9

Estimación e intervalos de confianza

La American Restaurant Association recopiló información sobre las veces que los matrimonios comen fuera de casa cada semana. Una encuesta de 60 parejas demostró que la cantidad media de comidas fuera de casa era de 2.76 por semana, con una desviación estándar de 0.75. Defina un intervalo de confianza de 97% para la media de la población. (Vea el objetivo 4 y el ejercicio 36).

Objetivos de aprendizaje

Al concluir el capítulo, será capaz de:

OA1 Definir un *estimador puntual.*

OA2 Definir *nivel de confianza.*

OA3 Construir el intervalo de confianza de la media poblacional cuando se conoce la desviación estándar de la población.

OA4 Construir el intervalo de confianza de una media poblacional cuando no se conoce la desviación estándar de la población.

OA5 Construir el intervalo de confianza de una proporción de la población.

OA6 Calcular el tamaño de la muestra necesario para estimar una proporción de la población o una media poblacional.

OA7 Ajustar el intervalo de confianza de poblaciones finitas.

9.1 Introducción

En el capítulo anterior se inició el estudio de la estadística inferencial. En él se presentaron las razones y métodos de muestreo. Las razones del muestreo son las siguientes:

- Entrar en contacto con toda la población consume demasiado tiempo.
- El costo de estudiar todos los elementos de la población es muy alto.
- Por lo general, los resultados de la muestra resultan adecuados.
- Algunas pruebas resultan negativas.
- Es imposible revisar todos los elementos.

Existen varios métodos de muestreo. El aleatorio simple es el que más se utiliza. En este tipo de muestreo, cada miembro de la población posee las mismas posibilidades de ser seleccionado como parte de la muestra. Otros métodos de muestreo son el sistemático, el estratificado y el muestreo por conglomerados.

El capítulo 8 presenta información relacionada con la media, la desviación estándar o la forma de la población. En la mayoría de las situaciones de negocios, dicha información no se encuentra disponible. En realidad, el propósito del muestreo es calcular de forma aproximada algunos de estos valores. Por ejemplo, se selecciona una muestra de una población y se utiliza la media de la muestra para aproximar la media de la población.

En este capítulo se estudian diversos aspectos importantes del muestreo. El primer paso es el estudio del **estimador puntual**. Un estimador puntual consiste en un solo valor (punto) deducido de una muestra para estimar el valor de una población. Por ejemplo, suponga que elige una muestra de 50 ejecutivos de nivel medio y le pregunta a cada uno de ellos la cantidad de horas que laboró la semana pasada. Se calcula la media de esta muestra de 50 y se utiliza el valor de la media muestral como estimador puntual de la media poblacional desconocida. Ahora bien, un estimador puntual es un solo valor. Un enfoque que arroja más información consiste en presentar un intervalo de valores del que se espera que se estime el parámetro poblacional. Dicho intervalo de valores recibe el nombre de **intervalo de confianza**.

En los negocios, a menudo es necesario determinar el tamaño de una muestra. ¿Con cuántos electores debe ponerse en contacto una compañía dedicada a realizar encuestas con el fin de predecir los resultados de las elecciones? ¿Cuántos productos se necesitan analizar para garantizar el nivel de calidad? En este capítulo también se explica una estrategia para determinar el tamaño adecuado de la muestra.

9.2 Estimadores puntuales e intervalos de confianza de una media

Un estimador puntual es un estadístico único para calcular un parámetro poblacional. Suponga que Best Buy, Inc., desea estimar la edad media de los compradores de televisores de plasma de alta definición; selecciona una muestra aleatoria de 50 compradores recientes, determina la edad de cada uno de ellos y calcula la edad media de los compradores de la muestra. La media de esta muestra es un estimador puntual de la media de la población.

ESTIMADOR PUNTUAL Estadístico calculado a partir de información de la muestra para estimar el parámetro poblacional.

Los siguientes ejemplos ilustran los estimadores puntuales de medias poblacionales.

1. El turismo constituye una fuente importante de ingresos para muchos países caribeños, como Barbados. Suponga que la Oficina de Turismo de Barbados desea un cálculo aproximado de la cantidad media que gastan los turistas que visitan el país. No resultaría viable ponerse en contacto con cada turista. Por consiguiente, se selecciona al azar a 500 turistas en el momento en que salen del país y se les pregunta los detalles de los gastos que realizaron durante su visita a la isla. La cantidad media que gastó la muestra de 500 turistas constituye un cálculo aproximado del parámetro poblacional desconocido. Es decir, la media muestral es el estimador puntual de la media poblacional.

Estadística en acción

En un lugar visible de la ventanilla de todos los automóviles nuevos aparece una calcomanía con un cálculo aproximado del ahorro de gasolina, según lo requiere la Environmental Protection Agency (EPA). Con frecuencia, el ahorro de gasolina constituye un factor importante para que el consumidor elija un automóvil nuevo, debido a los costos del combustible o a cuestiones ambientales. Por ejemplo, los cálculos aproximados del rendimiento de combustible de un BMW 328i Sedán 2010 (automático de 6 cilindros) son de 18 millas por galón (mpg) en la ciudad y de 28 mpg en carretera. La EPA reconoce que el verdadero ahorro de gasolina puede diferir de los cálculos aproximados: "Ninguna prueba puede simular todas las combinaciones de condiciones y clima posibles, del comportamiento del conductor y hábitos en el cuidado del automóvil. El millaje real depende de cómo, cuándo y dónde se maneje el vehículo. La EPA descubrió que las mpg que obtiene la mayoría de los conductores difieren de los cálculos aproximados por unas cuantas mpg." De hecho, la calcomanía del parabrisas también incluye una estimación del intervalo relativo al ahorro de combustible: 14 a 22 mpg en ciudad y de 23 a 33 mpg en carretera.

OA1 Definir un *estimador puntual.*

2. Litchfield Home Builders, Inc., construye casas en la zona sureste de Estados Unidos. Una de las principales preocupaciones de los compradores es la fecha en que concluirán las obras. Hace poco Litchfield comunicó a sus clientes: "Su casa quedará terminada en 45 días a partir de la fecha de instalación de los muros." El departamento de atención a clientes de Litchfield desea comparar este ofrecimiento con experiencias recientes. Una muestra de 50 casas terminadas este año reveló que el número medio de días de trabajo a partir del inicio de la construcción de los muros a la terminación de la casa fue de 46.7 días hábiles. ¿Es razonable concluir que la media poblacional aún es de 45 días y que la diferencia entre la media muestral (46.7 días) y la media de población propuesta es un error de muestreo? En otras palabras, ¿la media muestral difiere en forma significativa de la media poblacional?

3. Estudios médicos recientes indican que el ejercicio constituye una parte importante de la salud general de una persona. El director de recursos humanos de OCF, fabricante importante de vidrio, desea calcular la cantidad de horas semanales que los empleados dedican al ejercicio. Una muestra de 70 empleados revela que la cantidad media de horas de ejercicio de la semana pasada fue de 3.3. La media muestral de 3.3 horas aproxima la media poblacional desconocida, la media de horas de ejercicio de todos los empleados.

La media muestral, $\bar{X}$, no es el único estimador puntual de un parámetro poblacional. Por ejemplo, *p,* una proporción muestral, es un estimador puntual de π, la proporción poblacional; y *s,* la desviación estándar muestral, es un estimador puntual de σ, la desviación estándar poblacional.

9.3 Intervalos de confianza de una media poblacional

Ahora bien, un estimador puntual sólo dice parte de la historia. Aunque se espera que el estimador puntual se aproxime al parámetro poblacional, sería conveniente medir cuán próximo se encuentra en realidad. Un intervalo de confianza sirve para este propósito. Por ejemplo, se estima que el ingreso anual medio de los trabajadores de la construcción en el área de Nueva York a Nueva Jersey es de \$85 000. Un intervalo de este valor aproximado puede oscilar entre \$81 000 y \$89 000. Para describir cuánto es posible confiar en que el parámetro poblacional se encuentre en el intervalo se debe generar un enunciado probabilístico. Por ejemplo: se cuenta con 90% de seguridad de que el ingreso anual medio de los trabajadores de la construcción en el área de Nueva York a Nueva Jersey se encuentra entre \$81 000 y \$89 000.

OA2 Definir *nivel de confianza.*

INTERVALO DE CONFIANZA Conjunto de valores que se forma a partir de una muestra de datos de forma que exista la posibilidad de que el parámetro poblacional ocurra dentro de dicho conjunto con una probabilidad específica. La probabilidad específica recibe el nombre de *nivel de confianza.*

Para calcular el intervalo de confianza, consideraremos dos situaciones:

- Utilizamos los datos de la muestra para calcular μ con $\bar{X}$, mientras que la desviación estándar de la población (σ) es conocida.
- Utilizamos los datos de la muestra para calcular μ con $\bar{X}$, mientras que la desviación estándar de la población es desconocida. En este caso, sustituimos la desviación estándar de la(s) muestra(s) por la desviación estándar de la población (σ).

Existen diferencias importantes en las suposiciones entre estas dos situaciones. Consideraremos primero el caso donde se conoce σ.

Desviación estándar de la población conocida (σ)

Un intervalo de confianza se calcula con el empleo de dos estadísticos: la media muestral $\overline{X}$ y la desviación estándar. De los capítulos anteriores, usted sabe que la desviación estándar es un estadístico importante, porque mide la dispersión, o la amplitud, de una población o de una muestra de distribución. Cuando se calcula un intervalo de confianza, se utiliza la desviación estándar para estimar el rango del intervalo de confianza.

Para demostrar la idea del intervalo de confianza, se comienza con una suposición simple: que conocemos el valor de la desviación estándar de la población, σ. Conocerla permite simplificar el desarrollo del intervalo de confianza, porque podemos utilizar la distribución normal estándar que se estudió en el capítulo 8.

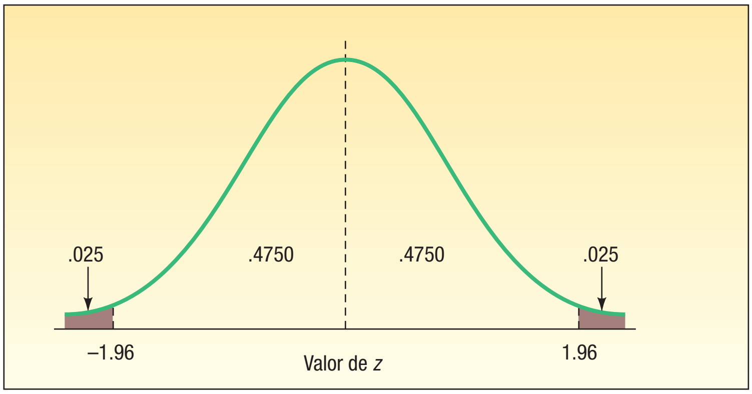
Recuerde que la distribución muestral de la media es la distribución de todas las medias muestrales, $\overline{X}$, con tamaño de la muestra, n, de una población. Se conoce la desviación estándar de la población, σ. A partir de esta información, y del teorema central del límite, sabemos que la distribución muestral sigue una distribución de probabilidad normal con una media μ y una desviación estándar $\sigma/\sqrt{n}$. Recuerde también que este valor recibe el nombre de error estándar.

Los resultados del teorema central del límite permiten afirmar lo siguiente con respecto a los intervalos de confianza utilizando el estadístico z:

1. Noventa y cinco por ciento de las medias muestrales seleccionadas de una población se encontrará dentro de 1.96 errores estándares (desviación estándar de las medias muestrales de la media poblacional, μ.
2. Noventa y nueve por ciento de las medias muestrales se encontrará a 2.58 errores estándares de la media poblacional.

Los intervalos calculados de esta manera proporcionan ejemplos de los *niveles de confianza* y reciben el nombre de **intervalo de confianza de 95%** e **intervalo de confianza de 99%**. Por lo tanto, *95%* y *99%* son los niveles de confianza y se refieren al porcentaje de intervalos similarmente construidos que incluirían el parámetro a calcular, en este caso, μ.

¿Cómo se obtienen los valores de 1.96 y 2.58? En el caso del intervalo de confianza de 95%, vea el siguiente diagrama y consulte el apéndice B.1 para determinar los valores z adecuados. Localice 0.4750 en el cuerpo de la tabla. Lea los valores del renglón y la columna correspondientes. El valor es 1.96. Por lo tanto, la probabilidad de hallar un valor z entre 0 y 1.96 es de 0.4750. Asimismo, la probabilidad de encontrar un valor z en el intervalo entre 0 y -1.96 también es de 0.4750. Al combinar ambos valores, la probabilidad de estar en el intervalo -1.96 y 1.96 es de 0.9500. En la siguiente página encontrará una porción del apéndice B.1. El valor z del nivel de confianza de 90% se determina de forma similar. Éste es de 1.65. En el caso de un nivel de confianza de 99%, el valor z es de 2.58.



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936

OA3 Construir el intervalo de confianza de la media poblacional cuando se conoce la desviación estándar de la población.

¿Cómo determinar el intervalo de confianza de 95%? La amplitud del intervalo se determina por medio del nivel de confianza y de la magnitud del error estándar de la media. Ya se ha descrito la forma de encontrar el valor *z* de un nivel de confianza particular. Recuerde que, según el capítulo anterior [vea la fórmula (8-1), p. 285], el error estándar de la media indica la variación de la distribución de las medias muestrales. Se trata, en realidad, de la desviación estándar de la distribución muestral de medias. La fórmula se repite en seguida:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

donde:

$\sigma_{\bar{x}}$ es el símbolo del error estándar de la media; se utiliza la letra griega porque se trata de un valor poblacional, y el subíndice $\bar{x}$ recuerda que se refiere a la distribución de las medias muestrales.

σ es la desviación estándar poblacional.

n es el número de observaciones en la muestra.

La magnitud del error estándar se ve afectada por dos valores. El primero es la desviación estándar de la población. Mientras mayor sea la desviación estándar de la población, σ, mayor será $\sigma/\sqrt{n}$. Si la población es homogénea, de modo que genere una desviación estándar poblacional pequeña, el error estándar también será pequeño. Sin embargo, la cantidad de observaciones de la muestra también afecta al error estándar. Una muestra grande generará un error estándar pequeño en la estimación, lo que indicará que hay menos variabilidad en las medias muestrales.

