CAPÍTULO VIII: PRUEBA DE HIPOTESIS

Una **Prueba de hipótesis** es una metodología inferencial que sirve para probar la validez de una afirmación acerca del valor de uno ó más parámetros de una población. Se llama **hipótesis estadística** a la afirmación cuantitativa acerca del o los parámetros poblacionales.

Un medio seguro para probar una hipótesis es realizar un examen de la población completa lo cual es prácticamente imposible por el costo y el tiempo que insumiría. La forma más práctica de hacerlo es a través de una muestra, mediante la cual se rechazará o no la hipótesis propuesta y se tomará una decisión al respecto.

Ejemplo: Supóngase que una repartición estatal abre una licitación de lámparas eléctricas. Un fabricante de tal producto decide presentarse a la misma para lo cual necesita conocer la vida útil media de las lámparas que posee en stock. De acuerdo a los datos que el fabricante dispone, por ejemplo alguna licitación anterior, supone que las lámparas de su producción tienen una vida útil promedio de 250 horas, afirmación que puede no ser verdadera para el presente. Para probar la validez de esta suposición, decide tomar una muestra de 25 lámparas. Si la media aritmética de la muestra, $\bar{x}$, fuese 240 horas es probable que el fabricante no rechace que su población de lámparas cumple con la afirmación supuesta mientras que si la media aritmética fuese 80 horas estaría predispuesto a rechazar su suposición.

Normalmente los resultados no son tan absolutos en su representación; las conclusiones no se pueden obtener con tanta facilidad. De aquí la necesidad de utilizar métodos estadísticos que permitan tomar decisiones respecto a hipótesis planteadas basándose en datos muestrales.

Se llama **hipótesis nula** y se nota $\mathbf{H_0}$, a aquella hipótesis referida a algún parámetro poblacional que se formula con el propósito de rechazar o no rechazar. En el ejemplo anterior, la hipótesis nula referida a la vida útil promedio de las lámparas eléctricas puede expresarse simbólicamente por

$$H_0: \mu = 250 \text{ horas.}$$

La hipótesis nula debe considerarse como verdadera a menos que exista suficiente evidencia en contra de ella.
Para construir una regla de decisión apropiada, en la prueba de hipótesis también es necesario establecer una **hipótesis alternativa** que refleje el valor posible o intervalo de valores del parámetro de interés si la hipótesis nula es falsa. Esto es, la hipótesis alternativa representa alguna forma de negación de la hipótesis nula y generalmente se nota $\mathbf{H_1}$. En el ejemplo anterior se formula como

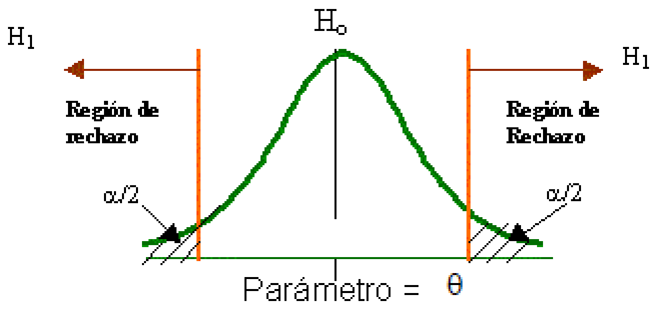
$$H_1: \mu \neq 250.$$

Por lo tanto, una **prueba de hipótesis estadística** con respecto a alguna característica desconocida de la población de interés es cualquier regla para decidir si se rechaza o no la hipótesis nula con base en una muestra aleatoria de la población.

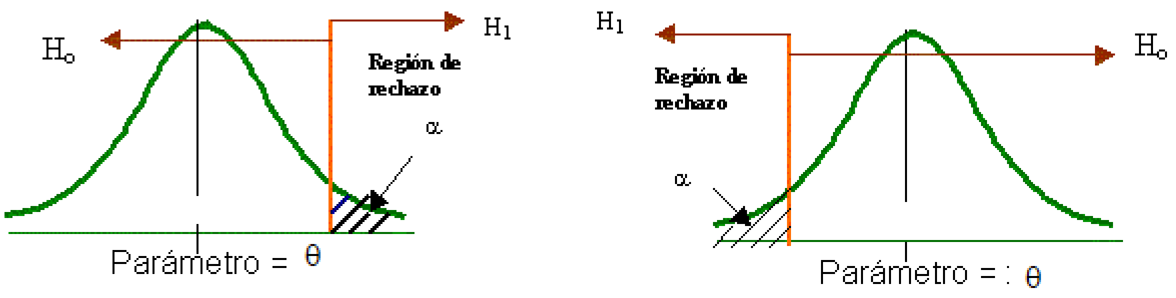
Una prueba donde la hipótesis alternativa especifique que el parámetro poblacional es sólo mayor o sólo menor que el valor especificado en la hipótesis nula constituye una **prueba unilateral**. Una hipótesis alternativa que especifique que el parámetro es distinto del valor indicado por H_0 se llama **prueba bilateral**.

