PRUEBAS DE HIPÓTESIS ESTADÍSTICAS

Hipótesis Nula e Hipótesis Alternativa

Muchos problemas de ingeniería requieren que se tome una decisión entre aceptar o rechazar alguna afirmación o conjetura que se formula respecto a algún parámetro. A esta conjetura se la identifica como "hipótesis" y al proceso de toma de decisión sobre esta hipótesis se le conoce como "prueba de hipótesis". Los parámetros que a menudo sometemos a prueba son: la media de una población, la varianza , la proporción de artículos de una población que pertenece a la clase de interés, la diferencia de medias de dos poblaciones , la diferencia entre dos proporciones , la razón de varianzas en dos poblaciones. . .
La prueba de hipótesis estadística corresponde a la etapa de análisis de datos de un "experimento comparativo" en el que el ingeniero está interesado, por ejemplo, en comparar la media de una población con un valor especificado. En estos experimentos pueden intervenir una o más poblaciones y la finalidad es probar hipótesis respecto a algunos de los parámetros de las poblaciones. Así es que una hipótesis estadística es una proposición (afirmación o conjetura) sobre los parámetros de una o más poblaciones.
Por ejemplo supongamos que se tiene interés en la rapidez de combustión de un agente propulsor sólido utilizado en el sistema de escape de emergencia para tripulaciones de aeronaves. La variable aleatoria X es la rapidez de combustión del agente propulsor y μ_X la rapidez promedio de combustión es el parámetro. Se debe decidir si la rapidez promedio de combustión es o no 50 cm/seg. Entonces la hipótesis estadística o hipótesis nula H_0 es:

→ Afirmo algo, igualdad

Hipótesis Nula $H_0: \mu_X = 50\ cm/seg$ → esta hip se somete a prueba / la acepto ó la rechazo

Como contrapartida puede ocurrir alguna de las siguientes proposiciones, llamada Hipótesis Alternativa :

$$\mu_X \neq 50\ cm/seg\ , \quad \mu_X > 50\ cm/seg, \quad \mu_X < 50\ cm/seg$$

→ rechazo a la Ho

A la Hipótesis Alternativa la identificamos con H_1 y resultan los siguientes casos:

Hipótesis Alternativa Bilateral → $H_1: \mu_X \neq 50\ cm/seg$ → ó de dos colas

Hipótesis Alternativa Unilateral Derecha → $H_1: \mu_X > 50\ cm/seg$ → cola a la derecha

Hipótesis Alternativa Unilateral Izquierda → $H_1: \mu_X < 50\ cm/seg$ → cola a la izq.

Las Hipótesis son proposiciones sobre la población y no sobre la muestra.

El valor del parámetro en la Hipótesis Nula (H_0) surge de

- Conocimiento previo.
- Alguna teoría o modelo relacionado con el proceso
- Especificaciones de diseño u obligaciones contractuales

Tengamos en cuenta que la prueba de hipótesis se formula para determinar si ha cambiado el valor del parámetro.

La hipótesis se somete a prueba a partir de los resultados de una muestra tomada de la población de interés. Si la información obtenida de la muestra es consistente con la hipótesis, se concluye que esta es verdadera; si no lo es, se concluye que es falsa. Para decidir si la concordancia es grande o no, en general, calculamos el valor de algún estadístico y comparamos el valor obtenido con la distribución muestral de ese estadístico, en el supuesto que la hipótesis sea verdadera. Sin embargo, la verdad o falsedad nunca puede establecerse a ciencia cierta, a menos que se examine a toda la población. Es necesario que la prueba contemple la probabilidad de llegar a una conclusión equivocada.

Las pruebas de Hipótesis tienen la siguiente estructura

- Una Hipótesis Nula , H_0 , la hipótesis que desea probarse y en la que se especifica un valor exacto de un parámetro : por ejemplo $\mu_X = \mu_0$
  → cual es la hipótesis a investigar
- Una Hipótesis Alternativa , H_1 . El rechazo de H_0 conduce a la aceptación de H_1. En esta hipótesis el parámetro toma valores en un intervalo y puede presentarse en alguna de estas formas: $\mu_X \neq \mu_0$ o $\mu_X > \mu_0$ o $\mu_X < \mu_0$
- Toma de una muestra
- Cálculo del estadístico de prueba a partir de la muestra
- Empleo de la distribución de probabilidades del estadístico para tomar la decisión respecto de H_0

En el caso del combustible sólido de los sistemas de escape de emergencia de las aeronaves una de las características importantes de este producto es la rapidez de combustión y según lo requieren las especificaciones, la rapidez promedio de combustión debe ser de 50 cm /seg. Se sabe que X: "rapidez de combustión de combustible sólido" tiene una distribución normal

$$X \sim N\left(\mu_X = 50\frac{cm}{seg},\ \sigma_X{}^2 = 4\frac{cm}{seg}\right)$$

Además el fabricante asegura que su combustible cumple holgadamente con las especificaciones y que la velocidad de combustión es mayor que la requerida y entonces quiere confirmarse lo afirmado por el proveedor.

En este caso las hipótesis a verificar son:

$$H_0:\ \mu_X = 50\ cm/seg$$

$$H_1:\ \mu_X > 50\ cm/seg$$

La Hipótesis Nula H_0 representa la conclusión que se obtendría si se cumplieran estrictamente las especificaciones.

La Hipótesis Nula se puede contemplar también desde la perspectiva del sistema legal, en el cual se presume inocencia hasta que se demuestre culpabilidad. Por lo tanto se supone que el valor esperado es de 50 cm /seg. A menos que se encuentre evidencia de "culpa " que muestre que ese valor ha cambiado.

El rechazo de H_0 da como resultado la aceptación de la Hipótesis Alternativa H_1 . La Hipótesis Alternativa es la declaración operacional de la hipótesis de investigación del experimentador. La hipótesis de investigación es la predicción derivada de la teoría sometida a prueba.

H_1 constituye la aseveración o hipótesis que se acepta si se rechaza H_0. Representa la conclusión que se obtendría si hubiera evidencia de "culpa".

La Hipótesis Nula es la que se prueba siempre. La Hipótesis Alternativa es la hipótesis de estudio o de investigación y representa una conclusión para la cual se busca evidencia, es la hipótesis que debe aceptarse en caso de rechazar H_0. Es la hipótesis que se desea apoyar con base a la información contenida en la muestra; en nuestro ejemplo $H_1: \mu_X > 50\ cm/seg$

Si se toma una muestra da tamaño $n = 25$, por ejemplo, y se obtiene la velocidad de combustión en cada unidad experimental que constituye la muestra y luego se calcula la media muestral este es el **valor del estadístico de prueba**. Este estadístico media muestral tiene una distribución muestral conocida que emplearemos para decidir por el rechazo o no rechazo de la H_0, cuestión que veremos más adelante.

Por ahora, si con la media obtenida en la muestra construyéramos un intervalo de confianza de $(1-\alpha)\ 100\%$ para μ_X y este contuviera únicamente valores mayores que 50 entonces se tendrá confianza en que la media excede el límite establecido.

Errores Tipo I y II

Hemos visto que una prueba de hipótesis estadística es un procedimiento sistemático basado en datos muestrales para optar o decidir entre la aceptación o el rechazo de una hipótesis estadística. Por desgracia no hay certeza de que no se cometerá una equivocación. De hecho hay dos tipos de errores que pueden cometerse:

- Error Tipo I que ocurrirá cuando la hipótesis establecida sea verdadera y la consideremos falsa.
- Error Tipo II que sucederá cuando la hipótesis sea falsa y la consideremos verdadera

Es decir que se cometerá un Error Tipo I cuando se rechace una hipótesis verdadera y un Error Tipo II cuando no se rechace (se acepte) una hipótesis falsa.

¿ Rechazo o no rechazo la Hipótesis Nula ?

Si se toman muestras independientes de tamaño n de una población normal con media y varianzas conocidas: $X \sim N(\mu_X, \sigma_X^2)$, y la media muestral $\bar{X}$ es el estadístico de la prueba de hipótesis entonces sabemos que tiene una distribución normal: $\bar{X} \sim N(\mu_X, \sigma_X^2/n)$.

Hay que advertir que hay más de una distribución de $\bar{X}$ en juego como s observa en la Figura 1 . Si H_0 es cierta es (1) si H_1 es cierta , la distribución puede ser cualquiera de las (2) . Una sola de estas hipótesis es la verdadera.