Para explicar estos conceptos, considere el siguiente ejemplo. Del Monte Foods, Inc., distribuye duraznos en trozo en latas de 4 onzas. Para asegurarse de que cada lata contenga por lo menos la cantidad que se requiere, Del Monte establece que el proceso de llenado debe verter 4.01 onzas de duraznos y almíbar en cada lata. Así, 4.01 es la media poblacional. Por supuesto, no toda lata contendrá exactamente 4.01 onzas de duraznos y almíbar. Algunas latas contendrán más y otras menos. Suponga que la desviación estándar del proceso es de 0.04 onzas. También suponga que el proceso se rige por la distribución de probabilidad normal. Ahora se selecciona una muestra aleatoria de 64 latas y se determina la media de la muestra. Ésta es de 4.015 onzas de duraznos y almíbar. El intervalo de confianza de 95% de la media poblacional de esta muestra particular es:

$$4.015 \pm 1.96(.04/\sqrt{64}) = 4.015 \pm .0098$$

El nivel de confianza de 95% se encuentra entre 4.0052 y 4.0248. Por supuesto, en este caso, la media de población de 4.01 onzas se encuentra en este intervalo. Pero no siempre será así. En teoría, si selecciona 100 muestras de 64 latas de la población, se calcula la media muestral y se crea un intervalo de confianza basado en cada media *muestral,* se esperaría encontrar una media *poblacional* de aproximadamente 95 de los 100 intervalos.

Los siguientes cálculos en el caso de un intervalo de confianza de 95% se resumen con la siguiente fórmula:

$$\bar{X} \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

De manera similar, un intervalo de confianza de 99% se calcula de la siguiente manera:

$$\bar{X} \pm 2.58 \frac{\sigma}{\sqrt{n}}$$

Como ya se señaló, los valores de 1.96 y 2.58 son valores z correspondientes a 95% medio y 99% medio de las observaciones, respectivamente.

No hay restricción a los niveles de confianza de 95 y 99%. Es posible seleccionar cualquier nivel de confianza entre 0 y 100% y encontrar el valor correspondiente de z. En general, un intervalo de confianza de la media poblacional, cuando se conoce la desviación estándar poblacional, se calcula de la siguiente manera:

INTERVALO DE CONFIANZA DE LA MEDIA POBLACIONAL CON UNA σ CONOCIDA

$$\bar{X} \pm z \frac{\sigma}{\sqrt{n}} \qquad \textbf{(9-1)}$$

En esta fórmula, z depende del nivel de confianza. Por consiguiente, con un nivel de confianza de 92%, el valor z en la fórmula (9-1) es de 1.75. El valor de z proviene del apéndice B.1. Esta tabla se basa en la mitad de la distribución normal, por lo que 0.9200/2 = 0.4600. El valor más próximo en el cuerpo de la tabla es de 0.4599, y el valor z correspondiente es de 1.75.

Con frecuencia, también se utiliza el nivel de confianza de 90%. En este caso, se desea que el área entre 0 y z sea de 0.4500, que se determina con la operación 0.9000/2. Para determinar el valor z con este nivel de confianza, descienda por la columna izquierda del apéndice B.1 hasta 1.6, y después recorra las columnas con los encabezamientos 0.04 y 0.05. El área correspondiente al valor z de 1.64 es 0.4495, y de 1.65, 0.4505. Para proceder con cautela, utilice 1.65. Intente buscar los siguientes niveles de confianza y verifique sus respuestas con los valores correspondientes de z indicados a la derecha.

Nivel de confianza	Probabilidad media más cercana	Valor z
80%	.3997	1.28
94%	.4699	1.88
96%	.4798	2.05

El siguiente ejemplo muestra los detalles para calcular un intervalo de confianza e interpreta el resultado.

Ejemplo

La American Management Association desea información acerca del ingreso medio de los gerentes de la industria del menudeo. Una muestra aleatoria de 256 gerentes revela una media muestral de $45 420. La desviación estándar de esta muestra es de $2 050. A la asociación le gustaría responder las siguientes preguntas:

1. ¿Cuál es la media de la población?
2. ¿Cuál es un conjunto de valores razonable de la media poblacional?
3. ¿Cómo se deben interpretar estos resultados?

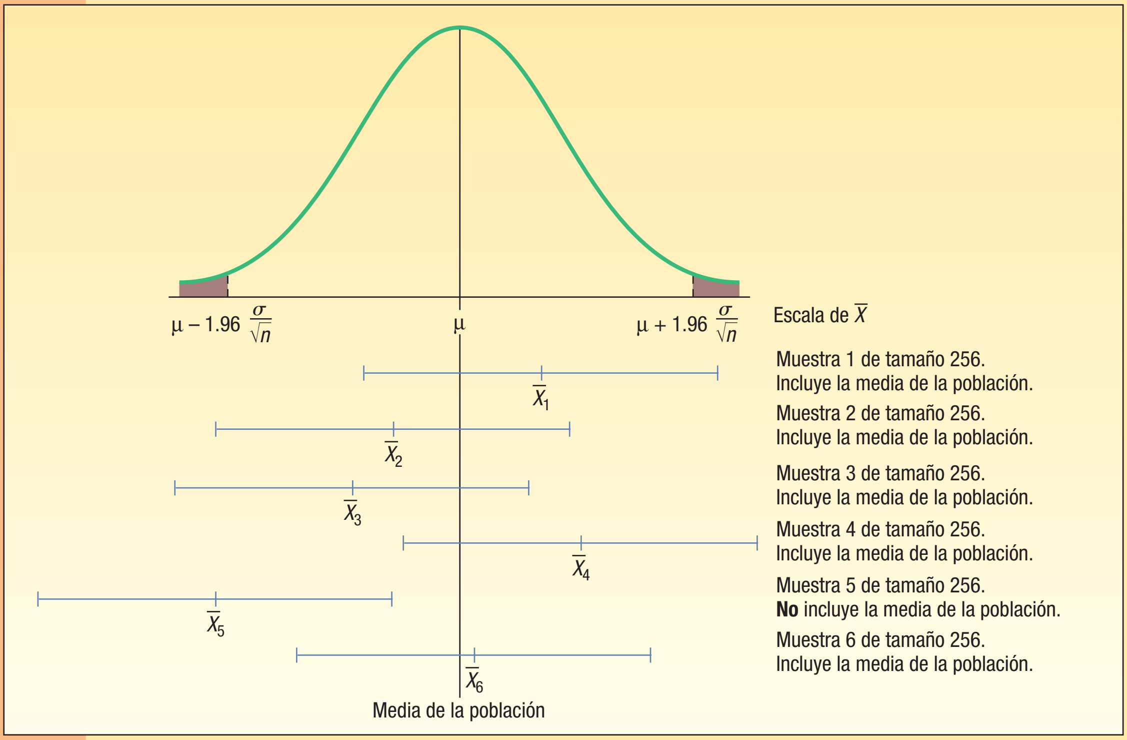
Solución

En general, las distribuciones de los salarios e ingresos tienen un sesgo positivo, pues unos cuantos individuos ganan considerablemente más que otros, lo cual sesga la distribución en dirección positiva. Por fortuna, el teorema central del límite estipula que, si se selecciona una muestra grande, la distribución de las medias muestrales tenderá a seguir la distribución normal. En este caso, una muestra de 256 gerentes es lo bastante grande para suponer que la distribución muestral tenderá a seguir la distribución normal. A continuación se responden las preguntas planteadas en el ejemplo.

1. **¿Cuál es la media de la población?** En este caso se ignora. Sí se sabe que la media de la muestra es de \$45 420. De ahí que la mejor estimación del valor de población sea el estadístico de la muestra correspondiente. Por consiguiente, la media de la muestra de \$45 420 constituye un *estimador puntual* de la media poblacional desconocida.
2. **¿Cuál es el conjunto de valores razonable de la media poblacional?** La asociación decide utilizar un nivel de confianza de 95%. Para determinar el intervalo de confianza correspondiente, se aplica la fórmula (9-1):

$$\overline{X} \pm z \frac{\sigma}{\sqrt{n}} = \$45\,420 \pm 1.96 \frac{\$2\,050}{\sqrt{256}} = \$45\,420 \pm \$251$$

   Es costumbre redondear estos puntos extremos a \$45 169 y \$45 671. Estos puntos extremos reciben el nombre de *límites de confianza.* El grado de confianza o *nivel de confianza* es de 95%, y el intervalo de confianza abarca de \$45 169 a \$45 671. Con frecuencia, ±\$251 se conoce como *margen de error.*
3. **¿Cómo se deben interpretar estos resultados?** Suponga que selecciona varias muestras de 256 gerentes, tal vez varios cientos. Para cada muestra, calcula la media y después construye un intervalo de confianza de 95%, como en la sección anterior. Puede esperar que alrededor de 95% de estos intervalos de confianza contenga la media de la *población.* Cerca de 5% de los intervalos no contendrán el ingreso anual medio poblacional, μ. No obstante, un intervalo de confianza particular contiene el parámetro poblacional o no lo contiene. El siguiente diagrama muestra los resultados de seleccionar muestras de la población de gerentes de la industria del menudeo: se calcula la media de cada una y, posteriormente, con la fórmula (9-1), se determina un intervalo de confianza de 95% de la media poblacional. Observe que no todos los intervalos incluyen la media poblacional. Los dos puntos extremos de la quinta muestra son inferiores a la media poblacional. Esto se debe al error de muestreo, que constituye el riesgo que se asume cuando se selecciona el nivel de confianza.



Simulación por computadora

Con ayuda de una computadora es posible seleccionar al azar muestras de una población, calcular con rapidez el intervalo de confianza y mostrar la frecuencia con que los intervalos de confianza incluyen, aunque no siempre, el parámetro de la población. El siguiente ejemplo aclarará esta cuestión.

Ejemplo

Tras varios años en el negocio de renta de automóviles, Town Bank sabe que la distancia media recorrida en un contrato de cuatro años es de 50 000 millas, y la desviación estándar, de 5 000. Suponga que desea encontrar la proporción de los intervalos de confianza de 95% que incluirán la media poblacional de 50 000 con el sistema de software de estadística de Minitab. Para facilitar los cálculos, trabaje en miles de millas, en lugar de unidades de milla. Seleccione 60 muestras aleatorias de tamaño 30 de una población con una media de 50, y una desviación estándar de 5.

Solución

Los resultados de 60 muestras aleatorias de 30 automóviles cada una se resumen en la captura de pantalla que aparece a continuación. De los 60 intervalos de confianza con un nivel de confianza de 95%, 2% o 3.33% no incluyen la media poblacional de 50. Se resaltan los intervalos (C3 y C59) que *no* incluyen la media poblacional. Con la cifra de 3.33% se aproxima al cálculo de que 5% de los intervalos no incluirán la media poblacional, y que 58 de 60, es decir, 96.67%, se aproxima a 95 por ciento.

Para explicar el primer cálculo con mayor detalle, Minitab comienza con la selección de una muestra aleatoria de 30 observaciones de una población con una media de 50 y una desviación estándar de 5. La media de estas 30 observaciones es de 50.053. El error muestral es de 0.053, que se determina por medio de $\bar{X} - \mu = 50.053 - 50.000$. Los puntos extremos del intervalo de confianza son 48.264 y 51.842. Estos puntos extremos se determinan con la fórmula (9-1):

$$\bar{X} \pm 1.96 \frac{\sigma}{\sqrt{n}} = 50.053 \pm 1.96 \frac{5}{\sqrt{30}} = 50.053 \pm 1.789$$

One-Sample Z:

The assumed sigma = 5

Variable	N	Mean	StDev	SE Mean	95.0% CI
C1	30	50.053	5.002	0.913	(48.264, 51.842)
C2	30	49.025	4.450	0.913	(47.236, 50.815)
C3	30	52.023	5.918	0.913	(50.234, 53.812)
C4	30	50.056	3.364	0.913	(48.267, 51.845)
C5	30	49.737	4.784	0.913	(47.948, 51.526)
C6	30	51.074	5.495	0.913	(49.285, 52.863)
C7	30	50.040	5.930	0.913	(48.251, 51.829)
C8	30	48.910	3.645	0.913	(47.121, 50.699)
C9	30	51.033	4.918	0.913	(49.244, 52.822)
C10	30	50.692	4.571	0.913	(48.903, 52.482)
C11	30	49.853	4.525	0.913	(48.064, 51.642)
C12	30	50.286	3.422	0.913	(48.497, 52.076)
C13	30	50.257	4.317	0.913	(48.468, 52.046)
C14	30	49.605	4.994	0.913	(47.816, 51.394)
C15	30	51.474	5.497	0.913	(49.685, 53.264)
C16	30	48.930	5.317	0.913	(47.141, 50.719)
C17	30	49.870	4.847	0.913	(48.081, 51.659)
C18	30	50.739	6.224	0.913	(48.950, 52.528)
C19	30	50.979	5.520	0.913	(49.190, 52.768)
C20	30	48.848	4.130	0.913	(47.059, 50.638)
C21	30	49.481	4.056	0.913	(47.692, 51.270)
C22	30	49.183	5.409	0.913	(47.394, 50.973)
C23	30	50.084	4.522	0.913	(48.294, 51.873)
C24	30	50.866	5.142	0.913	(49.077, 52.655)
C25	30	48.768	5.582	0.913	(46.979, 50.557)
C26	30	50.904	6.052	0.913	(49.115, 52.694)
C27	30	49.481	5.535	0.913	(47.691, 51.270)
C28	30	50.949	5.916	0.913	(49.160, 52.739)

Variable	N	Mean	StDev	SE Mean	95.0% CI
C29	30	49.106	4.641	0.913	(47.317, 50.895)
C30	30	49.994	5.853	0.913	(48.205, 51.784)
C31	30	49.601	5.064	0.913	(47.811, 51.390)
C32	30	51.494	5.597	0.913	(49.705, 53.284)
C33	30	50.460	4.393	0.913	(48.671, 52.249)
C34	30	50.378	4.075	0.913	(48.589, 52.167)
C35	30	49.808	4.155	0.913	(48.019, 51.597)
C36	30	49.934	5.012	0.913	(48.145, 51.723)
C37	30	50.017	4.082	0.913	(48.228, 51.806)
C38	30	50.074	3.631	0.913	(48.285, 51.863)
C39	30	48.656	4.833	0.913	(46.867, 50.445)
C40	30	50.568	3.855	0.913	(48.779, 52.357)
C41	30	50.916	3.775	0.913	(49.127, 52.705)
C42	30	49.104	4.321	0.913	(47.315, 50.893)
C43	30	50.308	5.467	0.913	(48.519, 52.097)
C44	30	49.034	4.405	0.913	(47.245, 50.823)
C45	30	50.399	4.729	0.913	(48.610, 52.188)
C46	30	49.634	3.996	0.913	(47.845, 51.424)
C47	30	50.479	4.881	0.913	(48.689, 52.268)
C48	30	50.529	5.173	0.913	(48.740, 52.318)
C49	30	51.577	5.822	0.913	(49.787, 53.366)
C50	30	50.403	4.893	0.913	(48.614, 52.192)
C51	30	49.717	5.218	0.913	(47.927, 51.506)
C52	30	49.796	5.327	0.913	(48.007, 51.585)
C53	30	50.549	4.680	0.913	(48.760, 52.338)
C54	30	50.200	5.840	0.913	(48.410, 51.989)
C55	30	49.138	5.074	0.913	(47.349, 50.928)
C56	30	49.667	3.843	0.913	(47.878, 51.456)
C57	30	49.603	5.614	0.913	(47.814, 51.392)
C58	30	49.441	5.702	0.913	(47.652, 51.230)
C59	30	47.873	4.685	0.913	(46.084, 49.662)
C60	30	51.087	5.162	0.913	(49.297, 52.876)

Autoevaluación 9-1

Bun-and-Run es una franquicia de comida rápida de la zona noreste, la cual se especializa en hamburguesas de media onza, y sándwiches de pescado y de pollo. También ofrece refrescos y papas a la francesa. El departamento de planeación de la firma informa que la distribución de ventas diarias de los restaurantes tiende a seguir la distribución normal. La desviación estándar de la distribución de ventas diarias es de $3 000. Una muestra de 40 mostró que las ventas medias diarias suman $20 000.

a) ¿Cuál es la media de la población?
b) ¿Cuál es la mejor estimación de la media de la población? ¿Qué nombre recibe este valor?
c) Construya un intervalo de confianza de 99% de la media poblacional.
d) Interprete el intervalo de confianza.

