ESTADÍSTICA APLICADA A LA INVESTIGACIÓN CIENTÍFICA

CONTRASTE DE HIPÓTESIS

HIPÓTESIS ESTADÍSTICA

En el Tema 6 (Estimación), se ha visto como puede estimarse un parámetro a partir de los datos contenidos en una muestra. Puede encontrarse ya sea un solo número (estimador puntual) o un intervalo de valores posibles (intervalo de confianza). Sin embargo, muchos problemas de ingeniería y ciencia requieren que se tome una decisión entre aceptar o rechazar una proposición sobre algún parámetro. Esta proposición recibe el nombre de **hipótesis**, y el procedimiento de toma de decisión sobre la hipótesis se conoce como **contraste de hipótesis**. Este es uno de los aspectos más útiles de la inferencia estadística, puesto que muchos tipos de problemas de toma de decisiones, pruebas o experimentos en el mundo de la ingeniería, pueden formularse como problemas de prueba de hipótesis.

Definición 1: Una hipótesis estadística es una proposición o supuesto sobre los parámetros de una o más poblaciones.

Un ejemplo de una hipótesis es la afirmación $\mu = 0{,}75$, donde μ es el diámetro interno promedio real de cierto tipo de tubo de PVC. Otro ejemplo es la expresión $p < 0{,}10$, donde p es la proporción de tarjetas de circuitos defectuosas entre las tarjetas de circuitos producidas por cierto fabricante. Si μ_1 y μ_2 denotan las resistencias a la ruptura promedio verdaderas de dos tipos de hilo, una hipótesis es la afirmación de que $\mu_1 - \mu_2 = 0$, y otra es la expresión $\mu_1 - \mu_2 > 5$.

Puesto que se emplean distribuciones de probabilidad para representar poblaciones, también es posible considerar una hipótesis estadística como una proposición sobre la distribución de probabilidad de una variable aleatoria. En este caso, un ejemplo de una hipótesis es la afirmación de que la distancia de frenado en ciertas condiciones tiene una distribución normal. Este tipo de análisis abordaremos cuando desarrollemos, dentro de este mismo tema, el punto “Pruebas de bondad de ajuste”.

En cualquier problema de contraste de hipótesis, hay dos hipótesis contradictorias en consideración. Una hipótesis podría ser la afirmación $\mu = 0{,}75$ y la otra $\mu \neq 0{,}75$, o bien, los dos enunciados contradictorios podrían ser $p \geq 0{,}10$ y $p < 0{,}10$. El objetivo es decidir con base en la información muestral, cuál de las dos hipótesis es correcta. Hay una analogía familiar con esto en un juicio penal. Una afirmación es la aseveración de que el individuo acusado es inocente. En el sistema judicial, ésta es la afirmación que al principio se cree cierta. Sólo ante la evidencia contundente de lo contrario, el jurado debe rechazar esta afirmación a favor de la aseveración alternativa de que el acusado es culpable. En este sentido, la afirmación de inocencia es la hipótesis favorecida o protegida, y la carga de demostración se coloca en quienes creen en la afirmación alternativa.

De manera similar, al probar hipótesis estadísticas, el problema se formulará de modo que desde el principio se favorezca una de las afirmaciones. Esta afirmación favorecida al inicio, no se rechazará a menos que la evidencia muestral la contradiga y proporcione un fuerte respaldo a la aseveración alternativa.

Definición 2: La ***hipótesis nula***, denotada por H_0, es la afirmación que se supone al principio como cierta (la afirmación de "creencia previa").

En el tratamiento del contraste de hipótesis, H_0 siempre será expresada como una afirmación de igualdad. Si θ denota el parámetro de interés, la hipótesis nula tendrá la forma $H_0 : \theta = \theta_0$, donde θ_0 es un número especificado denominado *valor nulo* del parámetro (valor que se afirma para θ en la hipótesis nula).

Es importante recordar que las hipótesis siempre son proposiciones sobre la población o distribución bajo estudio, no proposiciones sobre la muestra. Por lo general, el valor del parámetro de la población especificado en la hipótesis nula se determina en una de tres maneras diferentes:

1. Puede ser resultado de la experiencia pasada o del conocimiento del proceso, entonces el objetivo de la prueba de hipótesis usualmente es determinar si ha cambiado el valor del parámetro.
2. Puede obtenerse a partir de alguna teoría o modelo que se relaciona con el proceso bajo estudio. En este caso, el objetivo de la prueba de hipótesis es verificar la teoría o modelo.

3. Cuando el valor del parámetro proviene de consideraciones externas, tales como las especificaciones de diseño o ingeniería, o de obligaciones contractuales. En esta situación, el objetivo usual de la prueba de hipótesis es probar el cumplimiento de las especificaciones.

Definición 3: La ***hipótesis alternativa***, denotada por H_1, es la aseveración de que es contradictoria a H_0.

