos procesos. ¿Los datos presentan suficiente evidencia para indicar una diferencia en las mediciones de la brillantez para los dos procesos? Use una prueba paramétrica y una no paramétrica y compare sus resultados.

Proceso	Brillantez								
A	6.1	9.2	8.7	8.9	7.6	7.1	9.5	8.3	9.0
B	9.1	8.2	8.6	6.9	7.5	7.9	8.3	7.8	8.9

11.64 Instrumentos de precisión Suponga (como en el caso de mediciones producidas por dos instrumentos de medición bien calibrados) que las medias de dos poblaciones son iguales. Use el estadístico de la suma de rango de Wilcoxon para probar hipótesis respecto a las varianzas poblacionales como sigue:

- Ordene la muestra combinada.
- Numere las observaciones ordenadas “de afuera hacia adentro”; esto es, numere 1 la observación más

pequeña, la mayor 2, la siguiente a la más pequeña 3, la siguiente a la más grande 4, y así sucesivamente. Esta secuencia de números induce un ordenamiento en los símbolos A (objetos de la población A) y B (objetos de la población B). Si $\sigma_A^2 > \sigma_B^2$, uno esperaría hallar una preponderancia de las A cercana al comienzo de la secuencia y así una “suma de rangos” relativamente pequeña para las observaciones A.

- c. Dadas las mediciones de la tabla producidas por los instrumentos de precisión bien calibrados A y B, pruebe cerca del nivel $\alpha = .05$ para determinar si el instrumento B más costoso es más preciso que el A. (Observe que esto implica una prueba de una cola.) Use el estadístico de prueba de la suma de rango de Wilcoxon.

Instrumento A	Instrumento B
1060.21	1060.24
1060.34	1060.28
1060.27	1060.32
1060.36	1060.30
1060.40	

- d. Pruebe usando la igualdad de varianzas de la prueba F .

DATOS
EX1165

11.65 Ablandadores de carne Se realizó un experimento para comparar la suavidad de cortes de carne tratados con dos ablandadores de carne diferentes, A y B. Para reducir el efecto de las variables extrañas, los datos fueron pareados por el corte de carne específico al aplicar los ablandadores a dos cortes tomados de la misma res, al cocinar cortes en pares juntos y usar un solo juez para cada par. Después de la cocción, cada corte fue calificado por un juez en una escala de 1 a 10, con un 10 correspondiente a la carne más suave. Los datos mostrados en la tabla son de un solo juez. ¿Los datos dan suficiente evidencia para indicar que uno de los suavizadores tiende a recibir calificaciones más altas que el otro? ¿La prueba t de Student sería más apropiada para analizar estos datos? Explique.

Corte	Suavizador	
	A	B
Asado de paletilla	5	7
Asado de lomo	6	5
Filete de costilla	8	9
Pecho	4	5
Filete	9	9
Filete bola	3	5
Asado de pierna	7	6
Solomillo	8	8
Puntas de solomillo	8	9
Chuleta	9	10

DATOS
EX1166

11.66 Entrevista a prospectos de trabajo Una gran empresa selecciona graduados universitarios

para un empleo usando tanto entrevistas como un examen psicológico de aprovechamiento. Las entrevistas efectuadas en la casa matriz de la compañía son mucho más costosas que las efectuadas en el plantel. En consecuencia, la oficina de personal estaba interesada en determinar si las calificaciones del examen estaban correlacionadas con las calificaciones de entrevistas y si los exámenes podrían ser sustituidos por entrevistas. La idea no era eliminar las entrevistas sino reducir su número. Para determinar si las medidas estaban correlacionadas, se calificaron 10 prospectos durante las entrevistas y se examinaron. Las calificaciones en pares se dan a continuación:

Persona	Calificación de entrevista	Calificación de examen
1	8	74
2	5	81
3	10	66
4	3	83
5	6	66
6	1	94
7	4	96
8	7	70
9	9	61
10	2	86

Calcule el coeficiente de correlación de rango de Spearman r_s . La calificación 1 se asigna al candidato juzgado como el mejor.

11.67 Entrevistas, continúa Consulte el ejercicio 11.66. ¿Los datos presentan suficiente evidencia para indicar que la correlación entre las calificaciones de entrevista y las calificaciones de examen es menor a cero? Si esta evidencia existe, ¿se puede decir que los exámenes se pueden usar para reducir el número de entrevistas?

11.68 Experimentos de asociación de palabras Una comparación de tiempos de reacción, para dos estímulos diferentes en un experimento psicológico de asociación de palabras, produjo los resultados siguientes cuando se aplicó a una muestra aleatoria de 16 personas:

Estímulo	Tiempo de reacción (segundos)							
1	1	3	2	1	2	1	3	2
2	4	2	3	3	1	2	3	3

¿Los datos presentan suficiente evidencia para indicar una diferencia en la media de los tiempos de reacción para los dos estímulos? Use una prueba apropiada, no paramétrica, y explique sus conclusiones.

DATOS
EX1169

11.69 Matemáticas y arte La tabla muestra las calificaciones de un grupo de 15 estudiantes en matemáticas y arte. Use la prueba de rango con signo de Wilcoxon para determinar si las calificaciones medianas para estos estudiantes difieren significativamente para las dos materias.

Estudiante	Matemáticas	Arte	Estudiante	Matemáticas	Arte
1	22	53	9	62	55
2	37	68	10	65	74
3	36	42	11	66	68
4	38	49	12	56	64
5	42	51	13	66	67
6	58	65	14	67	73
7	58	51	15	62	65
8	60	71			

11.70 Matemáticas y arte, continúa Consulte el ejercicio 11.69. Calcule el coeficiente de correlación de rango de Spearman para estos datos y pruebe H_0 : no hay asociación entre los pares ordenados al nivel de significancia de 10%.

11.71 Producción de trigo El ejercicio 7.67 presentó un análisis de varianza de las producciones de cinco variedades diferentes de trigo, observadas cada una en un terreno, en cada uno de seis lugares diferentes (véase el conjunto de datos EX0767). Los datos de este diseño aleatorizado de bloques se dan a continuación:

Variedades	Lugar					
	1	2	3	4	5	6
A	35.3	31.0	32.7	36.8	37.2	33.1
B	30.7	32.2	31.4	31.7	35.0	32.7
C	38.2	33.4	33.6	37.1	37.3	38.2
D	34.9	36.1	35.2	38.3	40.2	36.0
E	32.4	28.9	29.2	30.7	33.9	32.1

- Use la prueba no paramétrica apropiada con el fin de determinar si los datos aportan suficiente evidencia para indicar una diferencia en las producciones, para las cinco diferentes variedades de trigo. Pruebe usando $\alpha = .05$.
- El ejercicio 7.68 presentó una salida impresa de computadora del análisis de varianza para comparar las producciones medias para las cinco variedades de trigo. ¿Cómo se comparan los resultados del análisis de varianza de la prueba F con la prueba de la parte **a**? Explique.

11.72 Aprendiendo a vender En el ejercicio 7.60 se compararon los números de ventas por aprendiz, después de terminar uno de cuatro programas diferentes de capacitación en ventas (véase el conjunto de datos EX0760). Seis aprendices completaron el programa de capacitación 1, ocho completaron el 2, y así sucesivamente. Los números de ventas por aprendiz se muestran en la tabla siguiente.

	Programa de capacitación			
	1	2	3	4
	78	99	74	81
	84	86	87	63
	86	90	80	71
	92	93	83	65
	69	94	78	86
	73	85		79
		97		73
		91		70
Total	482	735	402	588

- ¿Los datos presentan suficiente evidencia para indicar que la distribución del número de ventas por aprendiz difiere de un programa de capacitación a otro? Determine usando la prueba no paramétrica apropiada.
- ¿Cómo se comparan los resultados de la prueba de la parte **a** con los resultados del análisis de varianza de la prueba F del ejercicio 7.60?

11.73 Contaminación de plantas de productos químicos

En el ejercicio 7.65 efectuamos un análisis de varianza para comparar los niveles medios de descargas residuales en agua, en cuatro plantas industriales diferentes (véase el conjunto de datos EX0765). Se tomaron cinco muestras del desecho líquido a la salida de cada una de las cuatro plantas industriales; los datos se muestran en la tabla siguiente.

Planta	Descargas contaminantes (lb/gal de desechos)				
A	1.65	1.72	1.50	1.37	1.60
B	1.70	1.85	1.46	2.05	1.80
C	1.40	1.75	1.38	1.65	1.55
D	2.10	1.95	1.65	1.88	2.00

- ¿Los datos presentan suficiente evidencia para indicar una diferencia en los niveles de contaminantes para las cuatro plantas industriales diferentes? Pruebe usando la prueba no paramétrica apropiada.
- Encuentre el valor p aproximado para la prueba e interprete su valor.
- Compare los resultados de la prueba de la parte **a** con el análisis de varianza del ejercicio 7.65. ¿Los resultados concuerdan? Explique.

11.74 Investigación del SIDA Unos científicos han demostrado que una vacuna recién inventada puede proteger a monos Rhesus contra infecciones causadas por un virus estrechamente relacionado con el virus de inmunodeficiencia humana (VIH) causante del SIDA. En su trabajo, Ronald C. Resrosiers y sus colegas del Centro Regional de Investigación de Primates de Nueva Inglaterra, aplicaron a cada uno de $n = 6$ monos Rhesus cinco inoculaciones con la vacuna del virus de inmunodeficiencia de simios (VIS). Una semana después de la última vacuna, cada mono recibió una inyección de VIS vivo. Dos de los seis monos vacunados no mostraron evidencia de infección por el VIS hasta por año y medio después de la inyección con el VIS. Los científicos pudieron aislar el virus del VIS de los otros cuatro monos vacunados, aun cuando estos animales no mostraron signos de la enfermedad. ¿Esta información contiene suficiente evidencia para indicar que la vacuna es eficaz para proteger al menos contra el VIS? Use $\alpha = .10$.

DATOS

EX1175

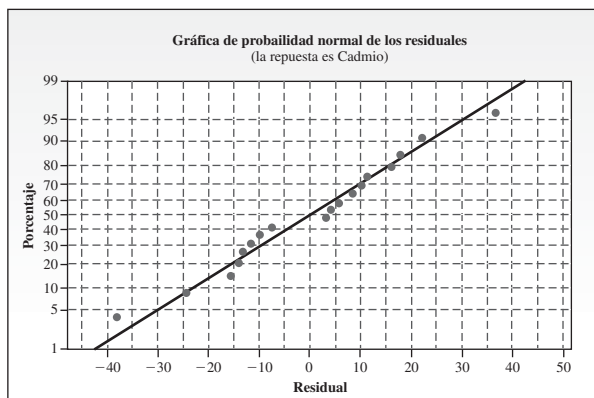
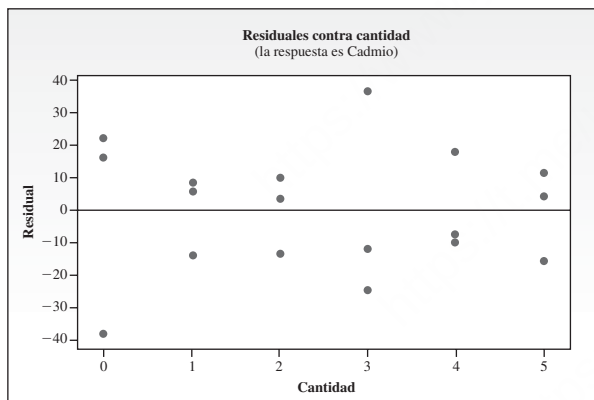
11.75 Metal pesado Se efectuó un experimento para determinar si hay acumulación de metales pesados en plantas crecidas en suelos mejorados

con lodo y en insectos que se alimentan de esas plantas.⁶ Los datos de la tabla siguiente son concentraciones de cadmio (en $\mu\text{g/kg}$), en plantas crecidas bajo seis cantidades diferentes de aplicación de lodo para tres cosechas distintas. Dichas cantidades son los tratamientos. Las tres cosechas representan bloques de tiempo en el diseño de dos vías.

Cantidad	Cosecha		
	1	2	3
Control	162.1	153.7	200.4
1	199.8	199.6	278.2
2	220.0	210.7	294.8
3	194.4	179.0	341.1
4	204.3	203.7	330.2
5	218.9	236.1	344.2

- a. Con base en las gráficas de diagnóstico, ¿está usted dispuesto a aceptar que las suposiciones de normalidad y varianza constante se satisfacen?

Gráficas de diagnóstico para el ejercicio 11.75



- b. Usando un método apropiado de análisis, estudie los datos para determinar si hay diferencias significativas entre las respuestas debido a las cantidades de aplicación.

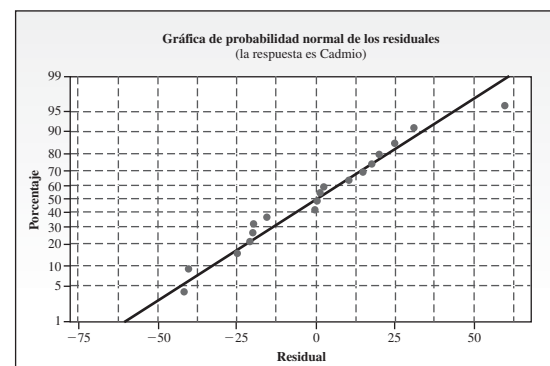
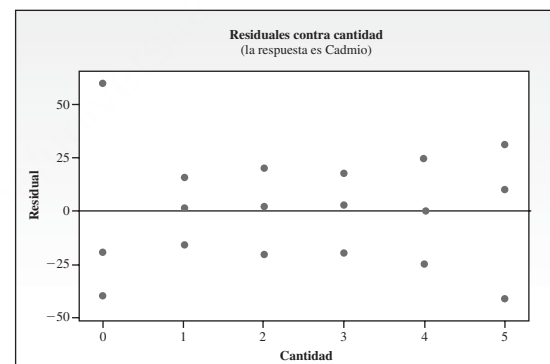
DATOS
EX1176

11.76 Consulte el ejercicio 11.75. Los datos de la tabla siguiente son concentraciones de cadmio encontradas en pulgones alimentados de las plantas crecidas en suelo mejorado con lodo.

Cantidad	Cosecha		
	1	2	3
Control	16.2	55.8	65.8
1	16.9	119.4	181.1
2	12.7	171.9	184.6
3	31.3	128.4	196.4
4	38.5	182.0	163.7
5	20.6	191.3	242.8

- a. Use las gráficas de diagnóstico para evaluar si las suposiciones de normalidad y varianza constante son razonables en este caso.
- b. Con base en las conclusiones de la parte a, use un método estadístico apropiado para probar si hay diferencias significativas en las concentraciones de cadmio para las seis cantidades de aplicación.

Gráficas de diagnóstico para el ejercicio 11.76



11.77 Calificación de solicitantes de profesor Antes de llenar varias nuevas posiciones de profesores en la secundaria, el director formó un consejo de revisión formado por cinco profesores a quienes se pidió que entrevistaran a 12 solicitantes y los clasificaran en orden de mérito. Siete de los 12 solicitantes tenían título universitario pero tenían poca experiencia en enseñanza. De los otros cinco solicitantes, todos tenían título universitario y considerable experiencia. Las clasificaciones de revisión del consejo se dan en la tabla siguiente.

Experiencia limitada	Experiencia considerable
4	1
6	2
7	3
9	5
10	8
11	
12	

¿Estas calificaciones indican que el consejo de revisión considera que la experiencia es un factor primordial en la selección de los mejores candidatos? Pruebe usando $\alpha = .05$.

DATOS
EX1178

11.78 Contaminantes en productos

químicos Un fabricante usa una gran cantidad de cierto producto químico. Como hay sólo dos proveedores de dicho producto, el fabricante desea probar si el porcentaje de contaminantes es el mismo para las dos fuentes, contra la alternativa de que hay una diferencia en los porcentajes de contaminantes para los dos proveedores. Los datos de muestras aleatorias independientes se dan a continuación:

Proveedor			
1		2	
.86	.65	.55	.58
.69	1.13	.40	.16
.72	.65	.22	.07
1.18	.50	.09	.36
.45	1.04	.16	.20
1.41	.41	.26	.15

- Use la prueba de la suma de rango de Wilcoxon para determinar si hay una diferencia en los porcentajes de contaminantes para los dos proveedores. Use $\alpha = .05$.
- Use la aproximación de muestra grande a la prueba de la suma de rango de Wilcoxon para determinar si hay diferencia en los porcentajes de contaminantes para los dos proveedores. Use $\alpha = .05$. Compare sus conclusiones con las conclusiones de la parte a.

11.79 Iluminación en el salón de clase Se observó y midió la productividad de 35 estudiantes tanto antes como después de la instalación de nuevo alumbrado en su salón de clases. Se observó que había mejorado la productividad de 21 de los 35 estudiantes, en tanto que la productividad del resto no pareció mostrar mejoría perceptible como resultado

de la nueva iluminación. Use la aproximación normal a la prueba del signo para determinar, al nivel de significancia de 5%, si la nueva iluminación fue o no eficaz para mejorar la productividad de los estudiantes.

DATOS
EX1180

11.80 Las marcas propias le ahorran \$ Ron

Marks⁷ visitó cinco cadenas de supermercados en Nueva York y Nueva Jersey y comparó los precios de las marcas propias y las marcas comerciales para los 30 artículos que se listan a continuación.

Producto	Marca comercial (\$)	Marca propia (\$)	Diferencia (\$)
Papel de aluminio	8.47	6.54	1.93
Frijoles cocidos	1.71	1.06	0.65
Migas de pan	2.03	1.22	0.81
Cuartos de mantequilla	4.03	3.03	1.00
Gajos de naranja enlatados	2.07	1.22	0.85
Aceite de canola	4.28	3.34	0.94
Jarabe sabor chocolate	4.13	3.22	0.91
Hisopos de algodón	3.53	1.98	1.55
Cóctel de jugo de arándano	2.82	2.42	0.40
Queso crema	2.05	1.35	0.70
Crema de cacahuete	2.85	2.05	0.80
Cuernitos congelados	2.71	1.98	0.73
Pasta seca	1.24	0.87	0.37
Cacahuates tostados sin aceite	3.53	2.79	0.74
Azúcar granulada	4.14	2.65	1.49
Jalea de uva	2.47	1.65	0.82
Pimienta negra molida	2.04	1.74	0.30
Half & Half (cuarto)	3.24	2.57	0.67
Harina para hot cakes	2.84	1.88	0.96
Jarabe para hot cakes	3.45	2.25	1.20
Pretzels	2.91	1.41	1.50
Arroz instantáneo	2.55	1.97	0.58
Cereal integral con pasas	3.65	2.63	1.02
Aderezo para ensaladas	2.55	2.02	0.53
Queso mozzarella rallado	3.29	2.33	0.96
Crema ácida	2.20	1.33	0.87
Mostaza parda condimentada	2.77	1.86	0.91
Salsa para filete	4.05	2.21	1.84
Sustituto de azúcar	2.70	2.11	0.59
Bolsas con cierre para emparedados	2.55	1.99	0.56

- Use la prueba del signo para determinar si los artículos de marca propia **cuestan menos** que sus contrapartes de marca comercial o no usando $\alpha = .01$.
- Use la prueba de rango con signo de Wilcoxon para determinar si los artículos de marca propia **cuestan menos** que sus contrapartes de marca comercial usando $\alpha = .01$.
- Use la prueba t en pares para determinar si el costo promedio de los artículos de marca propia **es menor** que el costo promedio de sus contrapartes de marca comercial usando $\alpha = .01$.
- ¿Las conclusiones son iguales para las tres pruebas? ¿Esperaría que lo fueran? ¿Por qué sí o por qué no?

DATOS
EX1181

11.81 Lego® El tiempo necesario para que niños de kínder ensamblen un juguete Lego específico se midió para niños que habían recibido instrucción durante cuatro periodos diferentes. Cuatro niños se asignaron al azar a cada grupo de instrucción, pero

dos fueron eliminados durante el experimento debido a enfermedad. En el experimento, el tiempo (en minutos) para ensamblar el juguete Lego fue registrado para cada niño.

Tiempo de instrucción (horas)			
.5	1.0	1.5	2.0
8	9	4	4
14	7	6	7
9	5	7	5
12		8	

Use la prueba H de Kruskal-Wallis para determinar si hay una diferencia en la distribución de tiempos para los cuatro tiempos de instrucción diferentes. Use $\alpha = .01$.

DATOS
EX1182

11.82 Fatiga de trabajadores Para investigar métodos para reducir la fatiga entre los empleados cuyos trabajos comprenden un monótono procedimiento de ensamblaje, a 12 empleados seleccionados al azar se les pidió que realizaran su trabajo normal en una de tres condiciones de prueba. Como medida de fatiga el experimentador usó el número de paradas de la línea de ensamblaje durante un periodo de cuatro horas para cada condición de prueba.

Empleado	Condiciones		
	1	2	3
1	31	22	26
2	20	15	23
3	26	21	18
4	31	22	32
5	12	16	18
6	22	29	34

7	28	17	26
8	15	9	12
9	41	31	46
10	19	19	25
11	31	34	41
12	18	11	21

- ¿Qué tipo de diseño experimental se ha empleado en este experimento?
- Use la prueba no paramétrica apropiada para determinar si la distribución de paradas de la línea de ensamble (y en consecuencia la fatiga del trabajador) difiere para estas tres condiciones. Pruebe al nivel de significancia de 5%.

11.83 Clasificación de mariscales de campo Se hizo una clasificación de mariscales de campo, de los ocho mejores equipos de la National Football League, al encuestar a varios periodistas de deportes y entrenadores profesionales. Esta “clasificación verdadera” se muestra a continuación, junto con “mi clasificación”.

	Mariscal de campo							
	A	B	C	D	E	F	G	H
Clasificación verdadera	1	2	3	4	5	6	7	8
Mi clasificación	3	1	4	5	2	8	6	7

- Calcule r_s .
- ¿Los datos indican una correlación positiva entre mi clasificación y la de los expertos? Pruebe al nivel de significancia de 5%.

CASO PRÁCTICO

DATOS Huevos

¿Cómo está su nivel de colesterol?

A medida que los consumidores se han interesado cada vez más en tomar alimentos saludables, muchos productos “light”, “sin grasa” y “sin colesterol” aparecen en el mercado. Uno de esos productos es el sustituto de huevo congelado, producto sin colesterol que se puede usar para cocinar y hornear en muchas de las mismas formas que un huevo común y corriente, aunque no todas. Algunos consumidores incluso usan sustitutos de huevo para elaborar un aderezo de ensalada César y otras recetas que piden huevos crudos porque estos productos están pasteurizados y así eliminan problemas de contaminación por bacterias.

Desafortunadamente, los productos que hay actualmente en el mercado muestran marcadas diferencias en sabor y textura cuando se prueban en su preparación básica de huevos revueltos. A cinco miembros de un grupo de discusión, todos ellos expertos en nutrición y preparación de alimentos, se les pidió que calificaran tres sustitutos de huevo con base en sabor, apariencia, textura y si ellos comprarían el producto.⁸ Los jueces probaron los tres sustitutos de huevo y los calificaron en una escala de 0 a 20. Los resultados, que se ven en la tabla siguiente, indican que la calificación más alta, por 23 puntos, fue para el producto llamado Healthy Choice Egg (Huevo Selecto Saludable) de ConAgra, que los probadores unánimemente acordaron como los que más se asemejan a los que produce una gallina. El producto que quedó en segundo lugar, Morningstar Farms’ Scramblers, impresionó a varios de los probadores por su “sabor singularmente dulce ... semejante a zanahoria”. Por último, ninguno de los probadores indicó que estarían dispuestos a comprar el de Egg Beaters de Fleishmann, que fue descrito por los expertos como “aguado”, “resbaloso” y “desagradable”. Por extraño que parezca, estos

resultados son contrarios a una prueba similar de sabor realizada 4 años antes, en donde Egg Beaters fue considerada mejor que los sustitutos de huevo de la competencia.

Probador	Healthy Choice	Scramblers	Egg Beaters
Dan Bowe	16	9	7
John Carroll	16	7	8
Donna Katzl	14	8	4
Rick O'Connell	15	16	9
Roland Passot	13	11	2
Totales	74	51	30

Fuente: Datos de "Egg Substitutes Range in Quality", por K. Sakekel, The San Francisco Chronicle, 10 de febrero, 1993, p. 8. Copyright © 1993 San Francisco Chronicle.

1. ¿Qué tipo de diseño se ha empleado en este experimento de prueba de sabor?
2. ¿Los datos satisfacen las suposiciones requeridas para un análisis paramétrico de varianza? Explique.
3. Use la técnica no paramétrica apropiada para determinar si hay una diferencia significativa entre el promedio de calificaciones para las tres marcas de sustitutos de huevo.

Teoría de la respuesta al ítem

OBJETIVOS GENERALES

El objetivo central del presente capítulo es que el lector aprenda las nociones básicas de la teoría de la respuesta al ítem (TRÍ).

De igual forma, al término del mismo el lector podrá:

- Comparar la TRÍ con la teoría clásica de los tests.
- Explicar los principales supuestos conceptuales y estadísticos de la TRÍ.
- Interpretar el modelo de la ojiva normal de Lord.
- Identificar los modelos logísticos de uno, dos y tres parámetros: Rasch, Birbaum de la habilidad.
- Describir los procedimientos de estimación y los parámetros del ítem.
- Evaluar la bondad de ajuste de los modelos de la TRÍ.
- Explicar el modelo politómico dentro de la TRÍ.
- Analizar ejemplos de modelos multicategoriales para el desarrollo de escalas de medida.
- Describir las ventajas que ofrece la TRÍ.
- Interpretar la información relevante al graficar los modelos anteriores.

ÍNDICE DEL CAPÍTULO

- Introducción (12.1)
- Teoría clásica de los tests en la psicometría (12.2)
- ¿Qué ofrece la teoría de la respuesta al ítem? (12.3)
- Principales supuestos de la TRÍ (12.4)
- Estimación de parámetros del examinado y los ítems (12.5)
- Función de información (12.6)
- Evaluación de bondad de ajuste del modelo (12.7)
- Modelos politómicos de la teoría de la respuesta al ítem (12.8)
- Resumen (12.9)

12.1

INTRODUCCIÓN

Uno de los objetivos centrales de la psicometría moderna es el desarrollo de modelos formales y métodos que posibiliten la medición de variables psicológicas y educativas. En esta disciplina se parte del principio de que es empíricamente posible medir desempeños psicológicos, conocimientos, habilidades, aptitudes, rasgos, procesos, etc., tanto en los ámbitos educativo y laboral, como en otras áreas de investigación en ciencias sociales y del comportamiento.

La *teoría de la respuesta al ítem* (TRÍ) es una parte muy importante, activa y en desarrollo de la psicometría moderna. Representa un nuevo enfoque que permite superar las limitaciones de la *teoría clásica de los tests psicométricos* (Hambleton, Swaminathan y Rogers, 1991; Van der Linden y Hambleton, 1997; Muñiz, 1997). Al igual que la teoría psicométrica clásica, la TRÍ pretende obtener una puntuación interpretable que corresponda al nivel de capacidad de una persona, en una habilidad o rasgo que busca medirse a partir de un modelo; en este caso, centrado en las propiedades de los ítems individuales, más que en las características globales de la puntuación en un test, como se hacía tradicionalmente.

La psicometría contemporánea incluye las siguientes áreas generales:

- Teoría de la medición.
- Teoría clásica de los tests.
- Teoría de la generalizabilidad.
- Teoría de la respuesta al ítem.

Como recordará el lector, en el primer capítulo del libro conoció el concepto de *medir* desde una perspectiva representacional de la medida. En el desarrollo de la *teoría de la medición*, el problema de medir variables psicológicas se ha abordado también desde otro enfoque muy interesante, la “*teoría de medición conjunta*”, la cual ofrece métodos para identificar la estructura cuantitativa de dichas variables. La postura básica de la medición conjunta es que no pueden darse por sentadas las propiedades cuantitativas de una variable, sino que deben someterse a prueba empírica y, además, satisfacer las condiciones numéricas de un sistema conjunto.

Por su parte, la *teoría de la generalizabilidad* ha intentado ofrecer una alternativa que permita liberar a la teoría clásica de los siguientes problemas: el manejo del error unidimensional e indiferenciado, el paralelismo estricto entre tests, así como las diferentes interpretaciones del concepto de confiabilidad, al que trata de sustituir por el de generalizabilidad de la medida.

Es claro que una explicación a profundidad de estos temas, de los avances en la teoría de la medida y de otros tópicos en la psicometría moderna, van más allá de los objetivos del libro, por lo que en este capítulo sólo se expondrán los supuestos básicos de la teoría clásica de los tests, las limitaciones que presenta y las ventajas que ofrece la teoría de la respuesta al ítem.

A continuación, se presentan los principales supuestos, concepciones y procedimientos de la teoría clásica de los tests y los problemas que intentan superarse con la teoría de la respuesta al ítem.

12.2

TEORÍA CLÁSICA DE LOS TESTS EN LA PSICOMETRÍA

La teoría clásica de los tests (TCT) constituye un conjunto de supuestos, teorías y técnicas *ad hoc* para analizar las puntuaciones de grupos de individuos en los tests. Dicho análisis de medición sobre las diferencias individuales se realiza habitualmente a partir de ciertos supuestos de la concepción tradicional, conocida como *teoría de la puntuación verdadera*.

Supuestos básicos de la teoría de la puntuación verdadera

Los principales supuestos de esta teoría se explican en los siguientes puntos:

1. La puntuación observada (X) de un sujeto en un test se compone de dos variables latentes:
 - a) La puntuación teórica o verdadera (T).
 - b) El error de medida (E).

Dicha puntuación observada se refiere a todo el test. Entonces: $X = T + E$

2. La variable latente (E) es una variable aleatoria con una distribución normal, media de cero y una varianza finita. Se puede expresar como:

$$E \sim N(0, \sigma^2 E)$$

3. Las variables latentes T y E no están correlacionadas:

$$\varepsilon(T, E) = 0 \Rightarrow (R_{TE} = 0)$$

Esto significa que al aumentar el valor de T , el valor de E sigue procediendo de la misma distribución. Por tanto, se asume que el error es independiente de la magnitud real del objeto medido.

4. Cuando se realizan dos mediciones, se supone que los errores de ambas son independientes entre sí.

$$\varepsilon(E_1, E_2) = 0 \Rightarrow (R_{E_1 E_2} = 0) \therefore \text{son independientes.}$$

5. El error que se comete en una medida es independiente de la puntuación teórica (verdadera) de una segunda medición.

$$\varepsilon(E_1, T_2) = 0 \Rightarrow (R_{E_1 T_2} = 0) \therefore \text{son independientes.}$$

En resumen, la teoría de la puntuación verdadera es una teoría de componentes latentes, los cuales se asumen como independientes o no correlacionados.

Confiabilidad de un test

A partir de los supuestos anteriores, la teoría clásica de los tests define uno de los conceptos centrales de la psicometría: *la confiabilidad de la medida*. Esto se explica en términos generales como una correlación entre dos conjuntos de puntuaciones obtenidas con el mismo test o con formas paralelas del test. Así, el concepto de confiabilidad en la TCT se define como una proporción de varianza entre puntuaciones observadas y la varianza de la puntuación verdadera.

Se expresa de la siguiente forma:

$$\rho^2(x, T) = \frac{\sigma^2(T)}{\sigma^2(x)}$$

Si supone dos medidas paralelas, entonces su correlación sería:

$$\rho(x, x') = \frac{\sigma(x, x')}{\sigma(x) \sigma(x')} = \frac{\varepsilon[(T+E)(T'+E')]}{\sigma(x) \sigma(x')}$$

$$\frac{\varepsilon(T^2) + \varepsilon(TE) + \varepsilon(T'E) + \varepsilon(EE')}{\sigma(x) \sigma(x')} = \frac{\varepsilon(T^2)}{\sigma^2(x)} = \frac{\sigma^2(T)}{\sigma^2(x)} = \rho^2_{XT} = \rho^2_{xx'}$$

La confiabilidad puede estimarse a partir de dos medidas equivalentes o paralelas, o bien, mediante el procedimiento test-retest. Sin embargo, en la TCT deberán cumplirse ciertas condiciones de paralelismo, misma que se expondrán a continuación.

Condiciones de paralelismo

Considere dos mediciones independientes X y X' :

Dado que $X = T + E$ y $X' = T' + E'$, entonces tiene las condiciones siguientes.

A. Medidas equivalentes:

$$T = T' \text{ y } F(E) \equiv F(E')$$

En este caso, las puntuaciones verdaderas son iguales y la distribución de errores es equivalente; es decir, los errores proceden de la misma distribución. Es una condición muy exigente y muy difícil de conseguir en la práctica psicométrica.

B. Medidas paralelas:

$$T = T' \text{ y } \sigma^2(E) = \sigma^2(E')$$

Las puntuaciones verdaderas son iguales y las distribuciones de los errores de las dos mediciones tienen varianzas iguales. Esta condición implica que la precisión de la medida debe ser la misma, o que la varianza del error es igual en las dos.

C. Medidas ζ – Equivalentes:

$$T = T' \text{ y } \sigma^2(E) \neq \sigma^2(E')$$

Las puntuaciones verdaderas son iguales, pero las dos medidas se llevan a cabo con distinta precisión; es decir, las varianzas de los errores son diferentes.

D. Medidas esencialmente ζ – Equivalentes

$$T = K + T', \text{ siendo } K = \text{Constante y } \sigma^2(E) \neq \sigma^2(E')$$

Las dos medidas tienen distinta precisión y se trata de la aplicación de tests de diferente longitud y también con distinto rango de dificultades de los ítems. Por tanto, en este caso las puntuaciones estarán en escalas distintas y entonces se requiere agregar una constante.

Características de los ítems en la TCT

Además de cumplir las condiciones de paralelismo para estimar la confiabilidad, el esquema básico en la TCT incluye un modelo correlacional, donde cada ítem se correlaciona con un criterio, generalmente con la puntuación total del test. Así, los parámetros que definen las propiedades de los ítems de un test en la TCT son:

a) La discriminación del ítem:

Índice de discriminación D = Correlación biserial puntual entre el ítem y la puntuación total del test.

b) La dificultad del ítem:

Índice de dificultad P = Proporción de sujetos que aciertan el ítem.

En la TCT puede haber variaciones importantes en los parámetros debido al tipo de ítems aplicados en una prueba o a cambios en la muestra; es decir, los valores pueden variar dependiendo de los ítems de acompañamiento que se incluyan o de las características del grupo de examinados. Como se explicará la estimación de parámetros de los ítems en la TRÍ no dependerán de la puntuación total del test, ni de los examinados en un grupo normativo que contesta cada ítem.

Principales limitaciones de la teoría clásica de los tests

El análisis de los estadísticos de la TCT ha resultado útil durante muchos años; sin embargo, se ha identificado una serie de problemas de carácter psicométrico:

1. La magnitud medida con un test no es invariante respecto del instrumento utilizado, ésta depende del test particular que se aplique. Por lo que la escala del rasgo medido varía de un test a otro similar.
2. Los parámetros de los ítems no son invariantes respecto de la muestra utilizada; el nivel de dificultad depende del nivel del grupo muestral al que se le aplique el test.
3. Los parámetros de un ítem no son invariantes respecto del resto de ítems del test utilizado. Dado que el criterio es el puntaje global del test, no existe independencia local entre los ítems de una prueba.
4. La puntuación verdadera (T) y el parámetro de dificultad (P) no forman parte del mismo modelo estadístico y no están en la misma escala. Por tanto, no son comparables.
5. La confiabilidad del test depende directamente de la variabilidad de las puntuaciones de la muestra utilizada. Entonces, las estimaciones de confiabilidad se obtienen con sesgo.
6. En la TCT, la confiabilidad está definida bajo la condición de formas paralelas, pero son muy difíciles de obtener en la práctica, por lo que en la aplicación de tests se obtienen cotas mínimas de confiabilidad.
7. En la TCT no se pueden comparar las puntuaciones del test con instrumentos diferentes, a menos que sean medidas estrictamente paralelas; por lo que se requiere un procedimiento de *equiparación*: hacer la transformación lineal en los dos tests para igualar la escala.
8. La varianza del error de medida se supone que es la misma para todos los sujetos, pero la varianza observada no es igual en el centro.
9. El análisis y la estimación de los parámetros de los ítems no forman parte de los supuestos estadísticos y procedimientos de la teoría de la puntuación verdadera.
10. En suma, la teoría clásica de los tests es útil, pero presenta, tanto en su aplicación como en el análisis teórico, un conjunto de problemas de medición que no ha podido resolver de manera satisfactoria.

¿QUÉ OFRECE LA TEORÍA DE LA RESPUESTA AL ÍTEM?

12.3

En la teoría de respuesta al ítem (TRÍ) los parámetros de cada reactivo no dependen de otros ítems con los que se presente en una prueba. La puntuación se obtiene en función de la respuesta del sujeto a cada ítem y de los parámetros del ítem. Así, esta teoría permite crear bancos de ítems previamente calibrados.

Una segunda ventaja es que en la TRÍ lo que se mide es la habilidad medida por el desempeño de una persona que contesta a un ítem. Por ello, se tiene que contar con las respuestas individuales a cada ítem, a fin de inferir el nivel de la variable psicológica que se está midiendo.

En tercer término, en la TRÍ un modelo de un rasgo latente (variable inferida) especifica una relación entre las respuestas observables de una persona en un test y el rasgo no observable o habilidad que es la que determina que responda correctamente en una prueba, esto es de su desempeño.

Otra aportación importante en los modelos de la TRÍ es que describen la relación entre el rasgo latente y la probabilidad de la respuesta correcta a un ítem, la cual se representa en una función llamada curva característica del ítem (CCI).

La probabilidad de que una persona proporcione una respuesta correcta a un ítem es independiente de la distribución del rasgo latente en la población de estudio. Es decir, la probabilidad de una respuesta correcta del examinado no dependerá de las respuestas de otros examinados en algún punto del continuo del rasgo. Por ello, la forma de la curva característica del ítem será invariante a través de diferentes muestras de examinados, cuando el modelo se ajusta a los datos.

En el siguiente cuadro se hace un resumen de los aspectos psicométricos que reflejan las principales diferencias de contraste entre la teoría clásica de los tests y la teoría de la respuesta al ítem.

	Teoría clásica de los tests	Teoría de respuesta al ítem
Criterio de medición	La puntuación global del test	La respuesta del examinado ante cada ítem
Medida del rasgo	La magnitud del rasgo medido varía de un test a otro	Invarianza en la medición de la habilidad con diferentes ítems
Tipo de modelo	Aplica un modelo lineal	Modelos logísticos y de ojiva normal
Estimación parámetros	Correlación biserial puntual y proporción de respuestas	Máxima verosimilitud conjunta, condicional o marginal
Parámetros (muestra)	Los valores de los parámetros varían según la muestra	Los parámetros de los ítems son invariantes al cambiar la muestra
Parámetros (entre ítems)	Los parámetros de los ítems varían con otros ítems	Los parámetros de cada ítem son independientes de otros ítems
Escala de parámetros	La puntuación del rasgo (T) y nivel de dificultad (P) del ítem no están en la misma escala	La habilidad medida (θ) se ubica en la misma escala de la dificultad del ítem (b)
Confiabilidad	La confiabilidad del test depende de la variabilidad de puntuaciones en la muestra (con sesgo) de información	Estimación estadística y gráfica de precisión de la medida mediante la función
Estimación del error	La varianza de error es la misma para todos los sujetos en todas las puntuaciones obtenidas	La estimación del error de medida es diferente para cada nivel de habilidad

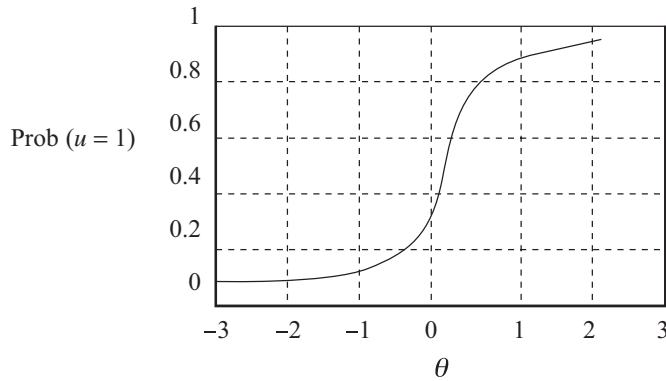
Curva característica del ítem (CCI)

La curva característica del ítem (CCI) es una función monotónica creciente que relaciona la probabilidad de éxito en un ítem con la habilidad medida (rasgo latente) mediante la aplicación de un test psicométrico. El rasgo latente es el responsable de que se responda bien o mal a un ítem. Como se sabe existe mayor o menor probabilidad para responder correcta o incorrectamente a un ítem particular, dependiendo del nivel de habilidad.

Un rasgo inferido, es decir la variable que se busca medir tiene origen y una unidad arbitrarios en un continuo que puede ir desde $-\infty$ hasta $+\infty$. El rasgo latente de habilidad que se designa como θ corresponde al eje X (véase figura 12.1). En el eje Y se representa la probabilidad de “theta” $P(\theta)$, es decir, la probabilidad de responder correctamente en un ítem, dado el nivel θ del rasgo medido. Como se señala en la gráfica este continuo se representa como (theta).

FIGURA 12.1

Curva característica del ítem (CCI)

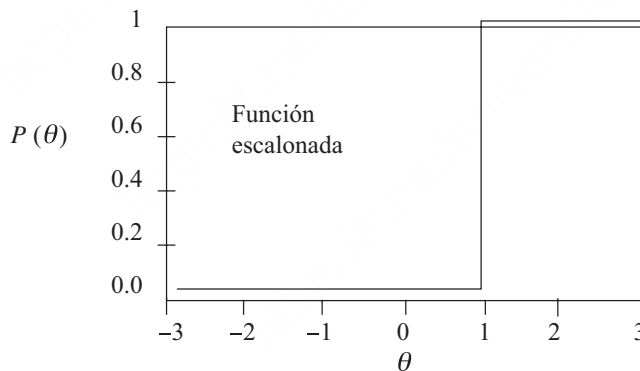


Modelo ideal de Guttman y parámetros de un ítem

A mediados del siglo xx Guttman planteó la llamada “escala perfecta”, en la cual un ítem ideal sería aquél que correspondiera a una función discontinua, escalonada, en la cual los sujetos con una capacidad menor a la dificultad del ítem tendrían una probabilidad de 0 de responder correctamente al ítem; y aquellos con capacidad mayor a ese ítem tendrían una probabilidad de 100% de contestarlo correctamente. A diferencia de la función psicofísica para medir umbrales sensoriales en la cual se observa un proceso gradual, en este modelo la probabilidad de responder correctamente al ítem pasaría de 0 a 1 en un punto.

FIGURA 12.2

Escala de Guttman



El modelo de Guttman y la función psicofísica constituyen antecedentes importantes y un punto de partida para entender los modelos de la TRÍ que representan funciones no lineales, monotónicas crecientes, en forma de “S”. En contraste con la TCT, en la teoría de respuesta al ítem los parámetros de los ítems adquieren una significación congruente con el modelo teórico, se encuentran en el mismo modelo, están en la misma escala y se estiman de manera diferente.

Índice de dificultad

El índice de dificultad de un ítem b_i es una estimación de la posición del ítem en la escala de θ . Se trata de un estadístico que informa en qué lugar de esta escala se parten las probabilidades de contestar correctamente ese ítem en particular. Este parámetro indica entonces la parte de la escala θ donde se da la transición desde una mayor probabilidad de responder incorrectamente

al ítem hasta una mayor probabilidad de hacerlo correctamente. El índice de dificultad b_i es el valor de θ para el que la probabilidad de responder correctamente el ítem i es de 0.5

$$P(x = 0 \mid \theta < B) < 0.5 \text{ y por encima de } b_i; P(x = 1 \mid \theta > B) > 0.5.$$

El parámetro b_i es entonces el punto de inflexión de la curva, y significa que el valor de la derivada primera es el máximo en ese punto dado. El índice de dificultad de un ítem señala en qué parte de la escala θ se encuentra el punto de inflexión de la CCI para ese ítem. Es el valor de la posición del ítem en la escala theta.

Discriminación de un ítem

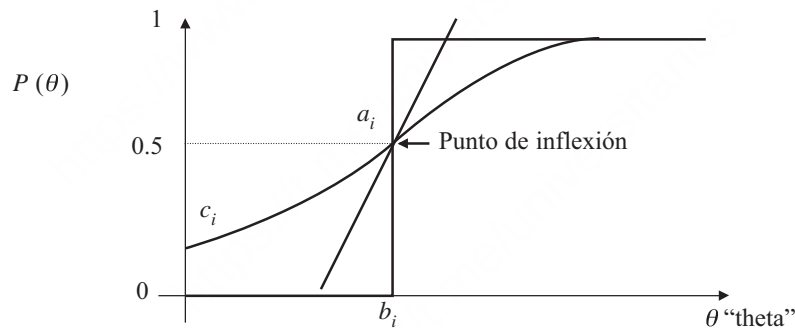
En la TRÍ, la discriminación de un ítem a_i es un valor directamente proporcional a la derivada primera de la curva en el punto de inflexión; es decir, se refiere a la pendiente de la curva en ese punto, definida como la tangente del ángulo α :

$$a_i = \tan \alpha \cdot \kappa$$

Parámetro de seudoadivinación

El tercer parámetro se refiere al parámetro C_i como la probabilidad de contestar el ítem por adivinación, o identificando una pista en la opción correcta del ítem, aún por abajo de la capacidad o conocimientos del sujeto examinado.

FIGURA 12.3
Parámetros de la curva característica del ítem



Modelo de ojiva normal

Considere una variable continua Y , la cual representa la respuesta de un sujeto ante un ítem que pretende medir la capacidad θ de ese sujeto, entonces la relación entre las dos variables estaría dada por el coeficiente β , que sería la pendiente de la recta que representa dicha relación. Así, dada la capacidad θ de un sujeto, ¿cuál sería la probabilidad de que responda correctamente al ítem? Esta misma pregunta puede expresarse de este modo:

Dada la capacidad θ , ¿cuál es la probabilidad de que el sujeto tenga una puntuación en Y superior al umbral y ?

Estará estimada por la integral definida de la curva normal de probabilidades, en el rango que va desde el umbral y hasta ∞ .

Por otro lado, si β representa la correlación entre las variables θ y Y , entonces β^2 sería la varianza explicada.

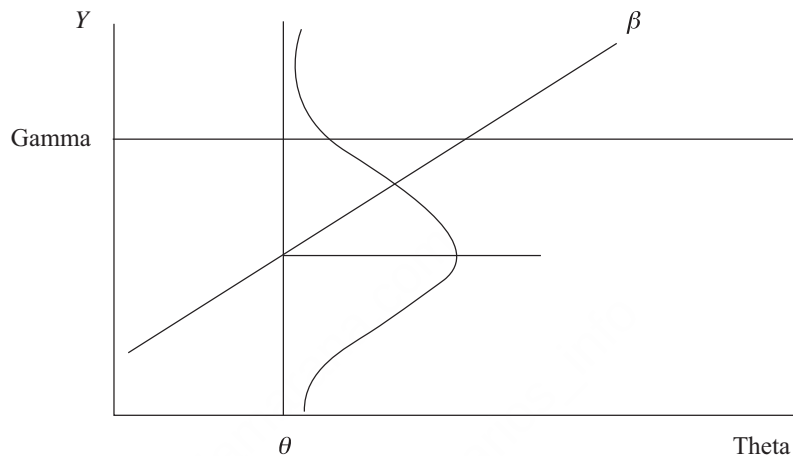
$$\therefore Y = 1 - \beta^2 \text{ sería la varianza No explicada} \Rightarrow \sigma^2 = 1 - \beta^2$$

Para un sujeto j , dada una puntuación en θ , ¿cuál es su probabilidad de responder correctamente al ítem i ? Será el área comprendida en la distribución por encima del umbral de corte y .

Se puede expresar así: $P(U_{ij} = 1 \mid \theta = \theta_j) = P(Y > y \mid \theta = \theta_j)$

$$= \int_y^{\infty} \frac{1}{\sigma \sqrt{2\pi}} \cdot \exp \left[\frac{-(Y - \beta\theta)^2}{2\sigma^2} \right] dY$$

FIGURA 12.4
Parámetros del modelo de
ojiva normal



La expresión del modelo de ojiva normal está en función de los parámetros y y β .

Reparametrización del modelo de ojiva normal

El parámetro y del modelo de ojiva normal tiene que ver con la dificultad del ítem, debido a que la probabilidad de responderlo disminuirá si se eleva ese valor; o bien, será más probable contestarlo correctamente si el valor de y baja. Pero también la forma en que aumente o disminuya ese valor depende de la pendiente de la recta de regresión β , por lo que los dos parámetros están interrelacionados.

El parámetro b_i tiene que ver así con la discriminación entre sujetos. Como recordará el lector, el parámetro de dificultad b_i se definió como la puntuación en θ para la probabilidad media de responder correctamente al ítem en 0.5.

Entonces $b_i = \text{valor de } \theta, \text{ tal que } P(U_{ij} = 1 \mid \theta) = 0.5)$

Cuando $\beta \theta = y$, tiene que $\beta b_i = y$

Entonces: $b_i = \frac{Y}{\beta}$

Otra forma de obtener una expresión más directa del modelo es cambiando o tipificando las variables.