Ejercicios

1. Se toma una muestra de 49 observaciones de una población normal con una desviación estándar de 10. La media de la muestra es de 55. Determine el intervalo de confianza de 99% de la media poblacional.

2. Se toma una muestra de 81 observaciones de una población normal con una desviación estándar de 5. La media de la muestra es de 40. Determine el intervalo de confianza de 95% de la media poblacional.

3. Se selecciona una muestra de 250 observaciones de una población normal en la cual la desviación estándar poblacional se sabe que es de 25. La media de la muestra es de 20.
   a) Determine el error estándar de la media.
   b) Explique por qué se debe utilizar la fórmula (9-1) para determinar el intervalo de confianza de 95 por ciento.
   c) Determine el intervalo de confianza de 95% de la media de la población.

4. Suponga que desea un nivel de confianza de 85%. ¿Qué valor utilizaría para multiplicar el error estándar de la media?

5. Una empresa de investigación llevó a cabo una encuesta para determinar la cantidad media que los fumadores gastan en cigarrillos durante una semana. La empresa descubrió que la distribución de cantidades que gastan por semana tendía a seguir una distribución normal, con una desviación estándar de $5. Una muestra de 49 fumadores reveló que $\overline{X} = \$20$.

a) ¿Cuál es el estimador puntual de la media de la población? Explique lo que indica.

b) Con el nivel de confianza de 95%, determine el intervalo de confianza de μ. Explique lo que significa.

6. Repase el ejercicio anterior. Suponga que se tomó una muestra de 64 fumadores (en lugar de 49). Suponga que la media muestral es la misma.

a) ¿Cuál es el estimador del intervalo de confianza de 95% de μ?

b) Explique por qué este intervalo de confianza es más reducido que el que se determinó en el ejercicio anterior.

7. Bob Nale es propietario de Nale's Quick Fill. A Bob le gustaría estimar la cantidad de galones de gasolina que vendió. Suponga que la cantidad de galones vendidos tiende a seguir una distribución normal, con una desviación estándar de 2.30 galones. De acuerdo con sus registros, selecciona una muestra aleatoria de 60 ventas y descubre que la cantidad media de galones vendidos es de 8.60.

a) ¿Cuál es el estimador puntual de la media poblacional?

b) Establezca un intervalo de confianza de 99% de la media poblacional.

c) Interprete el significado del inciso *b*).

8. La doctora Patton es profesora de inglés. Hace poco contó el número de faltas de ortografía que cometió un grupo de estudiantes en sus ensayos. Observó que la distribución de las faltas de ortografía por ensayo se regía por la distribución normal con una desviación estándar de 2.44 palabras por ensayo. En su clase de 40 alumnos de las 10 de la mañana, el número medio de palabras con faltas de ortografía fue de 6.05. Construya un intervalo de confianza de 95% del número medio de palabras con faltas de ortografía en la población de ensayos.

Desviación estándar poblacional σ desconocida

OA4 Construir el intervalo de confianza de una media poblacional cuando no se conoce la desviación estándar de la población.

En la sección anterior se supuso que se conocía la desviación estándar de la población. En el caso de las latas de duraznos de 4 onzas de Del Monte, quizá había una gran cantidad de mediciones del proceso de llenado. Por consiguiente, resulta razonable suponer que se dispone de la desviación estándar de la población. Sin embargo, en la mayoría de los casos de muestreo, no se conoce la desviación estándar de la población (σ). He aquí algunos ejemplos en los que se pretende estimar las medias poblacionales y es poco probable que se conozcan las desviaciones estándares. Suponga que cada uno de los siguientes estudios se relaciona con estudiantes de la West Virginia University.

- El decano de la Facultad de Administración desea estimar la cantidad media de horas de estudiantes de tiempo completo con trabajos remunerativos cada semana. Selecciona una muestra de 30 estudiantes; se pone en contacto con cada uno de ellos y les pregunta cuántas horas laboraron la semana pasada. De acuerdo con la información de la muestra, puede calcular la media muestral, pero no es probable que conozca o pueda determinar la desviación estándar *poblacional* (σ) que se requiere en la fórmula (9-1). Puede calcular la desviación estándar de la muestra y utilizarla como estimador, pero quizá no conocería la desviación estándar de la población.
- La docente a cargo del asesoramiento de los estudiantes desea estimar la distancia que el estudiante común viaja cada día de su casa a clases. Ella selecciona una muestra de 40 estudiantes, se pone en contacto con ellos y determina la distancia que recorre cada uno, de su casa al centro universitario. De acuerdo con los datos de la muestra, calcula la distancia media de viaje, es decir $\overline{X}$. No es probable que se conozca o se encuentre disponible la desviación estándar de la población, lo cual, nuevamente, torna obsoleta la fórmula (9-1).
- El director de créditos estudiantiles desea conocer el monto medio de créditos estudiantiles en el momento de la graduación. El director selecciona una muestra de 20 estudiantes graduados y se pone en contacto con cada uno para obtener la información. De acuerdo con la información con la que cuenta, puede estimar la cantidad media. Sin embargo, para establecer un intervalo de confianza con la fórmula (9-1), es necesaria la desviación estándar de la población. No es probable que esta información se encuentre disponible.

Estadística en acción

William Gosset nació en Inglaterra en 1876 y murió allí en 1937. Trabajó muchos años en Arthur Guinness, Sons and Company. En realidad, en sus últimos años estuvo a cargo de Guiness Brewery en Londres. Guinness prefería que sus empleados utilizaran seudónimos cuando publicaban trabajos, de modo que, en 1908, cuando Gosset escribió "The Probable Error of a Mean", utilizó el nombre de *Student.* En este artículo describió por primera vez las propiedades de la distribución *t.*

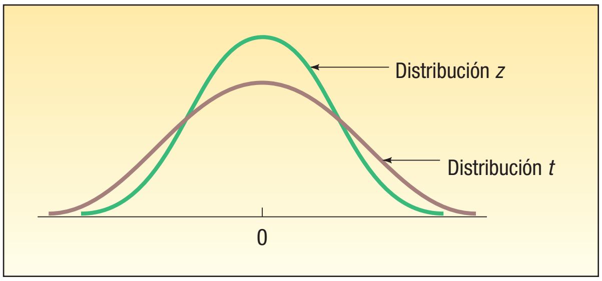
Por fortuna, se utiliza la desviación estándar de la muestra para estimar la desviación estándar poblacional. Es decir, se utiliza *s,* la desviación estándar de la muestra, para estimar σ, la desviación estándar de la población. No obstante, al hacerlo no es posible utilizar la fórmula (9-1). Como no conoce σ, no puede utilizar la distribución *z.* Sin embargo, hay una solución: utilizar la desviación estándar de la media y sustituir la distribución *z* con la distribución *t.*

La distribución *t* es una distribución de probabilidad continua, con muchas características similares a las de la distribución *z.* William Gosset, experto cervecero, fue el primero en estudiar la distribución *t.*

Estaba especialmente interesado en el comportamiento exacto de la distribución del siguiente estadístico:

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}}$$

Aquí, *s* es un estimador de σ. Le preocupaba en particular la discrepancia entre *s* y σ cuando *s* se calculaba a partir de una muestra muy pequeña. La distribución *t* y la distribución normal estándar se muestran en la gráfica 9-1. Observe que la distribución *t* es más plana y que se extiende más que la distribución normal estándar. Esto se debe a que la desviación estándar de la distribución *t* es mayor que la distribución normal estándar.

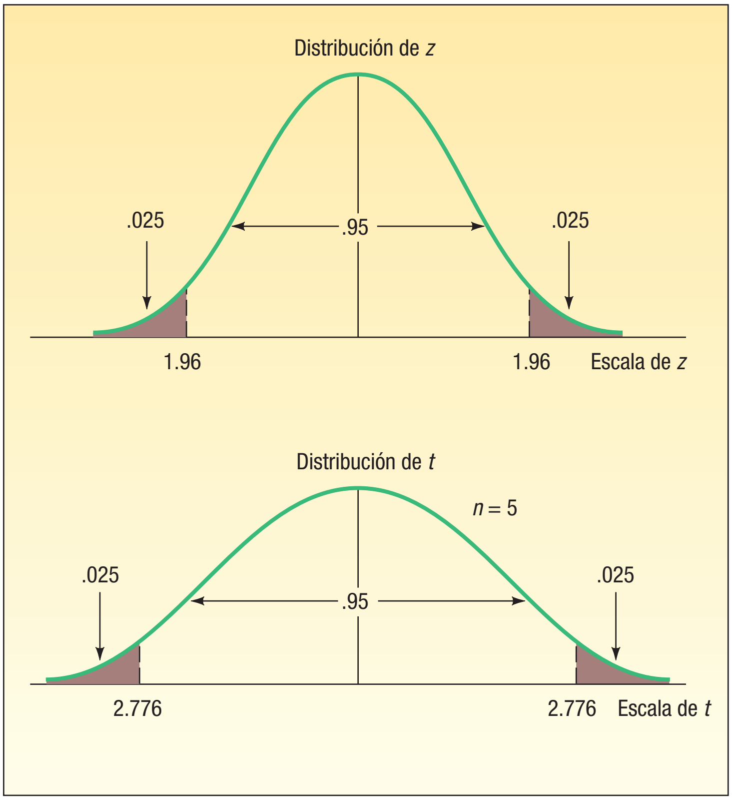


GRÁFICA 9-1 Distribución normal estándar y distribución *t* de Student

Las siguientes características de la distribución *t* se basan en el supuesto de que la población de interés es de naturaleza normal, o casi normal.

- Como en el caso de la distribución *z,* es una distribución continua.
- Como en el caso de la distribución *z,* tiene forma de campana y es simétrica.
- No existe una distribución *t,* sino una familia de distribuciones *t.* Todas las distribuciones *t* tienen una media de 0, y sus desviaciones estándares difieren de acuerdo con el tamaño de la muestra, *n.* Existe una distribución *t* para un tamaño de muestra de 20, otro para un tamaño de muestra de 22, etc. La desviación estándar de una distribución *t* con 5 observaciones es mayor que en el caso de una distribución *t* con 20 observaciones.
- La distribución *t* se extiende más y es más plana por el centro que la distribución normal estándar (vea la gráfica 9-1). Sin embargo, conforme se incrementa el tamaño de la muestra, la distribución *t* se aproxima a la distribución normal estándar, pues los errores que se cometen al utilizar *s* para estimar σ disminuyen con muestras más grandes.

Como la distribución *t* de Student posee mayor dispersión que la distribución *z,* el valor de *t* en un nivel de confianza dado tiene una magnitud mayor que el valor *z* correspondiente. La gráfica 9-2 muestra los valores de *z* para un nivel de confianza de 95% y de *t* para el mismo nivel de confianza cuando el tamaño de la muestra es de $n = 5$. En breve se explicará la forma como se obtuvo el valor real de *t.* Por el momento, observe que, con el mismo nivel de confianza, la distribución *t* es más plana o más amplia que la distribución normal estándar.



GRÁFICA 9-2 Valores de z y t para el nivel de confianza de 95 por ciento

Para crear un intervalo de confianza de la media poblacional con la distribución *t,* se ajusta la fórmula (9-1) de la siguiente manera.

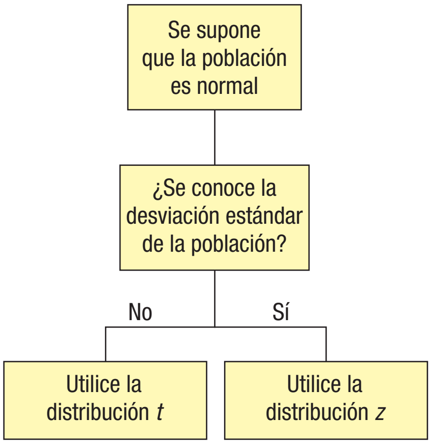
INTERVALO DE CONFIANZA DE LA MEDIA POBLACIONAL CON σ DESCONOCIDA	$\bar{X} \pm t \frac{s}{\sqrt{n}}$	**(9-2)**

Para crear un intervalo de confianza de la media poblacional con una desviación estándar desconocida:

1. Suponga que la población muestreada es normal o aproximadamente normal. De acuerdo con el teorema central del límite, sabemos que este supuesto es cuestionable en el caso de muestras pequeñas, y es más válida en el de muestras más grandes.
2. Estime la desviación estándar de la población (σ) con la desviación estándar de la muestra (s).
3. Utilice la distribución t en lugar de la distribución z.

Cabe hacer una aclaración en este momento. La decisión de utilizar t o z se basa en el hecho de que se conozca σ, la desviación estándar poblacional. Si se conoce, se utiliza z. Si no se conoce, se debe utilizar t. La gráfica 9-3 resume el proceso de toma de decisión.

El siguiente ejemplo ilustra un intervalo de confianza de una media poblacional cuando no se conoce la desviación estándar de la población y para determinar el valor apropiado de t en una tabla.



GRÁFICA 9-3 Cómo determinar cuándo se debe usar la distribución z o la distribución t

Ejemplo

Un fabricante de llantas desea investigar la durabilidad de sus productos. Una muestra de 10 llantas que recorrieron 50 000 millas reveló una media muestral de 0.32 pulgadas de cuerda restante con una desviación estándar de 0.09 pulgadas. Construya un intervalo de confianza de 95% de la media poblacional. ¿Sería razonable que el fabricante concluyera que después de 50 000 millas la cantidad media poblacional de cuerda restante es de 0.30 pulgadas?

