

Unidad 11

- Pruebas de significación de medias.

10

pruebas de significación de medias

Objetivos del capítulo

Después de estudiar este capítulo, se deberá estar en condiciones de:

1. Establecer el objeto general de las pruebas de significación para medias.
2. Listar tres tipos diferentes de proporciones que se puedan establecer con respecto a las medias.
3. Esbozar el procedimiento para una prueba de significación.
4. Utilizar datos muestrales para probar proposiciones (es decir, poder resolver problemas comunes).

Sinopsis del capítulo

Introducción

Prueba de medias de una muestra

(σ_x) conocida

(σ_x) no conocida

Pruebas de medias de dos muestras

Cálculo de la probabilidad de un error de tipo II

Resumen

INTRODUCCIÓN

EL OBJETO de las pruebas de significación de medias es evaluar las proporciones con respecto a medias de la población. Las diversas pruebas requieren *datos cuantitativos* (es decir, datos continuos o discretos).

Básicamente existen tres tipos de proposiciones que se plantean con respecto a las medias de la población, cada uno de los cuales requiere de un diferente procedimiento de evaluación. Una aseveración puede comprender la media de una población puede la evaluación requiere de una prueba de *una muestra*. O bien, una proposición puede ser que las medias de *dos* poblaciones son iguales, lo que requiere de una prueba de dos muestras. Por último, una aseveración puede ser que las medias de *más de dos* poblaciones sean iguales, y la evaluación requiere de una prueba de *k muestras*. En este capítulo se estudian las primeras dos pruebas, y en el siguiente se analizan las pruebas de *k muestras*.

En el capítulo anterior, el proceso básico de la prueba de significación fue comentado completamente. En resumen, en el caso de pruebas de significación es necesario seguir los siguientes pasos:

1. Establecer las hipótesis nula y alternativa.
2. Identificar una distribución de muestreo apropiada. La mayor parte de las pruebas de medias comprenden ya sea la distribución normal o la distribución *t*.

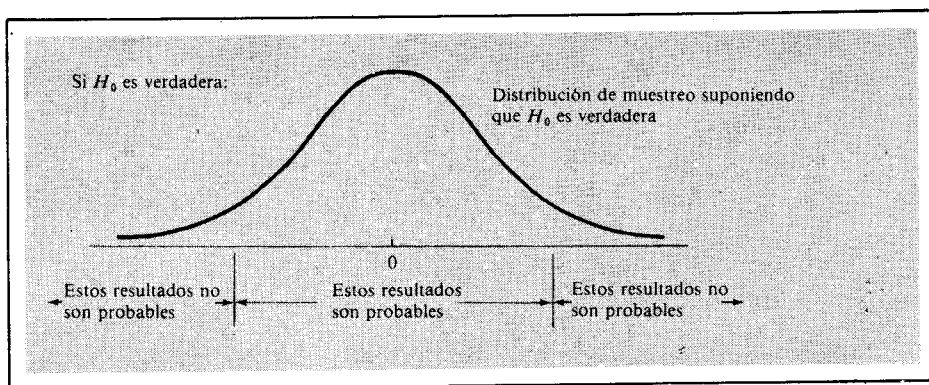


FIGURA 10.1 Concepto general de una prueba de significación de dos colas.

3. Dividir la distribución de muestreo en regiones de aceptación (probablemente variación aleatoria) y rechazo (tal vez sin que exista la casualidad). (Ver figura 10.1)
4. Calcular el valor estadístico de prueba.
5. Comparar el valor estadístico de la muestra con el valor crítico. Rechazar H_0 si ésta es mayor que el valor crítico.

PRUEBA DE MEDIAS DE UNA MUESTRA

Se utiliza una prueba de una muestra para probar una afirmación con respecto a una media de una población única. Se toma una muestra de n observaciones y se calcula el valor medio de la muestra. Posteriormente, la desviación entre el valor propuesto y este valor medio de la muestra se compara con la variabilidad de la distribución de muestreo basada en dicha afirmación. Desviaciones grandes indican que la afirmación es falsa; desviaciones pequeñas sustentarán dicha aseveración.

Por ejemplo, supóngase que se desea evaluar el aserto de un fabricante que establece que sus llantas radiales tienen un promedio de vida de 40 000 millas. La hipótesis nula se convierte en

$$H_0: \mu = 40\,000 \text{ millas}$$

A fin de ilustrarlo más claramente, considérense las tres hipótesis alternativas posibles, aunque de ordinario sólo se debería considerar una. Las tres son

$$H_1: \mu \neq 40\,000 \quad H_1: \mu > 40\,000 \quad H_1: \mu < 40\,000$$

La evaluación de cualquiera de las anteriores debe tener en cuenta el grado en el que el valor estadístico de la muestra puede variar o desviarse del parámetro propuesto, debido a la variación casual en el muestreo. Esto se representará mediante una distribución de muestreo que tenga una media igual al valor del parámetro propuesto. La distribución de muestreo será distribuida normalmente para muestras que se obtengan a partir de una población normal con una desviación estándar conocida, y presentará una distribución t cuando la desviación estándar de la población sea estimada a partir de la desviación estándar de la muestra s . Cuando el tamaño de la muestra es mayor que 30, se puede hacer a un lado la necesidad de suponer que se está muestreando a partir de una población normal.

Conceptualmente, los valores críticos se pueden establecer en valores que se relacionen específicamente con un problema dado. Por ejemplo, en este caso los valores críticos pueden ser 39 000 millas y 41 000 millas. Sin embargo, es mucho más sencillo trabajar con valores estadísticos de pruebas y valores críticos estandarizados, ya que la mayoría de las tablas de probabilidad se establecen en términos de valores estandarizados, como se señaló en el capítulo anterior.

Ahora es muy importante establecer un nivel de significación (el cual, a su vez, dará lugar a un valor crítico) *previo* al muestreo. De otra manera, existe la posibilidad de que la persona que esté evaluando los datos de la muestra seleccione un nivel de significación que lleve a una decisión que corresponda a su idea preconcebida de cómo “debe” ser el resultado de la prueba.

Una vez que se elija el nivel de significación, se pueden reunir los datos de la muestra y calcular el valor estadístico de prueba.

$$\text{valor estadístico de prueba} = \frac{\text{valor medio de la muestra} - \text{media propuesta}}{\text{desviación estándar de la distribución de muestreo}}$$

Si se tiene como dato la desviación estándar de la población, el valor estadístico de prueba es*

$$Z_{\text{prueba}} = \frac{\bar{X} - \mu_0}{\sigma_x / \sqrt{n}}$$

Si no se conoce σ_x , el valor estadístico de prueba es

$$t_{\text{prueba}} = \frac{\bar{X} - \mu_0}{s_x / \sqrt{n}}$$

σ_x conocida

Cuando se conoce la desviación estándar de la población, la distribución de muestreo adecuada es la distribución normal. Si la población que se muestra es normal, la distribución de muestreo será normal en el caso de *todos* los tamaños de la muestra. Si la población no es normal, o si se desconoce su forma, se puede utilizar una prueba de una muestra solamente para tamaños de muestra mayores de 30 observaciones. Por tanto, *en este método no se puede trabajar con muestras pequeñas de poblaciones no normales*.

Considérese ahora con mayor detalle el ejemplo del fabricante de neumáticos, que aseguraba que éstos tenían una vida útil de por lo menos 40 000 millas. Supóngase que los resultados de la prueba fueran los siguientes: una muestra de $n = 49$, con un valor medio muestral de 38 000 millas. Si se sabe que el recorrido de los neumáticos de la población tiene una desviación estándar de 3500 millas se puede proceder como sigue:

1. Establecer las hipótesis nula y alternativa:

$$H_0: \mu = 40\,000 \text{ millas}$$

Suponer que la prueba está siendo llevada a cabo por un grupo de consumidores. Naturalmente este grupo querrá comprobar que no están adquiriendo neumáticos con una vida útil más corta, por lo que la hipótesis alternativa sería

$$H_1: \mu < 40\,000 \text{ millas}$$

2. Seleccionar un nivel de significación y dividir la distribución de muestreo adecuada. Supóngase que el grupo decide aceptar un riesgo de 0.05 de rechazar H_0 si está es verdadera. Así, $\alpha = 0.05$. Como se conoce σ_x y es grande el tamaño de la muestra, se utiliza la distribución normal. El valor de z que deja 0.05 en la cola es -1.65 .