Si $t = \frac{Y - \beta\theta}{\sigma}$

$$Y = \sigma t + \beta \theta$$

$$dY = \sigma dt$$

$$\frac{dY}{dt} = \frac{d(\sigma t)}{dt} + \frac{d\beta\theta}{dt} = \sigma + 0 \Rightarrow \frac{dY}{dt} = \sigma$$

Sustituyendo en la expresión de la ojiva normal tendría:

$$P(U_{ij} = 1) | \theta = \theta_j = \int_{-\infty}^{\infty} \frac{-Y - \beta\theta}{\sigma} \frac{1}{\sigma \sqrt{2\pi}} \cdot \exp\left(\frac{(-t)^2}{2}\right) \sigma dt$$

Cambiando de orden el límite superior:

$$\begin{aligned} & \int_{-\infty}^{\infty} \frac{-Y - \beta\theta}{\sigma} \frac{1}{\sigma \sqrt{2\pi}} \cdot \exp\left(\frac{(-t)^2}{2}\right) \sigma dt \\ &= \frac{\beta\theta - y}{\sigma} = \frac{\beta\theta}{\sigma} - \frac{y}{\sigma} = \frac{\beta}{\sqrt{1-\beta^2}} \theta - \frac{y}{\sqrt{1-\beta^2}} \\ &= \frac{\beta}{\sqrt{1-\beta^2}} \theta - \frac{\beta}{\sqrt{1-\beta^2}} \frac{y}{\beta} = \frac{\beta}{\sqrt{1-\beta^2}} \left(\theta - \frac{y}{\beta} \right) = \frac{\beta}{\sqrt{1-\beta^2}} (\theta - b) \end{aligned}$$

Si al elemento $\frac{\beta}{\sqrt{1-\beta^2}}$ le llama “a”,

entonces $\frac{\beta}{\sqrt{1-\beta^2}} (\theta - b) = a (\theta - b)$ [límite superior de la integral definida]

Con los nuevos parámetros de dificultad del ítem y a el índice de discriminación, tiene las expresiones:

$$b_i = \frac{y}{\beta} \quad a = \frac{\beta}{\sqrt{1-\beta^2}}$$

Entonces, el modelo de ojiva normal, con los nuevos parámetros, sería *para dos parámetros*:

$$P(U_{ij} = 1 | \theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a(\theta-b)} \exp\left(\frac{(-t)^2}{2}\right) dt$$

Así, la probabilidad de responder correctamente a un ítem dado θ representaría la diferencia entre la capacidad de un sujeto para ello y la dificultad de un ítem, además del nivel de discriminación de éste.

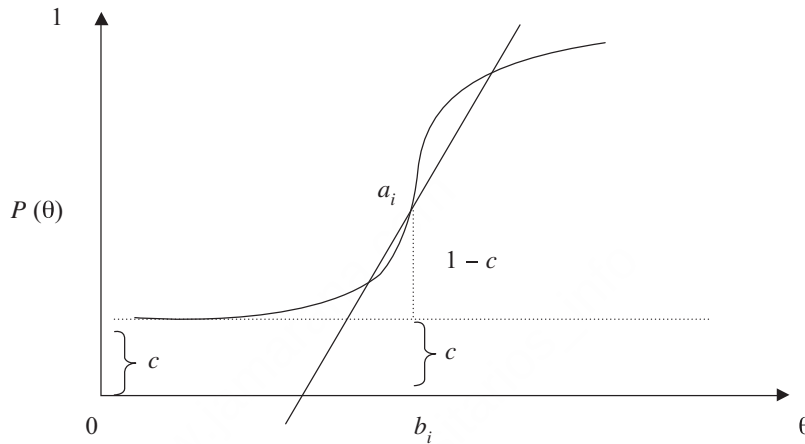
En el modelo de tres parámetros, tendría una expresión más general, donde se consideraría, además de los dos parámetros anteriores, que puede haber cierto nivel de “adivinación” o “seudo-azar”.

La expresión para el modelo de ojiva normal de tres parámetros sería:

$$P_i(\theta) = C_i + \frac{1 - C_i}{2\pi} \int_{-\infty}^{a(\theta - b)} \exp \frac{(-t^2)}{2} dt$$

El tercer parámetro, denominado “c”, refleja el hecho de que los sujetos de menor capacidad pueden tener cierta probabilidad (mayor que cero) de responder correctamente al ítem. Entonces, C_i sería la asíntota inferior de la función del ítem i .

FIGURA 12.5
Parámetro c en la curva característica del ítem



Modelo logístico de un parámetro o modelo de Rasch

El modelo de Rasch, o función logística de un parámetro, se ha utilizado ampliamente por las ventajas teóricas y prácticas que representa en comparación con el modelo de ojiva normal. Dicho modelo fue propuesto originalmente por el matemático George Rasch en la década de 1960 y posteriormente fue desarrollado por otros investigadores en psicometría en Estados Unidos.

En el modelo de Rasch se concibe la respuesta de un sujeto a un ítem como una conducta que depende de una sola característica del ítem: el nivel de dificultad. Y de una sola característica del sujeto: el nivel de capacidad o habilidad.

Como se sabe, las probabilidades de acertar a un ítem en relación con las de fallar, para un sujeto J , están dadas por el cociente de las dos cantidades: una, referida a la capacidad del sujeto y la otra por la dificultad del ítem i .

$$\frac{P_i}{Q_i} = \frac{\theta^*}{b^*} \quad \text{donde: } \theta^* \text{ es la cantidad del sujeto (su capacidad)} \\ b^* \text{ es la cantidad el ítem (su dificultad)}$$

El cociente anterior variará entre $\frac{0}{1}$ y $\frac{1}{0}$

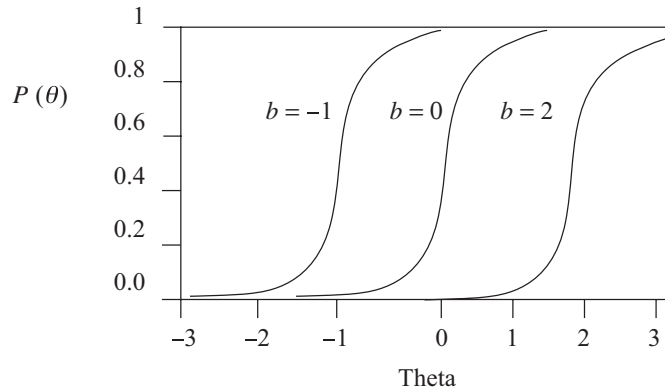
$$Q_i = 1 - P_i \quad \therefore \quad \frac{\theta^*}{b^*} = \frac{P_i}{1 - P_i} \Rightarrow P_i b^* = (1 - P_i) \theta^*$$

$$(1 - P_i) \theta^* = \theta^* - P_i \theta^* \Rightarrow \theta^* = \theta^* P_i + P_i b^* = P_i (\theta^* + b^*)$$

$$P_i(\theta^* + b^*) = \theta^* \Rightarrow \text{Despejando: } P_i = \frac{\theta^*}{(\theta^* + b^*)}$$

FIGURA 12.6

Modelo de un parámetro
(Rasch)



En la gráfica se observa una representación del modelo de un parámetro; se muestran tres ítems con distinto grado de dificultad, pero con la misma pendiente. El índice de discriminación es constante e igual que 1 en los tres ítems. En este modelo sólo se estima el nivel de dificultad en función del nivel de θ .

Transformación exponencial de la función logística en el modelo de Rash:

Si $\theta^* = \exp(\theta)$ y si $b^* = \exp(b) \Rightarrow$ sustituyendo:

$$P_i(\theta) = \frac{\exp(\theta)}{\exp(\theta) + \exp(b)} = \frac{\frac{\exp(\theta)}{\exp(\theta)}}{\frac{\exp(b)}{\exp(\theta)} + 1} = \frac{1}{1 + \exp(b - \theta)}$$

$$\frac{1}{1 + \exp(b - \theta)} = \frac{1}{1 + \exp[-(\theta - b)]} \Rightarrow$$

$$P_i(\theta) = \frac{\exp(\theta - b)}{1 + \exp(\theta - b)}$$

$$\text{Por tanto: } P_i(\theta) = \frac{\exp(D(\theta - b))}{1 + \exp(D(\theta - b))}$$

En esta última ecuación, la cual representa el modelo de Rasch, se ha incluido una constante de escalamiento D cuyo valor es 1.7. Esta constante hace que la función logística coincida con la curva obtenida en el modelo de ojiva normal.

El modelo de Rasch tiene la ventaja, a diferencia de la función de ojiva normal, de que es más sencillo y parsimonioso, porque no es necesario integrar para estimar los parámetros de la función y el número de elementos a estimar es menor.

El modelo de Rasch supone dos condiciones:

- Que no se producen respuestas acertadas por “adivinación”.
- Que todos los ítems de una prueba tienen el mismo nivel de discriminación.

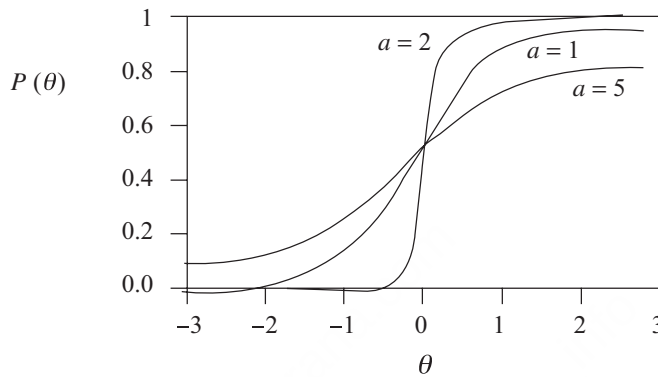
Modelo logístico de dos parámetros

En este modelo de dos parámetros, o modelo de Birnbaum, además del parámetro b_i que se refiere a la dificultad del ítem, en cada reactivo se tiene un índice diferente de discriminación a .

Siguiendo la derivación del modelo anterior, se tendría que:

$$P_i(\theta) = \frac{\exp [a (\theta - b)]}{1 + \exp [-a(\theta - b)]} = \frac{1}{1 + \exp [Da (\theta - b)]}$$

FIGURA 12.7
Modelo de dos parámetros



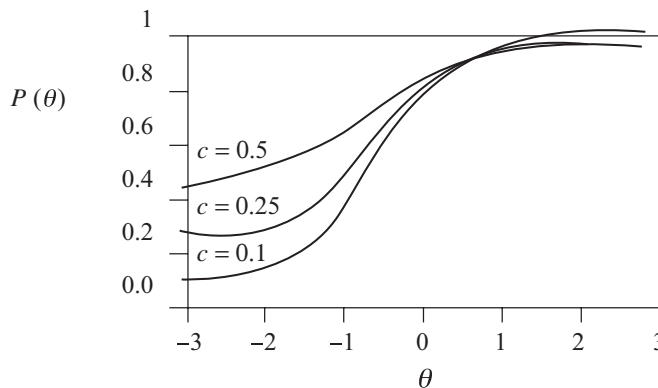
En estas curvas características del ítem se observan dos ítems con distinto índice de dificultad y diferentes índices de discriminación. Este modelo, como en el anterior, supone que no hay aciertos por adivinación y, con ello, que no existe en la ecuación el parámetro que refleje esta posibilidad.

Modelo logístico con tres parámetros

El tercer parámetro en los modelos logísticos supone el hecho de que los sujetos examinados de más baja capacidad tienen una probabilidad mayor que cero de responder correctamente al ítem. Lo anterior necesariamente debe interpretarse como elección al azar de la opción correcta; el valor de (c) suele ser menor que el valor esperado por simple azar. Por eso se le denomina parámetro de “seudoadivinación” o “seudo-azar”. La expresión del modelo de tres parámetros sería:

$$P_i(\theta) = c + (1 - c) \frac{1}{1 + \exp [Da(\theta - b)]}$$

FIGURA 12.8
Modelo de tres parámetros



En esta gráfica se ilustra el modelo de tres parámetros; se muestran tres ítems con un nivel de dificultad similar, pero con diferentes niveles de discriminación y con valores diferentes en el parámetro “ c ” de “seudoadivinación”.

PRINCIPALES SUPUESTOS DE LA TRÍ

La TRÍ asume que existe una relación funcional entre los valores de la variable que se busca medir con los ítems y la probabilidad de contestarlos correctamente; por lo que en los modelos de la TRÍ, la probabilidad de acertar a un ítem depende (si el modelo es adecuado) de un factor: el nivel de θ . Otros supuestos son:

Unidimensionalidad del test

En la teoría de la respuesta al ítem debe demostrarse que el conjunto de ítems que miden la variable θ constituye una sola dimensión, es decir, debe ser unidimensional. Entre los posibles métodos para comprobar que un conjunto de ítems son de esta índole, está el *análisis factorial*, que sigue siendo uno de los más utilizados. Mientras más alta sea la varianza que explique el primer factor, se identificará la existencia de *unidimensionalidad* de la prueba.

Indeterminación de la escala del rasgo latente

El rasgo latente θ es la característica de un sujeto que explica su desempeño al resolver un ítem. En la TRÍ, la escala en la que se define el rasgo está indeterminada. Es decir, el rasgo latente es una magnitud no observable que tiene origen y unidades arbitrarios. Para solucionar esta indeterminación, usualmente se atribuye a la escala de θ una media de cero y desviación estándar de uno.

Así, los valores de θ como los valores de b sirven para calcular $P(\theta)$. Y la distancia de $\theta - b$ es útil para obtener ese valor. Si se cambia el valor de θ , pero también el valor de b_i , la probabilidad será la misma. La implicación práctica es que debe haber elementos comunes para poder hacer una comparación entre medidas y usar una misma escala que sea comparable. Independientemente de la ubicación de la escala, en la TRÍ los parámetros de la curva característica del ítem se mantienen invariantes.

La independencia local

El supuesto de *independencia local* se refiere a que las respuestas de una persona ante los ítems a los que se somete en una prueba son estadísticamente independientes entre sí. Esto significa que el rendimiento de un sujeto a un ítem no está determinado por los resultados obtenidos en otros ítems de la prueba.

La respuesta de un sujeto a cada ítem sólo está relacionada con el rasgo que mide y nada más, por ninguna otra cosa. Por tanto, dos ítems no deben compartir ninguna característica que afecte la probabilidad de la respuesta correcta del sujeto, más allá de aquellas que supuestamente mide el ítem.

El supuesto de independencia local se cumple si la correlación parcial entre todos los posibles pares de ítems, eliminando el influjo común del factor medido, es 0. Un conjunto de ítems es unidimensional cuando el número de rasgos necesarios para hacer cero las correlaciones parciales entre los ítems, es igual a 1 (un rasgo).

Sea $\theta_j = (\theta_{1j}, \theta_{2j}, \theta_{3j}, \dots, \theta_{pj})$ Vector de rasgos latentes a un conjunto de ítems.

Si V_i = Variable que puede adoptar sólo dos valores: $u_{ij} = \begin{cases} 1 \\ 0 \end{cases}$

Si se cumple el supuesto de independencia local, entonces

$$P(V_1 = u_1, V_2 = u_2, V_3 = u_3, \dots, V_p = u_p \mid \theta)$$

Matemáticamente, puede expresarse como un producto de probabilidades:

$$P(V_1 = u_1 | \theta) \cdot P(V_2 = u_2 | \theta) \cdot P(V_3 = u_3 | \theta) \dots P(V_i = u_i | \theta) \dots P(V_p = u_p | \theta)$$

12.5

ESTIMACIÓN DE PARÁMETROS DEL EXAMINADO Y LOS ÍTEMS

La teoría de la respuesta al ítem supone que la respuesta a cada ítem depende sólo del nivel de destreza o habilidad de un sujeto que se expone a una prueba. Así, para cada sujeto es necesario estimar sólo un parámetro individual: θ

Entonces si aplica un test de cierto número de ítems a N sujetos, el número de parámetros de la habilidad que debe estimar, es N , uno por cada sujeto. El número de parámetros de los ítems dependerá del modelo elegido.

Método de estimación de máxima verosimilitud

En la TRÍ se cuenta con varios métodos de estimación, tales como la máxima verosimilitud conjunta, condicional o marginal. Cuando se cuenta con un banco de ítems calibrados, en el que se conocen las características de los parámetros de cada ítem, se puede utilizar entonces la función de verosimilitud para estimar los parámetros individuales.

$$P(U) = P_1^{u_1} \cdot Q_1^{(1-u_1)} \cdot P_2^{u_2} \cdot Q_2^{(1-u_2)} \cdot P_3^{u_3} \cdot Q_3^{(1-u_3)} \dots P_n^{u_n} \cdot Q_n^{(1-u_n)} = L$$

Por ejemplo, P es una función de θ , a y b en el modelo de dos parámetros. Entonces, el logaritmo neperiano de la función de verosimilitud viene dado de la siguiente forma:

$$\ln L = U_1 \cdot \ln P_1 + (1-U_1) \ln Q_1 + U_2 \ln P_2 + (1-U_2) \ln Q_2 + \dots U_n \cdot \ln P_n + (1-U_n)$$

$$\ln L = f(\theta) \text{ debido a que: } P_1 = P(U_1 = 1 | \theta)$$

$$Q_1 = P(U_1 = 0 | \theta)$$

Esta función tiene los máximos para los mismos valores de θ , a y b que la función de verosimilitud.

Para identificar el máximo de la función, se iguala la primera derivada de θ a cero.

$$\frac{d \ln L}{d \theta} = 0 \quad \text{Si se resuelve esta ecuación y obtiene el valor de } \theta \text{ que la satisfaga, se obtiene el estimador de máxima verosimilitud.}$$

Como se trata de estimar todos los valores de θ , hay que usar métodos numéricos, debido a que es difícil obtener soluciones analíticas de las ecuaciones. Así, los parámetros se estiman, por ejemplo, mediante el método de Newton-Raphson.

Otros métodos de estimación de parámetros, además del de máxima verosimilitud conjunta, condicional o marginal, son los bayesianos. En términos generales, a partir de la distribución *a priori* de θ y de la función de verosimilitud, se obtiene la distribución *a posteriori*, de $L(q) \cdot \varphi(q)$. Entonces, se obtiene una función con un máximo definido más preciso.

Estimación de los parámetros: a y b

$$1) \hat{a} = \frac{rbp}{\sqrt{1 - \sqrt{bp}}} . \text{ Sería la aproximación por índices de la teoría clásica}$$

$$2) \hat{b} = \frac{Z_i}{rbp} . \text{ En la TRÍ, es importante recordar que } a = \frac{\beta}{\sqrt{1 - \beta^2}} \text{ y que } b = \frac{y}{\beta} .$$

La estimación de los parámetros se realiza por aproximaciones sucesivas (iteraciones). Si se desconocen los valores de los parámetros de los ítems, se inicia a partir de un valor aproximado que puede darse como cierto. A partir de este valor inicial se estiman los valores de los parámetros de los ítems; y entonces, con esos valores se calcula el valor de los nuevos parámetros individuales. A partir de los parámetros individuales se estiman de forma reiterada los parámetros de los ítems. El proceso de iteraciones se repite hasta que los valores de los parámetros converjan y se detenga el proceso de estimación.

12.6

FUNCIÓN DE INFORMACIÓN

La teoría de la respuesta al ítem permite estimar el valor de θ y también puede proporcionar una medida de precisión de las estimaciones en la medida del rasgo.

La cantidad de información que da la CCÍ de un ítem en un punto dado, depende de dos aspectos básicos:

1. De qué tan bien el ítem es capaz de discriminar entre los sujetos alrededor de ese punto (pendiente de la curva), es decir, la derivada primera en ese punto.
2. De qué tan precisas son las predicciones de la curva característica del ítem en ese punto, es decir, de la varianza.

Mientras más pendiente, mayor discriminación; esto es, el ítem ofrece más información. Y mientras más varianza, mayor error y menos información.

$$i(\theta_j) = \frac{1}{\sigma_{\theta^2 j}}$$

Por tanto, el concepto de cantidad de información para un determinado valor de θ se define como el inverso de la varianza de sus errores de medida en ese valor.

Así, la función de información viene dada por:

$$i(\theta) = \frac{P'_i(\theta)^2}{P_i(\theta) Q_i(\theta)}$$

Usos de la función de información

La función de información en la TRÍ tiene una utilidad similar al coeficiente de confiabilidad de la teoría clásica, en cuanto a que ofrece una estimación sobre la precisión de la medida y el nivel de error. Pero aquí la función de información es un concepto más fuerte.

En el caso de la TRÍ, la función de información sólo depende de la relación de cada ítem con el rasgo que pretende medirse. Los principales usos de la $F(I)$ son:

1. La función de información es un indicador que permite conocer la precisión con que mide cada ítem en un test.
2. Al calcular todos los niveles de θ se obtienen las funciones de información que sirven para construir y seleccionar los tests más adecuados, con el máximo de información en el punto de corte.
3. La función de información también sirve para aplicar tests adaptativos por computadora. Este procedimiento se ahorra la presentación de ítems intermedios o extremos y sólo selecciona los más informativos en el nivel de rasgo de cada individuo que está evaluándose.
4. En la teoría de la respuesta al ítem, la función de información tiene una función similar al coeficiente de confiabilidad en la TCT. Como se ha señalado, este concepto es mucho más robusto y sustituye a la confiabilidad clásica.
5. La función de información de un test toma valores distintos para cada nivel del rasgo que se mide. Así, se cuenta con una herramienta más precisa y poderosa para la construcción de pruebas.

Función de información del test

La función de información de un test es la suma de las funciones de información de los ítems que lo conforman. Las funciones de información de un test y de cada ítem están dadas por:

En un test: $X = \sum_{i=1}^n U_{ij}$, donde n = número de ítems

Entonces, para el sujeto $J \Rightarrow$ como $U_{ij} \begin{matrix} < 0 \\ < 1 \end{matrix}$

$$\varepsilon(U_{ij} | \theta) = 1 P(U_{ij} = 1 | \theta_j) + 0 P(U_{ij} = 0 | \theta_j)$$

$$P(U_{ij} = 1 | \theta_j) \quad [\text{curva característica del ítem}]$$

$$\varepsilon(X | \theta_j) = \sum_{i=1}^n P_i(\theta_j) \quad [\text{curva característica del test}]$$

El inverso de la función de información = la varianza del estimador en ese punto

$$V(\hat{\theta} | \theta) = \frac{1}{I(\theta)}$$

$$\text{El error estándar de medida de } \theta \text{ es por tanto: } S_e = \frac{1}{\sqrt{I(\theta)}}$$

La función de información de los ítems, de una subescala o un test, constituye una de las herramientas psicométricas más importantes en la actualidad para la construcción, aplicación y análisis de los tests.

12.7

EVALUACIÓN DE BONDAD DE AJUSTE DEL MODELO

Las ventajas y potencial de los modelos de la TRÍ para resolver problemas de medición son muy prometedoras. Sin embargo, no pueden lograrse sólo con estimar los parámetros de un modelo con un programa de computadora disponible para ello. Dichas ventajas son posibles solamente cuando el ajuste, entre el modelo y los datos obtenidos de las respuestas a los ítems del test, es satisfactorio.

Entre los métodos que existen para evaluar el ajuste del modelo a la distribución de resultados obtenidos, se han utilizado diferentes procedimientos estadísticos e indicadores. Por ejemplo, las técnicas basadas en el estadístico ji cuadrada consisten en comparar valores pronosticados por el modelo con los valores reales obtenidos empíricamente. Para ello, se divide la distribución de θ en varias categorías o fractiles, se les calcula la frecuencia esperada y se comparan los valores pronosticados con los datos en cada una.

El estadístico ji cuadrada de Yen, conocido como Q_1 , es un procedimiento que se utiliza para comprobar la bondad de ajuste del modelo elegido. Está dado por la siguiente expresión:

$$Q_1 = \sum_{j=1}^m \frac{N_j (O_{ij} - E_{ij})^2}{E_{ij} (1 - E_{ij})}$$

La distribución se divide en:

- m = el número de categorías o fractiles sobre la habilidad estimada de θ .
- O_{ij} = la proporción de respuestas correctas al ítem i en la categoría j .
- E_{ij} = la proporción pronosticada de respuestas correctas a ese ítem i para cada j .
- N = el número de sujetos examinados dentro de cada fractil.

El estadístico de Yen se distribuye aproximadamente como una ji cuadrada con $m - k$ grados de libertad, donde k = número de parámetros del modelo.

Interpretación del índice de bondad de ajuste

En el caso de un buen ajuste de los datos al modelo, los valores del estadístico de Yen coincidirían con ji cuadrada en todas las categorías. Lo que indicará el valor de ji cuadrada es si la diferencia es estadísticamente significativa o no.

Es importante recordar que en la evaluación de la bondad de ajuste se espera que el valor de ji-cuadrada con $m - k$ grados de libertad no sea significativo, con el fin de considerar un buen ajuste del modelo.

En general, se ha observado que los resultados de las pruebas estadísticas de ajuste del modelo son sensibles al tamaño de la muestra utilizada. Sin embargo, al menos en dos situaciones extremas identificables, los resultados pueden tener una interpretación clara en la prueba de ajuste:

1. Cuando el tamaño de la muestra es pequeño y se obtiene un mal ajuste; es decir, si el valor del estadístico de Yen resulta muy significativo.

En este caso, el investigador puede tener la confianza de concluir que el modelo elegido no ajusta satisfactoriamente a los datos obtenidos y, por tanto, no es el adecuado, o bien, es necesario revisar el funcionamiento de ítems.

2. Cuando el tamaño de la muestra es grande y efectivamente se obtiene un buen ajuste. Esto es, cuando el valor de ji cuadrada de Yen no resulta significativo.

En este caso, el investigador puede tener la confianza para asumir que, a pesar de una muestra grande, el modelo sí se ajusta a los datos de manera satisfactoria, por lo que es adecuado para estimar los parámetros de ese conjunto de ítems.

12.8

MODELOS POLITÓMICOS DE LA TEORÍA DE LA RESPUESTA AL ÍTEM

Como se ha explicado en las secciones anteriores, uno de los propósitos centrales de la teoría de la respuesta al ítem es la especificación de una función matemática que relaciona la probabilidad de respuesta de un individuo a un ítem con una capacidad o habilidad que se busca medir. Durante los primeros esfuerzos de este enfoque se trabajaron modelos aplicables a instrumentos unidimensionales con respuestas dicotómicas mediante el análisis de funciones de uno, dos o tres parámetros, en formulaciones matemáticas de ojiva normal o logísticas.

El danés George Rasch (1960) fue el primer investigador que se interesó en desarrollar un modelo logístico de sólo un parámetro. El planteamiento básico de Rasch difería de la aproximación clásica para abordar el problema de la medida de variables psicológicas, pero también del tipo de manejo matemático; había desarrollado un modelo psicométrico que daba cuenta de manera objetiva del desempeño de un individuo (su capacidad) en función sólo del nivel de cada ítem (su dificultad) como unidad básica de medida.

A partir de los antecedentes expuestos, los modelos de medida de la teoría de la respuesta al ítem desarrollados en las últimas décadas, además de superar las limitaciones de la teoría clásica de los tests, han aprovechado y potenciado los avances matemáticos y tecnológicos de la psicometría moderna. Hoy día se cuenta con mayores ventajas estadísticas e informáticas y se han podido derivar generalizaciones y aplicaciones específicas dentro de los modelos básicos. Asimismo, se han logrado avances en los procedimientos para la estimación de parámetros y el análisis de datos no sólo de carácter dicotómico, sino también para diferentes formas de respuesta, escalas multidimensionales e instrumentos de medida multicategoriales con modelos politómicos de procesos psicológicos y conductuales (Masters, 1982; Van der Linden & Hambleton, 1997).

En efecto, en la evaluación e investigación educativa y social con frecuencia se han aplicado instrumentos con formatos de respuesta muy diferentes a los clásicos tests con ítems dicotómicos de aptitudes y aprovechamiento escolar. Entre los tipos de respuestas que se evalúan, a menudo se usan instrumentos con respuestas categóricas, de crédito parcial y escalas de valoración con respuestas graduadas. Algunos instrumentos miden actitudes, frecuencia de conductas o la resolución de problemas, entre otros procesos o componentes cognoscitivos y conductuales, mediante escalas tipo Likert. Por ello, en los últimos años también la investigación psicológica y social ha enfocado la atención en la aplicación de otros modelos psicométricos, a fin de analizar diversos tipos de ítems y desarrollar escalas con datos politómicos.

Los nuevos modelos se han diseñado para datos obtenidos bajo otras condiciones de control o estandarización, a fin de analizar respuestas multicategoriales, ordenadas y no ordenadas, de procesos y en formatos más abiertos y flexibles. Estas variaciones han implicado que las puntuaciones tengan que ajustarse a diferentes variables y por tanto nuevos parámetros a considerar y métodos para estimarlos. Por ejemplo, si el puntaje depende de la decisión entre

alternativas de respuesta en cada ítem, es claro que esto dependerá del tipo de opciones que deberán parametrizarse para este tipo de ítems.

El principio metodológico y estadístico general de esta clase de modelos se refiere al aspecto básico de identificar parámetros separados de cada factor del ítem, ajustando y estimando (controlando) los efectos de los parámetros de respuestas. En esta nueva etapa de la psicometría se han logrado avances en el desarrollo y aplicación de modelos dentro de la teoría de la respuesta al ítem que constituyen procedimientos adecuados para el análisis de datos politómicos. Por ejemplo, en el uso de escalas tipo Likert o en instrumentos con formatos multicategoriales de respuesta o politómicos.

Uno de los primeros desarrollos de los modelos politómicos fue el que realizó Fumiko Samejima entre 1969 y 1973, quien, siguiendo un enfoque thurstoniano de evaluación de umbrales de discriminación en juicios de elección categorial, aplicó un modelo de ojiva normal para analizar datos de categorías ordenadas. En esa época no se mostraba aún mucho interés por desarrollar nuevos modelos, sino en aprovechar las aplicaciones prácticas y el desarrollo de los modelos dicotómicos existentes para pruebas de aptitudes y desempeño escolar. Más adelante se empezaron a derivar dentro de la TRÍ modelos para escalas de valoración o de clasificación por Andrich y Andersen; de categorías nominales de Bock; modelos de crédito parcial de Masters o el de Muraki; modelos logísticos de respuesta graduada de Samejima; de alternativas múltiples de Thissen & Steinberg; modelos multidimensionales de Wilson & Wang; o de componentes y procesos cognoscitivos de Embretson, entre otros modelos psicométricos que se están aplicando actualmente (Van der Linden & Hambleton, 1997).

Modelos politómicos para categorías ordenadas

En la psicometría existen fundamentalmente dos tipos de modelos para analizar datos de respuestas para categorías ordenadas (Masters, 1982). Las dos aproximaciones se diferencian en los supuestos que manejan y en el enfoque metodológico. Por un lado, los modelos tipo Thurstone parten del supuesto de que los sujetos responden al ítem de una forma racional; es decir, suponen que el ítem induce en el sujeto una reacción subjetiva, identificada como proceso discriminacional. Esto significa que la probabilidad de respuesta a un ítem depende de la relación entre el valor de esa reacción específica y los parámetros del ítem. Otra característica que asumen estos modelos es que si se unen categorías adyacentes se suman sus probabilidades, mientras que las categorías de las demás permanecen invariantes. Un modelo derivado de esta tradición metodológica es el modelo de respuesta graduada de Samejima, el cual se mostrará como un ejemplo de desarrollo de un modelo para datos politómicos.

Por otro lado, existen desarrollos de los modelos tipo Rasch que parten de una lógica distinta. En términos generales, dentro de éstos, primero se considera el tipo de datos que se obtienen al aplicar un test o una escala, luego se analiza la matriz de los puntajes de los sujetos y de los ítems y, a partir de esos puntajes, se hace una transformación directa a los parámetros que permiten estimar los datos, por último se deduce la función de probabilidad. Este tipo de modelos debe cumplir una condición para estimar los parámetros: contar con *estadísticos suficientes* que contengan toda la información muestral referida a cada parámetro de los sujetos y los ítems. Los estadísticos suficientes se refieren a las puntuaciones de los sujetos en el test y a las de los ítems de las respuestas de los sujetos de la muestra.

En los modelos tipo Rasch se asume que la puntuación del sujeto en el test contiene toda la información disponible para estimar el nivel del rasgo (θ), lo cual permite contar con un valor resumido global de los patrones de respuestas. Es decir, todos los sujetos con la misma puntuación, independientemente del patrón específico de respuestas, tendrían el mismo valor estimado de θ . Por lo contrario, en los modelos tipo Thurstone los sujetos con la misma puntuación global en el test pueden tener valores estimados de θ diferentes. Por ello, se dice que en los modelos thurstonianos la puntuación del sujeto no es un estadístico suficiente de θ , lo cual, lejos de ser una limitación, evita que se pierda parte de la información que contiene el vector de respuestas al resumirlo en una puntuación global.

Otra diferencia entre modelos thurstonianos y modelos de Rasch consiste en que estos últimos asumen que la aditividad entre categorías adyacentes sí afecta la probabilidad de las demás categorías. Esto supone que, cuando el sujeto elige una respuesta, no se evalúa cada categoría por separado, sino todas en su conjunto. En cuanto al parámetro de los ítems, se considera similar la puntuación del ítem con todos los sujetos de la muestra como un estadístico suficiente para estimar los parámetros, lo cual no sucede con dichos modelos de límites categoriales tipo Thurstone. Ejemplos de modelos tipo Rasch son el modelo de escala de clasificación de Andrich y el de crédito parcial de Masters.

Modelo de respuesta graduada

El modelo de respuesta graduada (MRG) representa un conjunto o familia de modelos matemáticos para categorías politómicas ordenadas (Samejima, 1997). Dicho modelo se aplica principalmente en el análisis de datos psicométricos de escalas con categorías, por ejemplo de letras, de números en un intervalo, o categorías de respuesta que asumen cierta graduación ordenada en los valores de las alternativas posibles de la variable que se mide. Otros ejemplos serían las escalas de cantidades o frecuencia estimada, valoraciones cualitativas, de procesos o aptitudes parciales y en las escalas tipo Likert. El MRG puede usarse con medidas cognitivas, habilidades y actitudes. Este tipo de modelos tienen la ventaja de proporcionar al investigador la posibilidad de conocer la precisión de la medida lograda por la escala para cada individuo y desarrollar escalas independientes de la muestra (De Ayala, 1993).

La función de respuesta para cada categoría de la escala vendría dada por:

$$P_{ui}(\theta) = \text{Prob}[U_i = u_i | \theta]$$

En donde U_i se refiere a la variable aleatoria y denota la respuesta graduada al ítem i , de tal forma que u_i representa la respuesta discreta específica a ese ítem en una categoría dada. Además, en un conjunto de ítems podrían identificarse los patrones de respuesta que indicarían las secuencias de U_i para cada ítem, desde $i = 1$ hasta $i = n$. Así, para ese conjunto de ítems tiene la notación del siguiente patrón de respuestas:

$$V = \{U_1, U_2, U_3, \dots, U_n\}$$

Si se asume el principio de independencia local para un grupo de sujetos con el mismo valor de habilidad θ , la distribución de las respuestas sería independiente de un patrón a otro.

Entonces, la probabilidad condicional $P_v(\theta)$ para el patrón de respuestas V sería:

$$P_v(\theta) = \text{Prob}[V = v | \theta] = \prod_{u_i \in V} P_{ui}(\theta) = L(V | \theta)$$

Samejima (1969) propuso los primeros modelos de respuesta graduada, utilizando tanto el modelo de ojiva normal como el modelo logístico para categorías ordenadas. Posteriormente, esta misma investigadora desarrolló un marco general para distinguir dentro de los modelos logísticos, el caso homogéneo del caso heterogéneo (Samejima, 1973). En el caso homogéneo la función P_u tiene la misma forma para todas las alternativas del ítem. En el caso heterogéneo las funciones P_u pueden variar en las alternativas y no sólo con cada ítem.

El MRG se caracteriza como un modelo de diferencias; es decir, al analizar las probabilidades para cada categoría de respuesta, la probabilidad de una puntuación dada se calcula como la diferencia entre dos funciones. Otra característica del modelo se refiere a que asume el principio de aditividad; esto es, que cuando dos opciones de respuesta se juntan en una, la probabilidad de esta última sería igual a la suma de las probabilidades de las alternativas ori-

ginales. En contraste con los modelos tipo Rash se considera que este modelo no cuenta con estadísticos globales suficientes para estimar θ de un individuo.

El MRG parte del principio de que al responder a un ítem es necesario completar una serie de operaciones o pasos previos en forma secuencial; esta secuencia viene dada de tal forma que para completar el paso $(u + 1)$ es necesario haber completado los pasos $(u - 1)$ anteriores exitosamente. Por ejemplo, considere P^*u como la probabilidad de completar o cubrir satisfactoriamente los u o más pasos anteriores; entonces tendría que $P^*0 = 1$ debido a que todos los sujetos que eligen una categoría del ítem pueden completar al menos 0 pasos. En el otro extremo tendríamos que $P^*(m + 1) = 0$ en donde si m es el número máximo de categorías o pasos posibles del ítem, ningún sujeto podría completar más pasos u obtener un puntaje mayor del máximo posible de la escala.

La función $P^*u(\theta)$ se usa para estimar las probabilidades reales de la puntuación lograda en cada ítem. Esta función representa la probabilidad condicional con la cual los sujetos con cierta capacidad θ completan los procesos cognoscitivos, o pasos, con éxito, hasta el paso u_i . Para cada paso o respuesta se asume que la $P_{ui}(\theta)$ es no decreciente, de tal manera que cada ítem posee una relación directa y positiva con la capacidad o variable psicológica medida.

La función de respuesta para cada categoría $P_{ui}(\theta)$ en una escala graduada en un ítem i estaría dada por:

$$P_{ui}(\theta) = \left[\prod_{s \leq u_i} M_s(\theta) \right] 1 - Mu_i + 1(\theta)$$

$$\text{Entonces: } P^*u_i(\theta) = \prod_{s \leq u_i} M_s(\theta)$$

$$\text{También puede expresarse como: } Pu_i(\theta) = P^*_{ui}(\theta) - P^*_{u+1}(\theta)$$

Ésa sería la forma más general del modelo MRG que resumiría una de sus principales características: el análisis de probabilidades de cada categoría estará dado por la diferencia entre las probabilidades de las curvas límite o de frontera entre cada paso o categoría de respuesta. Dicha función puede calcularse con el modelo de ojiva normal o con un modelo logístico.

Es importante notar que $P^*u_i(\theta)$ viene a ser la función de respuesta positiva al ítem i , y donde el puntaje graduado del ítem se transforma en un puntaje binario provisional, asignando 0 a todas las demás categorías menores que u_i , y se asigna 1 a las categorías mayores o iguales a u_i . Es claro que si $P^*u_i(\theta)$ también es no decreciente en θ y asume la unidad para $u_i = 0$ y el cero para $u_i = m + 1$ en el rango completo de θ ; es decir, en la escala de los diferentes pasos o categorías del intervalo de respuestas.

La función logística del modelo politómico de respuesta graduada sería:

$$P^*u = \frac{\exp[D \alpha_i (\theta - \beta u_i)]}{1 + \exp[D \alpha_i (\theta - \beta u_i)]}$$

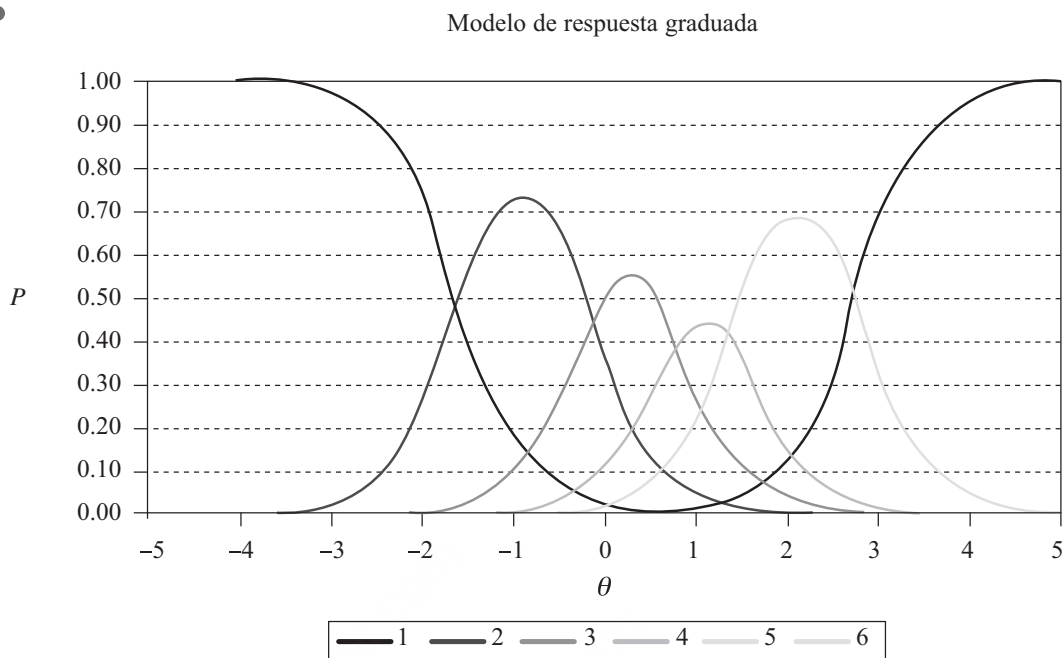
En donde, como se ha señalado anteriormente, la constante $D = 1.7$.

Así, las probabilidades serían las mismas con el modelo de ojiva normal.

A continuación se muestra un ejemplo gráfico del modelo politómico MRG para un ítem con seis categorías de respuesta graduada y sus probabilidades para cada categoría estimadas con la función logística.

FIGURA 12.9

Representación de
funciones de respuesta de
un modelo politómico



Ventajas de los modelos politómicos

En suma, las principales ventajas de los modelos politómicos de la TRÍ para estimar las propiedades psicométricas de una escala frente al análisis clásico de datos con puntuaciones globales del test, serían:

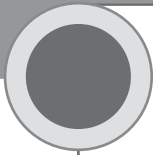
1. El análisis de un modelo politómico de la TRÍ implica un estudio previo de la dimensionalidad del test o de la escala que se está construyendo. Al aplicar el modelo no es suficiente con estimar sus parámetros; es necesario evaluar la bondad de ajuste del modelo a los datos, a fin de considerar que la escala es unidimensional y adecuada para evaluar a los sujetos.
2. Este tipo de modelos permite ubicar a los sujetos en una misma escala. Como se sabe el parámetro b del ítem puede interpretarse como el valor afectivo de preferencia del sujeto y permite compararlo con el parámetro de rasgo del sujeto. Mediante un análisis de frecuencias no se podría apreciar esto debido a que no proporciona información de la intensidad de preferencia de la categoría del ítem.
3. La construcción de escalas con estos modelos permite estimar predicciones sobre la probabilidad de respuesta de un sujeto ante la presentación de nuevos ítems, del mismo tipo. Esta propiedad permite construir instrumentos de medida seleccionando los ítems que sean máximamente informativos.
4. Los modelos de respuesta graduada para categorías ordenadas constituyen una alternativa metodológica útil para el desarrollo de escalas que miden rasgos y procesos psicológicos diferentes de las medidas tradicionales de conocimientos.
5. No obstante la limitada generalidad que pueden alcanzar algunos modelos como el MRG, representan aspectos positivos al restringir la interpretación del evaluador a una población, a un conjunto de ítems y a un contexto particular. De esta forma se fortalecería la validez metodológica y el investigador estaría obligado a buscar mayor evidencia empírica que apoye la validez de constructo de la escala.

RESUMEN

Un aspecto central en la teoría de la respuesta al ítem es la especificación de una función matemática que relaciona la probabilidad de respuesta de un individuo a un ítem con una habilidad que se busca medir. Inicialmente en este enfoque se desarrollaron modelos aplicables a instrumentos unidimensionales, con respuestas dicotómicas, mediante el análisis de funciones de uno, dos o tres parámetros, en formulaciones matemáticas de ojiva normal y logísticas.

En las últimas décadas, además de superar las limitaciones de la teoría clásica de los tests, se han aprovechado y potenciado los avances matemáticos y tecnológicos de la psicometría moderna. En consecuencia se cuenta con mejores métodos estadísticos y mayores ventajas informáticas. Además se han podido derivar generalizaciones y aplicaciones específicas dentro de los modelos básicos. Asimismo, se han logrado avances en los procedimientos para la estimación de parámetros y análisis de datos, pero también para diferentes formas de respuesta, e instrumentos de medida multicategoriales con modelos politómicos de procesos psicológicos y conductuales. El principio metodológico y estadístico general de los modelos psicométricos se refiere al aspecto básico de identificar parámetros separados de cada factor del Ítem, ajustando y estimando los efectos de los parámetros de respuestas.

En la investigación educativa y social con frecuencia se aplican instrumentos con formatos de respuesta diferentes a los clásicos tests con ítems de aptitudes o de aprovechamiento escolar. Los tipos de respuestas que se evalúan requieren instrumentos de respuestas categóricas, de crédito parcial y escalas de valoración con respuestas graduadas. Otros instrumentos miden actitudes, frecuencia de conductas o resolución de problemas, entre otros procesos o componentes cognoscitivos y conductuales. La investigación psicológica y social ha enfocado la atención en la aplicación de nuevos modelos estadísticos y psicométricos, a fin de analizar y desarrollar escalas con datos politómicos y multidimensionales.



Apéndice A

Los escritos científicos



INTRODUCCIÓN

Una vez hecho el análisis de la literatura (libros, capítulos de libros y artículos) sobre un tema determinado, habiéndose planteado un problema interesante y trascendente de investigación, seleccionado una metodología pertinente para resolverlo, habiendo colectado los datos y analizado estadísticamente los resultados, lo que queda es publicarlos. Nada de lo anterior tiene sentido si el resultado de los esfuerzos realizados para avanzar en la comprensión de un aspecto de la realidad, usando el método científico, no se hace público, es decir, se pone a la disposición de la comunidad científica, realizando una aportación, no importa cuán pequeña o grande sea ésta.

La ciencia es una empresa social, sobre la cual miles, tal vez en la actualidad millones de individuos, participan de un arduo proceso de generación de ideas, análisis, discusión de las mismas y de contrastación de éstas contra la evidencia.

Los congresos y, fundamentalmente las publicaciones, son los foros donde se lleva a cabo el juicio que en forma paulatina moldea el resultado de la ciencia. Para todo fin práctico, una investigación es un informe de los resultados de alguna investigación específica, mientras que otros, en su mayoría en la forma de libros o capítulos especializados en antologías y libros coordinados, son revisiones críticas del estado del arte del problema en cuestión. Otros más son discusiones teóricas donde se analiza, a la luz de la evidencia acumulada hasta el momento, la mejor explicación posible, tal vez estableciendo el debate entre diferentes aproximaciones o teorías.

Mientras un trabajo no llega a la prensa, éste no se encuentra sujeto al proceso natural y muy saludable de la crítica. Popper habla de tres mundos, el mundo objetivo de la realidad, el mundo de las percepciones e ideas acerca de este primer mundo, donde se dan conjeturas acerca de sus entes y naturaleza y el tercer mundo, el mundo de las ideas publicadas y, por tanto, susceptible de ser analizado y criticado por los demás y hasta por uno mismo. La ciencia se da sólo en ese tercer mundo.

No obstante lo importante que resulta publicar para la ciencia y el conocimiento, uno de los puntos que cuesta más trabajo de los artículos de investigación, divulgación, análisis de textos, antología o trabajos de tesis, es escribirlos: enfrentar la hoja en blanco requiere de tiempo para tomar el valor de hacerlo. Más difícil aún resulta preparar el escrito para su publicación. Es muy posible que usted, estimado lector, tenga algún trabajo en el cajón de su escritorio, el cual nunca se atrevió a enviar a alguna editorial, por temor al rechazo.

Tomando en cuenta que, dado lo que se señaló en los primeros párrafos, vale la pena tomarse el trabajo de preparar un escrito para enviarlo a una editorial, conviene eliminar prejuicios que muchas veces sólo están en el pensamiento de cada uno. Las razones, entre otras muy válidas pueden ser:

- Concluir una investigación, haciéndola llegar a los colegas.
- Ver de manera objetiva el esfuerzo realizado.
- Adquirir prestigio.
- Obtener alto valor curricular.

Por las razones señaladas, en este libro se incluye este apéndice que pretende orientar acerca de cómo preparar un artículo para su publicación, tomando como base las recomendaciones de la American Psychological Association (APA) para el caso de los escritos del área social, y las del Comité Internacional de Directores de Revistas Médicas (CIDRM) para el caso de los escritos biomédicos.

Qué escribir

Debido a que cada autor puede tener un motivo personal para escribir, conviene reflexionar acerca de que los buenos escritos resultan benéficos para el propio autor, los lectores, las instituciones y la ciencia. A fin de que tenga mayor probabilidad de ser publicado, un manuscrito debe cumplir varios requisitos, a saber:

- Buena calidad del contenido: intentar disfrazar una investigación planteada o manejada de manera deficiente, resulta en un trabajo que está destinado al rechazo. Para los editores, un artículo que refleje deficiencia en el método, resulta inaceptable.
- Originalidad: por desgracia, muchos autores, sólo por publicar, informan de cambios insignificantes en investigaciones previas de sí mismos o de otros. Esto tiene un costo para el conocimiento porque hace que avance con mucha lentitud o que no avance y además consume recursos de las instituciones y de los grupos de trabajo.
- Congruencia: entre las operaciones específicas del estudio, el diseño, el análisis, con la interpretación y discusión de resultados.

Tipo de escrito

El tipo de escrito que puede hacerse es muy diverso, a continuación se esbozan algunos.