Solución

Para comenzar, se supone que la distribución de la población es normal. En este caso no hay muchas evidencias, pero tal vez la suposición sea razonable. No se conoce la desviación estándar de la población, pero sí la desviación estándar de la muestra, que es de 0.09 pulgadas. Se aplica la fórmula (9-2):

$$\overline{X} \pm t \frac{s}{\sqrt{n}}$$

De acuerdo con la información dada, $\overline{X} = 0.32$, $s = 0.09$ y $n = 10$. Para hallar el valor de t, utilice el apéndice B.2, una parte del cual se reproduce en la tabla 9-1. El primer paso para localizar t consiste es desplazarse a lo largo de las columnas identificadas como "Intervalos de

TABLA 9-1 Una parte de la distribución t

	Intervalos de confianza				
	80%	**90%**	**95%**	**98%**	**99%**
	Nivel de significancia de una prueba de una cola				
gl	**0.10**	**0.05**	**0.025**	**0.010**	**0.005**
	Nivel de significancia de una prueba de dos colas				
	0.20	**0.10**	**0.05**	**0.02**	**0.01**
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169

confianza" hasta el nivel de confianza que se requiere. En este caso, desea el nivel de confianza de 95%, así que vaya a la columna con el encabezamiento "95%". La columna del margen izquierdo se identifica como "gl". Estas palabras se refieren al número de grados de libertad, esto es, el número de observaciones incluidas en la muestra menos el número de muestras, el cual se escribe $n - 1$. En este caso es de $10 - 1 = 9$. ¿Por qué se decidió que había 9 grados de libertad? Cuando se utilizan estadísticas de la muestra, es necesario determinar el número de valores que se encuentran *libres para variar.*

Para ilustrarlo, suponga que la media de cuatro números es de 5. Los cuatro números son 7, 4, 1 y 8. Las desviaciones respecto de la media de estos números deben sumar 0. Las desviaciones de $+2$, -1, -4 y $+3$ suman 0. Si se conocen las desviaciones de $+2$, -1 y -4, el valor de $+3$ se fija (se restringe) con el fin de satisfacer la condición de que la suma de las desviaciones debe totalizar 0. Por consiguiente, 1 grado de libertad se pierde en un problema de muestreo que implique la desviación estándar de la muestra, pues se conoce un número (la media aritmética). En el caso de un nivel de confianza de 95% y 9 grados de libertad, seleccione la fila con 9 grados de libertad. El valor de *t* es 2.262.

Para determinar el intervalo de confianza se sustituyen los valores en la fórmula (9-2):

$$\bar{X} \pm t \frac{s}{\sqrt{n}} = 0.32 \pm 2.262 \frac{0.09}{\sqrt{10}} = 0.32 \pm .064$$

Los puntos extremos del intervalo de confianza son 0.256 y 0.384. ¿Cómo interpretar este resultado? Si repitiéramos este estudio 200 veces, calculando el intervalo de confianza de 95% con cada media de la muestra y la desviación estándar, 190 intervalos incluirían la media poblacional. Diez intervalos no la incluirían. Éste es el efecto del error muestral. Otra interpretación es concluir que la media poblacional se encuentra en este intervalo. El fabricante puede estar seguro (95% seguro) de que la profundidad media de las cuerdas oscila entre 0.256 y 0.384 pulgadas. Como el valor de 0.30 se encuentra en este intervalo, es posible que la media de la población sea de 0.30 pulgadas.

He aquí otro ejemplo para explicar el uso de los intervalos de confianza. Suponga que un artículo publicado en el periódico local indica que el tiempo medio para vender una residencia de la zona es de 60 días. Usted selecciona una muestra aleatoria de 20 residencias que se vendieron en el último año y encuentra que el tiempo medio de venta es de 65 días. De acuerdo con los datos de la muestra, crea un intervalo de confianza de 95% de la media de la población. Usted descubre que los puntos extremos son 62 y 68 días. ¿Cómo interpreta este resultado? Puede confiar de manera razonable en que la media poblacional se encuentre dentro de este intervalo. El valor propuesto para la media poblacional, es decir, 60 días, no se incluye en el intervalo. No es probable que la media poblacional sea de 60 días. La evidencia indica que la afirmación del periódico local puede no ser correcta. En otras palabras, parece poco razonable obtener la muestra que usted tomó de una población que tenía un tiempo de venta medio de 60 días.

El siguiente ejemplo mostrará detalles adicionales para determinar e interpretar el intervalo de confianza. Se usó Minitab para realizar los cálculos.

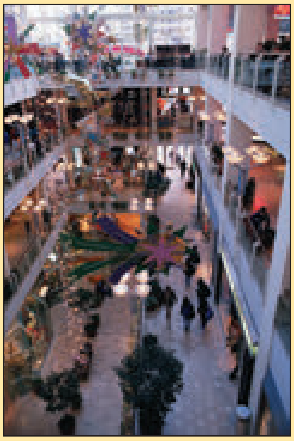
Ejemplo

El gerente de Inlet Square Mall, cerca de Ft. Myers, Florida, desea estimar la cantidad media que gastan los clientes que visitan el centro comercial. Una muestra de 20 clientes revela las siguientes cantidades.

$48.16	$42.22	$46.82	$51.45	$23.78	$41.86	$54.86
37.92	52.64	48.59	50.82	46.94	61.83	61.69
49.17	61.46	51.35	52.68	58.84	43.88	

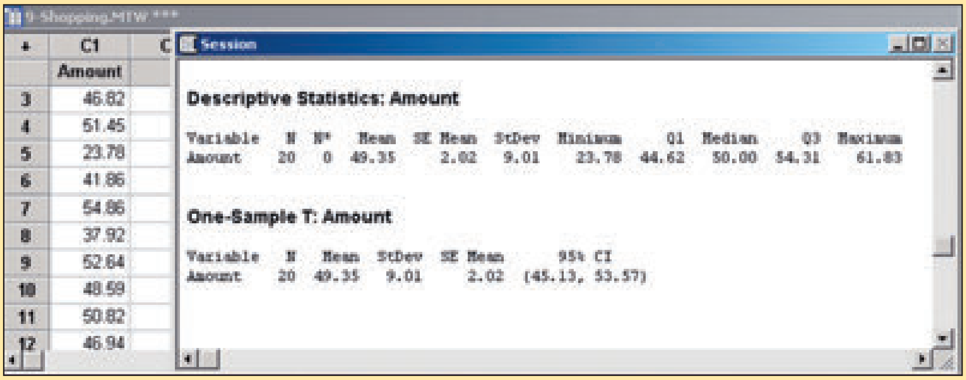
¿Cuál es la mejor estimación de la media poblacional? Determine un intervalo de confianza de 95%. Interprete el resultado. ¿Concluiría de forma razonable que la media poblacional es de $50? ¿Y de $60?

Solución



El gerente del centro comercial supone que la población de las cantidades gastadas sigue la distribución normal. En este caso es una suposición razonable. Además, la técnica del intervalo de confianza resulta muy poderosa y tiende a consignar cualquier error del lado conservador si la población no es normal. No cabe suponer una condición normal cuando la población se encuentra pronunciadamente sesgada o cuando la distribución tiene colas *gruesas.* En el capítulo 18 se exponen métodos para manejar este problema en caso de que no sea posible suponer una condición normal. En este caso, resulta razonable suponer una condición normal.

No se conoce la desviación estándar de la población. De ahí que resulte adecuado utilizar la distribución *t* y la fórmula (9-2) para encontrar el intervalo de confianza. Se utiliza el software Minitab para hallar la media y la desviación estándar de esta muestra. Los resultados aparecen a continuación.



El gerente del centro comercial no conoce la media poblacional. La media muestral constituye la mejor aproximación de dicho valor. De acuerdo con la captura de pantalla de Minitab, la media es de $49.35, que constituye la mejor aproximación, la *estimación puntual,* de la media poblacional desconocida.

Se aplica la fórmula (9-2) para determinar el intervalo de confianza. El valor de *t* se localiza en el apéndice B.2. Hay $n - 1 = 20 - 1 = 19$ grados de libertad. Al desplazarse por el renglón con 19 grados de libertad a la columna del intervalo de confianza de 95%, el valor de esta intersección es de 2.093. Se sustituyen estos valores en la fórmula 9-2 para encontrar el intervalo de confianza.

$$\bar{X} \pm t\frac{s}{\sqrt{n}} = \$49.35 \pm 2.093\frac{\$9.01}{\sqrt{20}} = \$49.35 \pm \$4.22$$

Los puntos extremos del intervalo de confianza son $45.13 y $53.57. Resulta razonable concluir que la media poblacional se encuentra en dicho intervalo.

El gerente de Inlet Square se preguntaba si la media poblacional podría haber sido $50 o $60. El valor de $50 se encuentra dentro del intervalo de confianza. Resulta razonable que la media poblacional sea de $50. El valor de $60 no se encuentra en el intervalo de confianza. De ahí que se concluya que no es probable que la media poblacional sea de $60.

Los cálculos para construir un intervalo de confianza también se encuentran disponibles en Excel. La captura de pantalla aparece a continuación. Observe que la media de la muestra ($49.35) y la desviación estándar de la muestra ($9.01) son las mismas que en los cálculos de Minitab. En la información de Excel, el último renglón de la salida también incluye el margen de error, que es la cantidad que se suma y se resta de la media muestral para formar los puntos extremos del intervalo de confianza. Este valor se determina a partir de la expresión

$$t\frac{s}{\sqrt{n}} = 2.093\frac{\$9.01}{\sqrt{20}} = \$4.22$$

Shopping [Compatibility Mode]

	A	B	C	D	E
1	Amount		Amount		
2	48.16				
3	42.22		Mean	49.35	
4	46.82		Standard Error	2.02	
5	51.45		Median	50.00	
6	23.78		Mode	#N/A	
7	41.86		Standard Deviation	9.01	
8	54.86		Sample Variance	81.22	
9	37.92		Kurtosis	2.26	
10	52.64		Skewness	-1.00	
11	48.59		Range	38.05	
12	50.82		Minimum	23.78	
13	46.94		Maximum	61.83	
14	61.83		Sum	986.96	
15	61.69		Count	20.00	
16	49.17		Confidence Level(95.0%)	4.22	
17	61.46				
18	51.35				
19	52.68				
20	58.84				
21	43.88				
22					

Autoevaluación 9-2

Dottie Kleman es la “Cookie Lady”. Hornea y vende galletas en 50 lugares del área de Filadelfia. La señora Kleman está interesada en el ausentismo de sus trabajadoras. La siguiente información se refiere al número de días de ausencias de una muestra de 10 trabajadoras durante el último periodo de pago de dos semanas.

4	1	2	2	1	2	2	1	0	3

a) Determine la media y la desviación estándar de la muestra.
b) ¿Cuál es la media de la población? ¿Cuál es la mejor estimación de dicho valor?
c) Construya un intervalo de confianza de 95% de la media poblacional.
d) Explique la razón por la que se utiliza la distribución *t* como parte del intervalo de confianza.
e) ¿Es razonable concluir que la trabajadora común no falta ningún día durante un periodo de pago?

Ejercicios

connect

9. Utilice el apéndice B.2 para localizar el valor *t* en las siguientes condiciones.
   a) El tamaño de la muestra es de 12, y el nivel de confianza, de 95 por ciento.
   b) El tamaño de la muestra es de 20, y el nivel de confianza, de 90 por ciento.
   c) El tamaño de la muestra es de 8, y el nivel de confianza, de 99 por ciento.

10. Utilice el apéndice B.2 para localizar el valor de *t* en las siguientes condiciones.
   a) El tamaño de la muestra es de 15, y el nivel de confianza, de 95 por ciento.
   b) El tamaño de la muestra es de 24, y el nivel de confianza, de 98 por ciento.
   c) El tamaño de la muestra es de 12, y el nivel de confianza, de 90 por ciento.

11. El propietario de Britten’s Egg Farm desea calcular la cantidad media de huevos que pone cada gallina. Una muestra de 20 gallinas indica que ponen un promedio de 20 huevos al mes, con una desviación estándar de 2 huevos al mes.
   a) ¿Cuál es el valor de la media de la población? ¿Cuál es el mejor estimador de este valor?
   b) Explique por qué necesita utilizar la distribución *t.* ¿Qué suposiciones necesita hacer?
   c) ¿Cuál es el valor de *t* en un intervalo de confianza de 95%?
   d) Construya el intervalo de confianza de 95% de la media de población.
   e) ¿Es razonable concluir que la media poblacional es de 21 huevos? ¿Y de 25 huevos?

12. La industria estadounidense de lácteos desea calcular el consumo medio de leche por año. Una muestra de 16 personas revela que el consumo medio anual es de 60 galones, con una desviación estándar de 20 galones.
   a) ¿Cuál es el valor de la media poblacional? ¿Cuál es el mejor estimador de este valor?
   b) Explique por qué necesita utilizar la distribución *t.* ¿Qué suposiciones necesita hacer?
   c) ¿Cuál es el valor de *t* en un intervalo de confianza de 90%?
   d) Construya el intervalo de confianza de 90% de la media de población.
   e) ¿Es razonable concluir que la media poblacional es de 63 galones?

13. Merrill Lynch Securities y Health Care Retirement, Inc., son dos grandes empresas ubicadas en el centro de Toledo, Ohio. Contemplan ofrecer de forma conjunta servicio de guardería para sus empleados. Como parte del estudio de viabilidad del proyecto, desean calcular el costo medio semanal por el cuidado de los niños. Una muestra de 10 empleados que recurren al servicio de guardería revela las siguientes cantidades gastadas la semana pasada.

$107	$92	$97	$95	$105	$101	$91	$99	$95	$104

Construya el intervalo de confianza de 90% de la media poblacional. Interprete el resultado.

14. Greater Pittsburgh Area Chamber of Commerce desea calcular el tiempo medio que los trabajadores que laboran en el centro de la ciudad utilizan para llegar al trabajo. Una muestra de 15 trabajadores revela las siguientes cantidades de minutos de viaje.

29	38	38	33	38	21	45	34
40	37	37	42	30	29	35	

Construya el intervalo de confianza de 98% de la media poblacional. Interprete el resultado.

9.4 Intervalo de confianza de una proporción

OA5 Construir el intervalo de confianza de una proporción de la población.

El material hasta ahora expuesto en este capítulo utiliza la escala de medición de razón. Es decir, se emplean variables como ingresos, pesos, distancias y edades. Ahora se considerarán casos como los siguientes:

- El director de servicios profesionales de Southern Technical Institute informa que 80% de sus graduados entra en el mercado laboral en un puesto relacionado con su área de estudio.
- Un representante de ventas afirma que 45% de las ventas de Burger King se lleva a cabo en la ventana de servicio para automóviles.
- Un estudio de las casas del área de Chicago indicó que 85% de las construcciones nuevas cuenta con sistema de aire acondicionado central.
- Una encuesta reciente entre hombres casados de entre 35 y 50 años de edad descubrió que 63% creía que ambos cónyuges deben aportar dinero.

Estos ejemplos ilustran la escala de medición nominal. Cuando se mide con una escala nominal, una observación se clasifica en uno de dos o más grupos mutuamente excluyentes. Por ejemplo, un graduado de Southern Tech entra al mercado laboral en un puesto relacionado con su campo de estudio o no lo hace. Un consumidor de Burger King hace una compra en la ventana de servicio para automóviles o no. Sólo hay dos posibilidades, y el resultado debe clasificarse en uno de los dos grupos.

PROPORCIÓN Fracción, razón o porcentaje que indica la parte de la muestra de la población que posee un rasgo de interés particular.