3. Calcular el valor estadístico de prueba:

$$Z_{\text{prueba}} = \frac{\bar{X} - \mu_0}{\sigma_x / \sqrt{n}} = \frac{38\,000 - 40\,000}{3\,500 / \sqrt{49}} = \frac{-2\,000}{500} = -4.0$$

* μ_0 = media propuesta

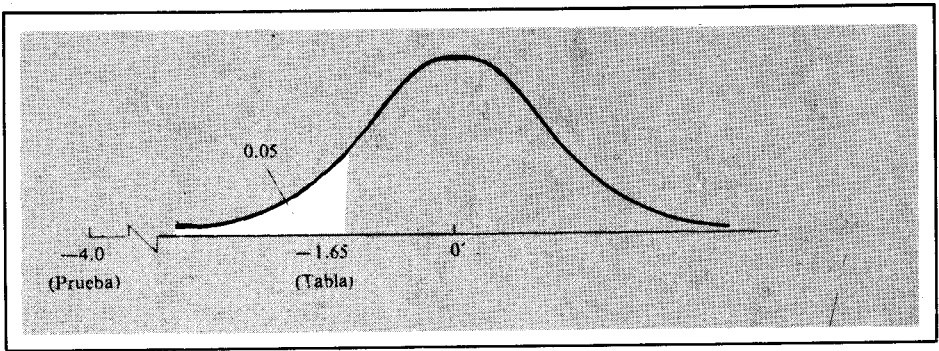


FIGURA 10.2 Se rechaza H_0 , debido a que el valor estadístico de prueba excede el valor crítico (tabla).

4. Comparar el valor estadístico de prueba con el valor crítico. Como -4.0 excede a -1.65 , H_0 es rechazada. Esto se ilustra en la figura 10.2. De este modo, se concluye que el promedio de vida útil es menor que 40 000 millas. (Por tanto, no se llevará a cabo la prueba para $\mu = 40\,000$.)

Cabe observar que esto *no* garantiza que la proposición sea falsa. Se puede estar cometiendo un error de tipo I. Por tanto, siempre que H_0 sea rechazada, esto significará que la evidencia de la muestra *sugiere* que H_0 no es verdadera; sin embargo, no hay forma de estar *absolutamente seguros* de que la proposición sea verdadera o no, a menos que se lleve a cabo un censo completo.

Como se señaló que la media verdadera no es de 40 000 millas, surge una pregunta lógica: ¿Cuál es el promedio verdadero de vida útil? Un método para responder a esta pregunta es utilizar los datos muestrales para estimar la media. Si se emplea un nivel de confianza de $(1 - \alpha)$ (es decir, 95%), se obtiene

$$\bar{x} \pm 1.96 \frac{\sigma_x}{\sqrt{n}} = 38\,000 \pm 1.96(500) = 38\,000 \pm 980$$

Supóngase ahora que es el fabricante el que realiza la prueba. Puede utilizar una prueba de dos colas, puesto que quiere vigilar si le están dando demasiada poca vida útil a los neumáticos que produce, lo que daría como resultado una publicidad falsa (entre otras cosas), y no desea aumentar significativamente la calidad en su fabricación, dado que también vende una llanta radial de 45 000 millas, a un precio más alto. De este modo, se observan los siguientes hechos:

$$\bar{x} = 38\,000 \text{ millas}$$

$$n = 49$$

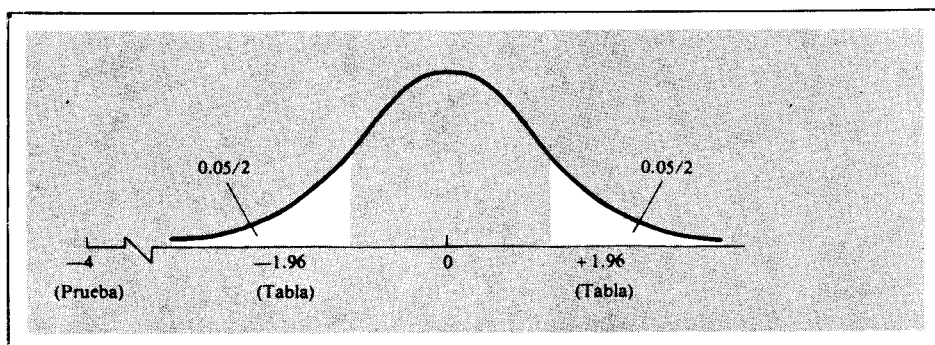
$$\sigma_x = 3500 \text{ millas}$$

El análisis se lleva a cabo como sigue:

1. Establecer H_0 y H_1 :

$$H_0: \mu = 40\,000$$

$$H_1: \mu \neq 40\,000$$

FIGURA 10.3 H_0 es rechazada.

2. Seleccionar α y dividir la distribución de muestreo. Si se utiliza $\alpha = 0.05$, los valores críticos de z son ± 1.96 .
3. Calcular el valor estadístico de prueba. Dado que se están utilizando los mismos datos, dicho valor permanece constante:

$$Z_{\text{prueba}} = \frac{38\,000 - 40\,000}{3\,500/\sqrt{49}} = -4.0$$

4. Como esto excede del valor crítico inferior, el fabricante rechazará H_0 y concluirá que la vida promedio es menor de 40 000.
La prueba se ilustra en la figura 10.3.

σ_x no conocida

Cuando no se conoce la desviación estándar de la población, esta debe ser estimada a partir de los datos muestrales, utilizando la desviación estándar de la muestra. Cuando esto sucede (en la *mayoría* de las situaciones reales, se desconoce σ_x), la distribución t resulta la distribución de muestreo más adecuada. Se considera práctico utilizar la distribución t solamente cuando se requiera que el tamaño de la muestra sea de 30 o menos observaciones, ya que para muestras más grandes los valores de t y z son aproximadamente iguales, y es posible emplear la distribución normal en lugar de la distribución t *

Supóngase que una vez más se considera el ejemplo de las llantas radiales de 40 000 millas. Sin embargo, por ahora se considerará que la desviación estándar de la población es desconocida. Asimismo, se quieren comparar los análisis de los casos para muestras grandes y pequeñas.

Como se hizo anteriormente, se debe considerar como hipótesis nula, H_0 , el hecho de que la media verdadera de la población es 40 000. Además, se deberán probar las tres hipótesis alternativas:

* Sin embargo, el valor estadístico de prueba siempre se representa como t cuando se utiliza s_x , independientemente del tamaño de la muestra.

1. $H_1: \mu < 40\,000$.
2. $H_1: \mu > 40\,000$.
3. $H_1: \mu \neq 40\,000$.

Ejemplo 1 Muestra grande: Supóngase que se toma una muestra de 36 observaciones, y que la media de la muestra resultante es 41 200 y la desviación estándar de 3000. Considérese también $\alpha = 0.05$. Ahora, como n es mayor que 30, el valor de z , tomado a partir de una tabla normal, se puede utilizar para aproximar el valor de t . De ahí que el valor crítico sea $+1.65$, o bien, -1.65 en el caso de una prueba de una cola, y ± 1.96 en el de una de dos colas. Para cualquier caso, el valor estadístico de prueba es

$$t_{\text{prueba}} = \frac{\bar{x} - \mu_0}{s_x / \sqrt{n}} = \frac{41\,200 - 40\,000}{3000 / \sqrt{36}} = +2.4$$

Las tres hipótesis alternativas posibles se comprueban en la figura 10.4. Cabe obser-