Artículos de investigación o informes de estudios empíricos, que son informes de investigaciones originales. Suelen incluir distintas secciones que son:

- **Introducción**, que resulta de la revisión de la literatura y en la cual se incluye el desarrollo del problema y se plasma el objetivo de la investigación.
- **Método**, apartado en el cual se describe con detalle cómo se llevó a cabo la investigación, de manera que otro investigador si lo desea, pueda repetirla.
- **Resultados**, importante sección en la cual se describen los resultados encontrados, tanto positivos como negativos o que contradigan a las hipótesis. Algunos autores temen que suene a fracaso el poner estos últimos; por lo contrario, permiten conocer caminos que no conducen a nada, evitan repeticiones innecesarias. Cabe recordar que no conviene hacer contrastaciones en este apartado, las cuales deben incluirse en el capítulo de discusión.
- **Discusión**, es la contrastación de lo encontrado en relación con lo que han dicho otros autores, así como la interpretación y análisis de las implicaciones de los resultados.

Artículos de reseña o recensión. Son evaluaciones críticas de material ya publicado. Este trabajo, realizado con cuidado por un experto, es de gran ayuda para el conocimiento y para otros investigadores o estudiantes, debido a que define y clarifica problemas para posible investigación, permite conocer el “estado del arte” de un campo del conocimiento, identifica relaciones, contradicciones, lagunas e inconsistencias y propone pasos adelante para la posible solución de problemas.

La estructura de este tipo de escrito suele ser libre, si bien debe incluir:

- **Introducción.** Que incluya el qué, por qué, cómo, cuándo, dónde, para qué y para quién.
- **Cuerpo del escrito.** En el cual se ordena por relación la temática tratada.
- **Conclusión.** Que integre y haga explícitos los puntos señalados en el párrafo anterior.

Artículos teóricos. En este tipo de documentos, el o los autores recopilan y analizan las teorías existentes acerca de un tópico y finalmente contrastan y dan nuevas teorías o nuevos

enfoques que permitirán disponer de distintos puntos de partida para la investigación. La estructura de este tipo de artículos y los de reseña es muy semejante, si bien en estos últimos se hace mayor énfasis en la conclusión.

Artículos metodológicos. Incluyen aproximaciones metodológicas nuevas, modificaciones de métodos existentes así como discusiones acerca de enfoques cuantitativos y de análisis de datos. Este tipo de artículo resulta muy útil al hacer el diseño de una investigación, sobre todo si tiene detalles suficientes como para que pueda valorarse la aplicabilidad de la metodología al diseño de un nuevo estudio.

Estudio de caso. Resultan de estudios acerca de un individuo o una organización, con el objetivo de ilustrar un problema, indicar algún modo de resolverlo o esclarecer una investigación o los elementos teóricos necesarios. Puede ser obligado utilizar material confidencial que quizá requiera autorización escrita del sujeto. Otra forma es alterar características específicas que tengan poca relación con las variables en estudio, limitar la descripción de rasgos particulares o encubrir los detalles del caso agregando material complementario.

Otros tipos de artículos. Menos frecuentes, pueden ser informes breves, comentarios y contestaciones a artículos ya publicados, discusiones acerca de métodos cuantitativos, historias de caso y monografías.

Cualquiera que sea el tipo de artículo que usted como autor desee publicar, se sugiere remitirse a la revista científica en la cual desee presentar su manuscrito, a fin de cubrir los requisitos que el comité editorial decide en cada revista.

Partes del manuscrito

Debido a que casi todos los artículos que se publican en ciencias sociales o biomédicas son informes de estudios empíricos, se presenta un esquema general de las partes que deben integrar un manuscrito. El autor debe seguir un esquema, dando por supuesto el enfoque de originalidad que lo hará único.

Hoja de portada o portadilla. Debe incluir el título, parte importante del escrito dado que además de informar a los lectores acerca del estudio, también se utiliza como presentación del contenido en bases de datos. Debe sintetizar la idea principal, ser conciso, por sí solo debe ser explicativo del artículo. Se recomienda una extensión máxima de 15 palabras (la APA recomienda entre 10 y 12; el Comité Internacional de Directores de Revistas Médicas [CIDRM] no pone extensión, sólo señala que debe ser “conciso pero informativo”) y se sugiere elaborarlo justo antes de enviarlo para publicación. Esto podría representar un problema debido a que en trabajos de tesis se exige un título al registrar el protocolo, sin embargo no existe impedimento para cambiar el título al publicar el trabajo.

Créditos. Deben incluir el o los nombres de los autores (sin grados académicos), completos, empezando por nombre(s) y apellidos, procurando siempre escribirlo de la misma manera, visualice su futuro como autor con muchas publicaciones y evite complicaciones a bibliotecarios y a otros investigadores. De igual manera anote la afiliación institucional, que es el lugar donde se realizó la investigación. Cuando no exista afiliación institucional, anotar la ciudad y la entidad de residencia bajo el nombre.

Cornisa. Es un renglón o título abreviado que se imprime en la parte superior de las páginas para identificar el artículo ante los lectores. Debe tener un máximo de 50 caracteres incluyendo letras, puntuación y espacios entre palabras.

Resumen. Es la carta de presentación del artículo, con frecuencia los lectores decidirán, con base en el resumen, si leerán el artículo completo y si les es útil como antecedente de su propia investigación. Por ello, el resumen debe ser: compacto en su información, legible, bien organizado, de corta extensión y completo. Es variable el número de palabras que cada revista acepta en el resumen, pero en términos generales es de 120 como máximo; revise lo que señala la revista en la cual desea publicar y apéguese a ello.

Una sugerencia para realizar su resumen: una vez concluido el artículo escriba las ideas principales, seleccione el párrafo, haga clic en el icono Herramientas y al desplegarse el menú haga clic en Contar palabras. Si el número excede al requerido, elimine algunos artículos, emplee dígitos para las cifras y evite frases trilladas o expresiones que no proporcionen información por sí mismas, como “Se concluye que...” o “Los resultados mostraron que...”.

Un buen resumen es **preciso**, no incluya en él información que no aparezca en el cuerpo del escrito. Es **completo**, no añada abreviaturas (excepto unidades de medida), escriba los nombres completos de pruebas y fármacos, defina términos poco comunes. Es **conciso** y **específico**, aproveche al máximo el lenguaje y haga que cada palabra y cada oración tengan en realidad contenido. **No evaluativo** sino informativo, no añada comentarios. Finalmente, debe ser **coherente** y **legible**, utilice verbos más que sustantivos equivalentes y emplee la voz activa en lugar de la pasiva (evite los gerundios).

El resumen de un artículo de investigación o informe de estudio empírico debe describir:

- Problema.
- Individuos participantes o sujetos especificando las características pertinentes como número, tipo, edad, sexo.
- El método experimental utilizado incluyendo mecanismos, forma de recopilación de datos, nombres de pruebas y nombres genéricos completos, así como dosis y vías de administración de los fármacos utilizados.
- Hallazgos incluyendo niveles de significado estadístico.
- Conclusiones e implicaciones o aplicaciones.

El resumen para un artículo de reseña o teórico debe incluir:

- Tema en un solo enunciado.
- Objetivo.
- Fuentes de información.
- Conclusiones.

El resumen de un artículo metodológico debe contener:

- Tipo general del método que se propone o discute.
- Características esenciales del método planteado.
- Rango de aplicación del método propuesto.
- Comportamiento del método, incluso su poder y la solidez de su estructura ante la violación de los supuestos teóricos.

El resumen de un estudio de caso debe describir:

- Al sujeto (individuo u organización) y sus características relevantes.
- La naturaleza del problema y su solución.
- Las preguntas que surgieron en relación con la investigación o fundamentación teórica.

Introducción. Por estar al principio del escrito, este apartado no se rotula, queda implícito. Debe describir el problema específico que se estudia y la estrategia de investigación. Resulta conveniente recordar que toda introducción debe incluir siempre en los primeros párrafos el **qué** (problema de estudio), **por qué** (justificación), **cómo** (estrategia de investigación o método), **dónde** (institución o lugar donde se lleva a cabo la investigación), **para qué** (trascendencia). Es posible que sea necesario agregar el **para quién** si la investigación es por encargo de una institución, un grupo político, etcétera.

Según la APA, en la introducción deben anotarse los antecedentes. Procure no hacer una revisión histórica exhaustiva, cite y refiera sólo los trabajos pertinentes al tema de su investigación; evite los detalles no esenciales, en cambio haga énfasis en los hallazgos pertinentes, los aspectos metodológicos relevantes y los principales resultados y conclusiones. La extensión de la introducción no debe exceder de dos cuartillas y entre más corto mejor. Cabe recordar que una cuartilla es una hoja de papel bond tamaño carta, escrita por una sola cara, con márgenes de 2.5 o 3 cm por lado y con 28 renglones a doble espacio.

Método. En este apartado se describe con detalle la forma en que se llevó a cabo el estudio. Esta descripción permite a quien lo lee evaluar si el método fue adecuado, así como la confiabilidad y validez de los resultados obtenidos. Además, como ya se mencionó, hace posible que otros investigadores repliquen el estudio si lo desean.

En este apartado es importante que se rotulen las subsecciones a fin de que queden muy claras. Algunas pueden ser:

- **Participantes o sujetos.** Si los sujetos son seres humanos, detalle la forma de selección y asignación, así como los acuerdos y pagos realizados, mencione si se obtuvieron consentimiento informado y acuerdos de confidencialidad. Mencione las principales características de las personas como sexo, edad, origen étnico, lugar de residencia y, si es pertinente, ocupación, nivel socioeconómico, grado de discapacidad, estado de salud y orientación sexual. Si alguna de estas características es variable en estudio, debe detallarla con precisión. En caso de que los sujetos no concluyan el estudio, mencione cuántos y las razones del abandono. En el caso de sujetos animales, mencionar el género, la especie y el número de cría y de dónde proviene, así como datos del proveedor. Debe indicar el número total de sujetos y cuántos de ellos fueron asignados a cada condición experimental.

Por último, indique al editor de la revista a la cual envía su artículo, que el tratamiento de los participantes, sean seres humanos o animales, estuvo de acuerdo con las normas éticas de la APA (“Ethical Principles of Psychologists and Code of Conduct”, APA 1992a) o en su caso del Comité Internacional de Directores de Revistas Médicas (CIDRM).

- **Herramientas.** En esta subsección se describen en forma breve los equipos o materiales empleados y la función en el experimento. Los equipos de línea deben mencionarse con número de serie o modelo y datos del proveedor. Si fueron hechos *ex profeso* para el experimento, cabe hacer una ilustración detallada o una fotografía. El equipo estándar de laboratorio sólo debe mencionarse sin detalles.
- **Procedimiento.** Aquí se resume paso a paso la ejecución de la investigación, el método debe informar al lector con detalle, *qué* hizo el investigador y *cómo* lo llevó a cabo: la aleatorización, contrabalanceo, recopilación de información en otro idioma. Si se emplean procedimientos nuevos o especiales, deben describirse puntualmente. En los artículos metodológicos, puede colocar en un apéndice las explicaciones más específicas como derivaciones y detalles de simulación de datos.

Resultados. Resuma los datos recolectados, el tratamiento estadístico o cualitativo y el producto obtenido. En esta sección es frecuente incluir contrastaciones de datos obtenidos por otros autores o analizar implicaciones de los mismos, en cuyo caso el apartado se denominará **Resultados y discusión**, o **Resultados y conclusiones**. Tal distribución dependerá mucho de la revista en la cual desea publicar. En opinión de quienes esto escriben, es mejor poner en apartados separados cada rubro, para mayor claridad de las ideas.

Como ya se mencionó, no omita los resultados que contradicen las hipótesis, quizá pensando en que se interprete como fracaso. Piense en el beneficio que aporta a otros investigadores el que usted les evite transitar por caminos equivocados.

El empleo de cuadros (tablas) y figuras es imprescindible en este apartado. Los aspectos a considerar para elegir uno u otro medio son: claridad, brevedad y economía. Evite repetir datos o emplear cuadros para información que puede decirse en dos o tres frases dentro del texto, como sería la distribución por sexo de los participantes.

Respecto de la presentación estadística, el enfoque puede ser muy diverso y las únicas sugerencias son: cuando presente datos provenientes de métodos estadísticos inferenciales (pruebas t , pruebas F , ji cuadrada) incluya información acerca de la magnitud o valor obtenidos de la prueba, los grados de libertad, la probabilidad de conseguir un valor tan extremo o más que el obtenido y la dirección del efecto.

Incluya estadística descriptiva (tamaño de muestra por celda, medias, correlaciones, desviaciones estándar) a fin de esclarecer la naturaleza del efecto que se está informando y para metaanálisis futuros. Por la misma razón es muy recomendable emplear intervalos de confianza, siempre el mismo en todo el artículo (por ejemplo 95 o 99 por ciento).

En el caso de sistemas de análisis de variables múltiples como los multivariados, análisis de regresión y de ecuaciones estructurales, son adecuadas la(s) media(s), el tamaño de la(s) muestra(s) y la matriz o matrices de varianza-covarianza o correlación. El principio consiste en proporcionar al lector no únicamente información sobre la significancia estadística, sino datos suficientes para evaluar la magnitud del efecto o de la relación observados.

Discusión. Inicie esta sección con una exposición amplia de la sustentación o falta de ella, para las hipótesis originales. Aquí se aclaran y confirman las semejanzas y diferencias entre los resultados propios y los obtenidos por otros autores. Reconozca las limitaciones y señale las explicaciones alternativas de los resultados.

Conclusiones. Si decide colocar este apartado o lo solicita la revista a la cual se enviará el artículo, incluya aquí aspectos como importancia del problema, qué cuestiones dependen de los hallazgos, las proposiciones que se confirman o no mediante extrapolación a temas de mayor importancia. De igual manera, cómo pueden vincularse los hallazgos con fenómenos a niveles más o menos complejos de análisis y qué se necesita saber para establecer los vínculos. Si los hallazgos son válidos y replicables, mencione qué fenómenos psicológicos de la vida real es posible explicar o modificar con base en los resultados y si existen aplicaciones fundamentadas en esta investigación. Cuando sea pertinente, concluya esta sección con un comentario acerca de la importancia de sus descubrimientos.

Fuentes de información

Son los soportes materiales de donde es posible obtener información del conocimiento disponible. Se clasifican en fuentes primarias y secundarias.

Las primeras son artículos escritos como resultado de investigación, que pueden encontrarse en: revistas, bancos de información (*PsysINFO*, *Medline*, entre otros), por internet y en libros que se publican con intervalos de no más de seis meses, por ejemplo las Clínicas Médicas de Norteamérica. En este tipo de publicaciones puede encontrar información de avanzada y actualizaciones que no rebasan el año desde que se generó el conocimiento.

Las fuentes de información secundarias se refieren a artículos de revisión, actualización y a libros que tienen por lo menos un año desde que se generó el conocimiento. Son importantes dado que es la forma en que la mayoría de las personas recibe la información; si bien como fuente para escribir un artículo, algunos editores (o lectores) podrían cuestionarlo. En todo caso, como autor debe buscar ambas fuentes, a fin de tener la solidez del material que ya pasó la prueba de la crítica y la contrastación; más aquella que tiene el avance y la frontera del conocimiento.

Estilo

La escritura científica debe basarse en el método científico. Por tanto, el estilo de redacción que debe utilizarse de manera genérica es el estilo científico que es: racional, claro, objetivo, coherente y preciso. Para ello, las unidades de pensamiento, que pueden ser una palabra, una frase, oración (enunciado) o un párrafo deben organizarse, ser continuas y coherentes.

Se sugiere organizar las ideas por párrafos colocando una idea central al inicio (por ser la más importante) y cuatro o cinco, como máximo, ideas secundarias. Para redactar el siguiente párrafo, si la idea inicial no tiene mucha relación con la última del párrafo anterior, debe utilizarse una palabra o frase conectora, como la frase **Si bien** en el ejemplo siguiente:

Zea mays es una planta monoica; sus inflorescencias masculinas y femeninas se encuentran en la misma planta. Las hojas toman una forma alargada íntimamente arrollada al tallo, del cual nacen las espigas o *mazorcas*. Cada mazorca está cubierta por filas de granos, cuyo número puede variar entre ocho y 30.

Si bien la planta es anual, su rápido crecimiento le permite alcanzar hasta los 2.5 m de altura, con un tallo erguido, rígido, sólido. El tallo está compuesto a su vez por tres capas: una epidermis exterior, impermeable y transparente, una *pared* por donde circulan las sustancias alimenticias y una *médula* de tejido esponjoso y blanco donde almacena reservas alimenticias, en especial azúcares.

No obstante, cuando los editores se refieren a estilo, suelen hablar del estilo editorial, que son las reglas o principios que aseguran la presentación clara y consistente de la palabra impresa. El estilo editorial es el empleo uniforme de signos de puntuación, abreviaturas, construcción de cuadros y figuras, selección de encabezados, citas y referencias, y en general todos los elementos que hacen legible y atractivo a un escrito publicado.

A continuación se desglosan los usos genéricos de signos de puntuación que recomienda el Manual de estilo de APA.

- **Punto:** se emplea al final de un enunciado (frase u oración) completo, al final de un párrafo y al final del escrito.
- **Coma:** se emplea entre enunciados, para resaltar una cláusula explicativa, para separar dos enunciados independientes unidos por una conjunción y para resaltar el año en fechas exactas, aun dentro de paréntesis (sobre todo en inglés).
- **Punto y coma:** separan dos enunciados independientes no relacionados mediante conjunción; también separa enunciados en un párrafo que ya tiene más de tres comas.
- **Dos puntos:** se colocan antes de mencionar una serie de sustantivos, adjetivos u orden de ideas, entre una idea introductoria gramaticalmente completa y la frase final que ilustra, extiende o amplifica la idea precedente. También deben usarse los dos puntos en razones y proporciones (la proporción sodio:cloro fue 1:1), y en las referencias entre el lugar de publicación y la casa editorial. En castellano, la palabra que sigue a los dos puntos, es minúscula a menos que se trate de nombre propio.
- **Giones:** si bien existe la tendencia a colocar guiones entre adjetivos relacionados como socio-económico y físico-químico, esto no es correcto; debe escribirse socioeconómico y fisicoquímico. En general emplee los guiones para yuxtaponer palabras sólo cuando sea absolutamente necesario, como en la frase: *vuelo México-Madrid*.
- **Uso de abreviaturas:** hay un abuso generalizado, producto de malas traducciones del inglés en el uso de abreviaturas. El lenguaje se vuelve críptico si aparecen siglas que el lector no sabe qué significan y, por tanto, se trunca la comunicación. En general, utilice una abreviatura sólo si a) es ampliamente conocida y el lector está más familiarizado con la abreviatura que con el nombre completo, como DNA o ADN (ácido desoxirribonucleico) o en psicología MMPI (Inventario Multifásico de la Personalidad Minnesota); o b) si se ahorra un espacio considerable y se evita una repetición engorrosa. Use sólo abreviaturas que le ayuden a comunicarse con sus lectores. Si finalmente se ve obligado a utilizar alguna abreviatura no común, mencione el término y enseguida las siglas entre paréntesis y a continuación sólo escriba la forma abreviada.
- **Unidades de medida:** utilice la nomenclatura del Sistema Internacional de Medidas, cuyas unidades básicas pueden verse en el cuadro siguiente. Observe que el símbolo no requiere punto; así la forma correcta de poner 18 m es así, sin punto, a menos que esté al final del párrafo.

Magnitud	Nombre	Símbolo
Longitud	metro	m
Masa	kilogramo	kg
Tiempo	segundo	s
Intensidad de corriente eléctrica	ampere	A
Temperatura termodinámica	kelvin	K
Cantidad de sustancia	mol	mol
Intensidad luminosa	candela	cd

Escriba los nombres completos de las abreviaturas para unidades métricas y no métricas que no estén acompañadas por valores numéricos, ejemplo: “Las unidades de longitud utilizadas fueron centímetros”.

Para el resto de signos de puntuación (raya, comillas, paréntesis, corchetes, diagonal) que puede resultar necesario emplear en un texto científico, se sugiere consultar un diccionario como el de la Real Academia de la Lengua Española o alguna de las publicaciones que la APA o el CIDRM tienen al respecto.

Lo mismo se propone para la **ortografía**, si bien la rápida proliferación del vocabulario con la internet y la World Wide Web, hace que los diccionarios no resulten tan efectivos. De cualquier manera, resulta imprescindible disponer de diccionarios y acceso a libros que despejen dudas, en tanto se adquiere la suficiente experiencia para un uso fluido del idioma.

Jerarquización. Es la organización de un manuscrito con encabezados y tiene como fundamento orientar al lector respecto de la importancia de los temas y su relación entre sí. Ejemplo:

EL AMBIENTE SOCIAL
Espacio y territorio humanos
Dominio territorial
Organización de la vida cotidiana
Identidad personal y de grupo.
El espacio personal: dimensiones psicológica, social y cultural
La dimensión psicológica
Teorías prevaletientes.

La APA recomienda este estilo de encabezado que consiste hasta en cinco posibles arreglos de formato, que pueden ser:

<i>Encabezado en mayúsculas centrado.</i>	nivel 5
<i>Encabezado en mayúsculas y minúsculas centrado.</i>	nivel 1
<i>Encabezado en mayúsculas y minúsculas centrado y en cursiva.</i>	nivel 2
<i>Encabezado secundario en mayúsculas y minúsculas, en cursiva y alineado a la izquierda.</i>	nivel 3
<i>Encabezado de párrafo con sangría, en minúsculas, en cursivas, alineado a la izquierda y que finaliza con punto.</i>	nivel 4

Puede ser que su escrito no requiera los cinco niveles sino sólo dos o tres. Usted es quien más sabe acerca de lo que escribe y conoce las relaciones que deben guardar las distintas secciones. Procure consultar las normas editoriales de la revista donde desea publicar o comuníquese con el editor de la misma a fin de plantear sus dudas. La recomendación es: no ponga números ni letras a los encabezados.

Números. Utilice guarismos en el número 10 y mayores, así como palabras para los números del cero al nueve. De igual manera emplee palabras cuando un número inicie el enunciado. En caso de decimales, coloque un cero antes del punto cuando los números sean menores al uno. No requiere el cero cuando el número no pueda ser mayor que uno (como en las correlaciones, proporciones y niveles de significación estadística).

Cuadros denominados también y aceptados por la APA como *tablas*, si bien este término es poco afortunado porque deriva de una mala traducción del inglés *table*, es un excelente recurso para presentar gran cantidad de datos en un espacio reducido. Las tablas de datos cualitativos pueden ser de análisis de varianza, regresión, de relaciones estructurales lineales, entre otros. Las de enunciados son muy útiles para agrupar datos cualitativos.

Errores frecuentes: colocar exceso de tablas y repetir datos en el texto.

Sugerencias: numérelas todas en orden consecutivo, con números arábigos y ubíquelas en el texto, ponga a cada una la cabeza (título) breve, claro y explicativo y quizá una anotación (o varias) al pie, que puede ser: nota general (se refiere a toda la tabla y se inicia con la palabra *Nota*), nota específica (se refieren a una columna o fila en particular y se señalizan mediante índices exponenciales en letras minúsculas, como ^{a, b, c...}) y notas de probabilidad (indican los resultados de pruebas de significación y se señalizan con asteriscos*). Cualquiera de ellas se incluirá sólo si es pertinente.

Figuras. Cualquier tipo de ilustración que no sea cuadro (tabla) se denomina figura, y puede ser diagrama, gráfica, fotografía, dibujo, esquemas, radiografía o mapa, entre otras. Este recurso visual es poderoso para comunicar de manera eficiente la información; sin embargo, pregúntese si realmente es necesaria una figura y qué tipo es mejor. Los estándares que prácticamente todas las revistas piden para considerar como buenas a las figuras, son sencillez, claridad y continuidad. Una buena figura enriquece al texto, no lo duplica; comunica sólo hechos esenciales; omite referencias entre el lugar de publicación y la casa editorial.

Problemas éticos de un reporte de investigación

Estos problemas pueden referirse a la información contenida en el artículo que se envía para edición o al derecho a la privacidad de los pacientes.

En el primer caso, la publicación redundante o duplicada consiste en la publicación de un artículo que coincide sustancialmente con otro ya publicado.

Los lectores de las revistas biomédicas deben tener la garantía de que aquello que están leyendo es original, a menos que se informe inequívocamente que el artículo es reedición, decidida por el autor o director de la revista. Esta decisión debe hallarse en consonancia con las leyes internacionales sobre los derechos de autor, con la conducta ética y con el uso eficiente de los recursos.

Cuando se envíe un original, el autor deberá informar al director de la revista acerca de cualquier presentación del documento a otras revistas, o cualquier trabajo anterior que pudiera considerarse publicación previa o duplicada de un trabajo idéntico o muy similar. El autor, también, debe advertir al director si el trabajo incluye cuestiones abordadas en trabajos ya publicados. Estos trabajos previos deben ser citados en el nuevo original y se incluirán copias, que junto con el manuscrito, se remitirán al director para ayudarle en la manera de abordar este asunto.

Protección del derecho a la intimidad de los pacientes

No debe infringirse el derecho a la intimidad de los pacientes sin su consentimiento informado. Por ello, no se publicará información que pueda identificar a alguna persona en textos, fotografías e historiales clínicos, a menos que dicha información sea esencial desde el punto de vista científico y el paciente (familiares o tutor) haya dado su consentimiento por escrito para su publicación. Dicho consentimiento se refiere a que el paciente tenga acceso al documento original que pretende publicarse.

Se omitirán los datos que permitan identificación alguna si no son esenciales, pero no deben alterarse o falsearse datos del paciente para lograr el anonimato. El total anonimato resulta difícil de lograr, y ante la duda se obtendrá el consentimiento informado. Por ejemplo, el hecho de ocultar los ojos en fotografías de pacientes no garantiza una adecuada protección del anonimato.

La obtención del consentimiento informado debe incluirse como requisito para la admisión de artículos y su obtención ha de mencionarse en el texto del artículo.

Organización de un manuscrito para su envío al editor

Algunas sugerencias generales que debe tener en mente antes incluso de escribir sus artículos, son: las características que sean, consúltelas en la revista que desea publicar desde antes de iniciar su trabajo de investigación, de manera que al recopilar la información y las referencias bibliohemerográficas, ya sepa los requerimientos que debe cubrir.

Algunas sugerencias como “receta” para organizar su manuscrito y enviarlo al editor de una revista:

1. De las consultas a las fuentes de información que realice, haga resúmenes y organícelos en orden lógico o cronológico según su investigación, anote en la misma tarjeta la referencia de la fuente consultada. Recuerde que las fuentes de información para investigación se clasifican en primarias cuando son revistas que publican artículos de investigación, sobre todo en revistas con arbitraje y secundarias cuando son de libros. Numere sus tarjetas y transcribalas armonizando la redacción a fin de que haya secuencia y continuidad en sus párrafos como se señaló al principio.
2. Haga archivos electrónicos y tenga SIEMPRE más de una copia en diversos medios (CD, *memory stick*, otra computadora) porque los accidentes, sobre todo cuando se está bajo estrés, no son raros.
3. Siga los lineamientos de la revista en donde desea publicar y obtenga todos los consentimientos relativos al trabajo (por ejemplo si utilizó cuadros o figuras de otro autor, o si hay coautores). En general los editores le solicitarán el consentimiento informado mediante escrito y firma.

Una vez que tenga su artículo completo:

4. Imprima el texto, marcando los lugares donde es más adecuado que vayan los cuadros y figuras ya numeradas en orden consecutivo.
5. Haga una buena impresión de sus cuadros (tablas) o figuras, numerándolas en orden consecutivo (Tabla 1, 2, 3... y Fig. 1, 2, 3...) y acomódelas al final del manuscrito impreso. Si envía fotografías, que sean de buena calidad, y en el reverso, con lápiz suave, coloque una flecha cuya punta señale la parte superior de la fotografía y añada el número de figura que le corresponda. Las figuras se numerarán consecutivamente según su primera mención en el texto. Si la figura ya fue anteriormente publicada, cite la fuente original y presente el permiso escrito del titular de los derechos de autor para la reproducción del material. Dicha autorización es necesaria, independientemente de quién sea el autor o la editorial; la única excepción se da en los documentos de dominio público.
6. Añada una hoja con el encabezado “pies de figura” y póngalos todos con su número consecutivo correspondiente.
7. Añada otra hoja con el encabezado “Cabezas de cuadro o tabla” y póngalos todos con su número consecutivo correspondiente.
8. Si envía algún “anexo” (carta de consentimiento informado, formato de encuesta, tablas estadísticas) imprímalo y añádalo al final de todo lo anterior.
9. Adjunte la hoja frontal con el título del artículo, el nombre completo de los autores, sus instituciones de procedencia, instituciones financiadoras y datos donde el editor pueda ponerse en contacto con el o los autores.

10. Indique al editor de la revista a la cual envía su artículo, que el tratamiento a los participantes, sean humanos o animales, estuvo de acuerdo con las normas éticas de la APA (“Ethical Principles of Psychologists and Code of Conduct”, APA, 1992a); o en su caso, las que recomienda el Comité Internacional de Directores de Revistas Médicas (CIDRM), que son las normas éticas del comité (institucional o regional) encargado de supervisar los ensayos en seres humanos y la declaración de Helsinki de 1975 modificada en 1983.
11. Coloque todo en un sobre de papel Manila al cual previamente se le escribió la editorial (si le es posible, el nombre del editor, que puede averiguar por teléfono o correo electrónico) y la dirección de la editorial, así como la dirección y nombre del remitente.
12. Envíelo por correo certificado o entréguelo en mano. Comuníquese lo antes posible con el editor para saber en qué tiempo puede tener una respuesta. Las posibles respuestas pueden ser: a) aceptado tal como está (ideal y prácticamente imposible); b) aceptado con la reserva de realizar los cambios sugeridos por los revisores y el editor (si sigue las sugerencias proporcionadas en este capítulo, será lo más frecuente), y c) rechazado.

Referencias. Son importantes para sustentar las afirmaciones realizadas respecto de lo que dijeron otros autores. Todas las citas del manuscrito deben aparecer en la lista de referencias, y todas las referencias deben ir citadas en el texto. La lista de referencias al final de un artículo en una revista científica proporciona a los lectores la información necesaria para identificar y localizar cada fuente. Es asunto ético incluir sólo aquellas fuentes que en realidad se utilizaron en la investigación y en la preparación del manuscrito. Cabe señalar la diferencia entre lista de referencias y bibliografía.

Lista de referencias. Cita trabajos que apoyan específicamente a un artículo en particular y a veces las revistas mencionan un número mínimo de referencias para apoyar un artículo. Los artículos de que trata este texto, requieren lista de referencias.

Bibliografía. Es más exhaustiva, cita trabajos que sirvieron de fundamento o son útiles para una lectura posterior. Puede incluir notas descriptivas y lecturas recomendadas.

Recuerde que los autores son responsables de su lista de referencias e incluya datos correctos y completos en cada entrada que sirvan para la busca de información en bibliotecas. En general, cada referencia debe contener los siguientes elementos: autor(es), año de publicación, título y datos de publicación. A continuación ejemplos de referencias con más detalle.

Artículo de revista:

King, J., Beals, J., Manson, S. M., & Trimble, J. E. (1992). A structural equation model of factors related to substance abuse among American Indian adolescents. *Drugs and Society*, 6(3-4), 253-268.

Libro completo:

Griffen, C. W., M. J., & Wirth, A. G. (1986). *Beyond acceptance: Parents of lesbians and gays talk about their experiences*. Englewood Cliffs, NJ: Prentice Hall.

Recuerde, no ponga referencias incompletas ni mal redactadas, sea cuidadoso al escribir los nombres de los autores, vale la pena tomarse el tiempo de redactar correctamente las referencias y seguir al pie de la letra las sugerencias de la revista donde desea publicar. De uno a seis autores menciónelos todos, de siete en adelante escriba: et al., (sin cursivas y con punto después de “al”).

Cualquier duda, se sugiere consultar: American Psychological Association (2002). Manual de estilo de publicaciones (adaptado para el español por Editorial El Manual Moderno), 2a. ed., México.

Requisitos para publicaciones del área de ciencias biomédicas

A continuación se expone lo que el Comité Internacional de Directores de Revistas Médicas establece como requisitos para publicar en este tipo de revistas. Se expresa sólo lo más frecuente y lo que difiere en cuanto a lo normado por la APA, a fin de que el texto sea realmente útil a nuestros lectores. Para información más detallada, se sugiere visitar la página: <http://www.icmje.org>.

Antecedentes. En 1978 el Comité Internacional de Directores de Revistas Médicas se reunió informalmente en Vancouver, Columbia Británica, para establecer las directrices que en cuanto a formato debían incluir los manuscritos enviados a sus revistas. El grupo se conoce como Grupo Vancouver. Sus requisitos para manuscritos, que incluían formatos para las referencias bibliográficas desarrollados por la National Library of Medicine (NLM) de Estados Unidos, se publicaron por primera vez en 1979. El Grupo Vancouver creció y se convirtió en el Comité Internacional de Directores de Revistas Médicas (CIDRM).

Los requisitos uniformes son instrucciones para los autores sobre cómo preparar sus manuscritos y los directores de las revistas que aceptan dichos requisitos, se comprometen a no devolver los manuscritos para que se realicen cambios de estilo. Sin embargo, en el proceso editorial las revistas pueden modificar los manuscritos aceptados para adecuarlos a su estilo de publicación.

De igual manera, los autores que remitan sus manuscritos a una revista que participe de esta normativa, no deben preparar los mismos según el estilo de la revista en concreto sino que debe seguir los requisitos uniformes.

Los autores seguirán también las instrucciones de cada revista respecto de qué temas son pertinentes y el tipo de artículos que admite: por ejemplo, originales, revisiones o notas clínicas. Además, es probable, que en dichas instrucciones figuren otros requisitos específicos de la publicación que deban seguirse, como el número de copias del manuscrito, idiomas aceptados, extensión del artículo y abreviaturas admitidas.

Requisitos para el envío de manuscritos

Resumen de los requisitos técnicos.

Doble espacio en todo el artículo.

Inicie cada sección o componente del artículo en una página.

Revise la ordenación: página del título, resumen y palabras clave, texto, agradecimientos, referencias bibliográficas, cuadros o tablas (en páginas por separado) y leyendas.

El tamaño de las ilustraciones, positivo sin montar, no debe superar los 203 × 254 mm (8 × 10 pulgadas).

Incluya las autorizaciones para la reproducción de material anteriormente publicado o para la utilización de ilustraciones que puedan identificar a personas.

Adjunte la cesión de los derechos de autor y formularios pertinentes.

Anexe el número de copias en papel que pida la revista en particular a la cual envía su trabajo.

Conserve una copia de todo el material enviado.

Preparación del original

Artículos en soporte electrónico.

Algunas revistas solicitan de los autores una copia en soporte electrónico (en disquete o memoria); pudiendo aceptar diversos formatos de procesadores o ficheros de textos (ASCII).

Al presentar los disquetes, los autores deben cerciorarse de que se ha incluido una versión del manuscrito en el disquete.

Incluir en el disquete solamente la versión última del manuscrito.

Especificar claramente el nombre del archivo.

Etiquetar el disquete con el formato y nombre del fichero.

Facilitar la información sobre el software y hardware utilizado.

Los autores deberán consultar en la sección de normas para los autores de la revista, las instrucciones en lo que se refiere a qué formatos se aceptan, las convenciones para denominar los archivos y disquetes, el número de copias que ha de enviarse y otros detalles.

Autoría

Todas las personas que figuren como autores habrán de cumplir con ciertos requisitos para recibir tal denominación. Cada autor deberá haber participado en grado suficiente para asumir la responsabilidad pública del contenido del trabajo. Uno o varios autores deberán responsabilizarse o encargarse de la totalidad del trabajo, desde el inicio hasta que el artículo haya sido publicado.

Para concederle a alguien el crédito de autor, hay que basarse únicamente en su contribución esencial en lo que se refiere a: a) la concepción y el diseño del estudio, acopio de los datos o el análisis y la interpretación de los mismos; b) la redacción del artículo o la revisión crítica de una parte sustancial de su contenido intelectual, y c) la aprobación final de la versión que será publicada. Los requisitos a), b) y c) tendrán que cumplirse simultáneamente. La participación exclusivamente en la obtención de fondos o en la recopilación de datos o la supervisión general del grupo de investigación no justifica la autoría.

Los directores de las revistas podrán solicitar a los autores que describan la participación de cada uno de ellos y esta información puede ser publicada. El resto de personas que contribuyan al trabajo y que no sean los autores deben citarse en la sección de agradecimientos.

Cada vez con mayor frecuencia, se realizan ensayos multicéntricos que se atribuyen a un autor corporativo. En estos casos, todos los miembros del grupo que figuren como autores deben satisfacer totalmente los criterios de autoría anteriormente citados. Los miembros del grupo que no satisfagan estos criterios deben ser mencionados, con su autorización, en la sección de agradecimientos o en apéndice (véase Agradecimientos). El orden de los autores dependerá de la decisión que de forma conjunta adopten los coautores. En todo caso, los autores deben ser capaces de explicar el mismo.

Palabras clave

Tras el resumen los autores deberán presentar e identificar como tales, de tres a 10 palabras clave que faciliten a los documentalistas el análisis del artículo y que se publicarán junto con el resumen. Utilice para este fin los términos del tesoro Medical Subject Headings (MeSH) del Index Medicus; en el caso de que se trate de términos de reciente aparición que aún no figuren en el MeSH pueden usarse los nuevos términos.

Nota: puede consultar una edición en español del MeSH elaborado por BIREME: “Descriptores de Ciencias de la Salud” [DeSC].

Estadística

Describa los métodos estadísticos con el suficiente detalle para permitir que un lector versado en el tema con acceso a los datos originales pueda verificar los resultados publicados. En la medida de lo posible, cuantifique los hallazgos y presente los mismos con los indicadores apropiados de error o de incertidumbre de la medición (como los intervalos de confianza). Se evitará la dependencia exclusiva de las pruebas estadísticas de verificación de hipótesis, tal como el uso de los valores P , que no aportan ninguna información cuantitativa importante. Analice los criterios de inclusión de los sujetos experimentales. Proporcione detalles sobre el proceso que se ha seguido en la distribución aleatoria. Describa los métodos de enmascaramiento utilizados. Haga constar las complicaciones del tratamiento. Especifique el número de

observaciones realizadas. Indique las pérdidas de sujetos de observación (como los abandonos en un ensayo clínico). Siempre que sea posible, las referencias sobre el diseño del estudio y métodos estadísticos serán de trabajos vigentes (indicando el número de las páginas) en lugar de los artículos originales donde se describieron por vez primera.

Especifique cualquier programa de computadora, de uso común, que se haya empleado.

Agradecimientos

Incluya la relación de todas aquellas personas que han colaborado pero que no cumplan los criterios de autoría, tales como ayuda técnica recibida, ayuda en la escritura del manuscrito o apoyo general prestado por el jefe del departamento. También se incluirá en los agradecimientos el apoyo financiero y los medios materiales recibidos.

Las personas que hayan colaborado en la preparación del original, pero cuyas contribuciones no justifiquen su acreditación como autores podrán ser citadas bajo la denominación de “investigadores clínicos” o “investigadores participantes” y su función o tipo de contribución debería especificarse, por ejemplo, “asesor científico”, “revisión crítica de la propuesta de estudio”, “recogida de datos” o “participación en el ensayo clínico”.

Dado que los lectores pueden deducir que las personas citadas en los agradecimientos de alguna manera avalan los datos y las conclusiones del estudio, se obtendrá la autorización por escrito de las personas citadas en dicha sección.

Referencias bibliográficas

Numere las referencias consecutivamente, según el orden en que se mencionen por primera vez en el texto. En éste, en las tablas y leyendas, las referencias se identificarán mediante números arábigos entre paréntesis. Las referencias citadas únicamente en las tablas o ilustraciones se numerarán siguiendo la secuencia establecida por la primera mención que se haga en el texto de la tabla o figura en concreto.

Se utilizará el estilo de los ejemplos que a continuación se ofrecen, que se basan en el estilo que utiliza la NLM en el Index Medicus. Abrevie los títulos de las revistas según el estilo que utiliza el Index Medicus. Consulte la List of Journals Indexed in Index Medicus (relación de revistas indizadas en el Index Medicus), que la NLM publica anualmente como parte del número de enero del Index Medicus, y como separata. Esta relación también puede obtenerse en la dirección Web de la NLM.

Evite citar resúmenes. Las referencias que se realicen de originales aceptados pero aún no publicados se indicará con expresiones del tipo “en prensa” o “próxima publicación”; los autores deberán obtener autorización escrita y tener constancia que su publicación está aceptada. La información sobre manuscritos presentados a una revista pero no aceptados cítela en el texto como “observaciones no publicadas”, previa autorización por escrito de la fuente.

Tampoco cite una “comunicación personal”, salvo cuando en la misma se facilite información esencial que no se halla disponible en fuentes públicamente accesibles, en estos casos se incluirán, entre paréntesis en el texto, el nombre de la persona y la fecha de la comunicación. En los artículos científicos, los autores que citen una comunicación personal deberán obtener la autorización por escrito.

Envío del manuscrito a la revista

Los manuscritos se acompañarán de una carta de presentación firmada por todos los autores. Esta carta debe incluir:

1. Información acerca de la publicación previa o duplicada o el envío de cualquier parte del trabajo a otras revistas, como se ha indicado anteriormente.

2. Una declaración de que el manuscrito ha sido leído y aprobado por todos los autores, que se ha cumplido con los requisitos de autoría expuestos y que cada autor cree que el artículo constituye un trabajo honesto.
3. Nombre, dirección y número de teléfono del autor encargado de la coordinación con los coautores en lo concerniente a las revisiones y aprobación final de las pruebas de imprenta del artículo en cuestión.

Ejemplos de las referencias bibliográficas

Nota: Los Requisitos de Uniformidad (actualización octubre 2005) contienen 41 ejemplos de diferentes documentos que pueden utilizarse como referencias bibliográficas (ejemplos disponibles en la National Library of Medicine NLM). En el caso del “Material electrónico” (ej. ref. 35-41), los ejemplos no abarcan la enorme variedad de documentos disponibles en la Red. En el supuesto de que en los ejemplos disponibles no se incluya el tipo de documento que usted requiere citar o surja una duda, se recomienda consultar el documento publicado por NLM en el 2001 sobre las citas bibliográficas en internet.

Artículos de revistas

(1) Artículo estándar

Autor(es). Título del artículo. Abreviatura internacional de la revista. Año; volumen (número): página inicial-final del artículo.

Medrano MJ, Cerrato E, Boix R, Delgado-Rodríguez M. Factores de riesgo cardiovascular en la población española: metaanálisis de estudios transversales. *Med Clin* (Barc). 2005; 124(16): 606-12.

(2) Organización o equipo como autor

Grupo de Trabajo de la SEPAR. Normativa sobre el manejo de la hemoptisis amenazante. *Arch Bronconeumol* 1997; 33: 31-40.

(3) Autoría compartida entre autores y un equipo

Jiménez Hernández MD, Torrecillas Narváez MD, Frieria Acebal G. Grupo Andaluz para el Estudio de Gabapentina y Profilaxis Migrañosa. Eficacia y seguridad de la gabapentina en el tratamiento preventivo de la migraña. *Rev Neurol*. 2002; 35: 603-6.

(4) No se indica autor

21st century heart solution may have a sting in the tail. *BMJ*. 2002; 325(7357): 184.

(5) Artículo en otro idioma distinto del inglés

Sartori CA, Dal Pozzo A, Balduino M, Franzato B. Exérèse laparoscopique de l'angle colique gauche. *J Chir* (París). 2004; 141: 94-105.

(6) Suplemento de un volumen

Plaza Moral V, Álvarez Gutiérrez FJ, Casan Clará P, Cobos Barroso N, López Viña A, Llauger Rosselló MA et al. Comité Ejecutivo de la GEMA. Guía Española para el Manejo del Asma (GEMA). *Arch Bronconeumol*. 2003; 39 Supl 5: 1-42.

(7) Suplemento de un número

Glauser TA. Integrating clinical trial data into clinical practice. *Neurology*. 2002; 58 (12 Suppl 7): S6-12.

(8) Parte de un volumen

Abend SM, Kulish N. The psychoanalytic method from an epistemological viewpoint. *Int J Psychoanal*. 2002; 83(Pt 2): 491-5.

(9) Parte de un número

Ahrar K, Madoff DC, Gupta S, Wallace MJ, Price RE, Wright KC. Development of a large animal model for lung tumors. *J Vasc Interv Radiol*. 2002; 13(9 Pt 1): 923-8.

(10) Número sin volumen

Fleta Zaragozano J, Lario Elboj A, García Soler S, Fleta Asín B, Bueno Lozano M, Ventura Faci P et al. Estreñimiento en la infancia: pauta de actuación. *Enferm Cient*. 2004; (262-263): 28-33.

(11) Sin número ni volumen

Outreach: bringing HIV-positive individuals into care. *HRSA Careaction*. 2002 Jun:1-6.

(12) Paginación en número romanos

Chadwick R, Schuklenk U. The politics of ethical consensus finding. *Bioethics*. 2002; 16(2): III-V.

(13) Artículo publicado electrónicamente antes que en versión impresa

Nota: Las citas *Epub ahead of print*, son referencias enviadas a PubMed por los editores de revistas que se publican en primera instancia on-line, adelantándose a la edición en papel. Posteriormente, cuando se publica en formato impreso, la referencia se modifica apareciendo los datos de la edición impresa, seguida de la electrónica Epub. Ejemplo de una referencia en PubMed publicada en edición electrónica y posteriormente cuando se publica impresa.

Sait KH, Ashour A, Rajabi M. Pregnancy outcome in non-gynecologic cancer. *Arch Gynecol Obstet*. 2004 Jun 2 [Epub ahead of print].

Sait KH, Ashour A, Rajabi M. Pregnancy outcome in non-gynecologic cancer. *Arch Gynecol Obstet*. 2005 Apr; 271(4): 346-9. Epub 2004 Jun 2.

Libros y otras monografías**(14) Autores individuales**

Autor(es). *Título del libro*. Edición. Lugar de publicación: Editorial; año.

Jiménez Murillo L, Montero Pérez FJ. *Compendio de Medicina de Urgencias: guía terapéutica*. 2ª ed. Madrid: Elsevier; 2005.

Nota: La primera edición no es necesario consignarla. La edición siempre se indica números arábigos y abreviatura: 2ª ed. Si la obra estuviera compuesta por más de un volumen, debe citarlo a continuación del título del libro Vol. 3.

(15) Director(es), compilador(es) como autor

Espinás Boquet J. coordinador. *Guía de actuación en Atención Primaria*. 2ª ed. Barcelona: Sociedad Española de Medicina; 2002.

Teresa E de, editor. *Cardiología en Atención Primaria*. Madrid: Biblioteca Aula Médica; 2003.

Nota: En la edición original figura “Editor” término inglés que se refiere al editor literario. En español este término debe traducirse como director (de una revista) o director, compilador o coordinador (de un libro). En español es frecuente que se utilice de manera incorrecta (anglicismo) el término inglés “Editor” como sinónimo de director o coordinador. Si figura ese término, debe conservarlo.

(16) Autor(es) y editor(es)

Breedlove GK, Schorfheide AM. *Adolescent pregnancy*. 2ª ed. Wiecezorek RR, editor. White Plains (NY): March of Dimes Education Services; 2001.

(17) Organización como autor

Comunidad de Madrid. *Plan de Salud Mental de la Comunidad de Madrid 2003-2008*. Madrid: Comunidad de Madrid, Consejería de Sanidad; 2002.

American Psychiatric Association. *Guías clínicas para el tratamiento de los trastornos psiquiátricos*. Barcelona: Ars MEDICA; 2004.

(18) Capítulo de libro

Autor(es) del capítulo. Título del capítulo. En*: Director/Coordinador/Editor del libro. Título del libro. Edición. Lugar de publicación: Editorial; año. Página inicial-final del capítulo.

Mehta SJ. Dolor abdominal. En: Friedman HH, coordinador. *Manual de Diagnóstico Médico*. 5ª ed. Barcelona: Masson; 2004. p. 183-90.

(19) Actas de congresos

Segundo Encuentro Latinoamericano de Psicología Ambiental; Grupo Entorno Comportamiento. 11-15 de Septiembre 2002. Los Reyes Iztacala, Tlalnepantla, Estado de México: Facultad de Estudios Superiores Iztacala, UNAM; 2002.

(20) Comunicación presentada a un congreso

Autor(es) de la Comunicación/Ponencia. Título de la Comunicación/Ponencia. En: Título oficial del Congreso. Lugar de Publicación: Editorial; año. página inicial-final de la comunicación/ponencia.

Castro Beiras A, Escudero Pereira J. El Área del Corazón del Complejo Hospitalario "Juan Canalejo". En: *Libro de Ponencias: V Jornadas de Gestión y Evaluación de Costes Sanitarios*. Bilbao; Ministerio de Sanidad y Consumo, Gobierno Vasco; 2000. p. 12-22.

Nota: Esta misma estructura se aplica a jornadas, simposios, reuniones científicas, etcétera.

(21) Informe científico o técnico

Autor(es). Título del informe. Lugar de publicación: Organismos/Agencia editora; año. Número o serie identificativa del informe.

Organización Mundial de la Salud. *Factores de riesgo de enfermedades cardiovasculares: nuevas esferas de investigación*. Informe de un Grupo Científico de la OMS. Ginebra: OMS; 1994. Serie de Informes Técnicos: 841.

