8

ESTIMACIÓN DE INTERVALOS

CONTENIDO

LA ESTADÍSTICA EN LA PRÁCTICA

DOLLAR GENERAL CORPORATION*

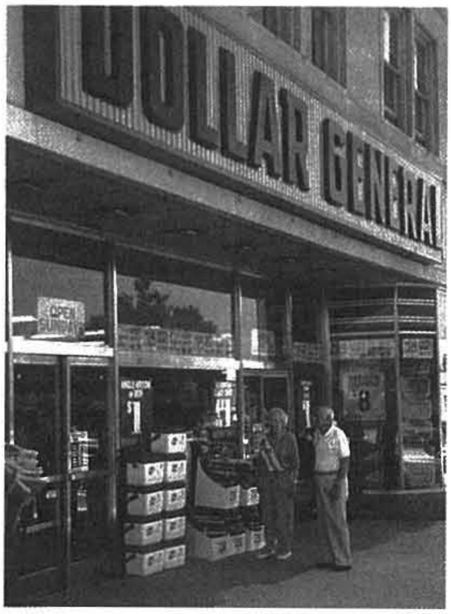
Nashville, Tennessee

Esta empresa fue fundada en 1939 como compañía vendedora de telas al mayoreo. Después de la Segunda Guerra Mundial, comenzó a operar locales en el centro y sur del estado de Kentucky con ventas al menudeo. Hoy, Dollar General Corporation tiene más de 2700 almacenes en 24 estados de la Unión Americana. Su clientela es, principalmente, de ingresos bajos y medios, y sus líneas son productos textiles y para la salud, belleza y limpieza, a precios bajos.

Siendo un negocio de gran movimiento de inventario con unos 17,000 productos distintos, Dollar General decidió adoptar el método LIFO (Last In, First Out: últimas llegadas, primeras salidas) para la valoración de sus inventarios. Este método compara los costos actuales con los ingresos actuales, con lo que se minimiza el efecto de cambios radicales en los precios sobre los resultados de pérdidas y ganancias. Además, el método LIFO reduce el ingreso neto, y con ello los impuestos al ingreso durante los periodos de inflación. Esto, a su vez, proporciona un flujo de efectivo generado por las ventas, según el ingreso, y permite remplazar el inventario a costos actuales.

Las prácticas contables requieren el establecimiento de un índice LIFO para el inventario, con el método de valoración LIFO. Por ejemplo, un índice LIFO de 1.028 indica que el valor de las existencias de la empresa a costos actuales refleja un 2.8% de aumento debido a la inflación en comparación con el último periodo anual.

Para establecer un índice LIFO se requiere que el inventario físico de cada producto a fin de año se evalúe el costo actual en ese fin de año, y el costo del fin del año precedente. Para evitar contar físicamente las existencias de cada producto en más de 2000 tiendas de menudeo, se selecciona una muestra aleatoria de 800 productos de 100 tiendas de menudeo y de tres bodegas. Se llevan a cabo los inventarios físicos de los 800 productos muestreados a final del año. El personal de contabilidad proporciona, entonces, los costos de ese fin de año y del fin de año anterior, necesarios para elaborar el índice LIFO.



Este almacén en Bedford, Indiana, es uno de entre más de 2700 de Dollar General Stores.

En uno de los últimos años, el índice LIFO fue de 1.030. Sin embargo, como este índice sólo es un estimado muestral del índice LIFO de población, se requiere conocer la precisión del estimado. Con base en los resultados de la muestra y con un 95% de nivel de confianza, se calculó que el margen de error era de .006. Así, el intervalo de 1.024 a 1.036 abarca el intervalo de 95% de confianza en el estimado poblacional del índice LIFO. Se juzgó la precisión como muy buena.

En este capítulo usted aprenderá cómo determinar una probabilidad acerca del error de muestreo asociado con la media de la muestra y la proporción de una muestra. A continuación verá cómo usar esa información para definir e interpretar estimados de intervalo de confianza de una media y una proporción poblacionales. También aprenderá a determinar el tamaño de muestra necesario para asegurar que el error de muestreo quede dentro de límites aceptables.

*Mr. Robert S. Knaul, contralor de Dollar General Corporation, proporcionó este ensayo.

En el capítulo 7 mostramos que el valor de la media de la muestra $\bar{x}$ y de la proporción de la muestra $\bar{p}$ dan como resultado estimados puntuales de la media μ y la proporción p de la población, respectivamente. Como los estimados puntuales se basan en una muestra de la población, no se espera que sean iguales al parámetro poblacional correspondiente. En este capítulo mostraremos cómo se elaboran estimados de intervalo de la media y de la proporción poblacionales, para proporcionar información sobre la precisión de un estimado. Como veremos, las distribuciones muestrales de $\bar{x}$ y $\bar{p}$ que presentamos en el capítulo 7 desempeñan un papel importante en el desarrollo de los estimados de intervalo de μ y p.

8.1 ESTIMACIÓN DE INTERVALO DE UNA MEDIA DE LA POBLACIÓN: CASO DE MUESTRA GRANDE

En esta sección mostraremos cómo usar la distribución muestral de $\bar{x}$ para determinar un estimado de intervalo de una media de población μ. Comenzaremos con el caso de una muestra grande ($n \geq 30$), en el que se conoce la desviación estándar σ de la población. Una vez logrado esto, indicaremos cómo hacerlo también para el caso de una muestra grande cuando se desconoce σ.

Comencemos con un estudio de muestreo que llevó a cabo CJW, Inc., una empresa de pedidos por correo que se especializa en artículos y accesorios deportivos. La empresa se esfuerza en proporcionar el mejor servicio posible a sus clientes. Para vigilar la calidad de su servicio, CJW selecciona una muestra aleatoria simple de clientes por correo cada mes; se establece contacto con cada cliente muestreado y se le hace una serie de preguntas acerca de la calidad del servicio al cliente. Las respuestas al cuestionario se usan para determinar una calificación de satisfacción por cada cliente muestreado; esas calificaciones van de 0 (la peor posible) a 100 (la mejor posible). Después se calcula una media de las calificaciones y se usa como estimado puntual de la media de satisfacción, para la población de todos los clientes de CJW.

Las encuestas mensuales anteriores han demostrado que, aunque la media de la muestra de calificación cambia de mes a mes, la desviación estándar de esas calificaciones ha tendido a estabilizarse en el valor 20. Por consiguiente, supondremos que la desviación estándar poblacional es $\sigma = 20$. La última encuesta de satisfacción del cliente produjo datos de 100 clientes ($n = 100$); la media de la muestra de calificaciones de satisfacción fue $\bar{x} = 82$. En la siguiente descripción usaremos los resultados de esta muestra para elaborar un estimado de intervalo de μ, la media de la población de calificación.

Error muestral

Siempre que se usa una media de muestra para proporcionar un estimado de punto de una media poblacional, alguien puede preguntar: ¿qué tan bueno es el estimado? La pregunta "¿qué tan bueno?", es una forma de indagar el error incurrido cuando se usa el valor de $\bar{x}$ como un estimado puntual de μ. En general, el valor absoluto de la diferencia entre un estimador puntual insesgado y el parámetro de población que estima se llama *error muestral*. Para el caso en el que la media de una muestra estima a una media poblacional, el error muestral es

El conocimiento de la distribución de $\bar{x}$ nos permite hacer afirmaciones probabilísticas acerca del error de muestreo, $|\bar{x} - \mu|$, aunque no se conozca el promedio poblacional μ.

$$\text{Error muestral} = |\bar{x} - \mu| \tag{8.1}$$

En la práctica no se puede determinar el valor del error muestral, porque no se conoce la media de la población μ. Sin embargo, se puede usar la distribución muestral de $\bar{x}$ para establecer márgenes de probabilidad respecto al tamaño del error muestral. Veremos cómo se hace con el estudio de muestreo de CJW.

Con una muestra de tamaño $n = 100$ y una desviación estándar de población $\sigma = 20$, el teorema del límite central que presentamos en el capítulo 7 nos permite llegar a la conclusión de que la distribución muestral de $\bar{x}$ se puede aproximar mediante una distribución normal de probabilidades con media μ y desviación estándar $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 20/\sqrt{100} = 2$. Esta distribución muestral aparece en la figura 8.1. Como la distribución muestral de $\bar{x}$ indica cómo se distribuyen sus valores en torno a μ, proporciona información acerca de las diferencias posibles entre $\bar{x}$ y μ. Podemos usar esa información para establecer aseveraciones probabilísticas acerca del error muestral.

Aseveraciones probabilísticas sobre el error muestral

Si usamos la terminología presentada en el capítulo 7, $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ se llama error estándar del promedio. En consecuencia, se puede decir que el 95% de los promedios muestrales deben quedar dentro de ± 1.96 errores estándar del promedio de la población.

Si usamos la tabla de áreas de la distribución normal estándar de probabilidades, veremos que el 95% de los valores de cualquier variable aleatoria con distribución normal quedan dentro de una distancia igual a ± 1.96 desviaciones estándar de la media. Por consiguiente, para la distribución muestral de la figura 8.1, el 95% de todos los valores de $\bar{x}$ debe estar a ± 1.96 desviaciones estándar o menos de μ. Como $1.96\ \sigma_{\bar{x}} = 1.96(2) = 3.92$, el 95% de las medias de muestra deben estar a ± 3.92 o menos de la media de la población.

La localización de las medias de muestra que resultan en un error muestral de 3.92 o menos se muestra en la figura 8.2. Observe que si una media de muestra está en la región identificada como "95% de los valores de $\bar{x}$", origina un error muestral de 3.92 o menos. Sin embargo, si una media de muestra está en la cola inferior o superior de la distribución, el error muestral será mayor de 3.92. Por consiguiente, podemos hacer la siguiente aseveración probabilística sobre el error muestral, para el problema de CJW.

> Hay una probabilidad de .95 de que la media de una muestra origine un error muestral de 3.92 o menos.

Esta aseveración probabilística acerca del error muestral es una *aseveración sobre la precisión*, que dice a CJW sobre el error que se puede esperar si se usa una muestra aleatoria simple de 100 clientes para estimar la media poblacional. Aunque se usa la probabilidad de .95, con frecuencia, para hacer aseveraciones sobre la precisión, también

Figura 8.1 DISTRIBUCIÓN DE LA MEDIA DE LA MUESTRA DE LA CALIFICACIÓN DETERMINADA A PARTIR DE MUESTRAS ALEATORIAS SIMPLES DE 100 CLIENTES

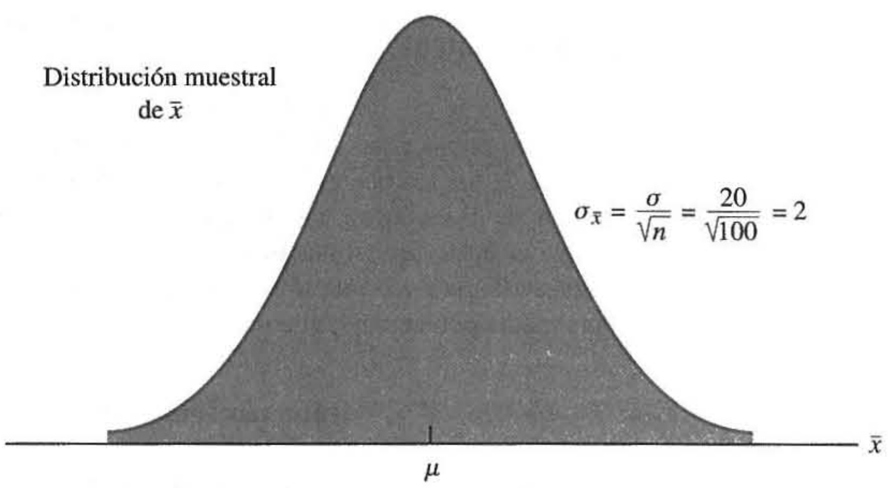


Figura 8.2 DISTRIBUCIÓN MUESTRAL DE $\bar{x}$ MOSTRANDO LA UBICACIÓN DE LAS MEDIAS DE MUESTRAS QUE ORIGINAN UN ERROR MUESTRAL DE 3.92 O MENOS

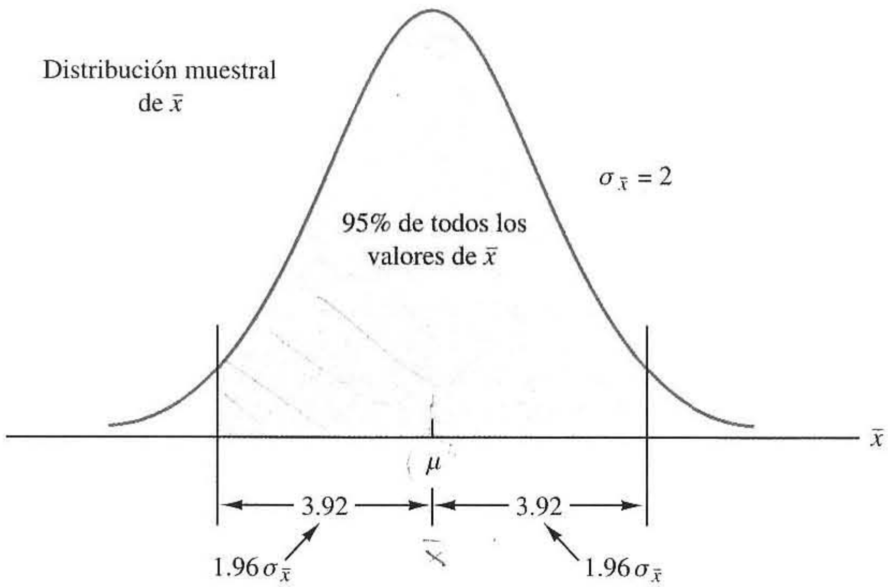
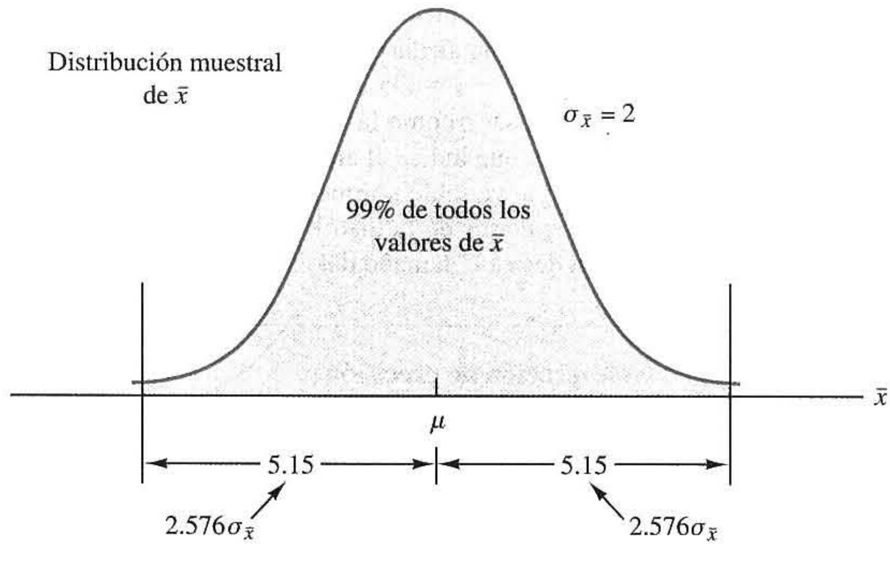
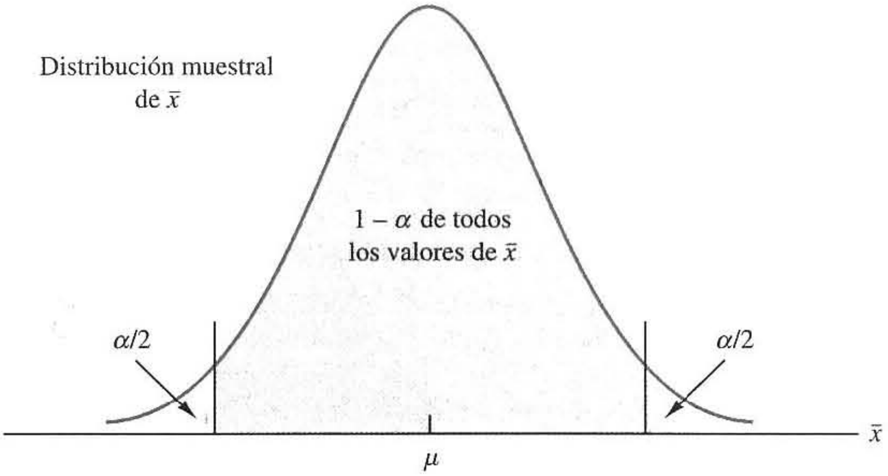