Como ejemplo de proporción, una encuesta reciente indicó que 92 de cada 100 entrevistados estaban de acuerdo con el horario de verano para ahorrar energía. La proporción de la muestra es de 92/100, o 0.92, o 92%. Si p representa la proporción de la muestra, X el número de *éxitos* y n el número de elementos de la muestra, se determina una proporción muestral de la siguiente manera:

PROPORCIÓN MUESTRAL $$p = \frac{X}{n} \tag{9-3}$$

La proporción de la población se define por medio de π. Por consiguiente, π se refiere al porcentaje de éxitos en la población. Recuerde, del capítulo 6, que π es la proporción de *éxitos* en una distribución binomial. Esto permite continuar la práctica de utilizar letras griegas para identificar parámetros de población y letras latinas para identificar estadísticas muestrales.

Para crear el intervalo de confianza de una proporción, es necesario cumplir con los siguientes supuestos:

1. Las condiciones binomiales, estudiadas en el capítulo 6, han quedado satisfechas. En resumen, estas condiciones son:
   a) Los datos de la muestra son resultado de conteos.
   b) Sólo hay dos posibles resultados (lo normal es referirse a uno de los resultados como *éxito* y al otro como *fracaso*).
   c) La probabilidad de un éxito permanece igual de una prueba a la siguiente.
   d) Las pruebas son independientes. Esto significa que el resultado de la prueba no influye en el resultado de otra.
2. Los valores $n\pi$ y $n(1 - \pi)$ deben ser mayores o iguales que 5. Esta condición permite recurrir al teorema central del límite y emplear la distribución normal estándar, es decir, z, para completar un intervalo de confianza.

El desarrollo del estimador puntual de la proporción de la población y el intervalo de confianza de una proporción de población es similar a hacerlo para una media. Para ilustrarlo, considere lo siguiente: John Gail es candidato para representar al tercer distrito de Nebraska ante el Congreso. De una muestra aleatoria de 100 electores en el distrito, 60 indican que planean votar por él en las próximas elecciones. La proporción de la muestra es de 0.60, pero no se conoce la proporción poblacional. Es decir, no se conoce qué proporción de electores de la *población* votará por Gail. El valor de la muestra, 0.60, es el mejor estimador del parámetro poblacional desconocido. Así, p, que es de 0.60, constituye un estimador de π, que no se conoce.

Para crear el intervalo de confianza de una proporción de población se aplica la fórmula:

INTERVALO DE CONFIANZA DE LA PROPORCIÓN DE UNA POBLACIÓN $$p \pm z\sqrt{\frac{p(1-p)}{n}} \tag{9-4}$$

Un ejemplo ayudará a explicar los detalles para determinar un intervalo de confianza y el resultado.

Estadística en acción

Los resultados de muchas encuestas que aparecen en periódicos, revistas de noticias y televisión utilizan intervalos de confianza. Por ejemplo, una encuesta reciente de 800 televidentes de Toledo, Ohio, reveló que 44% observaba las noticias de la noche en la estación local afiliada a CBS. El artículo también indicó que el margen de error fue de 3.4%. El margen de error es, en realidad, la cantidad que se suma y resta del estimador puntual para determinar los puntos extremos de un intervalo de confianza. De acuerdo con la fórmula (9-4) y el nivel de confianza de 95 por ciento:

$$z\sqrt{\frac{p(1-p)}{n}}$$

$$= 1.96\sqrt{\frac{.44(1-.44)}{800}}$$

$$= 0.034$$

Ejemplo

El sindicato que representa a Bottle Blowers of America (BBA) considera la propuesta de fusión con Teamsters Union. De acuerdo con el reglamento del sindicato de BBA, por lo menos tres cuartas partes de los miembros del sindicato deben aprobar cualquier fusión. Una muestra aleatoria de 2 000 miembros actuales de BBA revela que 1 600 planean votar por la propuesta. ¿Qué es el estimador de la proporción poblacional? Determine el intervalo de confianza de 95% de la proporción poblacional. Fundamente su decisión en esta información de la muestra: ¿puede concluir que la proporción necesaria de miembros del BBA favorece la fusión? ¿Por qué?

Solución

Primero calcule la proporción de la muestra de acuerdo con la fórmula (9-3). Ésta es de 0.80, que se calcula de la siguiente manera:

$$p = \frac{X}{n} = \frac{1\,600}{2\,000} = .80$$

Por consiguiente, se calcula que 80% de la población favorece la propuesta de fusión. Determine el intervalo de confianza de 95% con ayuda de la fórmula (9-4). El valor z correspondiente al nivel de confianza de 95% es de 1.96.

$$p \pm z\sqrt{\frac{p(1-p)}{n}} = .80 \pm 1.96\sqrt{\frac{.80(1-.80)}{2\,000}} = .80 \pm .018$$

Los puntos extremos del intervalo de confianza son 0.782 y 0.818. El punto extremo más bajo es mayor que 0.75. Así, es probable que se apruebe la propuesta de fusión, pues el estimador del intervalo incluye valores superiores a 75% de los miembros del sindicato.

Un repaso de la interpretación del intervalo de confianza: si la encuesta fue aplicada 100 veces con 100 muestras distintas, los intervalos de confianza construidos a partir de 95 de las muestras contendrán la verdadera proporción de la población. Además, la interpretación de un intervalo de confianza resulta de mucha utilidad en la toma de decisiones, y desempeña un papel muy importante en especial la noche de las elecciones. Por ejemplo, Cliff Obermeyer se postula para representar ante el Congreso al 6o. distrito de Nueva Jersey. Suponga que se entrevista a los electores que acaban de votar y 275 indican que votaron por Obermeyer. Considere que 500 electores es una muestra aleatoria de quienes votan en el 6o. distrito. Esto significa que 55% de los electores de la muestra votó por Obermeyer. De acuerdo con la fórmula (9-3):

$$p = \frac{X}{n} = \frac{275}{500} = .55$$

Ahora, para estar seguros de la elección, Obermeyer debe ganar *más de* 50% de los votos de la población de electores. En este momento se conoce un estimador puntual, que es de 0.55, de la población de electores que votarán por él. Ahora bien, no se conoce el porcentaje de la población que votará por el candidato. En estas circunstancias, la pregunta es: ¿es posible tomar una muestra de 500 electores de una población en la que 50% o menos de los electores apoye a Obermeyer para encontrar que 55% de la muestra lo apoya? En otras palabras, ¿el error de muestreo, que es $p - \pi = .55 - .50 = .05$, se debe al azar, o la población de electores que apoya a Obermeyer es superior a 0.50? Si se establece el intervalo de confianza de la proporción de la muestra y halla que 0.50 no se encuentra en el intervalo, concluirá que la proporción de electores que apoya a Obermeyer es mayor que 0.50. ¿Qué significa esto? Bien, significa que puede resultar electo. ¿Qué pasa si 0.50 pertenece al intervalo? Entonces concluirá que es posible que 50% o menos de los electores apoyen su candidatura y no es posible concluir que será electo a partir de de la información de la muestra. En este caso, si se utiliza el nivel de significancia de 95% y la fórmula (9-4), se tiene que:

$$p \pm z\sqrt{\frac{p(1-p)}{n}} = .55 \pm 1.96\sqrt{\frac{.55(1-.55)}{500}} = .55 \pm .044$$

Así, los puntos extremos del intervalo de confianza son: 0.55, $-0.044 = 0.506$ y $0.55 + 0.044 = 0.594$. El valor de 0.50 no pertenece al intervalo. Por lo tanto, se concluye que probablemente *más de 50%* de los electores apoya a Obermeyer, lo cual es suficiente para que sea elegido.

¿Siempre se utiliza este procedimiento? Sí. Es exactamente el procedimiento de las cadenas de televisión, revistas de noticias y sondeos en la noche de las elecciones.

Autoevaluación 9-3

Se llevó a cabo una encuesta de mercado para calcular la proporción de amas de casa que reconocerían el nombre de la marca de un limpiador a partir de la forma y color del envase. De las 1 400 amas de casa de la muestra, 420 identificaron la marca por su nombre.

a) Calcule el valor de la proporción de la población.
b) Construya el intervalo de confianza de 99% de la proporción poblacional.
c) Interprete sus conclusiones.

Ejercicios

15. El propietario de West End Kwick Fill Gas Station desea determinar la proporción de clientes que utilizan tarjeta de crédito o débito para pagar la gasolina en el área de las bombas. Entrevistó a 100 clientes y descubre que 80 pagaron en ella.
 ***a*)** Calcule el valor de la proporción de la población.
 ***b*)** Construya el intervalo de confianza de 95% de la proporción poblacional.
 ***c*)** Interprete sus conclusiones.

16. Maria Wilson considera postularse para la alcaldía de la ciudad de Bear Gulch, Montana. Antes de solicitar la postulación, decide realizar una encuesta entre los electores de Bear Gulch. Una muestra de 400 electores revela que 300 la apoyarían en las elecciones de noviembre.
 ***a*)** Calcule el valor de la proporción de la población. Calcule el error estándar de la proporción.
 ***b*)** Construya el intervalo de confianza de 99% de la proporción poblacional.
 ***c*)** Interprete sus resultados.

17. La televisora Fox TV considera reemplazar uno de sus programas de investigación ciminal, que se transmite durante las horas de mayor audiencia, por una nueva comedia orientada a la familia. Antes de tomar una decisión definitiva, los ejecutivos estudian una muestra de 400 telespectadores. Después de ver la comedia, 250 afirmaron que la verían y sugirieron reemplazar el programa de investigación criminal.
 ***a*)** Calcule el valor de la proporción de la población.
 ***b*)** Construya el intervalo de confianza de 99% de la proporción poblacional.
 ***c*)** Interprete los resultados que obtuvo.

18. Schadek Silkscreen Printing, Inc., compra tazas de plástico para imprimir en ellas logotipos de eventos deportivos, graduaciones, cumpleaños u otras ocasiones importantes. Zack Schadek, el propietario, recibió un envío grande esta mañana. Para asegurarse de la calidad del envío, seleccionó una muestra aleatoria de 300 tazas. Halló que 15 estaban defectuosas.
 ***a*)** ¿Cuál es la proporción aproximada de tazas defectuosas en la población?
 ***b*)** Construya el intervalo de confianza de 95% de la proporción de tazas defectuosas.
 ***c*)** Zack llegó con su proveedor al acuerdo de que devolverá lotes con 10% o más de artículos defectuosos. ¿Debe devolver este lote? Explique su decisión.

9.5 Elección del tamaño adecuado de una muestra

Una variable importante cuando se trabaja con intervalos de confianza es el tamaño de la muestra. Sin embargo, en la práctica, no es una variable, sino una decisión que se toma para que la estimación del parámetro de población sea bueno. Esta decisión se basa en tres variables:

OA6 Calcular el tamaño de la muestra necesario para estimar una proporción de la población o una media poblacional.

1. El margen de error que tolerará el investigador.
2. El nivel de confianza deseado.
3. La variabilidad o dispersión de la población que se estudia.

La primera variable es el *margen de error.* El máximo error admisible, designado *E,* es la magnitud que se suma y resta de la media muestral (o proporción muestral) para determinar los puntos extremos del intervalo de confianza. Por ejemplo, en un estudio de salarios, podemos decidir que deseamos estimar el salario promedio de la población con un margen de error de más o menos $1 000. O en una encuesta de opinión, podemos decidir que deseamos calcular la proporción de la población con un margen de error de más o menos 5%. El margen de error es la magnitud del error que se tolerará al estimar un parámetro poblacional. Quizás se pregunte por qué no elegir márgenes pequeños de error. Existe una compensación entre el margen de error y el tamaño de la muestra. Un margen de error pequeño requiere de una muestra más grande y de más tiempo y dinero para recolectarla. Un margen de error más grande permitirá tener una muestra más pequeña y un intervalo de confianza más amplio.

La segunda elección es el *nivel de confianza.* Al trabajar con un intervalo de confianza, lógicamente se elegirán niveles de confianza relativamente altos como de 95 y 99%, que son los más comunes. Para calcular el tamaño de la muestra, se necesitará un estadístico z que corresponda al nivel de confianza elegido. El nivel de confianza de 95% corresponde al valor z de 1.96, y el nivel de confianza de 99%, a un valor z de 2.58. Note que las muestras más grandes (con su consecuente requerimiento de más tiempo y dinero para recolectarlas) corresponden a niveles de confianza más altos. Asimismo, observe que utilizamos un estadístico z.

El tercer factor en la determinación del tamaño de una muestra es la *desviación estándar de la población.* Si la población se encuentra muy dispersa, se requiere una muestra grande. Por el contrario, si se encuentra concentrada (homogénea), el tamaño de muestra que se requiere será menor. No obstante, puede ser necesario utilizar un estimador de la desviación estándar de la población. He aquí algunas sugerencias para determinar dicho estimador.

1. **Realice un estudio piloto.** Éste es el método más común. Suponga que desea un cálculo aproximado de la cantidad de horas que trabajan a la semana los estudiantes matriculados en la Facultad de Administración de la University of Texas. Para probar la validez del cuestionario, se aplica a una pequeña muestra de estudiantes. A partir de esta pequeña muestra se calcula la desviación estándar de la cantidad de horas que trabajan y se utiliza este valor como la desviación estándar de la población.
2. **Utilice un estudio comparativo.** Aplique este enfoque cuando se encuentre disponible un estimador de la dispersión de otro estudio. Suponga que quiere calcular la cantidad de horas semanales que trabajan los recolectores de basura. La información de ciertas dependencias estatales o federales que normalmente estudian la fuerza de trabajo puede ser útil para obtener un cálculo aproximado de la desviación estándar.
3. **Emplee un enfoque basado en el intervalo.** Para aplicar este enfoque necesita conocer o contar con un cálculo de los valores máximo y mínimo de la población. Recuerde, del capítulo 3, en el que se explicó la regla empírica, que se podía esperar que casi todas las observaciones se encontraran a más o menos 3 desviaciones estándares de la media, si la distribución seguía la distribución normal. Por consiguiente, la distancia entre los valores máximo y mínimo es de 6 desviaciones estándares. Puede calcular la desviación estándar como un sexto del rango. Por ejemplo, la directora de operaciones del University Bank desea un cálculo aproximado del número de cheques que expiden cada mes los estudiantes universitarios. Ella cree que la distribución del número de cheques es normal. La cantidad mínima de cheques expedidos cada mes es de 2, y la máxima, de 50. El rango de la cantidad de cheques que se expiden por mes es de 48, que se determina al restar $50 - 2$. El estimador de la desviación estándar es entonces de 8 cheques mensuales: 48/6.