- Si H_0 es verdadera $\bar{X} \sim N(50, \sigma_X^2/n)$
- Si H_1 es verdadera $\bar{X} \sim N\left(\mu_X, \frac{\sigma_X^2}{n}\right)$ con $\mu_X > 50$

Se toma una muestra independiente de n observaciones y sean $x_1, x_2, \ldots, x_n$ los valores observados de la variable X en la muestra correspondiente y $\bar{x}$ su media aritmética. El valor $\bar{x}$ es más probable que ocurra bajo la hipótesis H_0 que bajo H_1; $\bar{x}$ sería un valor excepcional bajo H_1 y un valor corriente bajo H_0. Es decir, si H_0 es verdadera, la media $\bar{x}$ para una muestra, muy alejada de 50 se presentará pocas veces. Por eso, si el valor calculado de $\bar{x}$ se aproxima a 50 aceptaremos la hipótesis H_0. Si, en cambio el valor de $\bar{x}$ obtenido está muy alejado de 50 , lo rechazaremos.

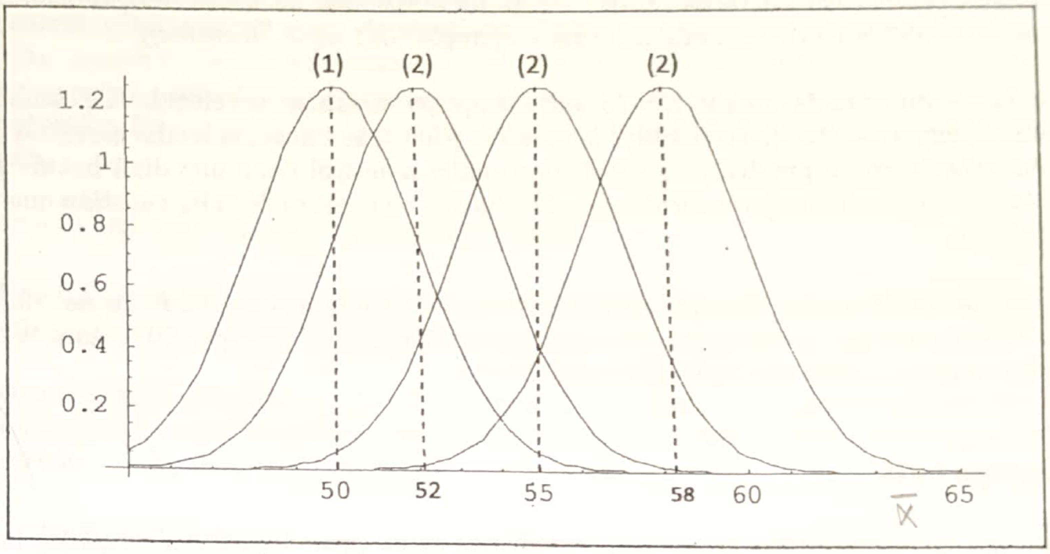

Figura 1: Distribuciones Muestrales de $\bar{X}$ bajo H_0 y bajo H_1

Supongamos que, efectivamente la media real es mayor que 50, la hipótesis alternativa H_1 es verdadera, por ejemplo supongamos que la media verdadera es 52

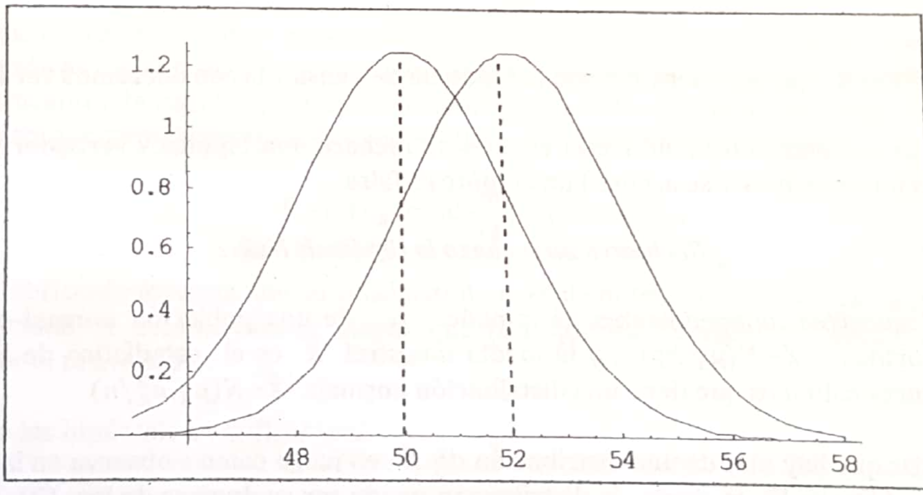

Figura 2: Distribuciones Muestrales de $\bar{X}$ para $\mu_X = 50$ y $\mu_X = 52$

¿ Qué pasa si $50 < \bar{x} < 52$? Esto nos llevaría a no rechazar la Hipótesis Nula y obtendríamos una conclusión incorrecta.

Al probar cualquier hipótesis estadística, se tiene cuatro posibilidades que determinan si la decisión es correcta o equivocada. Estas cuatro situaciones se resumen en la tabla siguiente:

condición real

Decisión	H_0 es VERDADERA	H_0 es FALSA
NO se RECHAZA H_0 Se ACEPTA H_0	DECISIÓN CORRECTA	ERROR TIPO II
Se RECHAZA H_0	ERROR TIPO I	DECISIÓN CORRECTA

Se acepta $H_0 : \mu_X = 50\ cm/seg$ es cierto

Se rechaza H_0 entonces $H_1 : \mu_X > 50\ cm/seg$ es cierta

Es decir que al utilizar una muestra para obtener conclusiones de una población existe el riesgo de llegar a una conclusión incorrecta.

- El error Tipo I se da si se rechaza H_0 cuando en realidad es verdadera.
- El error tipo II se da si se acepta H_0 cuando en realidad es falsa. En este caso se rechaza o no se acepta la hipótesis verdadera que es la alternativa, H_1.

Lo que haremos es tomar todos los valores $\bar{X}$ y encontrar un punto crítico $\bar{X}_c$ tal que :

Si $\bar{X} > \bar{X}_c$ rechazo H_0

Si $\bar{X} \leq \bar{X}_c$ no rechazo H_0

Los posibles valores del estadístico que nos hacen rechazar la hipótesis nula constituyen la **región de rechazo o región crítica (RR)**

Cometeremos Error Tipo I cuando el valor calculado en la muestra $\bar{x} > \bar{X}_c$ y me vea inclinado a rechazar H_0 cuando en realidad H_0 es verdadera.

Cometeremos Error Tipo II cuando el valor calculado en la muestra $\bar{x} \leq \bar{X}_c$ y me vea inclinado a no rechazar H_0 cuando en realidad H_0 es falsa.

A la probabilidad de cometer un error Tipo I se la designa con α y se la denomina **nivel de significación de la prueba**

probabilidad de cometer Error Tipo I

lo pone el que diseña la prueba ← **α = P(Error Tipo I) = P (Rechazar $H_0 \backslash H_0$ es verdadera)**

En nuestro ejemplo $\alpha = P(\bar{X} > \bar{X}_c \setminus \mu_X = 50)$

Tradicionalmente quien diseña la prueba, el estadístico, controla la tasa de errores Tipo I estableciendo el nivel de riesgos que está dispuesto a tolerar en términos de rechazar una hipótesis nula verdadera.

Puesto que el valor de α puede ser elegido por el que realiza el experimento y su elección determinará en parte la aceptación o no aceptación de la hipótesis, este debe fijarse antes de comenzar el experimento:

- Si es de gran importancia rechazar la hipótesis, el riesgo α de cometer ese error debe ser pequeño.
- Si es de gran importancia que una hipótesis sea rechazada si existe únicamente una pequeña evidencia en contra, puede convenir elegir α mayor.

Un convenio que se sigue con frecuencia es establecer el resultado **significativo** si la hipótesis se rechaza con α=0,05 y **muy significativo** si la hipótesis se rechaza con α=0,01.

A la probabilidad de cometer un Error Tipo II se la designa con β

β = P(Error Tipo II) = P (Aceptar $H_0 \backslash H_1$ es verdadera)

éste es el que se usa

Convenio: Hip. se rechaza con $\alpha = 0{,}05$ → resultado significativo

Hip. se ✓ con $\alpha = 0{,}01$ → ✓ muy significativo

En nuestro ejemplo $\beta = P\,(\bar{X} \le \bar{X}_c \setminus \mu_X = 52\,)$

Establecimos la región de rechazo como RR= { $\bar{X} > \bar{X}_c$ } , pero ¿ qué valor debería elegirse para $\bar{X}_c$? ¿ dónde tomamos el punto crítico ?