Figura 8.3 DISTRIBUCIÓN MUESTRAL DE $\bar{x}$ MOSTRANDO LA UBICACIÓN DEL 99% DE LOS VALORES DE $\bar{x}$



se pueden usar otros valores, por ejemplo .90 y .99. Por ejemplo, la figura 8.3 muestra la localización del 99% de las medias de muestra en el problema de CJW. Mediante la tabla de distribución normal estándar de probabilidades vemos que el 99% de los valores de $\bar{x}$ están a ± 2.576 desviaciones estándar de distancia respecto a μ, o menos. Como $2.576\sigma_{\bar{x}} = 2.576(2) = 5.15$, hay una probabilidad de .99 de que la media de muestra tenga un error muestral de 5.15 o menos. Igualmente, hay un .90 de probabilidad de que la media de muestra origine un error muestral de $1.645\sigma_{\bar{x}} = 1.645(2) = 3.29$ o menos.

Figura 8.4 ÁREAS DE UNA DISTRIBUCIÓN MUESTRAL DE $\bar{x}$ QUE SE USAN PARA ESTABLECER ASEVERACIONES PROBABILÍSTICAS SOBRE EL ERROR MUESTRAL



Generalicemos el procedimiento que usamos para hacer aseveraciones sobre la precisión del error muestral que se puede cometer siempre que se usa el valor de una media de muestra para estimar una media de población. Emplearemos la letra griega α para indicar la probabilidad de que el error muestral sea *mayor* que el correspondiente en la aseveración de precisión. En la figura 8.4, $\alpha/2$ representa el área, o la probabilidad, en cada cola de la distribución de muestreo, y $1 - \alpha$ el área, o la probabilidad, de que la media de la muestra origine un error muestral menor o igual al correspondiente de la aseveración de precisión. Por ejemplo, al afirmar que hay un .95 de probabilidad de que el valor de una media de la muestra origine un error muestral de 3.92 o menos, se basa en $\alpha = .05$ y $1 - \alpha = .95$. El área en cada cola de la distribución de muestreo es $\alpha/2 = .025$.

Para usar z como la variable aleatoria normal estándar colocamos cerca de ella un subíndice que indica el área en la cola superior de la distribución. En general, $z_{\alpha/2}$ es el valor de la variable normal estándar que corresponde a un área de $\alpha/2$ en la cola o extremo superior de la distribución. Con esa notación, la siguiente aseveración de precisión define el tamaño del error muestral cuando se usa $\bar{x}$ para estimar μ.

Aseveración de precisión

Hay una probabilidad $1 - \alpha$ de que el valor de una media de muestra origine un error muestral de $z_{\alpha/2}\sigma_{\bar{x}}$ o menos.

Cálculo de un estimado de intervalo: caso de muestra grande conociendo σ

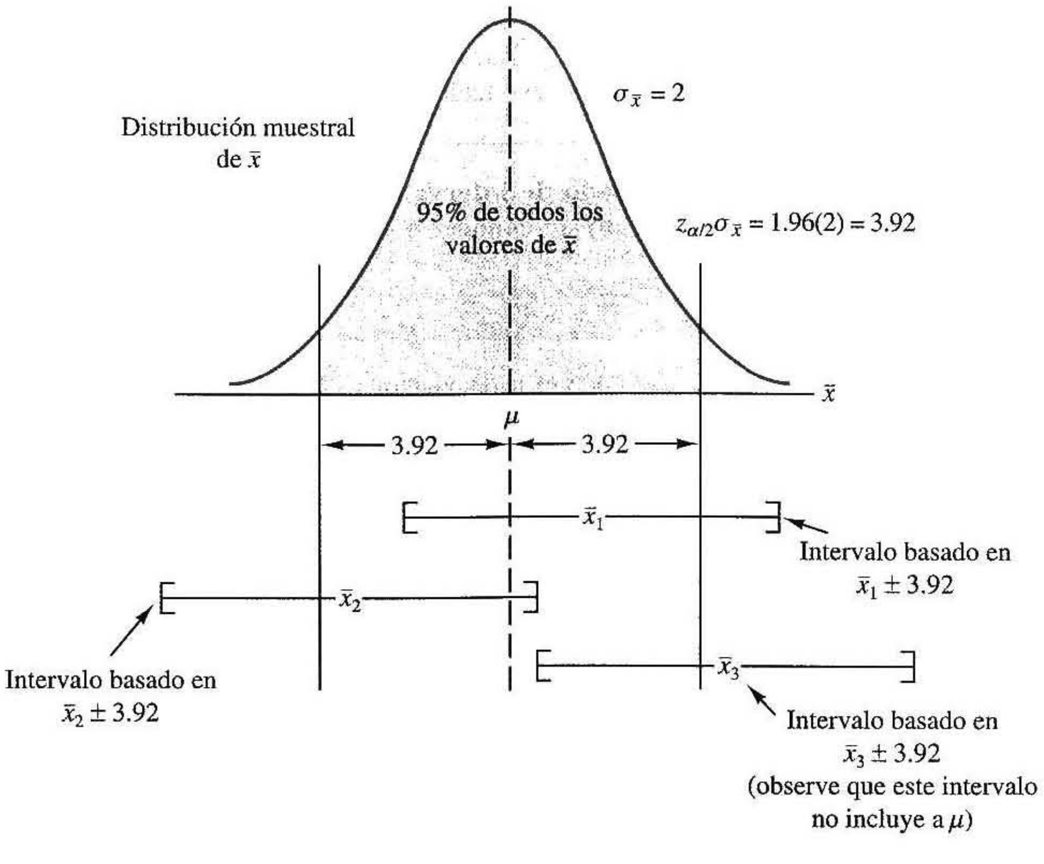
En el ejemplo de CJW dijimos que hay una probabilidad de .95 de que el valor de la media de muestra origine un error muestral de 3.92 o menos. Podemos definir un estimado de intervalo para μ restando 3.92 de $\bar{x}$ y sumando 3.92 más a $\bar{x}$; esto es, $\bar{x} \pm 3.92$. Para interpretar el estimado de intervalo de μ veamos los valores posibles de $\bar{x}$ que se podrían obtener a partir de tres muestras distintas, aleatorias simples, cada una formada por 100 clientes.

Supongamos que la primera media de muestra tiene el valor que vemos indicado en la figura 8.5 como $\bar{x}_1$. En este caso, vemos en la figura que el intervalo formado, al restar y sumar 3.92 a $\bar{x}_1$, abarca la media de población μ. Ahora veamos lo que sucede si resulta que la muestra tiene el valor que se identifica con $\bar{x}_2$ en la figura 8.5. Aunque esta media de muestra es distinta de la primera, vemos que el intervalo basado en $\bar{x}_2$ también incluye a la media de población μ. Sin embargo, el intervalo basado en la tercera media de muestra, identificada con $\bar{x}_3$, no incluye a la media de población; la razón es que $\bar{x}_3$ está en una cola de la distribución, a mayor distancia que 3.92 de μ. Por consiguiente, al restar y sumar 3.92 a $\bar{x}_3$ se define un intervalo que no incluye a μ.

En general, cualquier media de muestra $\bar{x}$ que está dentro de la región no sombreada de la figura 8.5 originará un intervalo que contiene a la media de población μ. Como el 95% de las medias de muestra posibles estará en esta región, el 95% de los intervalos definidos al sumar y restar 3.92 a $\bar{x}$ abarcarán a μ. En consecuencia, decimos que se tiene un 95% de confianza de que el intervalo definido como de $\bar{x} - 3.92$ a $\bar{x} + 3.92$ va a incluir la media de población. En la terminología estadística normal se dice que el intervalo es un *intervalo de confianza*. Como el 95% de las medias de muestra originarán un intervalo de confianza que abarca a la media μ de población, se dice que el intervalo de confianza se establece al *nivel de confianza* del 95%. El valor de .95 se llama *coeficiente de confianza*.

Un estimado de intervalo de confianza consta de dos partes: un estimado puntual y un valor $\pm$ que describe la precisión del estimado. Al valor $\pm$ se le llama *margen de error*. De esta forma, la media de muestra de la calificación de satisfacción en el ejemplo

Figura 8.5 INTERVALOS FORMADOS A PARTIR DE ALGUNAS MEDIAS DE MUESTRA EN LOS LUGARES $\bar{x}_1$, $\bar{x}_2$, y $\bar{x}_3$



El margen de error es el valor $\pm$ que determina la precisión de un estimado de punto. Al estimar un promedio de población, el margen de error es $z_{\alpha/2}\sigma_{\bar{x}}$.

de CJW tiene el margen de error 3.92. Usando el intervalo $\bar{x} \pm 3.92$, vemos que el 95% del estimado de intervalo de confianza de la media de población es 82 $\pm$3.92, o sea, de 78.08 a 85.92. Por consiguiente, con un nivel de confianza del 95%, CJW puede llegar a concluir que la calificación de satisfacción para todos sus clientes de pedidos por correo, está entre 78.08 y 85.92.

Ahora describiremos el procedimiento general para calcular un estimado de intervalo de una media de población para el caso de muestra grande, cuando se conoce σ. Como se dijo antes, hay una probabilidad de $1 - \alpha$ de que el valor de la media de la muestra origine un error muestral de $z_{\alpha/2}\sigma_{\bar{x}}$ o menos. Aprovechando que $\sigma_{\bar{x}} = \sigma/\sqrt{n}$, podemos escribir el procedimiento para calcular el *estimado de intervalo* de una media de población como sigue.

Estimado de intervalo de una media de población: caso de muestra grande ($n \geq 30$) con σ desconocida

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{8.2}$$

en donde $1 - \alpha$ es el coeficiente de confianza y $z_{\alpha/2}$ es el valor de z que origina un área de $\alpha/2$ en la cola o extremo superior de la distribución normal estándar de probabilidades.

En la tabla 8.1 vemos los valores de $z_{\alpha/2}$ de los niveles de confianza que se usan con más frecuencia.

Cálculo de un estimado de intervalo: caso de muestra grande sin conocer σ

Una dificultad para aplicar la ecuación (8.2) es que en la mayoría de los casos de muestreo se desconoce el valor de la desviación estándar poblacional σ. En el caso de muestra grande ($n \geq 30$) tan sólo se usa el valor de la desviación estándar de la muestra, s, como estimado puntual de σ para obtener el siguiente estimado de intervalo.

Estimado de intervalo de una media de población: caso de muestra grande ($n \geq 30$) sin conocer σ

$$\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}} \tag{8.3}$$

en donde s es la desviación estándar de la muestra, $1 - \alpha$ es el coeficiente de confianza y $z_{\alpha/2}$ es el valor de z que define un área de $\alpha/2$ en la cola o extremo superior de la distribución normal estándar de probabilidades.

Tabla 8.1 VALORES DE $z_{\alpha/2}$ PARA LOS NIVELES DE CONFIANZA DE USO MÁS COMÚN

Nivel de confianza	α	$\alpha/2$	$z_{\alpha/2}$
90%	.10	.05	1.645
95%	.05	.025	1.96
99%	.01	.005	2.576

Un ejemplo del procedimiento de estimación de intervalo es el estudio que llevó a cabo la Aseguradora Estatal. Supongamos que como parte de una revisión anual de las pólizas de seguros de vida, la aseguradora selecciona una muestra aleatoria simple de 36 asegurados. Se revisan las pólizas correspondientes en lo que se refiere a la cantidad asegurada, el valor de la póliza, opciones por invalidez, etc. Un gerente pidió, para ese estudio, un estimado de nivel de confianza de 90% para la edad promedio de la población de asegurados.

La tabla 8.2 muestra los datos por edad que se reunieron con una muestra aleatoria simple de 36 asegurados. La edad promedio obtenida, $\bar{x} = 39.5$ años, es el estimado puntual de la media de población de la edad. Además, la desviación estándar de la muestra de los datos en la tabla 8.2 es $s = 7.77$. Al nivel de confianza del 90%, $z_{.05} = 1.645$. Aplicando la ecuación (8.3) obtenemos

$$39.5 \pm 1.645 \frac{7.77}{\sqrt{36}}$$

$$39.5 \pm 2.13$$

Por lo tanto, el estimado de intervalo de la media de población, con 90% de confianza, es de 37.37 a 41.63 años. Así, el gerente puede tener el 90% de confianza en que la edad promedio de la población de los asegurados con esta empresa sea de 37.37 a 41.63 años.

Se puede usar el programa Minitab para obtener un estimado de intervalo de confianza para una media de población. Los resultados de cómputo para el problema de Aseguradora Estatal se muestran en la figura 8.6. La desviación estándar de la muestra es $s = 7.77$, el valor supuesto de la desviación estándar poblacional σ. Los demás resultados muestran que la muestra de 36 asegurados da 39.50 años como resultado una media de población, un error estándar del promedio de 1.30 años y un estimado de confianza de 90% desde 37.37 hasta 41.63 años.