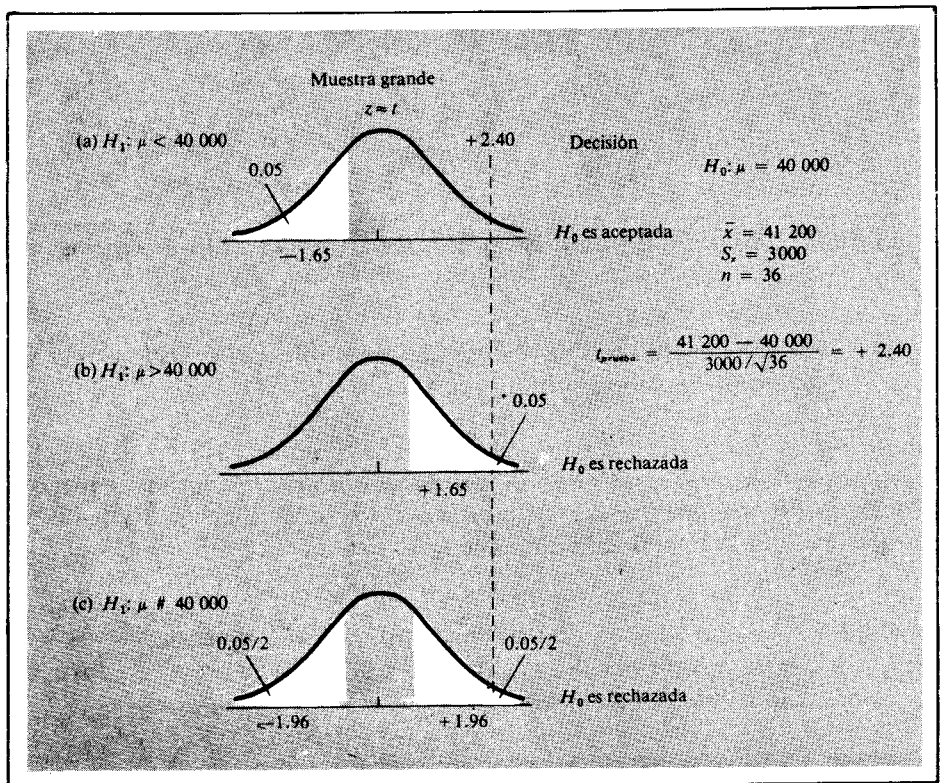


FIGURA 10.4 Comparación entre las tres alternativas en términos de la evaluación de los datos muestrales en el caso de una muestra grande.

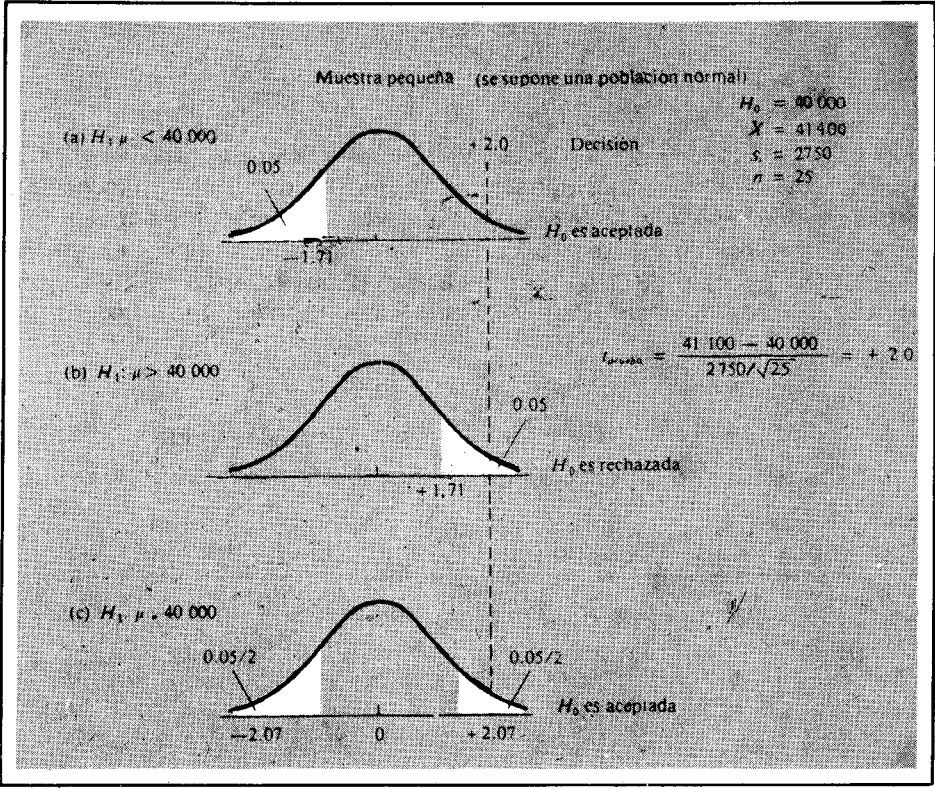


FIGURA 10.5 Para una muestra pequeña se ilustran tres pruebas de hipótesis alternativas.

var que el propósito de estas tres pruebas es ilustrar las diversas pruebas existentes; en la práctica, solamente se deberá utilizar *una* alternativa.

Ejemplo 2 Muestra pequeña: Cuando el tamaño de la muestra es 30 o menor, el valor de t se obtiene de una tabla t . Cabe observar que con tamaños pequeños de muestra, la población que se está muestreando debe estar distribuida normalmente o de lo contrario esta técnica no se podría emplear. Supóngase una muestra de $n = 25$ que produce

$$\bar{x} = 41\,100 \text{ millas} \quad s_x = 2\,750 \text{ millas}$$

Si se utiliza $\alpha = 0.05$ y $25 - 1 = 24$ grados de libertad, el valor crítico de t es $+1.71$, o bien, -1.71 para una prueba de una cola, y ± 2.07 para una prueba de dos colas. El valor estadístico de prueba es

$$t_{prueba} = \frac{\bar{x} - \mu_0}{s_x/\sqrt{n}} = \frac{41\,100 - 40\,000}{2\,750/\sqrt{25}} = +2.0$$

Las tres pruebas posibles y las conclusiones resultantes se ilustran en la figura 10.5. Una vez más, únicamente se deberá considerar una sola hipótesis alternativa como esa.

El valor estadístico para una prueba de medias de una muestra

a. Con σ_x conocida:

$$z = \frac{\text{valor medio de la muestra-media propuesta}}{\sigma_x/\sqrt{n}}$$

b. Con σ_x no conocida:

$$t = \frac{\text{valor medio de la muestra-media propuesta}}{s_x/\sqrt{n}}$$

EJERCICIOS

- En cada uno de los casos siguientes, decida si una prueba de una o dos colas es adecuada, y posteriormente trace una curva normal que ilustre cada prueba. Indique el área en la(s) cola(s).
 - $H_0: \mu = 10, H_1: \mu \neq 10, \alpha = 0.02$
 - $H_0: \mu = 0.037, H_1: \mu > 0.037, \alpha = 0.05$
 - $H_0: \mu = 3.2, H_1: \mu < 3.2, \alpha = 0.01$
 - $H_0: \mu = 17.45, H_1: \mu > 17.45, \alpha = 0.05$
 - $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2, \alpha = 0.02$
- Enuncie las hipótesis nula y alternativa para cada una de las siguientes situaciones.
 - Una organización que prueba productos duda sobre la aseveración de un fabricante, en la que establece que las pilas tienen una vida promedio de 25 horas de operación continua.
 - La tubería galvanizada debe tener una media de 2 pulgadas de diámetro para que sea aceptable.
 - No se pondrán en práctica las nuevas técnicas educativas a menos que demuestren que la razón promedio de aprendizaje aumentará en comparación con las técnicas que actualmente se están utilizando.
 - Una empacadora quiere evitar que los frascos de mermelada de 12 onzas se llenen arriba del límite.
 - La misma empacadora anterior quiere evitar también que esos mismos frascos de mermelada de 12 onzas queden llenos abajo del límite.
- Con referencia a las cuatro proposiciones anteriores de los incisos a-d del ejercicio 1, pruebe cada una en el nivel especificado con la información adicional que se proporciona a continuación. Establezca un intervalo de confianza $(1 - \alpha)$ para cualquier caso donde H_0 sea rechazada.

	a.	b.	c.	d.
Valor medio de la muestra	12.2	0.040	3.3	19.05
s_x	1.8	0.01	0.2	3.5
Tamaño de la muestra	13	81	25	50