Analicemos α y β si el punto crítico $\bar{X}_c$ está a la izquierda. Si movemos $\bar{X}_c$ a la derecha aumenta el error tipo II y disminuye el error tipo I como se observa en la Figura 3.

Error tipo II: necesitamos el valor de una media para poder calcular β.

"Es complicado poder calcular β."

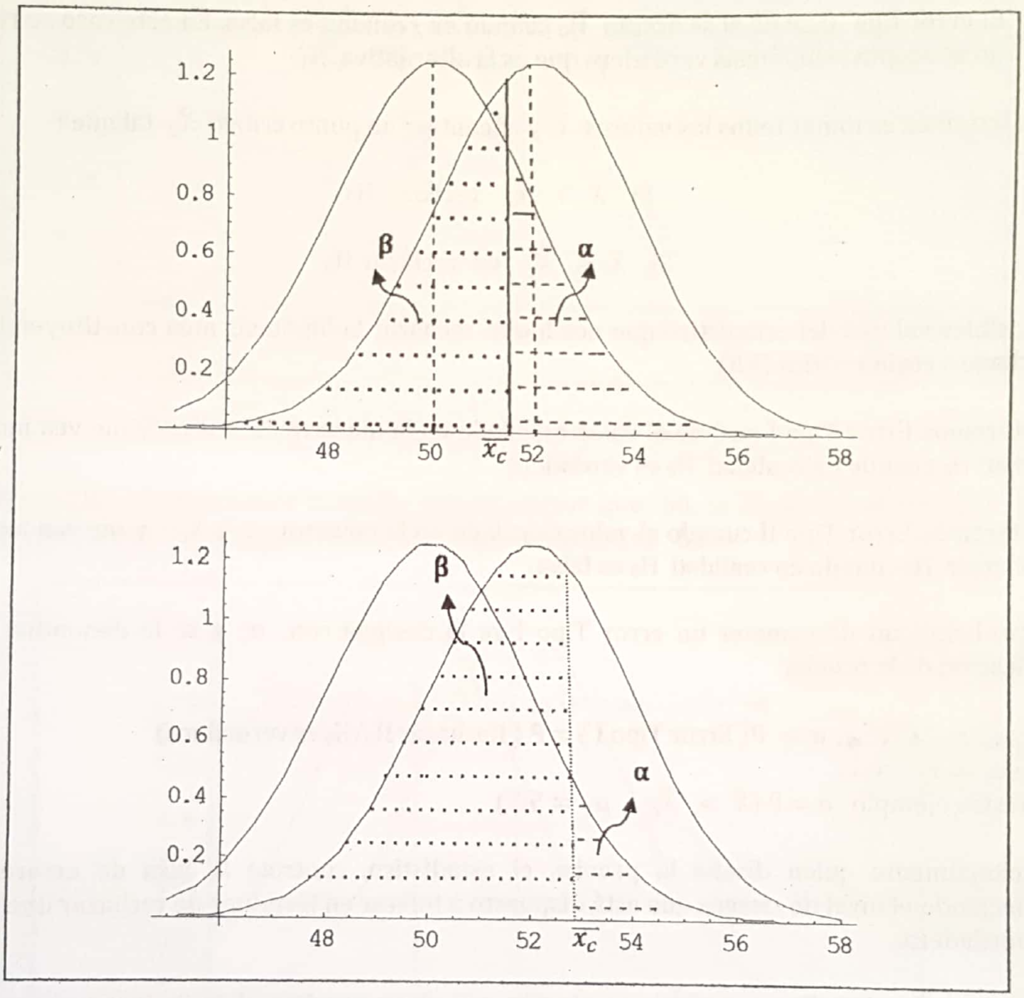

Figura 3: Probabilidad de los errores tipo I y II según el valor de la VA $\bar{X}_c$

Los errores de ambos tipos son competitivos, si aumenta uno se disminuye el otro.

"α uno lo establece, uno lo controla"

Determinación de la región de rechazo

En el caso que estamos analizando el estadístico de la prueba es

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma_X}{\sqrt{n}}} \qquad Z \sim N\,(0,1)$$

Sea la probabilidad de rechazar una hipótesis nula verdadera

$$P\,(\bar{X} > \bar{X}_c \setminus \mu_X = 50\,) = \alpha$$

- Aceptar Ho → Conclusión Débil (a menos que β sea pequeño)

- Rechazo Ho → Conclusión Fuerte (datos aportan evidencias de que Ho es falsa)

Si $\alpha = 0{,}05$ esto me dice que el 5% de todas las muestras me llevarán a rechazar $\mu_0 = 50$ cuando esto es verdadero.

Esta probabilidad se determina empleando el estadístico de prueba

$$P(Z > Z_\alpha) = P\left(\frac{\bar{X}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}} > \frac{\bar{X}_c-\mu_0}{\frac{\sigma_X}{\sqrt{n}}}\right) = \alpha$$

Donde $\dfrac{\bar{X}_c-\mu_0}{\frac{\sigma_X}{\sqrt{n}}} = Z_\alpha$ y despejamos $\bar{X}_c = \mu_0 + Z_\alpha \dfrac{\sigma}{\sqrt{n}}$

Donde $Z_\alpha = Z_{0,05}$ es el valor de Z que deja un área igual a $\alpha = 0{,}05$ a su derecha, es decir $Z_{0,05} = 1{,}645$ y σ_X es la desviación estándar de X y n es el tamaño de la muestra. Entonces para $\sigma_X = 2$ cm/seg y una muestra independiente de tamaño $n = 25$ resulta para el ejemplo dado :

$$\bar{X}_c = 50 + 1{,}645 \frac{2}{\sqrt{25}} = 50{,}658$$

La región de rechazo es $\bar{X} > 50{,}658$

Entonces resulta la siguiente regla para la prueba de hipótesis:

- Con los valores de la variable X obtenidos de una muestra de tamaño n: $x_1, x_2, \ldots, x_n$; se calcula $\bar{x}$ un valor de la variable $\bar{X}$ y se rechaza H_0 al nivel de significación α si $\bar{x} > \mu_0 + Z_\alpha \dfrac{\sigma}{\sqrt{n}}$. Esto es, en el ejemplo, si $\bar{x} > \bar{X}_c$ o $\bar{x} > 50{,}658$ se rechaza H_0 con un nivel de significación de 0,05. En caso contrario no se rechaza.

También la regla puede establecerse en términos del valor calculado a partir de la muestra del estadístico de prueba. En este caso la regla es:

- Con los valores de la variable X obtenidos de una muestra de tamaño n: $x_1, x_2, \ldots, x_n$; se calcula $\bar{x}$ y el valor del estadístico de prueba $z_{\text{de prueba}} = \dfrac{\bar{x}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}}$ y se rechaza H_0 al nivel de significación α si y solo si $z_{\text{de prueba}} > z_\alpha$ ($z_\alpha = z_{\text{crítico}}$) . En el ejemplo se rechaza H_0 con un nivel de significación de 0,05 si $z_{\text{de prueba}} > 1{,}645$. En caso contrario no se rechaza.

Cálculo de β

Recordemos que $\beta = P(\text{Error Tipo II}) = P(\text{Aceptar } H_0 \backslash H_1 \text{ es verdadera})$. Entonces

$$\beta = P(\bar{X} \leq \bar{X}_c \setminus \mu_X > 50) = P\left(\frac{\bar{X}-\mu_X}{\frac{\sigma_X}{\sqrt{n}}} \leq \frac{\bar{X}_c-\mu_X}{\frac{\sigma_X}{\sqrt{n}}}\right)$$ es una función del valor μ_X

Para poder estandarizar la variable y calcular la probabilidad, necesitamos el verdadero valor de μ_X. Supongamos que el valor verdadero de μ_X es 52 (Hipótesis alternativa específica verdadera) ¿ Cual es la probabilidad de cometer Error Tipo II ? Esto sucede cuando no rechazamos H_0 y H_0 es falsa . Como la probabilidad de no rechazar H_0 es $P(\bar{x} \leq \bar{X}_c)$, $\bar{X}_c$ calculado para el nivel de significación α; entonces

$$\beta = P(\bar{X} \leq \bar{X}_c \backslash \ \mu_X = \mu_1) = P\left(Z \leq \frac{\bar{X}_c - \mu_1}{\frac{\sigma_X}{\sqrt{n}}}\right)$$

Para el ejemplo

- $\beta = P(\bar{X} \leq 50{,}658 \backslash \ \mu_X = 52) = P\left(Z \leq \frac{50{,}658-52}{\frac{2}{\sqrt{25}}}\right) = P(Z \leq -3{,}36) = 0{,}0004$

Si la media fuera realmente 52 el 0,04% de todas las muestras de tamaño 25 me llevarían a aceptar $\mu_X = 50$ (H_0) cuando en realidad es $\mu_X = 52$ (H_1)

- $\beta = P(\bar{X} \leq 50{,}658 \backslash \ \mu_X = 51) = P\left(Z \leq \frac{50{,}658-51}{\frac{2}{\sqrt{25}}}\right) = P(Z \leq -0{,}86) = 0{,}1949$

Si la media fuera realmente 51 el 19,49% de todas las muestras de tamaño 25 me llevarían a aceptar $\mu_X = 50$ (H_0) cuando en realidad es $\mu_X = 51$(H_1).