Tabla 8.2 EDADES DE ASEGURADOS DE VIDA [illegible] MUESTRA ALEATORIA SIMPLE DE 36 ASEGURADOS [illegible]

LifeIns

Asegurado	Edad	Asegurado	Edad	Asegurado	Edad
1	32	13	39	25	23
2	50	14	46	26	36
3	40	15	45	27	42
4	24	16	39	28	34
5	33	17	38	29	39
6	44	18	45	30	34
7	45	19	27	31	35
8	48	20	43	32	42
9	44	21	54	33	53
10	47	22	36	34	28
11	31	23	34	35	49
12	36	24	48	36	39

Figura 8.6 INTERVALOS DE CONFIANZA DETERMINADOS CON MINITAB, PARA EL PROBLEMA DE MUESTREO DE ASEGURADORA ESTATAL

El supuesto sigma = 7.77

Variable	N	Media	Desviación estándar	Error estándar de la media	90.0 % CI
Edad	36	39.50	7.77	1.30	(37.37, 41.63)

NOTAS y comentarios

1. Al determinar un estimado de intervalo de la media de población se especifica el coeficiente deseado de confianza $(1-\alpha)$, antes de seleccionar la muestra. De esta forma, antes de seleccionarla, ya sabemos que hay una probabilidad de $1-\alpha$ de que el intervalo de confianza que vamos a calcular contendrá la media de población μ. Sin embargo, una vez tomada la muestra, se calcula la media muestral de $\bar{x}$ y se determina el estimado particular de intervalo; ese intervalo podrá o no podrá contener a μ. Si $1-\alpha$ es razonablemente grande, podemos confiar en que el intervalo resultante contiene a μ, porque sabemos que si aplicamos repetidamente el procedimiento, el $100(1-\alpha)$ por ciento de todos los intervalos posibles que se calculan con este método van a contener a μ.
2. Observe que el tamaño de muestra n aparece en el denominador de las ecuaciones (8.2) y la (8.3) de cálculo de intervalo. Así, si determinado tamaño de la muestra da como resultado un intervalo demasiado ancho para ser de utilidad, podremos considerar un aumento del tamaño de muestra. Con n en el denominador, un tamaño de la muestra mayor dará como resultado un intervalo más angosto y mayor precisión. El procedimiento para determinar el tamaño de una muestra aleatoria simple, necesaria para obtener determinada precisión, se describe en la sección 8.3.

EJERCICIOS

MÉTODOS

1. Una muestra aleatoria simple de 40 elementos dio como resultado una media de muestra de 25. La desviación estándar de la población es $\sigma = 5$.
   a. ¿Cuál es el error estándar de la media, $\sigma_{\bar{x}}$?
   b. Con 95% de probabilidades, ¿cuál es el margen de error?

AUTOEXAMEN

2. Una muestra aleatoria simple de 50 artículos originó una media de muestra 32 y una desviación estándar muestral 6.
   a. Determine un intervalo de confianza de 90% para la media de población.
   b. Determine un intervalo de confianza de 95% para la media de población.
   c. Determine un intervalo de confianza de 99% para la media de población.
3. Una muestra de 60 artículos tuvo una media de 80 y una desviación estándar de 15.
   a. Determine el intervalo de confianza de 95% para la media de población.
   b. Suponga que la media y la desviación estándar de la muestra se obtuvieron de una muestra de 120 artículos. Determine un intervalo de confianza de 95% para la media de población.
   c. ¿Cuál es el efecto de mayor tamaño de muestra sobre el estimado de intervalo de una media de población?
4. Se informa que el intervalo de confianza de 95% para una media de población es de 122 a 130. Si la media de la muestra es 126 y la desviación estándar de la muestra es 16.07, ¿qué tamaño de muestra se usó en la determinación?

APLICACIONES

AUTOEXAMEN

5. Para tratar de estimar la media de consumo por cliente, en un gran restaurante, se reunieron datos de una muestra de 49 clientes durante un periodo de 3 semanas.
   a. Suponga que la desviación estándar de población es de $2.50 dólares. ¿Cuál es el error estándar de la media?
   b. Con nivel de confianza de 95%, ¿cuál es el margen de error?
   c. Si la media de la muestra es de $22.60 dólares, ¿cuál es el intervalo de confianza de 95% para la media de la población?
6. Los ingresos semanales promedio de las personas que trabajan en varias industrias aparecieron en el *The New York Times 1988 Almanac*. Esos ingresos para quienes trabajan en los servicios fueron de $369 dólares. Suponga que este resultado se basó en una muestra de 250 personas dedicadas a los servicios, y que la desviación estándar de la muestra fue $50. Calcule el intervalo de confianza de 95% para la población de ingresos semanales de personas que trabajan en los servicios.
7. Una muestra de 532 suscriptores a *Business Week* mostró que el tiempo promedio que pasa un suscriptor en Internet y en servicios en línea es 6.7 horas semanales (*Business Week 1996 World Wide Subscriber Study*). Si la desviación estándar de la muestra es 5.8 horas, ¿Cuál es el intervalo de confianza de 95% de la población de tiempos promedio que pasan los suscriptores a *Business Week* en Internet y en servicios en línea?
8. En un estudio de préstamos a estudiantes, el Departamento de Educación informó que los beneficiarios del fondo Stafford Loan deberán un promedio de $12,168 al recibirse (*USA Today*, 5 de abril de 1995). Suponga que este promedio de deuda se basa en una muestra de 480 préstamos a estudiantes, y que la desviación estándar de la población de las deudas al recibirse es $2200.
   a. Determine un estimado de confianza de 90% del promedio poblacional de la deuda.
   b. Determine un estimado de confianza de 95% del promedio poblacional de la deuda.
   c. Determine un estimado de confianza de 99% del promedio poblacional de la deuda.
   d. Describa lo que sucede con el ancho del intervalo de confianza a medida que se aumenta el nivel de confianza. ¿Parece razonable? Explique su respuesta.

Workers

9. Una operación de llenado de envases tiene desviación estándar histórica de 5.5 onzas. Un inspector de control de calidad selecciona, periódicamente, 36 recipientes al azar, y emplea el peso de llenado promedio de la muestra para estimar el correspondiente a la población
   a. ¿Cuál es el error estándar del promedio, ?
   b. Con .75, .90 y .99 de probabilidad, ¿qué se puede afirmar acerca del error de muestreo? ¿Qué sucede a la declaración del error de muestreo cuando aumenta la probabilidad? ¿Por qué sucede así?
   c. ¿Cuál es el intervalo de confianza de 99% para el peso promedio de llenado de la población en el proceso, si el promedio muestral es 48.6 onzas?
10. Se determinó la rentabilidad de vender automóviles usados, en un estudio de la Asociación Nacional de Comerciantes de Automóviles (*USA Today*, 12 de abril de 1995). Suponga que con una muestra de 200 vendedores de coches usados se obtuvo una ganancia promedio de $300 y desviación estándar muestral $150. Con esta información defina un estimado de intervalo de confianza de 95% para la utilidad promedio de la población de ventas de automóviles usados.
11. La encuesta anual de calidad de automóviles, efectuada por J. D. Power & Associates, determinó que la cantidad promedio de defectos, en todas las marcas, por cada vehículo nuevo, es 1.07 (*The Wall Street Journal*, 27 de enero de 1994). Suponga que se toma una muestra de 30 automóviles nuevos de determinada marca y se obtienen las siguientes cantidades de defectos por vehículo.

0	1	1	2	1	0	2	3	2	1	0	4	3	1	1
0	2	0	0	2	3	0	2	0	2	0	3	1	0	2

   a. Con estos datos, ¿cuál es el promedio muestral de la cantidad de defectos por vehículo?
   b. ¿Cuál es la desviación estándar de la muestra?
   c. Determine un estimado de intervalo de confianza de 95% para la cantidad promedio de defectos por vehículo para la población de automóviles de esta marca.
   d. Después de revisar el estimado de confianza de la parte c), un analista estadístico sugirió que el fabricante revisara una mayor cantidad de automóviles nuevos antes de llegar a una conclusión al comparar la calidad de sus vehículos con el promedio general de J. D. Powers, de 1.07 defectos por vehículo. ¿Respalda usted esta idea? ¿Por qué?

12. La Asociación Internacional de Transporte Aéreo hace encuestas entre agentes viajeros para determinar calificaciones de calidad de los aeropuertos transatlánticos principales. La calificación máxima es 10. El aeropuerto con mayor calificación fue Amsterdam, con calificación promedio de 7.93, seguido por Toronto, con 7.17 (*Newsweek*, 13 de junio de 1994). Suponga que se toma una muestra aleatoria simple de 50 agentes viajeros y que a cada uno se le pide calificar al Aeropuerto Internacional de Miami. A continuación vemos las calificaciones obtenidas con la muestra de 50.

Miami

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

Determine un estimado del intervalo de confianza de 95% para la calificación promedio poblacional del aeropuerto de Miami.

8.2 ESTIMACIÓN DE INTERVALO DE UN PROMEDIO POBLACIONAL: CASO DE MUESTRA PEQUEÑA

En el caso de muestra pequeña ($n < 30$), la distribución de $\bar{x}$ depende de la distribución de probabilidades de la población. Si la población tiene una distribución normal de probabilidades, se puede usar la metodología que presentaremos en esta sección para establecer un intervalo de confianza de un promedio de población. Sin embargo, si no es adecuada la hipótesis de distribución normal de probabilidades para la población, la única alternativa es aumentar el tamaño muestral hasta $n \geq 30$, y confiar en los procedimientos de estimación de intervalo para muestra grande que usan las ecuaciones (8.2) y (8.3).

Si la población tiene distribución normal de probabilidades, la distribución de $\bar{x}$ será normal, independientemente del tamaño de la muestra. En este caso, si ya se conoce la desviación estándar poblacional σ, se puede usar la ecuación (8.2) para calcular un estimado de intervalo de promedio poblacional, aun con una muestra pequeña. Sin embargo, si se desconoce la desviación estándar poblacional, σ, se usa la desviación estándar s de la muestra para estimar σ, y el intervalo de confianza correspondiente se basa en una distribución de probabilidades que se llama *distribución t*.

Quien inventó la distribución t fue William Sealy Gosset, que escribía con el seudónimo "Student". Gosset se graduó en matemáticas en Oxford y trabajó en la Cervecería Guinness en Dublín, Irlanda. Desarrolló una nueva teoría estadística sobre muestras pequeñas al trabajar con muestras pequeñas y en experimentos donde intervenían temperaturas en esa cervecería.

La distribución t es una familia de distribuciones parecidas de probabilidades; una distribución t específica depende de un parámetro llamado *grados de libertad*. Esto es, hay una distribución t única con un grado de libertad, otra con dos grados de libertad, otra con tres grados de libertad, y así sucesivamente. A medida que aumenta la cantidad de grados de libertad, la diferencia entre la distribución t y la distribución normal estándar de probabilidades se hace más y más pequeña. En la figura 8.7 vemos distribuciones t con 10 y 20 grados de libertad, y su relación con la distribución normal estándar de probabilidades. Observe que una distribución t con más grados de libertad tiene menos dispersión y se parece cada vez más a la distribución normal estándar de probabilidades. También observe que el promedio de la distribución t es cero.

Usaremos un subíndice de t para indicar el área en la cola superior de la distribución t. Por ejemplo, así como usamos $z_{.025}$ para indicar el valor de z que determina un área de .025 en la cola superior de una distribución normal estándar de probabilidades, usaremos $t_{.025}$ para indicar un área de .025 en la cola superior de la distribución t. En general, usaremos la notación $t_{\alpha/2}$ para representar un valor de t que define un área de $\alpha/2$ de la cola superior de la distribución t. Vea la figura 8.8.

Al aumentar los grados de libertad, la distribución t tiende a la distribución normal. De hecho, los valores z de la distribución normal se pueden determinar en el renglón de la tabla de distribución t que corresponden a grados de libertad infinito.

La tabla 8.3 es de la distribución t. Observe, por ejemplo, que para una distribución t con 10 grados de libertad, $t_{.025} = 2.228$. De igual manera, para una distribución t con 20 grados de libertad, $t_{.025} = 2.086$. A medida que aumentan los grados de libertad, $t_{.025}$ tiende a $z_{.025} = 1.96$.

Ahora que tenemos idea de lo que es la distribución t, veamos cómo se usa para determinar un estimado del intervalo de un promedio poblacional. Suponga que la

Figura 8.7 COMPARACIÓN DE LA DISTRIBUCIÓN NORMAL ESTÁNDAR CON LAS DISTRIBUCIONES t CON 10 Y 20 GRADOS DE LIBERTAD

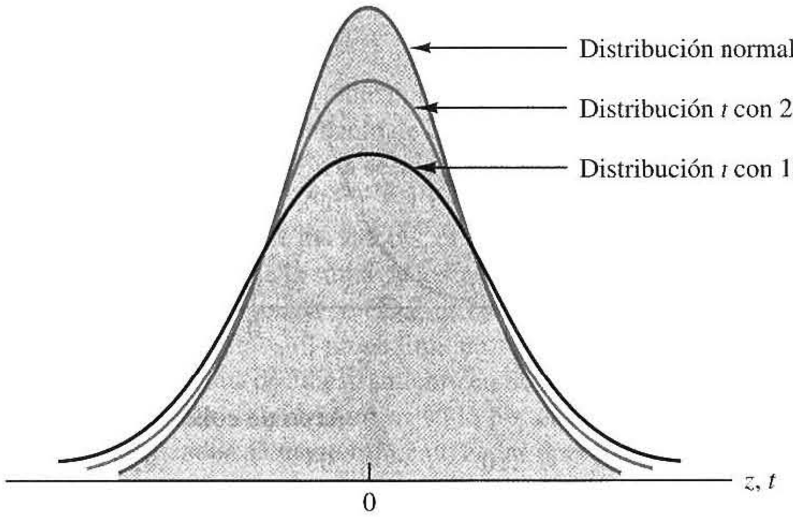
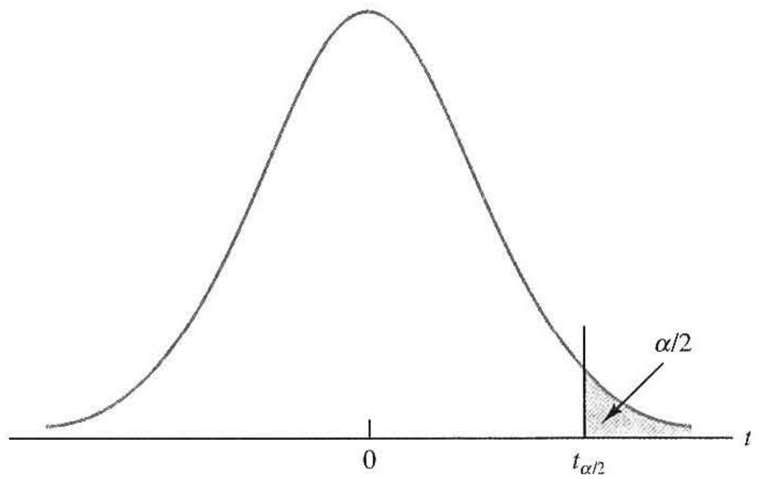


Figura 8.8 DISTRIBUCIÓN t CON EL ÁREA $\alpha/2$ DE PROBABILIDAD EN LA COLA SUPERIOR



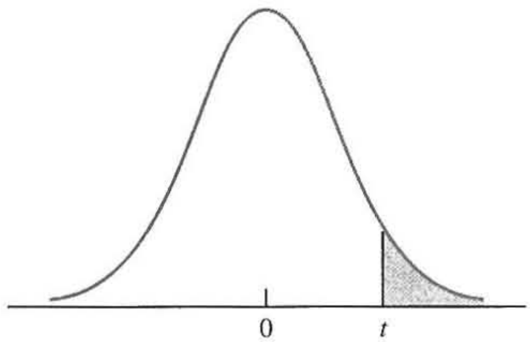
población tiene distribución normal de probabilidades, y que la desviación estándar de la muestra, s, se usa para estimar la desviación estándar de la población, σ. Se puede aplicar el siguiente procedimiento de estimación de intervalo (8.4).