4. Una agencia de empleos anuncia que los empleados que colocó en los últimos seis meses obtienen salarios que promedian \$ 9000 al mes. Una muestra aleatoria de ese grupo, tomada por una oficina gubernamental, obtiene un salario promedio de \$ 8000 y una desviación estándar de \$ 1000 sobre la base de 50 personas.
 - a. ¿Qué distribución de muestreo es la que *teóricamente* se puede utilizar correctamente? ¿Por qué?
 - b. ¿Qué distribución de muestreo se puede emplear para obtener una aproximación razonable?
 - c. Pruebe la aseveración de la agencia de empleos respecto a la alternativa de que el salario promedio es menor de \$ 9000, utilizando el nivel de significación de 0.05.
- ★ 5. Una compañía que vende tiras repelentes contra insectos asegura que su producto es eficaz, por lo menos durante 400 horas. Un análisis sobre nueve tiras seleccionadas aleatoriamente indicó un promedio de 380 horas.
 - a. Pruebe la aseveración de la compañía respecto a la alternativa de que el repelente sea eficaz menos de las 400 horas, a un nivel 0.01, si la desviación estándar de la muestra es 60 horas.
 - b. Repita el inciso a utilizando el conocimiento de que la desviación estándar de la población es de 90 horas.
 - c. ¿En cuáles de los incisos anteriores es necesario saber que la población es aproximadamente normal? ¿Por qué?
6. Nueve personas han seguido una dieta especial durante dos meses. En ese tiempo, sus pérdidas de peso individuales fueron de 1.2, 2.0, 1.0, 0.8, 1.1, 0.2, 0.5, 0.4 y 0.1 kgs. Pruebe la hipótesis nula de un promedio verdadero de pérdida de peso de 0 kilos contra la alternativa de una pérdida de peso mayor de 0, utilizando $\alpha = 0.01$. Suponga que la población es normal.
- ★ 7. Mediante un proceso de fabricación, se produce alambre de acero con una resistencia media a la tensión de 200 libras por pulgada cuadrada (psi). La desviación estándar del proceso es 20 psi. El ingeniero encargado del control de calidad quiere diseñar una prueba que le indique si ha habido o no un cambio en el promedio del proceso, utilizando un tamaño de muestra de 25 y un nivel de significación de $\alpha = 0.05$. Suponga que la población de la resistencia del alambre es aproximadamente normal.
 - a. Enuncie H_0 y H_1 para esta prueba.
 - b. ¿Para qué intervalo de resistencia del alambre, el proceso se considerará fuera de control (es decir, se concluirá que la media del proceso ha cambiado a partir de 200 psi)?
8. Un fabricante de automóviles asegura que sus autos de tipo familiar pueden resistir un choque de frente a una velocidad superior de 10 mph, si están equipados con un parachoques (defensa) con amortiguadores, cuyo costo de reparación es

de \$ 100.00 dólares o menos. Si una oficina investigadora independiente toma una muestra de 6 de esos autos, se observa que el promedio de costo de reparación es de \$ 150.00 por vehículo, con una desviación estándar muestral de \$ 30.00. Suponga que la distribución de los costos de reparación es aproximadamente normal.

- a. ¿Existe la evidencia suficiente en este momento para rechazar la afirmación de la empresa al nivel de 0.01?
 - b. ¿Es necesario saber que la población es aproximadamente normal? ¿Por qué?
9. Una compañía aseguradora empezará una extensa campaña publicitaria para vender seguros de vida si considera que el salario promedio obtenido por familia es menor de \$ 10 000. Una muestra aleatoria de 50 familias en esa área tiene un ingreso promedio de \$ 9600 y una desviación estándar de \$ 1000.
- a. Tomando como base la evidencia de muestra, ¿se aceptará o rechazará la afirmación al nivel 0.05?
 - b. ¿Puede la conclusión a la que se llegó ser errónea utilizando la evidencia de muestra? ¿De qué tipo de error se trataría? ¿Por qué?
10. Suponga que se cuenta con la siguiente información:

$$H_0: \text{media} = 75 \quad \alpha = 0.05$$

$$H_1: \text{media} > 75 \quad n = 64$$

$$\sigma_x = 8.0$$

- a. Antes de tomar una muestra, ¿qué es P (error de tipo I)?
 - b. Si el valor medio de la muestra es 76.0, ¿cuál es la probabilidad de un error de tipo I?
 - c. Si el valor medio de la muestra es 77.5, ¿cuál es la probabilidad de un error de tipo I?
11. En un esfuerzo por acelerar el tiempo que toma un calmante en entrar al torrente sanguíneo, un analista e investigador de fármacos agregó otro ingrediente a la fórmula normal. La fórmula original tenía un tiempo medio de 43 minutos. En 36 observaciones de la nueva mezcla, se obtuvo un tiempo medio de 42 minutos, con una desviación estándar de 6 minutos. Suponga que la distribución de tiempos es aproximadamente normal.
- a. Enuncie H_0 y H_1 .
 - b. ¿Qué se puede concluir con respecto a la eficacia del nuevo ingrediente al nivel 0.05?
 - c. ¿Cuál sería su respuesta en relación con la eficacia del nuevo ingrediente si se utiliza al nivel 0.01?
 - d. ¿Es necesario suponer que la población es normal? ¿Por qué?
12. El muestreo de una población finita requiere que la estimación del error estándar de la distribución de muestreo sea modificada mediante el factor de corrección $\sqrt{(N - n)/(N - 1)}$ si el tamaño de la muestra excede del 5% de la población. El gerente de crédito de una empresa considera que el saldo de crédito promedio de 400 clientes no excede de \$75.00. Evalúe lo anterior al nivel 0.05 para los siguientes casos:
- a. $n = 30$, $s_x = 5$, $\bar{x} = \$77$
 - b. $n = 50$, $s_x = 5.2$, $\bar{x} = \$76$

PRUEBAS DE MEDIAS DE DOS MUESTRAS

Las pruebas de dos muestras se utilizan para decidir si las medias de dos poblaciones son iguales. Se requieren dos muestras independientes, una de cada una de las dos poblaciones.* Considérese, por ejemplo, una compañía investigadora que experimenta con dos diferentes mezclas de pintura, para ver si se puede modificar el tiempo de secado de una pintura para uso doméstico. Cada mezcla es probada un determinado número de veces, y comparados posteriormente los tiempos medios de secado de las dos muestras. Una parece ser superior, ya que su tiempo medio de secado (muestra) es 30 minutos menor que el de la otra muestra.

Pero, ¿son realmente diferentes los tiempos medios de secado de las dos pinturas, o esta diferencia muestral es nada más la variación aleatoria que se espera, aun cuando las dos fórmulas presentan idénticos tiempos medios de secado? Una vez más, las diferencias *casuales* se deben distinguir de las diferencias *reales*.

Con frecuencia se utilizan pruebas de dos muestras para comparar dos métodos de enseñanza, dos marcas, dos ciudades, dos distritos escolares y otras cosas semejantes.

La hipótesis nula puede establecer que las dos poblaciones tienen medias iguales:

$$H_0: \mu_1 = \mu_2$$

Las alternativas pueden ser alguna de las siguientes:

$$H_1: \mu_1 \neq \mu_2 \quad H_1: \mu_1 > \mu_2 \quad H_1: \mu_1 < \mu_2$$

(Nota: $\mu_1 > \mu_2$ es equivalente a $\mu_2 < \mu_1$.)

La prueba se concentra en la diferencia relativa entre las medias de dos muestras, una de cada población. Esta diferencia se divide entre la desviación estándar de una distribución de muestreo. La desviación estándar se calcula suponiendo primero que H_0 es verdadera. En este caso, las dos muestras se pueden considerar como provenientes de la *misma* población, y mediante la combinación de las variancias de las dos poblaciones (o muestras, si no se conocen estas medidas de la población), se puede determinar la variancia de la población total. Cuando se conocen σ_1 y σ_2 , el valor estadístico de prueba es el siguiente:

$$z_{\text{prueba}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

* Cuidado: Estas pruebas de dos muestras requieren que las muestras sean independientes (es decir, que provengan de grupos diferentes). Esto significa que los datos previos y posteriores, como los que se pueden obtener de un estudio dietético, al examinar el peso promedio antes y después de la dieta del mismo grupo, no se pueden evaluar de esta manera. En el Capítulo 13 se presenta una prueba adecuada para casos de esta naturaleza.

Cabe suponer que el valor real de z , cuando H_0 es verdadera, está distribuido normalmente con una media de 0 y una desviación estándar de 1.0 (es decir, la distribución normal estandarizada) para casos en los que la suma $n_1 + n_2$ es mayor de 30. Para tamaños más pequeños de muestra, z estará distribuida normalmente sólo si las dos poblaciones que se muestrean también lo están.