Cuando la hipótesis nula es falsa, β aumenta a medida que el valor verdadero del parámetro tiende al valor hipotético propuesto por la hipótesis nula. El valor de β disminuye a medida que aumenta la diferencia entre el verdadero valor medio y el propuesto.

Potencia de la prueba

La potencia de un test esta dada por la probabilidad de rechazar H_0, dada una hipótesis alternativa específica verdadera. Esto es la probabilidad de decidir correctamente. Siendo H_1 verdadera es la probabilidad de no cometer un error tipo II y puede interpretarse como la probabilidad de rechazar de manera correcta una hipótesis nula falsa. Resulta

Potencia del test = $1 - \beta$

Para nuestro ejemplo; si el valor verdadero de la media fuera 51, la potencia de la prueba es 1-0,1949 = 0,8051. Esta prueba rechazará de manera correcta H_0 y detectará esa diferencia el 80,51 % de las veces.

Si se piensa que el valor de esta potencia es bajo, entonces el analista puede aumentar α o el tamaño de la muestra.

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.
Cuando la hipótesis nula es falsa, β aumenta a medida que el valor verdadero del parámetro tiende al valor hipotético propuesto por la hipótesis nula. El valor de β disminuye a medida que aumenta la diferencia entre el verdadero valor medio y el propuesto.

Al probar cualquier hipótesis las probabilidades de las decisiones, correctas o equivocadas se resumen en la siguiente tabla.

Probabilidades

Decisión	H_0 es VERDADERA	H_0 es FALSA
NO se RECHAZA H_0 Se ACEPTA H_0	$1 - \alpha$	β
Se RECHAZA H_0	α	$1 - \beta$

Conclusión fuerte y conclusión débil

El estadístico controla la probabilidad α del error Tipo I cuando selecciona los valores críticos. Así habitualmente es más fácil para el analista fijar la probabilidad del error Tipo I en casi cualquier valor deseado. Puesto que el analista puede controlar de manera directa la probabilidad de rechazar en forma errónea H_0, siempre puede considerarse el rechazo de la hipótesis nula H_0 como una conclusión fuerte.

La probabilidad β de un Error Tipo II no es constante sino que depende del valor verdadero del parámetro y también del tamaño de la muestra que se haya seleccionado. Se considera que la decisión de aceptar H_0 es una conclusión débil, a menos que se sepa que β es aceptablemente pequeño. Más que decir "se acepta H_0" se prefiere la expresión "no se rechaza H_0". La incapacidad para rechazar H_0 implica que no se ha encontrado evidencia suficiente para rechazar H_0, esto es para formular una proposición fuerte.

La incapacidad para rechazar H_0 no significa necesariamente que exista una probabilidad grande que H_0 sea cierta, sino que simplemente significa que se necesitan más datos, más evidencia para extraer una conclusión fuerte.

Hipótesis unilaterales y bilaterales

Una prueba de hipótesis tal como:

$$H_0: \quad \mu_X = \mu_0$$

$$H_1: \mu_X \neq \mu_0$$

Recibe el nombre de prueba bilateral debido a que es importante detectar diferencias a partir del valor hipotético del parámetro μ_0 que se encuentran en cualquier lado de μ_0.
En una prueba de este tipo la región crítica se separa en dos partes con la misma probabilidad $\alpha/2$ en cada cola de la distribución del estadístico de prueba.

Muchos problemas de pruebas de hipótesis pueden involucrar de manera natural hipótesis alternativas unilaterales, tales como:

$$H_0: \quad \mu_X = \mu_0$$

$$H_1: \mu_X > \mu_0$$

En este caso la región crítica se encuentra en la cola superior de la distribución del estadístico de prueba

$$H_0: \quad \mu_X = \mu_0$$

$$H_1: \mu_X < \mu_0$$

Para esta hipótesis alternativa la región crítica debe encontrarse en la cola inferior de la distribución del estadístico de prueba.

En este tipo de pruebas, llamadas pruebas de una cola, para localizar la región crítica, se visualiza el comportamiento del estadístico de prueba si la hipótesis nula es verdadera y se coloca la región crítica en el extremo o cola apropiada de la distribución. La desigualdad de la hipótesis alternativa H_1 apunta en la dirección de la región crítica.

Al construir una hipótesis siempre se plantea la hipótesis nula como una igualdad de modo que la probabilidad α del error tipo I pueda controlarse con un valor específico.

En cuanto a la hipótesis alternativa

- es unilateral si el objetivo es hacer una afirmación donde aparezcan proposiciones tales como *"mayor que" "Menor que" "Superior a" "Excede a" "Al menos"*

- es bilateral si la afirmación no implica ninguna dirección o es del tipo *"no es igual a"*

Procedimiento general para la prueba de hipótesis

El procedimiento general para formular una prueba de hipótesis puede describirse en los siguientes pasos

Antes de examinar los datos muestrales

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 significación α (error tipo I)

Usando los datos muestrales

5. Establecer el estadístico de prueba apropiado
6. Establecer la región de rechazo para el estadístico
7. Calcular las cantidades muestrales necesarias, sustituirlas en la expresión correspondiente para el estadístico de prueba y obtener el valor de prueba.
8. Decidir si debe o no rechazarse H_0 y notificar esto en el contexto del problema

Los cuatro primeros pasos deben cumplirse antes de examinar los datos muestrales

<u>Ejemplo :</u> Prueba de hipótesis sobre la media con varianza conocida.

Desarrollemos, teniendo en cuenta estos pasos, el ejemplo con el que venimos trabajando en el texto y al que le corresponde el siguiente enunciado:

Los sistemas de escape de emergencia para tripulaciones de aeronaves son impulsados por un combustible sólido. Una de las características de este producto es la rapidez de combustión . las especificaciones requieren que la rapidez promedio de combustión sea de 50 cm / seg . se sabe que la desviación estándar de esta rapidez es σ= 2 cm / seg . el experimentador decide especificar una probabilidad para el error Tipo I o nivel de significación α = 0,05 . Si se selecciona una muestra aleatoria de tamaño 25 y se obtiene una rapidez promedio de combustión de 51,3 cm / seg

¿ A qué conclusiones debe llegar con respecto a la rapidez de combustión de este producto?

1. El parámetro de interés es μ_X : la rapidez promedio de combustión del combustible sólido.
2. La Hipótesis Nula es H_0: $\mu_X = 50\ cm/seg$
3. La Hipótesis Alternativa es H_1 : $\mu_X > 50\ cm/seg$
4. El nivel de significación es α=0,05
5. El estadístico de prueba es $Z = \frac{\bar{X}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}}$ $Z \sim N(0,1)$

Z critico (de tabla)

6. Se rechaza Ho con un nivel de significación de 0,05 si $Z_{\text{de prueba}} > 1{,}645$

7. Con los valores de la variable X obtenidos de una muestra de tamaño n =25; se obtuvo

$\bar{x} = 51{,}3$ que permite determinar $Z_{\text{de prueba}} = \frac{51{,}3-50}{\frac{2}{\sqrt{25}}} = 3{,}25$

8. Como $Z_{\text{de prueba}} = 3{,}25 > 1{,}645$ se Rechaza H_0 con un nivel de significación de 0,05.

El rechazo de H_0 significa que se acepta que el valor promedio puede ser mayor que 50 con un nivel de significación de 0,05 en base al resultado obtenido para la muestra de 25 mediciones. Existe un evidencia fuerte que la rapidez promedio de combustión es mayor que 50 cm / seg.