Estimado de intervalo de un promedio de población: caso de muestra pequeña ($n < 30$) con σ desconocida

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} \tag{8.4}$$

en donde $1 - \alpha$ es el coeficiente de confianza, $t_{\alpha/2}$ es el valor de t que define un área de $\alpha/2$ en la cola o extremo superior de una distribución t con $n - 1$ grados de libertad, y s es la desviación estándar de la muestra. Se supone que la población tiene distribución normal estándar de probabilidades.

Tabla 8.3 TABLA DE LA DISTRIBUCIÓN t PARA ÁREAS DE COLA SUPERIOR. EJEMPLO: CON 10 GRADOS DE LIBERTAD, $t_{.025} = 2.228$



Grados de libertad	Área de cola superior (sombreada) .10	.05	.025	.01	.005
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
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

La razón de que la cantidad de grados de libertad asociada con el valor de t en la ecuación (8.4) sea $n - 1$ tiene que ver con el uso de s como estimado de la desviación estándar poblacional σ. La ecuación de la desviación estándar muestral es

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

Los grados de libertad indican la cantidad de elementos independientes de información que intervienen en el cálculo de $\Sigma(x_i - \bar{x})^2$. Esos n elementos son los siguientes: $x_1 - \bar{x}$, $x_2 - \bar{x}, \ldots, x_n - \bar{x}$. En la sección 3.2 indicamos que $\Sigma(x_i - \bar{x}) = 0$ para cualquier conjunto de datos. Así, sólo $n - 1$ de los valores de $x_i - \bar{x}$ son independientes; esto es, si conocemos $n - 1$ de los valores, el valor que resta se puede determinar con exactitud por la condición de que la suma de los $x_i - \bar{x}$ debe ser cero. Así, $n - 1$ es la cantidad de grados de libertad asociada con $\Sigma(x_i - \bar{x})^2$ y a esto se debe la distribución t que se usó en (8.4).

Demostraremos el procedimiento de estimación de intervalo con muestra pequeña con un programa de adiestramiento en Scheer Industries. Su director de manufactura desea contar con un programa asistido por computadora que se pueda usar para adiestrar a los empleados de mantenimiento en las operaciones de reparación de máquinas. Se espera que con el método de cómputo se reduzca el tiempo necesario para el adiestramiento. Para evaluarlo, el director de manufactura pide un estimado de tiempo promedio de adiestramiento necesario con el programa asistido por computadora.

Supongamos que la gerencia aceptó adiestrar a 15 empleados con el nuevo método. En la tabla 8.4 vemos los datos de días necesarios de adiestramiento por cada empleado de la muestra. El promedio y la desviación estándar muestrales, para esos datos, son los siguientes.

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{808}{15} = 53.87 \text{ días}$$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{651.73}{14}} = 6.82 \text{ días}$$

El estimado puntual del tiempo promedio de adiestramiento, para la población de empleados, es 53.87 días. Podemos obtener información acerca de la precisión de este estimado determinando un estimado de intervalo del promedio poblacional. Como se desconoce la desviación estándar poblacional, usaremos la desviación estándar muestral, $s = 6.82$ días, como estimado de punto de s. Con el tamaño pequeño de muestra, $n = 15$, emplearemos la ecuación (8.4) para determinar un estimado de intervalo del promedio poblacional con 95% de confianza. Si suponemos que la población de tiempos de adiestramiento tiene una distribución normal de probabilidades, la distribución t con $n - 1 = 14$ grados de libertad, es la adecuada para el procedimiento de estimación

Tabla 8.4 TIEMPO DE ADIESTRAMIENTO EN DÍAS PARA EL PROGRAMA ASISTIDO POR COMPUTADORA EN SCHEER INDUSTRIES

Empleado	Tiempo	Empleado	Tiempo	Empleado	Tiempo
1	52	6	59	11	54
2	44	7	50	12	58
3	55	8	54	13	60
4	44	9	62	14	62
5	45	10	46	15	63

La distribución t se basa en la hipótesis de que la población tiene distribución normal de probabilidades. Sin embargo, los intervalos de confianza que se basan en la distribución t se pueden usar, mientras que la distribución de la población no sea muy distinta de la distribución normal de probabilidades.

de intervalo. Vemos, en la tabla 8.3, que con 14 grados de libertad, $t_{\alpha/2} = t_{.025} = 2.145$. Sustituyendo en la ecuación (8.4),

$$\bar{x} \pm t_{.025} \frac{s}{\sqrt{n}}$$

$$53.87 \pm 2.145\left(\frac{6.82}{\sqrt{15}}\right)$$

$$53.87 \pm 3.78$$

Así, el estimado de intervalo de confianza de 95% para el promedio poblacional de tiempo de adiestramiento es de 50.09 a 57.65 días.

Con el programa Minitab se puede determinar el intervalo de confianza del promedio poblacional en el ejemplo de Industrias Scheer. La figura 8.9 muestra los resultados obtenidos. La muestra de 15 empleados produce un promedio de 53.87 días, desviación estándar muestral dc 6.82 días, error estándar de la media igual a 1.76 días e intervalo de confianza de 95% de 50.09 a 57.65 días.

Figura 8.9 INTERVALOS DE CONFIANZA GENERADOS CON MINITAB PARA EL PROBLEMA DE MUESTREO DE SCHEER INDUSTRIES

Variable	N	Media	Desviación estándar	Error estándar de la media	95.0 % CI
Tiempo	15	53.87	6.82	1.76	(50.09, 57.65)

NOTAS y comentarios

La distribución t no se restringe al caso de la muestra pequeña. En realidad se puede aplicar siempre que la población sea normal o casi normal, y siempre que se use la desviación estándar muestral para estimar la desviación estándar de población. Si se cumplen estas condiciones, se puede usar la distribución t para cualquier tamaño de muestra. Sin embargo, la ecuación (8.3) indica que con una muestra grande ($n \geq 30$), la estimación de intervalo del promedio poblacional se puede basar en la distribución normal estándar de probabilidades, y en el valor $z_{\alpha/2}$. Así, teniendo disponible la ecuación (8.3) para el caso de muestra grande, por lo general no se usa la distribución t sino cuando se tiene un caso de muestra pequeña.

EJERCICIOS

MÉTODOS

13. Para una distribución t con 12 grados de libertad, determine el área, o probabilidad, que hay en cada región.

- **a.** A la izquierda de 1.782.
- **b.** A la derecha de −1.356.
- **c.** A la derecha de 2.681.
- **d.** A la izquierda de −1.782.
- **e.** Entre −2.179 y +2.179.
- **f.** Entre −1.356 y +1.782.

14. Determine el o los valores t en cada uno de los siguientes ejemplos.

- **a.** Área de la cola superior de .05 con 18 grados de libertad.
- **b.** Área de la cola inferior de .10 con 22 grados de libertad.
- **c.** Área de la cola superior de .01 con 5 grados de libertad.
- **d.** El 90% del área está entre estos dos valores de t con 14 grados de libertad.
- **e.** El 95% del área está entre estos dos valores de t con 28 grados de libertad.

AUTOEXAMEN

15. Los siguientes datos se reunieron con una muestra de ocho artículos proveniente de una población normal: 10, 8, 12, 15, 13, 11, 6, 5.

a. ¿Cuál es el estimado de punto del promedio poblacional?

b. ¿Cuál es el estimado de punto de la desviación estándar poblacional?

c. ¿Cuál es el intervalo de confianza de 95% para el promedio de población?

16. Una muestra aleatoria simple de 20 elementos, procedente de una población normal, originó un promedio muestral de 17.25 y una desviación estándar muestral de 3.3.

a. Determine un intervalo de confianza del 90% para el promedio de población.

b. Determine un intervalo de confianza del 95% para el promedio de población.

c. Determine un intervalo de confianza del 99% para el promedio de población.

APLICACIONES

AUTOEXAMEN

17. Al ensayar un nuevo método de producción, se seleccionaron 18 empleados al azar, y se les pidió lo probaran. La tasa de producción promedio muestral para los 18 empleados fue 80 partes por hora, y la desviación estándar muestral fue 10 partes por hora. Determine intervalos de confianza del 90 y 95% de la tasa de producción promedio poblacional con el nuevo método, suponiendo que la población tiene una distribución normal de probabilidades.

18. La sección Dinero e Inversión, de *The Wall Street Journal*, contiene un resumen diario de las inversiones en la Bolsa de Nueva York, la American Stock Exchange, el extranjero, opciones, bienes, futuros, etcétera. En la sección de la Bolsa de Nueva York, aparece información de precio máximo por acción durante 52 semanas, y el precio mínimo en ese lapso, dividendos, rendimiento, relación P/E, volumen diario, precios máximo y mínimo diarios por acción, precio al cierre por acción y cambio neto diario. La relación P/E (precio a rendimiento) para cada acción se calcula dividiendo el precio por acción entre las ganancias por acción que informa la empresa durante los cuatro últimos trimestres. Se tomó una muestra de 10 acciones en *The Wall Street Journal* (19 de marzo de 1998) y se obtuvieron los datos siguientes de relación P/E: 5, 7, 9, 10, 14, 23, 20, 15, 3, 26.

a. ¿Cuál es el estimado de punto de la relación P/E promedio para la población de todas las acciones de la lista de la Bolsa de Nueva York?

b. ¿Cuál es el estimado de punto de la desviación estándar de las relaciones P/E para la población de todas esas acciones?

c. Con un coeficiente de confianza de .95, ¿cuál es el estimado de intervalo para la relación P/E promedio para la población de esas acciones? Suponga que la población tiene distribución normal.

d. Comente la precisión de los resultados.

19. La Asociación Americana de Agencias de Publicidad tiene un registro de datos sobre minutos de anuncios por cada media hora de programas principales de TV. En la tabla siguiente vemos una lista de datos representativos de una muestra de programas preferentes en cadenas principales a las 8:30 P.M.

TVtime

6.0	6.6	5.8
7.0	6.3	6.2
7.2	5.7	6.4
7.0	6.5	6.2
6.0	6.5	7.2
7.3	7.6	6.8
6.0	6.2	

Determine un estimado de punto y un intervalo de confianza de 95% para la cantidad promedio de minutos de anuncios en los principales espectáculos televisivos a las 8:30 P.M.

20. Se pidió al personal de ventas de Distribuidora González que presentara informes semanales con los clientes llamados durante la semana. En una muestra de 61 informes semanales se determinó un promedio de 22.4 llamadas a clientes por semana, y que la desviación estándar era 5 llamadas.

a. Suponga el caso de muestra grande, ecuación (8.3) para determinar un intervalo de confianza de 95% para la cantidad promedio de llamadas semanales a clientes para la población del personal de ventas.
b. Suponga que esa población tiene distribución normal. Aplique la distribución *t*, con 60 grados de libertad, para determinar un intervalo de confianza de 95% para la cantidad promedio de llamadas semanales a clientes.
c. Compare sus respuestas en los incisos *a* y *b*. Comente la razón por la que en el caso de muestra grande se pueden basar los estimados de intervalo en el procedimiento seguido en el inciso *a*, aun cuando también se pueda aplicar la distribución *t*.

21. En Estados Unidos, el Departamento de Transporte informó las millas que viajan diariamente los residentes de áreas metropolitanas en automóvil (*1994 Information Please Environmental Almanac*). Suponga que con una muestra aleatoria de 15 residentes de Cleveland se obtuvieron los siguientes datos de millas diarias en automóvil.

20	20	28	16	11	17	23	16	22	18	10	22	29	19	32

a. Calcule un estimado de intervalo de confianza de 95% para la cantidad promedio de millas diarias de la población de residentes de Cleveland.
b. ¿Qué hipótesis acerca de la población fue necesaria para llegar a la respuesta en el inciso *a*?
c. Suponga que deseamos estimar las millas recorridas diariamente por la población, con precisión de ± 2 millas con 95% de confianza. ¿Permiten los datos este nivel de precisión? ¿Qué acción, si es el caso, recomendaría usted tomar?

22. La cantidad de horas que duermen los estadounidenses cada noche varía mucho, desde el 12% de la población que duerme menos de 6 horas hasta el 3% que duerme más de 8 horas (*The Macmillan Visual Almanac,* 1996). A continuación vemos una muestra de las horas que duermen cada noche 25 personas.

Sleep

6.9	7.6	6.5	6.2	5.3
7.8	7.0	5.5	7.6	6.7
7.3	6.6	7.1	6.9	6.0
6.8	6.5	7.2	5.8	8.6
7.6	7.1	6.0	7.2	7.7

a. ¿Cuál es el estimado puntual de la media de población de la cantidad de horas que se duerme cada noche?
b. Suponiendo que la población tiene distribución normal, determine un intervalo de confianza de 95% para la cantidad de la media de población de horas de sueño cada noche.

8.3 DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA

En la sección 8.1 pudimos afirmar lo siguiente acerca del error muestral, cuando se usa un promedio muestral para obtener un estimado puntual de la media de población:

> Hay una probabilidad $1 - \alpha$ de que el valor de la media de muestra origine un error muestral igual a $z_{\alpha/2}\sigma_{\bar{x}}$ o menos.

Como $\sigma_{\bar{x}} = \sigma/\sqrt{n}$, este enunciado se puede expresar como sigue:

> Hay una probabilidad $1 - \alpha$ de que el valor de la media de la muestra origine un error muestral igual a $z_{\alpha/2}(\sigma/\sqrt{n})$ o menos.

La cantidad $z_{\alpha/2}\ (\sigma/\sqrt{n})$ se llama margen de error. Según este enunciado vemos que los valores de $z_{\alpha/2}$, σ, y el tamaño muestral n se combinan para determinar el error de muestreo. Una vez seleccionado un coeficiente de confianza o probabilidad de $1 - \alpha$, se puede determinar $z_{\alpha/2}$. Dados los valores de $z_{\alpha/2}$ y σ, podemos determinar el tamaño de la muestra, n, necesario para originar cualquier error muestral. A continuación presentamos la deducción de la fórmula para determinar el tamaño requerido de muestra, n.