Es una afirmación sobre algún parámetro de una determinada distribución, cuyo valor representa una situación nueva que se sospecha puede ocurrir. La hipótesis alternativa tendrá la forma $H_1 : \theta = \theta_1$ y se asemejará a una de las tres afirmaciones siguientes: (1) $H_1 : \theta_1 > \theta_0$ (en cuyo caso la hipótesis nula implícita es $\theta \leq \theta_0$), (2) $H_1 : \theta_1 < \theta_0$ (así la hipótesis nula implícita establece que $\theta \geq \theta_0$), o bien (3) $H_1 : \theta_1 \neq \theta_0$ (donde la hipótesis nula implícita es $\theta = \theta_0$).

La hipótesis nula sólo se rechaza si la evidencia muestral hace pensar que H_0 es falsa. Si la muestra no contradice de forma contundente a H_0, se continua con la creencia de que la hipótesis nula es verdadera. Las dos conclusiones posibles de un análisis de prueba de hipótesis son entonces *rechazar* H_0 ó *no rechazar* H_0.

Un procedimiento de prueba, con base en datos muestrales, para decidir si se rechaza H_0, recibe el nombre de **contraste de hipótesis**. Si esta información es consistente con la hipótesis, se concluye que ésta es verdadera; sin embargo si esta información es inconsistente con la hipótesis, se concluye que esta es falsa. Debe hacerse hincapié en que la verdad o falsedad de una hipótesis en particular nunca puede conocerse con certidumbre, a menos que pueda examinarse a toda la población. Usualmente esto es imposible en muchas situaciones prácticas. Por tanto, es necesario desarrollar un procedimiento de prueba de hipótesis teniendo en cuenta la probabilidad de llegar a una conclusión equivocada.

Por ejemplo, se analiza la rapidez de combustión de un agente propulsor sólido utilizado en los sistemas de salida de emergencia para la tripulación de aeronaves. El interés se centra sobre la rapidez de combustión promedio. De manera específica, la atención recae en decir si la rapidez de combustión promedio es o no 50 cm/s. Esto puede expresarse de manera formal como:

$$H_0 : \mu = 50 \text{ cm/s}$$

$$H_1 : \mu \neq 50 \text{ cm/s}$$

La proposición $H_0 : \mu = 50$ cm/s, se conoce como ***hipótesis nula*** , mientras que la proposición $H_1 : \mu \neq 50$ cm/s, recibe el nombre de ***hipótesis alternativa***. Puesto que la hipótesis alternativa especifica valores de μ que pueden ser mayores o menores que 50 cm/s, también se conoce como ***hipótesis alternativa bilateral***. En algunas situaciones, lo que se desea es formular una ***hipótesis alternativa unilateral***, como en

$$H_0 : \mu = 50 \text{ cm/s} \qquad\qquad H_0 : \mu = 50 \text{ cm/s}$$

o

$$H_1 : \mu < 50 \text{ cm/s} \qquad\qquad H_1 : \mu > 50 \text{ cm/s}$$

Los procedimientos de prueba de hipótesis dependen del empleo de la información contenida en la muestra aleatoria de la población de interés. Si esta información es consistente con la hipótesis, se concluye que ésta es verdadera; sin embargo si esta información es inconsistente con la hipótesis, se concluye que ella es falsa.

La estructura de los problemas de prueba de hipótesis es idéntica en todas las aplicaciones que se han de considerar en este curso. Este procedimiento involucra la toma de una muestra aleatoria, el cálculo de una ***estadística de prueba*** a partir de los datos muestrales, y luego el uso de esta estadística para tomar una decisión sobre la hipótesis nula.

CONTRASTE DE UNA HIPÓTESIS ESTADÍSTICA

Para ilustrar los conceptos generales, considere el problema de la rapidez de combustión del agente propulsor presentado con anterioridad. La hipótesis nula establecida es que la rapidez promedio de combustión es 50 cm/s, mientras que la hipótesis alternativa es que ésta no es igual a 50 cm/s. Esto es, se desea probar:

$$H_0 : \mu = 50 \text{ cm/s}$$

$$H_1 : \mu \neq 50 \text{ cm/s}$$

Supóngase que se realiza una prueba sobre una muestra de 10 especimenes, y que se observa cual es la rapidez de combustión promedio muestral. La media muestral es un estimador de la media verdadera de la población μ. Un valor de la media muestral $\bar{x}$ que esté próximo al valor hipotético μ = 50 cm/s es una evidencia de que el verdadero valor de la media μ es realmente 50 cm/s; esto es, tal evidencia apoya la hipótesis nula H_0. Por otra parte, una media muestral muy diferente de 50 cm/s constituye una evidencia que apoya la hipótesis alternativa H_1. Por tanto, en este caso, la media muestral $\bar{x}$ es una ***estadística de prueba***.

La media muestral puede tomar muchos valores diferentes. Supóngase que si $48{,}5 \leq \bar{x} \leq 51{,}5$; entonces no se rechaza la hipótesis nula $H_0 : \mu$ = 50 cm/s, y que si $\bar{x} < 48{,}5$ ó $\bar{x} > 51{,}5$; entonces se rechaza la hipótesis nula $H_0 : \mu$ = 50 cm/s.