Tamaño de la muestra para calcular una media poblacional

Para calcular una media poblacional, se puede expresar la interacción entre estos tres factores y el tamaño de la muestra se expresa con la fórmula siguiente. Note que esta fórmula es el margen de error que se utiliza para calcular los puntos extremos de los intervalos de confianza para estimar una media poblacional.

$$E = z\frac{\sigma}{\sqrt{n}}$$

Al despejar *n* en esta ecuación se obtiene el siguiente resultado:

TAMAÑO DE LA MUESTRA PARA ESTIMAR LA MEDIA DE LA POBLACIÓN

$$n = \left(\frac{z\sigma}{E}\right)^2 \qquad \textbf{(9-5)}$$

donde:

n es el tamaño de la muestra.
z es el valor normal estándar correspondiente al nivel de confianza deseado.
σ es la desviación estándar de la población.
E es el error máximo admisible.

El resultado de este cálculo no siempre es un número entero. Cuando el resultado no es un entero, se acostumbra redondear *cualquier* resultado fraccionario. Por ejemplo, 201.21 se redondearía a 202.

Ejemplo

Un estudiante de administración pública desea determinar la cantidad media que ganan al mes los miembros de los consejos ciudadanos de las grandes ciudades. El error al calcular la media debe ser inferior a $100, con un nivel de confianza de 95%. El estudiante encontró un informe del Departamento del Trabajo en el que la desviación estándar es de $1 000. ¿Cuál es el tamaño de la muestra que se requiere?

Solución

El error máximo admisible, *E,* es de $100. El valor *z* de un nivel de confianza de 95% es de 1.96, y el estimador de la desviación estándar, $1 000. Al sustituir estos valores en la fórmula (9-5) se obtiene el tamaño de la muestra que se requiere:

$$n = \left(\frac{z\sigma}{E}\right)^2 = \left(\frac{(1.96)(\$1\,000)}{\$100}\right)^2 = (19.6)^2 = 384.16$$

El valor calculado de 384.16 se redondea a 385. Se requiere una muestra de 385 para satisfacer las especificaciones. Si el estudiante desea incrementar el nivel de confianza, por ejemplo, a 99%, se requerirá una muestra más grande. El valor *z* correspondiente al nivel de confianza de 99% es 2.58.

$$n = \left(\frac{z\sigma}{E}\right)^2 = \left(\frac{(2.58)(\$1\,000)}{\$100}\right)^2 = (25.8)^2 = 665.64$$

Se recomienda una muestra de 666. Observe cuánto modificó el tamaño de la muestra el cambio en el nivel de confianza. Un incremento del nivel de confianza de 95% al de 99% dio como resultado un incremento de 281 observaciones o 73% [(666/385)*100]. Esto puede incrementar mucho el costo del estudio, en términos de tiempo y dinero. De ahí que deba considerarse con cuidado el nivel de confianza.

Tamaño de la muestra para calcular la proporción de una población

Para determinar el tamaño de la muestra en el caso de una proporción, es necesario especificar estas mismas tres variables:

1. El margen de error.
2. El nivel de confianza deseado.
3. La variación o dispersión de la población a estudiar.

En el caso de la distribución binomial, el margen de error es:

$$E = z\sqrt{\frac{\pi(1 - \pi)}{n}}$$

Si se resuelve la ecuación para despejar n se obtiene lo siguiente:

TAMAÑO DE LA MUESTRA DE LA PROPORCIÓN DE LA POBLACIÓN

$$n = \pi(1 - \pi)\left(\frac{z}{E}\right)^2 \qquad \textbf{(9-6)}$$

donde:

n es el tamaño de la muestra.
z es el valor normal estándar correspondiente al nivel de confianza deseado.
π es la proporción de la población.
E es el máximo error tolerable.

Las elecciones del estadístico z y el margen de error E son las mismas que para calcular la media poblacional. Sin embargo, en este caso la desviación estándar de la población de una distribución normal está representada por $\pi(1 - \pi)$. Para encontrar el valor de una proporción de la población, podemos hallar un estudio similar o conducir un estudio piloto. Si no se puede encontrar un valor confiable, entonces se debe usar un valor de π de 0.50. Observe que $\pi(1 - \pi)$ tiene el mayor valor utilizando 0.50 y, por lo tanto, sin una buena estimación de la proporción de la población, se sobrestima el tamaño de la muestra. Esta diferencia no afectará el estimador de la proporción de la población.

Ejemplo

En el estudio del ejemplo anterior también se calcula la proporción de ciudades que cuentan con recolectores de basura privados. El estudiante desea que el margen de error se encuentre a 0.10 de la proporción de la población; el nivel de confianza deseado es de 90%, y no se encuentra disponible ningún estimador de la proporción de la población. ¿Cuál es el tamaño de la muestra que se requiere?

Solución

El estimador de la proporción de la población se encuentra a 0.10, por lo que $E = 0.10$. El nivel de confianza deseado es de 0.90, que corresponde a un valor z de 1.65. Como no se encuentra disponible ningún estimador de la población, se utiliza 0.50. El número de observaciones que se sugiere es

$$n = (.5)(1 - .5)\left(\frac{1.65}{.10}\right)^2 = 68.0625$$

El investigador necesita una muestra aleatoria de 69 ciudades.

Autoevaluación 9-4

El secretario académico de la universidad desea calcular el promedio aritmético de las calificaciones de los estudiantes que se graduaron durante los pasados 10 años. Los promedios oscilan entre 2.0 y 4.0. El promedio se va a calcular a 0.05 más o menos de la media poblacional. La desviación estándar se calcula que es de 0.279. Utilice el nivel de confianza de 99%. ¿Ayudaría al secretario a determinar cuántas boletas tiene que estudiar?

Ejercicios

connect™

19. Se calcula que una población tiene una desviación estándar de 10. Desea estimar la media de la población a menos de 2 unidades del error máximo admisible, con un nivel de confianza de 95%. ¿De qué tamaño debe ser la muestra?

20. Quiere estimar la media de la población a menos de 5, con un nivel de confianza de 99%. Se calcula que la desviación estándar es de 15. ¿De qué tamaño debe ser la muestra?

21. El estimador de la proporción poblacional debe estar a más o menos 0.05, con un nivel de confianza de 95%. El mejor estimador de la proporción poblacional es de 0.15. ¿De qué tamaño debe ser la muestra que se requiere?

22. El estimador de la proporción poblacional debe estar a más o menos de 0.10, con un nivel de confianza de 99%. El mejor estimador de la proporción poblacional es de 0.45. ¿De qué tamaño debe ser la muestra que se requiere?

23. Se planea llevar a cabo una encuesta para determinar el tiempo medio que ven televisión los ejecutivos corporativos. Una encuesta piloto indicó que el tiempo medio por semana es de 12 horas, con una desviación estándar de 3 horas. Se desea calcular el tiempo medio que se ve televisión menos de un cuarto de hora. Se utilizará el nivel de confianza de 95%. ¿A cuántos ejecutivos debe entrevistarse?

24. Un procesador de zanahorias corta las hojas, lava las zanahorias y las inserta en un paquete. En una caja se guardan veinte paquetes para enviarse. Para controlar el peso de las cajas, se revisaron unas cuantas. El peso medio fue de 20.4 libras, y la desviación estándar, de 0.5 libras. ¿Cuántas cajas debe tener la muestra para conseguir una confianza de 95% de que la media de la muestra no difiere de la media de la población por más de 0.2 libras?

25. Suponga que el presidente de Estados Unidos desea un cálculo de la proporción de la población que apoya su actual política relacionada con las revisiones del sistema de seguridad social. El presidente quiere que el cálculo se encuentre a menos de 0.04 de la proporción real. Suponga un nivel de confianza de 95%. Los asesores políticos del presidente calculan que la proporción que apoya la actual política es de 0.60.

a) ¿De qué tamaño debe ser la muestra que se requiere?

b) ¿De qué tamaño debe ser una muestra si no hubiera disponible ningún estimador de la proporción que apoya la actual política?

26. Las encuestas anteriores revelan que 30% de los turistas que van a Las Vegas a jugar durante el fin de semana gasta más de $1 000 cada uno. La gerencia desea actualizar este porcentaje.

a) El nuevo estudio utilizará el nivel de confianza de 90%. El estimador estará a menos de 1% de la proporción de la población. ¿Cuál es el tamaño necesario de la muestra?

b) La gerencia indicó que el tamaño de la muestra determinado es demasiado grande. ¿Qué se puede hacer para reducir la muestra? Con base en su sugerencia, vuelva a calcular el tamaño de la muestra.

9.6 Factor de corrección de una población finita

Las poblaciones de las que se han tomado muestras hasta ahora han sido muy grandes o infinitas. ¿Qué sucedería si la población de la que se toma la muestra no fuera muy grande? Es necesario realizar algunos ajustes en la forma de calcular el error estándar de las medias muestrales y del error estándar de las proporciones muestrales.

OA7 Ajustar el intervalo de confianza de poblaciones finitas.

Una población con un límite superior es *finita.* Por ejemplo, hay 12 179 estudiantes en la matrícula de la Eastern Illinois University; hay 40 empleados en Spence Sprockets; Chrysler ensambló 917 Jeeps Wrangler en la planta de Alexis Avenue el día de ayer; o había 65 pacientes programados para cirugía en St. Rose Memorial Hospital en Sarasota el día de ayer. Una población finita puede ser muy pequeña; puede constar de todos los estudiantes registrados para este curso. También puede ser muy grande, como todas las personas de la tercera edad que viven en Florida.

En el caso de una población finita, en la que el número total de objetos o individuos es N y el número de objetos o individuos incluidos en la muestra es n, es necesario ajustar los errores muestrales en las fórmulas de los intervalos de confianza. En otras palabras, para determinar el intervalo de confianza de la media, se ajusta el error estándar de la media en las fórmulas (9-1) y (9-2). Si quiere determinar el intervalo de confianza de una proporción, necesita ajustar el error estándar de la proporción en la fórmula (9-3).

Este ajuste recibe el nombre de **factor de corrección de una población finita**. Con frecuencia se le abrevia *FCP,* el cual es:

$$FPC = \sqrt{\frac{N - n}{N - 1}}$$

¿Por qué es necesario aplicar un factor y cuál es el efecto de hacerlo? Por lógica, si la muestra es un porcentaje significativo de la población, el estimador es más preciso. Observe el efecto del término $(N - n)/(N - 1)$. Suponga que la población es de 1 000 y que la muestra es de 100. Entonces esta razón es de (1 000 − 100)/(1 000 − 1), o 900/999. Al extraer la raíz cuadrada se obtiene el factor de corrección 0.9492. Al multiplicar este factor de corrección por el error estándar, se *reduce* el error estándar aproximadamente 5% (1 − 0.9492 = 0.0508). Esta reducción de la magnitud del error estándar da como resultado un intervalo menor de valores al calcular la media poblacional o la proporción poblacional. Si la muestra es de 200, el factor de corrección es de 0.8949, lo cual significa que el error estándar se redujo más de 10%. La tabla 9-2 muestra los efectos de diversos tamaños de muestras.

TABLA 9-2 Factor de corrección de una población finita de muestras seleccionadas cuando la población es de 1 000.

Tamaño de la muestra	Fracción de la población	Factor de corrección
10	.010	.9955
25	.025	.9879
50	.050	.9752
100	.100	.9492
200	.200	.8949
500	.500	.7075

Así, si quisiera construir un intervalo de confianza de una media a partir de una población finita sin conocer la desviación estándar de la población, la fórmula (9-2) se ajusta de la siguiente manera:

$$\bar{X} \pm t \frac{s}{\sqrt{n}}\left(\sqrt{\frac{N - n}{N - 1}}\right)$$

Haría un ajuste similar en la fórmula (9-3), en caso de una proporción.

El siguiente ejemplo resume los pasos para determinar un intervalo de confianza de una media.

Ejemplo

Hay 250 familias en Scandia, Pennsylvania. Una muestra aleatoria de 40 de estas familias revela que la contribución anual media a la iglesia fue de $450, y la desviación estándar, de $75. ¿La media poblacional puede ser de $445 o $425?

1. ¿Cuál es la media de la población? ¿Cuál es el mejor estimador de la media poblacional?
2. Construya el intervalo de confianza de 90% de la media de la población. ¿Cuáles son los puntos extremos del intervalo de confianza?
3. Interprete el intervalo de confianza.

Solución

Primero observe que la población es finita. Es decir, existe un límite para el número de personas que hay en Scandia, en este caso, 250.

1. No conoce la media poblacional, que es el valor que quiere calcular. El mejor estimador de la media poblacional es la media de la muestra, que es de $450.
2. La fórmula para determinar el intervalo de confianza de una media de población es la siguiente:

$$\bar{X} \pm t \frac{s}{\sqrt{n}}\left(\sqrt{\frac{N-n}{N-1}}\right)$$

En este caso, sabe que $\bar{X} = 450$, $s = 75$, $N = 250$ y que $n = 40$. No conoce la desviación estándar de la población, así que utiliza la distribución *t*. Para hallar el valor apropiado de *t* recurra al apéndice B.2, recorra la parte superior del renglón hasta la columna con el encabezamiento de 90%. Los grados de libertad son: $gl = n - 1 = 40 - 1 = 39$; así, vaya a la celda en la que el renglón de *gl* de 39 interseca la columna con el encabezamiento de 90%. El valor es de 1.685. Al sustituir estos valores en la fórmula, se obtiene:

$$\bar{X} \pm t \frac{s}{\sqrt{n}}\left(\sqrt{\frac{N-n}{N-1}}\right)$$

$$= \$450 \pm 1.685 \frac{\$75}{\sqrt{40}}\left(\sqrt{\frac{250-40}{250-1}}\right) = \$450 \pm \$19.98\sqrt{.8434} = \$450 \pm \$18.35$$

Los puntos extremos del intervalo de confianza son $431.65 y $468.35.

3. Es probable que la media poblacional sea de más de $431.65 e inferior a $468.35. En otras palabras, ¿la media de la población puede ser de $445? Sí, pero no es probable que sea de $425. ¿Por qué? Porque el valor de $445 se encuentra dentro del intervalo de confianza y $425 no pertenece al intervalo de confianza.

Autoevaluación 9-5

El mismo estudio relacionado con las contribuciones para la iglesia en Scandia reveló que 15 de las 40 familias tomadas de la muestra asiste regularmente a la iglesia. Construya el intervalo de confianza de 95% de la población de familias que asiste a la iglesia con regularidad.

Ejercicios

connect™

27. Se seleccionan al azar 36 artículos de una población de 300. La media de la muestra es de 35, y la desviación estándar, de 5. Construya el intervalo de confianza de 95% de la media poblacional.

28. Se seleccionan al azar 45 elementos de una población de 500. La media muestral es de 40 y la desviación estándar de la muestra es de 9. Construya el intervalo de confianza de 99% de la media poblacional.

29. La asistencia al juego de béisbol de la liga menor de Savannah Colts de la noche anterior fue de 400. Una muestra aleatoria de 50 asistentes reveló que la cantidad media de refrescos consumidos por persona fue de 1.86, con una desviación estándar de 0.50. Construya el intervalo de confianza de 99% de la cantidad media de refrescos consumidos por persona.