Sea E = el error máximo de muestreo

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Si se ha seleccionado un margen deseado de error antes de muestrear, se pueden aplicar los procedimientos de esta sección para determinar el tamaño de muestra necesario para satisfacer el requisito de margen de error.

Despejamos a $\sqrt{n}$, y obtenemos

$$\sqrt{n} = \frac{z_{\alpha/2}\sigma}{E}$$

Elevamos al cuadrado ambos lados de la ecuación y llegamos a la siguiente fórmula del tamaño de la muestra.

Tamaño de la muestra para un estimado de intervalo de una media de población

$$n = \frac{(z_{\alpha/2})^2\sigma^2}{E^2} \tag{8.5}$$

Con este tamaño de muestra se obtiene el margen deseado de error con el nivel de confianza elegido.

En la ecuación (8.5), el valor de E es el margen de error que se puede aceptar al nivel de confianza dado, y el valor de $z_{\alpha/2}$ es consecuencia directa del nivel de confianza que se usa para determinar el estimado de intervalo. Aunque se debe tomar en cuenta la preferencia del usuario, lo que se escoge con más frecuencia es el 95% de confianza, $z_{.025} = 1.96$.

Por último, para aplicar la ecuación (8.5) se requiere un valor de la desviación estándar σ de la población. En la mayoría de los casos se desconoce σ. Sin embargo, podemos aplicar esa ecuación si contamos con un *valor preliminar*, o valor de planeación, de σ. En la práctica se puede optar por uno de los siguientes procedimientos:

Se debe conocer un valor de planeación de la desviación estándar poblacional σ para poder determinar el tamaño de la muestra. Aquí se describen varios métodos para obtener ese valor de planeación.

1. Usar la desviación estándar muestral de una muestra previa de las mismas unidades, o de otras parecidas.
2. Usar un estudio piloto para seleccionar una muestra preliminar de unidades. La desviación estándar muestral de ella se puede usar como el valor de planeación de σ.
3. Usar el juicio o una "mejor estimación" del valor de σ. Por ejemplo, podríamos comenzar estimando los valores máximo y mínimo de los datos en la población. La diferencia entre ellos proporciona un estimado del rango de los datos. Por último, se sugiere tomar el rango dividido entre cuatro como una aproximación de la desviación estándar para contar con un valor de planeación aceptable para σ.

Regresemos al ejemplo de Scheer Industries de la sección 8.2 para ver cómo se usa la ecuación (8.5) en la determinación del tamaño de la muestra para un estudio. Antes habíamos demostrado que, con un nivel de confianza de 95%, una muestra de 15 empleados de Scheer generaba un estimado de la media de población del tiempo de adiestramiento de 53.87 ± 3.78 días. Suponga que, después de conocer esos resultados, el director de manufactura de Scheer no queda satisfecho con el grado de precisión, al sentir que un error muestral de ± 3.78 días es demasiado. Además, supongamos que el director hace la siguiente aseveración acerca de la precisión que desea: "me gustaría un intervalo de confianza de 95% de que el error muestral sea de dos días o menos". Podemos ver que el director especifica que $E = 2$ días. Con un nivel de confianza de 95%, $z_{\alpha/2} = z_{.025} = 1.96$. Necesitamos un valor de planeación de σ para usarlo en la ecuación (8.5), para determinar el tamaño de la muestra. ¿Contamos con un valor de planeación de σ en este ejemplo? Aunque se desconoce σ, aprovecharemos la ventaja de los datos recopilados para los 15 empleados, que vimos en la sección 8.2. Podemos considerar que estos datos fueron obtenidos a partir de un estudio piloto, y que la desviación estándar de la muestra $s = 6.82$ días es el valor de planeación de σ. Así, al aplicar la ecuación (8.5), tenemos

La ecuación (8.5) determina el tamaño de la muestra *mínimo* que satisface el requisito del margen de error. Si el tamaño de la muestra calculado es fraccionario, el redondeo al siguiente entero determinará el tamaño de la muestra recomendado. Con frecuencia, los estadísticos recomiendan un tamaño todavía mayor, para aumentar la probabilidad de que los resultados muestrales satisfagan el requisito del margen de error.

$$n = \frac{(z_{\alpha/2})^2\sigma^2}{E^2} = \frac{(1.96)^2(6.82)^2}{2^2} = 44.67$$

Esta ecuación nos dice que el tamaño mínimo de la muestra debe ser 44.67 empleados, para satisfacer el requisito del margen de error. En los casos en que n calculado sea una fracción, se redondea hacia arriba, hasta el siguiente valor entero; en consecuencia, el tamaño de la muestra recomendado para el ejemplo de Scheer Industries es de 45 empleados.

Por último, observe que en ese ejemplo se usó $z_{.025}$ para determinar el tamaño de la muestra, aunque los cálculos originales, con 15 empleados, hayan empleado la distribución t. La razón por la que usamos $z_{.025}$ es que, como todavía falta por determinar el tamaño de la muestra, estamos anticipando que n será mayor de 30, lo que hace que $z_{.025}$ sea el valor adecuado. Además, si todavía se debe determinar n, todavía no conocemos los $(n - 1)$ grados de libertad necesarios para usar la distribución t. Por consiguiente, el uso de la ecuación (8.5) para determinar el tamaño de muestra siempre se basará en un valor de z y no en uno de t.

EJERCICIOS

MÉTODOS

23. ¿De qué tamaño debe ser una muestra para poder tener el 95% de confianza en que el error muestral es de 5 o menor? Suponga que la desviación estándar de la población es de 25.

AUTOEXAMEN

24. Se estima que el rango, para un conjunto de datos, es 36.

a. ¿Cuál es el valor de planeación para la desviación estándar de la población?

b. ¿De qué tamaño es la muestra que se debe tomar para tener el 95% de confianza de que el error muestral sea de 3 o menor?

c. ¿De qué tamaño es la muestra que se debe tomar para tener el 95% de confianza de que el error muestral sea de 2 o menor?

APLICACIONES

AUTOEXAMEN

25. Acerca del ejemplo de Scheer Industries de esta sección. Emplee $\sigma = 6.82$ días como valor de planeación para la desviación estándar poblacional.

a. Suponiendo una confianza del 95%, ¿de qué tamaño se requeriría la muestra para un margen de error de 1.5 minutos?

b. Si se determinó la precisión con el 90% de confianza, ¿qué tamaño de muestra se requerirá para obtener un margen de error de 2 minutos?

26. En la sección 8.1, la Aseguradora Estatal usó una muestra aleatoria simple de 36 asegurados para estimar la media de la población de las edades de los asegurados. Con una probabilidad de 95%, el margen de error fue de 2.35 años. Este resultado se basó en una desviación estándar muestral de 7.2 años.

a. ¿Qué tamaño debería tener una muestra aleatoria simple para reducir el margen de error a dos años? ¿A 1.5 años? ¿A un año?

b. ¿Recomendaría usted que la aseguradora trate de estimar la media de la población de las edades de sus asegurados con $E = 1$ año? Explique su respuesta.

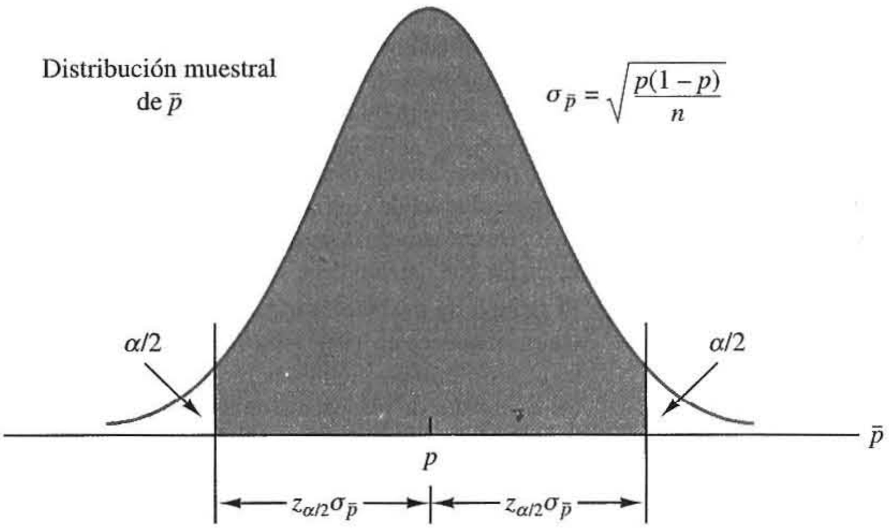
27. Se cree que los sueldos anuales iniciales de egresados de licenciatura en administración de empresas pueden tener una desviación estándar aproximada de \$2000 dólares. Suponga que se desea un estimado de intervalo de 95% de nivel de confianza para la media del sueldo anual inicial. ¿De qué tamaño debe tomarse la muestra, si el margen de error es
   a. \$500? **b.** \$200? **c.** \$100?
28. El Departamento de Vivienda y Desarrollo Urbano de Estados Unidos publica datos acerca del alquiler mensual de viviendas con una recámara en áreas metropolitanas (*The Federal Register*, 30 de abril de 1997). La desviación estándar de la renta mensual es, aproximadamente, de \$80 dólares. Suponga que se debe seleccionar una muestra de áreas metropolitanas para estimar la media de la población de renta mensual de viviendas con una recámara. Emplee el nivel de confianza de 95%.
   a. ¿De qué tamaño debe ser la muestra para que el margen de error deseado sea de \$25 dólares?
   b. ¿De qué tamaño debe ser para que sea de \$15 dólares?
29. El tiempo de traslado al trabajo, para residentes de las 15 ciudades mayores de Estados Unidos, aparece en 1998 *Information Please Almanac*. Suponga que se emplea una muestra aleatoria simple preliminar de residentes de San Francisco y se determina que 6.25 minutos es el valor de planeación de la desviación estándar poblacional.
   a. Si se desea estimar la media de la población del tiempo de traslado para los residentes de San Francisco, con 2 minutos de margen de error, ¿qué tamaño de la muestra se debe usar? Suponga una confianza de 95%.
   b. Si se desea que el margen de error sea de 1 minuto, ¿qué tamaño de la muestra se debe usar? Suponga una confianza de 95%.
30. De acuerdo con el ejercicio 18, la desviación estándar de la muestra de las relaciones P/G para acciones en la Bolsa de Valores de Nueva York es $s = 7.8$ (*The Wall Street Journal*, 19 de marzo de 1998). Suponga que nos interesa estimar la media de la relación P/G de las acciones de la Bolsa de Valores de Nueva York. ¿Cuántas acciones deben incluirse en la muestra, si deseamos que el margen de error sea 2? Suponga una confianza de 95%.

8.4 ESTIMACIÓN DEL INTERVALO DE UNA PROPORCIÓN DE LA POBLACIÓN

En la sección 8.2 presentamos el ejemplo de Scheer Industries, implicando la estimación de la media del tiempo de capacitación de empleados para un programa nuevo de adiestramiento en reparación de máquinas. Para evaluar el programa desde una perspectiva diferente, la gerencia ha pedido determinar alguna medida de la calidad del programa. El grado de éxito se había medido antes con las calificaciones de los empleados en una prueba, al final del adiestramiento. De acuerdo con la experiencia, la empresa sabe que una persona que aprueba el examen tiene una alta posibilidad de desempeñarse bien en su puesto. Después de algunas discusiones, la administración convino en basar la evaluación de calidad del nuevo método de adiestramiento sobre la proporción de empleados que aprobaron el examen. Supongamos que Scheer ha implantado la recomendación del tamaño muestral que recomendábamos en la sección anterior. Así, tenemos ahora una muestra de 45 empleados que podemos usar para establecer un estimado de intervalo para la proporción de población que aprueba el examen.

En el capítulo 7 demostramos que una proporción muestral $\bar{p}$ es un estimador insesgado de una proporción poblacional p, y que para muestras grandes, la distribución muestral de probabilidades de $\bar{p}$ se puede aproximar con una distribución normal de probabilidades, como se ve en la figura 8.10. Recuérdese que el empleo de la distribución normal como aproximación de la distribución muestral de $\bar{p}$ se basa en la condición de que tanto np como $n(1 - p)$ valen 5 o más. Usaremos la distribución muestral de $\bar{p}$ para hacer aseveraciones probabilísticas acerca del error muestral, siempre que usemos una proporción muestral $\bar{p}$ para estimar una proporción poblacional p. En este caso, se define el error de muestreo como el valor absoluto de la diferencia entre $\bar{p}$ y p, y se representa con $|\bar{p} - p|$.

Figura 8.10 APROXIMACIÓN NORMAL DE LA DISTRIBUCIÓN MUESTRAL DE $\bar{p}$ CUANDO $np \geq 5$ y $n(1 - p) \geq 5$



Las aseveraciones probabilísticas que se pueden hacer acerca del error muestral para la proporción tienen la siguiente forma.

> Hay una probabilidad de $1 - \alpha$ de que el valor de la proporción poblacional origine un error muestral igual a $z_{\alpha/2}\sigma_{\bar{p}}$ o menos.

Por consiguiente, la cantidad $z_{\alpha/2}\sigma_{\bar{p}}$ es el margen de error.

El razonamiento en el que se basa esta afirmación es el mismo que describimos cuando se usó el valor de una media de muestra como estimado de una media de población, y éste es el siguiente: como ya sabemos que la distribución muestral de $\bar{p}$ se puede aproximar con una distribución normal de probabilidades, podemos usar el valor de $z_{\alpha/2}$ y el del error estándar de la población, $\sigma_{\bar{p}}$, para formar la aseveración con respecto al margen de error.

Una vez que veamos que la aseveración probabilística acerca del error muestral se basa en $z_{\alpha/2}\sigma_{\bar{p}}$, podemos restar y sumar ese valor a $\bar{p}$ para obtener un estimado de intervalo de la proporción poblacional. Ese estimado de intervalo es

$$\bar{p} \pm z_{\alpha/2}\sigma_{\bar{p}} \tag{8.6}$$

en donde $1 - \alpha$ es el coeficiente de confianza. Como $\sigma_{\bar{p}} = \sqrt{p(1 - p)/n}$, la ecuación (8.6) se puede escribir en esta forma:

$$\bar{p} \pm z_{\alpha/2}\sqrt{\frac{p(1 - p)}{n}} \tag{8.7}$$

Para usar la ecuación (8.7) en la determinación de un estimado de intervalo de una proporción poblacional p, deber conocerse el valor de p. Como ese valor es lo que se trata de estimar, tan sólo sustituimos la proporción $\bar{p}$ de población con p. La ecuación

Al establecer intervalos de confianza para proporciones, a $\sigma_{\bar{p}}$ se le llama error estándar de la proporción, y la cantidad $z_{\alpha/2}\sqrt{\bar{p}(1-\bar{p})/n}$ determina el margen de error.

general que resulta, para un estimado del intervalo de confianza de una proporción poblacional es la siguiente.*

Estimado de intervalo de una proporción de la población

$$\bar{p} \pm z_{\alpha/2}\sqrt{\bar{p}(1-\bar{p})/n} \tag{8.8}$$

en donde $1-\alpha$ es el coeficiente de confianza y $z_{\alpha/2}$ es el valor de z que origina un área de $\alpha/2$ en la cola o extremo superior de la distribución normal estándar de probabilidades.