Patrocinado por un organismo o institución:

Ahn N, Alonso Meseguer J, Herce San Miguel JA. *Gasto sanitario y envejecimiento*. Madrid: Fundación BBVA; 2003. Documentos de trabajo: 7.

(22) Tesis doctoral

Autor. Título de la tesis [tesis doctoral].† Lugar de publicación: Editorial; año.

Muñiz García J. *Estudio transversal de los factores de riesgo cardiovascular en población infantil del medio rural gallego* [tesis doctoral]. Santiago: Servicio de Publicaciones e Intercambio Científico, Universidad de Santiago; 1996.

Otros trabajos publicados**(23) Artículo de periódico**

Autor del artículo†. Título del artículo. Nombre del periódico††. Día mes año; Sección†††: página (columna)††††.

† Autor del artículo (si figurase).

†† Los nombres de periódicos no se facilitan abreviados.

††† Si existiera identificada como tal.

†††† Si aparece identificada.

Carrasco D. Avalado el plazo de cinco años para destruir parte de la HC. *Diario Médico*. Viernes 23 de julio de 2004; Normativa: 8.

Espiño I. ¿Le va mejor al paciente que participa en un ensayo clínico? *El Mundo sábado* 31 de enero de 2004. Salud: S6 (Oncología).

(24) Material audiovisual

Autor(es). Título de la videocinta [videocinta]. Lugar de edición: Editorial; año. Aplicable a todos los soportes audiovisuales.

Borrel F. La entrevista clínica. Escuchar y preguntar. [video] Barcelona: Doyma; 1997.

(25) Documentos legales: Leyes/Decretos/Órdenes.

Título de la ley/decreto/orden. (Nombre del Boletín Oficial, número, fecha de publicación)

Estatuto Marco del personal estatutario de los servicios de salud. Ley 55/2003 de 16 de diciembre. Boletín Oficial del Estado, nº 301 (17-12-2003).

(26) Diccionarios y obras de consulta

Dorland Diccionario Enciclopédico Ilustrado de Medicina. 28ª ed. Madrid: McGraw-Hill, Interamericana; 1999. Afasia; p. 51.

Material no publicado

(27) En prensa

Nota: NLM prefiere “de próxima aparición” (en inglés: forthcoming) debido a que no todos los temas serán publicados.

Leshner AI. Molecular mechanisms of cocaine addiction. *N Engl J Med*. En prensa 1997.

Material electrónico

Autor(es). Título [CD-ROM, DVD, Disquete]. Edición. Lugar: Editorial; año.

Best CH. Bases fisiológicas de la práctica médica [CD-ROM]. 13ª ed. Madrid: Editorial Médica Panamericana; 2003.

(28) Artículo de revista en internet

Autor(es) del artículo. Título del artículo. Nombre de la revista [revista en internet, revista en línea o revista on-line] año [fecha de consulta]; volumen (número): [Extensión/páginas*]. Dirección electrónica.

Francés I, Barandiarán M, Marcellán T, Moreno L. Estimulación psicocognoscitiva en las demencias. *An Sist Sanit Navar* [revista en internet]* 2003 septiembre-diciembre. [Acceso 19 de octubre de 2005]; 26(3). Disponible en: <http://www.cfnavarra.es/salud/anales/textos/vol26/n3/revis2a.html>

* Si constasen.

Recomendaciones para escribir referencias bibliográficas

Las referencias o citas bibliográficas constituyen una sección destacada en un trabajo científico. La selección cuidadosa de documentos relevantes, es un elemento que da solidez a la exposición teórica del texto, a la vez que constituye una importante fuente de información para el lector. Se facilita una serie de indicaciones para elaborar las referencias bibliográficas basadas en los requisitos de uniformidad (estilo Vancouver).

Deben numerarse consecutivamente según el orden en que se mencionen por primera vez en el texto. Algunas revistas en sus instrucciones para autores recomiendan que se utilicen números arábigos en superíndice y sin paréntesis.

Cuando hay más de una cita, éstas deben separarse mediante comas, pero si fueran correlativas, se menciona la primera y la última separadas por un guión.

Cuando en el texto se menciona un autor, el número de la referencia se pone tras el nombre del autor. Si se tratase de un trabajo realizado por más de dos autores, se cita el primero de ellos seguido de la abreviatura “et al.” y su número de referencia.

Los documentos que se citen deben ser actuales. Algunas revistas señalan que no deben tener más de cinco años y preferiblemente que sean de los dos últimos.

Se recurre a citar documentos que tengan más años por motivos históricos o si no se encuentran referencias actualizadas como alternativa.

Para citar adecuadamente los documentos electrónicos, es recomendable consultar el documento sobre las citas bibliográficas en internet publicado por la National Library of Medicine de Estados Unidos, o la norma de la International Standards Organization (ISO 690-2) para documentos electrónicos.

Respecto del número de citas a incluir en cada trabajo, las revistas suelen recomendar que los trabajos originales incluyan entre 20-30 referencias; los originales breves y notas clínicas, entre 10 y 20 referencias; las cartas al director un máximo de 10. Para otras secciones: revisiones, editoriales, se recomienda consultarlo en las instrucciones para autores o al comité de redacción.

Los títulos de las revistas deben abreviarse según el estilo que utiliza la National Library of Medicine (NLM). Puede consultarse el Journals Database de PubMed. Para comprobar las abreviatura de revistas españolas, puede consultarse el catálogo C17 (Catálogo colectivo de publicaciones periódicas de las Bibliotecas de Ciencias de la Salud Españolas). En el supuesto de no localizar una abreviatura, puede consultarse la “List of Serial Title Word Abbreviations Internacional” conforme a la norma ISO 4, o bien el “The List of Title Word Abbreviations” de la agencia ISSN.

Una vez finalizada la lista de fuentes consultadas, tiene que asegurarse de la correspondencia de las citas en el texto y el número asignado en la referencia.

Nuestra sugerencia final es no darse por vencido, practicar, preparar los manuscritos, consultar en libros o a los editores cuando haya dudas y... publicar. Suerte.



REFERENCIAS

Popper K.R. (1974). *Conocimiento objetivo*. Madrid: Tecnos.

American Psychological Association (APA). (2002). *Manual de estilo de publicaciones* (adaptado para el español) 2a. edición. México: Editorial El Manual Moderno.

American Psychological Association (APA). (1994). *Publication Manual of the American Psychological Association* (4a. ed.), Washington, D.C.: Author.

Bobenrieth Astete MA. *El artículo científico original. Estructura, estilo y lectura crítica*. Granada: Juan de Andalucía, Escuela Andaluza de Salud Pública, 1994.

International Committee of Medical Journal Editors. Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication. Updated October 2005 [Internet]. CMJE; 2005 [acceso 15 de julio de 2006]. Disponible en: <http://www.icmje.org/>

Luna Castillo, Antonio (2002). *Metodología de la tesis*. México: Trillas.

National Library of Medicine Recommended Formats for Bibliographic Citation [Internet]. Bethesda: National Library of Medicine, diciembre 2003 [acceso 17 de diciembre de 2005]. Disponible en: <http://www.nlm.nih.gov/pubs/formats/recommendedformats.html>

Ramos, M.M., Catena, A. y Trujillo, H. M. (2004). *Manual de métodos y técnicas de investigación en ciencias del comportamiento*. Madrid: Biblioteca Nueva.

Real Academia Española (1992). *Diccionario de la Lengua Española* (vigésima primera edición, vol. 2). Madrid: España.

Rodríguez Bonache MJ. Cómo se debe citar un artículo científico? Rehabilitación (Madrid). 2002; 36:67-69.

Tamayo y Tamayo Mario (2004). *El proceso de la investigación científica*. México: Limusa, Noriega Editores.

Apéndice B

Matrices



El manejo del álgebra matricial es importante para la resolución de diversos problemas de estadística y para la comprensión de algunos capítulos posteriores. Por ello se abunda en el desarrollo de la teoría elemental de las matrices y su relación con algunos modelos estadísticos lineales.

Concepto Una matriz es un conjunto (arreglo) de elementos ordenados tanto en renglones como en columnas. Esos elementos se denotan por a_{ij} , los renglones mediante la letra m y las columnas por la letra n . Así:

Sea una matriz que tiene los renglones $-1, 2$ y $0, 1$; entonces, $m = 2$ y $n = 2$ es una matriz 2×2 , y se representa en la forma siguiente:

$$A_{2 \times 2} = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}$$

Orden de una matriz

El orden de una matriz estará expresado mediante un par ordenado, donde el primer componente indica los renglones que tiene la matriz y, el segundo, las columnas.

EJEMPLO

1

$$B_{2 \times 3} = \begin{bmatrix} 3 & 1 & 0 \\ -2 & -4 & -1 \end{bmatrix} \text{ La matriz es de orden } 2 \times 3$$

$$C_{1 \times 3} = [0, 1, -1] \text{ La matriz es de orden } 1 \times 3$$

$$E_{3 \times 1} = \begin{bmatrix} 6 \\ 7 \\ -9 \end{bmatrix} \text{ La matriz es de orden } 3 \times 1$$

Una matriz de orden $m \times n$, será:

$$A_{m \times n} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}$$

Las matrices son básicas en la teoría económica contemporánea, en psicología, en la administración moderna y, en general, en la investigación (por ejemplo, el proceso industrial y control de la calidad total). En otros campos, como las ciencias de la salud, los médicos y biólogos utilizan las matrices como un método muy útil en el estudio de interrelaciones genéticas de la herencia; donde se utilizan ampliamente las operaciones y propiedades de las matrices, entre estos métodos pueden mencionarse los análisis factoriales.

Tipos de matrices

Las matrices toman diferentes denominaciones, atendiendo a su orden.

Matriz renglón o vector renglón $A_{1 \times n}$

EJEMPLO

2

$$A_{1 \times 4} = (2, -1, 0, 1)$$

Matriz columna $A_{m \times 1}$

$$A_{4 \times 1} = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2 \end{bmatrix} \quad A_{3 \times 1} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} \quad A_{2 \times 1} = \begin{bmatrix} 5 \\ -4 \end{bmatrix}$$

También se conoce como *vector columna*.

Matriz cuadrada

Se denomina así a las matrices que tienen el mismo número de renglones y de columnas, o sea $m = n$.

EJEMPLO

3

$$A_2 = \begin{bmatrix} -2 & 1 \\ 0 & 2 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 1 & 3 \\ 2 & -1 & 0 \\ 1 & -2 & 0 \end{bmatrix}$$

Matriz diagonal

Es una matriz cuadrada, cuyos elementos $a_{ij} = 0$ para toda $i \neq j$.

EJEMPLO

4

$$\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

En este tipo de matrices, los únicos valores diferentes de cero se encuentran sobre la diagonal principal (de izquierda a derecha).

Matriz escalar

Es una matriz diagonal, con a_{ij} iguales.

EJEMPLO

5

$$\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \quad \begin{bmatrix} -4 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

Por lo general, la matriz escalar se representa de la siguiente manera:

$$\begin{bmatrix} \lambda & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix} = \lambda \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & 1 \end{bmatrix} = 1$$

Matriz identidad

Llamada también *matriz unitaria*, es una matriz escalar donde $\lambda = 1$, y se representa por I , de tal manera que para conocer el orden se le coloca el subíndice respectivo.

EJEMPLO

6

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Matriz nula

Es aquella matriz en la cual todos sus elementos son iguales a cero.

Igualdad de matrices

Dos matrices, A y B , son iguales cuando tienen el mismo orden, y todos y cada uno de sus elementos cuentan con el mismo valor algebraico.

EJEMPLO

7

$$A = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 \\ -1 & x \end{bmatrix}$$

$$A = B \text{ si y sólo si } x = 2$$

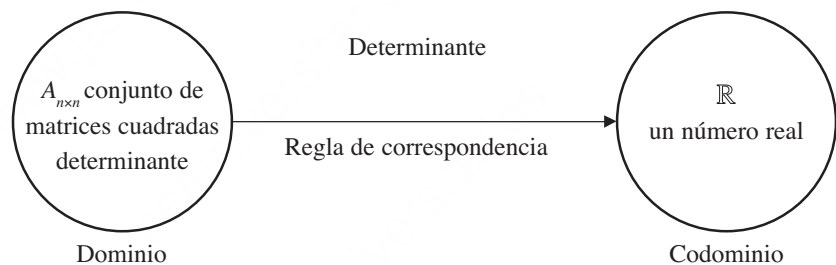
$$C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad D = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$$

$$C = D \text{ si y sólo si } w = 1, \quad x = 2, \quad y = 3, \quad z = 4$$

Determinantes

Antes de abordar el desarrollo del álgebra de matrices, se explicará el concepto y algunas propiedades de los determinantes, que se estudiarán más adelante.

Un determinante es un número real asociado con una matriz cuadrada, mediante una función llamada *determinante* y que se denota $\det A$, $|A|$, o Δ , siendo A una matriz cuadrada. El dominio es el conjunto de las matrices cuadradas y el codominio son los reales.



Determinantes para matrices de 2×2

Ahora se explica cómo se calcula el determinante de una matriz 2×2 : es el producto de los elementos de la diagonal principal, *menos* el producto de los elementos de la otra diagonal.

EJEMPLO

8

$$\begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix} = A \quad \det A = \begin{vmatrix} 2 & 1 \\ 3 & 0 \end{vmatrix} = 2 \times 0 - 3 \times 1 = -3$$

$$\det A = -3$$

Se obtuvo como resultado un número real, y es importante mencionar que este número real es único para cada matriz.

Determinantes para matrices de 3×3

A continuación se explica cómo se calcula el determinante para matrices de 3×3 . Para ello, es necesario dar a conocer dos conceptos.

Sea A una matriz de 3×3

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Como puede observar, cada elemento tiene su posición ij , donde i = número de renglón, también llamada fila, y j = número de columna.

Definición Si A es una matriz cuadrada de orden 3, entonces el menor M_{ij} de un elemento a_{ij} es el determinante de la matriz de orden 2, que se obtiene al omitir el renglón i y la columna j de A .

Definición El cofactor A_{ij} del elemento a_{ij} se define como

$$A_{ij} = (-1)^{i+j} M_{ij}$$

Las definiciones anteriores se explican con el ejemplo siguiente. Sea

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -8 & -1 & 7 \\ -6 & 4 & 0 \end{bmatrix}$$

Como primer paso, se calculan los menores, es decir, M_{ij} :

$$M_{11} = \begin{bmatrix} -1 & 7 \\ 4 & 0 \end{bmatrix} = (-1)(0) - (4)(7) = -28$$

$$M_{12} = \begin{bmatrix} -8 & 7 \\ -6 & 0 \end{bmatrix} = (-8)(0) - (-6)(7) = +42$$

$$M_{13} = \begin{bmatrix} -8 & -1 \\ -6 & 4 \end{bmatrix} = (-8)(4) - (-1)(-6) = -32 - 6 = -38$$

El segundo paso consiste en calcular los cofactores A_{ij} :

$$A_{11} = (-1)^{1+1}$$

$$A_{12} = (-1)^{1+2}$$

$$A_{13} = (-1)^{1+3}$$

$$M_{11} = (-1)^2 (-28) = -28$$

$$M_{12} = (-1)^3 (42) = -42$$

$$M_{13} = (-1)^4 (-38) = -38$$

El tercer y último paso es obtener el determinante por medio de la siguiente definición:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} A_{11} + a_{12} A_{12} + a_{13} A_{13}$$

En notación de suma, se expresa de esta forma:

$$\det A = \sum_{k=1}^3 a_{1k} A_{1k}$$

Sustituya los valores y obtiene:

$$\det A = 1(-28) + 2(-42) + 3(-38) = -28 - 84 - 114 = -226$$

$$\det A = -226$$

Álgebra de matrices

Transpuesta de una matriz

Es la matriz que resulta de intercambiar el primer renglón con la primera columna, el segundo renglón con la segunda columna, y así sucesivamente. El orden de la matriz queda intercambiado: si es de orden (2×3) , quedará (3×2) . La transpuesta de una matriz A se representa por A^T .

EJEMPLO

9

$$A_{2 \times 2} = \begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix} \quad A_{2 \times 2}^T = \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}$$

$$E_{2 \times 3} = \begin{bmatrix} 1 & 3 & -2 \\ 0 & 1 & -1 \end{bmatrix} \quad E_{3 \times 2}^T = \begin{bmatrix} 1 & 0 \\ 3 & 1 \\ -2 & -1 \end{bmatrix}$$

Matriz simétrica

Se denomina así una matriz cuando es igual a su transpuesta $A = A^T$.

EJEMPLO

10

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix} \quad A^T = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix}$$

Observe que esto sólo es posible cuando las matrices son del mismo orden, o sea, una matriz simétrica es cuadrada.

Operaciones con matrices

Suma de matrices

Para que la suma de dos matrices pueda realizarse, hace falta que sean del mismo orden.

EJEMPLO

11

$$\begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix} + \begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} -2+3 & 1+(-1) \\ 0+0 & 3+(-2) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

La suma algebraica se realiza con los elementos correspondientes. Sean:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ -1 & 3 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 3 & -1 & 1 \\ 1 & 2 & 2 \\ 1 & 0 & 4 \end{bmatrix}$$

$$A + B = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ -1 & 3 & -1 \end{bmatrix} + \begin{bmatrix} 3 & -1 & 1 \\ 1 & 2 & 2 \\ 1 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1+3 & 2+(-1) & -1+1 \\ 0+1 & 1+2 & 0+2 \\ -1+1 & 3+0 & -1+4 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & 3 & 5 \end{bmatrix}$$

Propiedades de la suma de matrices:

1. $A + B = B + A$ Propiedad conmutativa
2. $(A + B) + C = A + (B + C)$ Propiedad asociativa

Multiplicación de matrices

Producto de una matriz por un escalar

El resultado de multiplicar una matriz por cualquier número real (escalar) es el producto de cada uno de los componentes por dicho escalar.

EJEMPLO

12

Multiplicar el escalar 5 por la matriz A, donde:

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 3 & 0 \\ -1 & 1 & 4 \end{bmatrix} \quad 5A = \begin{bmatrix} 10 & 0 & -5 \\ 5 & 15 & 0 \\ -5 & 5 & 20 \end{bmatrix}$$

Matriz opuesta

Se obtiene cuando una matriz se multiplica por (-1) .

EJEMPLO

13

$$(-1) \begin{bmatrix} 5 & -3 \\ 2 & 6 \end{bmatrix} = \begin{bmatrix} -5 & 3 \\ -2 & -6 \end{bmatrix}$$

Resta de matrices

Es la suma algebraica de una matriz dada con la matriz opuesta de la matriz que se quiere restar, o sea:

$$A - B = A + (-1)B$$

EJEMPLO

14

Sean las matrices A y B , tales que:

$$\begin{bmatrix} 7 & 3 \\ 4 & -2 \end{bmatrix} - \begin{bmatrix} 5 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 7 & 3 \\ 4 & -2 \end{bmatrix} + (-1) \begin{bmatrix} 5 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 7 & 3 \\ 4 & -2 \end{bmatrix} + \begin{bmatrix} -5 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 3 & -4 \end{bmatrix} = A - B$$

Producto de dos matrices

En la multiplicación de dos matrices ($A \times B$), el número de columnas de A debe ser igual al número de renglones de B , es decir, son matrices conformables.

EJEMPLO

15

Sean A y B , tales que:

$$\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \times \begin{bmatrix} 1 & -1 & 3 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2(1)+1(2) & 2(-1)+1(1) & 2(3)+1(0) \\ 3(1)+4(2) & 3(-1)+4(1) & 3(3)+4(0) \end{bmatrix} = \begin{bmatrix} 4 & -1 & 6 \\ 11 & 1 & 9 \end{bmatrix} = C_{2 \times 3}$$

De este modo, si A es de orden $m \times n$ y B es de orden $n \times p$, el producto es una matriz C de orden $m \times p$. Dadas las siguientes matrices, realizar sus productos:

$$(-1, 2, 1) \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} = [(-1) 3 + 2(1) + 1(4)] = [3]$$

$$\begin{bmatrix} 5 \\ -2 \end{bmatrix} [1, -1, 3] = \begin{bmatrix} 5(1) & 5(-1) & 5(3) \\ -2(1) & -2(-1) & -2(3) \end{bmatrix} = \begin{bmatrix} 5 & -5 & 15 \\ -2 & 2 & -6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1(1)+0(0) & 1(0)+0(1) \\ 2(1)+3(0) & 2(0)+3(1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$$

Cualquier matriz A multiplicada por la matriz unitaria o identidad, será la misma matriz A .
Propiedades del producto de matrices.

i)	$A \times B \neq B \times A$	No cumple la propiedad conmutativa
ii)	$A \times (B \times C) = (A \times B) \times C$	Propiedad asociativa
iii)	$A \times (B + C) = (A \times B) + (A \times C)$ $A \times (B - C) = (A \times B) - (A \times C)$	Propiedad distributiva del producto respecto de la suma y la resta

Inversa de una matriz

Es importante señalar que la división entre matrices no está definida. Por lo que es necesario definir la inversa multiplicativa, también llamada *inversa de una matriz* que, multiplicada por la original, da como resultado la unitaria o identidad (I); o sea:

$$AA^{-1} = I \quad AB = I \quad B = A^{-1}$$

El producto de la matriz A , por su inversa, es la unitaria I

$B = A^{-1}$, siempre que la matriz inversa de A exista, ya que no todas las matrices tienen inversa. Las condiciones que debe cumplir una matriz para tener inversa son las siguientes:

1. Que sea matriz cuadrada ($A_{m \times m}$).
2. Que el determinante de A sea diferente de cero ($\det A \neq 0$); cuando esto ocurre, se dice que la matriz A es no singular.
3. Si $\det A = 0$, A es una matriz singular y no tiene inversa.

Procedimiento para formar la inversa de A por el método de cofactores:[†]

1. Reemplazar cada elemento de A por su cofactor:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$$

Este caso es para una matriz de orden 3×3 .

2. Dividir cada elemento de la nueva matriz, formada por sus cofactores, entre el determinante de A , ($|A|$).
3. Obtener su transpuesta, A^T
4. Comprobar que $AA^{-1} = I$

[†] El cofactor de un determinante es el determinante de orden menor que se forma al anular el renglón y la columna de cada uno de sus elementos, tomando en cuenta la relación de signos que se muestra.

EJEMPLO

16

Obtener la inversa de

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & -3 \\ 3 & -2 & 1 \end{bmatrix}$$

Paso 1. Calcular $|A|$ Como se tiene un renglón con dos ceros, entonces el determinante de A es simplemente:

$$|A| = -(-3) \begin{vmatrix} 1 & -1 \\ 3 & -2 \end{vmatrix} = 3(-2 + 3) = 3$$

considerando los signos de referencia

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$$

Paso 2. Obtener la matriz de cofactores:

$$\begin{bmatrix} \begin{vmatrix} 0 & -3 \\ -2 & 1 \end{vmatrix} - \begin{vmatrix} 0 & -3 \\ 3 & 1 \end{vmatrix} \begin{vmatrix} 0 & 0 \\ 3 & -2 \end{vmatrix} \\ - \begin{vmatrix} -1 & 2 \\ -2 & 1 \end{vmatrix} \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} - \begin{vmatrix} 1 & -1 \\ 3 & -2 \end{vmatrix} \\ \begin{vmatrix} -1 & 2 \\ 0 & 3 \end{vmatrix} - \begin{vmatrix} 1 & 2 \\ 0 & -3 \end{vmatrix} \begin{vmatrix} 1 & -1 \\ 0 & -0 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} -6 & -9 & 0 \\ -3 & -5 & -1 \\ 3 & 3 & 0 \end{bmatrix}$$

Paso 3. La matriz de cofactores se divide entre el determinante de A , o sea $|A| = 3$, y se obtiene

$$\begin{bmatrix} -2 & -3 & 0 \\ -1 & -\frac{5}{3} & -\frac{1}{3} \\ 1 & 1 & 0 \end{bmatrix}$$

Paso 4. Se obtiene la matriz transpuesta, es decir, la matriz inversa:

$$\begin{bmatrix} -2 & -1 & 1 \\ -3 & -\frac{5}{3} & 1 \\ 0 & -\frac{1}{3} & 0 \end{bmatrix}$$

Paso 5. Se comprueba:

$$AA^{-1} = I$$

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & -3 \\ 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} -2 & -1 & 1 \\ -3 & -\frac{5}{3} & 1 \\ 0 & -\frac{1}{3} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Anexo

Tablas

CONTENIDO

Tabla 1	Probabilidades binomiales acumulativas	644
Tabla 2	Probabilidades acumulativas de Poisson	649
Tabla 3	Áreas bajo la curva normal	651
Tabla 4	Valores críticos de t	653
Tabla 5	Valores críticos de ji cuadrada	654
Tabla 6	Puntos porcentuales de la distribución F	656
Tabla 7	Valores críticos de T para la prueba de suma de rango de Wilcoxon, $n_1 \leq n_2$	664
Tabla 8	Valores críticos de T para la prueba de rango con signo de Wilcoxon, $n = 5(1)50$	666
Tabla 9	Valores críticos del coeficiente de correlación de rango de Spearman para una prueba de una cola	667
Tabla 10	Números aleatorios	668
Tabla 11	Puntos porcentuales del rango de Student, $q_\alpha(k, g)$	670
Tabla 12	Porcentiles de la distribución normal estandarizada	674
Tabla 13	Valores críticos de Q (Tukey-Snedecor) para comparaciones múltiples	675
Tabla 14	Valores críticos para la prueba de Hartley (homogeneidad de varianzas)	676
Tabla 15	Valores críticos para la prueba de Cochran (homogeneidad de variantes)	677
Tabla 16	Valores de $ n $ para la prueba de MacNemar	678
Tabla 17	Valores críticos de A (prueba de Sandler)	679
Tabla 18	Tabla de Dunnet 1a para $p = 95\%$	680
Tabla 19	Tabla de Dunnet 1b para $p = 99\%$	681
Tabla 20	Tabla de Dunnet 2a para $p = 95\%$	682
Tabla 21	Tabla de Dunnet 2b para $p = 95\%$	683

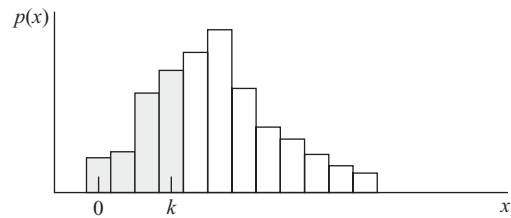


TABLA 1 Probabilidades binomiales acumulativas
Los valores tabulados son $P(x \leq k) = p(0) + p(1) + \dots + p(k)$.
(Los cálculos están redondeados al tercer lugar decimal.)

<i>n</i> = 2														
<i>p</i>														
<i>k</i>	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	<i>k</i>
0	.980	.902	.810	.640	.490	.360	.250	.160	.090	.040	.010	.002	.000	0
1	1.000	.998	.990	.960	.910	.840	.750	.640	.510	.360	.190	.098	.020	1
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	2

<i>n</i> = 3														
<i>p</i>														
<i>k</i>	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	<i>k</i>
0	.970	.857	.729	.512	.343	.216	.125	.064	.027	.008	.001	.000	.000	0
1	1.000	.993	.972	.896	.784	.648	.500	.352	.216	.104	.028	.007	.000	1
2	1.000	1.000	.999	.992	.973	.936	.875	.784	.657	.488	.271	.143	.030	2
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	3

<i>n</i> = 4														
<i>p</i>														
<i>k</i>	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	<i>k</i>
0	.961	.815	.656	.410	.240	.130	.062	.026	.008	.002	.000	.000	.000	0
1	.999	.986	.948	.819	.652	.475	.312	.179	.084	.027	.004	.000	.000	1
2	1.000	1.000	.996	.973	.916	.821	.688	.525	.348	.181	.052	.014	.001	2
3	1.000	1.000	1.000	.998	.992	.974	.938	.870	.760	.590	.344	.185	.039	3
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	4

TABLA 1 (continuación)

 $n = 5$

p														
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.951	.774	.590	.328	.168	.078	.031	.010	.002	.000	.000	.000	.000	0
1	.999	.977	.919	.737	.528	.337	.188	.087	.031	.007	.000	.000	.000	1
2	1.000	.999	.991	.942	.837	.683	.500	.317	.163	.058	.009	.001	.000	2
3	1.000	1.000	1.000	.993	.969	.913	.812	.663	.472	.263	.081	.023	.001	3
4	1.000	1.000	1.000	1.000	.998	.990	.969	.922	.832	.672	.410	.226	.049	4
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	5

 $n = 6$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.941	.735	.531	.262	.118	.047	.016	.004	.001	.000	.000	.000	.000	0
1	.999	.967	.886	.655	.420	.233	.109	.041	.011	.002	.000	.000	.000	1
2	1.000	.998	.984	.901	.744	.544	.344	.179	.070	.017	.001	.000	.000	2
3	1.000	1.000	.999	.983	.930	.821	.656	.456	.256	.099	.016	.002	.000	3
4	1.000	1.000	1.000	.998	.989	.959	.891	.767	.580	.345	.114	.033	.001	4
5	1.000	1.000	1.000	1.000	.999	.996	.984	.953	.882	.738	.469	.265	.059	5
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	6

 $n = 7$

Table 1: Cumulative probabilities $P(X \leq k)$ for $X \sim \text{Bin}(n=7, p)$														
	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.932	.698	.478	.210	.082	.028	.008	.002	.000	.000	.000	.000	.000	0
1	.998	.956	.850	.577	.329	.159	.062	.019	.004	.000	.000	.000	.000	1
2	1.000	.996	.974	.852	.647	.420	.227	.096	.029	.005	.000	.000	.000	2
3	1.000	1.000	.997	.967	.874	.710	.500	.290	.126	.033	.003	.000	.000	3
4	1.000	1.000	1.000	.995	.971	.904	.773	.580	.353	.148	.026	.004	.000	4
5	1.000	1.000	1.000	1.000	.996	.981	.938	.841	.671	.423	.150	.044	.002	5
6	1.000	1.000	1.000	1.000	1.000	.998	.992	.972	.918	.790	.522	.302	.068	6
7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	7

 $n = 8$

p														
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.923	.663	.430	.168	.058	.017	.004	.001	.000	.000	.000	.000	.000	0
1	.997	.943	.813	.503	.255	.106	.035	.009	.001	.000	.000	.000	.000	1
2	1.000	.994	.962	.797	.552	.315	.145	.050	.011	.001	.000	.000	.000	2
3	1.000	1.000	.995	.944	.806	.594	.363	.174	.058	.010	.000	.000	.000	3
4	1.000	1.000	1.000	.990	.942	.826	.637	.406	.194	.056	.005	.000	.000	4
5	1.000	1.000	1.000	.999	.989	.950	.855	.685	.448	.203	.038	.006	.000	5
6	1.000	1.000	1.000	1.000	.999	.991	.965	.894	.745	.497	.187	.057	.003	6
7	1.000	1.000	1.000	1.000	1.000	.999	.996	.983	.942	.832	.570	.337	.077	7
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	8

TABLA 1 (continuación) $n = 9$

Table 1: Cumulative probabilities $P(X \leq k)$ for $X \sim \text{Bin}(n=10, p)$														
	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.914	.630	.387	.134	.040	.010	.002	.000	.000	.000	.000	.000	.000	0
1	.997	.929	.775	.436	.196	.071	.020	.004	.000	.000	.000	.000	.000	1
2	1.000	.992	.947	.738	.463	.232	.090	.025	.004	.000	.000	.000	.000	2
3	1.000	.999	.992	.914	.730	.483	.254	.099	.025	.003	.000	.000	.000	3
4	1.000	1.000	.999	.980	.901	.733	.500	.267	.099	.020	.001	.000	.000	4
5	1.000	1.000	1.000	.997	.975	.901	.746	.517	.270	.086	.008	.001	.000	5
6	1.000	1.000	1.000	1.000	.996	.975	.910	.768	.537	.262	.053	.008	.000	6
7	1.000	1.000	1.000	1.000	1.000	.996	.980	.929	.804	.564	.225	.071	.003	7
8	1.000	1.000	1.000	1.000	1.000	1.000	.998	.990	.960	.866	.613	.370	.086	8
9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	9

 $n = 10$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.904	.599	.349	.107	.028	.006	.001	.000	.000	.000	.000	.000	.000	0
1	.996	.914	.736	.376	.149	.046	.011	.002	.000	.000	.000	.000	.000	1
2	1.000	.988	.930	.678	.383	.167	.055	.012	.002	.000	.000	.000	.000	2
3	1.000	.999	.987	.879	.650	.382	.172	.055	.011	.001	.000	.000	.000	3
4	1.000	1.000	.998	.967	.850	.633	.377	.166	.047	.006	.000	.000	.000	4
5	1.000	1.000	1.000	.994	.953	.834	.623	.367	.150	.033	.002	.000	.000	5
6	1.000	1.000	1.000	.999	.989	.945	.828	.618	.350	.121	.013	.001	.000	6
7	1.000	1.000	1.000	1.000	.998	.988	.945	.833	.617	.322	.070	.012	.000	7
8	1.000	1.000	1.000	1.000	1.000	.998	.989	.954	.851	.624	.264	.086	.004	8
9	1.000	1.000	1.000	1.000	1.000	1.000	.999	.994	.972	.893	.651	.401	.096	9
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	10

 $n = 11$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.895	.569	.314	.086	.020	.004	.000	.000	.000	.000	.000	.000	.000	0
1	.995	.898	.697	.322	.113	.030	.006	.001	.000	.000	.000	.000	.000	1
2	1.000	.985	.910	.617	.313	.119	.033	.006	.001	.000	.000	.000	.000	2
3	1.000	.998	.981	.839	.570	.296	.113	.029	.004	.000	.000	.000	.000	3
4	1.000	1.000	.997	.950	.790	.533	.274	.099	.022	.002	.000	.000	.000	4
5	1.000	1.000	1.000	.988	.922	.754	.500	.246	.078	.012	.000	.000	.000	5
6	1.000	1.000	1.000	.998	.978	.901	.726	.467	.210	.050	.003	.000	.000	6
7	1.000	1.000	1.000	1.000	.996	.971	.887	.704	.430	.161	.019	.002	.000	7
8	1.000	1.000	1.000	1.000	.999	.994	.967	.881	.687	.383	.090	.015	.000	8
9	1.000	1.000	1.000	1.000	1.000	.999	.994	.970	.887	.678	.303	.102	.005	9
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.980	.914	.686	.431	.105	10
11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	11

TABLA 1 (continuación)

 $n = 12$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.886	.540	.282	.069	.014	.002	.000	.000	.000	.000	.000	.000	.000	0
1	.994	.882	.659	.275	.085	.020	.003	.000	.000	.000	.000	.000	.000	1
2	1.000	.980	.889	.558	.253	.083	.019	.003	.000	.000	.000	.000	.000	2
3	1.000	.998	.974	.795	.493	.225	.073	.015	.002	.000	.000	.000	.000	3
4	1.000	1.000	.996	.927	.724	.438	.194	.057	.009	.001	.000	.000	.000	4
5	1.000	1.000	.999	.981	.882	.665	.387	.158	.039	.004	.000	.000	.000	5
6	1.000	1.000	1.000	.996	.961	.842	.613	.335	.118	.019	.001	.000	.000	6
7	1.000	1.000	1.000	.999	.991	.943	.806	.562	.276	.073	.004	.000	.000	7
8	1.000	1.000	1.000	1.000	.998	.985	.927	.775	.507	.205	.026	.002	.000	8
9	1.000	1.000	1.000	1.000	1.000	.997	.981	.917	.747	.442	.111	.020	.000	9
10	1.000	1.000	1.000	1.000	1.000	1.000	.997	.980	.915	.725	.341	.118	.006	10
11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.986	.931	.718	.460	.114	11
12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	12

 $n = 15$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.860	.463	.206	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000	0
1	.990	.829	.549	.167	.035	.005	.000	.000	.000	.000	.000	.000	.000	1
2	1.000	.964	.816	.398	.127	.027	.004	.000	.000	.000	.000	.000	.000	2
3	1.000	.995	.944	.648	.297	.091	.018	.002	.000	.000	.000	.000	.000	3
4	1.000	.999	.987	.836	.515	.217	.059	.009	.001	.000	.000	.000	.000	4
5	1.000	1.000	.998	.939	.722	.403	.151	.034	.004	.000	.000	.000	.000	5
6	1.000	1.000	1.000	.982	.869	.610	.304	.095	.015	.001	.000	.000	.000	6
7	1.000	1.000	1.000	.996	.950	.787	.500	.213	.050	.004	.000	.000	.000	7
8	1.000	1.000	1.000	.999	.985	.905	.696	.390	.131	.018	.000	.000	.000	8
9	1.000	1.000	1.000	1.000	.996	.966	.849	.597	.278	.061	.002	.000	.000	9
10	1.000	1.000	1.000	1.000	.999	.991	.941	.783	.485	.164	.013	.001	.000	10
11	1.000	1.000	1.000	1.000	1.000	.998	.982	.909	.703	.352	.056	.005	.000	11
12	1.000	1.000	1.000	1.000	1.000	1.000	.996	.973	.873	.602	.184	.036	.000	12
13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.833	.451	.171	.010	13
14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.794	.537	.140	14
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	15

TABLA 1 (continuación) $n = 20$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.818	.358	.122	.012	.001	.000	.000	.000	.000	.000	.000	.000	.000	0
1	.983	.736	.392	.069	.008	.001	.000	.000	.000	.000	.000	.000	.000	1
2	.999	.925	.677	.206	.035	.004	.000	.000	.000	.000	.000	.000	.000	2
3	1.000	.984	.867	.411	.107	.016	.001	.000	.000	.000	.000	.000	.000	3
4	1.000	.997	.957	.630	.238	.051	.006	.000	.000	.000	.000	.000	.000	4
5	1.000	1.000	.989	.804	.416	.126	.021	.002	.000	.000	.000	.000	.000	5
6	1.000	1.000	.998	.913	.608	.250	.058	.006	.000	.000	.000	.000	.000	6
7	1.000	1.000	1.000	.968	.772	.416	.132	.021	.001	.000	.000	.000	.000	7
8	1.000	1.000	1.000	.990	.887	.596	.252	.057	.005	.000	.000	.000	.000	8
9	1.000	1.000	1.000	.997	.952	.755	.412	.128	.017	.001	.000	.000	.000	9
10	1.000	1.000	1.000	.999	.983	.872	.588	.245	.048	.003	.000	.000	.000	10
11	1.000	1.000	1.000	1.000	.995	.943	.748	.404	.113	.010	.000	.000	.000	11
12	1.000	1.000	1.000	1.000	.999	.979	.868	.584	.228	.032	.000	.000	.000	12
13	1.000	1.000	1.000	1.000	1.000	.994	.942	.750	.392	.087	.002	.000	.000	13
14	1.000	1.000	1.000	1.000	1.000	.998	.979	.874	.584	.196	.011	.000	.000	14
15	1.000	1.000	1.000	1.000	1.000	1.000	.994	.949	.762	.370	.043	.003	.000	15
16	1.000	1.000	1.000	1.000	1.000	1.000	.999	.984	.893	.589	.133	.016	.000	16
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.965	.794	.323	.075	.001	17
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.992	.931	.608	.264	.017	18
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.988	.878	.642	.182	19
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	20

 $n = 25$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.778	.277	.072	.004	.000	.000	.000	.000	.000	.000	.000	.000	.000	0
1	.974	.642	.271	.027	.002	.000	.000	.000	.000	.000	.000	.000	.000	1
2	.998	.873	.537	.098	.009	.000	.000	.000	.000	.000	.000	.000	.000	2
3	1.000	.966	.764	.234	.033	.002	.000	.000	.000	.000	.000	.000	.000	3
4	1.000	.993	.902	.421	.090	.009	.000	.000	.000	.000	.000	.000	.000	4
5	1.000	.999	.967	.617	.193	.029	.002	.000	.000	.000	.000	.000	.000	5
6	1.000	1.000	.991	.780	.341	.074	.007	.000	.000	.000	.000	.000	.000	6
7	1.000	1.000	.998	.891	.512	.154	.022	.001	.000	.000	.000	.000	.000	7
8	1.000	1.000	1.000	.953	.677	.274	.054	.004	.000	.000	.000	.000	.000	8
9	1.000	1.000	1.000	.983	.811	.425	.115	.013	.000	.000	.000	.000	.000	9
10	1.000	1.000	1.000	.994	.902	.586	.212	.034	.002	.000	.000	.000	.000	10
11	1.000	1.000	1.000	.998	.956	.732	.345	.078	.006	.000	.000	.000	.000	11
12	1.000	1.000	1.000	1.000	.983	.846	.500	.154	.017	.000	.000	.000	.000	12
13	1.000	1.000	1.000	1.000	.994	.922	.655	.268	.044	.002	.000	.000	.000	13
14	1.000	1.000	1.000	1.000	.998	.966	.788	.414	.098	.006	.000	.000	.000	14
15	1.000	1.000	1.000	1.000	1.000	.987	.885	.575	.189	.017	.000	.000	.000	15
16	1.000	1.000	1.000	1.000	1.000	.996	.946	.726	.323	.047	.000	.000	.000	16
17	1.000	1.000	1.000	1.000	1.000	.999	.978	.846	.488	.109	.002	.000	.000	17
18	1.000	1.000	1.000	1.000	1.000	1.000	.993	.926	.659	.220	.009	.000	.000	18
19	1.000	1.000	1.000	1.000	1.000	1.000	.998	.971	.807	.383	.033	.001	.000	19
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.910	.579	.098	.007	.000	20
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.967	.766	.236	.034	.000	21
22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.902	.463	.127	.002	22
23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.973	.729	.358	.026	23
24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.928	.723	.222	24
25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	25

TABLA 2 Probabilidades acumulativas de Poisson

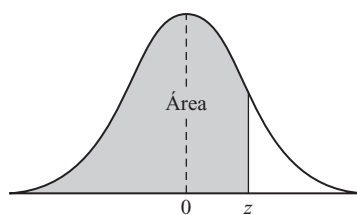
Los valores tabulados son $P(x \leq k) = p(0) + p(1) + \dots + p(k)$.
(Los cálculos están redondeados al tercer lugar decimal.)

k	μ										
	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0	1.5
0	.905	.819	.741	.670	.607	.549	.497	.449	.407	.368	.223
1	.995	.982	.963	.938	.910	.878	.844	.809	.772	.736	.558
2	1.000	.999	.996	.992	.986	.977	.966	.953	.937	.920	.809
3		1.000	1.000	.999	.998	.997	.994	.991	.987	.981	.934
4				1.000	1.000	1.000	.999	.999	.998	.996	.981
5							1.000	1.000	1.000	.999	.996
6										1.000	.999
7											1.000

k	μ										
	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
0	.135	.082	.050	.033	.018	.011	.007	.004	.003	.002	.001
1	.406	.287	.199	.136	.092	.061	.040	.027	.017	.011	.007
2	.677	.544	.423	.321	.238	.174	.125	.088	.062	.043	.030
3	.857	.758	.647	.537	.433	.342	.265	.202	.151	.112	.082
4	.947	.891	.815	.725	.629	.532	.440	.358	.285	.224	.173
5	.983	.958	.916	.858	.785	.703	.616	.529	.446	.369	.301
6	.995	.986	.966	.935	.889	.831	.762	.686	.606	.563	.450
7	.999	.996	.988	.973	.949	.913	.867	.809	.744	.673	.599
8	1.000	.999	.996	.990	.979	.960	.932	.894	.847	.792	.729
9		1.000	.999	.997	.992	.983	.968	.946	.916	.877	.830
10			1.000	.999	.997	.993	.986	.975	.957	.933	.901
11				1.000	.999	.998	.995	.989	.980	.966	.947
12					1.000	.999	.998	.996	.991	.984	.973
13						1.000	.999	.998	.996	.993	.987
14							1.000	.999	.999	.997	.994
15								1.000	.999	.999	.998
16									1.000	1.000	.999
17											1.000

TABLA 2 (continuación)

<i>k</i>	μ								
	7.5	8.0	8.5	9.0	9.5	10.0	12.0	15.0	20.0
0	.001	.000	.000	.000	.000	.000	.000	.000	.000
1	.005	.003	.002	.001	.001	.000	.000	.000	.000
2	.020	.014	.009	.006	.004	.003	.001	.000	.000
3	.059	.042	.030	.021	.015	.010	.002	.000	.000
4	.132	.100	.074	.055	.040	.029	.008	.001	.000
5	.241	.191	.150	.116	.089	.067	.020	.003	.000
6	.378	.313	.256	.207	.165	.130	.046	.008	.000
7	.525	.453	.386	.324	.269	.220	.090	.018	.001
8	.662	.593	.523	.456	.392	.333	.155	.037	.002
9	.776	.717	.653	.587	.522	.458	.242	.070	.005
10	.862	.816	.763	.706	.645	.583	.347	.118	.011
11	.921	.888	.849	.803	.752	.697	.462	.185	.021
12	.957	.936	.909	.876	.836	.792	.576	.268	.039
13	.978	.966	.949	.926	.898	.864	.682	.363	.066
14	.990	.983	.973	.959	.940	.917	.772	.466	.105
15	.995	.992	.986	.978	.967	.951	.844	.568	.157
16	.998	.996	.993	.989	.982	.973	.899	.664	.221
17	.999	.998	.997	.995	.991	.986	.937	.749	.297
18	1.000	.999	.999	.998	.996	.993	.963	.819	.381
19		1.000	.999	.999	.998	.997	.979	.875	.470
20			1.000	1.000	.999	.998	.988	.917	.559
21					1.000	.999	.994	.947	.644
22						1.000	.997	.967	.721
23							.999	.981	.787
24							.999	.989	.843
25							1.000	.994	.888
26								.997	.922
27								.998	.948
28								.999	.966
29								1.000	.978
30									.987
31									.992
32									.995
33									.997
34									.999
35									.999
36									1.000

**TABLA 3 Áreas bajo la curva normal**

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

TABLA 3 (continuación)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

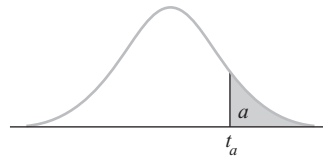


TABLA 4
Valores críticos
de t

gl	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	df
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10
11	1.363	1.796	2.201	2.718	3.106	11
12	1.356	1.782	2.179	2.681	3.055	12
13	1.350	1.771	2.160	2.650	3.012	13
14	1.345	1.761	2.145	2.624	2.977	14
15	1.341	1.753	2.131	2.602	2.947	15
16	1.337	1.746	2.120	2.583	2.921	16
17	1.333	1.740	2.110	2.567	2.898	17
18	1.330	1.734	2.101	2.552	2.878	18
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20
21	1.323	1.721	2.080	2.518	2.831	21
22	1.321	1.717	2.074	2.508	2.819	22
23	1.319	1.714	2.069	2.500	2.807	23
24	1.318	1.711	2.064	2.492	2.797	24
25	1.316	1.708	2.060	2.485	2.787	25
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
∞	1.282	1.645	1.960	2.326	2.576	∞

FUENTE: De "Table of Percentage Points of the t-Distribution", *Biometrika*, 32 (1941): 300. Reproducida con permiso de los fideicomisarios de *Biometrika*.

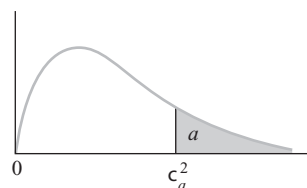


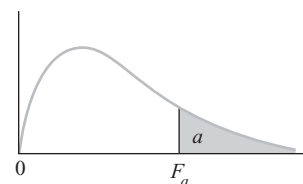
TABLA 5
Valores críticos
de ji cuadrada

gl	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$
1	.0000393	.0001571	.0009821	.0039321	.0157908
2	.0100251	.0201007	.0506356	.102587	.210720
3	.0717212	.114832	.215795	.351846	.584375
4	.206990	.297110	.484419	.710721	1.063623
5	.411740	.554300	.831211	1.145476	1.61031
6	.675727	.872085	1.237347	1.63539	2.20413
7	.989265	1.239043	1.68987	2.16735	2.83311
8	1.344419	1.646482	2.17973	2.73264	3.48954
9	1.734926	2.087912	2.70039	3.32511	4.16816
10	2.15585	2.55821	3.24697	3.94030	4.86518
11	2.60321	3.05347	3.81575	4.57481	5.57779
12	3.07382	3.57056	4.40379	5.22603	6.30380
13	3.56503	4.10691	5.00874	5.89186	7.04150
14	4.07468	4.66043	5.62872	6.57063	7.78953
15	4.60094	5.22935	6.26214	7.26094	8.54675
16	5.14224	5.81221	6.90766	7.96164	9.31223
17	5.69724	6.40776	7.56418	8.67176	10.0852
18	6.26481	7.01491	8.23075	9.39046	10.8649
19	6.84398	7.63273	8.90655	10.1170	11.6509
20	7.43386	8.26040	9.59083	10.8508	12.4426
21	8.03366	8.89720	10.28293	11.5913	13.2396
22	8.64272	9.54249	10.9823	12.3380	14.0415
23	9.26042	10.19567	11.6885	13.0905	14.8479
24	9.88623	10.8564	12.4011	13.8484	15.6587
25	10.5197	11.5240	13.1197	14.6114	16.4734
26	11.1603	12.1981	13.8439	15.3791	17.2919
27	11.8076	12.8786	14.5733	16.1513	18.1138
28	12.4613	13.5648	15.3079	16.9279	18.9392
29	13.1211	14.2565	16.0471	17.7083	19.7677
30	13.7867	14.9535	16.7908	18.4926	20.5992
40	20.7065	22.1643	24.4331	26.5093	29.0505
50	27.9907	29.7067	32.3574	34.7642	37.6886
60	35.5346	37.4848	40.4817	43.1879	46.4589
70	43.2752	45.4418	48.7576	51.7393	55.3290
80	51.1720	53.5400	57.1532	60.3915	64.2778
90	59.1963	61.7541	65.6466	69.1260	73.2912
100	67.3276	70.0648	74.2219	77.9295	82.3581

FUENTE: De "Tables of the Percentage Points of the χ^2 Distribution", *Biometrika Tables for Statisticians*, vol. 1, 3a. ed. (1966).
Reproducida con permiso de los fideicomisarios de *Biometrika*.