30. Hay 300 soldadores en Maine Shipyards Corporation. Una muestra de 30 de ellos reveló que 18 se graduaron en un curso de soldadura certificado. Construya el intervalo de confianza de 95% de la proporción de soldadores graduados en un curso de soldadura certificado.

Resumen del capítulo

I. Un estimador puntual es un solo valor (estadístico) para estimar un valor de la población (parámetro).

II. Un intervalo de confianza es un conjunto de valores entre los cuales se espera que ocurra el parámetro de la población.

- **A.** Los factores que determinan la magnitud de un intervalo de confianza de una media son:
  - **1.** El número de observaciones en la muestra, *n.*
  - **2.** La variabilidad en la población, normalmente calculada por la desviación estándar de la muestra, *s.*
  - **3.** El nivel de confianza.
    - ***a*)** Para determinar los límites de confianza cuando se conoce la desviación estándar de la población se utiliza la distribución *z.* La fórmula es:

$$\bar{X} \pm z\frac{\sigma}{\sqrt{n}} \quad \textbf{(9-1)}$$

    - ***b*)** Para determinar los límites de confianza cuando no se conoce la desviación estándar de la población se utiliza la distribución *t.* La fórmula es:

$$\bar{X} \pm t\frac{s}{\sqrt{n}} \quad \textbf{(9-2)}$$

III. Las principales características de la distribución *t* son:

- **A.** Es una distribución continua.
- **B.** Tiene forma de campana y es simétrica.
- **C.** Es plana, o más amplia, que la distribución normal estándar.
- **D.** Existe una familia de distribuciones *t,* según el número de grados de libertad.

IV. Una proporción es una razón, fracción o porcentaje que indica la parte de la muestra o población que posee una característica particular.

- **A.** Una proporción muestral se determina por medio de *X,* el número de éxitos, dividido entre *n,* el número de observaciones.
- **B.** Se construyó un intervalo de confianza de una proporción muestral con la siguiente fórmula:

$$p \pm z\sqrt{\frac{p(1-p)}{n}} \quad \textbf{(9-4)}$$

V. Es posible determinar un tamaño apropiado de muestra para calcular tanto medias como proporciones.

- **A.** Hay tres factores que determinan el tamaño de una muestra cuando desea calcular la media.
  - **1.** El margen de error máximo, *E.*
  - **2.** El nivel de confianza deseado.
  - **3.** La variación en la población.
  - **4.** La fórmula para determinar el tamaño muestral de la media es:

$$n = \left(\frac{z\sigma}{E}\right)^2 \quad \textbf{(9-5)}$$

- **B.** Hay tres factores que determinan el tamaño de una muestra cuando desea calcular una proporción.
  - **1.** El margen de error, *E.*
  - **2.** El nivel de confianza deseado.
  - **3.** Un valor de π para calcular la variación en la población.
  - **4.** La fórmula para determinar el tamaño muestral de una proporción es:

$$n = \pi(1-\pi)\left(\frac{z}{E}\right)^2 \quad \textbf{(9-6)}$$

VI. En el caso de una población finita, el error estándar se ajusta con el factor $\sqrt{\frac{N-n}{N-1}}$.

Ejercicios del capítulo

connect™

31. Una muestra aleatoria de 85 líderes de grupo, supervisores y personal similar de General Motors reveló que, en promedio, pasan 6.5 años en su trabajo antes de ascender. La desviación estándar de la muestra fue de 1.7 años. Construya el intervalo de confianza de 95 por ciento.

32. A un inspector de carne del estado de Iowa se le encargó calcular el peso neto medio de los paquetes de carne molida con la etiqueta “3 libras”. Por supuesto, se da cuenta de que los paquetes no pesan precisamente 3 libras. Una muestra de 36 paquetes revela que el peso medio es de 3.01 libras, con una desviación estándar de 0.03 libras.

a) ¿Cuál es la media poblacional estimada?

b) Determine el intervalo de confianza de 95% de la media poblacional.

33. Como parte de su paquete promocional, la Cámara de Comercio de Milwaukee desea tener una estimación del costo medio mensual de un apartamento de una recámara. Una muestra aleatoria de 40 apartamentos disponibles para renta reveló que el costo medio mensual era de $323. La desviación estándar de la muestra es $25.

a) Determine un intervalo de confianza de 98% para el precio medio de la población.

b) ¿Es razonable concluir que la media poblacional fue de $350 por mes?

34. Una encuesta reciente a 50 ejecutivos despedidos reveló que tardaron 26 semanas en colocarse en otro puesto. La desviación estándar de la muestra fue de 6.2 semanas. Construya el intervalo de confianza de 95% de la media de población. ¿Es razonable que la media poblacional sea de 28 semanas? Justifique su respuesta.

35. Marty Rowatti recién asumió el puesto de director de la YMCA de South Jersey. Le gustaría contar con datos recientes sobre el tiempo que han pertenecido a la YMCA sus miembros actuales. Para investigarlo, suponga que selecciona una muestra aleatoria de 40 miembros actuales. El tiempo medio de membresía de quienes se encuentran en la muestra es de 8.32 años, y la desviación estándar, de 3.07 años.

a) ¿Cuál es la media de la población?

b) Construya un intervalo de confianza de 90% para la media poblacional.

c) La directora anterior, en el breve informe que preparó al retirarse, indicó que ahora el tiempo medio de membresía era de “casi 10 años”. ¿Confirma la información esta aseveración? Cite evidencias.

36. La American Restaurant Association reunió información sobre la cantidad de veces que los matrimonios jóvenes comen fuera de casa a la semana. Una encuesta de 60 parejas indicó que la cantidad media de comidas fuera de casa es de 2.76 comidas semanales, con una desviación estándar de 0.75, también por semana. Construya el intervalo de confianza de 97% de la media poblacional.

37. La National Collegiate Athletic Association (NCAA) informó que la cantidad media de horas semanales que los asistentes de los entrenadores de fútbol invierten en entrenamiento y reclutamiento durante la temporada es de 70. Una muestra aleatoria de 50 asistentes indicó que la media de la muestra es de 68.6 horas, con una desviación estándar de 8.2 horas.

a) De acuerdo con los datos de la muestra, construya el intervalo de confianza de 99% de la media de la población.

b) ¿Incluye el intervalo de confianza de 99% el valor que sugiere la NCAA? Interprete este resultado.

c) Suponga que decidió cambiar el intervalo de confianza de 99 a 95%. Sin realizar cálculos, ¿aumentará el intervalo, se reducirá o permanecerá igual? ¿Qué valores de la fórmula cambiarán?

38. El Departamento de Recursos Humanos de Electronics, Inc., desea incluir un plan dental como parte del paquete de prestaciones. La pregunta que se plantea es: ¿cuánto invierte un empleado común y su familia en gastos dentales al año? Una muestra de 45 empleados revela que la cantidad media que se invirtió el año pasado fue de $1 820, con una desviación estándar de $660.

a) Construya el intervalo de confianza de 95% de la media poblacional.

b) Al presidente de Electronics, Inc., se le proporcionó la información del inciso *a).* Éste indicó que podía pagar $1 700 de gastos dentales por empleado. ¿Es posible que la media poblacional pudiera ser de $1 700? Justifique su respuesta.

39. Un estudiante llevó a cabo un estudio e informó que el intervalo de confianza de 95% de la media variaba de 46 a 54. Estaba seguro de que la media de la muestra era de 50; de que la desviación estándar de la muestra era de 16, y de que la muestra era de por lo menos 30 elementos, pero no recordaba el número exacto. ¿Puede usted ayudarle?

40. Un estudio reciente llevado a cabo por la American Automobile Dealers Association reveló que la cantidad media de utilidades por automóvil vendido en una muestra de 20 concesionarias fue de $290, con una desviación estándar de $125. Construya el intervalo de confianza de 95% de la media poblacional.

41. Un estudio de 25 graduados de universidades de cuatro años llevado a cabo por la American Banker’s Association reveló que la cantidad media que debía un estudiante por concepto de crédito estudiantil era de $14 381. La desviación estándar de la muestra fue de $1 892. Construya el intervalo de confianza de 90% de la media poblacional. ¿Es razonable concluir que la media de la población en realidad es de $15 000? Indique por qué.

42. Un factor importante en la venta de propiedades residenciales es la cantidad de personas que le echan un vistazo a las casas. Una muestra de 15 casas vendidas recientemente en el área de Buffalo, Nueva York, reveló que el número medio de personas que ven las casas fue de 24, y la desviación estándar de la muestra, de 5 personas. Construya el intervalo de confianza de 98% de la media poblacional.

43. Warren County Telephone Company afirma en su informe anual que "el consumidor habitual gasta $60 mensuales en el servicio local y de larga distancia". Una muestra de 12 abonados reveló las cantidades que gastaron el mes pasado.

$64	$66	$64	$66	$59	$62	$67	$61	$64	$58	$54	$66

a) ¿Cuál es el estimador puntual de la media poblacional?
b) Construya el intervalo de confianza de 90% de la media poblacional.
c) ¿Es razonable la afirmación de la compañía de que el "consumidor habitual" gasta $60 mensuales? Justifique su respuesta.

44. El fabricante de una nueva línea de impresoras de inyección de tinta desea incluir, como parte de su publicidad, el número de páginas que el usuario puede imprimir con un cartucho. Una muestra de 10 cartuchos reveló el siguiente número de páginas impresas.

2 698	2 028	2 474	2 395	2 372	2 475	1 927	3 006	2 334	2 379

a) ¿Cuál es el estimador puntual de la media poblacional?
b) Construya el intervalo de confianza de 95% de la media poblacional.

45. La doctora Susan Benner es psicóloga industrial. En este momento estudia el estrés en los ejecutivos de las compañías de internet. Elaboró un cuestionario que cree que mide el estrés. Un resultado de 80 indica un nivel de estrés peligroso. Una muestra aleatoria de 15 ejecutivos reveló los siguientes niveles de estrés.

94	78	83	90	78	99	97	90	97	90	93	94	100	75	84

a) Determine el nivel medio de estrés de esta muestra. ¿Cuál es el estimador puntual de la media poblacional?
b) Construya el intervalo de confianza de 95% de la media poblacional.
c) ¿Es razonable concluir que los ejecutivos de internet tienen un nivel medio de estrés peligroso, según el cuestionario de la doctora Benner?

46. Como requisito para obtener el empleo, los candidatos de Fashion Industries deben pasar por una prueba de drogas. De los últimos 220 solicitantes, 14 reprobaron. Construya el nivel de confianza de 99% de la proporción de solicitantes que no pasan la prueba. ¿Es razonable concluir que más de 10% de los solicitantes no la superan?

47. Fashion Industries aplica pruebas aleatorias a sus empleados a lo largo del año. El año pasado, de las 400 pruebas aleatorias aplicadas, 14 empleados no pasaron. ¿Es razonable concluir que menos de 5% de los empleados no pasan la prueba aleatoria de drogas? Explique su respuesta.

48. Durante un debate nacional sobre cambios en el sistema de salud, un servicio de noticias por cable realizó una encuesta de opinión entre 500 pequeños propietarios de empresas. Se reveló que 65% de estos pequeños empresarios no aprueban los cambios. Construya el intervalo de confianza de 95% de la proporción que se opone a dichos cambios en el sistema de salud. Comente los resultados.

49. En York County, Carolina del Sur, hay 20 000 votantes. Una muestra aleatoria de 500 votantes de esa localidad reveló que 350 planean votar por el regreso al senado de Louella Millar. Construya el intervalo de confianza de 99% de la proporción de votantes en el condado que planea votar por Millar. A partir de la información de esta muestra, ¿es posible confirmar su reelección?

50. En una encuesta para medir la popularidad del presidente, se pidió a una muestra aleatoria de 1 000 electores que marcara una de las siguientes afirmaciones:

1. El presidente hace un buen trabajo.
2. El presidente realiza un trabajo deficiente.
3. Prefiero no opinar.

Un total de 560 entrevistados eligió la primera afirmación e indicó que considera que el presidente realiza un buen trabajo.

a) Construya el intervalo de confianza de 95% de la proporción de entrevistados que piensan que el presidente hace un buen trabajo.

b) Con base en el intervalo del inciso *a*), ¿es razonable llegar a la conclusión de que la mayoría (más de la mitad) de la población considera que el presidente realiza un buen trabajo?

51. Edward Wilkin, jefe de la policía de River City, informa que hubo 500 infracciones de tránsito el mes pasado. Una muestra de 35 de estas infracciones mostró que la suma media de las multas fue de \$54, con una desviación estándar de \$4.50. Construya el intervalo de confianza de 95% de la suma media de una infracción en River City.

52. El First National Bank de Wilson tiene 650 clientes con cuentas de cheques. Una encuesta reciente de 50 de estos clientes mostró que 26 tenían una tarjeta Visa con el banco. Construya el intervalo de confianza de 99% de la proporción de clientes con cuenta de cheques que tienen una tarjeta Visa con el banco.

53. Se estima que 60% de los hogares en Estados Unidos contrata televisión por cable. A usted le gustaría verificar esta afirmación para su clase de comunicación masiva. Si desea que su estimador se encuentre a menos de 5 puntos porcentuales con un nivel de confianza de 95%, ¿qué tamaño de muestra requiere?

54. Usted necesita calcular la cantidad media de días que viajan al año los vendedores. La media de un pequeño estudio piloto fue de 150 días, con una desviación estándar de 14 días. Si usted debe calcular la media poblacional a menos de 2 días, ¿a cuántos vendedores debe incluir en la muestra? Utilice un intervalo de confianza de 90 por ciento.

55. Usted va a llevar a cabo el sondeo de una muestra para determinar el ingreso medio familiar en un área rural del centro de Florida. La pregunta es: ¿a cuántas familias se debe incluir en la muestra? En una muestra piloto de 10 familias, la desviación estándar de la muestra fue de \$500. El patrocinador de la encuesta desea que usted utilice un nivel de confianza de 95%. El estimador debe estar dentro de un margen de \$100. ¿A cuántas familias debe entrevistar?

56. *Families USA,* revista mensual que trata temas relacionados con la salud y sus costos, encuestó a 20 de sus suscriptores. Encontró que las primas anuales de seguros de salud para una familia con cobertura de una empresa promediaron \$10 979. La desviación estándar de la muestra fue de \$1 000.

a) Con base en la información de esta muestra, construya el intervalo de confianza de 90% de la prima anual media de la población.

b) ¿De qué tamaño debe ser la muestra para que la media poblacional se encuentre dentro de un margen menor a \$250, con 99% de confianza?

57. La presurización en la cabina del avión influye en la comodidad de los pasajeros. Una presurización más alta permite un ambiente más cercano a lo normal y un vuelo más relajado. Un estudio que llevó a cabo un grupo de usuarios de aerolíneas registró la presión de aire correspondiente a 30 vuelos elegidos de forma aleatoria. El estudio reveló una presión equivalente media de 8 000 pies, con una desviación estándar de 300 pies.

a) Establezca un intervalo de confianza de 99% para la presión equivalente de la media poblacional.

b) ¿De qué tamaño necesita ser la muestra para que la media de la población se encuentre dentro de un margen de 25 pies, con una confianza de 95 por ciento?