Los valores de $\bar{x}$ que son menores que 48,5 o mayores que 51,5 constituyen la ***región de rechazo*** de la prueba, mientras que todos los valores que están en el intervalo $48{,}5 \leq \bar{x} \leq 51{,}5$ forman la ***región de No rechazo***. Las fronteras entre las regiones crítica y de no rechazo reciben el nombre de ***valores críticos.*** En este ejemplo, los valores críticos son 48,5 y 51,5. Esto se ilustra en la Figura 1.

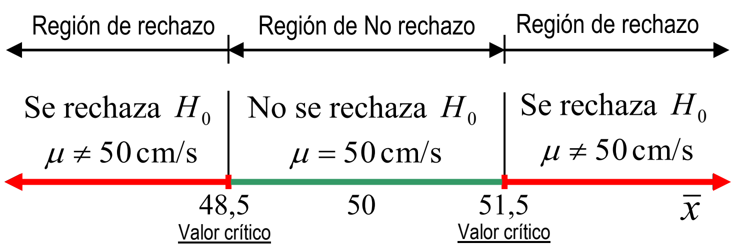

Figura 1: Criterios de decisión para la prueba de $H_0 : \mu = 50$ cm/s contra $H_1 : \mu \neq 50$ cm/s

Este procedimiento tiene dos constituyentes: 1) una *estadística de prueba* o función de los datos muestrales utilizada para tomar una decisión y 2) una *región de rechazo* que consiste en los valores de x para los cuales H_0 será rechazada.

Un procedimiento de prueba se especifica mediante lo siguiente:

1. Una ***estadística de prueba***, una función de los datos muestrales en la que se basará la decisión (rechazar H_0).
2. Una ***región de rechazo***, el conjunto de los valores de la estadística de prueba para los que se rechazará H_0.

Entonces, la hipótesis nula será rechazada si y sólo si el valor de la estadística de prueba observada o calculada cae en la región de rechazo.

Errores en el contraste de hipótesis

La base para elegir una determinada región de rechazo radica en entender los errores que uno podría enfrentar al sacar una conclusión. La costumbre es establecer conclusiones con respecto a la hipótesis nula H_0. Por tanto, se rechaza H_0 si el estadístico de prueba cae en la región de rechazo, de lo contrario, no se rechaza H_0.

Este procedimiento de decisión puede conducir a una de dos conclusiones erróneas. Por ejemplo, es posible que el valor verdadero de la rapidez promedio de combustión del agente propulsor sea igual a 50 cm/s. Sin embargo, para todos los especimenes bajo prueba, bien puede observarse un valor del estadístico de prueba $\bar{x}$ que cae en la región de rechazo. En este caso, la hipótesis nula H_0 será rechazada, pero de hecho, H_0 en realidad es verdadera. Este tipo de conclusión equivocada se conoce como ***error tipo I***.

Definición 4: El ***error tipo I*** consiste en rechazar hipótesis nula H_0 cuando es verdadera.

Ahora supóngase que la verdadera rapidez promedio de combustión es diferente de 50 cm/s, aunque la media muestral $\bar{x}$ caiga dentro de la región de no rechazo. En este caso no se rechaza H_0 cuando ésta es falsa. Este tipo de conclusión recibe el nombre de ***error tipo II***.

Definición 5: El ***error tipo II*** requiere no rechazar la hipótesis nula H_0 cuando ésta es falsa.

En el mejor de los casos, se podrían desarrollar procedimientos de prueba para los cuales ningún tipo de error es posible. Sin embargo éste ideal se puede lograr sólo si una decisión se basa en un examen de la población completa, lo cual casi siempre resulta irrealizable. La dificultad con el uso de un procedimiento basado en datos muestrales, es que debido a la variabilidad del muestreo, podría resultar una muestra no representativa. Por tanto, al probar cualquier hipótesis estadística, existen cuatro situaciones diferentes que determinan si la decisión final es correcta o errónea. Las decisiones asociadas a las condiciones de H_0 se dan en la tabla 1.

Condición \ Decisión	No rechazar H_0	Rechazar H_0
H_0 es verdadera	**Decisión correcta**	**Error tipo I**
H_0 es falsa	**Error tipo II**	**Decisión correcta**

Tabla 1: Decisiones en el contraste de hipótesis

En vez de demandar procedimientos sin errores, se deben considerar procedimientos para los que sea poco probable que ocurra cualquier tipo de error. Es decir, un buen procedimiento es aquel para el que la probabilidad de cometer cualquier tipo de error es pequeña. En general, se fija la probabilidad de cometer el error tipo I y se trata de minimizar la probabilidad de cometer el error tipo II. La elección de un valor de corte particular para una región de rechazo fija las probabilidades de los errores tipo I y II. Estas probabilidades de error se denotan por lo común con α y β, respectivamente. Debido a que H_0 especifica un valor único del parámetro, hay un solo valor de α. Sin embargo, hay un valor de β diferente para cada valor del parámetro consistente con H_1.

La P(cometer error tipo I) = P(rechazar H_0/H_0 es cierta) = α , se denomina ***nivel de significancia*** de la prueba, alguna veces ***tamaño de la región crítica***.

La P(cometer error tipo II) = P(no rechazar H_0/H_0 es falsa) = β

Se debe destacar que α y β no son complementarios, salvo el caso teórico en que $H_0 = H_1$.