Cuando no se conocen las desviaciones estándar de la población, el valor estadístico de prueba es como el que se presenta a continuación.

$$t_{\text{prueba}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

El valor de t , suponiendo que H_0 sea verdadera, se puede aproximar mediante z si $n_1 + n_2$ excede de 30.

Cuando los tamaños de las dos muestras no son iguales, y su suma es menor de 30, la fórmula para el valor estadístico de prueba se convierte en:

$$t_{\text{prueba}} \approx \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \right] \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

El valor de t cuando H_0 es verdadera, tiene una distribución t con $n_1 + n_2 - 2$ grados de libertad, si se puede suponer que ambas poblaciones son aproximadamente normales.

Los siguientes ejemplos ilustran pruebas de medias de dos muestras para a) muestras grandes en relación con muestras pequeñas y b) situaciones con desviaciones estándar conocidas y no conocidas.

Ejemplo 3 Se conoce σ_x . Utilice $\alpha = 0.05$

$$H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 > \mu_2$$

$$\bar{x} = 20 \quad \bar{x}_2 = 18$$

$$\sigma_1 = \sigma_2 = 3$$

$$n_1 = n_2 = 36$$

Calcule el valor estadístico de prueba:

$$z_{\text{prueba}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{20 - 18}{\sqrt{\frac{9}{36} + \frac{9}{36}}} = \frac{2}{\sqrt{\frac{18}{36}}} = \frac{2}{\sqrt{0.5}} = \frac{2}{0.7} = 2.9$$

Como el valor estadístico de prueba es mayor de 1.65, rechace H_0 y concluya que la media de la población 1 es mayor que la media de la población 2. Ver figura 10.6.

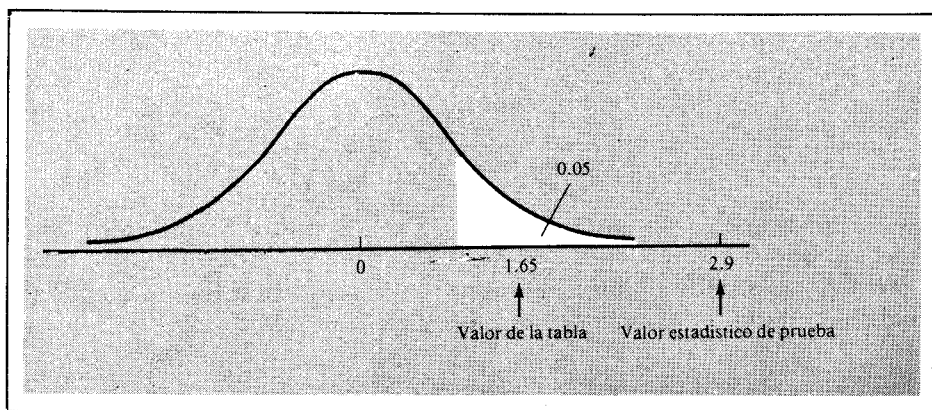


FIGURA 10.6 Se rechaza H_0 , ya que $z_{\text{prueba}} > z_{\text{tabla}}$.

Ejemplo 4 No se conoce σ_x . Pruebe la hipótesis

$$H_0: \mu_1 = \mu_2$$

cuando la alternativa es

$$H_1: \mu_1 \neq \mu_2$$

utilice $\alpha = 0.05$ y

$$s_{x_1} = 1.1 \quad s_{x_2} = 1.2$$

$$\bar{x}_1 = 5.4 \quad \bar{x}_2 = 5.0$$

a. Utilice $n_1 = n_2 = 36$.

b. Utilice $n_1 = 15$, $n_2 = 3$ (debemos ser capaces de suponer una población normal)

Solución:

a. El valor estadístico de prueba es

$$t_{\text{prueba}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_{x_1}^2}{n_1} + \frac{s_{x_2}^2}{n_2}}} = \frac{5.4 - 5.0}{\sqrt{\frac{1.21}{36} + \frac{1.44}{36}}} = \frac{.4}{\sqrt{0.2 + 0.24}} = \frac{0.4}{0.663} = .6$$

Como $n_1 + n_2 > 30$ $z \approx t$

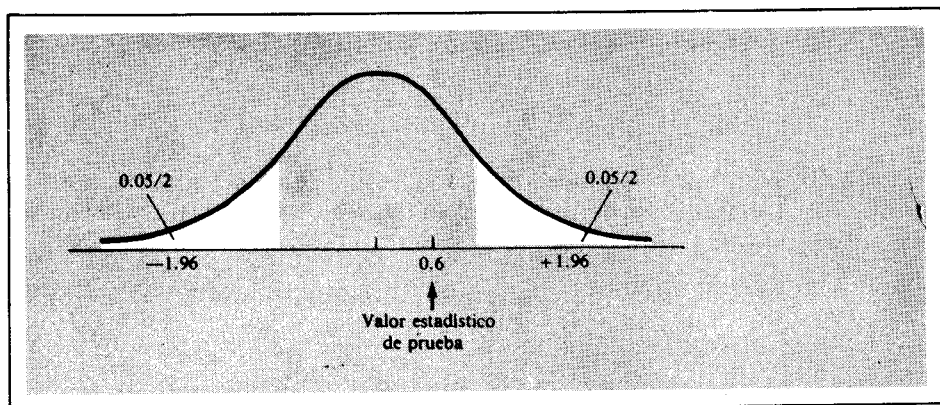


FIGURA: 10.7 Se acepta H_0 , dado que t_{prueba} se encuentra dentro del intervalo de aceptación.

Como el valor estadístico de prueba está en el intervalo de aceptación, no se puede rechazar la hipótesis nula (ver figura 10.7). Se concluye que la diferencia entre las dos medias de la muestra es probablemente el resultado de la variación casual debido al muestreo aleatorio.

b. El valor estadístico de prueba es

$$\begin{aligned}
 t_{prueba} &= \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \right] \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \\
 &= \frac{5.4 - 5.0}{\sqrt{\left[\frac{(14)1.1^2 + (2)1.2^2}{15 + 3 - 2} \right] \left(\frac{1}{15} + \frac{1}{3} \right)}} = \frac{0.4}{0.704} = 0.568
 \end{aligned}$$

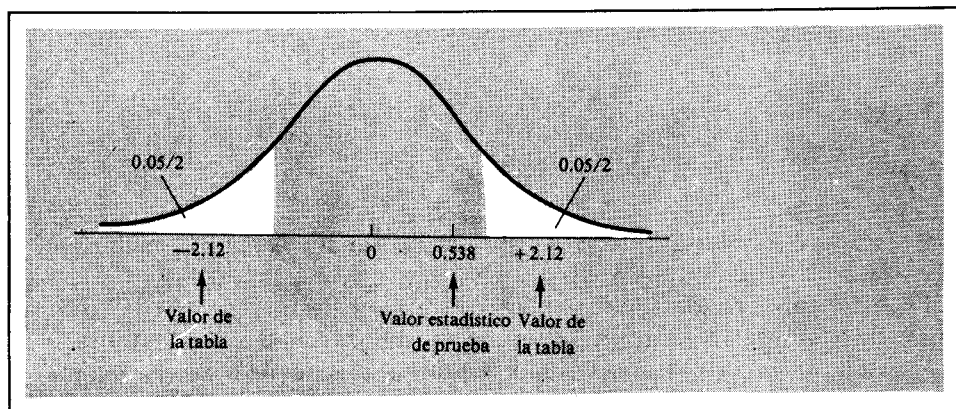


FIGURA: 10.8 H_0 es aceptada.

Cuando la suma de n_1 y n_2 es menor de 30, se debe utilizar una tabla de t y $n_1 + n_2 - 2$ grados de libertad para encontrar el valor crítico. De la tabla H se obtiene que

$$t_{0.025, 16 \text{ g.l.}} = 2.12$$

Como el valor estadístico de prueba cae en la región de aceptación, no se puede concluir que las dos poblaciones de las que se han obtenido las muestras tengan medias diferentes. Ver figura 10.8.