58. Una muestra aleatoria de 25 personas empleadas por las autoridades del estado de Florida estableció que ganaban un salario promedio (con prestaciones) de \$65.00 por hora. La desviación estándar es de \$6.25 por hora.

a) ¿Cuál es la media de la población? ¿Cuál es el mejor estimador de la media poblacional?

b) Construya el intervalo de confianza de 99% del salario medio de la población (con prestaciones) de estos empleados.

c) ¿De qué tamaño debe ser la muestra para calcular la media de la población con un error admisible de \$1.00, con una confianza de 95 por ciento?

59. Una alianza cinematográfica utilizó una muestra aleatoria de 50 ciudadanos estadounidenses para calcular que el estadounidense común vio videos y películas en DVD 78 horas el año pasado. La desviación estándar de esta muestra fue de 9 horas.

a) Construya el intervalo de confianza de 95% de la cantidad media poblacional de horas empleadas en ver videos y películas en DVD el año pasado.

b) ¿De qué tamaño debe ser la muestra para que resulte 90% confiable de que la media de la muestra se encuentra dentro de un margen de 1.0 hora de la media de la población?

60. Dylan Jones lleva registros meticulosos de la eficiencia en el gasto de combustible de su nuevo auto. Después de las primeras nueve veces que llenó el tanque, encontró que la media era de 23.4 millas por galón (mpg) con una desviación estándar muestral de 0.9 mpg.

a) Calcule el intervalo de confianza del 95% para su mpg.

b) ¿Cuántas veces debe llenar el tanque de gasolina para obtener un margen de error por debajo de 0.1 mpg?

61. Una encuesta a 36 propietarios de iPhone seleccionados al azar mostró que el precio de compra tiene una media de \$416, con una desviación estándar muestral de \$180.

a) Calcule el error estándar de la media muestral.

b) Calcule el intervalo de confianza de 95% de la media.

c) ¿De qué tamaño debe ser la muestra para estimar la media poblacional dentro de \$10?

62. Usted planea llevar a cabo una encuesta para hallar la proporción de fuerza laboral con dos o más trabajos. Decide con base en un nivel de confianza de 95%, y establece que la proporción estimada debe encontrarse en un margen de menos de 2% de la proporción poblacional. Una encuesta piloto revela que 5 de 50 de los entrevistados tenían dos o más trabajos. ¿A cuántos trabajadores debe entrevistar para satisfacer los requisitos?

63. La proporción de contadores públicos que cambiaron de empresa en los últimos tres años se debe calcular con un margen de 3%. Es necesario utilizar el nivel de confianza de 95%. Un estudio que se realizó hace varios años reveló que el porcentaje de contadores públicos que cambió de compañía en tres años fue de 21.

a) Para actualizar el estudio, ¿cuál es el número de expedientes de contadores públicos que se deben estudiar?

b) ¿Con cuántos contadores públicos es necesario ponerse en contacto si no se cuenta con estimadores anteriores de la proporción poblacional?

64. Como parte de una revisión anual de sus cuentas, un corredor selecciona una muestra aleatoria de 36 clientes. Al revisar sus cuentas, calculó una media de \$32 000, con una desviación estándar muestral de \$8 200. ¿Cuál es el intervalo de confianza de 90% del valor medio de las cuentas de la población de clientes?

65. El Registro Nacional de Control de peso trata de obtener secretos de éxito de gente que ha perdido cuando menos 30 libras y mantuvo su peso por al menos un año. La dependencia reporta que de 2 700 registrados, 459 estuvieron en una dieta baja en carbohidratos (menos de 90 gramos al día).

a) Construya el intervalo de confianza de 95% de esta fracción.

b) ¿Es posible que el porcentaje de la población sea 18 por ciento?

c) ¿Qué tan grande debe ser la muestra para estimar la proporción dentro de 0.5 por ciento?

66. Cerca ya de las elecciones, un servicio de noticias por cable conduce una encuesta de opinión de 1 000 probables votantes. El resultado muestra que el contendiente republicano tiene una ventaja de 52 a 48 por ciento.

a) Construya el intervalo de confianza de 95% de la proporción que favorece al candidato republicano.

b) Calcule la probabilidad de que el candidato demócrata sea el líder real.

c) Repita el análisis anterior basándose en una muestra de 3 000 probables votantes.

67. Una muestra de 352 suscriptores de la revista *Wired* indicó que el tiempo medio invertido en el uso de internet es de 13.4 horas a la semana, con una desviación estándar de 6.8 horas. Determine un intervalo de confianza de 95% del tiempo medio que pasan los suscriptores en la red.

68. El Tennessee Tourism Institute (TTI) planea hacer un muestreo de la información que proporcione una muestra de los visitantes que ingresan al estado para saber cuántos de ellos van a acampar. Los cálculos actuales indican que acampa 35% de los visitantes. ¿De qué tamaño debe ser la muestra para calcular la proporción de la población con un nivel de confianza de 95% y un error admisible de 2 por ciento?

Ejercicios de la base de datos

69. Consulte los datos de Real State, con información sobre las casas vendidas en Goodyear, Arizona, el año pasado.

a) Construya el intervalo de confianza de 95% del precio de venta medio de las casas.

b) Construya el intervalo de confianza de 95% de la distancia media de la casa al centro de la ciudad.

c) Construya el intervalo de confianza de 95% de la proporción de casas con garage.

d) Para reportar sus hallazgos, redacte un memo de negocios a Gary Loftus, presidente de la Cámara de Comercio de Goodyear.

70. Consulte los datos Baseball 2009, con información sobre los 30 equipos de la Liga Mayor de Béisbol de la temporada 2009.

a) Construya el intervalo de confianza de 95% de la cantidad media de cuadrangulares por equipo.

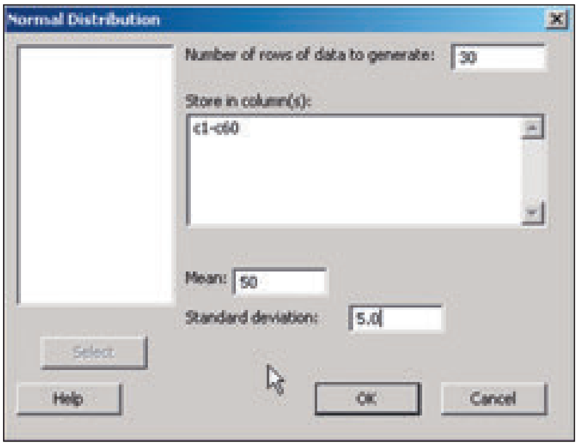
b) Construya el intervalo de confianza de 95% de la cantidad media de errores que cometió cada equipo.

c) Construya el intervalo de confianza de 95% de la cantidad media de robos de base de cada equipo.

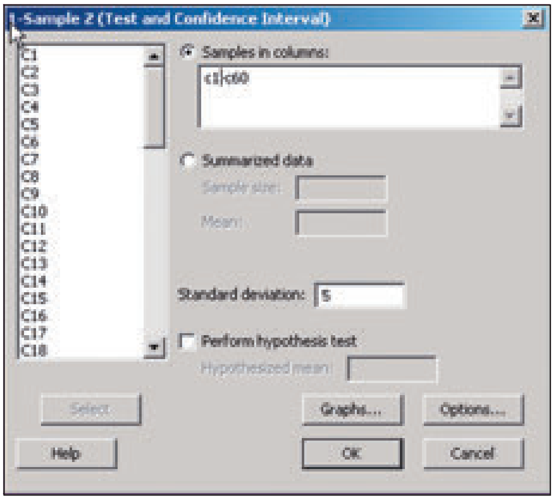
71. Consulte los datos de los autobuses del Distrito Escolar Buena.
- ***a)*** Construya el intervalo de confianza de 95% del mantenimiento medio de los autobuses.
- ***b)*** Construya el intervalo de confianza de 95% del millaje medio de los autobuses.
- ***c)*** Redacte un memo de negocios para el oficial estatal de transporte para reportar sus resultados.

Comandos de software

1. Los comandos de Minitab de las 60 columnas de 30 números aleatorios del ejemplo con solución de la página 304 son los siguientes**:**
- ***a)*** Seleccione **Calc**, **Random Data** y haga clic en **Normal**.
- ***b)*** En el cuadro de diálogo, haga clic en **Generate**; escriba *30* para el número de hileras de datos; *C1-C60* en **Store in column(s)**; *50,* en **Mean**; *5.0* en **Standard deviation**, y finalmente haga clic en **OK**.

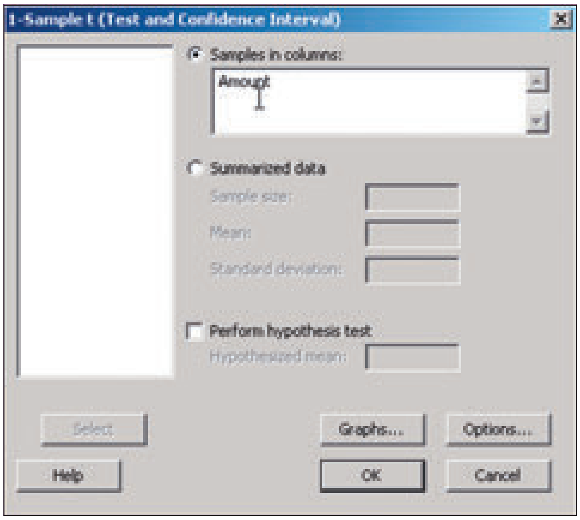


2. A continuación se presentan los comandos Minitab para los 60 intervalos de confianza de la página 304.
- ***a)*** Seleccione **Stat**, **Basic Statistics** y haga clic en **1-Sample Z**.
- ***b)*** En el cuadro de diálogo indique que las **Variables** son *C1-C60* y que la **Standard Duration** es de *5.* En seguida haga clic en **Options**, en la esquina inferior izquierda; en el siguiente cuadro de diálogo indique que el **Confidence level** es de *95* y haga clic en **OK**. Haga clic en **OK** en el cuadro de diálogo principal.

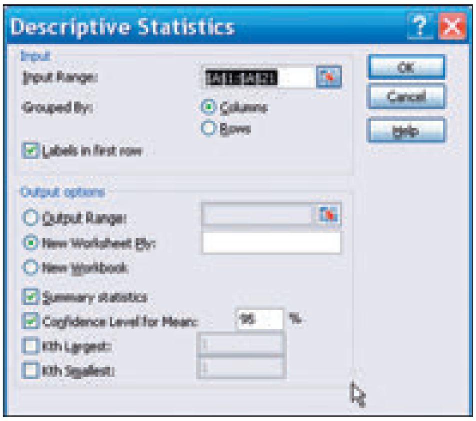


3. A continuación aparecen los comandos Minitab correspondientes a la estadística descriptiva de la página 311. Introduzca los datos en la primera columna y rotúlela *Amount.* En la barra de herramientas seleccione **Stat**, **Basic Statistics** y **Display Descriptive Statistics**. En el cuadro de diálogo seleccione *Amount* como **Variable** y haga clic en **OK**.

4. Los comandos Minitab para el intervalo de confianza de la cantidad que se gasta en el centro comercial de Inlet Square de la página 311 son:
- ***a)*** Introduzca las 20 cantidades gastadas en la columna *C1* y dé a la variable el nombre de *Amount*. Éste se llama **Shopping** y se localiza en la carpeta para el capítulo 9.
- ***b)*** En la barra de herramientas, seleccione **Stat**, **Basic Statistics** y haga clic en **1-Sample t**.
- ***c)*** Seleccione **Samples in columns:**, seleccione **Amount** y haga clic en **OK**.



5. Los comandos de Excel para el intervalo de confianza de las cantidades que se gastan en el centro comercial de Inlet Square de la página 312 son los siguientes:
   a) De la barra de menú, seleccione **Data**. En el extremo derecho, seleccione, **Data Analysis** y **Descriptive Statistics**, y haga clic en **OK**.
   b) Para el **Input Range** escriba *A1:A21,* haga clic en **Labels in first row**, escriba *C1* como **Output Range**, haga clic en **Summary statistics** y **Confidence Level for Mean**, y, en seguida, en **OK**.



Capítulo 9 Respuestas a las autoevaluaciones

9-1 ***a)*** Desconocido. Se trata del valor que se desea calcular.
b) \$20 000, estimador puntual.
c) $\$20\ 000 \pm 2.58\dfrac{\$3\ 000}{\sqrt{40}} = \$20\ 000 \pm \$1\ 224$
d) Los puntos extremos del intervalo de confianza son \$18 776 y \$21 224. Aproximadamente 99% de los intervalos construidos de forma similar incluirían la media poblacional.

9-2 ***a)*** $\bar{X} = \dfrac{18}{10} = 1.8 \qquad s = \sqrt{\dfrac{11.6}{10 - 1}} = 1.1353$
b) La media poblacional no se conoce. El mejor estimador es la media de la muestra, 1.8 días.
c) $1.80 \pm 2.262\dfrac{1.1353}{\sqrt{10}} = 1.80 \pm 0.81$

Los puntos extremos son 0.99 y 2.61.
d) Se utiliza *t* porque no se conoce la desviación estándar.
e) El valor de 0 no se encuentra en el intervalo. No es razonable concluir que la cantidad media de días de ausencias laborales sea de 0 por empleado.

9-3 ***a)*** $p = \dfrac{420}{1\ 400} = .30$
b) $30 \pm 2.58(.0122) = .30 \pm .03$
c) El intervalo se encuentra entre 0.27 y 0.33. Alrededor de 99% de los intervalos construidos de forma similar incluirían la media poblacional.

9-4 $n = \left(\dfrac{2.58(.279)}{.05}\right)^2 = 207.26$. La muestra debe redondearse a 208.

9-5 $.375 \pm 1.96\sqrt{\dfrac{.375(1 - .375)}{40}}\sqrt{\dfrac{250 - 40}{250 - 1}} =$

$.375 \pm 1.96(.0765)(.9184) = .375 \pm .138$

Repaso de los capítulos 8 y 9

El capítulo 8 comenzó con la descripción de las razones por las que es necesario el muestreo. Se hacen muestreos porque con frecuencia es imposible estudiar cada elemento o individuo que integran algunas poblaciones. Resultaría muy costoso y consumiría demasiado tiempo, por ejemplo, ponerse en contacto con todos los ejecutivos de bancos de Estados Unidos y registrar sus ingresos anuales. Asimismo, el muestreo con frecuencia destruye el producto. Un fabricante de medicamentos no puede probar las propiedades de cada tableta elaborada, pues no le quedaría nada para vender. Por consiguiente, para calcular un parámetro poblacional, se selecciona una muestra de la población. Una muestra forma parte de la población. Debe tenerse cuidado en garantizar que cada miembro de la población tenga la misma oportunidad de que se le elija; de otra manera, las conclusiones pueden estar sesga-