Trabajando con el ejemplo de la rapidez promedio de combustión del agente propulsor, se presenta el error tipo I cuando $\bar{x}$ < 48,5 ó $\bar{x}$ > 51,5 y la verdadera rapidez promedio de combustión es μ = 50 cm/s. Suponiendo que la desviación estándar de la rapidez de combustión es $\sigma = 2{,}5$ cm/s, y que la rapidez de combustión tiene una distribución para la que se aplican las condiciones del teorema central del límite, de modo que la distribución de la media muestral es aproximadamente normal con media μ = 50 cm/s y desviación estándar $\sigma/\sqrt{n} = 2{,}5/\sqrt{10} = 0{,}79$. La probabilidad de cometer un error tipo I es igual a la suma de las áreas que aparecen con color azul en las colas de la de la distribución normal de la figura 2. Esta probabilidad puede calcularse de la siguiente manera:

$$\alpha = P(\bar{X} < 48{,}5 \mid \mu = 50) + P(\bar{X} > 51{,}5 \mid \mu = 50)$$

Los valores de z que corresponden a los valores críticos 48,5 y 51,5 son

$$z_1 = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{48{,}5-50}{0{,}79} = -1{,}90 \quad \text{y} \quad z_2 = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{51{,}5-50}{0{,}79} = 1{,}90$$

Por lo tanto,

$$\alpha = P(Z < -1{,}90) + P(Z > 1{,}90)$$

$$\alpha = 0{,}0288 + 0{,}0288$$

$$\alpha = 0{,}0576$$

Esto implica que el 5,76% de todas las muestras aleatorias conducirán al rechazo de la hipótesis $H_0 : \mu = 50$ cm/s cuando la verdadera rapidez promedio de combustión es en realidad 50 cm/s.

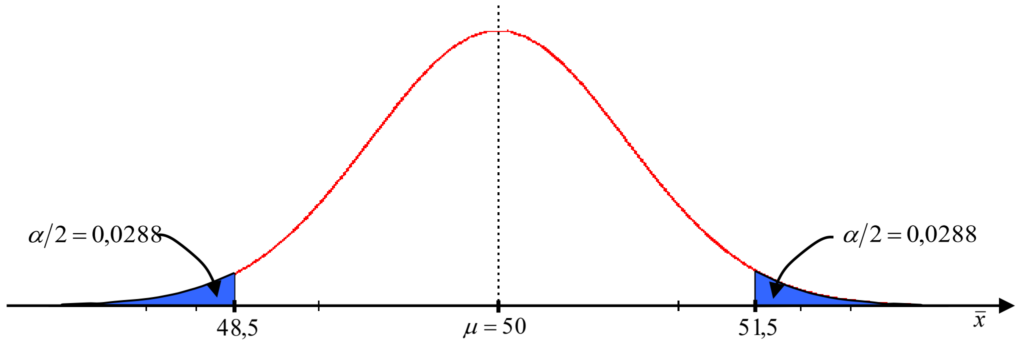

Figura 2: Región critica para $H_0 : \mu = 50$ cm/s contra $H_1 : \mu \neq 50$ cm/s y $n = 10$

Al analizar la Figura 2, se observa que es posible reducir α al aumentar la región de aceptación. Por ejemplo, si se toman como valores críticos 48 y 52, el valor de α es

$$\alpha = P\left(Z < \frac{48-50}{0{,}79}\right) + P\left(Z > \frac{52-50}{0{,}79}\right)$$

$$\alpha = P(Z < -2{,}53) + P(Z > 2{,}53)$$

$$\alpha = 0{,}0057 + 0{,}0057$$

$$\alpha = 0{,}0114$$

También se puede reducir el valor de α mediante el incremento del tamaño de la muestra. Si $n = 16$, entonces $\sigma/\sqrt{n} = 2{,}5/\sqrt{16} = 0{,}625$ y, al utilizar la región crítica original de la Figura 2, se tiene que

$$\alpha = P\left(Z < \frac{48{,}5-50}{0{,}625}\right) + P\left(Z > \frac{51{,}5-50}{0{,}625}\right)$$

$$\alpha = P(Z < -2{,}40) + P(Z > 2{,}40)$$

$$\alpha = 0{,}0082 + 0{,}0082$$

$$\alpha = 0{,}0164$$

En el procedimiento de prueba de hipótesis, también es importante analizar la probabilidad del error tipo II, el cual se denota por β. Esto es,

$$\beta = P(\text{error tipo II}) = P(\text{aceptar } H_0 \mid H_0 \text{ es falsa})$$

Para calcular β se debe tener una hipótesis alternativa específica; esto es, debe tenerse un valor particular de μ.Por ejemplo, supóngase que es importante rechazar la hipótesis nula $H_0 : \mu = 50$ cada vez que la rapidez promedio de combustión μ es mayor que 52 cm/s. Para ello, puede calcularse la probabilidad β

de un error tipo II para el valor $\mu = 52$, o sea, encontrar la probabilidad de no rechazar la hipótesis nula $H_0 : \mu = 50$ cm/s cuando el valor verdadero es $\mu = 52$ cm/s.