TABLA 5
(continuación)

$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$	gl
2.70554	3.84146	5.02389	6.63490	7.87944	1
4.60517	5.99147	7.37776	9.21034	10.5966	2
6.25139	7.81473	9.34840	11.3449	12.8381	3
7.77944	9.48773	11.1433	13.2767	14.8602	4
9.23635	11.0705	12.8325	15.0863	16.7496	5
10.6446	12.5916	14.4494	16.8119	18.5476	6
12.0170	14.0671	16.0128	18.4753	20.2777	7
13.3616	15.5073	17.5346	20.0902	21.9550	8
14.6837	16.9190	19.0228	21.6660	23.5893	9
15.9871	18.3070	20.4831	23.2093	25.1882	10
17.2750	19.6751	21.9200	24.7250	26.7569	11
18.5494	21.0261	23.3367	26.2170	28.2995	12
19.8119	22.3621	24.7356	27.6883	29.8194	13
21.0642	23.6848	26.1190	29.1413	31.3193	14
22.3072	24.9958	27.4884	30.5779	32.8013	15
23.5418	26.2962	28.8485	31.9999	34.2672	16
24.7690	27.8571	30.1910	33.4087	35.7185	17
25.9894	28.8693	31.5264	34.8053	37.1564	18
27.2036	30.1435	32.8523	36.1908	38.5822	19
28.4120	31.4104	34.1696	37.5662	39.9968	20
29.6151	32.6705	35.4789	38.9321	41.4010	21
30.8133	33.9244	36.7807	40.2894	42.7956	22
32.0069	35.1725	38.0757	41.6384	44.1813	23
33.1963	36.4151	39.3641	42.9798	45.5585	24
34.3816	37.6525	40.6465	44.3141	46.9278	25
35.5631	38.8852	41.9232	45.6417	48.2899	26
36.7412	40.1133	43.1944	46.9630	49.6449	27
37.9159	41.3372	44.4607	48.2782	50.9933	28
39.0875	42.5569	45.7222	49.5879	52.3356	29
40.2560	43.7729	46.9792	50.8922	53.6720	30
51.8050	55.7585	59.3417	63.6907	66.7659	40
63.1671	67.5048	71.4202	76.1539	79.4900	50
74.3970	79.0819	83.2976	88.3794	91.9517	60
85.5271	90.5312	95.0231	100.425	104.215	70
96.5782	101.879	106.629	112.329	116.321	80
107.565	113.145	118.136	124.116	128.299	90
118.498	124.342	129.561	135.807	140.169	100

**TABLA 6 Puntos porcentuales de la distribución F**

g_2	a	g_1								
		1	2	3	4	5	6	7	8	9
1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	.050	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
	.025	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3
	.010	4052	4999.5	5403	5625	5764	5859	5928	5982	6022
	.005	16211	20000	21615	22500	23056	23437	23715	23925	24091
2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
	.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	.005	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.4
3	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
	.010	34.12	30.82	29.46	28.71	28.24	27.91	27.64	27.49	27.35
	.005	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88
4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
	.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	.005	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14
5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	.005	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77
6	.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
	.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	.005	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39
7	.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
	.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	.005	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51
8	.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	.025	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
	.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	.005	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34
9	.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	.025	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
	.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	.005	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54

FUENTE: Parte de "Tables of Percentage Points of the Inverted Beta (F) Distribution", *Biometrika*, vol. 33 (1943); por M. Merrington y C.M. Thompson y de la tabla 18 de *Biometrika Tables for Statisticians*, vol. 1, Cambridge University Press, 1954, editado por E.S. Pearson y H.O. Hartley. Reproducida con permiso de los autores, editores y los fideicomisarios de *Biometrika*.

TABLA 6 (continuación)

$g/1$										a	$g/2$
10	12	15	20	24	30	40	60	120	∞		
60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33	.100	1
241.9	243.9	245.9	248.0	249.1	250.1	251.2	252.2	253.3	254.3	.050	
968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018	.025	
6056	6106	6157	6209	6235	6261	6287	6313	6339	6366	.010	
24224	24426	24630	24836	24940	25044	25148	25253	25359	25465	.005	
9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49	.100	2
19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50	.050	
39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50	.025	
99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50	.010	
199.4	199.4	199.4	199.4	199.5	199.5	199.5	199.5	199.5	199.5	.005	
5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13	.100	3
8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53	.050	
14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90	.025	
27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13	.010	
43.69	43.39	43.08	42.78	42.62	42.47	42.31	42.15	41.99	41.83	.005	
3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76	.100	4
5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63	.050	
8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26	.025	
14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46	.010	
20.97	20.70	20.44	20.17	20.03	19.89	19.75	19.61	19.47	19.32	.005	
3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10	.100	5
4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36	.050	
6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02	.025	
10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02	.010	
13.62	13.38	13.15	12.90	12.78	12.66	12.53	12.40	12.27	12.14	.005	
2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72	.100	6
4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67	.050	
5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85	.025	
7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88	.010	
10.25	10.03	9.81	9.59	9.47	9.36	9.24	9.12	9.00	8.88	.005	
2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47	.100	7
3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23	.050	
4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14	.025	
6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65	.010	
8.38	8.18	7.97	7.75	7.65	7.53	7.42	7.31	7.19	7.08	.005	
2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29	.100	8
3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93	.050	
4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67	.025	
5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86	.010	
7.21	7.01	6.81	6.61	6.50	6.40	6.29	6.18	6.06	5.95	.005	
2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16	.100	9
3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71	.050	
3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33	.025	
5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31	.010	
6.42	6.23	6.03	5.83	5.73	5.62	5.52	5.41	5.30	5.19	.005	

TABLA 6 (continuación)

$g/2$	$g/1$									
	a	1	2	3	4	5	6	7	8	9
10	.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
	.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	.005	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97
11	.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	.025	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59
	.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	.005	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54
12	.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	.025	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
	.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	.005	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20
13	.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	.025	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31
	.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	.005	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94
14	.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	.025	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21
	.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	.005	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72
15	.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	.025	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
	.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	.005	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54
16	.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	.025	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05
	.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	.005	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38
17	.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	.025	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98
	.010	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
	.005	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25
18	.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	.050	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	.025	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93
	.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	.005	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.14
19	.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	.050	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	.025	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88
	.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	.005	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04
20	.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	.025	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84
	.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	.005	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96

TABLA 6 (continuación)

$g/1$										a	$g/2$
10	12	15	20	24	30	40	60	120	∞		
2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06	.100	10
2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54	.050	
3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08	.025	
4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91	.010	
5.85	5.66	5.47	5.27	5.17	5.07	4.97	4.86	4.75	4.64	.005	
2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97	.100	11
2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40	.050	
3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88	.025	
4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60	.010	
5.42	5.24	5.05	4.86	4.76	4.65	4.55	4.44	4.34	4.23	.005	
2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90	.100	12
2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30	.050	
3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72	.025	
4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36	.010	
5.09	4.91	4.72	4.53	4.43	4.33	4.23	4.12	4.01	3.90	.005	
2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85	.100	13
2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21	.050	
3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60	.025	
4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17	.010	
4.82	4.64	4.46	4.27	4.17	4.07	3.97	3.87	3.76	3.65	.005	
2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80	.100	14
2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13	.050	
3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49	.025	
3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00	.010	
4.60	4.43	4.25	4.06	3.96	3.86	3.76	3.66	3.55	3.44	.005	
2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76	.100	15
2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07	.050	
3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40	.025	
3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87	.010	
4.42	4.25	4.07	3.88	3.79	3.69	3.58	3.48	3.37	3.26	.005	
2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72	.100	16
2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01	.050	
2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32	.025	
3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75	.010	
4.27	4.10	3.92	3.73	3.64	3.54	3.44	3.33	3.22	3.11	.005	
2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69	.100	17
2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96	.050	
2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25	.025	
3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65	.010	
4.14	3.97	3.79	3.61	3.51	3.41	3.31	3.21	3.10	2.98	.005	
1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66	.100	18
2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92	.050	
2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19	.025	
3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57	.010	
4.03	3.86	3.68	3.50	3.40	3.30	3.20	3.10	2.99	2.87	.005	
1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63	.100	19
2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88	.050	
2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13	.025	
3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49	.010	
3.93	3.76	3.59	3.40	3.31	3.21	3.11	3.00	2.89	2.78	.005	
1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61	.100	20
2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84	.050	
2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09	.025	
3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42	.010	
3.85	3.68	3.50	3.32	3.22	3.12	3.02	2.92	2.81	2.69	.005	

TABLA 6 (continuación)

$g/2$	$g/1$									
	a	1	2	3	4	5	6	7	8	9
21	.100	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
	.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	.025	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80
	.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	.005	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.88
22	.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	.025	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76
	.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	.005	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81
23	.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	.025	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73
	.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	.005	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75
24	.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	.025	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
	.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	.005	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69
25	.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	.025	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68
	.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	.005	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64
26	.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	.025	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65
	.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	.005	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60
27	.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	.025	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63
	.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	.005	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56
28	.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	.025	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61
	.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	.005	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.52
29	.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	.025	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59
	.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	.005	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48
30	.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.005	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45

TABLA 6 (continuación)

g_{l_1}										a	g_{l_2}
10	12	15	20	24	30	40	60	120	∞		
1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59	.100	21
2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81	.050	
2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04	.025	
3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36	.010	
3.77	3.60	3.43	3.24	3.15	3.05	2.95	2.84	2.73	2.61	.005	
1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57	.100	22
2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78	.050	
2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00	.025	
3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31	.010	
3.70	3.54	3.36	3.18	3.08	2.98	2.88	2.77	2.66	2.55	.005	
1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55	.100	23
2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76	.050	
2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97	.025	
3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26	.010	
3.64	3.47	3.30	3.12	3.02	2.92	2.82	2.71	2.60	2.48	.005	
1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53	.100	24
2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73	.050	
2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94	.025	
3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21	.010	
3.59	3.42	3.25	3.06	2.97	2.87	2.77	2.66	2.55	2.43	.005	
1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52	.100	25
2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71	.050	
2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91	.025	
3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17	.010	
3.54	3.37	3.20	3.01	2.92	2.82	2.72	2.61	2.50	2.38	.005	
1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50	.100	26
2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69	.050	
2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88	.025	
3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13	.010	
3.49	3.33	3.15	2.97	2.87	2.77	2.67	2.56	2.45	2.33	.005	
1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49	.100	27
2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67	.050	
2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85	.025	
3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10	.010	
3.45	3.28	3.11	2.93	2.83	2.73	2.63	2.52	2.41	2.29	.005	
1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48	.100	28
2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65	.050	
2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83	.025	
3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06	.010	
3.41	3.25	3.07	2.89	2.79	2.69	2.59	2.48	2.37	2.25	.005	
1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47	.100	29
2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64	.050	
2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81	.025	
3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03	.010	
3.38	3.21	3.04	2.86	2.76	2.66	2.56	2.45	2.33	2.21	.005	
1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46	.100	30
2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62	.050	
2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79	.025	
2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01	.010	
3.34	3.18	3.01	2.82	2.73	2.63	2.52	2.42	2.30	2.18	.005	

TABLA 6 (continuación)

$g/2$	$g/1$									
	a	1	2	3	4	5	6	7	8	9
40	.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	.025	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45
	.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	.005	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22
60	.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.025	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.005	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01
120	.100	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
	.050	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
	.025	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22
	.010	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	.005	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81
∞	.100	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63
	.050	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.63
	.025	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11
	.010	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41
	.005	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62

TABLA 6 (continuación)

$g/1$										a	$g/2$
10	12	15	20	24	30	40	60	120	∞		
1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38	.100	40
2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51	.050	
2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64	.025	
2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80	.010	
3.12	2.95	2.78	2.60	2.50	2.40	2.30	2.18	2.06	1.93	.005	
1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29	.100	60
1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39	.050	
2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48	.025	
2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60	.010	
2.90	2.74	2.57	2.39	2.29	2.19	2.08	1.96	1.83	1.69	.005	
1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19	.100	120
1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25	.050	
2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31	.025	
2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38	.010	
2.71	2.54	2.37	2.19	2.09	1.98	1.87	1.75	1.61	1.43	.005	
1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00	.100	∞
1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00	.050	
2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00	.025	
2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00	.010	
2.52	2.36	2.19	2.00	1.90	1.79	1.67	1.53	1.36	1.00	.005	

TABLA 7 Valores críticos de T para la prueba de suma de rango de Wilcoxon, $n_1 \leq n_2$

TABLA 7a)															
Valores		n_1													
críticos de cola	n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15
izquierda a 5%	3	—	6												
	4	—	6	11											
	5	3	7	12	19										
	6	3	8	13	20	28									
	7	3	8	14	21	29	39								
	8	4	9	15	23	31	41	51							
	9	4	10	16	24	33	43	54	66						
	10	4	10	17	26	35	45	56	69	82					
	11	4	11	18	27	37	47	59	72	86	100				
	12	5	11	19	28	38	49	62	75	89	104	120			
	13	5	12	20	30	40	52	64	78	92	108	125	142		
	14	6	13	21	31	42	54	67	81	96	112	129	147	166	
	15	6	13	22	33	44	56	69	84	99	116	133	152	171	192

TABLA 7b)															
Valores		n_1													
críticos de cola	n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15
izquierda a 2.5%	4	—	—	10											
	5	—	6	11	17										
	6	—	7	12	18	26									
	7	—	7	13	20	27	36								
	8	3	8	14	21	29	38	49							
	9	3	8	14	22	31	40	51	62						
	10	3	9	15	23	32	42	53	65	78					
	11	3	9	16	24	34	44	55	68	81	96				
	12	4	10	17	26	35	46	58	71	84	99	115			
	13	4	10	18	27	37	48	60	73	88	103	119	136		
	14	4	11	19	28	38	50	62	76	91	106	123	141	160	
	15	4	11	20	29	40	52	65	79	94	110	127	145	164	184

Fuente: Datos de "An Extended Table of Critical Values for the Mann-Whitney (Wilcoxon) Two-Sample Statistic" por Roy C. Milton, pp. 925-934 en *Journal of the American Statistical Association*, vol. 59, núm. 307, septiembre de 1964. Reimpresión con permiso de *Journal of the American Statistical Association*. Copyright 1964 por la American Statistical Association. Todos los derechos reservados.

TABLA 7c)
Valores
críticos de cola
izquierda a 1%

n_2	n_1													
	2	3	4	5	6	7	8	9	10	11	12	13	14	15
3	—	—												
4	—	—	—											
5	—	—	10	16										
6	—	—	11	17	24									
7	—	6	11	18	25	34								
8	—	6	12	19	27	35	45							
9	—	7	13	20	28	37	47	59						
10	—	7	13	21	29	39	49	61	74					
11	—	7	14	22	30	40	51	63	77	91				
12	—	8	15	23	32	42	53	66	79	94	109			
13	3	8	15	24	33	44	56	68	82	97	113	130		
14	3	8	16	25	34	45	58	71	85	100	116	134	152	
15	3	9	17	26	36	47	60	73	88	103	120	138	156	176

TABLA 7d)
Valores
críticos de cola
izquierda a .5%

	n_1													
n_2	3	4	5	6	7	8	9	10	11	12	13	14	15	
3	—													
4	—	—												
5	—	—	15											
6	—	10	16	23										
7	—	10	16	24	32									
8	—	11	17	25	34	42								
9	6	11	18	26	35	45	56							
10	6	12	19	27	37	47	58	71						
11	6	12	20	28	38	49	61	73	87					
12	7	13	21	30	40	51	63	76	90	105				
13	7	13	22	31	41	53	65	79	93	109	125			
14	7	14	22	32	43	54	67	81	96	112	129	147		
15	8	15	23	33	44	56	69	84	99	115	133	151	171	

TABLA 8
Valores críticos
de T para la
prueba de rango
con signo de
Wilcoxon,
 $n = 5(1)50$

Una cola	Dos colas	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
$\alpha = .050$	$\alpha = .10$	1	2	4	6	8	11
$\alpha = .025$	$\alpha = .05$		1	2	4	6	8
$\alpha = .010$	$\alpha = .02$			0	2	3	5
$\alpha = .005$	$\alpha = .01$				0	2	3
Una cola	Dos colas	$n = 11$	$n = 12$	$n = 13$	$n = 14$	$n = 15$	$n = 16$
$\alpha = .050$	$\alpha = .10$	14	17	21	26	30	36
$\alpha = .025$	$\alpha = .05$	11	14	17	21	25	30
$\alpha = .010$	$\alpha = .02$	7	10	13	16	20	24
$\alpha = .005$	$\alpha = .01$	5	7	10	13	16	19
Una cola	Dos colas	$n = 17$	$n = 18$	$n = 19$	$n = 20$	$n = 21$	$n = 22$
$\alpha = .050$	$\alpha = .10$	41	47	54	60	68	75
$\alpha = .025$	$\alpha = .05$	35	40	46	52	59	66
$\alpha = .010$	$\alpha = .02$	28	33	38	43	49	56
$\alpha = .005$	$\alpha = .01$	23	28	32	37	43	49
Una cola	Dos colas	$n = 23$	$n = 24$	$n = 25$	$n = 26$	$n = 27$	$n = 28$
$\alpha = .050$	$\alpha = .10$	83	92	101	110	120	130
$\alpha = .025$	$\alpha = .05$	73	81	90	98	107	117
$\alpha = .010$	$\alpha = .02$	62	69	77	85	93	102
$\alpha = .005$	$\alpha = .01$	55	68	68	76	84	92
Una cola	Dos colas	$n = 29$	$n = 30$	$n = 31$	$n = 32$	$n = 33$	$n = 34$
$\alpha = .050$	$\alpha = .10$	141	152	163	175	188	201
$\alpha = .025$	$\alpha = .05$	127	137	148	159	171	183
$\alpha = .010$	$\alpha = .02$	111	120	130	141	151	162
$\alpha = .005$	$\alpha = .01$	100	109	118	128	138	149
Una cola	Dos colas	$n = 35$	$n = 36$	$n = 37$	$n = 38$	$n = 39$	
$\alpha = .050$	$\alpha = .10$	214	228	242	256	271	
$\alpha = .025$	$\alpha = .05$	195	208	222	235	250	
$\alpha = .010$	$\alpha = .02$	174	186	198	211	224	
$\alpha = .005$	$\alpha = .01$	160	171	183	195	208	
Una cola	Dos colas	$n = 40$	$n = 41$	$n = 42$	$n = 43$	$n = 44$	$n = 45$
$\alpha = .050$	$\alpha = .10$	287	303	319	336	353	371
$\alpha = .025$	$\alpha = .05$	264	279	295	311	327	344
$\alpha = .010$	$\alpha = .02$	238	252	267	281	297	313
$\alpha = .005$	$\alpha = .01$	221	234	248	262	277	292
Una cola	Dos colas	$n = 46$	$n = 47$	$n = 48$	$n = 49$	$n = 50$	
$\alpha = .050$	$\alpha = .10$	389	408	427	446	466	
$\alpha = .025$	$\alpha = .05$	361	379	397	415	434	
$\alpha = .010$	$\alpha = .02$	329	345	362	380	398	
$\alpha = .005$	$\alpha = .01$	307	323	339	356	373	

FUENTE: De "Some Rapid Approximate Statistical Procedures" (1964) 28, por F. Wilcoxon y R.A. Wilcox. Reproducida con el bondadoso permiso de Lederle Laboratories, división de American Cyanamid Company.

TABLA 9
Valores críticos
del coeficiente
de correlación
de rango de
Spearman para
una prueba de
una cola

<i>n</i>	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$
5	.900	—	—	—
6	.829	.886	.943	—
7	.714	.786	.893	—
8	.643	.738	.833	.881
9	.600	.683	.783	.833
10	.564	.648	.745	.794
11	.523	.623	.736	.818
12	.497	.591	.703	.780
13	.475	.566	.673	.745
14	.457	.545	.646	.716
15	.441	.525	.623	.689
16	.425	.507	.601	.666
17	.412	.490	.582	.645
18	.399	.476	.564	.625
19	.388	.462	.549	.608
20	.377	.450	.534	.591
21	.368	.438	.521	.576
22	.359	.428	.508	.562
23	.351	.418	.496	.549
24	.343	.409	.485	.537
25	.336	.400	.475	.526
26	.329	.392	.465	.515
27	.323	.385	.456	.505
28	.317	.377	.448	.496
29	.311	.370	.440	.487
30	.305	.364	.432	.478

FUENTE: De "Distribution of Sums of Squares of Rank Differences for Small Samples" por E.G. Olds, *Annals of Mathematical Statistics* 9 (1938). Reproducida con el permiso del editor, *Annals of Mathematical Statistics*.

TABLA 10 Números aleatorios

Fila	Columna													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	10480	15011	01536	02011	81647	91646	69179	14194	62590	36207	20969	99570	91291	90700
2	22368	46573	25595	85393	30995	89198	27982	53402	93965	34095	52666	19174	39615	99505
3	24130	48360	22527	97265	76393	64809	15179	24830	49340	32081	30680	19655	63348	58629
4	42167	93093	06243	61680	07856	16376	39440	53537	71341	57004	00849	74917	97758	16379
5	37570	39975	81837	16656	06121	91782	60468	81305	49684	60672	14110	06927	01263	54613
6	77921	06907	11008	42751	27756	53498	18602	70659	90655	15053	21916	81825	44394	42880
7	99562	72905	56420	69994	98872	31016	71194	18738	44013	48840	63213	21069	10634	12952
8	96301	91977	05463	07972	18876	20922	94595	56869	69014	60045	18425	84903	42508	32307
9	89579	14342	63661	10281	17453	18103	57740	84378	25331	12566	58678	44947	05585	56941
10	84575	36857	53342	53988	53060	59533	38867	62300	08158	17983	16439	11458	18593	64952
11	28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	01547	85590	91610	78188
12	63553	40961	48235	03427	49626	69445	18663	72695	52180	20847	12234	90511	33703	90322
13	09429	93969	52636	92737	88974	33488	36320	17617	30015	08272	84115	27156	30613	74952
14	10365	61129	87529	85689	48237	52267	67689	93394	01511	26358	85104	20285	29975	89868
15	07119	97336	71048	08178	77233	13916	47564	81056	97735	85977	29372	74461	28551	90707
16	51085	12765	51821	51259	77452	16308	60756	92144	49442	53900	70960	63990	75601	40719
17	02368	21382	52404	60268	89368	19885	55322	44819	01188	65255	64835	44919	05944	55157
18	01011	54092	33362	94904	31273	04146	18594	29852	71585	85030	51132	01915	92747	64951
19	52162	53916	46369	58586	23216	14513	83149	98736	23495	64350	94738	17752	35156	35749
20	07056	97628	33787	09998	42698	06691	76988	13602	51851	46104	88916	19509	25625	58104
21	48663	91245	85828	14346	09172	30168	90229	04734	59193	22178	30421	61666	99904	32812
22	54164	58492	22421	74103	47070	25306	76468	26384	58151	06646	21524	15227	96909	44592
23	32639	32363	05597	24200	13363	38005	94342	28728	35806	06912	17012	64161	18296	22851
24	29334	27001	87637	87308	58731	00256	45834	15398	46557	41135	10367	07684	36188	18510
25	02488	33062	28834	07351	19731	92420	60952	61280	50001	67658	32586	86679	50720	94953
26	81525	72295	04839	96423	24878	82651	66566	14778	76797	14780	13300	87074	79666	95725
27	29676	20591	68086	26432	46901	20849	89768	81536	86645	12659	92259	57102	80428	25280
28	00742	57392	39064	66432	84673	40027	32832	61362	98947	96067	64760	64585	96096	98253
29	05366	04213	25669	26422	44407	44048	37937	63904	45766	66134	75470	66520	34693	90449
30	91921	26418	64117	94305	26766	25940	39972	22209	71500	64568	91402	42416	07844	69618
31	00582	04711	87917	77341	42206	35126	74087	99547	81817	42607	43808	76655	62028	76630
32	00725	69884	62797	56170	86324	88072	76222	36086	84637	93161	76038	65855	77919	88006
33	69011	65795	95876	55293	18988	27354	26575	08625	40801	59920	29841	80150	12777	48501
34	25976	57948	29888	88604	67917	48708	18912	82271	65424	69774	33611	54262	85963	03547
35	09763	83473	73577	12908	30883	18317	28290	35797	05998	41688	34952	37888	38917	88050
36	91567	42595	27958	30134	04024	86385	29880	99730	55536	84855	29080	09250	79656	73211
37	17955	56349	90999	49127	20044	59931	06115	20542	18059	02008	73708	83517	36103	42791
38	46503	18584	18845	49618	02304	51038	20655	58727	28168	15475	56942	53389	20562	87338
39	92157	89634	94824	78171	84610	82834	09922	25417	44137	48413	25555	21246	35509	20468
40	14577	62765	35605	81263	39667	47358	56873	56307	61607	49518	89656	20103	77490	18062
41	98427	07523	33362	64270	01638	92477	66969	98420	04880	45585	46565	04102	46880	45709
42	34914	63976	88720	82765	34476	17032	87589	40836	32427	70002	70663	88863	77775	69348
43	70060	28277	39475	46473	23219	53416	94970	25832	69975	94884	19661	72828	00102	66794
44	53976	54914	06990	67245	68350	82948	11398	42878	80287	88267	47363	46634	06541	97809
45	76072	29515	40980	07391	58745	25774	22987	80059	39911	96189	41151	14222	60697	59583
46	90725	52210	83974	29992	65831	38857	50490	83765	55657	14361	31720	57375	56228	41546
47	64364	67412	33339	31926	14883	24413	59744	92351	97473	89286	35931	04110	23726	51900
48	08962	00358	31662	25388	61642	34072	81249	35648	56891	69352	48373	45578	78547	81788
49	95012	68379	93526	70765	10592	04542	76463	54328	02349	17247	28865	14777	62730	92277
50	15664	10493	20492	38391	91132	21999	59516	81652	27195	48223	46751	22923	32261	85653

FUENTE: De Handbook of Tables for Probabiliy and Statistics, 2a. ed., por William H. Beyer (CRC Press). Usada con permiso de William H. Beyer.

TABLA 10 (continuación)

Fila	Columna													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
51	16408	81899	04153	53381	79401	21438	83035	92350	36693	31238	59649	91754	72772	02338
52	18629	81953	05520	91962	04739	13092	97662	24822	94730	06496	35090	04822	86774	98289
53	73115	35101	47498	87637	99016	71060	88824	71013	18735	20286	23153	72924	35165	43040
54	57491	16703	23167	49323	45021	33132	12544	41035	80780	45393	44812	12515	98931	91202
55	30405	83946	23792	14422	15059	45799	22716	19792	09983	74353	68668	30429	70735	25499
56	16631	35006	85900	98275	32388	52390	16815	69298	82732	38480	73817	32523	41961	44437
57	96773	20206	42559	78985	05300	22164	24369	54224	35033	19687	11052	91491	60383	19746
58	38935	64202	14349	82674	66523	44133	00697	35552	35970	19124	63318	29686	03387	59846
59	31624	76384	17403	53363	44167	64486	64758	75366	76554	31601	12614	33072	60332	92325
60	78919	19474	23632	27889	47914	02584	37680	20801	72152	39339	34806	08930	85001	87820
61	03931	33309	57047	74211	63445	17361	62825	39908	05607	91284	68833	25570	38818	46920
62	74426	33278	43972	10119	89917	15665	52872	73823	73144	88662	88970	74492	51805	99378
63	09066	00903	20795	95452	92648	45454	09552	88815	16553	51125	79375	97596	16296	66092
64	42238	12426	87025	14267	20979	04508	64535	31355	86064	29472	47689	05974	52468	16834
65	16153	08002	26504	41744	81959	65642	74240	56302	00033	67107	77510	70625	28725	34191
66	21457	40742	29820	96783	29400	21840	15035	34537	33310	06116	95240	15957	16572	06004
67	21581	57802	02050	89728	17937	37621	47075	42080	97403	48626	68995	43805	33386	21597
68	55612	78095	83197	33732	05810	24813	86902	60397	16489	03264	88525	42786	05269	92532
69	44657	66999	99324	51281	84463	60563	79312	93454	68876	25471	93911	25650	12682	73572
70	91340	84979	46949	81973	37949	61023	43997	15263	80644	43942	89203	71795	99533	50501
71	91227	21199	31935	27022	84067	05462	35216	14486	29891	68607	41867	14951	91696	85065
72	50001	38140	66321	19924	72163	09538	12151	06878	91903	18749	34405	56087	82790	70925
73	65390	05224	72958	28609	81406	39147	25549	48542	42627	45233	57202	94617	23772	07896
74	27504	96131	83944	41575	10573	08619	64482	73923	36152	05184	94142	25299	84387	34925
75	37169	94851	39117	89632	00959	16487	65536	49071	39782	17095	02330	74301	00275	48280
76	11508	70225	51111	38351	19444	66499	71945	05422	13442	78675	84081	66938	93654	59894
77	37449	30362	06694	54690	04052	53115	62757	95348	78662	11163	81651	50245	34971	52924
78	46515	70331	85922	38329	57015	15765	97161	17869	45349	61796	66345	81073	49106	79860
79	30986	81223	42416	58353	21532	30502	32305	86482	05174	07901	54339	58861	74818	46942
80	63798	64995	46583	09785	44160	78128	83991	42865	92520	83531	80377	35909	81250	54238
81	82486	84846	99254	67632	43218	50076	21361	64816	51202	88124	41870	52689	51275	83556
82	21885	32906	92431	09060	64297	51674	64126	62570	26123	05155	59194	52799	28225	85762
83	60336	98782	07408	53458	13564	59089	26445	29789	85205	41001	12535	12133	14645	23541
84	43937	46891	24010	25560	86355	33941	25786	54990	71899	15475	95434	98227	21824	19585
85	97656	63175	89303	16275	07100	92063	21942	18611	47348	20203	18534	03862	78095	50136
86	03299	01221	05418	38982	55758	92237	26759	86367	21216	98442	08303	56613	91511	75928
87	79626	06486	03574	17668	07785	76020	79924	25651	83325	88428	85076	72811	22717	50585
88	85636	68335	47539	03129	65651	11977	02510	26113	99447	68645	34327	15152	55230	93448
89	18039	14367	61337	06177	12143	46609	32989	74014	64708	00533	35398	58408	13261	47908
90	08362	15656	60627	36478	65648	16764	53412	09013	07832	41574	17639	82163	60859	75567
91	79556	29068	04142	16268	15387	12856	66227	38358	22478	73373	88732	09443	82558	05250
92	92608	82674	27072	32534	17075	27698	98204	63863	11951	34648	88022	56148	34925	57031
93	23982	25835	40055	67006	12293	02753	14827	23235	35071	99704	37543	11601	35503	85171
94	09915	96306	05908	97901	28395	14186	00821	80703	70426	75647	76310	88717	37890	40129
95	59037	33300	26695	62247	69927	76123	50842	43834	86654	70959	79725	93872	28117	19233
96	42488	78077	69882	61657	34136	79180	97526	43092	04098	73571	80799	76536	71255	64239
97	46764	86273	63003	93017	31204	36692	40202	35275	57306	55543	53203	18098	47625	88684
98	03237	45430	55417	63282	90816	17349	88298	90183	36600	78406	06216	95787	42579	90730
99	86591	81482	52667	61582	14972	90053	89534	76036	49199	43716	97548	04379	46370	28672
100	38534	01715	94964	87288	65680	43772	39560	12918	86737	62738	19636	51132	25739	56947

TABLA 11 a)
Puntos
porcentuales
del rango
de Student,
 $q_{.05}(k, gl)$
puntos
de 5%
superior

<i>gl</i>	<i>k</i>									
	2	3	4	5	6	7	8	9	10	11
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51
13	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36
15	3.01	3.67	4.08	4.37	4.60	4.78	4.94	5.08	5.20	5.31
16	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26
17	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11	5.21
18	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17
19	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14
20	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01
30	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.92
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73
120	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64
∞	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55

TABLA 11a)
(continuación)

<i>k</i>									<i>gl</i>
12	13	14	15	16	17	18	19	20	
51.96	53.20	54.33	55.36	56.32	57.22	58.04	58.83	59.56	1
14.75	15.08	15.38	15.65	15.91	16.14	16.37	16.57	16.77	2
9.95	10.15	10.35	10.52	10.69	10.84	10.98	11.11	11.24	3
8.21	8.37	8.52	8.66	8.79	8.91	9.03	9.13	9.23	4
7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	8.21	5
6.79	6.92	7.03	7.14	7.24	7.34	7.43	7.51	7.59	6
6.43	6.55	6.66	6.76	6.85	6.94	7.02	7.10	7.17	7
6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87	8
5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.64	9
5.83	5.93	6.03	6.11	6.19	6.27	6.34	6.40	6.47	10
5.71	5.81	5.90	5.98	6.06	6.13	6.20	6.27	6.33	11
5.61	5.71	5.80	5.88	5.95	6.02	6.09	6.15	6.21	12
5.53	5.63	5.71	5.79	5.86	5.93	5.99	6.05	6.11	13
5.46	5.55	5.64	5.71	5.79	5.85	5.91	5.97	6.03	14
5.40	5.49	5.57	5.65	5.72	5.78	5.85	5.90	5.96	15
5.35	5.44	5.52	5.59	5.66	5.73	5.79	5.84	5.90	16
5.31	5.39	5.47	5.54	5.61	5.67	5.73	5.79	5.84	17
5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79	18
5.23	5.31	5.39	5.46	5.53	5.59	5.65	5.70	5.75	19
5.20	5.28	5.36	5.43	5.49	5.55	5.61	5.66	5.71	20
5.10	5.18	5.25	5.32	5.38	5.44	5.49	5.55	5.59	24
5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.47	30
4.90	4.98	5.04	5.11	5.16	5.22	5.27	5.31	5.36	40
4.81	4.88	4.94	5.00	5.06	5.11	5.15	5.20	5.24	60
4.71	4.78	4.84	4.90	4.95	5.00	5.04	5.09	5.13	120
4.62	4.68	4.74	4.80	4.85	4.89	4.93	4.97	5.01	∞

FUENTE: De *Biometrika Tables for Statisticians*, vol. 1, 3a. ed., editado por E.S. Pearson y H.O. Hartley (Cambridge University Press, 1966). Reproducida con permiso de los fideicomisarios de *Biometrika*.

TABLA 11b)
Puntos
porcentuales
del rango de
Student,
 $q_{01}(k, g!);$
puntos de 1%
superior

	k										
gl	2	3	4	5	6	7	8	9	10	11	
1	90.03	135.0	164.3	185.6	202.2	215.8	227.2	237.0	245.6	253.2	
2	14.04	19.02	22.29	24.72	26.63	28.20	29.53	30.68	31.69	32.59	
3	8.26	10.62	12.17	13.33	14.24	15.00	15.64	16.20	16.69	17.13	
4	6.51	8.12	9.17	9.96	10.58	11.10	11.55	11.93	12.27	12.57	
5	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48	
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55	
8	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86	8.03	
9	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49	7.65	
10	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36	
11	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	
12	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	
14	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66	
15	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55	
16	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	
19	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25	
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	
24	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	
30	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	
40	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60	5.69	
60	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45	5.53	
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.37	
∞	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	

TABLA 11b)
(continuación)

<i>k</i>									<i>gl</i>
12	13	14	15	16	17	18	19	20	
260.0	266.2	271.8	277.0	281.8	286.3	290.0	294.3	298.0	1
33.40	34.13	34.81	35.43	36.00	36.53	37.03	37.50	37.95	2
17.53	17.89	18.22	18.52	18.81	19.07	19.32	19.55	19.77	3
12.84	13.09	13.32	13.53	13.73	13.91	14.08	14.24	14.40	4
10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81	11.93	5
9.48	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54	6
8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65	7
8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03	8
7.78	7.91	8.03	8.13	8.23	8.33	8.41	8.49	8.57	9
7.49	7.60	7.71	7.81	7.91	7.99	8.08	8.15	8.23	10
7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95	11
7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73	12
6.90	7.01	7.10	7.19	7.27	7.35	7.42	7.48	7.55	13
6.77	6.87	6.96	7.05	7.13	7.20	7.27	7.33	7.39	14
6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26	15
6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15	16
6.48	6.57	6.66	6.73	6.81	6.87	6.94	7.00	7.05	17
6.41	6.50	6.58	6.65	6.72	6.79	6.85	6.91	6.97	18
6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89	19
6.28	6.37	6.45	6.52	6.59	6.65	6.71	6.77	6.82	20
6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61	24
5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41	30
5.76	5.83	5.90	5.96	6.02	6.07	6.12	6.16	6.21	40
5.60	5.67	5.73	5.78	5.84	5.89	5.93	5.97	6.01	60
5.44	5.50	5.56	5.61	5.66	5.71	5.75	5.79	5.83	120
5.29	5.35	5.40	5.45	5.49	5.54	5.57	5.61	5.65	∞

FUENTE: De *Biometrika Tables for Statisticians*, vol. 1, 3a. ed., editado por E.S. Pearson y H.O. Hartley (Cambridge University Press, 1966). Reproducida con permiso de los fideicomisarios de *Biometrika*.

TABLA 12
Porcentiles de
la distribución
normal
estandarizada

$P(Z \leq z)$	z
0.001	-3.0902
0.005	-2.5758
0.01	-2.3263
0.02	-2.0537
0.03	-1.8808
0.04	-1.7507
0.05	-1.6449
0.10	-1.2816
0.15	-1.0364
0.20	-0.8416
0.30	-0.5244
0.40	-0.2533
0.50	0
0.60	0.2533
0.70	0.5244
0.80	0.8414
0.85	1.0364
0.90	1.2816
0.95	1.6449
0.96	1.7507
0.97	1.8808
0.98	2.0537
0.99	2.3263
0.995	2.5758
0.999	3.0902

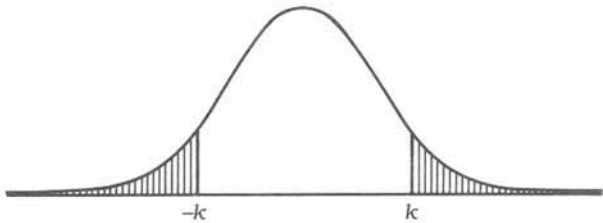


TABLA 13
Valores críticos
de Q (Tukey-
Snedecor) para
comparaciones
múltiples

gl grados de libertad	Números de grupos							
	2	3	4	5	6	7	8	9
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36
2	6.08	8.28	9.80	10.89	11.73	12.43	13.03	13.54
3	4.50	5.91	6.83	7.51	8.04	8.47	8.85	9.18
4	3.93	5.04	5.76	6.29	6.70	7.06	7.35	7.60
5	3.64	4.60	5.22	5.67	5.93	6.38	6.58	6.80
6	3.46	4.34	4.90	5.31	5.63	5.89	6.12	6.32
7	3.34	4.16	4.68	5.06	5.35	5.59	5.82	5.99
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77
9	3.20	3.95	4.42	4.76	5.02	5.24	5.43	5.60
10	3.15	3.88	4.33	4.66	4.91	5.12	5.30	5.46
11	3.11	3.82	4.26	4.58	4.82	5.03	5.20	5.35
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27
13	3.06	3.73	4.15	4.46	4.69	4.88	5.05	5.19
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13
15	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08
16	3.00	3.65	4.05	4.34	4.56	4.74	4.90	5.03
17	2.98	3.62	4.02	4.31	4.52	4.70	4.86	4.99
18	2.97	3.61	4.00	4.28	4.49	4.67	4.83	4.96
19	2.96	3.59	3.98	4.26	4.47	4.64	4.79	4.92
20	2.95	3.58	3.96	4.24	4.45	4.62	4.77	4.90
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81
30	2.89	3.48	3.84	4.11	4.30	4.46	4.60	4.72
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55
120	2.80	3.36	3.69	3.92	4.10	4.24	4.36	4.47
∞	2.77	3.32	3.63	3.86	4.03	4.17	4.29	4.39

Fuente: *Biometrika Tables for Statisticians*, vols. I y II, editado por E. S. Pearson y H. O. Hartley, Biometrika Trust, Oxford University Press, 1976.

TABLA 14
Valores críticos
para la prueba
de Hartley
(homogeneidad
de varianzas)

gl para χ^2_{α} σ_j^2	α	$k = \text{número de varianzas}$											
		2	3	4	5	6	7	8	9	10	11	12	
4	.05	9.60	15.5	20.6	25.2	29.5	33.6	37.5	41.4	44.6	48.0	51.4	
	.01	23.2	37.	49.	59.	69.	79.	89.	97.	106.	113.	120.	
5	.05	7.15	10.8	13.7	16.3	18.7	20.8	22.9	24.7	26.5	28.2	29.9	
	.01	14.9	22.	28.	33.	38.	42.	46.	50.	54.	57.	60.	
6	.05	5.82	8.38	10.4	12.1	13.7	15.0	16.3	17.5	18.6	19.7	20.7	
	.01	11.1	15.5	19.1	22.	25.	27.	30.	32.	34.	36.	37.	
7	.05	4.99	6.94	8.44	9.70	10.8	11.8	12.7	13.5	14.3	15.1	15.8	
	.01	8.89	12.1	14.5	16.5	18.4	20.	22.	23.	24.	26.	27.	
8	.05	4.43	6.00	7.18	8.12	9.03	9.78	10.5	11.1	11.7	12.2	12.7	
	.01	7.50	9.9	11.7	13.2	14.5	15.8	16.9	17.9	18.9	19.8	21.	
9	.05	4.03	5.34	6.31	7.11	7.80	8.41	8.95	9.45	9.91	10.3	10.7	
	.01	6.54	8.5	9.9	11.1	12.1	13.1	13.9	14.7	15.3	16.0	16.6	
10	.05	3.72	4.85	5.67	6.34	6.92	7.42	7.87	8.28	8.66	9.01	9.34	
	.01	5.85	7.4	8.6	9.6	10.4	11.1	11.8	12.4	12.9	13.4	13.9	
12	.05	3.28	4.16	4.79	5.30	5.72	6.09	6.42	6.72	7.00	7.25	7.48	
	.01	4.91	6.1	6.9	7.6	8.2	8.7	9.1	9.5	9.9	10.2	10.6	
15	.05	2.86	3.54	4.01	4.37	4.68	4.95	5.19	5.40	5.59	5.77	5.93	
	.01	4.07	4.9	5.5	6.0	6.4	6.7	7.1	7.3	7.5	7.8	8.0	
20	.05	2.46	2.95	3.29	3.54	3.76	3.94	4.10	4.24	4.37	4.49	4.59	
	.01	3.32	3.8	4.3	4.6	4.9	5.1	5.3	5.5	5.6	5.8	5.9	
30	.05	2.07	2.40	2.61	2.78	2.91	3.02	3.12	3.21	3.29	3.36	3.39	
	.01	2.63	3.0	3.3	3.4	3.6	3.7	3.8	3.9	4.0	4.1	4.2	
60	.05	1.67	1.85	1.96	2.04	2.11	2.17	2.22	2.26	2.30	2.33	2.36	
	.01	1.96	2.2	2.3	2.4	2.4	2.5	2.5	2.6	2.6	2.7	2.7	
∞	.05	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
	.01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	

Fuente: *Biometrika Tables for Statisticians*, vols. I y II, editado por E.S. Pearson y H. O. Hartley, Biometrika Trust, Oxford University Press, 1976.

TABLA 15
Valores críticos
para la prueba
de Cochran
(homogeneidad
de variantes)

$$C = (\text{la } s^2 \text{ más grande } \sum S_j^2)$$

gl para s_j^2	$1 - \alpha$	$k = \text{número de varianzas}$										
		2	3	4	5	6	7	8	9	10	15	20
1	0.95	0.9985	0.9669	0.9065	0.8412	0.7808	0.7271	0.6798	0.6385	0.6020	0.4709	0.3894
	0.99	0.9999	0.9933	0.9676	0.9279	0.8828	0.8376	0.7945	0.7544	0.7175	0.5747	0.4799
2	0.95	0.9750	0.8709	0.7679	0.6838	0.6161	0.5612	0.5157	0.4775	0.4450	0.3346	0.2705
	0.99	0.9950	0.9423	0.8643	0.7885	0.7218	0.6644	0.6152	0.5727	0.5358	0.4069	0.3297
3	0.95	0.9392	0.7977	0.6841	0.5981	0.5321	0.4800	0.4377	0.4027	0.3733	0.2758	0.2205
	0.99	0.9794	0.8831	0.7814	0.6957	0.6258	0.5685	0.5209	0.4810	0.4469	0.3317	0.2654
4	0.95	0.9057	0.7457	0.6287	0.5441	0.4803	0.4307	0.3910	0.3584	0.3311	0.2419	0.1921
	0.99	0.9586	0.8335	0.7212	0.6329	0.5635	0.5080	0.4627	0.4251	0.3934	0.2882	0.2288
5	0.95	0.8772	0.7071	0.4447	0.5895	0.5065	0.3974	0.3595	0.3286	0.3029	0.2195	0.1735
	0.99	0.9373	0.7933	0.6761	0.5875	0.5195	0.4559	0.4226	0.3870	0.3572	0.2593	0.2048
6	0.95	0.8534	0.6771	0.5598	0.4783	0.4184	0.3726	0.3362	0.3067	0.2823	0.2034	0.1602
	0.99	0.9172	0.7606	0.6410	0.5531	0.4866	0.4347	0.3932	0.3592	0.3308	0.2386	0.1877
7	0.95	0.8332	0.6530	0.5365	0.4564	0.3980	0.3535	0.3185	0.2901	0.2666	0.1911	0.1501
	0.99	0.8988	0.7335	0.6129	0.5259	0.4608	0.4105	0.3704	0.3378	0.3106	0.2228	0.1748
8	0.95	0.8159	0.6333	0.5175	0.4387	0.3817	0.3384	0.3043	0.2768	0.2541	0.1815	0.1422
	0.99	0.8823	0.7107	0.5897	0.5037	0.4401	0.3911	0.3522	0.3207	0.2945	0.2104	0.1646
9	0.95	0.8010	0.6167	0.5017	0.4241	0.3682	0.3259	0.2926	0.2659	0.2439	0.1736	0.1357
	0.99	0.8674	0.6912	0.5702	0.4854	0.4229	0.3751	0.3373	0.3067	0.2813	0.2002	0.1567
16	0.95	0.7341	0.5466	0.4366	0.3645	0.3135	0.2756	0.2462	0.2226	0.2032	0.1429	0.1108
	0.99	0.7949	0.6059	0.4884	0.4094	0.3529	0.3105	0.2779	0.2514	0.2297	0.1612	0.1248
36	0.95	0.6602	0.4748	0.3720	0.3066	0.2612	0.2278	0.2022	0.1820	0.1655	0.1144	0.0879
	0.99	0.7067	0.5153	0.4057	0.3351	0.2858	0.2494	0.2214	0.1992	0.1811	0.1251	0.0960
144	0.95	0.5813	0.4031	0.3093	0.2513	0.2119	0.1833	0.1616	0.1446	0.1308	0.0889	0.0675
	0.99	0.6062	0.4230	0.3251	0.2644	0.2229	0.1929	0.1700	0.1521	0.1376	0.0934	0.0709

Fuente: Journal of the American Statistical Association, vol. 75, 1980.

TABLA 16
Valores de $|n|$
para la prueba
de MacNemar

n	Nivel de significancia a una cola					
	0.100	0.050	0.025	0.010	0.005	0.0005
	Nivel de significancia a dos colas					
	0.200	0.100	0.050	0.020	0.010	0.001
4	0					
5	0	0				
6	0	0	0			
7	1	0	0	0		
8	1	1	0	0	0	
9	2	1	1	0	0	
10	2	1	1	0	0	
11	2	2	1	1	0	0
12	3	2	2	1	1	0
13	3	3	2	1	1	0
14	4	3	2	2	1	0
15	4	3	3	2	2	1
16	4	4	3	2	2	1
17	5	4	4	3	2	1
18	5	5	4	3	3	1
19	6	5	4	4	3	2
20	6	5	5	4	3	2
21	7	6	5	4	4	2
22	7	6	5	5	4	3
23	7	7	6	5	4	3
24	8	7	6	5	5	3
25	8	7	7	6	5	4
26	9	8	7	6	6	4
27	9	8	7	7	6	4
28	10	9	8	7	6	5
29	10	9	8	7	7	5
30	10	10	9	8	7	5
31	11	10	9	8	7	6
32	11	10	9	8	8	6
33	12	11	10	9	8	6
34	12	11	10	9	9	7
35	13	12	11	10	9	7
36	13	12	11	10	9	7
37	14	13	12	10	10	8
38	14	13	12	11	10	8
39	15	13	12	11	11	8
40	15	14	13	12	11	9
41	15	14	13	12	11	9
42	16	15	14	13	12	10
43	16	15	14	13	12	10
45	17	16	15	14	13	11
46	17	16	15	14	13	11
47	18	17	16	15	14	11
48	19	17	16	15	14	12
49	19	18	17	15	15	12
50	19	18	17	16	15	13

Fuente: Ignacio Méndez Ramírez *et al.*, *El protocolo de la investigación. Lineamientos para su elaboración y análisis*, 2a ed., 5a. reimpr., Trillas, México, 1997.