Los pasos 6. al 8. Pueden realizarse en la siguiente forma alternativa:

6. Se rechaza Ho con un nivel de significación de 0,05 si $\bar{x}$ obtenido de una muestra de tamaño n =25 resulta $\bar{x} > \bar{X}_c$,

$$\bar{X}_c = \mu_0 + Z_\alpha \frac{\sigma}{\sqrt{n}} = 50 + 1{,}645 \frac{2}{\sqrt{25}} = 50{,}658$$

7. Con los valores de la variable X obtenidos de una muestra de tamaño n =25; se obtuvo $\bar{x} = 51{,}3$ que permite determinar $\bar{x} > \bar{X}_c$ ($51{,}3 > 50{,}658$)

8. Como $\bar{x} > \bar{X}_c$ se Rechaza H_0 con un nivel de significación de 0,05.

Resolvamos el mismo problema proponiendo una prueba bilateral.

En este caso para determinar la región de rechazo planteamos:

$$P(\bar{X} > \bar{X}_{c1}) = \alpha/2 \quad \text{o} \quad P(\bar{X} < \bar{X}_{c2}) = \alpha/2$$

Estas probabilidades se determinan empleando el estadístico de prueba

$$P(\bar{X} > \bar{X}_{c1}) = P(Z > Z_{\alpha/2}) = P\left(\frac{\bar{X}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}} > \frac{\bar{X}_{c1}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}}\right) = \alpha/2$$

Donde $\frac{\bar{X}_{c1}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}} = Z_{\alpha/2}$ y despejamos $\bar{X}_{c1} = \mu_0 + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

$$P(\bar{X} < \bar{X}_{c2}) = P(Z < -Z_{\alpha/2}) = P\left(\frac{\bar{X}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}} < \frac{\bar{X}_{c2}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}}\right) = \alpha/2$$

Donde $\frac{\bar{X}_{c2}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}} = -Z_{\alpha/2}$ y despejamos $\bar{X}_{c2} = \mu_0 - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Si $\bar{x} > \bar{X}_{c1} = \mu_0 + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ o $\bar{x} < \bar{X}_{c2} = \mu_0 - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ se rechaza H_0

También se rechaza H_0 si $Z_{\text{de prueba}} = \frac{\bar{x}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}} > Z_{\alpha/2}$ o $Z_{\text{de prueba}} = \frac{\bar{x}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}} < -Z_{\alpha/2}$

Resumamos los pasos de la prueba bilateral para el problema dado:

1. El parámetro de interés es μ_X : la rapidez promedio de combustión del combustible sólido.
2. La Hipótesis Nula es H_0: $\mu_X = 50\ cm/seg$
3. La Hipótesis Alternativa es H_1 : $\mu_X \neq 50\ cm\ /seg$
4. El nivel de significación es α=0,05
5. El estadístico de prueba es $Z = \frac{\bar{X}-\mu_0}{\frac{\sigma_X}{\sqrt{n}}}$ $\quad Z \sim N(0,1)$
6. Se rechaza Ho con un nivel de significación de 0,05 si $Z_{\text{de prueba}} > Z_{0,025}$=1,96 o si $Z_{\text{de prueba}} < -Z_{0,025} = -1,96$, $Z_{0,025}$ es el valor de Z que deja un área de 0,025 a su derecha (0,975 a su izquierda).

7. Con los valores de la variable X obtenidos de una muestra de tamaño n =25; se obtuvo $\bar{x} = 51,3$ que permite determinar $Z_{\text{de prueba}} = \frac{51,3-50}{\frac{2}{\sqrt{25}}} = 3,25$
8. Como $Z_{\text{de prueba}}$ =3,25 > 1,96 se Rechaza H_0 con un nivel de significación de 0,05.

Para esta prueba bilateral también pueden proponerse los pasos 6. al 8. del siguiente modo:

6. Se rechaza Ho con un nivel de significación de 0,05 si $\bar{x}$ obtenido de una muestra de tamaño n =25 resulta $\bar{x} > \bar{X}_c$ o $\bar{x} < -X_c$

$$\bar{X}_c = \mu_0 + Z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}} = 50 + 1,96\ \frac{2}{\sqrt{25}} = 50,784$$

$$\bar{X}_c = \mu_0 - Z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}} = 50 - 1,96\ \frac{2}{\sqrt{25}} = 49,216$$

7. Con los valores de la variable X obtenidos de una muestra de tamaño n =25; se obtuvo $\bar{x} = 51,3$ que permite determinar $\bar{x} > \bar{X}_c$ (51,3 > 50,784)

8. Como $\bar{x} > \bar{X}_c$ se Rechaza H_0 con un nivel de significación de 0,05.

¿ Cuál prueba es la más apropiada para este problema la de una cola o la de dos colas? ¿ Porqué ?

Uso del *valor P* en la prueba de hipótesis

Hemos visto una manera de comunicar los resultados de una prueba de hipótesis que consiste en establecer el rechazo o no de la hipótesis nula fue con un nivel de significación α . Este planteo de las conclusiones de la prueba no brinda al tomador de decisiones ninguna idea sobre si el valor calculado del estadístico de prueba estaba apenas en la región de rechazo o bien adentro de ella. Además establecer de esta manera los resultados impone a otros usuarios de la información un nivel de significación predeterminado. Este enfoque puede ser poco satisfactorio ya que algunos tomadores de decisiones se podrían sentir incómodos con los riesgos implicados si $\alpha = 0{,}05$, por ejemplo. Para evitar estas dificultades se adopta en la práctica el enfoque del valor P.

P es la probabilidad de que el estadístico de prueba tome un valor que sea al menos tan extremo como el valor observado del estadístico de prueba cuando la Hipótesis Nula H_0 es verdadera.

El valor P acarrea información sobre el peso de la evidencia contra H_0, de modo que el tomador de decisiones puede llegar a una conclusión para cualquier nivel de significación especificado.

El valor P o nivel de significación observado de una prueba estadística es el valor más pequeño de α que conduce al rechazo de la hipótesis Nula H_0. Es el riego real de cometer un error Tipo I si se rechaza la Hipótesis Nula con base en el valor observado del estadístico de prueba. El valor P mide la fuerza de la evidencia contra H_0. Cualquier nivel de significación mayor a él que se considere apropiado también conduce al rechazo de la hipótesis nula.
Veamos el cálculo de P para pruebas donde el estadístico de prueba tiene una distribución normal:

Si $z_{\text{de prueba}}$ es el valor del estadístico de prueba obtenido a partir de la muestra resultan los siguientes valores de P , donde $\phi(z_{\text{de prueba}}) = P(Z \leq z_{\text{de prueba}})$

- Para una prueba bilateral o de dos colas $P = 2\,(1 - \phi(z_{\text{de prueba}}))$. El valor P es el área a la derecha de $z_{\text{de prueba}}$ más el área a la izquierda de $-z_{\text{de prueba}}$
- Para una prueba unilateral derecha o de cola superior $P = 1 - \phi(z_{\text{de prueba}})$. El valor P es el área a la derecha de $z_{\text{de prueba}}$
- Para una prueba unilateral izquierda o de cola inferior $P = \phi(z_{\text{de prueba}})$. El valor P es el área a la izquierda de $-z_{\text{de prueba}}$

Para el problema desarrollado como ejemplo el valor de P es :

- Para la prueba unilateral derecha $P = 1 - \phi(3{,}25) = 1 - 0{,}9994 = 0{,}0006$

Por tanto H_0: $\mu_X = 50$ será rechazada con cualquier nivel de significación $\alpha \geq P = 0{,}0006$.
Por ejemplo H_0 será rechazada si $\alpha = 0{,}01$ también será rechazada si $\alpha = 0{,}001$ pero no será rechazada si $\alpha = 0{,}0005$

- Para la prueba bilateral $P = 2\,(1 - \phi(3{,}25)) = 0{,}0012$

Por tanto H_0: $\mu_X = 50$ será rechazada con cualquier nivel de significación $\alpha \geq P = 0{,}0012$. Por ejemplo H_0 será rechazada si $\alpha = 0{,}01$ pero no será rechazada si $\alpha = 0{,}001$

No siempre resulta sencillo calcular P para una prueba pero es un valor que calcula muchos programas de computadora.

Si el valor P brinda un nivel de significación adecuado, se rechaza la Hipótesis Nula. El valor P es el valor de la probabilidad con que se rechaza la hipótesis con el valor del estadístico que resulta de la muestra. El valor P es el nivel de significación observado.

Un valor de P pequeño indica que el valor observado del estadístico de prueba está lejos del valor hipotético dado del parámetro. Esto es una fuerte evidencia de que H_0 es falsa y debe rechazarse.