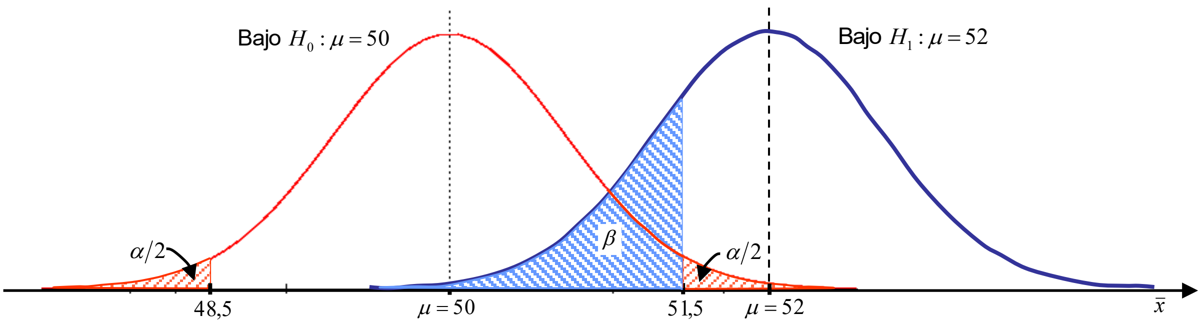

Figura 3: Probabilidad del error tipo II cuando $\mu = 52$ y $n = 10$

La figura 3 resulta de gran ayuda para calcular e interpretar la probabilidad β de un error de tipo II. La distribución normal de la parte izquierda de la figura 3 es la distribución del estadístico de prueba $\bar{X}$ cuando la hipótesis nula $H_0 : \mu = 50$ cm/s es verdadera (esto es lo que se entiende con la expresión "bajo $H_0 : \mu = 50$"), y la distribución normal de de la derecha es la distribución de $\bar{X}$ cuando la hipótesis alternativa es verdadera y el valor de la media es 52 (o "bajo $H_1 : \mu = 52$"). Se comete un error tipo II si la media muestral $\bar{x}$ cae entre 48,5 y 51,5 (que son las fronteras de la región crítica) cuando $\mu = 52$. Como puede verse en la figura 3, esto es sólo la probabilidad de que $48{,}5 \leq \bar{x} \leq 51{,}5$ cuando la media verdadera es $\mu = 52$, lo que corresponde al área rayada en color azul bajo la distribución normal de la derecha. Por lo tanto, con respecto a la figura 3, se tiene

$$\beta = P(48{,}5 \leq \bar{X} \leq 51{,}5 \mid \mu = 52)$$

Los valores de z que corresponden a 48,4 y 51,5 cuando $\mu = 52$ son

$$z_1 = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{48{,}5 - 52}{0{,}79} = -4{,}43 \quad \text{y} \quad z_2 = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{51{,}5 - 52}{0{,}79} = -0{,}63$$

Por tanto,

$$\beta = P(-4{,}43 \leq Z \leq -0{,}63)$$

$$\beta = P(Z \leq -0{,}63) - P(Z \leq -4{,}43)$$

$$\beta = 0{,}2643 - 0{,}000$$

$$\beta = 0{,}2643$$

En consecuencia, si se prueba $H_0 : \mu = 50$ contra $H_1 : \mu = 52$ con $n = 10$, y el valor verdadera de la media es $\mu = 52$, la probabilidad de que se acepte la hipótesis nula falsa es 0,2643.

La probabilidad β de cometer error tipo II aumenta rápidamente a medida que el valor verdadero μ tiende al valor hipotético. Por ejemplo, véase la figura 4, donde el valor verdadero de la media es $\mu = 50{,}5$ y el valor hipotético es $H_0 : \mu = 50$. El valor verdadero de μ está muy próximo a 50, y el valor de β es,

$$\beta = P(48{,}5 \leq \overline{X} \leq 51{,}5 \mid \mu = 50{,}5) = 0{,}8923$$

Por tanto, la probabilidad del error tipo II es mucho mayor para el caso donde la media verdadera es 50,5 cm/s, que para el caso donde ésta es 52 cm/s. Claro está que, en las situaciones prácticas el interés recae en detectar diferencias grandes entre la media verdadera y el valor especificado en la hipótesis nula, y no en analizar el error tipo II si la media está "próxima" al valor hipotético de ésta.