TABLA 17
Valores críticos
de A (prueba de
Sandler)

Para cualquier valor de $N-1$, la tabla muestra los valores de A , correspondientes a varios niveles de probabilidad R.D. Si $A_c(\text{calculada}) \leq A_{\alpha(\text{tabla})}$. Entonces A_c es significativa.

Niveles de significancia para una prueba direccional						
	0.05	0.025	0.01	0.005	0.0005	
Niveles de significancia para una prueba bidireccional						
N-1*	0.10	0.05	0.02	0.01	0.001	N-1
1	0.5125	0.5031	0.50049	0.50012	0.5000012	1
2	0.412	0.369	0.347	0.340	0.334	2
3	0.385	0.324	0.286	0.272	0.254	3
4	0.376	0.304	0.257	0.238	0.211	4
5	0.372	0.293	0.240	0.218	0.184	5
6	0.370	0.286	0.230	0.205	0.167	6
7	0.369	0.281	0.222	0.196	0.155	7
8	0.368	0.278	0.217	0.190	0.146	8
9	0.368	0.276	0.213	0.185	0.139	9
10	0.368	0.274	0.210	0.181	0.134	10
11	0.368	0.273	0.207	0.178	0.130	11
12	0.368	0.271	0.205	0.176	0.126	12
13	0.368	0.270	0.204	0.174	0.124	13
14	0.368	0.270	0.202	0.172	0.121	14
15	0.368	0.269	0.201	0.170	0.119	15
16	0.368	0.268	0.200	0.169	0.117	16
17	0.368	0.268	0.199	0.168	0.116	17
18	0.368	0.267	0.198	0.167	0.114	18
19	0.368	0.267	0.197	0.166	0.113	19
20	0.368	0.266	0.197	0.165	0.112	20
21	0.368	0.266	0.196	0.165	0.111	21
22	0.368	0.266	0.196	0.164	0.110	22
23	0.368	0.266	0.195	0.163	0.109	23
24	0.368	0.265	0.195	0.163	0.108	24
25	0.368	0.265	0.194	0.161	0.108	25
26	0.368	0.265	0.194	0.162	0.107	26
27	0.368	0.265	0.193	0.161	0.107	27
28	0.368	0.265	0.193	0.161	0.106	28
29	0.368	0.264	0.193	0.161	0.106	29
30	0.368	0.264	0.193	0.160	0.105	30
40	0.368	0.263	0.191	0.158	0.102	40
60	0.369	0.262	0.189	0.155	0.099	60
120	0.369	0.261	0.187	0.153	0.095	120
∞	0.370	0.260	0.185	0.151	0.092	∞

* N = el número de pares.

Tabla generada por J. Carlos Medina (MacStat)

TABLA 18
Tabla de Dunnett
1a para $p = 95\%$

<i>gl</i>	1	2	3	4	5	6	7	8	9
5	2.02	2.44	2.68	2.85	2.98	3.08	3.16	3.24	3.30
6	1.94	2.34	2.56	2.71	2.83	2.92	3.00	3.07	3.12
7	1.89	2.27	2.48	2.62	2.73	2.82	2.89	2.95	3.01
8	1.86	2.22	2.42	2.55	2.66	2.74	2.81	2.87	2.92
9	1.83	2.18	2.37	2.50	2.60	2.68	2.75	2.81	2.86
10	1.81	2.15	2.34	2.47	2.56	2.64	2.70	2.76	2.81
11	1.80	2.13	2.31	2.44	2.53	2.60	2.67	2.72	2.77
12	1.78	2.11	2.29	2.41	2.50	2.58	2.64	2.69	2.74
13	1.77	2.09	2.27	2.39	2.48	2.55	2.61	2.66	2.71
14	1.76	2.08	2.25	2.37	2.46	2.53	2.59	2.64	2.69
15	1.75	2.07	2.24	2.36	2.44	2.51	2.57	2.62	2.67
16	1.75	2.06	2.23	2.34	2.43	2.50	2.56	2.61	2.65
17	1.74	2.05	2.22	2.33	2.42	2.49	2.54	2.59	2.64
18	1.73	2.04	2.21	2.32	2.41	2.48	2.53	2.58	2.62
19	1.73	2.03	2.20	2.31	2.40	2.47	2.52	2.57	2.61
20	1.72	2.03	2.19	2.30	2.39	2.46	2.51	2.56	2.60
24	1.71	2.01	2.17	2.28	2.36	2.43	2.48	2.53	2.57
30	1.70	1.99	2.15	2.25	2.33	2.40	2.45	2.50	2.54
40	1.68	1.97	2.13	2.23	2.31	2.37	2.42	2.47	2.51
60	1.67	1.95	2.10	2.21	2.28	2.35	2.39	2.44	2.48
120	1.66	1.93	2.08	2.18	2.26	2.32	2.37	2.41	2.45
∞	1.64	1.92	2.06	2.16	2.23	2.29	2.34	2.38	2.42

Fuente: Journal of the American Statistical Association, vol. 50, pp. 1096-1121.

TABLA 19
Tabla de Dunnet
1b para $p = 99\%$

Tabla de Dunnet 1b para $p = 99\%$

g^l	1	2	3	4	5	6	7	8	9
5	3.37	3.90	4.21	4.43	4.60	4.73	4.85	4.94	5.03
6	3.14	3.61	3.88	4.07	4.21	4.33	4.43	4.51	4.59
7	3.00	3.42	3.66	3.83	3.96	4.07	4.15	4.23	4.30
8	2.90	3.29	3.51	3.67	3.79	3.88	3.96	4.03	4.09
9	2.82	3.19	3.40	3.55	3.66	3.75	3.82	3.89	3.94
10	2.76	3.11	3.31	3.45	3.56	3.64	3.71	3.78	3.83
11	2.72	3.06	3.25	3.38	3.48	3.56	3.63	3.69	3.74
12	2.68	3.01	3.19	3.32	3.42	3.50	3.56	3.62	3.67
13	2.65	2.97	3.15	3.27	3.37	3.44	3.51	3.56	3.61
14	2.62	2.94	3.11	3.23	3.32	3.40	3.46	3.51	3.56
15	2.60	2.91	3.08	3.20	3.29	3.36	3.42	3.47	3.52
16	2.58	2.88	3.05	3.17	3.26	3.33	3.39	3.44	3.48
17	2.57	2.86	3.03	3.14	3.23	3.30	3.36	3.41	3.45
18	2.55	2.84	3.01	3.12	3.21	3.27	3.33	3.38	3.42
19	2.54	2.83	2.99	3.10	3.18	3.25	3.31	3.36	3.40
20	2.53	2.81	2.97	3.08	3.17	3.23	3.29	3.34	3.38
24	2.49	2.77	2.92	3.03	3.11	3.17	3.22	3.27	3.31
30	2.46	2.72	2.87	2.97	3.05	3.11	3.16	3.21	3.24
40	2.42	2.68	2.82	2.92	2.99	3.05	3.10	3.14	3.18
60	2.39	2.64	2.78	2.87	2.94	3.00	3.04	3.08	3.12
120	2.36	2.60	2.73	2.82	2.89	2.94	2.99	3.03	3.06
∞	2.33	2.56	2.68	2.77	2.84	2.89	2.93	2.97	3.00

TABLA 20
Tabla de Dunnett
2a para $p = 95\%$

gl	1	2	3	4	5	6	7	8	9
5	2.57	3.03	3.39	3.66	3.88	4.06	4.22	4.36	4.49
6	2.45	2.86	3.18	3.41	3.60	3.75	3.88	4.00	4.11
7	2.36	2.75	3.04	3.24	3.41	3.54	3.66	3.76	3.86
8	2.31	2.67	2.94	3.13	3.28	3.40	3.51	3.60	3.68
9	2.26	2.61	2.86	3.04	3.18	3.29	3.39	3.48	3.55
10	2.23	2.57	2.81	2.97	3.11	3.21	3.31	3.39	3.46
11	2.20	2.53	2.76	2.92	3.05	3.15	3.24	3.31	3.38
12	2.18	2.50	2.72	2.88	3.00	3.10	3.18	3.25	3.32
13	2.16	2.48	2.69	2.84	2.96	3.06	3.14	3.21	3.27
14	2.14	2.46	2.67	2.81	2.93	3.02	3.10	3.17	3.23
15	2.13	2.44	2.64	2.79	2.90	2.99	3.07	3.13	3.19
16	2.12	2.42	2.63	2.77	2.88	2.96	3.04	3.10	3.16
17	2.11	2.41	2.61	2.75	2.85	2.94	3.01	3.08	3.13
18	2.10	2.40	2.59	2.73	2.84	2.92	2.99	3.05	3.11
19	2.09	2.39	2.58	2.72	2.82	2.90	2.97	3.04	3.09
20	2.09	2.38	2.57	2.70	2.81	2.89	2.96	3.02	3.07
24	2.06	2.35	2.53	2.66	2.76	2.84	2.91	2.96	3.01
30	2.04	2.32	2.50	2.62	2.72	2.79	2.86	2.91	2.96
40	2.02	2.29	2.47	2.58	2.67	2.75	2.81	2.86	2.90
60	2.00	2.27	2.43	2.55	2.63	2.70	2.76	2.81	2.85
120	1.98	2.24	2.40	2.51	2.59	2.66	2.71	2.76	2.80
∞	1.96	2.21	2.37	2.47	2.55	2.62	2.67	2.71	2.75

TABLA 21
Tabla de Dunnet
2b para $p = 95\%$

<i>gl</i>	1	2	3	4	5	6	7	8	9
5	4.03	4.63	5.09	5.44	5.73	5.97	6.18	6.36	6.53
6	3.71	4.22	4.60	4.88	5.11	5.30	5.47	5.61	5.74
7	3.50	3.95	4.28	4.52	4.71	4.87	5.01	5.13	5.24
8	3.36	3.77	4.06	4.27	4.44	4.58	4.70	4.81	4.90
9	3.25	3.63	3.90	4.09	4.24	4.37	4.48	4.57	4.65
10	3.17	3.53	3.78	3.95	4.10	4.21	4.31	4.40	4.47
11	3.11	3.45	3.68	3.85	3.98	4.09	4.18	4.26	4.33
12	3.05	3.39	3.61	3.76	3.89	3.99	4.08	4.15	4.22
13	3.01	3.33	3.54	3.69	3.81	3.91	3.99	4.06	4.13
14	2.98	3.29	3.49	3.64	3.75	3.84	3.92	3.99	4.05
15	2.95	3.25	3.45	3.59	3.70	3.79	3.86	3.93	3.99
16	2.92	3.22	3.41	3.55	3.65	3.74	3.82	3.88	3.93
17	2.90	3.19	3.38	3.51	3.62	3.70	3.77	3.83	3.89
18	2.88	3.17	3.35	3.48	3.58	3.67	3.74	3.80	3.85
19	2.86	3.15	3.33	3.46	3.55	3.64	3.70	3.76	3.81
20	2.85	3.13	3.31	3.43	3.53	3.61	3.67	3.73	3.78
24	2.80	3.07	3.24	3.36	3.45	3.52	3.58	3.64	3.69
30	2.75	3.01	3.17	3.28	3.37	3.44	3.50	3.55	3.59
40	2.70	2.95	3.10	3.21	3.29	3.36	3.41	3.46	3.50
60	2.66	2.90	3.04	3.14	3.22	3.28	3.33	3.38	3.42
120	2.62	2.84	2.98	3.08	3.15	3.21	3.25	3.30	3.33
∞	2.58	2.79	2.92	3.01	3.08	3.14	3.18	3.22	3.25

Fuentes de datos

Capítulo 1

1. "Fortune 500," http://money.cnn.com/magazines/fortune/fortune500/2010/full_list/, 8 July 2010.
2. "Birth Order and the Baby Boom," *American Demographics* (Trend Cop), March 1997, p. 10.
3. "Tuna Goes Upscale," *Consumer Reports*, June 2001, p. 19.
4. "Nintendo Wii 45496880019 Gaming Console," <http://videogames.pricegrabber.com/wii-console-accessories/Nintendo-Console/m27737371.html/st=pop/sv=title>, 8 July 2010.
5. A. Tubb, A.J. Parker, and G. Nickless, "The Analysis of Romano-British Pottery by Atomic Absorption Spectrophotometry," *Archaeometry* 22 (1980):153.
6. A. Azzalini and A.W. Bowman, "A Look at Some Data on the Old Faithful Geyser," *Applied Statistics* (1990):57.
7. "Aaron Rodgers #12 QB," <http://sports.espn.go.com/nfl/players/gamelog?playerId=8439> and "Drew Brees #9 QB," <http://sports.espn.go.com/nfl/players/gamelog?playerId=2580>, 18 March 2011.
8. D.G. Altman and J.M. Bland, "Time to Survival Data," *British Medical Journal BMJ* 1998; 3 17:468–469 (15 August) at <http://bmj.bmjjournals.com/cgi/content/full/317/7156/4687>.
9. P.A. Mackowiak, S.S. Wasserman, and M.M. Levine, "A Critical Appraisal of 98.6 Degrees F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wunderlich," *Journal of the American Medical Association* (268):1578–1580.
10. Allen L. Shoemaker, "What's Normal? Temperature, Gender, and Heart Rate," *Journal of Statistics Education* (1996).
11. Allen L. Shoemaker, "What's Normal? Temperature, Gender, and Heart Rate," *Journal of Statistics Education* (1996).
12. "New SUV Ratings & Reliability," www.consumerreports.org/cro/cars/new-cars/suvs/ratings-reliability/specs.htm, 20 August 2010.
13. www.mlb.com, 25 August 2010.
14. "How Much We Volunteer," <http://www.usatoday.com/news/snapshot.htm>, 25 August 2010.
15. "How Old is Your Vehicle?" <http://www.usatoday.com/news/snapshot.htm>, 6 July 2010.
16. "What Fans are Willing to Pay for a Concert Ticket," <http://www.usatoday.com/news/snapshot.htm>, 25 August 2010.

Capítulo 2

1. "Racial and Ethnic Distribution of ABO Bloodtypes: Bloodbook.com," <http://www.bloodbook.com/world-abo.html>, 25 August 2010.
2. Bruce E. Morgan and Michael A. Oberlander, "An Examination of Injuries in Major League Soccer," *The American Journal of Sports Medicine*, 29(4), 2001, pp. 426–429.
3. <http://sports.espn.go.com/nba/teams/stats>, 24 August 2010.

4. Adapted from “Demo Memo,” *American Demographics*, May 1997, and http://factfinder.census.gov/servlet/STTable?_bm=y&-geo_id=01000US&-qr_name=ACS_2008_1YR_G00_S0101&-ds_name=ACS_2008_1YR_G00_&-_lang=en&-redoLog=false&-state=st&-CONTEXT=st, 26 August 2010.
5. Andrew S. Levy, M.J. Wetzler, M. Lewars, and W. Laughlin, “Knee Injuries in Women Collegiate Rugby Players,” *The American Journal of Sports Medicine* 25, no. 3 (1997):360.
6. P.D. Franklin, R.A. Lemon, and H.S. Barden, “Accuracy of Imaging the Menisci on an In-Office, Dedicated, Magnetic Resonance Imaging Extremity System,” *The American Journal of Sports Medicine* 25, no. 3 (1997):382.

Capítulo 3

1. “Health: Latest Findings,” *Time*, 5 July 2010, p. 18.
2. “Coffee Breaks Daily,” <http://www.usatoday.com/news/snapshot.htm?section=L&label=2006-10-26-cell>, 26 October 2006.
3. http://testprep.about.com/od/sat/f/SATFAQ_GoodSAT.htm, 27 August 2010.
4. “How back pain limits sports” <http://www.usatoday.com/news/snapshot.htm?section=L&label=2006-10-27-wash>.
5. “U.S. Pet Ownership Statistics,” http://www.humanesociety.org/issues/pet_overpopulation/facts/pet_ownership_statistics.html, 27 August 2010.
6. <http://www.foodsafetynews.com/2010/07/three-daycares-closed-due-to-e-coli-outbreak/>, 26 August 2010.
7. <http://www.cnn.com/2010/HEALTH/04/15/foodborne.illness.cdc/index.html?iref=allsearch>, 26 August 2010.
8. Adapted from A.M. Goodman and A.R. Ennos, “The Response of Field-Grown Sunflower and Maize to Mechanical Support,” *Annals of Botany* 79 (1997):703.
9. John Fetto, “Shop Around the Clock,” *American Demographics*, September 2003, p. 18.
10. “Medical Encyclopedia: Pulse,” *Medline Plus: Trusted Health Information for You*, <http://www.nlm.nih.gov/medlineplus/ency/article/003399.htm#Normal%20Values>, 2 April 2004.
11. “Call It in the Air,” *The Press-Enterprise* (Riverside, CA), 19 October 1992.
12. Mark A. Atkinson, “Diet, Genetics, and Diabetes,” *Food Technology* 51, no. 3 (March 1997), p. 77.
13. “SquareTrade’s Report on Simple Vs. Smartphone Reliability,” <http://blog.squaretrade.com/2008/05/smartphone-reli.html>, 26 August 2010.
14. “Most popular chocolate,” <http://usatoday.com/news/snapshot.htm?section=L&label=2006-10-27-wash>.
15. “Upgrades Drivers Say Would Make Their Communities More Drivable,” <http://www.usatoday.com/news/snapshot.htm>, 6 July 2010.
16. “Consumers Who are Committed to Living with Fewer Credit Cards,” <http://www.usatoday.com/news/snapshot.htm>, 25 August 2010.
17. Philip A. Altman and D.S. Dittmer, *The Biology Data Book*, 2nd ed., Vol I. (Bethesda, MD: Federation of American Societies for Experimental Biology, 1964), p. 137.
18. <http://lib.stat.cmu.edu/DASL/Datafiles/MilitiamenChests.html>, 27 September 2010.
19. “How polite are cellphone users?” <http://usatoday.com/news/snapshot.htm?section=L&label=2006-10-27>.

Capítulo 4

1. Alice Park, “Omega-3 May Reduce Heart Risks Less Than Thought,” <http://www.time.com/time/health/article/0,8599,2014603,00.html>, 30 August 2010.
2. Liz Szabo, “Study: Tai Chi Could Ease Fibromyalgia Pain,” *The Press-Enterprise* (Riverside, CA), 19 August 2010, p. 4D.

3. “USDA National Nutrient Database: Bananas, raw,” http://www.nal.usda.gov/fnic/foodcom/cgi-bin/list_nut_edit.pl, 1 September 2010.
4. Allen L. Shoemaker, “What’s Normal? Temperature, Gender, and Heart Rate,” *Journal of Statistics Education* (1996).
5. Adam Fernandez, “Nuts About You,” *American Demographics* 26, no. 1 (February 2004):14.
6. P.C. Karalekas, Jr., C.R. Ryan, and F.B. Taylor, “Control of Lead, Copper, and Iron Pipe Corrosion in Boston,” *American Water Works Journal*, February 1983.
7. *Science News* 136 (19 August 1989):124.
8. Catherine M. Santaniello and R.E. Koning, “Are Radishes Really Allelopathic to Lettuce?” *The American Biology Teacher* 58, no. 2 (February 1996):102.
9. J. Hackl, *Journal of Quality Technology*, April 1991.
10. Daniel Seligman, “The Road to Monte Carlo,” *Fortune*, 15 April 1985.

Capítulo 5

1. Adapted from “Polar Bear,” http://en.wikipedia.org/wiki/Polar_Bear#Size_and_weight, 9 November 2006.
2. *Science News* 136 (19 August 1989):124.
3. “Immigration,” *ABC News/Washington Post Poll*, 3–6 June 2010, <http://www.pollingreport.com/immigration.htm>, 2 September 2010.
4. William Leonard, Barbara Speziale, and John Pernick, “Performance Assessment of a Standards-Based High School Biology Curriculum,” *The American Biology Teacher* 63, no. 5 (2001):310–316.
5. “Engineering, Computer Science Students Have Highest Salary Expectations,” http://www.naceweb.org/so/08022010/computer_science_engineering_salary/, 2 September 2010.
6. “Problems and Priorities,” *FOX News/Opinion Dynamics Poll*, 4–5 May 2010, <http://www.pollingreport.com/prioriti3.htm>, 4.
7. Mark Gillespie, “Baseball Fans Overwhelmingly Want Mandatory Steroid Testing,” Gallup News Service, <http://gallup.com/content/print.aspx?ci=11245>, 14 February 2004.
8. David L. Wheeler, “More Social Roles Means Fewer Colds,” *Chronicle of Higher Education* XLIII, no. 44 (11 July 1997):A13.
9. Adapted from A.M. Goodman and A.R. Ennos, “The Responses of Field-Grown Sunflower and Maize to Mechanical Support,” *Annals of Botany* 79 (1997):703.
10. G. Wayne Marino, “Selected Mechanical Factors Associated with Acceleration in Ice Skating,” *Research Quarterly for Exercise and Sport* 54, no. 3 (1983).

Capítulo 6

1. “Pricing of Tuna,” *Consumer Reports*, June 2001.
2. W.B. Jeffries, H.K. Voris, and C.M. Yang, “Diversity and Distribution of the Pedunculate Barnacles *Octolasmis* Gray, 1825 Epizoic on the Scyllarid Lobster *Thenus orientalis* (Lund, 1793),” *Crustaceana* 46, no. 3 (1984).
3. Wendy K. Baell and E.H. Wertheim, “Predictors of Outcome in the Treatment of Bulimia Nervosa,” *British Journal of Clinical Psychology* 31 (1992):330–332.
4. “L.A. Heart Data.” Adapted from data found at <http://www-unix.oit.umass.edu/~statdata/statdata/data/laheart.dat>.
5. Jan D. Lindhe, “Clinical Assessment of Antiplatelet Agents,” *Compendium of Continuing Education in Dentistry*, Supplement 5 (1984).
6. Susan J. Beckham, W.A. Grana, P. Buckley, J.E. Breasile, and P.L. Claypool, “A Comparison of Anterior Compartment Pressures in Competitive Runners and Cyclists,” *American Journal of Sports Medicine* 21, no. 1 (1992):36.
7. Michael A. Brehm, J.S. Buguliskis, D.K. Hawkins, E.S. Lee, D. Sabapathi, and R.A. Smith, “Determining Differences in Efficacy of Two Disinfectants Using *t*-tests,” *The American Biology Teacher* 58, no. 2 (February 1996):111.

8. A. Tubb, A.J. Parker, and G. Nickless, "The Analysis of Romano-British Pottery by Atomic Absorption Spectrophotometry," *Archaeometry* 22 (1980):153.
9. "2010 College-Bound Seniors: Total Group Profile Report," <http://professionals.collegeboard.com/profdownload/2010-total-group-profile-report-cbs.pdf>, 14 September 2010.
10. "Aaron Rodgers #12 QB," <http://sports.espn.go.com/nfl/players/gamelog?playerId=8439> and "Ben Roethlisberger #7 QB," <http://sports.espn.go.com/nfl/players/gamelog?playerId=5536>, 18 March 2011.
11. Carlos E. Macellari, "Revision of Serpulids of the Genus *Rotularia* (Annelida) at Seymour Island (Antarctic Peninsula) and Their Value in Stratigraphy," *Journal of Paleontology* 58, no. 4 (1984).
12. T.M. Casey, M.L. May, and K.R. Morgan, "Flight Energetics of Euglossine Bees in Relation to Morphology and Wing Stroke Frequency," *Journal of Experimental Biology* 116 (1985).
13. Karl J. Niklas and T.G. Owens, "Physiological and Morphological Modifications of *Plantago Major* (Plantaginaceae) in Response to Light Conditions," *American Journal of Botany* 76, no. 3 (1989):370–382.
14. John Fetto, "Shop Around the Clock," *American Demographics* 25, no. 7 (September 2003):18.

Capítulo 7

1. "Pricing of Tuna," *Consumer Reports*, June 2001.
2. "2010–2011: Interpreting Your GRE Scores" http://www.ets.org/s/gre/pdf/gre_interpreting_scores.pdf, 16 September 2010.
3. "2011 Automobile Insurance," <http://interactive.web.insurance.ca.gov/survey/survey?type=autoSurvey&event=autoStart>, 22 March 2011.
4. H.F. Barsam and Z.M. Simutis, "Computer-Based Graphics for Terrain Visualization Training," *Human Factors*, no. 26, 1984. Copyright 1984 by the Human Factors Society, Inc. Reproduced by permission.
5. Ciril Rebetez, Mireille Betrancourt, Mirweis Sangin, and Pierre Dillenbourg, "Learning from Animation Enabled by Collaboration," *Instructional Science*, V 35, No. 5, pp. 471–485, September 2010.
6. Russell R. Pate, Chia-Yih Wang, Marsha Dowda, Stephen W. Farrell, and Jennifer R. O'Neill, "Cardiorespiratory Fitness Levels Among U.S. Youth 12 to 19 Years of Age," *Archives of Pediatric Adolescent Medicine*, Vol. 160, October 2006, pp. 1005–1011.
7. Based on data from "Average Salary of Full Time Instructional Faculty on 9-Month Contracts in Degree-granting Institutions," *Digest of Educational Statistics*, http://nces.ed.gov/programs/digest/d09/tables/dt09_257.asp, 5 July 2010.
8. A. Tubb, A.J. Parker, and G. Nickless, "The Analysis of Romano-British Pottery by Atomic Absorption Spectrophotometry," *Archaeometry* 22 (1980):153.
9. "Smart Phone Ratings," <http://www.consumerreports.org/cro/electronics-computers/phones-mobile-devices/cell-phones-services/smart-phone-ratings/ratings-overview.htm>, 23 August 2010.
10. Roberta R. Bailey and Craig Idlebrook, "Save Money on Groceries," www.MotherEarthNews.com, August/September 2010.

Capítulo 8

1. Stellan Ohlsson, "The Learning Curve for Writing Books: Evidence from Professor Asimov," *Psychological Science* 3, no. 6 (1992):380–382.
2. Daniel C. Harris, *Quantitative Chemical Analysis*, 3rd ed. (New York: Freeman, 1991).
3. "2009–2010 Accountability Progress Reporting (APR)," <http://api.cde.ca.gov/AcntRpt2010/2010GrthAPICo.aspx?cYear=2009-10&cSelect=33>, RIVERSIDE, 23 September 2010.

4. Sarah Janssen, Ed., *The World Almanac and Book of Facts 2011* (New York, NY: Infobase Publishing), 2011.
5. Lawrence E. Levine and Victorina Wasmuth, "Laptops, Technology, and Algebra 1: A Case Study of an Experiment," *Mathematics Teacher* 97, no. 2 (February 2004):136.
6. "LCD TV Ratings & Reliability," <http://www.consumerreports.org/cro/electronics-computers/tvs-services/tvs/lcd-tv-ratings/ratings-overview.htm>, 13 July 2010.
7. "Drew Brees #9 QB," <http://sports.espn.go.com/nfl/players/gamelog?playerId=2580>, 22 March 2011.
8. W.B. Jeffries, H.K. Voris, and C.M. Yang, "Diversity and Distribution of the Pedunculate Barnacles *Octolasmis* Gray, 1825 Epizoic on the Scyllarid Lobster, *Thenus orientalis* (Lund, 1793)," *Crustaceana* 46, no. 3 (1984).
9. Gregory K. Torrey, S.F. Vasa, J.W. Maag, and J.J. Kramer, "Social Skills Interventions Across School Settings: Cast Study Reviews of Students with Mild Disabilities," *Psychology in the Schools* 29 (July 1992):248.
10. G. Wayne Marino, "Selected Mechanical Factors Associated with Acceleration in Ice Skating," *Research Quarterly for Exercise and Sport* 54, no. 3 (1983).
11. A.J. Ellis, "Geothermal Systems," *American Scientist*, September/October 1975.
12. "Ice Cream Nutritional Statement," *Coldstone Creamery*, http://www.coldstonecreamery.com/nutritional/nutrition_information.html, 27 September 2010.
13. Allen L. Shoemaker, "What's Normal? Temperature, Gender, and Heart Rate," *Journal of Statistics Education* (1996).
14. http://espn.go.com/mlb/stats/team/_/stat/batting/year/2010/seasontype/2, 5 November 2010.
15. David R. McAllister et al., "A Comparison of Preoperative Imaging Techniques for Predicting Patellar Tendon Graft Length before Cruciate Ligament Reconstruction," *The American Journal of Sports Medicine*, 20(4):461–465.
16. Henry Gleitman, *Basic Psychology*, 4th ed. (New York: Norton, 1996).
17. <http://www.the-numbers.com/charts/daily/201/20100923.php>, 27 September 2010.
18. "Lexus GX," http://en.wikipedia.org/wiki/Lexus_GX, 27 September 2010.
19. "Explore Our Menu," http://www.starbucks.com/menu/catalog/nutrition?paging=false&drink=all&page=2#view_control=nutrition, 27 September 2010.
20. *Automotive News: 1997 Market Data Book*, 28 May 1997, and Sarah Janssen, Ed., *The World Almanac and Book of Facts 2011* (New York, NY: Infobase Publishing), 2011.

Capítulo 9

1. W.S. Good, "Productivity in the Retail Grocery Trade," *Journal of Retailing* 60, no. 3 (1984).
2. "Digital Cameras," *Consumer Reports*, August 2010, p. 50.
3. Kimberly Pierceall, "S.B. Airport Plans for Big Future," *The Press-Enterprise* (Riverside, CA), 24 September 2010, p. D1.
4. "2009–2010 Accountability Progress Reporting (APR)," <http://api.cde.ca.gov/AcctRpt2010/2010GrthAPICo.aspx?cYear=2009-10&cSelect=33>, RIVERSIDE, 23 September 2010.
5. R. Blair and R. Miser, "Biotin Bioavailability from Protein Supplements and Cereal Grains for Growing Broiler Chickens," *International Journal of Vitamin and Nutrition Research* 59 (1989):55–58.
6. "2010–2011: Interpreting Your GRE Scores," http://www.ets.org/s/gre/pdf/gre_interpreting_scores.pdf, 16 September 2010.
7. "All Season Tire Ratings," *Consumer Reports*, www.consumerreports.org/cro/cars/tires-auto-parts/tires/all-season-tire-ratings/ratings-overview.htm, 28 September 2010.
8. "Tuna Goes Upscale," *Consumer Reports*, June 2001, p. 19.

Capítulo 10

1. Daniel Q. Haney, "Mondays May Be Hazardous," *The Press-Enterprise* (Riverside, CA), 17 November 1992, p. A16.
2. <http://www.pollingreport.com/health.htm>, *Newsweek Poll* conducted by Princeton Survey Research Associates International. 20–21 October, 2010.
3. Adapted from Linda Schmittroth, Ed., *Statistical Record of Women Worldwide* (Detroit and London: Gale Research, 1991).
4. Jason Malloy, "Gene Expression: NLSY Blogging: Eye and Hair Color of Americans," <http://www.gnyp.com/blog/2008/12/nlsy-blogging-eye-and-hair-color-of.php>, 31 December 2008.
5. Siobhan Reilly, Michele Abendstern, Jane Hughes, David Challis, Dan Venables, and Irene Pedersen, "Quality in Long-Term Care Home for People with Dementia: An Assessment of Specialist Provision," *Aging and Society*, 26(2006):649–668.
6. W.W. Menard, "Time, Chance and the Origin of Manganese Nodules," *American Scientist*, September/October, 1976.
7. Thomas Lord and Terri Orkwiszewski, "Moving from Didactic to Inquiry-Based Instruction in a Science Laboratory," *American Journal of Primatology*, 68 (October 2006).
8. Jonathan W. Jantz, C.D. Blosser, and L.A. Fruechting, "A Motor Milestone Change Noted with a Change in Sleep Position," *Archives of Pediatric Adolescent Medicine* 151 (June 1997):565.
9. "Next Generation of Americans," *CBS News Poll*, www.pollingreport.com, 17–22 December 2009.
10. Kim Marie McGoldrick, Gail Hoyt, and David Colander, "The Professional Development of Graduate Students for Teaching Activities: The Students' Perspective," *Journal of Economic Education*, 41(2): 194–201, 2010.
11. "Food Safety," *CBS News Poll*, www.pollingreport.com, 17–22 December 2009.
12. Sarah Janssen, Ed., *The World Almanac and Book of Facts 2011* (New York, NY: Infobase Publishing), 2011.
13. Doreen Matsui, R. Lim, T. Tschen, and M.J. Rieder, "Assessment of the Palatability of b-Lactamase-Resistant Antibiotics in Children," *Archives of Pediatric Adolescent Medicine* 151 (June 1997):599.
14. Andrew S. Levy, M.J. Wetzler, M. Lewars, and W. Laughlin, "Knee Injuries in Women Collegiate Rugby Players," *The American Journal of Sports Medicine* 25, no. 3 (1997):360.
15. Adapted from David L. Wheeler, "More Social Roles Means Fewer Colds," *Chronicle of Higher Education* XLIII, no. 44 (July 11, 1997):A13.

Capítulo 11

1. T.M. Casey, M.L. May, and K.R. Morgan, "Flight Energetics of Euglossine Bees in Relation to Morphology and Wing Stroke Frequency," *Journal of Experimental Biology* 116 (1985).
2. "Alzheimer's Test Set for New Memory Drug," *The Press-Enterprise* (Riverside, CA), 18 November 1997, p. A-4.
3. *Science News* 136 (August 1989):126.
4. D. Matsui et al., "Assessment of the Palatability of b-Lactamase-Resistant Antibiotics in Children," *Archives of Pediatric Adolescent Medicine* 151 (1997):559–601.
5. *Science News*, 1989, p. 116.
6. G. Merrington, L. Winder, and I. Green, "The Uptake of Cadmium and Zinc by the Birdcherry Oat Aphid *Rhopalosiphum padi* (Homoptera:Aphididae) Feeding on Wheat Grown on Sewage Sludge Amended Agricultural Soil," *Environmental Pollution* 96, no. 1 (1997):111–114.
7. Ron Marks, "Store brands save up to 52%," <http://www.consumerreports.org/cro/magazine-archive/2010/october/shopping/store-brands-vs-name-brands/store-brands-saved/index.htm>, October 2010.
8. Karola Sakekel, "Egg Substitutes Range in Quality," *San Francisco Chronicle*, 10 February 1993, p. 8.

Respuestas a ejercicios seleccionados

Capítulo 1

- 1.1** b. $\bar{x} = 2$; $m = 1$; moda = 1 c. sesgadas
- 1.3** a. 5.8 b. 5.5 c. 5 y 6
- 1.4** a. ligeramente sesgada a la derecha c. $\bar{x} = 1.08$; $m = 1$; moda = 1
- 1.6** 2.5 es un número promedio calculado (o estimado) para todas las familias de una categoría particular.
- 1.8** La mediana, porque la distribución está altamente sesgada a la derecha.
- 1.11** a. 2.4 b. 2.8 c. 1.673
- 1.13** a. 3 b. 2.125 c. $s^2 = 1.2679$; $s = 1.126$
- 1.15** a. 1.11 b. $s^2 = .19007$, $s = .436$ c. $R \approx 2.5s$
- 1.17** a. $s \approx 1.67$ b. $s = 1.75$ c. no d. sí e. no
- 1.19** a. aproximadamente .68 b. aproximadamente .95 c. aproximadamente .815 d. aproximadamente .16
- 1.21** a. $s \approx .20$ b. $\bar{x} = .76$; $s = .165$
- 1.23** a. aproximadamente .68 b. aproximadamente .95 c. aproximadamente .003
- 1.25** a. relativamente plana; $\bar{x} \approx 4.5$ b. ≈ 2.25 c. $\bar{x} = 4.586$; $s = 2.892$
- 1.27** a. sesgada a la derecha b. 0 a 104 días
- 1.29** b. $\bar{x} = 7.729$ c. $s = 1.985$

k	$\bar{x} \pm ks$	Actual	Chebyshev	Regla empírica
1	(5.744, 9.714)	.71	At least 0	Aprox. .68
2	(3.759, 11.699)	.96	At least 3/4	Aprox. .95
3	(1.774, 13.684)	1.00	At least 8/9	Aprox. .997

- 1.31** a. 42 b. $s \approx 10.5$ c. $s = 13.10$ d. 1.00; 1.00; sí
- 1.32** a. $s \approx .444$ b. $s = .436$
- 1.33** a–b. $\bar{x} = 1.4$; $s^2 = 1.4$
- 1.35** a. $\bar{x} = 2.04$; $s = 2.806$ b–c.

k	$\bar{x} \pm ks$	Actual	Chebyshev	Regla empírica
1	(- .766, 4.846)	.84	Al menos 0	Aprox. .68
2	(-3.572, 7.652)	.92	Al menos 3/4	Aprox. .95
3	(-6.378, 10.458)	1.00	Al menos 8/9	Aprox. .997

- 1.37** mín = 0, $Q_1 = 6$, $m = 10$, $Q_3 = 14$, máx = 19; IQR = 8
- 1.39** a. $Q_1 = .3125$; $Q_3 = .7150$ b. .4025 c. límites inferior y superior: -.29125 y 1.31875; no
- 1.41** cuartil superior e inferior: -2.25 y 15.75; $x = 22$ es un resultado atípico
- 1.43** a. mín = 1.70, $Q_1 = 130.5$, $m = 246.5$, $Q_3 = 317.5$, máx = 485 b. cuartil superior e inferior: -150 y 598 c–d. No, pero hay cuatro observaciones extremadamente pequeñas, no identificadas por la gráfica de caja como resultados atípicos.
- 1.45** a.
- | Variable | Minimum | Q1 | Median | Q3 | Maximum |
|----------|---------|-------|--------|-------|---------|
| Rodgers | 7.00 | 18.00 | 21.00 | 26.00 | 34.00 |
| Brees | 21.00 | 24.00 | 27.50 | 32.25 | 37.00 |
- b. Rodgers: límites inferior y superior: 6 y 38; sin resultados atípicos; relativamente simétrica. Brees: límites superior e inferior: 11.625 y 44.625; sin resultados atípicos; relativamente simétrica.
- 1.46** a. sesgada a la izquierda b. $\bar{x} = 108.15$; $m = 123.5$; media < mediana implica sesgada a la izquierda c. cuartil superior e inferior: -43.125 y 259.875; sesgada a la izquierda, no hay resultados atípicos.
- 1.47** Las temperaturas en mujeres tienen un centro (mediana) más alto y son más variables; tres resultados atípicos en el grupo de mujeres.
- 1.49** a. Genérica: $m = 26$, $Q_1 = 25$, $Q_3 = 27.25$, IQR = 2.25; Sunmaid: $m = 26$, $Q_1 = 24$, $Q_3 = 28$, IQR = 4 b. Genérica: cuartil superior e inferior: 21.625 y 30.625; Sunmaid: cuartil superior e inferior: 18 y 34 c. sí d. El tamaño promedio es casi igual; los tamaños individuales de pasas son más variables para pasas Sunmaid.

1.51 a. $R = 32.1$ b. $s \approx 8.025$ c. $s = 7.671$

1.53 $m = 6.35$, $Q_1 = 2.325$, $Q_3 = 12.825$; cuartil inferior y superior: -13.425 y 28.575 ; un resultado atípico ($x = 32.3$).

1.55 a–b.

k	$\bar{x} \pm ks$	Chebyshev	Regla empírica
1	(.16, .18)	Al menos 0	Aprox. .68
2	(.15, .19)	Al menos 3/4	Aprox. .95
3	(.14, .20)	Al menos 8/9	Aprox. .997

c. No, la distribución de $n = 4$ mediciones no puede ser de forma de montículo.

1.57 68%; 95%

1.59 a. 27; 20.2; 6.8 b. ligeramente sesgada a la izquierda c. 23.96; 1.641 d. máxima $x = 27$, puntaje $z = 1.85$; mínima $x = 20.2$, puntaje $z = -2.29$; no e. 24.3 f. 22.95 y 24.85

1.61 a. $s \approx 7.75$ b. $\bar{x} = 59.2$; $s = 10.369$
c. $m = 60$, $Q_1 = 51.25$, $Q_3 = 69.75$; cuartiles inferiores y superiores: 23.5 y 97.5; no hay resultados atípicos.

1.62 $\sigma \approx 100$

1.63 a. 16% b. 81.5%

1.65 a. .9735 b. .16

1.67 a. .025 b. .84

1.69 a. Al menos 3/4 tienen entre 145 y 205 profesores. b. .16

1.75 b. sí c. más de 2 o 3 desviaciones estándar desde la media

1.77 a. 2.5, 3.75, 4.2, 4.75, 5.7 b. cuartiles inferiores y superiores: 2.25 y 6.25 c. no d. en forma de montículo; sí

Capítulo 2

2.1 a. {1, 2, 3, 4, 5, 6} c. 1/6

d. $P(A) = 1/6$; $P(B) = 1/2$; $P(C) = 2/3$;
 $P(D) = 1/6$; $P(E) = 1/2$; $P(F) = 0$

2.3 $P(E_1) = .45$; $P(E_2) = .15$; $P(E_i) = .05$ para $i = 3, 4, \dots, 10$

2.5 a. {NDQ, NDH, NQH, DQH} b. 3/4
c. 3/4

2.9 a. .58 b. .14 c. .46

2.11 a. seleccionar al azar tres personas y registrar sus géneros b. {FFF, FMM, MFM, MMF, MFF, FMF, FFM, MMM} c. 1/8 d. 3/8 e. 1/8

2.13 a. ordenar A, B, C

b. {ABC, ACB, BAC, BCA, CAB, CBA}

d. 1/3, 1/3

2.15 a. .467 b. .513 c. .533

2.16 80

2.18 a. 60 b. 3 628 800 c. 720 d. 20

2.20 6720

2.22 216

2.24 120

2.25 720

2.26 a. 140 608 b. 132 600 c. .00037 d. .943

2.28 a. 2 598 960 b. 4 c. .000001539

2.30 $5.720645 \times (10^{12})$

2.32 1/56

2.35 a. 3/5 b. 4/5

2.37 a. 1 b. 1/5 c. 1/5

2.39 a. .05 b. sí c. no d. no

2.41 a. no; no b. no; sí

2.43 a. .08 b. .52

2.45 a. .3 b. no c. sí

2.47 a. no, porque $P(A \cap B) \neq 0$
b. no, porque $P(A) \neq P(A|B)$

2.48 a. .14 b. .56 c. .30

2.52 a. $P(A) = .9918$; $P(B) = .0082$

b. $P(A) = .9836$; $P(B) = .0164$

2.53 .05

2.55 a. .99 b. .01

5.57 a. 154/256 b. 155/256 c. 88/256

d. 88/154 e. 44/67 f. 23/35

g. 12/101 h. 189/256

2.59 a. .7225 b. .4712 c. .1043

2.60 a. .23 b. .6087; .3913

2.62 .38

2.63 .012

2.65 .999999

2.66 a. .3582 b. .4883 c. .4467

2.68 a. 1/8 b. 1/64 c. No necesariamente; podrían haber estudiado juntos, y así sucesivamente.

2.70 a. 5/6 b. 25/36 c. 11/36

2.71 a. .8 b. .64 c. .36

2.73 .0256; .1296

2.75 .2; .1

2.76 a. .5182 b. .1136 c. .7091
d. .3906

2.78 a.

x	0	1	2
$p(x)$	6/15	8/15	1/15

b. 1/15 c. $\mu = 2/3$; $\sigma^2 = 16/45$

Capítulo 3

- 3.2** a. continua b. continua
c. discreta d. discreta e. continua
- 3.4** a. .2 c. $\mu = 1.9$; $\sigma^2 = 1.29$; $\sigma = 1.136$
d. .3 e. .9
- 3.5** 1.5
- 3.7** a. $p(x) = C_x^3(.47)^x(.53)^{3-x}$
c. .396 d. $\mu = 1.41$; $\sigma = .864$
- 3.8** a. $p(0) = 3/10$; $p(1) = 6/10$; $p(2) = 1/10$
- 3.9** a. .1; .09; .081 b. $p(x) = (.9)^{x-1}(.1)$
- 3.11** a. 4.0656 b. 4.125 c. 3.3186
d. $E(x)$ disminuye conforme $P(A)$ se incrementa
- 3.13** a. .28 b. .18 c. $\mu = 1.32$; $\sigma = 1.199$
d. .94
- 3.14** \$20,500
- 3.15** a. .058 b. .989 c. .011 d. .047
e. .437
- 3.17** a. .2965 b. .8145 c. .1172
d. .3670
- 3.19** a. .097 b. .329 c. .671 d. 2.1
e. 1.212
- 3.21** $p(0) = .000$; $p(1) = .002$; $p(2) = .015$;
 $p(3) = .082$; $p(4) = .246$; $p(5) = .393$;
 $p(6) = .262$
- 3.23** a. .251 b. .618 c. .367 d. .633
e. 4 f. 1.549
- 3.25** a. .901 b. .015 c. .002 d. .998
- 3.27** a. .748 b. .610 c. .367 d. .966
e. .656
- 3.29** a. 1; .99 b. 90; 3 c. 30; 4.58
d. 70; 4.58 e. 50; 5
- 3.31** a. .9568 b. .957 c. .9569
d. $\mu = 2$; $\sigma = 1.342$ e. .7455; .9569; .9977
f. sí; sí
- 3.33** binomial; $n = 2$; $p = .6$
- 3.35** no; la variable no es el número de éxitos en n intentos.
En cambio, el número n de intentos es variable.
- 3.37** a. 1.000 b. .997 c. .086
- 3.39** a. .098 b. .991 c. .098 d. .138
e. .430 f. .902
- 3.40** a. .0081 b. .4116 c. .2401
- 3.42** a. $\mu = 10$ b. 4 a 16 c. Si este improbable
valor se observara en realidad, podría ser que los
intentos (campos) no sean independientes.
- 3.44** a. .016796 c. .98320
- 3.45** a. .107 b. .762
- 3.47** a. .082085 b. .205212 c. .256516
d. .543813
- 3.49** a. .647 b. .353 c. .224 d. .493
- 3.51** a. .135335 b. .27067 c. .593994
d. .036089
- 3.53** a. .677 b. .6767 c. sí
- 3.55** a. .0067 b. .1755 c. .560
- 3.57** a. .271 b. .594 c. .406
- 3.59** $P(x > 5) = .017$; improbable
- 3.61** a. 2/3 b. 1/15 c. 1/2
- 3.63** a. .6 b. .5143 c. .0714
- 3.65** a. $p(0) = .36$; $p(1) = .48$; $p(2) = .15$; $p(3) = .01$
c. $\mu = .8$, $\sigma^2 = .50286$ d. .99; .99; sí
- 3.66** $p(0) = .2$; $p(1) = .6$; $p(2) = .2$
- 3.68** a. hipergeométrica b. .1786 c. .01786
d. .2857
- 3.70** a. .9772 b. .1230 c. .9802 d. .9699
- 3.72** a. .9452 b. .9664 c. .8159 d. ≈ 1.0000
- 3.74** a. .6753 b. .2401 c. .2694 d. .0901
e. ≈ 0
- 3.75** a. 1.96 b. 1.44
- 3.78** a. 1.65 b. -1.645
- 3.80** a. 1.28 b. 1.645 c. 2.05 d. 2.33
- 3.82** a. .1596 b. .1151 c. .1359
- 3.84** 58.3
- 3.86** $\mu = 8$; $\sigma = 2$
- 3.87** a. .4586 b. .0526 c. .0170
- 3.89** .1562; .0012
- 3.91** a. .0475 b. .00226 c. 29.12 a 40.88
d. 38.84
- 3.93** a. .9938 b. .0301
- 3.94** .0475
- 3.96** 63 550
- 3.97** a. .3085 b. .2417 c. .0045
- 3.99** a. sí b. 15; 2.449 c. .9878
- 3.101** a. sí b. $\mu = 7.5$; $\sigma = 2.291$ c. .6156
d. .618
- 3.103** a. .2676 b. .3520 c. .3208 (use la regla de
pares e impares para redondear) d. .9162
- 3.105** a. .178 b. .392
- 3.107** a. .245 b. .2483
- 3.109** a. .0006 b. .3050 c. .5675
- 3.111** .9441
- 3.114** a. 31 b. 3.432 c. no; $x = 25$ está a sólo 1.75
desviaciones estándar debajo de la media.