EJERCICIOS

- ★ 1. Suponga que se tienen los siguientes datos de muestras de dos poblaciones:
- | | Agencia A | Agencia B |
|---------------------------|-----------|-----------|
| Tiempo medio de respuesta | 4 hr | 5 hr |
| Desviación estándar | 1 hr | 1.2 hr |
| Número de observaciones | 30 | 24 |
- a. Calcule la estimación combinada de variancia para las dos muestras.
b. Utilizando el nivel de significación de 0.01, pruebe la afirmación de que las dos agencias presentan la misma razón media de respuesta.
c. Empleando el nivel de significación de 0.01, pruebe la proposición de que el tiempo medio de respuesta de la Agencia B es mayor que el de la Agencia A.
d. ¿Es necesario hacer una suposición de muestreo a partir de dos poblaciones normales? Explique brevemente la respuesta.
2. Dos encuestas independientes sobre salarios, realizadas en dos áreas metropolitanas muy distantes entre sí, revelaron la siguiente información con respecto a los sueldos promedio de operadores de equipo pesado:

	Area	
	A	B
$\bar{x}$	\$6.50/hr	\$7.00/hr
s_x	\$1.50/hr	\$1.00/hr
n	25	25

¿Se puede concluir que los sueldos promedio son diferentes al nivel de significación de 0.05?

3. Al comparar métodos de enseñanza, 12 niños de preprimaria del grupo de control armaron un rompecabezas en un tiempo medio de 3.2 minutos ($s_x = 30$ segundos). Los 10 niños del grupo de prueba, después de observar una película relacionada con el armado de rompecabezas, realizaron tareas semejantes en un promedio de 2.8 minutos, con $s_x = 30$.

- a. ¿Se deberá establecer esto como una prueba de una cola? ¿o de dos colas? ¿Por qué?
 - b. ¿Qué se puede concluir con respecto a la eficacia de la película, utilizando el nivel 0.05?
 - c. ¿Es necesario suponer que la población de tiempo es normal? ¿Por qué sí o por qué no?
 - d. ¿Es necesario suponer que los niños fueron asignados aleatoriamente a los grupos? Explíquelo.
4. Una gran cadena de tiendas está interesada en saber si la cantidad media de ventas es mayor en su local ubicado en el centro de la ciudad que en el que se encuentra en un gran centro comercial. Pruebe la proposición de que las ventas en ambos locales son iguales contra la alternativa de que no lo sean, utilizando el nivel 0.01. Una muestra aleatoria de transacciones comerciales en los dos locales proporciona los siguientes datos:

	Centro de la ciudad	Centro comercial
$\bar{x}$	\$ 45.00	\$ 43.50
Tamaño de la muestra	100	100

(Suponga que ambas desviaciones estándar de la población son de \$ 10.00.)

5. Utilizando el nivel de significación 0.05, determine si la resistencia a la tensión media en dos remesas de hilo de nilón son iguales, dadas estas observaciones muestrales.

Remesa L	Remesa N
30 psi	32 psi
28	33
27	31
28	30
32	29

CÁLCULO DE LA PROBABILIDAD DE UN ERROR DE TIPO II*

Se comete un error de tipo II, al aceptar la hipótesis nula cuando es falsa. A fin de calcular la probabilidad de que esto ocurra, es útil considerar la probabilidad como la cantidad que la distribución de muestreo, con base en la media real de población, superpone en la región de aceptación. Esto se ilustra en la figura 10.9.

El procedimiento para determinar la probabilidad de error de tipo II es como sigue (ver figura 10.10):

1. Establecer la región de aceptación para H_0 , utilizando 1) la media de la pobla-

*Esta sección es optativa y puede suprimirse sin que se pierda continuidad.

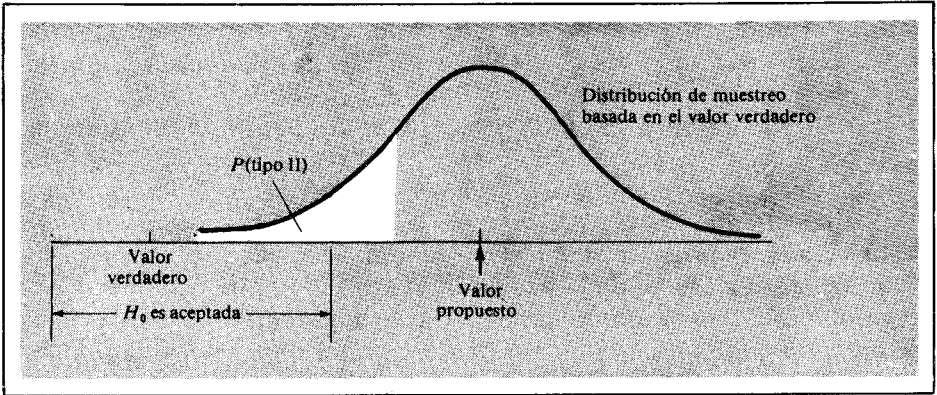


FIGURA 10.9 La probabilidad de un error de tipo II es la cantidad que la distribución de muestreo, con base en el parámetro real de la población, superpone en la región de aceptación para la hipótesis nula.

- ción propuesta (supuesta) y 2) los datos específicos del problema (es decir, la desviación estándar, el tamaño de la muestra, etc.) [figura 10.10 (a)].
2. Agregar una distribución de muestreo que tenga como base la media real de la población [figura 10.10 (b)].
 3. Utilizando el valor verdadero como punto de referencia, determinar el área [que se indica como Q en la figura 10.10 (c)] entre aquél y la regla de decisión.
 4. Sumar 50% o restar esta cantidad, para obtener $P(\text{tipo II})$, dependiendo de si el valor verdadero está dentro o fuera de la región de aceptación. En la figura 10.10 el valor verdadero está fuera de la región, por lo que el área Q es restada del 50%. El área Q se tendría que sumar al 50% si el valor verdadero hubiera estado dentro.

Ejemplo 5 Con base en la siguiente información, calcule la probabilidad de cometer un error de tipo II si la media real de una población es 14.88.

media supuesta	14.00
desviación estándar de la población	2.00
tamaño de la muestra	36
nivel de significación para una prueba de dos colas	10%

Solución:

Como se conoce σ_x , es apropiado utilizar la distribución normal. El valor z en el caso de un 10%, en una prueba de dos colas, es ± 1.65 .

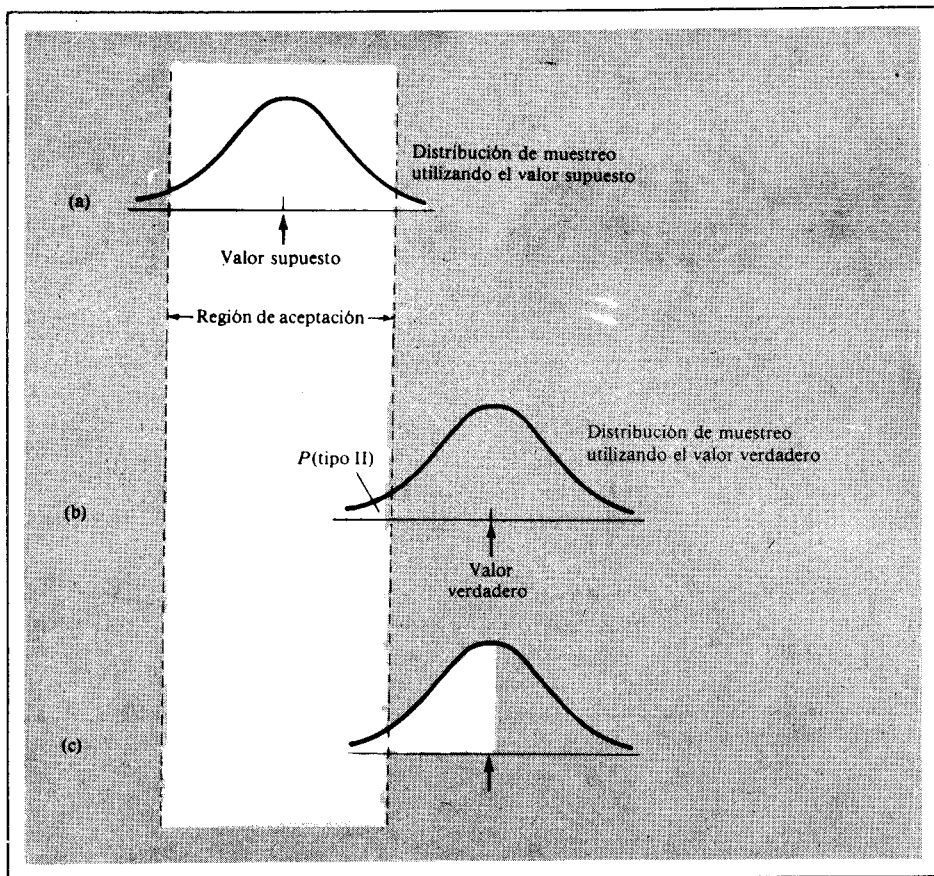


FIGURA: 10.10 Pasos para calcular la probabilidad de un error de tipo II.