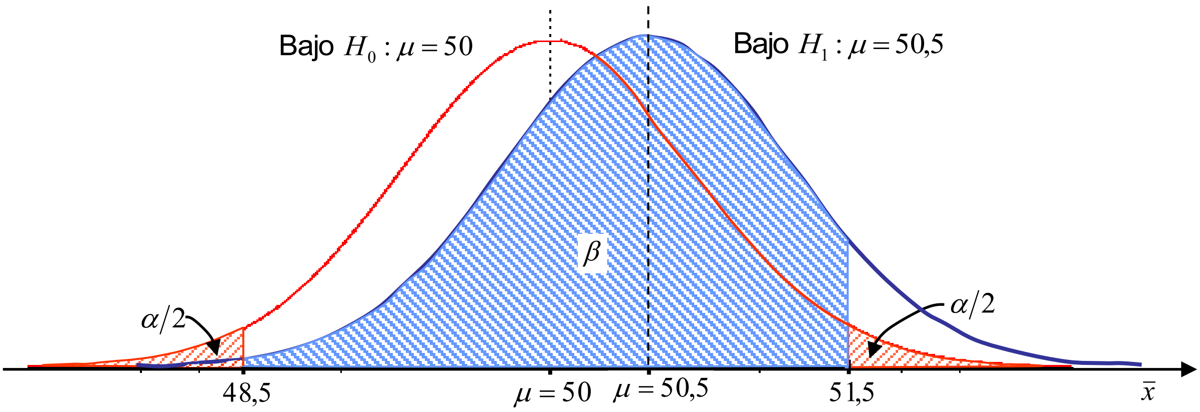

Figura 4: Probabilidad del error tipo II cuando $\mu = 50{,}5$ y $n = 10$

La probabilidad del error tipo II también depende del tamaño de la muestra n. Supóngase que la hipótesis nula es $H_0 : \mu = 50$cm/s y que el valor verdadero de la media es $\mu = 52$. Si el tamaño de la muestra aumenta de $n = 10$ a $n = 16$, entonces la probabilidad del error tipo II es

$$\beta = P(48{,}5 \leq \overline{X} \leq 51{,}5 \mid \mu = 52)$$

Cunado $n = 16$, la desviación estándar de $\overline{X}$ es $\sigma/\sqrt{n} = 2{,}5/\sqrt{16} = 0{,}625$ y los valores de z que corresponden a 48,5 y 51,5 cuando $\mu = 52$ son

$$z_1 = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{48{,}5 - 52}{0{,}625} = -5{,}60 \quad \text{y} \quad z_2 = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{51{,}5 - 52}{0{,}625} = -0{,}80$$

Por tanto,

$$\beta = P(-5{,}60 \leq Z \leq -0{,}80)$$
$$\beta = P(Z \leq -0{,}80) - P(Z \leq -5{,}60)$$
$$\beta = 0{,}2119 - 0{,}000$$
$$\beta = 0{,}2119$$

Se tiene entonces que, el aumento en el tamaño de la muestra da como resultado una disminución en la probabilidad del error tipo II.

Se pueden establecer algunas consideraciones respecto de los últimos conceptos presentados:

1. El tamaño de la región crítica y, en consecuencia, la probabilidad α de un error tipo I, siempre pueden reducirse mediante una selección apropiada de los valores críticos.
2. Los errores tipo I y tipo II están relacionados. Una disminución en la probabilidad en un tipo de error siempre da como resultado un aumento en la probabilidad del otro, siempre y cuando el tamaño de la muestra n no cambie.
3. En general, un aumento en el tamaño muestral n reducirá α y β, siempre y cuando los valores críticos se mantengan constantes.
4. Si la hipótesis nula es falsa, β es un máximo cuando el valor real del parámetro se aproxima al hipotético propuesto por la hipótesis nula. El valor de β disminuye a medida que aumenta la diferencia entre el verdadero valor medio y el hipotético propuesto.

Puesto que en general, el analista puede controlar de manera directa la probabilidad de rechazar de manera errónea H_0, siempre puede considerarse el rechazo de la hipótesis nula H_0 como una ***conclusión fuerte***.

Por otra parte, dado que la probabilidad β del error tipo II depende tanto del tamaño de muestra como del punto en el cual la hipótesis nula H_0 es falsa, es costumbre considerar la decisión de no rechazar H_0 como una ***conclusión débil***, a menos que se sepa que β es aceptablemente pequeño.

Definición 6: La ***Potencia*** de una prueba estadística es la probabilidad de rechazar la hipótesis nula H_0 cuando es falsa.

El valor de la potencia es "$1-\beta$", y la potencia puede interpretarse como la *probabilidad de rechazar de manera correcta un hipótesis nula falsa*. A menudo las pruebas estadísticas se analizan comparando sus propiedades de potencia. En nuestro ejemplo de la rapidez de combustión del agente propulsor cuando se prueba $H_0 : \mu = 50$ contra $H_1 : \mu \neq 50$ con $n = 10$, y suponiendo que el verdadero valor de la media es $\mu = 52$, se tiene que $\beta = 0{,}2643$ de modo que la potencia de la prueba es $1-\beta = 1-0{,}2643 = 0{,}7357$.

La potencia es una medida muy descriptiva y concisa de la sensibilidad de una prueba estadística, donde por sensibilidad se entiende la capacidad de una prueba para detectar diferencias. En este caso, la sensibilidad de la prueba para detectar la diferencia entre una rapidez promedio de combustión de 50 cm/s y otra de 52 cm/s, es 0,7357. Esto es, si el valor verdadero de la media es en realidad 52 cm/s, esta prueba rechazará de manera correcta $H_0 : \mu = 50$ y "detectará" esta diferencia el 73,57% de las veces. Si se piensa que el valor de potencia es bajo, entonces se puede aumentar α o el tamaño de la muestra n. En general, puede considerarse como aceptable un valor de potencia de entre 0,8 ó 0,85 en adelante.