- 3.122** a. .228 b. no es indicación de que sea más probable que las personas escojan números de en medio
- 3.123** a. 20 b. 4 c. .006 d. El psiquiatra está equivocado
- 3.124** a. $\mu = 50$; $\sigma = 6.124$ b. El valor $x = 35$ está 2.45 desviaciones estándar debajo de la media. Es un tanto improbable que la cifra de 25% sea representativa de este plantel.
- 3.126** a. .5 b. $\mu = 12.5$; $\sigma = 2.5$ c. Hay preferencia por el segundo diseño.
- 3.128** a. sí; $n = 10$; $p = .25$ b. .2440 c. .0000296
d. Sí; el modelo genético no se comporta como se esperaba.
- 3.130** a. sí b. $1/8192 = .00012$
- 3.132** a. $p(x) = \frac{C_x^M C_{10-x}^{50-M}}{C_{10}^{50}}$ donde $M = \#$ de piezas defectuosas en la caja y $x = \#$ de piezas defectuosas en la muestra. b. .6367 c. .3968; .2415
- 3.134** a. .015625 b. .421875 c. .25
- 3.135** a. $p = 1/3$ b. .3292 c. .8683
- 3.137** a. .135335 b. .676676
- 3.138** .655
- 3.139** a. .794 b. .056 c. -0.82 a 3.82 o 0 a 3
- 3.140** a. 36 b. 4.8 c. Sí, porque $x = 49$ está a 2.71 desviaciones estándar arriba de la media.
- 3.142** a. 240 b. 9.798 c. 221 a 259
d. $x = 200$ se encuentra a más de 4 desviaciones estándar debajo de la media; quizá la cifra de 60% es demasiado alta.
- 3.147** .0713
- 3.149** $P(A) = 1/2$; $P(B) = 2/3$; $P(A \cap B) = 1/3$;
 $P(A \cup B) = 5/6$; $P(C) = 1/6$; $P(A \cap C) = 0$;
 $P(A \cup C) = 2/3$
- 3.152** a. .9544 b. .0561
- 3.154** a. $z_0 = -1.96$ b. $z_0 = .36$
- 3.156** a. .9651 b. .1056 c. .0062
- 3.158** a. .8849 b. .1841 c. .9279 d. .3372
- 3.160** a. .7734 b. .9115
- 3.162** .8612
- 3.164** a. .1056 b. .8944 c. .1056
- 3.165** .16
- 3.166** .0344
- 3.168** .3859
- 3.171** 383.5 horas
- 3.172** .8980

- 3.173** a. binomial, $n = 100$, $p = .75$ b. sí
c. .7960 d. ≈ 0

3.175 no; para $x = 19$, $z = 1.461$

3.177 87.48

3.179 a. .9107 b. 35.812 y 43.848 c. .921; .953

3.181 .0446

3.182 a. 1.27 b. .1020

3.183 .1251

Capítulo 4

4.1 1/500

4.8 a. $\mu = 10$; $\sigma/\sqrt{n} = .5$

b. $\mu = 5$; $\sigma/\sqrt{n} = .2$

c. $\mu = 120$; $\sigma/\sqrt{n} = .3536$

4.10 c. aproximadamente en forma de montículo

4.12 a. 1 b. .707 c. .500 d. .333 e. .250

f. .200 g. .100

4.14 a. aproximadamente normal b. 53; 3

4.16 a. aproximadamente normal b. 100; 3.16

4.18 a. 106; 2.4 b. .0475 c. .9050

4.19 b. un número grande de réplicas

4.23 a. 1266; 22.517 b. .0655

4.24 a. ≈ 0 b. sí; el valor $\bar{x} = 98.25$ está casi 5 desviaciones estándar debajo de la media supuesta, $\mu = 98.6$

4.26 a. $p = .3$; SE = .0458 b. $p = .1$; SE = .015

c. $p = .6$; SE = .0310

4.28 a. .7019 b. .5125

4.30 a. .0099 b. .03 c. .0458 d. .05

e. .0458 f. .03 g. .0099

4.32 a. aproximadamente normal b. .25; .0484
c. .9265

4.34 a. sí; $\mu = .78$ y $\sigma = .0414$ b. .2358

c. .2090 d. $z = -3.14$; quizá p es menor a .78.

4.35 a. aproximadamente normal con media .75 y desviación estándar .0306 b. .0516 c. .69 a .81

4.37 a. LCL = 150.13; UCL = 161.67

4.39 a. LCL = 0; UCL = .090

4.41 a. LCL = 8598.7; UCL = 12 905.3

4.43 LCL = .078; UCL = .316

4.45 LCL = .0155; UCL = .0357

4.47 media demasiado grande a las horas 2, 3 y 4

4.51 a. .4938 b. .0062 c. .0000

4.53 a. ≈ 12.5 b. .9986 c. Es probable que sean correctas

4.55 c. no

4.59 a. muestra de conglomerado b. muestra sistemática de 1 en 10 c. muestra estratificada
d. muestra sistemática de 1 en 10
e. muestra aleatoria simple

4.61 a. 131.2; 3.677 b. sí c. .1515

4.63 a. LCL = 0; UCL = .0848 b. $\hat{p} > .0848$

4.65 sí

4.68 a. aproximadamente normal con media de 288 y desviación estándar .9798 b. .0207 c. .0071

4.70 UCL = .2273; LCL = 0

Capítulo 5

5.3 a. .160 b. .339 c. .438

5.5 a. .554 b. .175 c. .055

5.7 a. .179 b. .098 c. .049 d. .031

5.9 a. .0588 b. .0898 c. .098 d. .0898
e. .0588 f. $p = .5$

5.11 $\hat{p} = .728$; margen de error (MOE) = .029

5.13 $\hat{p} = .90$; MOE = .0263

5.15 $\bar{x} = 39.8$; MOE = 4.768

5.17 $\bar{x} = 7.2\%$; MOE = .776

5.18 a. $\hat{p} = .75$; MOE = .0268

b. $1.96 \sqrt{\frac{.5(.5)}{1004}} = .031$; no, a menos que el sondeo

haya sido redondeado al siguiente medio punto porcentual

5.19 La estimación puntual es $\bar{x} = 19.3$ con margen de error = 1.86

5.21 a. (.797, .883) b. (21.469, 22.331) c. Los intervalos construidos de este modo encierran el verdadero valor de μ 90% del tiempo en muestreo repetido

5.23 (.846, .908)

5.25 a. 3.92 b. 2.772 c. 1.96

5.27 a. 3.29 b. 5.16 c. El ancho aumenta.

5.29 (3.496, 3.904); muestra aleatoria

5.30 a. (.106, .166) b. Aumente el tamaño muestral o disminuya el nivel de confianza o ambas cosas

5.31 a. $98.085 < \mu < 98.415$ b. no; quizá el valor 98.6 no es el verdadero promedio de temperatura corporal para personas sanas

5.32 a. (4.61, 5.99) b. sí

5.34 a. (-.77, 3.77) b. no

5.36 (15.463, 36.937)

5.38 a. (17.676, 19.324) b. (15.710, 17.290)

c. (.858, 3.142) d. sí

5.40 a. $\bar{x}_1 - \bar{x}_2 = 5545$; MOE = 902.08 b. sí

5.43 a. (-.528, -.032); sí, porque $\mu_1 - \mu_2 = 0$ no está en el intervalo.

5.45 a. (-.203, -.117) b. muestras aleatorias e independientes de distribuciones binomiales

5.48 a. (-.118, -.002) b. Sí, porque $p_1 - p_2 = 0$ no está en el intervalo

5.49 a. (.095, .445) b. sí

5.50 (.061, .259)

5.51 a. (-.082, .022) b. No, porque $p_1 - p_2 = 0$ no está en el intervalo.

5.52 a. $\mu < 76.63$ b. $\mu < 1.89$

5.54 $\mu_1 - \mu_2 < 4$

5.56 505

5.58 $n_1 = n_2 = 1086$

5.59 b. 9604

5.60 $n_1 = n_2 = 360$

5.63 $n_1 = n_2 = 136$

5.64 $n_1 = n_2 = 98$

5.66 a. $\bar{x} = 29.1$; MOE = .9555 b. (28.298, 29.902) c. $\mu > 28.48$ d. 234

5.68 $n_1 = n_2 = 224$

5.70 1083

5.72 $n_1 = n_2 = 925$

5.74 (8.087, 11.313)

5.77 (33.41, 34.59)

5.79 al menos 1825

5.81 .3874; .651

5.83 a. (2.837, 3.087) b. 276

5.85 (2.694, 2.716)

5.87 (.161, .239)

5.89 al menos 97

Capítulo 6

6.1 a. 2.015 b. 2.306 c. 1.330 d. 1.96

6.3 a. $.02 < \text{valor } p < .05$ b. valor $p < .005$
c. valor $p > .20$ d. valor $p < .005$

6.5 a. $\bar{x} = 7.05$; $s = .4994$ b. 7.496
c. Rechazar H_0 ; $t = -2.849$ d. Sí.

6.7 no; $t = -1.195$

6.9 a. sí; $t = -3.044$ b. 98.316; sí

6.10 (3.652, 3.912)

6.12 a. Rechazar H_0 ; $t = -4.31$. b. (23.23, 29.97) c. La media de tratamiento previo se ve más pequeña que las otras dos medias

- 6.16** (233.98, 259.94)
- 6.18** a. 3.775 b. 21.2258
- 6.20** a. $H_0: \mu_1 - \mu_2 = 0$; $H_a: \mu_1 - \mu_2 \neq 0$
 b. $|t| > 2.771$ c. $t = 2.795$ d. valor $p < .01$
 e. Rechazar H_0
- 6.22** a. sí; s^2 mayor/ s^2 menor = 1.36
 b. $t = .06$ con valor $p = .95$
 c. 19.1844 d. No rechazar H_0 .
 e. (-5.223, 5.503); sí
- 6.24** a. $H_0: \mu_1 - \mu_2 = 0$ contra $H_a: \mu_1 - \mu_2 > 0$
 b. sí; $t = 2.806$ c. $.005 < \text{valor } p < .01$
- 6.26** a. No rechazar H_0 ; $t = 1.92$ b. (-3.32, 4.72)
 c. no agrupada
- 6.28** a. No rechazar H_0 ; $t = -1.68$
 b. (-.0029, .0003); sí
- 6.30** a. no; no rechazar H_0 ; $t = 1.606$
 b. (-.061, .341)
- 6.32** $\mu_1 - \mu_2 > -.118$; sí
- 6.33** no; $t = .79$
- 6.35** a. $H_0: \mu_1 - \mu_2 = 0$ contra $H_a: \mu_1 - \mu_2 > 0$
 b. No rechazar H_0 ; $t = 1.511$
- 6.39** análisis en pares
- 6.41** no; $t = 2.29$
- 6.43** sí; $t = 11.32$
- 6.44** no; $\chi^2 = 34.24$
- 6.46** a. $s^2 = .6990476$ b. (.291, 3.390)
 c. no rechazar H_0 ; $\chi^2 = 5.24$ d. valor $p > .20$
- 6.48** a. no; $t = -.232$ b. sí; $\chi^2 = 20.18$
- 6.50** a. no b. sí; $z = 3.262$
- 6.52** no; $\chi^2 = 29.433$
- 6.54** (.667, 4.896)
- 6.56** $F = 1.22$ con valor $p > .20$; no rechazar H_0 ; $= \sigma_1^2 = \sigma_2^2$
- 6.58** a. no; $F = 2.66$ b. sí
- 6.60** Resto: $F = 1.03$ con valor $p > .20$; 80% máximo O_2 ; $F = 2.01$ con valor $p > .20$; máximo O_2 ; $F = 14.29$ con valor $p < .01$; use la prueba t no agrupada para un máximo de O_2
- 6.66** a. valor $p > .10$ b. $.05 < \text{valor } p < .10$
 c. $.005 < \text{valor } p < .01$ d. valor $p > .10$
- 6.68** sí; $t = 2.108$; $.025 < \text{valor } p < .05$
- 6.70** (56.223, 99.303)
- 6.72** (3.545, 4.975)
- 6.74** a. (169.1, 199.9) b. (69.43, 76.57)
 c. (2.28, 2.80) d. no
- 6.76** no; $t = -1.49$
- 6.78** a. sí; $t = -3.354$ b. valor $p < .01$
 c. (-10.246, -2.354); sí
- 6.80** a. Rechazar H_0 ; $\sigma_1^2 = \sigma_2^2$; $F = 3.88$
 b. Rechazar H_0 ; $\mu_1 - \mu_2 = 300$; $t = 2.13$; hay evidencia suficiente para indicar que $(\mu_1 - \mu_2) > 300$
- 6.82** a. (.02698, .02808)
- 6.84** a. no; $F = 2.21$ b. (.975, 5.03); sí
- 6.86** (35.845, 48.405)
- 6.87** a. Es válida la suposición de normalidad.
 b. (5.12, 5.67)
- 6.89** sí; $t = -2.39$
- 6.91** a. $t = 9.5641$ con valor $p = .0000$; hay evidencia suficiente para indicar una diferencia en el promedio de resistencias
- 6.94** a. (-11.414, -8.958) b. sí
- 6.96** a. no; $F = 1.922$ con valor $p > .20$
 b. (.452, 8.1685)
- 6.97** no; $t = 3.038$
- 6.99** (24.582, 73.243)
- 6.101** a. sí; $\chi^2 = 24.73$ b. (.0284, .1318)
- 6.103** (3.873, 4.519)
- 6.104** a. sí; $F = 1.21$ b. $t = 1.65$; hay evidencia insuficiente para indicar una diferencia en las dos medias poblacionales

Capítulo 7

7.1 Fuente	gl
Tratamientos	5
Error	54
Total	59

7.3 a. (2.731, 3.409) b. (.07, 1.03)

7.5 a. Fuente	gl	SS	MS	F
Tratamientos	3	339.8	113.267	16.98
Error	20	133.4	6.67	
Total	23			

b. $gl_1 = 3$ y $gl_2 = 20$ c. $F > 3.10$
 d. sí, $F = 16.98$ e. valor $p < .005$; sí

7.7 a. CM = 103.142857; SS Total = 26.8571
 b. SST = 14.5071; MST = 7.2536
 c. SSE = 12.3500; MSE = 1.1227

d. Analysis of Variance					
Source	DF	SS	MS	F	P
Trts	2	14.51	7.25	6.46	0.014
Error	11	12.35	1.12		
Total	13	26.86			

f. $F = 6.46$; rechazar H_0 con $.01 < \text{valor } p < .025$

7.9 a. (1.95, 3.65) b. (.27, 2.83)

- 7.11** a. (67.86, 84.14) b. (55.82, 76.84)
c. (-3.629, 22.963) d. No, no son independientes

- 7.13** a. Cada observación es la longitud media de 10 hojas b. sí, $F = 57.38$ con valor $p = .000$
c. Rechazar H_0 ; $t = 12.09$ d. (1.810, 2.924)

7.15

Analysis of Variance for Percent

Source	DF	SS	MS	F	P
Method	2	0.0000041	0.0000021	16.38	0.000
Error	12	0.0000015	0.0000001		
Total	14	0.0000056			

- 7.17** a. diseño completamente aleatorizado

b.

Source	DF	SS	MS	F	P
State	3	3272.2	1090.73	26.44	0.000
Error	16	660.0	41.25		
Total	19	3932.2			

- c. $F = 26.44$; rechazar H_0 con valor $p = .005$.

- 7.19** Las medias muestrales deben ser independientes; tamaños muestrales iguales

- 7.21** a. 1.878s b. 2.1567s

- 7.23** $\bar{x}_1$ $\bar{x}_2$ $\bar{x}_3$ $\bar{x}_4$

- 7.25** a. no; $F = .60$ con valor $p = .562$ b. no hay diferencias

- 7.27** a. sí; $F = 4.47$, valor $p < .05$ b. (-128.946, 38.946) c. $\bar{x}_{SS}$ $\bar{x}_{LS}$ $\bar{x}_{PS}$

Fuente	gl	SS	MS	F
Tratamientos	2	11.4	5.70	4.01
Bloques	5	17.1	3.42	2.41
Error	10	14.2	1.42	
Total	17	42.7		

- 7.31** (-3.833, -7.767)

- 7.33** a. sí; $F = 19.19$ b. sí; $F = 135.75$

- c. $\bar{x}_1$ $\bar{x}_3$ $\bar{x}_4$ $\bar{x}_2$ d. (-5.332, -2.668) e. sí

- 7.35** a. 7 b. 7 c. 5 e. sí; $F = 9.68$

- f. sí; $F = 8.59$

7.37

ANOVA de dos vías y contra bloques, productos químicos

Analysis of Variance for y

Source	DF	SS	MS	F	P
Blocks	2	7.1717	3.5858	40.21	0.000
Chemical	3	5.2000	1.7333	19.44	0.002
Error	6	0.5350	0.0892		
Total	11	12.9067			

- 7.39** a. sí; $F = 10.06$ b. sí; $F = 10.88$
c. $\omega = 2.98$; las preparaciones 1 y 3 no son significativamente diferentes d. (1.12, 5.88)

7.41

ANOVA de dos vías: costo contra estimador, trabajo

Analysis of Variance for Cost

Source	DF	SS	MS	F	P
Estimator	2	10.862	5.431	7.20	0.025
Job	3	37.607	12.536	16.61	0.003
Error	6	4.528	0.755		
Total	11	52.997			

- 7.43** a. Los bloques son artículos; los tratamientos son tiendas. b. sí, $F = 25.53$; valor $p = .000$
c. sí, $F = 29.99$; valor $p = .000$

Fuente	gl
A	3
B	4
AB	12
Error	40
Total	59

- 7.47** (-1.11, 5.11)

- 7.49** a. hay fuerte interacción

- b. $F = 37.85$ con valor $p = .000$; sí d. no

- 7.50** b. sí c. Como la interacción es significativa, la atención debe concentrarse en medias para las combinaciones individuales de nivel de factor
d. Capacitación: $.05 < \text{valor } p < .10$; capacidad: valor $p < .005$; interacción: $.01 < \text{valor } p < .025$

- 7.52** a. factorial de 2×4 ; estudiantes, género en los dos niveles, escuelas a cuatro niveles
c. no; $F = 1.19$ e. El principal efecto para escuelas es significativo; $F_{\hat{y}} = 27.75$; ω de Tukey = 82.63

- 7.54** a.

Source	DF	SS	MS	F	P
Training	1	4489.00	4489.00	117.49	0.000
Situation	1	132.25	132.25	3.46	0.087
Interaction	1	56.25	56.25	1.47	0.248
Error	12	458.50	38.21		
Total	15	5136.00			

- b. No. $F = 1.47$; valor $p = .248$.

- c. no; $F = 3.46$; valor $p = .087$.

- d. sí; $F = 117.49$; valor $p = .000$.

- 7.56** diferencias significativas entre los tratamientos A y C, B y C, C y E, y D y E

- 7.58** a. diferencia significativa en medias de tratamiento; $F = 27.776$ b. ω de Tukey = .190
c. sí; $F = 6.588$

7.60

ANOVA de una vía: ventas contra programa

Analysis of Variance for Sales

Source	DF	SS	MS	F	P
Program	3	1385.8	461.9	9.84	0.000
Error	23	1079.4	46.9		
Total	26	2465.2			

- 7.62** a. no; $F = 1.40$ b. valor $p > .10$
 c. sí; $F = 6.51$ d. sí; $F = 7.37$

- 7.64** a. experimento factorial de 2×3
 b. no; $F = .452$ con valor $p = .642$
 d. $(-22.56, -5.84)$

- 7.66** a. diseño de bloques aleatorizado (o diferencia en pares)

b.

ANOVA de dos vías: total contra semana, tienda

Source	DF	SS	MS	F	P
Week	3	139.708	46.569	0.74	0.593
Store	1	562.298	562.298	8.99	0.058
Error	3	187.654	62.551		
Total	7	889.660			

- c. no, $F = 8.99$ con valor $p = .058$

- 7.68** a. experimento factorial b. sí; $F = 7.61$
 c. $\omega = 2.67$

- 7.70** a. diseño completamente aleatorizado

- b. Sí, hay una diferencia significativa.
 $F = 126.85$, valor $p = .000$

Source	DF	SS	MS	F	P
Site	2	132.277	66.139	126.85	0.000
Error	21	10.950	0.521		
Total	23	143.227			

- 7.72** No hay evidencia de no normalidad. Parece haber una diferencia en la variabilidad dentro de algunas de las combinaciones de niveles de factores

Capítulo 8

- 8.1** intersección con eje $y = 1$, pendiente = 2

- 8.3** $y = 3 - x$

- 8.7** a. $\hat{y} = 6.00 - .557x$ c. 4.05

- d. Analysis of Variance

Source	DF	SS	MS
Regression	1	5.4321	5.4321
Residual Error	4	0.1429	0.0357
Total	5	5.5750	

- 8.9** a. $\hat{y} = 195.90 + 67x$

c. Source	df	SS	MS
Regression	1	43146.9296	43146.9296
Error	3	1860.2704	620.0901
Total	4	45007.2000	

- 8.11** a. 10 b. 9

- c. **Análisis de regresión: y contra x**

The regression equation is

$$y = 3.00 + 0.475x$$

Predictor	Coef	SE Coef	T	P
Constant	3.000	2.127	1.41	0.196
x	0.4750	0.1253	3.79	0.005

$S = 2.24165$ $R\text{-Sq} = 64.2\%$ $R\text{-Sq(adjust)} = 59.8\%$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	72.200	72.200	14.37	0.005
Residual Error	8	40.200	5.025		
Total	9	112.400			

- d. $\hat{y} = 3.00 + .475x$ e. 7.75

- 8.13** a. $y = \text{API}; x = \text{EL}$ b. sí
 $\hat{y} = 847.468 - 1.7045x$ d. sí

- 8.15** a. sí b. $= -11.665 + .755x$ c. $\hat{y} = 52.51$

- 8.17** a. fuerte relación lineal positiva
 b. aproximadamente 1
 c. $\hat{y} = 12.221 + .815x$ d. $\hat{y} = 62.75$

- 8.19** a. sí, $t = 5.20$ b. $F = 27.00$
 c. $t_{.025} = 3.182$; $F_{.05} = 10.13$

- 8.21** a. sí, $F = 152.10$ con valor $p = .000$
 b. $r^2 = .974$

- 8.23** a. $\hat{y} = -.257 + .215x$ b. sí; $t = 5.90$
 c. $r^2 = .813$

- 8.25** a. No. Rechazar H_0 ; $t = 8.34$ con valor $p < .005$
 b. $r^2 = .959$ c. el patrón indica que la relación puede ser curvilínea

- 8.27** a. $MSE = .08333$ b. sí; $t = -12.124$
 c. $r^2 = .98$ d. la variación total se ha reducido en 98%

- 8.29** a. $r^2 = .757$ b. 75.7%

- 8.31** gráfica de residuales de probabilidad normal; los puntos deberán aproximar una línea recta con pendiente hacia arriba

- 8.33** gráfica de residuales contra ajustes; dispersión aleatoria de puntos, sin patrones

- 8.35** no hay violaciones extremas de las suposiciones de regresión

- 8.37** b. $MSE = 58.1$ c. ligera desviación de la normalidad; posiblemente una observación inusual, pero no hay violaciones extremas de las suposiciones de regresión

- 8.39** a. (3.259, 5.141) b. (2.24, 6.16)

- 8.41** a. $\hat{y} = 4.3 + 1.5x$ c. $s^2 = 1.53$
 d. sí, $t = 3.83$ e. valor $p < .01$ f. no hay violaciones de las suposiciones de regresión
 g. (8.11, 9.49)

- 8.43** a. (21.613, 33.199) b. (304 676, 307 360)
 c. 167.739 d. (295 826, 304 151)

- 8.45** a. $\hat{y} = 156.13 + 4.844x$ b. $r^2 = .163$ c. no hay violaciones obvias

- 8.49** a. $r = 1$ b. $r = -1$

- 8.51** a. $r = -.982$ c. 96.47%

- 8.53** a. correlacionado negativamente
 b. $H_a: r < 0$ c. Rechazar H_0 ; $t = -1.872$

- 8.55** a. sí, $t = -3.260$ b. valor $p < .01$

- 8.57** no, $t = .92$ con valor $p > .10$

8.59 a. $r = .1741$ b. no, $t = .559$

8.61 a. $\hat{y} = 46 - .317x$

c. Source	df	SS	MS
Regression	1	601.6667	601.6667
Error	10	190.3333	19.0333
Total	11	792.0000	

d. ligeras irregularidades e. $(-.442, -.192)$

f. $(27.09, 33.24)$ g. $(19.97, 40.36)$

8.63 Las respuestas variarán

8.65 a. los tres son significativos b. radiografías, resonancia magnética tridimensional, resonancia magnética estándar (usando coeficiente de determinación)
c. conclusiones consistentes (no pueden diferenciar entre radiografías y resonancia magnética tridimensional)

8.67 Las respuestas variarán.

8.71 Sí; $r = .562$ con valor $p = .036$

8.75 a. $\hat{y} = 7 + 15.4x$

b. Source	df	SS	MS
Regression	1	2371.6	2371.6
Error	6	50.4	8.4
Total	7	2422.0	

c. sí; $t = 16.80$ d. $(42.99, 48.01)$

e. $(13.16, 17.64)$ f. $r^2 = .979$

8.77 a. curvilínea

b. $\hat{y} = 2\,309\,189.75 - 1140.595x$

c. no, $t = -.56$ d. se ha ajustado el modelo equivocado

Capítulo 9

9.1 b. rectas paralelas

9.3 a. sí, $F = 57.44$ con valor $p < .005$ b. $R^2 = .94$

9.5 a. cuadrático b. $R^2 = .815$; relativamente buen ajuste c. sí, $F = 37.37$ con valor $p = .000$

9.7 a. $b_0 = 10.5638$ b. sí, $t = 15.20$ con valor $p = .000$

9.9 b. $t = -8.11$ con valor $p = .000$; rechazar H_0 ; $\beta_2 = 0$ a favor de H_a ; $\beta_2 < 0$

9.11 a. $R^2 = .9955$ b. $R^2(\text{adj}) = 99.25\%$
c. El modelo cuadrático ajusta ligeramente mejor

9.13 a. Usar variables x_1 , x_3 y x_5 b. no

9.14 a. $\hat{y} = -8.177 + 292x_1 + 4.434x_2$
b. Rechazar H_0 , $F = 16.28$ con valor $p = .002$. El modelo contribuye con información significativa para la predicción de y c. sí, $t = 5.54$ con valor $p = .001$ d. $R^2 = .823$; 82.3%

9.16 a. cuantitativa b. cuantitativa c. cualitativa; $x_1 = 1$ si la planta B, 0 de otro modo; $x_2 = 1$ si la planta C, 0 de otro modo d. cuantitativa
e. cualitativa; $x_1 = 1$ si turno de día, 0 si turno de noche

9.18 a. x_2 b. $\hat{y} = 12.6 + 3.9x_2^2$ o $\hat{y} = 13.14 - 1.2x_2 + 3.9x_2^2$

9.20 a. $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \epsilon$ con $x_2 = 1$ si pepino, 0 si algodón c. No, la prueba para interacción da $t = .63$ con valor $p = .533$ d. sí

9.22 $y = \beta_0 + \beta_1x_1 + \beta_2x_1^2 + \beta_3x_2 + \beta_4x_1x_2 + \beta_5x_1^2x_2 + \epsilon$

9.24 a. $\hat{y} = 8.585 + 3.8208x - 0.21663x^2$
b. $R^2 = .944$ c. sí; $F = 33.44$ d. sí; $t = -4.93$ con valor $p = .008$ e. no

9.26 b. $\hat{y} = 4.10 + 1.04x_1 + 3.53x_2 + 4.76x_3 - 0.43x_1x_2 - 0.08x_1x_3$ c. sí; $t = -2.61$ con valor $p = .028$
d. no; $F = 3.86$; considere eliminar los términos de interacción.

9.28 a. $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1^2 + \beta_4x_1x_2 + \beta_5x_1^2x_2 + \epsilon$

b. $F = 25.85$; $R^2 = .768$

c. $\hat{y} = 4.51 + 6.394x_1 + .1318$

d. $\hat{y} = -46.34 + 23.458x_1 - .3707$

e. no; $t = .78$ con valor $p = .439$

9.30 a. sí, precio y puntuación general están correlacionados; y está correlacionada con x_1 y x_3
b. $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + \epsilon$
c. $R^2 = .471$ y $R^2(\text{adj}) = .339$; no
d. x_1 y x_3 ; $y = b_0 + b_1x_1 + b_2x_3 + \epsilon$; $R^2 = .468$ y $R^2(\text{adj}) = .421$; ligeramente mejor que el modelo completo

9.34 a. 99.85% b. sí; $F = 1676.61$ con valor $p = .000$ c. sí; $t = -2.65$ con valor $p = .045$
d. sí; $t = 15.14$ con valor $p = .000$
e. Lineal: $R^2(\text{adj}) = 91.87\%$, cuadrático: $R^2(\text{adj}) = 99.79\%$, el término cuadrático es significativo
f. El término cuadrático falta

Capítulo 10

10.3 a. $\chi^2 > 12.59$ b. $\chi^2 > 21.666$

10.5 a. $H_0: p_1 = p_2 = p_3 = p_4 = p_5 = 1/5$

b. 4 c. 9.4877 d. $\chi^2 = 8.00$

e. No rechazar H_0

10.7 Sí, $\chi^2 = 24.48$; los automovilistas tienden a preferir los carriles interiores

- 10.9** no, $\chi^2 = 3.63$
- 10.11** no, $\chi^2 = 13.58$
- 10.13** sí; rechazar H_0 ; $\chi^2 = 28.386$
- 10.15** 8
- 10.17** Rechazar H_0 ; $\chi^2 = 18.352$ con valor $p = .000$.
- 10.19** a. sí; $\chi^2 = 7.267$ b. $.025 < \text{valor } p < .05$
- 10.20** a. no, $\chi^2 = 10.207$ con valor $p = .07$
b. sí, $\chi^2 = 10.207$ con valor $p = .037$
- 10.21** a. $\chi^2 = 10.597$ b. $\chi^2 > 13.2767$
c. No rechazar H_0 d. $.025 < \text{valor } p < .05$
- 10.24** a. Cada tipo de atención representa una población binomial en la que medimos la presencia o ausencia de servicios EMI b. sí; $\chi^2 = 18.446$
- 10.26** Sí, rechazar H_0 ; $\chi^2 = 36.499$
- 10.27** no, $\chi^2 = 4.4$ con valor $p > .10$
- 10.29** no, $\chi^2 = 1.89$ con valor $p > .10$
- 10.31** a. no, $\chi^2 = 1.815$ b. valor $p > .10$
- 10.35** no; no rechazar H_0 ; $\chi^2 = 1.311$.
- 10.37** a. Rechazar H_0 ; $\chi^2 = 18.527$.
b. Rechazar H_0 ; $z = 4.304$; sí.
- 10.40** sí, $\chi^2 = 7.488$ con $.005 < \text{valor } p < .01$
- 10.42** sí, $\chi^2 = 6.190$ con $.025 < \text{valor } p < .05$;
(.347, .483)
- 10.44** a. sí, $\chi^2 = 33.017$ b. valor $p < .005$
- 10.46** sí, $\chi^2 = 17.395$ con valor $p = .008$
- 10.48** a. No rechazar H_0 ; $\chi^2 = 3.660$ con valor $p = .454$;
sí.
- 10.49** a. sí; $\chi^2 = 12.182$ con valor $p = .002$
b. La susceptibilidad de un resfrío parece disminuir cuando aumenta el número de parentescos
- 11.16** a. $H_0: p = \frac{1}{2}$; $H_a: p \neq \frac{1}{2}$; región de rechazo: $\{0, 1, 7, 8\}$; $x = 6$; no rechazar H_0 en $\alpha = .07$; valor $p = .290$
- 11.18** $z = 3.15$; rechazar H_0 .
- 11.20** b. $T = \min\{T^+, T^-\}$ c. $T \leq 137$
d. No rechazar H_0 .
- 11.22** No rechazar H_0 ; $z = -.34$.
- 11.24** a. Rechazar H_0 ; $T = 1.5$ b. Los resultados no concuerdan.
- 11.26** a. no; $T = 6.5$
- 11.28** a. No rechazar H_0 ; $x = 8$.
b. No rechazar H_0 ; $T = 14.5$.
- 11.30** a. prueba de diferencia en pares, prueba del signo, prueba de rango con signo de Wilcoxon
b. Rechazar H_0 con ambas pruebas; $x = 0$ y $T = 0$.
- 11.32** sí, $H = 13.90$
- 11.34** a. no; $H = 2.63$ b. valor $p > .10$
c. valor $p > .10$
- 11.36** no; $H = 2.54$ con valor $p > .10$
- 11.38** a. Rechazar H_0 ; $F_r = 21.19$. b. valor $p < .005$
d. $F = 75.43$ e. valor $p < .005$
f. Los resultados son idénticos
- 11.40** a. No rechazar H_0 ; $F_r = 5.81$
b. $.05 < \text{valor } p < .10$
- 11.42** a. $r_s \geq .425$ b. $r_s \geq .601$
- 11.44** a. $|r_s| \geq .400$ b. $|r_s| \geq .526$
- 11.46** a. $-.593$ b. sí
- 11.48** a. $r_s = .811$ b. sí
- 11.50** sí
- 11.52** sí, $r_s = .9118$
- 11.59** a. No rechazar H_0 ; $x = 2$
b. No rechazar H_0 ; $t = -1.646$
- 11.61** a. No rechazar H_0 ; $x = 7$
b. No rechazar H_0 ; $x = 7$
- 11.63** No rechazar H_0 con la prueba de la suma de rango de Wilcoxon ($T = 77$) o la prueba de diferencia en pares ($t = .30$)
- 11.65** No rechazar H_0 usando la prueba del signo ($x = 2$); no
- 11.67** sí; $r_s = -.845$
- 11.69** Rechazar H_0 ; $T = 14$
- 11.71** a. Rechazar H_0 ; $F_r = 20.13$ b. Los resultados son iguales

Capítulo 11

- 11.1** a. T_1^* b. $T \leq 31$ c. $T \leq 27$
- 11.3** a. H_0 : las distribuciones poblacionales son idénticas; H_a : la población 1 corrida a la izquierda de la población 2 b. $T_1 = 16$; $T_1^* = 39$
c. $T \leq 19$ d. sí; rechazar H_0
- 11.5** No rechazar H_0 ; $z = -1.59$.
- 11.7** No rechazar H_0 ; $T = 102$.
- 11.9** sí; rechazar H_0 ; $T = 45$
- 11.10** sí; rechazar H_0 ; $T = 44$
- 11.12** b. $\alpha = .002, .007, .022, .054, .115$
- 11.14** de una cola: $\mathbf{n} = 10$: $\alpha = .001, .011, .055$;
 $\mathbf{n} = 15$: $\alpha = .004, .018, .059$; $\mathbf{n} = 20$: $\alpha = .001, .006, .021, .058, .132$; de dos colas: $\mathbf{n} = 10$:

11.73 a. Rechazar H_0 ; $H = 9.08$

b. $.025 < \text{valor } p < .05$

c. Los resultados son iguales

11.75 a. no **b.** diferencias significativas entre las respuestas a las tres tasas de aplicación;

$F_r = 10.33$ con valor $p = .006$

11.77 $T = 19$; $T_{.05} = 21$ ($T_{.01} = 18$). Rechazar H_0 .

11.79 $z = 1.18 < z_{.05} = 1.645$; la iluminación no es efectiva

11.81 $H = 7.43$ $gl = 3$; $.05 < \text{valor } p < .10$; no hay diferencia significativa

11.82 a. $r_s = .738$. **b.** valor $p = .025 < .05$; sí, correlación positiva

Glosario

Alfa (α). 1. En las pruebas de significancia estadística, el nivel que designa la probabilidad de cometer el error tipo I; es también conocido como valor p . 2. En las estimaciones de uniformidad u homogeneidad interna, un coeficiente de confiabilidad, como en el alfa de Cronbach.

Alfa de Cronbach. Media de fiabilidad o índice de confiabilidad, utilizada respecto de un conjunto de dos o más indicadores de un constructo. Los valores van de 0 a 1. Los valores de 0.60 a 0.70 se consideran el límite inferior de aceptabilidad.

Algoritmo. Conjunto de reglas o procedimientos; parecido a una ecuación.

Análisis cluster. Técnica multivariante cuyo objetivo es agrupar a los encuestados o los casos con perfiles similares sobre una serie de características definidas. Similar al del análisis factorial Q .

Análisis confirmatorio. Uso de una técnica multivariable para contrastar (confirmar) una relación preestablecida. Por ejemplo, suponga la hipótesis de que sólo dos variables deberían ser las predictoras de una variable dependiente. Si contrasta empíricamente la significancia de esos dos predictores y la no significancia del resto, este contraste es un análisis confirmatorio. Es el opuesto a un análisis exploratorio.

Análisis cualitativo. Organización e interpretación de información no numérica con el propósito de descubrir dimensiones subyacentes y esquemas de relación importantes.

Análisis cuantitativo. La manipulación de datos numéricos por medio de procedimientos estadísticos con el fin de describir fenómenos o estimar la magnitud y confiabilidad de las relaciones entre ellos.

Análisis de contenido. Técnica para describir en forma sistemática y objetiva la forma y el contenido de materiales escritos, verbales o visuales; se usa con frecuencia en estudios cuantitativos de medios masivos de comunicación.

Análisis de correspondencias. Aproximación de composición a la elaboración de mapas conceptuales que relacionan categorías de una tabla de contingencia. La mayoría de las aplicaciones contienen un conjunto de objetos y atributos.

Análisis de costo/beneficio. 1. En la investigación evaluadora, comparación de los costos financieros de un programa o intervención, con los retornos financieros atribuibles a éste. 2. En un proyecto de uso o aplicación, estimación de los costos o riesgos de una nueva práctica o procedimiento en comparación con sus beneficios.

Análisis de covarianza (ANCOVA). Procedimiento estadístico utilizado para evaluar las diferencias medias entre grupos, que se cree que pueden afectar la acción de la(s) variable(s) independiente(s) en lo referente a una variable dependiente, al mismo tiempo que se controlan una o más variables ajenas (covariantes o covariables).

Análisis de función discriminatorio. Procedimiento estadístico usado para predecir la pertenencia a un grupo o el estado de una variable categórica (nivel nominal) con base en dos o más variables independientes.

Análisis de potencia. Procedimiento para estimar la probabilidad de cometer un error tipo II o los requisitos de tamaño de la muestra.

Análisis de reactivos. Muestra el grado de consistencia de varios reactivos.

Análisis de varianza (ANOVA). Técnica estadística empleada para determinar si las muestras provienen de poblaciones con medias iguales o corroborar el efecto de uno o más tratamientos, comparando la variabilidad entre grupos con la variabilidad dentro de los mismos. El análisis univariante de la varianza utiliza una variable dependiente, mientras que el análisis multivariante de la varianza compara muestras basadas en dos o más variables dependientes.

Análisis de varianza multivariada (MANOVA). Procedimiento estadístico usado para evaluar la significancia de las diferencias entre las medias de dos o más grupos en dos o más variables dependientes consideradas en forma simultánea.

Análisis del contenido. Procedimiento para analizar comunicaciones verbales o escritas en forma sistemática y objetiva, generalmente con el fin de estimar cuantitativamente las variables.

Análisis en función discriminativa. Técnica estadística utilizada para predecir la pertenencia a un grupo o el nivel dentro de una variable categórica (nominal), con base en dos o más variables independientes.

Análisis estadístico. 1. Organización y análisis de datos cuantitativos mediante procedimientos estadísticos, incluyendo estadísticas descriptivas e inferenciales. 2. Método para recopilar, organizar, concentrar, reducir, presentar, analizar, generalizar y contrastar los resultados (datos) de las observaciones directas o indirectas de un estudio, investigación o experimento; de tal manera que puedan responderse las interrogantes planteadas de antemano.

Análisis exploratorio. Análisis que establece posibles relaciones sólo de la forma más general y a continuación deja a las técnicas multivariantes la estimación de la(s) relación(es). Opuesto al análisis confirmatorio, el investigador no busca "confirmar" cualquier estimación especificada antes del análisis, sino que deja al método y a los datos definir la naturaleza de las relaciones. Un ejemplo de este tipo de análisis es la regresión múltiple por etapas, en la que el método consiste en añadir variables pronosticadoras hasta que se cumpla algún criterio.

Análisis factorial. Analiza relaciones entre las variables para identificar grupos de variables que forman dimensiones laten-

tes (factores). Procedimiento estadístico que permite reducir un gran número de variables a un conjunto más pequeño de variables con características o dimensiones subyacentes en común.

Análisis factorial común. Modelo factorial en el que los factores se basan en una matriz de correlación reducida.

Análisis multivariante. Análisis de varias variables en una única relación o un conjunto de relaciones.

Análisis por clasificación múltiple. Variante de la regresión múltiple y del análisis de varianza (ANCOVA) que produce medias sobre la variable dependiente, ajustadas según los efectos de las covariantes.

Análisis secundario. Formas de investigación de los datos reunidos por un investigador, que son analizados de nuevo por otro, regularmente para corroborar nuevas hipótesis.

Anonimato. Protección del participante en un estudio, al grado que el mismo investigador no puede vincularlo con la información generada.

Asignación aleatoria. Asignación de sujetos a condiciones de estudio o control, por azar, es decir en un procedimiento regido únicamente por el azar o la casualidad.

Autocomunicado. Todo método para reunir datos que entraña el comunicado directo de información por la persona que es estudiada (como sería en una entrevista o cuestionario).

Autorización con conocimiento informado. Principio ético que obliga a los investigadores a solicitar la participación voluntaria de los sujetos, después de informarles de los posibles riesgos y beneficios.

Bimodal. Distribución de frecuencias que tiene dos modas.

Cálculo de la probabilidad máxima. Método de cálculo (usado a veces en lugar del método de los mínimos cuadrados) en el que los factores de evaluación son aquellos que estiman los parámetros que tienen las mayores probabilidades de haber generado los resultados observados.

Cálculo de los mínimos cuadrados. Método de estimación estadística de uso común en que la solución minimiza las sumas de los cuadrados del valor de los errores; también se le llama prueba de los mínimos cuadrados ordinarios.

Calificación estándar o Z. Modalidad de calificación expresada en unidades de desviación estándar.

Capacidad de generalizar. Grado en el cual las técnicas de investigación justifican la inferencia de que los hallazgos representan algo que va más allá de las observaciones específicas en que se basaron; en particular, la inferencia o deducción de que es posible generalizar los hallazgos a toda la población partiendo de la muestra.

Catálogo de códigos. Documentación utilizada en el procesamiento de datos que indica el sitio y valores de todas las variables archivadas.

Celda. 1. La intersección de una hilera y una columna en una tabla con dos o más dimensiones. **2.** En un diseño experimental, la representación de una condición experimental dentro de un diagrama.

Censo. Estudio que abarca a una población completa.

Codificación. El proceso de transformar datos en bruto en otros estandarizados (por lo regular, numéricos) para el procesamiento y análisis estadístico.

Coefficiente beta (β_n). Coeficiente de regresión tipificado (estandarizado) que permite una comparación directa entre coeficientes de correlación con su capacidad explicativa relativa de la variable de criterio. Mientras los coeficientes de regresión se expresan en términos de unidades de la variable asociada, realizando por tanto comparaciones inapropiadas, los coeficientes beta utilizan datos tipificados y pueden ser comparados directamente.

Coefficiente de confiabilidad. Índice cuantitativo cuyo valor varía generalmente de 0.0 a 1.00 y que proporciona una idea de confiabilidad de un instrumento; se calcula mediante procedimientos como la técnica del Alfa de Cronbach, la técnica de dos mitades, el método de test-retest, etcétera.

Coefficiente de correlación (R_{xy}). Indica la magnitud de la asociación entre la variable independiente y la variable criterio o dependiente. El signo (+/-) indica la dirección de la relación. El coeficiente de correlación va desde +1.00 (para una relación directamente proporcional perfecta, a través de 0.00, es decir ausencia de relación hasta -1.00, para una relación inversamente proporcional perfecta.

Coefficiente de correlación múltiple. Índice que resume el grado de relación entre dos o más variables independientes y una dependiente; se representa con el símbolo R .

Coefficiente de correlación parcial. Medidas de la fuerza de la relación entre la variable criterio y una variable predictora simple, donde los efectos de las otras variables que intervienen en el modelo se mantienen constantes. Por ejemplo, r_{xy} , x^2 , $x+$, miden la variación en Y asociada con x_2 cuando el efecto de x_1 sobre x_2 e Y es constante. Usado en la estimación del modelo en el método de regresión para la selección secuencial de una variable con el fin de identificar la variable independiente con mayor poder de pronóstico, incrementado al de las variables de pronóstico ya presentes en el modelo.

Coefficiente de correlación producto-momento (r). El coeficiente de correlación más ampliamente utilizado, el cual designa la magnitud de la relación entre dos variables medidas por lo menos en una escala de intervalos; también se denomina r de Pearson.

Coefficiente de determinación (R^2). Medida de la proporción de la varianza de la variable criterio sobre su media aritmética, es explicada por las variables independientes o de pronóstico. El coeficiente puede variar entre cero y uno si el modelo de regresión es estimado y aplicado apropiadamente; el investigador puede asumir que cuanto mayor sea el valor de R^2 , mayor será el poder explicativo de la ecuación de regresión y, por tanto, mejor pronóstico de la variable criterio.

Coefficiente de regresión (b_n). Valor numérico del parámetro estimado directamente asociado con las variables independientes. Por ejemplo, en el modelo $Y = b_0 + b_1 X_1$, el valor b_0 es el coeficiente de regresión de la variable X_1 . En el modelo de regresión múltiple (es decir, $Y = b_0 + b_1 X_1 + b_2 X_2$), los coeficientes de regresión son parciales porque cada uno tiene en

cuenta no sólo las relaciones entre Y y X_1 y entre Y y X_2 , sino también entre X_1 y X_2 . El coeficiente no está limitado en su rango, ya que se basa tanto en el grado de asociación como en las unidades de escala de la variable de pronóstico. Por ejemplo, dos variables con la misma asociación a Y tendrían diferentes coeficientes si una variable de pronóstico se midió en escala sobre 7 puntos y otra basada en una escala de 100 puntos.

Coefficiente de validez. Índice cuantitativo que generalmente va de 0.0 a 1.00, con el cual se estima qué tan válido es un instrumento. Por lo general se calcula junto al método de validación de instrumentos por medio de criterios.

Coefficiente phi (π). Índice que describe la magnitud de la relación entre dos variables dicotómicas.

Comparaciones por pares. Contrastes en que se presentan dos elementos y se pide a un experto que los compare.

Concepto. Abstracción basada en observaciones de algunas conductas o características, como serían “estrés” o muerte.

Confiabilidad. Consistencia de las mediciones: posibilidad de repetir o replicar los hallazgos, estabilidad de la medición en el tiempo. Grado de congruencia o fidelidad con la que un instrumento cuantifica el atributo que pretende medir.

Confiabilidad de los calificadores (u observadores). Grado en que los dos calificadores u observadores, que trabajan independientemente, asignan a las mismas calificaciones o valores al atributo que están cuantificando. Tales calificaciones ocurren normalmente dentro del contexto de la investigación u observación.

Confiabilidad de test-retest. Evaluación de la estabilidad de un instrumento mediante la correlación de los resultados obtenidos en aplicaciones reiterativas. También se aplica a una misma persona la prueba en dos ocasiones y se comparan los resultados.

Confiabilidad por mitades. Método consistente en dividir una escala o una prueba en dos mitades que se comparan estadísticamente.

Confianza, intervalo de. Límites de dos valores, máximo y mínimo, dentro de los cuales se considera que un parámetro de la población está incluido con cierta confianza o probabilidad.

Consentimiento informado. Principio ético que exige a los investigadores obtener la participación voluntaria de los sujetos después de haberles informado acerca de los posibles riesgos y beneficios del estudio. Los participantes lo otorgan libremente.

Constructo. Noción que el investigador puede definir en términos conceptuales, pero que no puede ser directamente medido (es decir, el encuestado no puede articular una respuesta única que proporcionaría total y perfectamente una medida del concepto) o medido sin error. Los constructos son las bases para formar las relaciones causales, en la medida que son las representaciones más “puras” posibles de un concepto. Es inventado (construido) por los investigadores para una finalidad específica. Sin embargo, cualquiera que sea su nivel de especificidad, un constructo no puede ser medido directa y perfectamente aunque sí debe hacerlo aproximadamente por indicadores.

Contraste *post hoc*. Contraste de las diferencias de medias, realizado después de que la hipótesis nula sea rechazada al aplicar un análisis de varianza. Por lo general, los contrastes *post hoc* no utilizan un único contraste, sino que prueban las

diferencias entre todas las combinaciones posibles de los grupos. Aunque proporcionan una información diagnóstica abundante, inflan el error tipo I con el empleo de contrastes estadísticos múltiples y, por tanto, tienen que utilizar niveles de confianza muy estrictos.

Contraste (prueba) *t* de Student. Estadístico que valora la significancia estadística de las diferencias entre dos medias muestrales para una sola variable dependiente. Es un caso especial del ANOVA para dos grupos o niveles de la variable de tratamiento.

Control. El proceso de conservar constantes las posibles influencias en la variable dependiente bajo estudio.

Correlación. Asociación entre dos conjuntos de puntuaciones; con frecuencia se expresa en términos de un coeficiente de correlación. Tendencia de la fluctuación (variación) en una variable a relacionarse con la fluctuación (variación) en otra variable.

Correlación biserial. Medida de correlación utilizada para reemplazar la correlación de momento-producto cuando una variable medida métricamente está asociada con una medida binaria (cero, uno) no métrica.

Correlación canónica. Procedimiento estadístico que permite examinar la relación entre dos o más variables independientes y dos o más variables dependientes.

Correlación negativa o relación inversa. Estado en el que, a medida que crece una variable, la otra decrece.

Correlación parcial bivalente. Correlación simple (dos variables) entre dos series de residuos (variación no explicada) que queda tras eliminar la influencia de otras variables independientes.

Correlación positiva. Situación en que un incremento de variable se acompaña por uno de la otra.

Correlación tetracórica. Medida de asociación utilizada para relacionar dos medidas binarias.

Covariante. Variable que se controla estadísticamente (mantenida constante) en el análisis de la covarianza. La covariante suele deberse a una influencia ajena que introduce confusión en la variable dependiente.

Covarianza o análisis de covarianza. Empleo de procedimientos similares a la regresión, para eliminar la variación extraña (ajena) de las variables dependientes debida a una o más variables independientes no-controladas (covariaciones o covariantes). Se supone que las covariaciones están linealmente asociadas con las variables dependientes. Después del ajuste debido a la influencia de las covariaciones, se realiza un ANOVA (o MANOVA) estándar. Este proceso de ajuste (conocido como ANCOVA y MANCOVA) permite generalmente el uso de contrastes más sensibles de los efectos de tratamiento.