Regresemos al ejemplo de Scheer Industries. Supongamos que en el ejemplo de 45 empleados que terminaron el programa de adiestramiento, 36 aprobaron el examen. Así, el estimado puntual de la proporción poblacional de quienes aprueban el examen es $\bar{p} = 36/45 = .80$. Usando la ecuación (8.8) y un coeficiente de confianza de .95, vemos que el estimado del intervalo para la proporción poblacional es

$$\bar{p} \pm z_{.025}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$$.80 \pm 1.96\sqrt{\frac{.80(1-.80)}{45}}$$

$$.80 \pm .12$$

Así, a un nivel de confianza de 95%, el estimado del intervalo de la proporción de población es de .68 a .92.

Determinación del tamaño de la muestra

Ahora veremos el asunto del tamaño que debe tener la muestra para obtener un estimado de una proporción poblacional con determinado nivel de precisión. Los argumentos de la determinación del tamaño de la muestra para establecer estimados de intervalo de p se parecen mucho a los que usamos en la sección 8.3 para determinar el tamaño de la muestra con el cual estimar una media de la población.

Anteriormente, en esta sección, hicimos notar que el margen de error asociado con un estimado de proporción poblacional es $z_{\alpha/2}\sigma_{\bar{p}}$. Con $\sigma_{\bar{p}} = \sqrt{p(1-p)/n}$, el margen de error se basa en los valores de $z_{\alpha/2}$, de la proporción poblacional p, del tamaño de la muestra n. Para determinado coeficiente de confianza $1-\alpha$, se puede determinar $z_{\alpha/2}$. Entonces, como el valor de la proporción de población es fijo, el margen de error está determinado por el tamaño de la muestra n. Los tamaños mayores de muestra proporcionan una mejor precisión.

Sea $E =$ el margen de error deseado

$$E = z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

Al despejar a n de la ecuación se obtiene la siguiente ecuación para el tamaño de muestra.

*Un estimado insesgado de $\sigma_{\bar{p}}^2$ es $\bar{p}(1-\bar{p})/(n-1)$, lo cual sugiere que $\sqrt{\bar{p}(1-\bar{p})/(n-1)}$ se debe usar en lugar de $\sqrt{\bar{p}(1-\bar{p})/n}$ en la ecuación (8.8). Sin embargo, el prejuicio o sesgo introducido al usar n en el denominador no causa dificultad alguna, porque por lo general las muestras grandes se usan para hacer estimados acerca de las proporciones poblacionales. En esos casos, la diferencia numérica entre los resultados obtenidos al usar n y los obtenidos al usar $n-1$ es despreciable.

Tamaño de la muestra para un estimado de intervalo de una proporción poblacional

$$n = \frac{(z_{\alpha/2})^2 p(1 - p)}{E^2} \tag{8.9}$$

Como la proporción poblacional p es lo que se trata de estimar a partir de la muestra, debe poder tenerse una aproximación preliminar de p para poder usar la ecuación (8.9). Aquí se describen cuatro formas de obtener esta aproximación.

En esta ecuación, el usuario debe especificar el valor del error muestral E; en la mayoría de los casos, E es .10 o menor. También, la preferencia del usuario especifica el nivel de confianza, y con él el valor correspondiente de $z_{\alpha/2}$. Por último, el empleo de esta ecuación requiere un valor de planeación para la proporción p de la población. En la práctica, este valor de planeación se puede elegir mediante uno de los procedimientos siguientes:

1. Usar la proporción muestral de una muestra anterior de las mismas unidades.
2. Llevar a cabo un estudio piloto para seleccionar una muestra preliminar de unidades. La proporción muestral a partir de esta muestra se puede usar como valor de planeación para p.
3. Usar el juicio o un "estimado mejor" del valor de p.
4. Si no se aplica alguna de las alternativas anteriores, usar $p = .50$.

Regresemos al ejemplo de Scheer Industries, donde queríamos estimar la proporción de empleados que pasan el examen del programa de adiestramiento. ¿De qué tamaño debe ser una muestra de empleados, si el director de manufactura de Scheer desea estimar la proporción de población con un margen de error igual a .10 o menos, con un nivel de confianza de 95%? Con $E = .10$ y $z_{.025} = 1.96$, necesitamos un valor de planeación para p, para poder determinar el tamaño de la muestra. Antes, en esta sección, dijimos que 36 de los 45 empleados que presentaron el examen, lo aprobaron. En consecuencia, $\bar{p} = 36/45 = .80$ se puede usar como valor de planeación de p. Al aplicar la ecuación (8.9), obtenemos

$$n = \left[\frac{(1.96)^2 .80(1 - .80)}{(.10)^2}\right] = 61.47$$

Por consiguiente, el tamaño de muestra mínimo debe ser de 61.47 empleados, para llenar el requisito del margen de error. Al redondear al siguiente entero se ve que lo recomendado es un tamaño de la muestra de 62 empleados.

La cuarta alternativa sugerida para seleccionar un valor de planeación es usar $p = .50$. Este valor de p se usa con frecuencia cuando no se dispone de información. Para comprender por qué, observemos que el numerador de la ecuación (8.9) indica que el tamaño de muestra es proporcional a la cantidad $p(1 - p)$. Un valor de esa cantidad dará como resultado un tamaño de la muestra mayor. En la tabla 8.5 vemos algunos valores posibles de $p(1 - p)$. Observe que el mayor valor de $p(1 - p)$ se presenta cuando $p = .50$. Así, si hay alguna incer-

Tabla 8.5 ALGUNOS VALORES POSIBLES DE $p(1 - p)$

p	$p(1 - p)$	
.10	(.10)(.90) = .09	
.30	(.30)(.70) = .21	
.40	(.40)(.60) = .24	
.50	(.50)(.50) = .25	← Valor máximo de $p(1 - p)$
.60	(.60)(.40) = .24	
.70	(.70)(.30) = .21	
.90	(.90)(.10) = .09	

Con frecuencia se diseñan las encuestas que contienen muchas preguntas con un valor de planeación $p = .50$ para calcular el tamaño de muestra. Al emplear este valor de planeación se obtiene un tamaño de muestra que garantiza que todos los estimados de proporciones cumplan con sus requisitos de margen de error.

tidumbre acerca de un valor adecuado de planeación para p, ya sabemos que $p = .50$ dará como resultado la recomendación del tamaño máximo de muestra. De hecho, nos situamos del lado de la seguridad o del lado conservador, al recomendar el mayor tamaño muestral posible. Si sucede que la proporción es distinta del valor de planeación de .50, la aseveración de precisión será mejor que lo anticipado. En cualquier caso, al usar $p = .50$ estamos garantizando que el tamaño de la muestra será suficiente para obtener el margen de error deseado.

En el ejemplo de Scheer Industries, un valor de planeación de $p = .50$ hubiera dado como resultado el siguiente tamaño de la muestra recomendado.

$$n = \frac{(1.96)^2.50(1 - .50)}{(.10)^2} = 96$$

Este mayor tamaño de muestra refleja la precaución inherente al usar el valor conservador de planeación para la proporción poblacional.

NOTAS y comentarios

El margen de error para estimar una proporción de población es casi siempre .10 o menor. En las encuestas nacionales de opinión que llevan a cabo organizaciones como Gallup y Harris, generalmente se establece un margen de error de .03 o .04. El empleo de esos márgenes de error [E en la ecuación 8.9], por lo general, dará un tamaño de muestra suficientemente grande como para satisfacer los requisitos del teorema del límite central, que $np \geq 5$ y $n(1 - p) \geq 5$.

EJERCICIOS

MÉTODOS

31. Una muestra aleatoria simple de 400 artículos contiene 100 respuestas Sí.
- **a.** ¿Cuál es el estimado puntual de la proporción de la población que tiene respuestas Sí?
- **b.** ¿Cuál es el error estándar de la población?
- **c.** Determine el intervalo de confianza de 95% para la proporción poblacional.

32. Una muestra aleatoria simple de 800 unidades genera una proporción $\bar{p} = .70$.
- **a.** Determine un intervalo de confianza de 90% para la proporción de población.
- **b.** Determine un intervalo de confianza de 95% para la proporción de población.

33. En una encuesta se dice que el valor de planeación para la proporción poblacional p es de .35. ¿De qué tamaño se debe tomar la muestra para tener un intervalo de confianza de 95% con margen de error igual a .05?

34. ¿De qué tamaño se debe tomar una muestra para tener el 95% de confianza de que el margen de error para la estimación de una proporción poblacional sea de .03? Suponga que no dispone de datos históricos para establecer un valor de planeación para p.

APLICACIONES

AUTOEXAMEN

35. En una encuesta de *Time*/CNN se pidió a 814 adultos que contestaran un cuestionario acerca de sus ideas sobre el estado general interno de Estados Unidos. A la pregunta: ¿Cree usted que todo va bien en Estados Unidos en la actualidad? 562 adultos contestaron Sí (*Time*, 11 de agosto de 1997).
- **a.** ¿Cuál es el estimado puntual de la proporción poblacional de adultos que creen que las cosas van bien en Estados Unidos?
- **b.** ¿Cuál es el margen de error, con 90% de confianza?
- **c.** ¿Cuál es el intervalo de confianza de 90% para la proporción de adultos que creen que todo va bien en Estados Unidos?

36. La empresa Bureau of National Affairs, Inc., seleccionó una muestra de 617 empresas y encontró que 56 de ellas pedían a sus empleados cederles los premios que ganarán en los programas de viajeros frecuentes ofrecidos por las aerolíneas, cuando las distancias recorridas se deberían a viajes de negocios (*The Wall Street Journal*, 28 de marzo de 1994).
   a. ¿Cuál es el estimado puntual de la proporción de las empresas que piden a sus empleados cederles los premios por distancia recorrida?
   b. Determine un estimado de intervalo de confianza de 95% de la proporción poblacional.

37. Whirtlin Worldwide reunió datos sobre las actitudes acerca de la calidad del servicio a clientes en tiendas de ventas al menudeo. La encuesta determinó que el 28% de los estadounidenses creen que el servicio a clientes es mejor en la actualidad que dos años atrás (*USA Today*, 20 de enero de 1998). Si en la muestra participaron 650 adultos, determine un intervalo de confianza de la proporción poblacional de adultos que creen que el servicio a clientes es mejor actualmente que hace dos años.

38. Los datos sobre el perfil de la audiencia del sitio de la Red ESPN SportsZone indicaron que el 26% de los usuarios eran mujeres (*USA Today*, 21 de enero de 1998). Suponga que este porcentaje se basó en una muestra de 400 usuarios.
   a. Con un 95% de nivel de confianza, ¿cuál es el margen de error asociado con la proporción estimada de mujeres?
   b. ¿Cuál es el intervalo de confianza de 95% de la proporción poblacional de usuarios mujeres?
   c. ¿Qué tamaño debe tener la muestra para un margen deseado de error de 3%?

AUTOEXAMEN

39. El Instituto de Turismo para el Estado de Florida va a muestrear visitantes en las principales playas del estado para estimar la proporción de quienes no son residentes de Florida. Las estimaciones anteriores son que el 55% de los visitantes en las playas no son residentes.
   a. ¿De qué tamaño se debe tomar la muestra para estimar la proporción de visitantes de otros estados con precisión de 3% respecto al valor real? Use un nivel de confianza de 95%.
   b. ¿De qué tamaño debe ser la muestra para que el error aumente a 6%?

40. Ferrel Calvillo Communications llevó a cabo una encuesta nacional de 902 golfistas mujeres, para estudiar cómo se consideran tratadas en los campos de golf (*USA Today*, 3 de junio de 1997). En la encuesta se encontró que 397 mujeres están satisfechas con los tiempos disponibles de los *tees*, 307 estaban satisfechas con los reglamentos de membresía y 234 estaban satisfechas con las instalaciones de los vestidores.
   a. Determine el estimado puntual de la proporción poblacional, y el intervalo de confianza de 95% para cada una de las tres preguntas de la encuesta.
   b. El artículo de *USA Today* donde se mencionó la encuesta se titulaba: "Encuesta: A las mujeres las tratan bajo par." ¿Cree usted que la información estadística del inciso *a* justifica este título?

41. Una encuesta de *The Wall Street Journal* y NBC News reunió datos acerca de cómo consideran los estadounidenses la calidad de la información en los diarios y en TV (*The Wall Street Journal*, 27 de junio de 1997). Una de las preguntas fue si el encuestado cree que lo que se dice de la economía de Estados Unidos es equilibrado, demasiado negativo o demasiado positivo. Los estimados preliminares son de que un 50% de la población cree que la información es equilibrada.
   a. ¿Qué tamaño de muestra se recomienda para que el margen deseado de error sea de 3.3%? Emplee el 95% de nivel de confianza.
   b. ¿Qué tamaño se recomienda para que sea 2.5%, con 95% de confianza?

42. Una encuesta entre mujeres ejecutivas, llevada a cabo por Louis Harris & Associates indicó que el 33% de las encuestadas evaluaba a su propia empresa como un lugar excelente para el trabajo de las mujeres ejecutivas (*Working Woman*, noviembre de 1994). Suponga que la revista *Working Woman* desea llevar a cabo una encuesta anual para dar seguimiento a esa proporción. Si $p = .33$ como valor de planeación de la proporción poblacional, ¿cuántas mujeres ejecutivas se deben muestrear para tener cada uno de los márgenes de error de la siguiente lista?
   a. 10%
   b. 5%
   c. 2%
   d. 1%
   e. En general, ¿qué sucede con el tamaño de muestra al disminuir el margen de error?

43. La Liga de Teatros y Productores Americanos recurre a una encuesta permanente para obtener información actualizada acerca del público en los teatros de Broadway (*Playbill*, invierno de 1997). Cada semana, la liga distribuye un cuestionario de una página en asientos aleatorios en teatros determinados por rotación, donde se representan espectáculos en Broadway. Sólo se necesitan 5 minutos para contestar el cuestionario, y permite que el público comunique sus ideas acerca de las actividades teatrales.

   a. ¿De qué tamaño debe ser la muestra para que el margen deseado de error para cualquier proporción poblacional sea de .04? Emplee un intervalo de confianza de 95% y un valor de planeación $p = .50$.

   b. Suponga que el tamaño muestral recomendado en el inciso a aumenta, y que en una semana hubo 445 personas que dijeron no vivir en la ciudad de Nueva York. ¿Cuál es el estimado puntual de los espectadores de Broadway que no viven en la ciudad de Nueva York?

   c. Use los datos del inciso *b* para determinar el intervalo de confianza de 95% para la proporción poblacional de los espectadores en Broadway que no viven en la ciudad de Nueva York.