Tipos de Ensayo

A continuación se presentan los tres tipos de pruebas de hipótesis que se pueden dar, a saber:

- Unilateral Derecho
- Unilateral Izquierdo
- Bilateral

Unilateral Derecho

Se desea comprobar la hipótesis de un aumento en el parámetro, en este caso el nivel de significancia se carga todo hacia el lado derecho, para definir las regiones de aceptación y de rechazo.

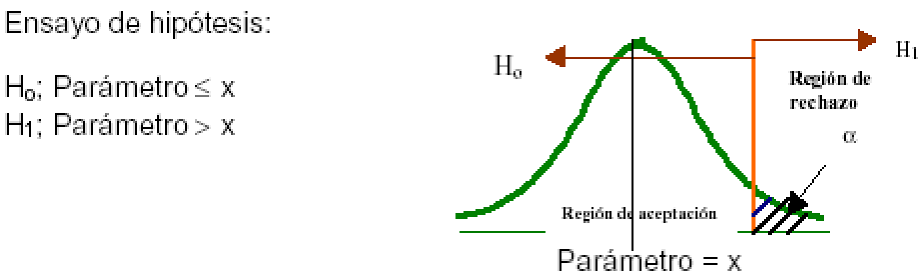

Figura 5: Ensayo unilateral derecho

Unilateral Izquierdo

Se desea comprobar la hipótesis de una disminución en el parámetro, en este caso el nivel de significancia se carga todo hacia el lado izquierdo, para definir las regiones de aceptación y de rechazo.

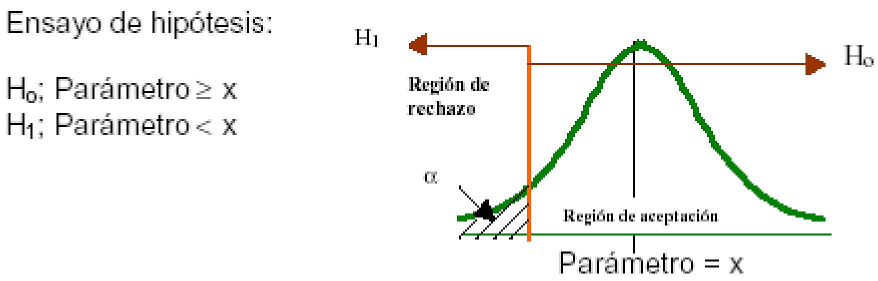

Figura 6: Ensayo unilateral izquierdo

Bilateral

Se desea comprobar la hipótesis de un cambio en el parámetro. El nivel de significancia se divide en dos y existen dos regiones de rechazo.

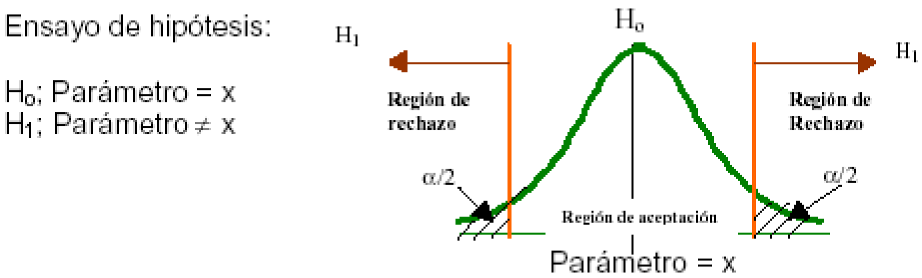

Figura 7: Ensayo bilateral

Pasos generales para la prueba de hipótesis

Se recomienda utilizar los siguientes pasos al aplicar el procedimiento de prueba de hipótesis.

1. Del contexto del problema, identificar el parámetro de interés.
2. Establecer la hipótesis nula, H_0.
3. Especificar una apropiada hipótesis alternativa, H_1.
4. Seleccionar un nivel de significancia α.
5. Establecer un estadístico de prueba apropiado.
6. Establecer la región de rechazo para el estadístico.
7. Calcular todas las cantidades muestrales necesarias, sustituirlas en la ecuación para el estadístico de prueba, y calcular el valor correspondiente.
8. Decidir si debe o no rechazar H_0 y notificar esto en el contexto del problema.

PRUEBAS RELACIONADAS CON UNA MEDIA POBLACIONAL

Caso 1: población normal con σ^2 conocida

Aunque la suposición de que el valor de σ^2 rara vez se cumple en la práctica, este caso proporciona un buen punto de partida debido a la facilidad con la que se pueden desarrollar el procedimiento de manera general.

Las suposiciones para esta prueba son mínimas. La población o distribución de interés tiene una media μ y varianza σ^2, con σ^2 conocida. El estadístico de prueba se basa en la media muestral $\overline{X}$, por lo que también se supondrá que la población está distribuída de manera normal o que se aplican las condiciones del teorema central del límite. Esto significa que la distribución de $\overline{X}$ es aproximadamente normal con media μ y varianza σ^2/n.