Las reglas de decisión son

$$\begin{aligned}
 14.00 \pm z \frac{\sigma_x}{\sqrt{n}} &= 14.00 \pm 1.65 \frac{2}{\sqrt{36}} = 14.00 \pm 1.65 \left(\frac{1}{3} \right) \\
 &= 14.00 \pm 0.55 \quad \text{o} \quad 13.45 \text{ a } 14.55
 \end{aligned}$$

Ver la figura 10.11.

Si la media verdadera es 14.88, la probabilidad de cometer un error de tipo II sería semejante a la que se muestra en el diagrama que se presenta en la figura 10.12.

Para determinar el área en el extremo o cola (que es la probabilidad de caer en un error de tipo II), es necesario calcular cuán lejos está la regla de decisión (14.55) del valor verdadero:

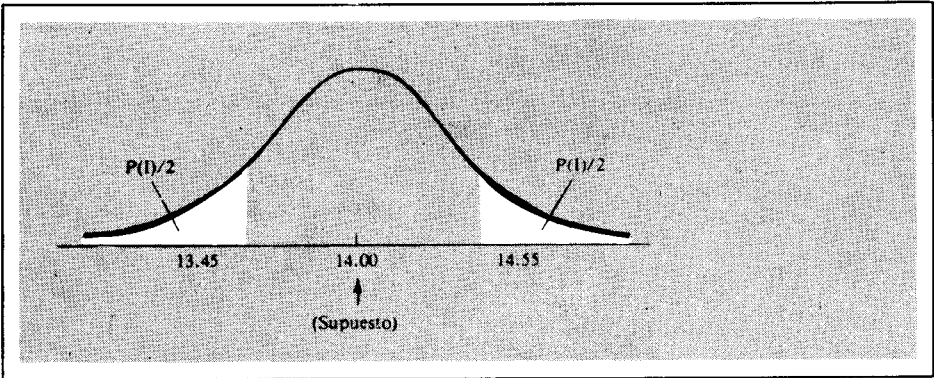


FIGURA 10.11 Regla de decisión

$$z = \frac{14.55 - 14.88}{2/\sqrt{36}} = \frac{-0.33}{0.33} = -1.0$$

El área bajo la curva normal entre la media y una desviación estándar de 1.0 es 0.3413. Por tanto, el área más allá de una desviación estándar de 1.0 es 0.5000 — 0.3413 = 0.1587; que es la probabilidad de un error de tipo II.

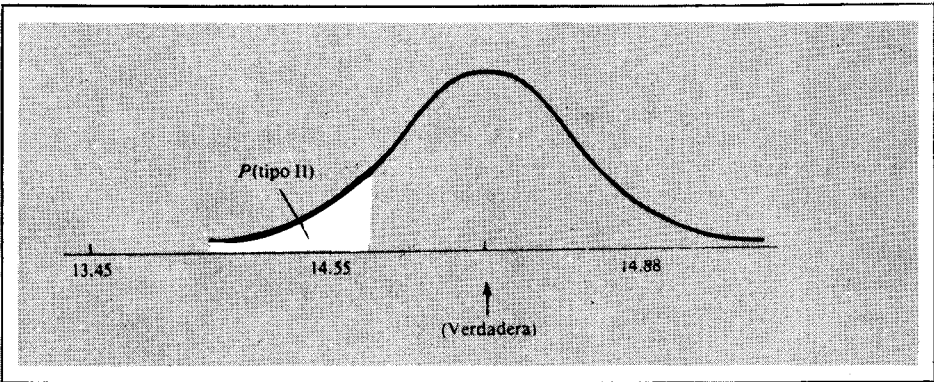


FIGURA 10.12 P(Error de tipo II) si la media verdadera es 14.88.

Ejemplo 6 Utilizando la misma información para el ejemplo 5, obtenga la probabilidad de cometer un error de tipo II si el valor verdadero es 14.33 en lugar de 14.88.

Solución:

La región de aceptación sería la misma que anteriormente se mencionó, pero ahora el valor verdadero se encuentra dentro de ésta (cabe observar que H_0 todavía es falsa, ya que el valor verdadero no es igual al valor propuesto). Ver figura 10.13.

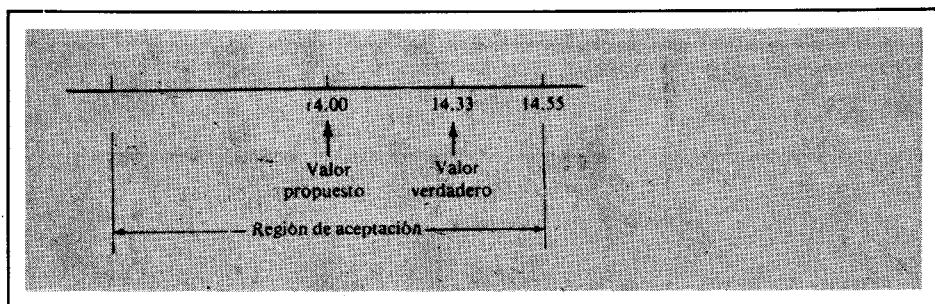


FIGURA 10.13 El valor verdadero se encuentra dentro de la región de aceptación.

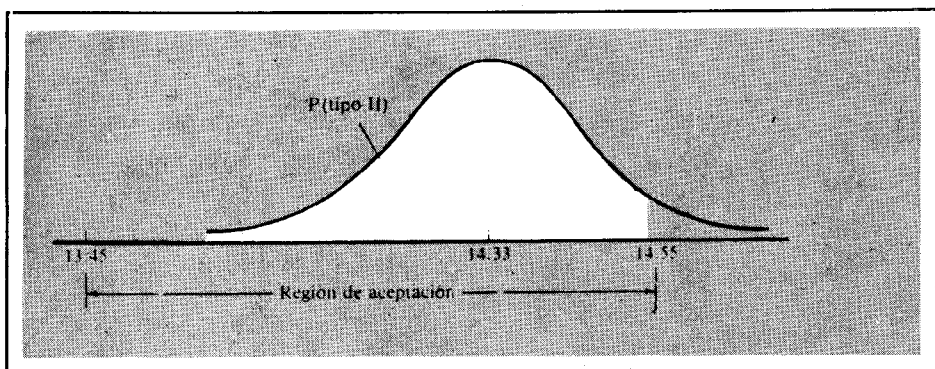


FIGURA 10.14 $P(\text{II})$ es grande con una media real de 14.33.

Agregando la distribución de muestreo cuando se utiliza la media real de 14.33 se obtiene la gráfica que se indica en la figura 10.14.

La probabilidad de un error de tipo II es evidentemente mayor del 50%, dado que más de la mitad de la distribución de muestreo está dentro de la región de aceptación. Establecer exactamente cuánto más excede del 50%, dependerá de cuán alejado esté 14.55 de la media verdadera:

$$z = \frac{14.55 - 14.33}{2/\sqrt{36}} = \frac{0.22}{0.33} = \frac{2}{3}$$

El área entre la media y $\frac{2}{3}$ de desviaciones estándar es aproximadamente 0.2486, por lo que la probabilidad de un error de tipo II es $0.5000 + 0.2486 = 0.7486$.

El tamaño de un error de tipo II depende de cuán falsa sea H_0 . Si es ligeramente falsa, la probabilidad de cometer un error de este tipo será mucho mayor que cuando H_0 sea muy falsa (es decir, que los valores supuestos y verdaderos estén muy alejados). En la figura 10.15 se muestra lo anterior. Cabe observar que, para calcular la probabilidad de cometer un error de tipo II, es necesario tener en mente un valor verdadero en particular, puesto que cada valor verdadero diferente producirá una $P(\text{error de tipo II})$ distinta.