Cuantificación basal. Cuantificación de la variable dependiente antes de introducir una intervención experimental.

Cuantificación nominal. El nivel más bajo de cuantificación, que consiste en asignar categorías a las características (por ejemplo, hombres = 1; mujeres = 2).

Cuantificación ordinal. Nivel de cuantificación que produce un ordenamiento jerárquico de la variable a lo largo de alguna dimensión.

Cuantificación proporcional. Nivel de cuantificación en el que hay distancias iguales entre las unidades de medida y que tienen un punto cero significativo (por ejemplo, el peso corporal es el nivel de cuantificación más alto).

Cuasi-experimento o experimento natural. Técnica de asignación no aleatoria de sujetos a condiciones. En ella, el investigador carece de control directo sobre la variable independiente.

Cuestionario. Serie de preguntas escritas sobre un tema respecto del cual se buscan las opiniones personales.

Curtosis. Medida del apuntamiento o llanura de una distribución cuando se compara con una distribución normal.

Curva normal. Curva simétrica en forma de campana que suele aproximarse a la frecuencia de ocurrencia de los eventos de la naturaleza.

Definición operacional. Definición de una variable por la forma en que se mide, la inteligencia se define operacionalmente como la calificación de una prueba de CI.

Desviación estándar (DE , s o σ). El concepto estadístico más utilizado para estimar el grado de variabilidad de un conjunto de resultados.

Diagrama de dispersión. Gráfica de una correlación o relación lineal entre dos variables.

Diagrama de Gantt. Diagrama que representa la programación de las actividades (tareas) de una investigación, estudio o experimento, y en el cual se destacan el orden secuencial y las interrelaciones de las actividades.

Diagrama de tallo y hojas. Una variante del histograma que proporciona una representación visual de la distribución de la variable, así como una enumeración de los valores efectivos de los datos.

Diferencial semántico. Técnica usada para cuantificar actitudes, en la cual se pide a los respondientes que califiquen un concepto de interés situándolo en una serie de escalas bipolares de siete puntos.

Desarrollado por C. Osgood, para medir el significado de los conceptos.

Diseño antes-después. Diseño experimental en el cual se obtienen datos de los sujetos de la investigación antes y después de introducir las condiciones experimentales, también se conoce como diseño de test-retest.

Diseño de experimentos. Plan de investigación en el cual el investigador manipula o controla directamente una o más variables de pronóstico y valora sus efectos sobre las variables dependientes. Ya habitual en la ciencia física, está ganando popularidad en las ciencias económicas y sociales. Por ejemplo, a los encuestados se les muestran anuncios publicitarios distintos que varían temáticamente en una característica, tales como diferentes pretensiones (emocional frente a racional).

Diseño de la investigación. Plan global para obtener y analizar datos, el cual incluye especificaciones para dar mayor validez interna y externa al estudio o experimento.

Diseño de medidas repetidas. Diseño experimental o estudio en el que el mismo grupo de sujetos se expone a más de una condición o tratamiento.

Diseño de muestreo proporcional. Estrategia de muestreo en la que el investigador toma muestras de los diferentes estratos de la población, en proporción directa a su representatividad dentro de dicha población.

Diseño factorial. Diseño experimental en el que se manipulan simultáneamente dos o más variables independientes. Este diseño permite analizar por separado los efectos principales de las variables independientes más los efectos de interacción de dichas variables.

Diseño multivariado. Enfoque donde participa más de una variable dependiente.

Diseño por series temporales. Diseño cuasi-experimental que consiste en obtener información a través de un periodo prolongado, durante el cual se establecen muchos puntos de recolección de datos antes y después de la introducción de un tratamiento.

Diseño pre-experimental. Diseño de investigación que no incluye controles o testigos para compensar la ausencia de la asignación aleatoria o de un grupo control (testigo).

Distribución asimétrica. Distribución de valores sesgada (sus dos mitades no son simétricas o especulares mutuas).

Distribución bimodal. Distribución de valores con dos picos modas (frecuencias altas).

Distribución de frecuencia. Conjunto sistematizado de valores numéricos, desde los más bajos hasta los más altos, junto al número de veces (frecuencia) en que se obtuvo tal valor.

Distribución de muestreo. Distribución teórica de un estadístico que utiliza un número infinito de muestras como base y los valores del estadístico calculado a partir de tales muestras como los puntos de la distribución.

Distribución multimodal. Distribución de valores con más de dos picos (frecuencia alta).

Distribución normal. Distribución teórica, simétrica y con forma de campana. También se llama curva normal o de Gauss. El eje horizontal representa todos los posibles valores de una variable y el eje vertical la probabilidad de que ocurran dichos valores. Aquéllos están agrupados alrededor de su media aritmética en forma simétrica y unimodal.

Distribución unimodal. Distribución de valores con una sola moda (un pico, frecuencia más alta).

Efecto de Hawthorne. Efecto de la variable dependiente debido a que los sujetos están conscientes de que están sometidos a estudio.

Efecto de interacción. Efecto que tienen sobre una variable dependiente dos o más variables independientes que actúan en combinación (en forma interactiva), al contrario de hacerlo como factores independientes entre sí.

Efectos principales. En un estudio con múltiples variables independientes, los efectos de cada variable independiente sobre la variable dependiente.

Empirismo. Proceso en el cual los datos provenientes de la realidad, y que se reúnen a través de los sentidos, se utilizan como base para generar conocimientos.

Encuesta. Tipo de investigación no experimental que se orienta a tener información respecto del “estado real” de alguna si-

tuación por medio de interrogatorio directo de una muestra de respondientes.

Entrevista. Método de reunión de datos en la cual una persona (entrevistador) formula preguntas a otra (entrevistado); se hace en forma directa o por teléfono.

Entrevista semi-estructurada. Técnica en que se formulan las mismas preguntas a todos los consultados, pero el orden de éstas puede diferir.

Error de medición. El grado de desviación entre los resultados verdaderos y los obtenidos al cuantificar o evaluar una característica.

Error de muestreo. Variación debida al azar entre muestras elegidas de una sola población.

Error de predicción. Diferencia entre los valores reales y de predicción de la variable criterio para cada observación en la muestra.

Error de selección. Peligro para la validez interna del estudio, que es resultado de las transferencias previas en tratamiento entre los grupos experimental y control.

Error estándar de la media. Medida de la dispersión de las medias o de las frecuencias de las medias esperadas debido a la variación muestral. Denota la gama esperada del coeficiente a través de las muestras múltiples de los datos. Resulta útil en pruebas estadísticas para ver si el coeficiente es sustancialmente diferente de cero (es decir, si la gama esperada del coeficiente contiene el valor cero a un nivel de confianza específico). El valor t de un coeficiente de regresión es el coeficiente dividido por su error estándar.

Error tipo I. Probabilidad de rechazar la hipótesis nula siendo cierta, es decir, concluir que dos medias son significativamente diferentes cuando de hecho son iguales. Valores de alfa (por ejemplo, 0.05 o 0.01) llevan a que el rechazo de la hipótesis nula sea insostenible, y el no rechazo de la hipótesis alternativa, de que las medias poblacionales son distintas.

Error tipo II. Probabilidad de no rechazar la hipótesis nula cuando es falsa, es decir, concluir que las dos medias no son significativamente diferentes cuando de hecho lo son. También conocido como beta (β). En términos más sencillos, la probabilidad de no encontrar correlación o diferencia de medias cuando existen. Está inversamente relacionado con el error tipo I. El valor 1 menos el error tipo II se define como la potencia o poder de la prueba estadística.

Escala de Likert. Tipo de cuantificación mixta de las actitudes, el cual implica sumar los resultados de un conjunto de reactivos (enunciados) a los que los respondientes deben asignar su grado de acuerdo o desacuerdo.

Escala ordinal. Las características a evaluar pueden ordenarse en una dimensión subyacente, pero no se brinda información respecto de la distancia entre los puntos, sino sólo a la magnitud o a la dirección en orden creciente o decreciente.

Estadística bivariable. Estadísticas derivadas del análisis de dos variables simultáneamente, con el fin de evaluar la relación empírica que priva entre ellas.

Estadística de prueba. Estadísticas usadas para evaluar la significancia estadística de las relaciones entre variables. Sus dis-

tribuciones de muestreo se conocen por las circunstancias en las que la hipótesis nula es verdadera; como ejemplos pueden citarse la *ji-cuadrada*, la razón F , el valor t y la r de Pearson.

Estadística descriptiva. Técnicas utilizadas para describir y resumir el conjunto de datos obtenidos por el investigador; por ejemplo, media aritmética, mediana, moda, desviación estándar, varianza, porcentaje de las frecuencias, gráficas, etcétera.

Estadística inferencial (o deductiva). Estadística que permite al investigador deducir si las relaciones, diferencias en una o varias muestras, pueden ocurrir en una población, de donde fue extraída la o las muestras.

Estadística multivariable. Técnica estadística que analiza relaciones entre tres o más variables. Dentro de las estadísticas multivariadas más usadas están: regresión múltiple, análisis de función discriminativa y análisis factorial.

Estadística no paramétrica. Clase general de estadística inferencial sin suposiciones rigurosas sobre la distribución normal de las variables; suele utilizarse en muestras pequeñas o también cuando los datos se miden en escalas nominal u ordinal.

Estadística paramétrica. Una clase de estadística inferencial que consta de: 1. suposiciones acerca de la distribución de las variables; 2. estimación de un parámetro, y 3. el uso de cuantificaciones intervalares o proporcionales.

Estadística univariada. Técnica para analizar una sola variable con fines descriptivos.

Estadísticas multivariadas. Procedimientos estadísticos diseñados para analizar las relaciones entre tres o más variables. Dentro de las estadísticas multivariadas en uso cabe mencionar la regresión múltiple, el análisis de covarianza y el análisis de factores.

Estadístico. Utilizado para corroborar la significancia estadística de relaciones entre variables. Las distribuciones muestrales de los estadísticos de prueba se conocen para circunstancias en que la hipótesis nula es verdadera; entre los ejemplos están *ji-cuadrada*, pruebas F , t y r de Pearson.

Estimación de máxima verosimilitud (MLE). Método de estimación habitualmente empleado en modelos de ecuaciones estructurales, incluyendo LISRELL y EQS.

Estimación por intervalos. Método de estimación estadística en el que el investigador establece límites para los valores que con mayor probabilidad, dentro de cierto nivel de confianza, representarán el verdadero parámetro de la población.

Estimación puntual. Procedimiento de estimación estadística en que el investigador utiliza la información de una muestra para calcular el valor aislado (estadístico) que mejor represente el valor del parámetro de la población.

Estrato. Subdivisión de la población con arreglo a alguna característica (por ejemplo, varones y mujeres).

Estructura muestral. Lista de todos los elementos en la población a partir de la cual se extrae la muestra.

Estructura o esquema conceptual. Conceptos o abstracciones interrelacionados, que están ensamblados en algún esquema racional por su importancia con el tema común (véase también teoría).

Estudio de campo. Estudio en el que los datos se toman en el campo, a partir de individuos en sus actividades cotidianas y no

en una situación artificial, esto con el objeto de comprender las costumbres, formas de comportamiento y creencias que tienen los individuos o grupos mientras están inmersos en la vida real.

Estudio de caso. Investigación profunda de una sola instancia. La unidad puede ser tan pequeña como un individuo o tan grande como toda una comunidad.

Estudio de casos y controles. Diseño de investigación, típico de las investigaciones retrospectivas efectuadas después de ocurridos los hechos, que consiste en comparar un “caso” (es decir, un sujeto con el trastorno en escrutinio, por ejemplo, cáncer pulmonar) y un control o testigo equivalente (es decir, una persona sin dicho trastorno).

Estudio de casos. Método de investigación que comprende el análisis minucioso de un individuo, grupo, comunidad o institución, considerándola como una unidad social.

Estudio de cohorte. Estudio de tendencias que se orienta a una muestra específica (a menudo un subgrupo catalogado por edades), a partir de las cuales se escogen muestras diferentes en puntos cronológicos sucesivos (como serían estudiantes universitarios graduados en el periodo determinado).

Estudio de panel. Técnica usada con frecuencia en la investigación de opinión pública: se entrevista en repetidas ocasiones a la misma muestra de personas.

Estudio de tendencia. Tipo de estudio longitudinal en el que se analizan diferentes muestras de una población en relación con el tiempo y algún fenómeno (como sería la serie de encuestas de preferencias políticas y electorales).

Estudio descriptivo univariado. Estudio en el que se obtiene información sobre la ocurrencia, su frecuencia o el valor promedio de las variables de interés, considerándolas una por una.

Estudio en el panel. Tipo de estudio longitudinal en el que participan los mismos sujetos para obtener datos en dos o más puntos del tiempo.

Estudio exploratorio. Estudio que comienza con un análisis de las causas supuestas (por ejemplo, el tabaquismo) y luego avanza para observar los efectos, también supuestos (por ejemplo, cáncer pulmonar).

Estudio longitudinal. Investigación sobre un individuo o grupo durante un periodo prolongado, a diferencia del estudio transversal.

Estudio piloto. Método de uso preliminar a pequeña escala de un procedimiento, diseñado para identificar problemas y omisiones antes de conducir el estudio real.

Estudio prospectivo. Investigación que comienza con el examen de las supuestas causas, como serían, por ejemplo, las del tabaquismo, y sigue adelante cronológicamente para observar los efectos supuestos (como sería el cáncer del pulmón).

Estudio retrospectivo. Estudio que comienza con la manifestación de la variable dependiente en el presente (por ejemplo, cáncer pulmonar) y donde luego se relaciona este efecto con alguna causa presunta en el pasado (por ejemplo, fumar tabaco).

Estudio transversal. Comparación de diferentes grupos en un solo momento particular.

Ética. La calidad de técnicas de investigación en lo que respecta a su cumplimiento de obligaciones profesionales, legales y sociales, para los sujetos del estudio.

Evaluación de necesidades. Estudio en el cual un investigador reúne datos para calcular las necesidades de un grupo, comunidad u organización; a menudo se emplea como guía para asignación de recursos.

Evaluación de programas. Conjunto de procedimientos sistemáticos para determinar la eficacia de un plan.

Evaluación formativa. Valoración dinámica de un producto o programa en su fase de desarrollo, con el fin de optimizar la calidad definitiva del mismo.

Evaluación psicométrica. Evaluación de la calidad de un instrumento, basada principalmente en las evidencias de su confiabilidad y validez.

Evaluación retrospectiva. Investigación diseñada para examinar la utilidad o el valor de un programa o práctica después de que ya está en operación.

Evaluación sumativa. Investigación que tiene como fin examinar la utilidad o valía de un programa o la práctica, después de que está ya en operación.

Experimento. Estudio de investigación en el que el investigador controla (manipula) la variable independiente y asigna sujetos en forma aleatoria a cada condición diferente.

Experimento doble ciego. Experimento en el que ni a los sujetos ni quienes administran el tratamiento saben cuál es el grupo experimental y cuál es el grupo control.

Factor (análisis factorial). Combinación lineal (valor teórico) de las variables originales. Los factores también representan las dimensiones subyacentes (construcciones) que resumen o justifican la serie original de variables en observación.

Fórmula de Kuder-Richardson (KR-20). Método para calcular un coeficiente de confiabilidad de la uniformidad interna del conjunto de reactivos de una escala; se usa cuando los reactivos son dicotómicos.

Frecuencia. Número de veces que ocurre una puntuación, dato o nivel de una categoría.

Frecuencia esperada (f_e). Repeticiones que representan la hipótesis nula o lo que se esperaría.

Frecuencia observada (f_o). Datos en bruto (frecuencia, no porcentajes).

Grados de libertad (gl). Concepto usado en las pruebas de significancia estadística que se refiere al número de valores de la muestra que no pueden ser calculados a partir del conocimiento de otros valores y una característica determinada (por ejemplo, si se conoce la media de una muestra, todos los valores, excepto uno, pueden variar libremente). Aunque por lo común el número de grados de libertad es igual a $N - 1$, existen diferentes fórmulas para las distintas pruebas.

Gráfica de barras. Figura construida con barras.

Gráfico de cajas y bigotes (box plot). Método de representación de la distribución de la variable. Una caja representa la mayor parte de la distribución, mientras que las extensiones, llamadas bigotes, llegan al final de la distribución. Es muy útil para realizar comparaciones de su media aritmética o mediana, de una o varias variables, considerando también uno o varios grupos.

Gráfico de dispersión. Representación de las relaciones entre dos variables métricas que muestran los valores conjuntos de cada observación en un gráfico de dos dimensiones.

Gráfico de distribución normal. Comparación gráfica de la forma de la distribución respecto de la distribución normal. Ésta se representa por una línea recta con un ángulo de 45 grados. La distribución efectiva se dibuja contra esa línea, de tal forma que cualquier diferencia se muestra como desviación de la línea recta, haciendo muy obvia e interpretable la identificación de estas diferencias.

Grupo control. Conjunto de sujetos homogéneos y parecidos al grupo experimental en todos los aspectos, excepto en que no recibe el tratamiento y no está expuesto a la variable independiente. Se usa para controlar los efectos de variables extrañas en la dependiente; constituye una base con la cual se medirán los efectos del tratamiento; también es llamado grupo testigo.

Grupo de comparación (control o testigo). Grupo de sujetos cuyos resultados respecto de una variable dependiente sirven como base para evaluar los resultados del grupo experimental o de interés primordial. En general, el término grupo de comparación se utiliza en vez de grupo control o testigo cuando en la investigación no se usa un diseño verdaderamente experimental.

Grupo de tratamiento o experimental. Conjunto de sujetos expuestos a los niveles de la variable independiente, de intervención o tratamiento.

Guttman, escala de. Método para medir actitudes, que utiliza un conjunto de puntos acumulativos (de un solo tono) respecto de los cuales se pide a los participantes que estén de acuerdo o en desacuerdo.

Hawthorne, efecto. Efecto de una variable dependiente, causado por el conocimiento, por parte de los sujetos, de que son participantes experimentales “especiales” en el estudio.

Hipótesis. Afirmación comprobable derivada lógicamente de la teoría o de la observación, puede confirmarse (no rechazarse) o rechazarse. Se somete a comprobación.

Hipótesis nula (H_0). Suposición de que las diferencias producidas por la manipulación de la investigación se deben a fluctuaciones al azar y que la variable independiente no tiene efecto en la dependiente.

Hipótesis alternativa. Hipótesis de trabajo o de investigación, alternativa a la hipótesis nula. Es una proposición escrita en terminos afirmativos (es decir, que habrá un efecto).

Hipótesis direccional. Hipótesis que hace una predicción específica sobre la dirección y la naturaleza de la reacción entre dos variables.

Hipótesis no direccional. Hipótesis de investigación que no estipula por adelantado la dirección y naturaleza de la relación entre las variables.

Histograma. Representación gráfica de la distribución de una variable. Al formar la distribución de frecuencias en categorías, puede verse el perfil de la distribución de la variable. Se utiliza para realizar una comparación visual con la distribución normal.

Homogeneidad. 1. En términos de la confiabilidad de un instrumento, el grado con el cual sus partes integrantes (subpartes)

muestran congruencia interna, esto es, miden el mismo atributo crítico. **2.** En términos más generales, el grado de semejanza de los objetos, es decir, el caracterizarse por su poca variabilidad.

Homoscedasticidad. Descripción de datos en los que la varianza del término de error aparece constante sobre un rango de variables independientes. El supuesto de igual varianza del error de la población E (estimado) es decisivo para la apropiada aplicación de la regresión lineal y de las pruebas paramétricas. Cuando el término de error tiene una varianza en aumento u ondulante, se dice que los datos son heteroscedásticos. La discusión de los residuos ilustra mejor este punto.

Igualamiento o ajuste. El “emparejamiento” de sujetos en un grupo con otros de otro grupo, basado en su semejanza en una o más dimensiones hechas para mejorar la susceptibilidad global de los grupos a ser comparados. Cuando se hace el igualamiento en el marco de un experimento, el método genera un diseño “aleatorizado” en bloques.

Investigación clínica. Investigación diseñada para generar conocimientos que guíen la práctica de la medicina o la psicología.

Investigación con encuestas. Recolección sistemática de información sobre creencias, actitudes, valores y comportamiento personales.

Investigación correlativa. Estudios en los que se exploran las interrelaciones de las variables de interés sin que haya ninguna intervención activa por parte del investigador.

Investigación descriptiva. Estudios que tienen como principal objetivo la representación precisa de las características de individuos, situaciones o grupos, y expresar la frecuencia con que determinados fenómenos ocurren.

Investigación empírica. Análisis que implica la medición de eventos observables.

Investigación evaluativa. Investigación cuyo objetivo es indagar qué tan bien funciona un programa, una práctica o una política.

Investigación *ex post facto* (después de los hechos). Investigación efectuada después de que ocurrieron los cambios en la variable independiente durante el curso natural de los eventos. Es una forma de investigación no experimental en que las aplicaciones causales son inferidas “después de los hechos”.

Investigación exploratoria. Tipo de investigación no experimental enfocada en obtener información relacionada con el *statu quo* de determinada situación. A menudo, se da por medio de la interrogación directa de una muestra, sirve para establecer hipótesis.

Investigación histórica. Estudios sistemáticos diseñados con el fin de establecer hechos y relaciones concernientes a ciertos acontecimientos pasados.

Investigación instrumental. Estudio realizado como requisito académico, vocacional o profesional cuyo objetivo es demostrar la competencia para la investigación.

Investigación metodológica. Investigación diseñada para crear o refinar procedimientos de obtención, organización o análisis de datos.

Investigación no experimental. Estudio en el que el investigador recaba sus datos sin introducir tratamiento o cambio alguno.

Investigación observacional. Estudios en que se reúnen datos por medio de observación y registro de conductas o actividades de interés.

Investigación secundaria. Estudio realizado en archivos con el uso de datos primarios existentes.

Ji-cuadrada (X^2). 1. Prueba estadística utilizada con datos categóricos para probar si una distribución de puntuaciones obtenida difiere confiablemente de lo que se esperaría debido al azar. 2. Método de estandarización de datos en una tabla de contingencia comparando la frecuencia observada con la esperada en cada una de las casilla (ésta se obtiene mediante el producto de las probabilidades marginales de su fila o renglón con las de su columna, dividiendo dicho producto entre la frecuencia total).

Lambda de Wilk. Índice usado en el análisis de funciones discriminatorias para indicar la proporción de la varianza de la variable dependiente que no es explicada por los predictores; $(\lambda) = 1 - R_2$.

Likert, escala de. Medición compuesta de actitudes que comprende la suma de puntuaciones obtenidas con un conjunto de proposiciones respecto a las cuales se pide a los respondientes indicar su grado de aprobación o desaprobación.

Literatura, revisión de la. Resumen crítico de investigación sobre un tema de interés, que se prepara, en forma general, para colocar al problema estudiado en el debido contexto o identificar deficiencias y faltas en estudios anteriores, para así justificar la nueva investigación.

Maduración. Amenaza para la validez interna de un estudio. Aparece cuando los factores influyen en la medición de la culminación (variable dependiente), como resultado del transcurso del tiempo.

Manipulación. Intervención o tratamiento introducido por el investigador en un estudio experimental o cuasi experimental. Él manipula la variable independiente para variar su impacto en la investigación.

Media (aritmética) o promedio. Estadística descriptiva que mide la tendencia central considerada como el punto de equilibrio o centro de gravedad de un conjunto de datos. Se calcula sumando todos los resultados y dividiendo el resultado entre el número de casos o sujetos.

Mediana (Me). Punto medio de una distribución cuando todas las calificaciones están acomodadas de mayor a menor. La mitad de las calificaciones (50%) se halla por encima de la mediana y el resto por debajo de ella.

Medición de la razón. Nivel de medición en el cual existen igual distancia entre las unidades cuantitativas y un cero significativo real. Es el máximo nivel de medición (como sería la edad).

Medida de intervalo. Nivel de medición donde un atributo de una variable se ordena por jerarquías o rangos, en una escala que tiene iguales distancias entre los puntos de la misma (como serían los grados Fahrenheit, por ejemplo).

Medida nominal. El nivel menor de medición que comprende la asignación de características a categorías (como sería asignar a los varones la categoría 1 y a las mujeres la categoría 2).

Medida ordinal. Nivel de medición que genera ordenamiento por jerarquías (rangos) de una variable, siguiendo alguna dimensión.

Medidas nominales. Características asignadas a las categorías; no hay una dimensión continua subyacente.

Métodos de investigación. Son las fases técnicas y estrategias para reunir y analizar los datos en una investigación.

Mínimos cuadrados. Procedimiento de estimación utilizados en la regresión simple y múltiple por el que se estiman los coeficientes de regresión para aminorar la suma total de los residuos cuadrados.

Moda. Calificación que ocurre con mayor frecuencia en una distribución unimodal, si hay dos, la distribución se denomina bimodal.

Modelo. Conjunto especificado de relaciones de dependencia que puede ser contrastado empíricamente por medio de la operacionalización de una teoría. El propósito de un modelo es proporcionar concisamente una representación amplia de las relaciones a examinar. El modelo puede ser formalizado en un diagrama de secuencias o en un conjunto de relaciones estructurales.

Mortalidad. Amenaza de la validez interna de un estudio que se refiere a la pérdida diferencial de sujetos (atrición), de diferentes grupos.

Muestra. Subconjunto de una población seleccionado para participar en un estudio de investigación.

Muestra aleatoria. Tipo de muestra probabilística en que todos los individuos de la población estudiada tienen la misma probabilidad de ser seleccionados.

Muestra con propósito. Muestra no probabilística en que se elige para su inclusión a los individuos considerados de mayor relevancia para lo estudiado.

Muestra de bola de nieve. Tipo de muestra con propósito (no probabilística) en la que el investigador pregunta a sus entrevistados a quién más debería considerar.

Muestra estratificada. Muestra probabilística donde se seleccionan sus características para que sean proporcionales a las presentes en la población total.

Muestra probabilística. Es la extraída de tal forma que pueda estimarse la probabilidad de inclusión de cualquier individuo determinado. Existen dos tipos generales: la aleatoria y la estratificada.

Muestras no probabilísticas. Son aquellas en las que el investigador desconoce la probabilidad de selección. Los tres tipos generales de éstas son: por cuotas, intencional y accidental.

Muestras por cuotas. Tipo de muestras no probabilísticas en que el investigador establece deliberadamente proporciones de una muestra distinta de las existentes en la población.

Muestreo. Proceso encaminado a la selección de un fragmento de la población que represente a la población entera.

Muestreo accidental. Selección de los individuos (o unidades) más fácilmente accesibles para un estudio; también se conoce como muestreo por conveniencia.

Muestreo aleatorio. Selección de una muestra de manera que todo miembro de una población (o subpoblación) tenga las mismas probabilidades de ser elegido.

Muestreo aleatorio estratificado. Selección aleatoria de sujetos en dos o más estratos independientes de la misma población.

Muestreo aleatorio simple. Tipo más sencillo de muestreo probabilístico en el que se crea un marco de muestreo enumerando a todos los miembros de una población de interés y seleccionando después una muestra, dentro de este marco, a través de procedimientos completamente aleatorios.

Muestreo cronológico. En la investigación por observación, la selección de periodos durante los cuales se harán las observaciones.

Muestreo desproporcionado. Estrategia de muestreo mediante la cual el investigador muestrea diferentes proporciones de sujetos, en distintos estratos de la población, con el fin de asegurar una representación adecuada de los estratos comparativamente menores.

Muestreo en el tiempo. En investigación observatoria, la elección de intervalos temporales durante los cuales se efectuarán las observaciones.

Muestreo intencionado (intencional). Método no probabilístico de muestreo donde el investigador escoge sujetos para el estudio; con base en su criterio personal, respecto de aquellos que tendrán mayor representatividad o productividad. También se le conoce como muestreo por criterio.

Muestreo multifásico. Estrategia de muestreo que se realiza a través de una serie de fases que van de unidades de muestreo grandes a pequeñas (por ejemplo, de las entidades federativas a las escuelas de enfermería y luego a los miembros del profesorado).

Muestreo no probabilístico. La selección de sujetos o unidades muestrales de una población por medio de técnicas no aleatorias; entre los ejemplos están los muestreos accidental, voluntario y por cuota.

Muestreo por cuotas. Selección no aleatoria de sujetos en la que el investigador especifica de antemano las características de la muestra con el propósito de incrementar su representatividad.

Muestreo por eventos. En los estudios por observación, plan de muestreo consistente en la selección de formas de comportamiento o eventos integrales.

Muestreo conglomerados. Forma de muestreo en varias etapas, en la que primero se seleccionan grandes conjuntos (por ejemplo, escuelas de enfermería) y luego se toman submuestras sucesivas de menor tamaño (por ejemplo, estudiantes de enfermería).

Muestreo por recomendación. Muestreo de sujetos con base en las referencias proporcionadas por otros sujetos que ya están en la muestra; también se llama muestreo en cascada.

Muestreo probabilístico. Selección de sujetos o unidades muestrales de una población por empleo de técnicas aleatorias; entre los ejemplos están muestreo aleatorio simple, en cúmulos y sistemático.

Muestreo sistemático. Selección de sujetos de tal forma que se escoja un individuo u objeto de cada determinado número de personas u objetos (cada décima persona, o elemento en una estructura muestral o lista).

Muestreo teórico. En estudios cualitativos, la selección de miembros de la muestra con base en descubrimientos que van surgiendo según avanza el estudio, con el fin de asegurar una representación adecuada de los temas importantes.

Muestreo, error en el. Fluctuación del valor de una estadística entre distintas muestras tomadas de la misma población.

Muestreo, marco de. La lista de todos los elementos de la población de la cual se toma la muestra.

Muestreo, vicio sistemático (error) de. Distorsiones que provienen de la selección de una muestra pero que ésta no es representativa de la población de la cual se extrajo.

Nivel de confianza. La probabilidad estimada de que un parámetro de la población esté dentro de un intervalo de confianza dado.

Nivel de significancia. La probabilidad de que una relación observada puede ser causada por el azar, esto es, por error muestral. Significancia al nivel de 0.05 (5%) que indica la probabilidad de que alguna relación de la magnitud observada aparezca por azar cinco veces de cada 100. También se conoce como la probabilidad de cometer el error tipo I.

Normas. Estándares de rendimiento de una prueba, basados en la obtención de información sobre los resultados de la prueba a partir de una muestra grande y representativa.

Observación participante. Método de reunir datos por observación de un grupo; organización en la cual el investigador participa como miembro.

Paradigma. Una manera de ver los fenómenos naturales que abarca un conjunto de suposiciones filosóficas y sirve como guía para abordar la indagación.

Parámetro. Una característica de la población (por ejemplo, la talla promedio de los ciudadanos mexicanos).

Población. El conjunto completo de individuos u objetos que tienen alguna característica en común (como serían todas las enfermeras en un estado o país). A veces se le conoce como universo.

Población accesible. La población de sujetos disponibles para un estudio particular; a menudo, un subgrupo no aleatorio de la población blanco.

Población blanco. Toda la población en la que el investigador está interesado y a la cual intenta generalizar los resultados de un estudio.

Población de interés o de estudio. La población entera en la que el investigador está interesado y en la cual quisiera generalizar los resultados de su estudio.

Predicción. Uno de los objetivos del método científico. Es el uso de la evidencia empírica para hacer predicciones en cuanto a cómo se comportarán las variables de interés en un escenario nuevo y con sujetos diferentes.

Pregunta o planteamiento abierto. Pregunta o planteamiento en una entrevista o cuestionario que no restringe las respuestas de los participantes a alternativas preestablecidas.

Pregunta o planteamiento cerrado. Pregunta que brinda a los participantes un conjunto de respuestas alternativas que son mutuamente excluyentes y exhaustivas, de las cuales hay que escoger la que más se acerque a la respuesta “verdadera”.

Procedimiento de comparación múltiple. Pruebas estadísticas, normalmente aplicadas después de que los resultados del análisis de varianza (ANOVA) indican diferencias de grupos estadísticamente significativas, que permiten comparar distintos pares de grupos.

Procedimiento de muestreo por tiempo. Método en que se seleccionan momentos específicos conforme a un plan de muestreo para registrar actividades observadas.

Protocolo. Documento que especifica lo que se propone estudiar el investigador; comunica el problema en estudio, su importancia, las técnicas planeadas para resolver la interrogante y, cuando se busca apoyo económico, los costos de la investigación.

Prueba. Procedimiento sistemático para comparar rendimiento, sentimientos, actitudes o valores personales.

Prueba clínica. Experimento destinado a probar la eficacia de un tratamiento clínico y que, por lo general, se efectúa con una muestra grande y heterogénea de sujetos.

Prueba de Friedman. Análogo no paramétrico del análisis de varianza (ANOVA para muestras repetidas) que se emplea cuando el investigador está trabajando con un solo grupo medido varias veces.

Prueba de Kruskal-Wallis. Prueba no paramétrica (análoga al ANOVA para muestras independientes) que se usa para estimar la diferencia en tres o más grupos independientes con base en sus resultados jerarquizados.

Prueba de la mediana. Prueba no paramétrica que consiste en comparar las medianas de dos grupos independientes para averiguar si éstos provienen de poblaciones con medianas distintas.

Prueba de McNemar. Análisis estadístico que permite comparar las diferencias en proporciones cuando los valores se obtuvieron a partir de grupos emparejados (no independientes).

Prueba de signos. Es la prueba estadística no paramétrica que permite comparar dos grupos emparejados; usa como base la calificación relativa de valores efectuada por uno y otro grupo.

Prueba de una cola. Método de significación también estadística en el cual se consideran solamente valores en un extremo o cola de la distribución, para precisar la significación. Se usa cuando el investigador ha conocido anticipadamente la dirección de una relación; véase direccional, hipótesis.

Prueba de Wilcoxon. Prueba de estadística no paramétrica que permite comparar dos grupos pareados, usando para ello la calificación relativa asignada a ciertos valores por uno y otro grupo.

Prueba estadística. Procedimiento analítico que permite al investigador determinar las probabilidades de que los resultados obtenidos de una muestra reflejen los verdaderos resultados de la población, conforme a las leyes de la probabilidad.

Prueba estandarizada. Tipo de prueba publicada con datos normativos y que se aplica a la forma preestablecida.

Prueba exacta de Fisher. Procedimiento estadístico utilizado para evaluar la significancia de la diferencia en proporciones; se usa cuando el tamaño de la muestra es muy pequeño o cuando las celdas de la tabla de contingencia no tienen observaciones.

Prueba sesgada. Prueba de significancia estadística en la cual los valores en un extremo (cola) de una distribución son considerados al determinar la significancia; se usan cuando el investigador ha predicho la dirección de una relación (véase hipótesis direccional).

Prueba *t* de Student. Prueba estadística paramétrica utilizada para analizar la diferencia entre dos medias.

Prueba *U* de Mann-Whitney. Prueba no paramétrica que se usa para evaluar la diferencia entre dos grupos independientes con base en resultados jerarquizados.

Pruebas de dos colas. Prueba de significación estadística en la cual se consideran los valores en ambos extremos de la distribución (colas), para precisar la significación. Utilizada cuando el investigador no ha predicho la dirección de una relación (véase hipótesis no direccional).

Psicometría. Teoría en la cual se basan los principios de cuantificación mental y la aplicación de esta teoría para la creación de instrumentos cuantificacionales. Puede tomar valores entre -1 y $+1$, con $+1$ indicando una relación positiva perfecta, 0 indicando la ausencia de relación y -1 indicando una relación inversa o negativa perfecta (a medida que una crece, otra disminuye).

Puntos de indagación, planteamientos o reactivos. Términos para denotar alguna pregunta de un cuestionario o prueba, o alguna proposición sobre actitud u otra escala (como sería, por ejemplo, el examen final, que consistió en 100 puntos).

Puntuación verdadera. Puntuación o valor hipotético que se obtendría si una medida fuera infalible; es la parte del valor observado que no depende del error aleatorio ni de errores o vicios de medición.

Puntuaciones estándar (puntuación *Z*). Puntuaciones típicas o normalizadas; valores expresados en términos de desviaciones estándar con respecto de la media aritmética. Los valores en bruto son transformados en otros con una media de cero y desviación estándar de uno.

***R*.** Símbolo usado para designar el coeficiente de correlación múltiple, el cual indica la magnitud (pero no la dirección) de la relación entre la variable dependiente y múltiples variables independientes, tomadas a la vez.

***R* de Pearson.** Coeficiente de correlación más extensamente utilizado que representa la magnitud de la relación entre dos variables cuantificadas por lo menos en el nivel de intervalos; también se conoce como correlación producto-momento.

***Rxy* (*r*).** Símbolo usado típicamente para designar un coeficiente de correlación o la relación lineal simple bivariada, el cual resume la magnitud y dirección de una relación entre dos variables (independiente y dependiente).

***R*² (cuadrado del coeficiente de correlación lineal simple o múltiple).** Indica la proporción de la varianza en la variable dependiente, que es explicada o justificada por un grupo de variables independientes. También se conoce como coeficiente de determinación.

Rango (o intervalo). Medida de dispersión o variabilidad, para calcularla se resta la menor calificación de la mayor. Medición de variabilidad que comprende la diferencia entre las cifras máxima y mínima en una distribución cuantitativa.

Razonamiento deductivo. Proceso de hacer predicciones específicas partiendo de principios generales (véase también razonamiento inductivo).

Razonamiento inductivo. El proceso de razonamiento que va de observaciones específicas hasta reglas más generales (véase también razonamiento deductivo).

Reactividad. Distorsión cuantitativa que nace del conocimiento que tiene el sujeto de ser observado o, en forma general, del efecto del propio método de medición.

Regresión. 1. Correlación, representa las relaciones entre conjuntos de calificaciones. 2. Procedimiento estadístico que permite predecir los valores de una variable dependiente con base en los valores de una o más variables independientes.

Regresión logística. Forma especial de regresión en la que la variable dependiente es una variable dicotómica (binaria) no métrica. Aunque existen algunas diferencias, la forma general de interpretación es bastante similar a la de la regresión lineal.

Regresión múltiple. Procedimiento estadístico que permite analizar los efectos simultáneos de dos o más variables independientes (o ajenas) sobre una variable dependiente; ésta debe ser medida sobre una escala de intervalos o de razón.

Regresión simple. Modelo de regresión con variable independiente única.

Relación causal. Relación entre dos variables, de tal forma que la presencia o ausencia de una (causa) rige la presencia o ausencia (el valor) de la otra (efecto).

Relación *F*. Valor obtenido en diversas pruebas estadísticas (por ejemplo, el análisis de varianza paramétrico), con el cual se compara la variación atribuible a diferentes fuentes (entre grupos y dentro de ellos).

Relación funcional. Relación o asociación entre dos variables en la que no puede presuponerse que una variable sea causa de la otra; sin embargo, cabe decir que la variable *X* cambia de valor en función de las alteraciones de la variable *Y*.

Relación negativa (inversamente proporcional). Relación entre dos variables en las que existe tendencia a que los valores altos de una se relacionen con valores bajos de la otra (por ejemplo, conforme la temperatura aumenta, la productividad de la gente disminuye). También se denomina relación inversa.

Relación positiva (directamente proporcional). Relación entre dos variables en la que hay tendencia a que los valores altos en una estén asociados con valores altos en la otra (por ejemplo, conforme la actividad física aumenta, el ritmo cardíaco se eleva).

Relación riesgo/beneficio. Riesgos y beneficios relativos, para el sujeto individual o para la sociedad en pleno, como consecuencia de su participación en un estudio científico.

Replicación. Duplicación de los procedimientos de investigación en un segundo estudio, con el fin de averiguar si los resultados previos se repiten.

Revista de arbitraje. Revista en la que se decide la aceptación o rechazo de artículos originales con fundamento en las recomendaciones de colegas revisores.

Sesgo de la muestra. Error introducido por un proceso de muestreo que favorece ciertas características sobre otras.

Significancia estadística. Término indicador de que es poco probable que los resultados obtenidos mediante un análisis de datos de una muestra se deban al azar, esto dentro de un nivel de probabilidad especificado.

Solomon, diseño de cuatro grupos. Diseño experimental en el que se usa la técnica previa y posterior para un par de grupos experimental y control, y la técnica de sólo retest para un segundo par.

Sondeo. Obtención de información más útil o detallada de un respondiente, en comparación con la que ha expresado voluntariamente durante la primera indagación.

Spearman, *rho* de. Coeficiente de correlación que indica la magnitud de una relación entre variables cuantificadas en la escala ordinal.

Spearman-Brown, fórmula de profecía de. Ecuación para hacer correcciones en un estimado de confiabilidad que se calculó por un método de “dos mitades”.

Sujeto. Persona que participa y aporta datos para un estudio.

Supuestos. Principios básicos aceptados como verdaderos, con arreglo a la lógica o razón, sin prueba o verificación.

Tabla de contingencia. Tabla bidimensional que permite tabular en forma cruzada las frecuencias observadas de dos variables cuantificadas en los niveles nominal u ordinal.

Tabla de números aleatorios. Tabla que contiene centenares de dígitos (de 0 a 9) dispuestos de tal manera que cada uno tenga la misma probabilidad de ir después de cualquier otro; se usa para el muestreo o asignación al azar.

Tabulación cruzada. Cuantificación de los casos que se presentan al considerar en forma simultánea los valores de dos o más variables (por ejemplo, género masculino/femenino, tabulado en forma cruzada con el grado de tabaquismo fumador/no fumador). Típicamente, los resultados se presentan como una tabla cuyas hileras y columnas dependen de los valores de las variables.

Tasa de respuesta. Tasa de participación en una encuesta, calculada al dividir el número de participantes reales entre el número de personas de la muestra.

Tau de Kendall. Coeficiente de correlación que se usa para indicar la magnitud de la relación entre datos de nivel ordinal.

Técnica de Delphi. Método para obtener las opiniones de un panel de expertos. Se interroga en forma individual a los expertos y se circula entre todos ellos un resumen de las opiniones. Luego se vuelve a interrogar a los expertos, proceso que se repite tantas veces como sea necesario hasta lograr determinado consenso.

Técnica de grupos conocidos. Método para estimar la validez de cierto constructo de un instrumento por medio del análisis del grado con el cual el dispositivo separa los grupos que, con base en la predicción, difieren con arreglo a alguna teoría o característica conocida.

Técnicas de “dos mitades”. Método para estimar la congruencia interna (confiabilidad) de un instrumento al correlacionar los valores de una mitad con la otra.

Técnicas de proyección. Métodos para medir atributos psicológicos (valores, aptitudes, personalidad) al dar a los respondientes estímulos no estructurados a los cuales reaccionen.

Tendencia central. Índice estadístico que representa lo más típico de un conjunto de resultados y que surge del punto central de la distribución de dichos resultados. Los tres índices más comunes de tendencia central son moda, mediana y media.

Teoría. Generalización abstracta que ofrece una explicación sistemática acerca de las relaciones entre fenómenos.

Thurstone, escala de. Tipo de escala actitudinal en la que un conjunto de “jueces”, en primer lugar, califica el grado de fa-

vorabilidad de un grupo de proposiciones respecto de algún objeto actitudinal (como sería el aborto) y, después, los sujetos identifican las proposiciones con que concuerdan más.

Tratamiento. Término que denota una intervención o manipulación experimental.

Tratamientos. Niveles de la variable independiente.

V de Cramer. Índice que describe la magnitud de la relación entre datos de nivel nominal y que se utiliza cuando la tabla de contingencia a la cual es aplicado es mayor que 2×2 .

Validez. Grado en que un test mide lo que pretende medir. Correlación entre puntuaciones verdaderas y observadas de la prueba o instrumento de medición y del criterio externo.

Validez concomitante. Grado en que los resultados de un instrumento se correlacionan con determinado criterio externo que fue cuantificado en forma simultánea.

Validez concurrente. Correlación de una prueba con el comportamiento presente o con otras pruebas o medidas existentes. Es un tipo de validez de criterio y predictiva.

Validez convergente. Método de validación de ideas que consiste en evaluar en qué grado son similares (es decir, qué tanto convergen) dos métodos para cuantificar la misma idea.

Validez de constructo. Vinculación de la medición de la prueba con constructos teóricos específicos.

Relación de la prueba con la teoría. La naturaleza de los datos a recoger depende de la concepción teórica del constructo. Es un concepto unificador de validez que integra consideraciones de contenido y de criterio, en un marco general para probar hipótesis racionales acerca de relaciones teóricamente relevantes.

Validez de contenido. Grado en que los reactivos de una prueba evalúan el dominio que ésta pretende cubrir. Relevancia de los reactivos en la conducta por medir.

Validez de criterio. Relación de las calificaciones de la prueba con otras medidas de la misma característica.

Validez externa. Grado en el que los resultados de un estudio o experimento pueden ser generalizados para abarcar escenarios o muestras distintas de las estudiadas.

Validez interna. El grado con el cual puede deducirse que el tratamiento experimental o estudio (variable independiente) y no los factores extraños no controlados, es la causa de los efectos observados.

Validez intrínseca. Postula la existencia de constructos alternativos y se examinan los datos de varias medidas, considerando sus covarianzas.

Validez predictiva o de predicción. 1. Grado en el que un instrumento puede predecir determinado criterio observado en un futuro. 2. Capacidad de una medida para vaticinar una conducta, las dos

mediciones se realizan simultáneamente. Tipo secundario de validez de criterio.

Valor p . En pruebas estadísticas, la probabilidad de que los resultados obtenidos se deban exclusivamente al azar, la probabilidad de cometer un error tipo I.

Variabilidad. Grado con el cual los valores de un conjunto de puntuaciones difieren extensamente o están “dispersos”, por

ejemplo, cabría esperar mayor variabilidad de la edad dentro de un hospital que dentro de un asilo.

Variable. Cualquier característica o atributo susceptible de medirse, es decir, adopta valores diferentes, dentro de la población en estudio (por ejemplo, la temperatura del cuerpo, la edad o el ritmo cardíaco).

Variable activa. Variable que el investigador crea o manipula.

Variable ajena o extraña. Variable que vuelve confusa la relación entre las variables independiente y dependiente y que debe ser controlada, ya sea a través del diseño de la investigación o mediante procedimientos estadísticos.

Variable categórica. Variable cuyos valores son escalonados y definidos (por ejemplo, el estado civil de una persona) en vez de ubicarse en cualquier punto de un continuo.

Variable continua. Variable que puede tomar cualquier valor dentro de los límites de una escala continua (por ejemplo, la estatura).

Variable de criterio (media del criterio). La cualidad o atributo usado para medir el efecto de una variable independiente; se usa a veces como un sinónimo de variable dependiente.

Variable dependiente. Variable de interés resultante. Es la que, según la hipótesis, depende de otra variable o es causada por ésta (llamada variable independiente). También se le denomina variable de criterio.

Variable dicotómica. Variable que tiene sólo dos valores o categorías (como sería el caso del sexo).

Variable endógena. En análisis de trayectorias, variable cuya fluctuación depende de otras variables en el modelo.

Variable exógena. En el análisis de trayectorias, variable cuyos determinantes se encuentran fuera del modelo.

Variable extraña. Variable que confunde la relación entre las variables independiente y dependiente, y que es necesario controlar por el diseño de la investigación o por medio de técnicas estadísticas (por ejemplo, en un estudio del efecto de la edad de la mujer en el número de partos prematuros, la clase social y el origen cultural serían variables extrañas).

Variable independiente. Variable que, según el investigador, causa o influye en la variable dependiente; en investigación, esta variable es la que se somete a manipulación.

Variable mediadora. Variable que media o actúa como “correve y dile” en una cadena que relaciona a otras dos variables (por ejemplo, cabe decir que la capacidad de adaptación media la relación entre los acontecimientos estresantes y la angustia).

Variable simulada. Variables dicotómicas creadas para muchos análisis estadísticos multivariados; utilizan típicamente los códigos 0 y 1 (por ejemplo, mujer = 1 y varón = 0).

Variables de atributo. Características preexistentes de la entidad en investigación, las cuales son simplemente observadas y cuantificadas por quien investiga.

Varianza. Medida de variabilidad o dispersión que es igual al cuadrado de la desviación estándar.

“Vicio” (tendenciosidad, error o parcialidad). Cualquier influencia que distorsiona los resultados de un estudio.



PROBABILIDAD Y ESTADÍSTICA

para las ciencias sociales
del comportamiento y la salud

Probabilidad y estadística para las ciencias sociales del comportamiento y la salud, primera edición en español, es una versión adaptada del libro *Introducción a la probabilidad y estadística*, decimocuarta edición, de William Mendenhall III, Robert J. Beaver y Barbara M. Beaver. En ella se han incluido temas que son indispensables para los profesionales de las ciencias sociales del comportamiento y la salud que, ya sea en su trabajo cotidiano o en la realización de estudios e investigaciones, deben manejar, analizar e interpretar grandes cantidades de datos: estadística y probabilidad dentro del contexto de la investigación, medición estándar, distribuciones, variables aleatorias discretas (binomial, de Poisson e hipergeométrica), estimación estadística de muestras, métodos para estimar los parámetros poblacionales y para datos categóricos, entre otros.

La obra conserva de la versión original algunas características importantes, pero en esta ocasión enfocadas en el estudio de la estadística desde la perspectiva de las ciencias sociales.



Visite nuestro sitio en <http://latinoamerica.cengage.com>

https://t.me/universitarios_info

<https://www.jamarana.com>

<https://t.me/universitarios>

ISBN-13: 978-607-526-310-6
ISBN-10: 607-526-310-1



9 786075 263106