<u>Ejemplo 1</u>:

Una empresa eléctrica fabrica focos que tienen una duración que se distribuye de forma aproximadamente normal con una media de 800 horas y una desviación estándar de 40 horas. Si una muestra aleatoria de 30 focos tiene una duración promedio de 788 horas, ¿muestran los datos suficiente evidencia para decir que la duración media ha cambiado? Utilice un nivel de significancia del 0,04.

Datos:

μ = 800 horas
σ = 40 horas
$\bar{x}$ = 788 horas
n = 30

Siguiendo los pasos generales para una prueba de hipótesis se tiene:

1. El parámetro de interés es μ, duración promedio en horas.
2. $H_0 : \mu = 800$ horas
3. $H_1 : \mu \neq 800$ horas
4. $\alpha = 0{,}04$
5. La estadística de prueba es

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

6. Región de rechazo:

   Vamos a rechazar H_0 cuando $\overline{X} < C_1$ ó $\overline{X} > C_2$, donde C_1 y C_2 son los denominados valores críticos.

   Se establece la región de rechazo:

$$\alpha = P(\overline{X} < C_1 \mid \mu = 800) + P(\overline{X} > C_2 \mid \mu = 800)$$

7. Como $\overline{X} \sim N(\mu, \sigma^2/n)$ en nuestro caso $\overline{X} \sim N(800, 1600/30)$, entonces

$$(1)\quad P(\overline{X} < C_1 \mid \mu = 800) = \alpha/2$$

$$(2)\quad P(\overline{X} > C_2 \mid \mu = 800) = \alpha/2$$

   Estandarizando tendremos

$$(1)\ P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{C_1-\mu_0}{\sigma/\sqrt{n}}\right)_{H_0} = 0{,}02 \ \rightarrow\ P\left(z < \frac{C_1-800}{40/\sqrt{30}}\right)_{H_0} = 0{,}02 \ \rightarrow\ \frac{C_1-800}{40/\sqrt{30}} = -2{,}052$$

$$\rightarrow\ C_1 = -2{,}052\frac{40}{\sqrt{30}} + 800 \ \rightarrow\ C_1 = 785{,}02$$

$$(2)\ P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{C_2-\mu_0}{\sigma/\sqrt{n}}\right)_{H_0} = 0{,}02 \ \rightarrow\ P\left(z > \frac{C_2-800}{40/\sqrt{30}}\right)_{H_0} = 0{,}02 \ \rightarrow\ \frac{C_2-800}{40/\sqrt{30}} = 2{,}052$$

$$\rightarrow\ C_2 = 2{,}052\frac{40}{\sqrt{30}} + 800 \ \rightarrow\ C_2 = 814{,}98$$

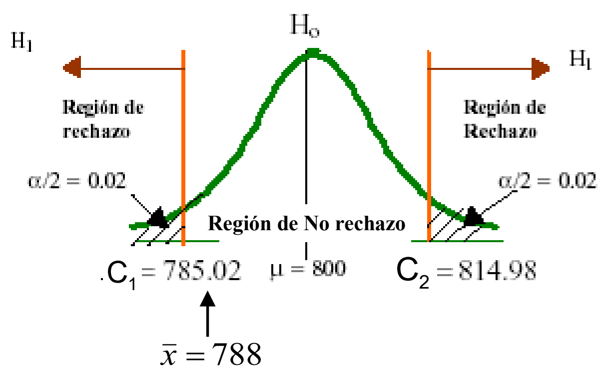

Figura 8: Regiones de rechazo y de no rechazo para $H_0{:}\mu = 800$ contra $H_1{:}\mu \neq 800\ \ y\ \ n = 30$

8. Justificación y decisión:

   Dado que la región de rechazo está dada para: $\overline{X} < 785{,}02$ y $\overline{X} > 814{,}98$, y como $\overline{x} = 788$ horas, no se rechaza H_0 y se concluye con un nivel de significancia del 0,04 que, no hay evidencia suficiente que indique que la duración media de los focos ha cambiado.

Podemos hallar la probabilidad β del error de tipo II considerando que la media verdadera es $\mu = 850$, en tal caso será:

$$\beta = P(785{,}02 \leq \overline{X} \leq 814{,}98 \,|\, \mu = 822)$$

Los valores de z que corresponden a 785,02 y 814,98 cuando $\mu = 822$ son

$$z_1 = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} = \frac{785{,}02 - 822}{7{,}3} = -5{,}04 \quad \text{y} \quad z_2 = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} = \frac{814{,}98 - 822}{7{,}3} = -0{,}96$$

Por tanto,

$$\beta = P(-5{,}04 \leq Z \leq -0{,}96)$$

$$\beta = P(Z \leq -0{,}96) - P(Z \leq -5{,}04)$$

$$\beta = 0{,}1685 - 0{,}000$$

$$\beta = 0{,}1685$$

En este caso se prueba $H_0 : \mu = 800$ contra $H_1 : \mu \neq 800$ con $n = 30$, y suponiendo que el verdadero valor de la media es $\mu = 822$, se tiene que $\beta = 0{,}1685$ de modo que la potencia de la prueba es $1 - \beta = 1 - 0{,}1685 = 0{,}8315$.

==============================