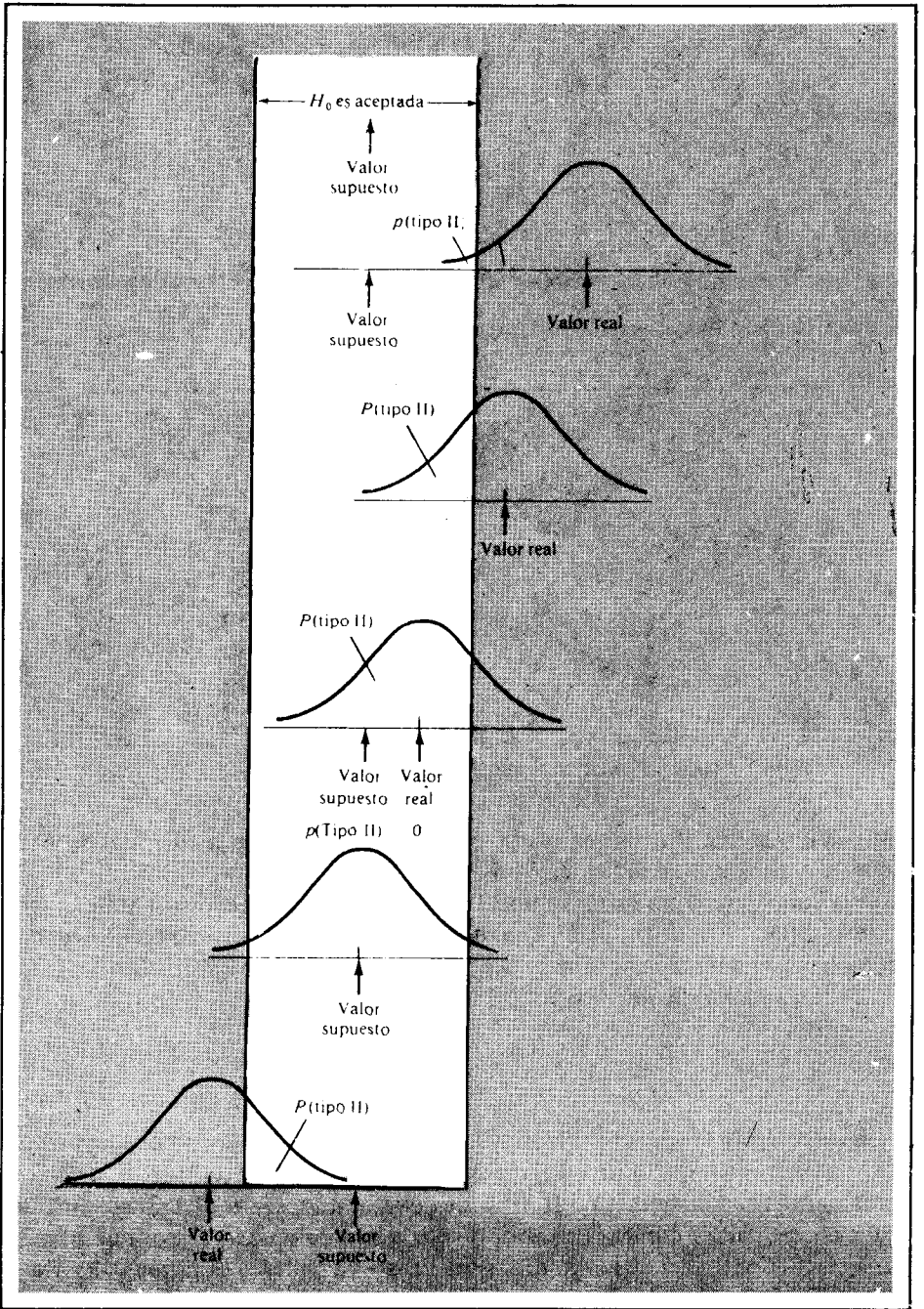


FIGURA 10.15 La probabilidad de un error de tipo II depende de cuán lejos de la media supuesta se encuentre la media verdadera (real).

Obsérvese en la figura 10.15 que en un caso la probabilidad de un error de tipo II es cero. La razón es que la media verdadera y la media supuesta son iguales, de manera que la hipótesis nula es verdadera. Cuando esto sucede, el único tipo de error posible es el error de tipo I.

Ejercicios

- Defina un error de tipo I.
 - Defina un error de tipo II.
- Suponga que se cuenta con la siguiente información:

$$\begin{array}{ll} H_0: \mu = 20 & H_1: \mu \neq 20 \\ n = 100 & \sigma_x = 10 \\ \alpha = 0.05 \end{array}$$

Obtenga P (tipo II) en los siguientes casos:

- $\mu = 19$
- $\mu = 19.5$
- $\mu = 20$
- $\mu = 22$

- Suponga que se tiene la siguiente información:

$$\begin{array}{ll} H_0: \mu = 5 & H_1: \mu > 5 \\ n = 25 & \sigma_x = 1 \\ \alpha = 0.04 \end{array}$$

Encuentre P (tipo II) si μ es como sigue:

- 5.2
- 5.0
- 5.4
- 6.0

- Suponga que se cuenta con la siguiente información:

$$\begin{array}{ll} H_0: \mu = 25 & H_1: \mu < 25 \\ n = 36 & s_x = 3 \\ \alpha = 0.01 \end{array}$$

Calcule P (tipo II) si μ es como sigue:

- 24.5
- 24
- 23.5

RESUMEN

La finalidad general de las pruebas de medias es evaluar las afirmaciones o proposiciones con respecto a las medias de poblaciones. Se utiliza una prueba de una muestra si la proposición es acerca de un valor medio de una población única; en tanto que se emplea una prueba de dos muestras si la afirmación establece que dos poblaciones tienen la misma media; y una prueba de muestra k cuando la proposición se relaciona con las medias de más de dos poblaciones. Las afirmaciones son evaluadas mediante 1) la suposición de que son verdaderas y, posteriormente 2) por medio de la elaboración

de una distribución de muestreo utilizando dicha proposición, la cual 3) puede ser utilizada para considerar determinado resultado de la muestra. Se considera que los valores medios de las muestras que dan como resultado pequeñas desviaciones a partir del valor esperado (propuesto) se cree que se deben a la variación casual en el muestreo; las grandes diferencias se toman como prueba de que la proposición es falsa. La desviación estándar de la distribución de muestreo es utilizada como medida de desviación respecto del valor esperado, y el nivel de significación (probabilidad de un error de tipo I) es empleado para determinar valores críticos que separan los resultados casuales de los no casuales.

TEMA DE REPASO

1. ¿Cuál es la finalidad general de una prueba de medias?
2. Mencione tres tipos de afirmaciones que se puedan hacer con respecto a las medias.
3. Esboze el procedimiento de examinación para las pruebas de significación.
4. ¿En qué condiciones es necesario suponer que la población que se está muestreando tiene una distribución aproximadamente normal?
5. ¿Cuándo se deberá utilizar la distribución t ?

EJERCICIOS SUPLEMENTARIOS

Supóngase que las muestras se obtienen de población normales en el caso de cada uno de los siguientes problemas.

1. Un proceso industrial produce piezas de cierto material que tienen una media de 40 cm cuando el proceso está bajo control. El proceso tiene una desviación estándar *conocida* de 2.2 cm, que permanece constante independientemente de si el proceso está o no bajo control. Suponiendo un nivel de significación de 0.05, encuentre los promedios máximo y mínimo de la muestra que indicarían que el proceso está bajo control para muestras de
 - a. 14
 - b. 22
- ★ 2. Determine si es igual el promedio de quejas por día en estas dos oficinas

										Totales
A	18	14	10	13	9	13	8	7	16	108
B	14	15	11	14	20	21	12	10	18	135

3. En un estudio comparativo del tiempo medio de escolaridad para una muestra aleatoria de 50 hombres y 50 mujeres en una industria, se obtuvieron los siguientes valores estadísticos de muestra:

Hombres	Mujeres
$\bar{x} = 3.2$ años	$\bar{x} = 3.7$ años
$s_x = 0.8$ años	$s_x = 0.9$ años