Si los valores de P son grandes significa que el estadístico de prueba observado no está lejos del valor hipotético del parámetro y no apoya el rechazo de H_0.

¿ Qué tan pequeño necesita ser el valor P para que H_0 pueda rechazarse ?

- Si el valor P es menor que un nivel de significación α preasignado, entonces puede rechazarse la hipótesis nula, e informar que los resultados son estadísticamente significativos en el nivel α.

- Muchos investigadores usan la siguiente "escala móvil" para clasificar sus resultados

  - Si el valor P es menor que 0,01 , se rechaza H_0. Los resultados son **muy significativos.**

  - Si el valor P está entre 0,01 y 0,05 , se rechaza H_0. Los resultados son **estadísticamente significativos**

  - Si el valor P está entre 0,05 y 0,10 , generalmente no se rechaza H_0. Los resultados sólo **tienden hacia la significación estadística**

  - Si el valor P es mayor que 0,10, H_0 no se rechaza. Los resultados no son **estadísticamente significativos**

Relación de las pruebas de hipótesis con la estimación del intervalo de confianza

La estimación del Intervalo de Confianza implica el cálculo de límites para los cuales es "razonable" que el parámetro en cuestión esté dentro de ellos.

Para el caso de la media poblacional μ_X con varianza σ_X^2 conocida, la prueba de hipótesis y la estimación del Intervalo de Confianza se basan en la Variable Aleatoria

$$Z = \frac{\bar{X} - \mu_X}{\frac{\sigma_X}{\sqrt{n}}}$$

Resulta que la prueba H_0: $\mu_X = \mu_0$ H_1 : $\mu_X \neq \mu_0$ en un nivel de significación α es equivalente a calcular un intervalo de confianza del $(1-\alpha)$ 100 % de μ_X y rechazar la Hipótesis Nula si μ_0 no está dentro del intervalo de confianza. Si, en cambio, μ_0 está dentro del intervalo no se rechaza la hipótesis.

El intervalo de confianza del $(1-\alpha)$ 100 % para μ_X resulta

$$\bar{x} - z_{\alpha/2}\frac{\sigma_X}{\sqrt{n}} < \mu_X < \bar{x} + z_{\alpha/2}\frac{\sigma_X}{\sqrt{n}}$$

Donde $\bar{x}$ es la media obtenida de una muestra de tamaño n

Para el ejemplo que hemos desarrollado el intervalo de confianza resulta:

$$51{,}3 - 1{,}96\frac{2}{\sqrt{25}} < \mu_X < 51{,}3 + 1{,}96\frac{2}{\sqrt{25}} \qquad 50{,}516 < \mu_X < 52{,}084$$

Se observa que $\mu_0 = 50$ no está incluido en el intervalo de confianza; entonces se rechaza H_0

La equivalencia del intervalo de confianza con las pruebas de hipótesis bilaterales se extiende a la diferencia de medias, proporciones, diferencia de proporciones, varianzas, cociente entre varianzas, etc .

En este capítulo y en el anterior se examinaron dos componentes importantes de la inferencia estadística, la estimación de intervalos de confianza y las pruebas de hipótesis. Aunque están basados en los mismos conceptos, se usan con propósitos diferentes. Se emplean los intervalos de confianza para estimar parámetros y las pruebas de hipótesis se usan para tomar decisiones acerca de los valores específicos de los parámetros de la población.

Hemos visto que para determinar la zona de rechazo en un test unilateral o bilateral en base a un nivel de significación dado o para calcular el valor P, se recurre a un estadístico de prueba con una distribución muestral conocida. La selección de esta variable aleatoria depende del parámetro de la población sobre el que se formula la prueba, las características de la población o poblaciones involucradas, el tamaño de la muestra, el tipo de muestreo, etc. Para identificarla se puede recurrir a la Tabla de Intervalos de Confianza en donde se establecen las variables aleatorias cuya distribución muestral permitirá establecer las zonas de rechazo en función del parámetro poblacional y de las características de la población y de la muestra.

Resolveremos algunos ejemplos

Ejemplos

1.- Si un estudio reciente reportó que de 899 negocios que operan desde el hogar, las mujeres eran dueñas de 369. ¿Existe evidencia de que es igual el número de hombres y mujeres propietarios de negocios que operan desde su casa?. Responde estableciendo una prueba de hipótesis. Utiliza un nivel de significación α=0,05. Calcula un valor P para esta prueba y establece su significado.

El parámetro respecto al que se desea probar una hipótesis es la proporción p de hombres o de mujeres que son propietarios de negocios que operan desde su casa.

Teniendo en cuenta que nos interesa probar si *existe evidencia de que es igual el número de hombres y mujeres propietarios de negocios que operan desde su casa,* se propone la siguiente Hipótesis Nula

$$H_0:\ p = 0{,}5$$

La proporción p de propietarios de negocios que operan desde su casa que son mujeres es 0,5. Hay tantos hombres como mujeres propietarios de negocios que operan desde su casa.

La Hipótesis Alternativa resulta

$$H_1:\ p \neq 0{,}5$$

Hay mayor proporción de personas de un género en particular entre los propietarios de negocios que operan desde su casa.

El nivel de significación que se propone para la prueba es α=0,05

El estadístico de prueba apropiado es $\hat{P} = \dfrac{X}{n}$ donde X es el número de éxitos en n intentos.

Sabemos que si $n\,\hat{p} \geq 5$ y $n\,\hat{q} \geq 5$ su distribución muestral es aproximadamente normal con media $\mu_{\hat{P}} = p$ y varianza

$$\sigma^2_{\hat{P}} = \frac{p\,q}{n} \quad , \quad q = 1 - p$$

$$\hat{P} \approx N\left(p, \frac{p\,q}{n}\right) \quad y \quad Z = \frac{\overline{P} - p}{\sqrt{\dfrac{\widehat{P}\,\widehat{Q}}{n}}}$$

Tal como se encuentra en el renglón 9 y columna 5 de la Tabla de Intervalos de Confianza

Con los datos muestrales se obtienen :

$$\hat{p} = \frac{369}{899} = 0{,}41 \text{ (proporción muestral de mujeres),}$$

n= 899 tamaño de la muestra y el número de éxitos en n intentos es 369

$$\hat{q} = 1 - \hat{p} = 0{,}59 \text{ proporción de hombres obtenida en la muestra}$$

$$z_{prueba} = \frac{\hat{p}-p}{\sqrt{\frac{\hat{p}\hat{q}}{n}}} = \frac{0{,}41-0{,}5}{\sqrt{\frac{0{,}41\ \ 0{,}59}{899}}} = -5.486$$

La regla para tomar la decisión es : Se rechaza H_0 con un nivel de significación de 0,05 si

$$z_{prueba} < -Z_{0,025} = -1{,}96 \quad \text{o} \quad z_{prueba} > Z_{0,025} = 1{,}96$$

$Z_{0,025}$ es el valor de Z que deja un área de 0,025 a su derecha,

$-Z_{0,025} = Z_{0,975}$ es el valor de Z que deja un área de 0,025 a su izquierda

Como si $z_{prueba} = -5.486 < -1{,}96$, se Rechaza H_0

Existe evidencia de que no es igual el número de hombres y de mujeres de negocios que operan desde su casa. Son más los hombres que las mujeres.

Como camino alternativo podemos determinar los valores críticos de $\hat{P}$ y rechazar H_0 con un nivel de significación de 0,05 si $\hat{p} < \hat{P}_{c_1}$ o $\hat{p} > \hat{P}_{c_2}$. Resulta

$$P(\hat{P} < \hat{P}_{c_1}) = P(Z < Z_{0,025}) = P\left(Z < \frac{\hat{P}_{c_1} - 0{,}5}{\sqrt{\frac{0{,}41\,0{,}59}{899}}}\right) = 0{,}025$$

$$Z_{0,025} = -1{,}96 \Rightarrow \frac{\hat{P}_{c_1} - 0{,}5}{\sqrt{\frac{0{,}41\,0{,}59}{899}}} = -1{,}96 \Rightarrow \hat{P}_{c_1} = 0{,}46$$

$$\hat{p} = 0{,}43 < \hat{P}_{c_1} = 0{,}46 \qquad \Rightarrow \qquad \text{se rechaza } H_0$$

Si proponemos el intervalo de confianza para el parámetro p

$$\bar{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\,\hat{q}}{n}} < p < \bar{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\,\hat{q}}{n}}$$

Se obtiene

$$0{,}41 - 1{,}96\sqrt{\frac{0{,}41\ \ 0{,}59}{899}} < p < 0{,}41 + 1{,}96\sqrt{\frac{0{,}41\ \ 0{,}59}{899}} \quad \Rightarrow \quad 0{,}3778 < p < 0{,}4422$$

Se observa que el intervalo de confianza para la proporción p no contiene al valor p= 0,5. Entonces se rechaza esta hipótesis.