RESUMEN

En este capítulo presentamos métodos para determinar un intervalo de confianza para una media de población μ y para una proporción poblacional p. El objetivo que se persigue al definir un intervalo de confianza es dar al usuario una mejor comprensión del error muestral que pueda presentarse. Un intervalo amplio de confianza indica mala precisión; en esos casos se puede aumentar el tamaño de la muestra para reducir el ancho del intervalo y mejorar la precisión del estimado.

La figura 8.11 condensa los procedimientos de estimación del intervalo para la media de población y es una guía práctica para calcular el estimado. La figura muestra que la ecuación usada para calcular un estimado de intervalo depende de si el tamaño muestral es grande ($n \geq 30$) o pequeño ($n < 30$), o de si se conoce la desviación estándar poblacional, y en algunos casos de si la población tiene una distribución normal o aproximadamente normal de probabilidades. Si el tamaño de muestra es grande, no se requiere hipótesis acerca de la distribución de la población, y se usa $z_{\alpha/2}$ en los cálculos del estimado de intervalo. Si el tamaño de muestra es pequeño, la población debe tener una distribución normal o aproximadamente normal de probabilidades, para poder determinar un estimado de intervalo para μ. En este caso, $z_{\alpha/2}$ se usa en el cálculo cuando se conoce σ, mientras que $t_{\alpha/2}$ se usa cuando se estima σ mediante la desviación estándar σ de la muestra. Por último, si el tamaño de muestra es pequeño y no es adecuada la hipótesis de una población distribuida normalmente, se recomienda aumentar a $n \geq 30$ el tamaño de muestra para determinar un estimado de intervalo de la media de población con muestra grande.

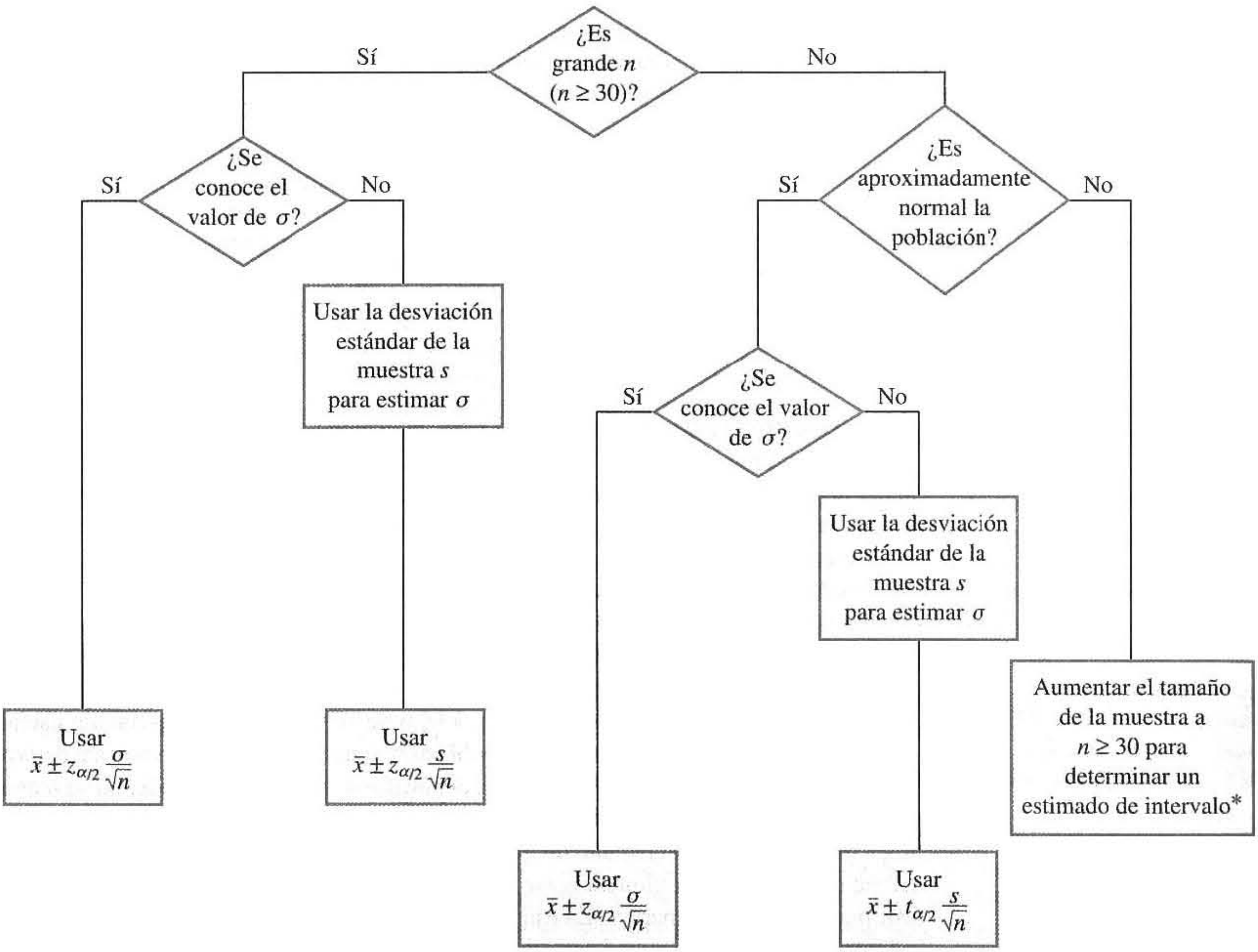
Además, mostramos cómo determinar el tamaño de muestra para que los estimados de intervalo de μ y de p tengan un nivel especificado de precisión. En la práctica, los tamaños de muestra requeridos para estimados de intervalo de una proporción poblacional son grandes, generalmente. En consecuencia, mostramos las fórmulas para estimar intervalos con muestra grande, para una proporción poblacional en la que $np \geq 5$ y $n(1 - p) \geq 5$, al mismo tiempo.

GLOSARIO

Estimado de intervalo Un estimado de un parámetro de población que define un intervalo dentro del que se cree está contenido el valor del parámetro.

Error muestral El valor absoluto de la diferencia entre el valor de un estimador puntual insesgado, como la media de muestra $\bar{x}$, y el valor del parámetro poblacional que estima, como la media de población μ; en este caso, el error muestral es $|\bar{x} - \mu|$. En el caso de la proporción poblacional, el error muestral es $|\bar{p} - p|$.

Figura 8.11 RESUMEN DE PROCEDIMIENTOS DE ESTIMACIÓN DE INTERVALO PARA UNA MEDIA DE POBLACIÓN



*A veces es posible usar métodos para estadísticas no paramétricas en casos de muestras pequeñas, para desarrollar intervalos confiables para localizar parámetros de una población.

Precisión Una aseveración probabilística acerca del error muestral.

Nivel de confianza La confianza asociada con un estimado de intervalo. Por ejemplo, si un procedimiento de estimación de intervalo da resultados tales que en el 95% de los intervalos formados usando ese procedimiento esté situado el parámetro poblacional, se dice que el estimado de intervalo está determinado con el nivel de confianza de 95%. Observe que .95 se llama coeficiente de confianza.

Margen de error El valor ± sumado a y restado de un punto estimado a fin de determinar un intervalo de confianza.

Distribución *t* Una familia de distribuciones de probabilidades que se puede usar para determinar estimados de intervalo de una media de población cuando se desconozca la desviación estándar de la población, y que ésta tenga una distribución normal o casi normal de probabilidades.

Grados de libertad Un parámetro de la distribución t. Cuando se usa la distribución t para determinar un estimado de intervalo de una media de población, tiene $n - 1$ grados de libertad, siendo n el tamaño de la muestra aleatoria simple.

FÓRMULAS CLAVE

Error muestral al estimar μ

$$|\bar{x} - \mu| \tag{8.1}$$

Estimado de intervalo de una media de población: caso de muestra grande y σ conocida

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \tag{8.2}$$

Estimado de intervalo de una media de población: caso de muestra grande y σ desconocida

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} \tag{8.3}$$

Estimado de intervalo de un promedio poblacional: caso de muestra pequeña y σ desconocida

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} \tag{8.4}$$

Tamaño de muestra para un estimado de intervalo de una media de población

$$n = \frac{(z_{\alpha/2})^2 \sigma^2}{E^2} \tag{8.5}$$

Estimado de intervalo de una proporción poblacional

$$\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} \tag{8.8}$$

Tamaño de muestra para un estimado de intervalo de una proporción poblacional

$$n = \frac{(z_{\alpha/2})^2 p(1 - p)}{E^2} \tag{8.9}$$

EJERCICIOS COMPLEMENTARIOS

44. Una encuesta que hizo Day-Timers determinó que los hombres en Estados Unidos trabajan un promedio de 52 horas a la semana (*The Cincinnati Enquirer*, 23 de octubre de 1995). Suponga que la encuesta se basó en una muestra de 650 hombres, y que la desviación estándar fue de 8.2 horas. ¿Cuál es el intervalo de confianza de 95% de la cantidad promedio de horas trabajadas por semana para la población estadounidense de hombres?

45. Una encuesta que realizó la Asociación Automotriz Estadounidense mostró que, para una familia de cuatro miembros, la media del gasto diario es de \$215.60 dólares cuando está de vacaciones. Suponga que se toma una muestra de 64 familias de cuatro miembros en cierto lugar, y se obtuvo una media de muestra de \$252.45 dólares diarios y una desviación estándar muestral de \$74.50 dólares.

- **a.** Determine un estimado del intervalo de confianza de 95% para la media de los gastos diarios de una familia de cuatro miembros que visita ese lugar.
- **b.** Con el intervalo de confianza determinado en el inciso *a*, ¿parece que la media de la población del gasto diario es distinto a la media que cita la Asociación Automotriz Estadounidense? Explique su respuesta.

ActTemps

46. Una empresa eléctrica observa que una muestra de cuentas morosas tiene una media de la deuda de \$131.44 dólares, con una desviación estándar de la muestra de \$16.19 dólares. Determine un intervalo de confianza del 90% para la cantidad promedio de deuda.

- **a.** Use ActTemps para determinar un punto estimado del número de minutos por día que pierden los oficinistas al tratar de localizar documentos mal etiquetados, mal archivados o extraviados
- **b.** ¿Cuál es la desviación estándar de la muestra?
- **c.** ¿Cuál es el intervalo de confianza de 95% para la media del número de minutos perdidos por día?

Mail

47. La empresa Arthur D. Little, Inc., estima que un 70% de la correspondencia que recibe una familia consiste en propaganda (*Time*, 14 de julio de 1997). En una muestra de 20 familias se recabaron los siguientes datos de la cantidad de propaganda que recibieron y la cantidad total de piezas postales recibidas durante una semana.

Familia	Propaganda	Correo total	Familia	Propaganda	Correo total
1	24	35	11	13	19
2	9	14	12	16	28
3	18	30	13	20	27
4	9	12	14	17	22
5	15	28	15	21	24
6	23	33	16	21	33
7	13	20	17	15	25
8	17	20	18	15	24
9	20	23	19	18	24
10	20	25	20	12	16

- **a.** ¿Cuál es el estimado puntual para la cantidad media de propaganda recibida por semana? ¿Cuál es el intervalo de confianza de 95% la media de población?
- **b.** ¿Cuál es el estimado puntual para la cantidad media de piezas postales recibidas durante una semana? ¿Cuál es el intervalo de confianza de 95% para la media de población?
- **c.** Emplee los estimados puntuales de los incisos *a* y *b*; ¿están de acuerdo con la aseveración de que un 70% de la correspondencia consiste en propaganda?

48. Los tiempos de armado en una muestra de determinada parte manufacturada fueron 8, 10, 10, 12, 15 y 17 minutos. Si se usa la media de la muestra para estimar la media de la población de tiempos de armado, determine un estimado puntual y un intervalo de confianza del 90% para la media de población. Suponga que la población tiene distribución normal.

49. Muchos estadounidenses que trabajan en grandes oficinas también trabajan en casa o en su oficina durante los fines de semana (*USA Today*, 18 de junio de 1997). ¿De qué tamaño debe ser una muestra para estimar la media de población del tiempo que se trabaja en los fines de semana, con un margen de error de 10 minutos? Emplee un nivel de confianza de 95% y suponga que el valor de planeación de la desviación estándar poblacional es de 45 minutos.

50. Para determinado modelo de automóvil se llevan a cabo pruebas de rendimiento de gasolina. Si la precisión que se desea es un intervalo de confianza de 98% con un margen de error de 1 milla por galón, ¿cuántos automóviles deben participar en la prueba? Suponga que las pruebas preliminares de rendimiento indican que la desviación estándar es de 2.6 millas por galón.

51. Al determinar la programación de las citas con pacientes, un centro médico desea un estimado de la media del tiempo que pasa un miembro de su personal con cada paciente. ¿De qué tamaño se debe tomar una muestra para que el margen de error sea 2 minutos a un nivel de confianza del 95%? ¿De qué tamaño debe ser la muestra para tener un nivel de confianza de 99%? Emplee un valor de planeación de 8 minutos para la desviación estándar poblacional.

52. En la 47ª Encuesta Anual de Pagos que se presenta en *Business Week*, se ven los datos de salario anual y bonos para los directores ejecutivos (*Business Week*, 21 de abril de 1997). En una muestra preliminar se vio que la desviación estándar es de $675 dólares, estando los datos en miles de dólares. ¿Cuántos directores ejecutivos deben estar en la muestra si deseamos estimar el salario o bono anual de la media de población, con un margen de error de $100,000 dólares? (Nota: El margen de error sería $E = 100$, porque los datos están en miles de dólares.) Emplee un intervalo de confianza de 95%.

53. El Centro Nacional de Estadísticas Educativas de Estados Unidos informó que el 47% de los alumnos de licenciatura trabajan para pagarse sus gastos (*The Tampa Tribune*, 22 de enero de 1997). Suponga que en esa encuesta se usó una muestra de 450 alumnos de licenciatura.

a. Determine un intervalo de confianza de 95% para la proporción poblacional de alumnos de licenciatura que trabajan para pagarse sus gastos.

b. Determine el intervalo de confianza de 99%.

c. ¿Qué sucede con el margen de error al aumentar la confianza de 95 a 99%?

54. Una encuesta de *USA Today* y CNN Gallup, entre 369 padres que trabajan, determinó que 200 de ellos dijeron pasar muy poco tiempo con sus niños, debido a compromisos en el trabajo (*USA Today*, 10 de abril de 1995).

a. ¿Cuál es el estimado puntual de la proporción poblacional de padres que trabajan que creen pasar muy poco tiempo con sus hijos debido a sus compromisos en el trabajo?

b. ¿Cuál es el margen de error, con 95% de confianza?

c. ¿Cuál es el estimado de intervalo de confianza de 95% para la proporción poblacional de padres que trabajan y creen pasar muy poco tiempo con sus hijos?