Determinación del valor P

$$P = 2\ \emptyset_{(-5{,}486)} = 2\left(1-\emptyset_{(5{,}486)}\right) < 2\left(1-\emptyset_{(3{,}99)}\right) = 2(1-0{,}99997) = 0{,}00006$$

El mayor valor de z para el que se registra un valor de la función de distribución acumulada en la tabla es $z = 3{,}99$. Empleamos este valor para obtener una cota superior del valor P. Si observamos en la Figura 4 el área a la izquierda de valores menores a -4 (o a la derecha da valores mayores que 4) es muy pequeña.

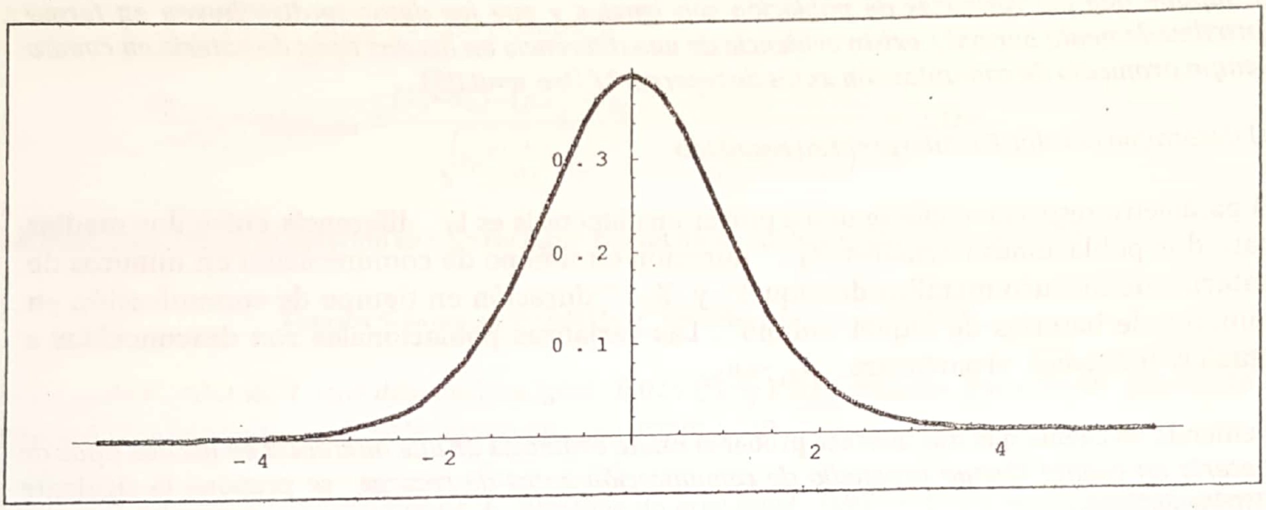

Figura 4: Distribución normal estándar

La hipótesis nula será rechazada para todo $\alpha \geq 0{,}00006$. La probabilidad de obtener este z prueba u otro menor es de menos que 0,00006 cuando la hipótesis nula es verdadera.

También puedes proponer una prueba de una cola o unilateral para este problema:

$$H_0: \ p = 0{,}5$$

La proporción p de propietarios de negocios que operan desde su casa que son mujeres es 0,5.

La Hipótesis Alternativa resulta

$$H_1: \ p < 0{,}5$$

Hay menor proporción de mujeres entre los propietarios de negocios que operan desde su casa.

2.- Un fabricante desarrolla una batería de hidruro metálico de níquel para usarse en teléfonos celulares en lugar de la de níquel cadmio. El director de control de calidad decide evaluar el desempeño de la batería recién desarrollada contra la de amplio uso de níquel-cadmio. Una muestra aleatoria de 25 baterías de níquel cadmio y una muestra aleatoria de 25 baterías desarrolladas se colocan en teléfonos celulares de la misma marca y modelo. La medida de desempeño de interés es el tiempo de comunicación en minutos antes de recargar. Los datos se muestran en la siguiente tabla

Batería de Níquel Cadmio					*Batería de hidruro metálico de níquel*				
54,5	71,0	67,0	67,8	41,7	78,3	103,0	79,8	95,4	81,3
56,7	64,5	69,7	86,8	70,4	91,1	69,4	46,4	82,8	87,3
40,8	74,9	72,5	75,4	76,9	82,3	71,8	62,5	83,2	77.5
64,9	81,0	104,4	83,3	90,4	85,0	85,3	74,3	85,3	85,5
82,0	72,8	71,8	58,7	68,8	86,1	72,1	112,3	74,1	41,1

a) Supone que las varianzas de población son iguales y que los datos se distribuyen en forma aproximadamente normal ¿ existe evidencia de una diferencia en los dos tipos de batería en cuanto tiempo promedio de comunicación antes de recargar? (Use $\alpha=0.05$)

b) determina el valor P e interpreta su resultado

El parámetro respecto al que se desea probar una hipótesis es la **diferencia entre dos medias** para dos poblaciones normales X_1 : " duración en tiempo de comunicación en minutos de baterías de hidruro metálico de níquel" y X_2 : " duración en tiempo de comunicación en minutos de baterías de níquel cadmio" . Las varianzas poblacionales son **desconocidas e iguales.** Indicamos al parámetro $\mu_{X_1} - \mu_{X_2}$

Teniendo en cuenta que nos interesa probar si *existe evidencia de una diferencia en los dos tipos de batería en cuanto tiempo promedio de comunicación antes de recarga,* se propone la siguiente Hipótesis Nula

$$H_0: \ \mu_{X_1} - \mu_{X_2} = 0 \quad (\mu_{X_1} = \mu_{X_2})$$

No hay diferencia en cuanto a la duración promedio de ambos tipos de baterías

La Hipótesis Alternativa resulta

$$H_1: \ \mu_{X_1} - \mu_{X_2} \neq 0 \quad (\mu_{X_1} \neq \mu_{X_2})$$

Hay diferencia entre la duración media de ambos tipos de baterías (Hay un tipo que dura más)

También puede proponerse

$$H_1: \ \mu_{X_1} - \mu_{X_2} > 0 \quad (\mu_{X_1} > \mu_{X_2})$$

Las baterías hidruro metálico de níquel tienen más duración que las de níquel cadmio

El nivel de significación que se propone para la prueba es $\alpha=0,05$

El estadístico de prueba apropiado es $\overline{X_1} - \overline{X_2}$ la diferencia entre las medias muestrales. En el caso de poblaciones normales y muestras de tamaño $n<30$ con varianzas iguales la distribución muestral para la diferencia de medias corresponde a la variable aleatoria T con distribución t de student con $\nu = n_1 + n_2 - 2$ siguiente

$$T = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_{X_1} - \mu_{X_2})}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Tal como se encuentra en el renglón 6 y columna 5 de la Tabla de Intervalos de Confianza

Con los datos muestrales se obtienen las medias y varianzas muestrales según la siguiente tabla

	Baterías de Hidruro Metálico De níquel	Baterías de Níquel Cadmio
Tamaño de la muestra	$n_1 = 25$	$n_2 = 25$
Media muestral	$\overline{x_1}$=79,73	$\overline{x_2}$=70,75
Varianza muestral	$s_1^2 = 226{,}03$	$s_2^2 = 195{,}79$

El valor muestral de S_p^2 es $s_p{}^2 = \frac{(n_1-1)s_1{}^2+(n_2-1)s_2{}^2}{n_1+n_2-2} = \frac{24(226{,}03+195{,}03)}{48} = 210{,}91$ ν=48

El valor de t para los datos de esta prueba es

$$t_{\text{ prueba}} = \frac{(\overline{x_1}-\overline{x_2})-(\mu_{X_1}-\mu_{X_2})}{\sqrt{s_p{}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{(79{,}73-70{,}75)}{\sqrt{210{,}91\left(\frac{2}{25}\right)}} = 2{,}186$$

La regla para tomar la decisión es : Se rechaza H_0 con un nivel de significación de 0,05 si

$$t_{\text{prueba}} < t_{0{,}975\ ,48} = \text{-}2{,}016 \text{ o } t_{\text{prueba}} > t_{0{,}025\ ,48} = 2{,}016$$

$t_{0{,}025\ ,48}$ es el valor de t que deja un área igual 0,025 ($1-\frac{\alpha}{2}$) a su derecha para ν=48 grados de libertad y se encuentra en la tabla ingresando con 0,025 y 48.

$t_{0{,}975\ ,48} = -\,t_{0{,}025\ ,48}$ es el valor de t que deja un área igual 0,975 ($1-\frac{\alpha}{2}$) a su derecha para ν=48 grados de libertad

Como t_{prueba}= 2,186 > 2,016 , se Rechaza H_0. Existe evidencia de que las baterías no tienen la misma duración. Hay diferencia en los tiempos de duración ¿ Cuál es el tipo de baterías que más dura ?

Como camino alternativo podemos determinar la región de rechazo determinando los valores críticos de $\overline{X_1}-\overline{X_2}$ y rechazar H_0 con un nivel de significación de 0,05 si

$$\overline{x_1}-\overline{x_2} < (\overline{X_1}-\overline{X_2})_{c_1} \quad \text{o si} \quad \overline{x_1}-\overline{x_2} > (\overline{X_1}-\overline{X_2})_{c_2}$$

Resulta

$$P\left(\overline{x_1}-\overline{x_2} > (\overline{X_1}-\overline{X_2})_{c_2}\right) = \frac{\alpha}{2} = 0{,}025$$

$$P\left(\overline{x_1}-\overline{x_2} > (\overline{X_1}-\overline{X_2})_{c_2}\right) = 1-P(t < t_{\alpha/2}, \nu) = 1-P(t < 2{,}016)$$

$$t_{0{,}025\ ,48} = \frac{(\overline{X_1}-\overline{X_2})_{c_2}}{\sqrt{210{,}91\left(\frac{2}{25}\right)}} = 2{,}016 \quad \Rightarrow (\overline{X_1}-\overline{X_2})_{c_2} = 2{,}016\sqrt{210{,}91\left(\frac{2}{25}\right)} = 8{,}28$$

La diferencia de medias para las muestras obtenidas es $\overline{x_1}-\overline{x_2}$ =79,73-70.75= 8,98 > 8,28 , entonces se rechaza H_0

En este caso, no necesitamos construir la cola izquierda de la zona de rechazo pues se pudo tomar la decisión con la región a la derecha.

Si proponemos el intervalo de confianza del $(1-\alpha)100\%$ para la diferencia de medias $\mu_{X_1}-\mu_{X_2}$ en dos poblaciones normales con varianzas iguales pero de valor desconocido $\sigma_{X_1}^2=\sigma_{X_2}^2=\sigma^2$

$$(\bar{x}_1-\bar{x}_2)-t_{\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}<\mu_{X_1}-\mu_{X_2}<(\bar{x}_1-\bar{x}_2)+t_{\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

Para $\alpha=0{,}05$ y los valores obtenidos en la muestra se tiene

$$(8{,}98)-2{,}016\sqrt{210{,}91\left(\frac{2}{25}\right)}<\mu_{X_1}-\mu_{X_2}<(8{,}98)+2{,}016\sqrt{210{,}91\left(\frac{2}{25}\right)}$$

Reemplazamos y resulta el intervalo de confianza para el parámetro poblacional

$$0{,}88<\mu_{X_1}-\mu_{X_2}<21{,}19$$

Se observa que el valor $\mu_{X_1}-\mu_{X_2}=0$ no está incluido en este intervalo de confianza y esto nos lleva a rechazar la igualdad de medias.

Determinación del valor P

Como $t_{prueba}=2{,}186$, se busca en la Tabla de los valores críticos de t en el renglón para $\nu=48$ un par de valores que contenga a este valor. Los valores encontrados son 2,0086 y 2,4033 que dejan respectivamente un área de 0,025 y 0,01 a su derecha. Podemos decir que

$$0{,}01<P<0{,}025$$

El valor exacto de P, el área a la derecha de t_{prueba} es proporcionado por los software estadísticos.

La hipótesis nula será rechazada para todo $\alpha \geq P$. La probabilidad de obtener este t_{prueba} u otro mayor es menor que P cuando la hipótesis nula es verdadera.

3) Para el problema anterior ¿ existen indicios de una diferencia en las varianzas del tiempo de comunicación, en minutos, antes de recargar entre los dos tipos de baterías? Emplea 0,05 de nivel de significación

El parámetro respecto al que se desea probar una hipótesis es $\frac{\sigma_{X_1}^2}{\sigma_{X_2}^2}$ el cociente de dos varianzas poblacionales, con las poblaciones X_1 y X_2 Normales. Se emplea como estimador o función de decisión al estadístico $\frac{S_1^2}{S_2^2}$ el cociente de las varianzas muestrales para muestras aleatorias independientes de tamaños n_1 y n_2 con una distribución muestral F de Fisher con ν_1,ν_2 grados de libertad, $\nu_1=n_1-1$ y $\nu_2=n_2-1$.

$$F=\frac{S_1^2}{S_2^2}\frac{\sigma_{X_2}^2}{\sigma_{X_1}^2}$$

Esto se encuentra en el renglón 11, columna 5 de la Tabla de Intervalos de confianza.

Teniendo en cuenta que nos interesa probar si *existen indicios de una diferencia en las varianzas del tiempo de comunicación, en minutos, antes de recargar entre los dos tipos de baterías,* se propone la siguiente Hipótesis Nula

$H_0: \sigma_{X_1} - \sigma_{X_2} = 0 \quad (\sigma_{X_1} = \sigma_{X_2})$ La varianzas son iguales

La Hipótesis Alternativa resulta

$H_1: \sigma_{X_1} - \sigma_{X_2} \neq 0 \quad (\sigma_{X_1} \neq \sigma_{X_2})$ Las varianzas no son iguales

El nivel de significación que se propone para la prueba es $\alpha=0{,}05$

Con los datos muestrales, se obtiene que $\frac{s_1^2}{s_2^2} = \frac{226{,}03}{195{,}79} = 1{,}154$

El valor de f para los datos de esta prueba es

$$f_{\text{prueba}} = \frac{s_1^2}{s_2^2}\,\frac{\sigma_{X_2}^2}{\sigma_{X_1}^2} = 1{,}154$$

La regla para tomar la decisión es : Se rechaza H_0 con un nivel de significación de 0,05 si

$$f_{\text{prueba}} < f_{0{,}975\,,24{,}24} = \frac{1}{f_{0{,}025\,,24{,}24}} = 0{,}44 \qquad \text{o} \quad f_{\text{prueba}} > f_{0{,}025\,,24{,}24} = 2{,}27$$

$f_{0{,}025\,,24{,}24}$ es el valor de f que deja un área igual 0,025 ($1-\frac{\alpha}{2}$) a su derecha para ν_1, ν_2 grados de libertad y se encuentra en la tabla ingresando con los grados de libertad y con 0,025 .

$f_{0{,}975\,,24{,}24}$ es el valor de t que deja un área igual 0,975 ($1-\frac{\alpha}{2}$) a su derecha para ν_1, ν_2 grados de libertad . Recordemos que $\frac{1}{f_{1-(\alpha/2),\nu_1,\nu_2}} = f_{\alpha/2,\nu_2,\nu_1}$

Como $f_{\text{prueba}} = 1{,}154$ está entre los valores 0,44 y 2,27 No se Rechaza H_0. NO existe evidencia de que las poblaciones de baterías tengan varianzas distintas. La suposición de varianzas iguales hecha en el ejemplo 2 puede mantenerse.

Se puede determinar el intervalo de confianza del 95% para $\frac{\sigma_{X_1}^2}{\sigma_{X_2}^2}$

$$\frac{s_1^2}{s_2^2}\,\frac{1}{f_{\alpha/2,\,\nu_1,\,\nu_2}} < \frac{\sigma_{X_1}^2}{\sigma_{X_2}^2} < \frac{s_1^2}{s_2^2}\, f_{\alpha/2,\,\nu_2,\,\nu_1}$$

Y se obtiene, reemplazando

$$1{,}154\ 0{,}44 < \frac{\sigma_{X_1}^2}{\sigma_{X_2}^2} < 1{,}154\ 2{,}27 \quad \Rightarrow \quad 0{,}508 < \frac{\sigma_{X_1}^2}{\sigma_{X_2}^2} < 2{,}61$$

Se observa que el valor $\frac{\sigma_{X_1}^2}{\sigma_{X_2}^2} = 1$ está incluido en este intervalo de confianza y esto nos lleva a no rechazar la igualdad de las varianzas.