55. En una encuesta telefónica de *Time* y CNN entre 1400 adultos estadounidenses se preguntó "¿Dónde va usted en su tiempo libre?" La principal respuesta, de 504 adultos, fue: a un centro comercial.

a. ¿Cuál es el estimado puntual de la proporción de adultos que prefieren ir a un centro comercial en su tiempo libre?

b. Con 95% de confianza, ¿cuál es el error muestral asociado con este estimado?

56. Una conocida empresa bancaria de tarjetas de crédito tiene interés en estimar la proporción de tarjetahabientes cuyo saldo es distinto de cero a final del mes, e incurren en intereses. Suponga que el margen de error deseado es de .03 con un intervalo de confianza de 98%.

a. ¿De qué tamaño se debe seleccionar una muestra si se cree que, más o menos, el 70% de los tarjetahabientes llegan con un saldo distinto de cero al final del mes?

b. ¿De qué tamaño se debe seleccionar una muestra si no se puede especificar un valor de planeación para la proporción poblacional?

57. Se pidió a una muestra de 200 personas identificar su principal fuente de información de noticias; 110 dijeron que esa fuente es los noticiarios televisivos.

a. Determine un intervalo de confianza de 95% para la proporción de las personas en la población que consideran a la televisión como su principal fuente de información noticiosa.

b. ¿Qué tamaño debe tener una muestra para estimar la proporción de la población, con un margen de error igual a .05, a nivel de confianza de 95%?

58. Una encuesta de 502 mujeres ejecutivas, llevada a cabo por Louis Harris & Associates indicó que 166 de ellas valorizaban a su propia empresa como lugar excelente para trabajo como el de ellas (*Working Woman*, noviembre de 1994). Determine un intervalo de confianza de 95% para la población de todas las ejecutivas que valorizan a su propia empresa como un excelente lugar de trabajo para mujeres ejecutivas.

59. En el *1997 Statistical Abstract of the United States* se menciona el porcentaje de personas de 18 años y mayores que fuman. Suponga que un estudio se diseña para reunir nuevos datos de fumadores y no fumadores. El mejor estimado preliminar de la proporción poblacional de quienes fuman es 30%.

a. ¿De qué tamaño debe tomarse una muestra para estimar la proporción de fumadores en la población, un margen de error igual a .02? Emplee un nivel de confianza de 95%.

b. Suponga que el estudio usa su recomendación de tamaño de muestra del inciso *a*, y ve que hay 520 fumadores. ¿Cuál es la estimación puntual de la proporción de fumadores en la población?

c. ¿Cuál es el intervalo de confianza de 95% para la proporción de fumadores en la población?

60. Aunque los horarios y los costos de las aerolíneas son factores importantes que toman en cuenta los viajeros por negocios para seleccionar una aerolínea, una encuesta de la revista *USA Today* indicó que dichos viajeros consideran la existencia de un programa de viajeros frecuentes como el factor más importante (*USA Today*, 11 de abril de 1995). De una muestra de viajeros de negocios que participaron en la encuesta, 618 de 1993 dijeron que el factor más importante son los programas de viajeros frecuentes.

a. ¿Cuál es el estimado puntual de la proporción de la población de viajeros de negocios que creen que un programa para usuario frecuente es el factor más importante al seleccionar una aerolínea?

b. Determine un intervalo de confianza del 95% para la proporción poblacional.

c. ¿De qué tamaño se requiere una muestra para informar que el margen de error es de .01 con 95% de confianza? ¿Recomendaría usted que la revista *USA Today* tratara de llegar a este grado de precisión? ¿Por qué?

CASO A RESOLVER 1

BOCK INVESTMENT SERVICES

Lisa Rae Bock inició la empresa Bock Investmente Services (BIS) en 1994, con el objetivo de llegar a ser el principal servicio de asesoría sobre mercado de dinero en Carolina del Sur. Para dar mejor servicio a sus clientes actuales y atraer nuevos, ha publicado un boletín semanal. Desea agregar una nueva sección al boletín, donde se vean los resultados de una encuesta telefónica semanal entre los gerentes de agencias financieras. Para investigar la posibilidad de ofrecer ese servicio, y determinar qué tipo de información incluir en el boletín, Lisa seleccionó una muestra aleatoria de 45 fondos de mercado de dinero. En la tabla 8.6 vemos una parte de los datos que obtuvo, donde aparecen los activos y rendimientos durante los siete y los 30 días pasados (*Barrons*, 3 de octubre de 1994). Antes de llamar a los gerentes y obtener datos adicionales, decidió Lisa llevar a cabo un análisis preliminar de los datos ya reunidos (véase la tabla 8.6).

Informe gerencial

1. Aplique los estadísticos adecuados para resumir los datos sobre activos y rendimientos del mercado de dinero.
2. Determine un estimado del intervalo de confianza del 95% las medias de los activos, del rendimiento a siete días y el de rendimiento a 30 días para la población de fondos de mercado de dinero. Ofrezca una interpretación gerencial de cada estimado de intervalo.
3. Describa la implicación de sus determinaciones en función de cómo Lisa podría usar este tipo de información para preparar su boletín semanal.
4. ¿Qué otra información recomendaría usted que reunirá Lisa para proporcionar el material más útil a sus clientes?

Tabla 8.6 [illegible]

Bock

Inversora	Activos (millones $)	Rendimiento a 7 días (%)	Rendimiento a 30 días (%)
Amcore	103.9	4.10	4.08
Alger	156.7	4.79	4.73
Arch MM/Trust	496.5	4.17	4.13
BT Instit Treas	197.8	4.37	4.32
Benchmark Div	2755.4	4.54	4.47
Bradford	707.6	3.88	3.83
Capital Cash	1.7	4.29	4.22
Cash Mgt Trust	2707.8	4.14	4.04
Composite	122.8	4.03	3.91
Cowen Standby	694.7	4.25	4.19
Cortland	217.3	3.57	3.51
Declaration	38.4	2.67	2.61
Dreyfus	4832.8	4.01	3.89
Elfun	81.7	4.51	4.41
FFB Cash	506.2	4.17	4.11
Federated Master	738.7	4.41	4.34
Fidelity Cash	13272.8	4.51	4.42
Flex-fund	172.8	4.60	4.48
Fortis	105.6	3.87	3.85
Franklin Money	996.8	3.97	3.92
Freedom Cash	1079.0	4.07	4.01
Galaxy Money	801.4	4.11	3.96
Government Cash	409.4	3.83	3.82
Hanover Cash	794.3	4.32	4.23
Heritage Cash	1008.3	4.08	4.00
Infinity/Alpha	53.6	3.99	3.91
John Hancock	226.4	3.93	3.87
Landmark Funds	481.3	4.28	4.26
Liquid Cash	388.9	4.61	4.64
MarketWatch	10.6	4.13	4.05
Merrill Lynch Money	27005.6	4.24	4.18
NCC Funds	113.4	4.22	4.20
Nationwide	517.3	4.22	4.14
Overland	291.5	4.26	4.17
Pierpont Money	1991.7	4.50	4.40
Portico Money	161.6	4.28	4.20
Prudential MoneyMart	6835.1	4.20	4.16
Reserve Primary	1408.8	3.91	3.86
Schwab Money	10531.0	4.16	4.07
Smith Barney Cash	2947.6	4.16	4.12
Stagecoach	1502.2	4.18	4.13
Strong Money	470.2	4.37	4.29
Transamerica Cash	175.5	4.20	4.19
United Cash	323.7	3.96	3.89
Woodward Money	1330.0	4.24	4.21

Fuente: *Barron's*, 3 de octubre de 1994.

CASO A RESOLVER 2

METROPOLITAN RESEARCH, INC.

Esta empresa es una organización dedicada a la investigación de consumo que lleva a cabo encuestas con objeto de evaluar una gran variedad de productos y servicios que se ofrecen a los consumidores. En un estudio, a Metropolitan le interesaba conocer la satisfacción del consumidor con el funcionamiento de los automóviles producidos por uno de los principales fabricantes de Detroit. Mandó un cuestionario a los propietarios de uno de los modelos grandes y halló varias quejas sobre problemas prematuros de la transmisión. Para investigar más sobre las fallas de las transmisiones, Metropolitan usó una muestra de reparaciones de transmisión realizadas en un taller de transmisiones del área de Detroit. Los siguientes datos muestran la cantidad real de millas recorridas por 50 vehículos en el momento de la falla de la transmisión.

Auto

85,092	32,609	59,465	77,437	32,534	64,090	32,464	59,902
39,323	89,641	94,219	116,803	92,857	63,436	65,605	85,861
64,342	61,978	67,998	59,817	101,769	95,774	121,352	69,568
74,276	66,998	40,001	72,069	25,066	77,098	69,922	35,662
74,425	67,202	118,444	53,500	79,294	64,544	86,813	116,269
37,831	89,341	73,341	85,288	138,114	53,402	85,586	82,256
77,539	88,798						

Informe gerencial

1. Aplique estadísticos adecuados para resumir los datos de fallas de la transmisión.
2. Determine un intervalo de confianza del 95% para la media de las millas recorridas hasta la falla, para la población de automóviles que tuvieron descompostura de transmisión. Presente una interpretación gerencial del estimado del intervalo.
3. Describa la implicación de sus resultados estadísticos en términos de la hipótesis de que algunos propietarios han tenido fallas prematuras de la transmisión.
4. ¿Cuántos casos de reparación deben muestrearse si la empresa desea estimar la media de la población de la cantidad de millas recorridas hasta la falla con precisión de 5000 millas con 95% de confianza?
5. ¿Qué otra información le gustaría reunir para evaluar con más detalle el problema de fallas de transmisión?

APÉNDICE 8.1 ESTIMACIÓN DEL INTERVALO DE CONFIANZA CON MINITAB

Caso de muestra grande

En la sección 8.1 mostramos el uso de estimados del intervalo de confianza generados en computadora, indicando cómo se puede usar el programa Minitab para obtener estimados de intervalo para la media de la edad en el estudio de la Aseguradora Estatal. Los datos aparecen en la tabla 8.2. Con la desviación estándar de la muestra, $s = 7.77$, como estimado de la desviación estándar poblacional σ, se pueden usar los pasos siguientes para producir el resultado del intervalo de confianza de 90% que se ve en la figura 8.6 (suponga que se han capturado los datos de la edad en la columna C1 de la hoja de cálculo de Minitab).

Paso 1. Seleccione el menú desplegable **Stat**
Paso 2. Seleccione el menú desplegable **Basic Statistics**

Paso 3. Seleccione la opción **1-Sample z**
Paso 4. Cuando aparezca el cuadro de diálogo:
Teclee C1 en el cuadro **Variables**
Teclee 90 en el cuadro **Confidence interval Level**
Teclee 7.77 en el cuadro **Sigma**
Seleccione **OK**

Caso de muestra pequeña

En la sección 8.2 mostramos el intervalo de confianza generado con Minitab, para la media de la población en el problema de Scheer Industries. Con los datos de la tabla 8.4 capturados en la columna C1, se emplean los siguientes pasos para determinar un intervalo de confianza el 95%, que aparece en la figura 8.9.

Paso 1. Seleccione el menú desplegable **Stat**
Paso 2. Seleccione el menú desplegable **Basic Statistics**
Paso 3. Seleccione la opción **1-Sample t**
Paso 4. Cuando aparezca el cuadro de diálogo:
Teclee C1 en el cuadro **Variables**
Teclee 95 en el cuadro **Confidence interval Level**
Seleccione **OK**

APÉNDICE 8.2 ESTIMACIÓN DEL INTERVALO DE CONFIANZA CON EXCEL

Caso de muestra grande

Ahora mostraremos cómo usar Excel para determinar intervalos de confianza para una media de la población en el caso de muestra grande, con el caso de un intervalo de confianza del 90% para la media de la edad de la población en el estudio de la Aseguradora Estatal, en la sección 8.1. Usaremos los datos de la tabla 8.2 y pondremos la Edad como nombre de la variable en la celda A1, y las 36 edades en las celdas A2 a A37.

Paso 1. Seleccione el menú desplegable **Herramientas**
Paso 2. Seleccione la opción **Análisis de datos**
Paso 3. Cuando aparezca el cuadro de diálogo Análisis de datos:
Seleccione **Estadística descriptiva**
Oprima **Aceptar**
Paso 4. Cuando aparezca el cuadro de diálogo Estadística Descriptiva:
Teclee A1:A37 en el cuadro **Rango de entrada**
Seleccione **Rótulos en primera fila**
Seleccione **Resumen de estadísticas**
Seleccione **Rango de salida** y seleccione B1 en el cuadro de diálogo
Seleccione **OK**
El valor del error estándar aparece en la celda C4
Paso 5. Seleccione la celda D3
Teclee la fórmula = 1.645 * C4

La media de la muestra 39.5 aparece en la celda C3 y el margen de error 2.13 aparece en la celda D3. Con la media de la muestra de 39.5 el intervalo de confianza del 90% se obtiene al restar 2.13 de 39.5 y sumando 2.13 a 39.5. El intervalo de confianza es de

37.37 a 41.63. Observe que se usó 1.645 en la fórmula del paso 5, porque se pidió un intervalo de confianza del 90%. En general, el usuario debe teclear el valor de z que corresponde al nivel deseado de confianza.

Caso de muestra pequeña

Para ilustrar cómo determinar un intervalo de confianza para el caso de muestra pequeña usaremos el intervalo de confianza 95% en el caso de Scheer Industries que describimos en la sección 8.2. El lector debe capturar los datos de tiempo de adiestramiento, de la tabla 8.4, en los renglones 2 a 16 de la columna A de la hoja de trabajo. Se usan los siguientes pasos para determinar el intervalo de confianza del 95%.

Paso 1. Seleccione el menú desplegable **Herramientas**
Paso 2. Seleccione la opción **Análisis de datos**
Paso 3. Cuando aparezca el cuadro de diálogo Análisis de datos:
Seleccione **Estadística descriptiva**
Seleccione **Aceptar**
Paso 4. Cuando aparezca el cuadro de diálogo Estadística descriptiva:
Teclee A1:A16 en el cuadro **Rango de entrada**
Seleccione **Rótulos en primera fila**
Seleccione **Resumen de estadísticas**
Seleccione **Nivel de confianza para la media** y teclee 95 en el cuadro
Seleccione **Rango de salida** y teclee B1 en el cuadro
Oprima **Aceptar**

La media de la muestra 53.87 aparece en la celda C3 y el margen de error 3.78 en la celda C16; observe que el nombre en Excel para el margen de error es Nivel de Confianza (*Confidence Level*) (95.0%). Con la media de la muestra de 53.87, el intervalo de confianza de 95% se obtiene restando 3.78 de 53.87 y sumando 3.78 a 53.87. Este intervalo de confianza va de 50.09 a 57.65.