Prueba Bilateral	Pruebas Unilaterales	
H_0: $\theta = \theta_0$	H_0: $\theta = \theta_0$ (ó $\theta \leq \theta_0$)	H_0: $\theta = \theta_0$ (ó $\theta \geq \theta_0$)
H_1: $\theta \neq \theta_0$	H_1: $\theta > \theta_0$	H_1: $\theta < \theta_0$

Prueba Bilateral



Pruebas Unilaterales



Los términos región de rechazo, región de aceptación y α serán definidos a continuación.
La siguiente tabla especifica las posibles decisiones que pueden tomarse con respecto a la hipótesis nula:

Hipótesis / Decisión	H_0 es verdadera	H_0 es falsa
No rechazar H_0	Decisión correcta	Decisión incorrecta **Error de Tipo II**
Rechazar H_0	Decisión incorrecta **Error de Tipo I**	Decisión correcta

La **probabilidad de cometer un Error de Tipo I** generalmente es denotada por α y la **probabilidad de cometer un Error de Tipo II** por β. Simbólicamente se puede expresar:

$$P \text{ (Error de Tipo I)} = P \text{ (rechazar } H_0 \text{ / } H_0 \text{ verdadera)} = \alpha$$

$$P \text{ (Error de Tipo II)} = P \text{ (No rechazar } H_0 \text{ / } H_0 \text{ falsa)} = \beta.$$

En el ejemplo de la vida útil promedio de las lámparas eléctricas, si se rechaza la hipótesis nula de 250 hs. de acuerdo a los resultados de la prueba, pero realmente el valor medio es de 250 hs. se comete un error de Tipo I. Por otra parte, si la hipótesis H_0: $\mu = 250$ horas no es rechazada según los datos de la prueba, pero realmente el valor medio es $\mu = 230$ hs., se comete un error de Tipo II.
La **probabilidad α** del error de Tipo I se conoce con el nombre de **nivel de significación**.

El enfoque general en prueba de hipótesis estadísticas es aceptar la premisa de que el error de Tipo I es más serio que el error de Tipo II y por lo tanto construir la prueba de modo que α sea pequeña. Este objetivo, sin embargo, ignora la probabilidad de cometer un error de Tipo II. Las probabilidades α y β no son independientes entre sí, y tampoco lo son del tamaño de la muestra *n*. Cuando disminuye el tamaño de α, aumenta el de β, y viceversa (si el tamaño *n* permanece fijo).
El complemento de β es la probabilidad

$$1- \beta = P \text{ (Rechazar } H_0 \text{ / } H_0 \text{ falsa)},$$

que recibe el nombre de **potencia de la prueba** puesto que es la *capacidad* (o “poder”) que tiene la prueba de reconocer correctamente que la hipótesis nula es falsa. Es decir, representa la probabilidad de detectar desviaciones respecto de H_0 cuando existen, y por lo tanto, es deseable que una prueba tenga una potencia grande (cercana a uno). La potencia suele notarse con las letras griegas γ o π según la bibliografía.
Para cualquier nivel de alfa dado, el aumento del tamaño de la muestra siempre produce una mayor potencia de la prueba estadística por lo tanto existe un disminución de β.

Planteadas las hipótesis nula y alternativa, la decisión de rechazar o no H_0 se basa en algún estadístico apropiado el cual recibe el nombre de **estadístico de prueba** y cuya **distribución de probabilidad es conocida bajo H_0**.

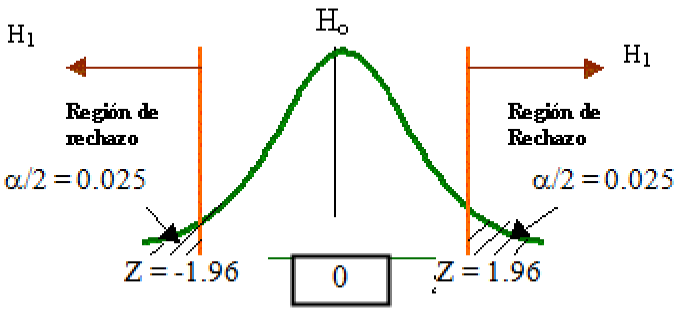
Prueba de hipótesis para la media de una población normal con varianza conocida

En el ejemplo de las lámparas eléctricas, el parámetro al cual se refiere la hipótesis nula es la media poblacional μ. El mejor estimador puntual de μ es la media muestral $\overline{X}$. Por lo tanto al probar H_0: $\mu = 250$ versus H_1: $\mu \neq 250$ se desearía determinar la proximidad de $\overline{X}$ respecto de 250. Un valor de $\overline{X}$ que sea mucho menor o mucho mayor que 250 conducirá a rechazar H_0. Mientras que un valor ligeramente inferior o superior a 250 conducirá a no rechazar H_0. Para decidir si $\overline{X}$ está suficientemente próxima a 250 es necesario estandarizar $\overline{X}$. Si suponemos que la distribución de la vida útil de las lámparas eléctricas es normal con desvío estándar conocido, entonces la estandarización de $\overline{X}$ está dada por

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

donde μ_0 es el valor de μ especificado en la hipótesis nula, ya que esta hipótesis se considera como verdadera hasta que por medio de una muestra se exhiba evidencia en contra de ella si la hubiere.

Los valores del estadístico de prueba para los cuales la decisión será rechazar la hipótesis nula constituyen lo que se conoce como la **región crítica o región de rechazo** de la prueba. El **valor crítico** es el punto que separa la región crítica del conjunto de valores para los cuales no se rechaza H_0. El problema consiste en determinar la posición exacta del o los valores críticos. El método tradicional para elegir una región crítica que lleve al rechazo de H_0 consiste en establecer un valor para α. El valor de α indica la importancia que se atribuye a las consecuencias asociadas con un rechazo incorrecto de H_0. Es habitual asignar a α el valor 0.05 o 0.01.
Si en el ejemplo anterior se elige α = 0.05, la región crítica incluye el 5% del área bajo la función de densidad del estadístico que se está utilizando en la prueba, que es la distribución normal estándar. Como la hipótesis alternativa es bilateral, la región óptima separará un área $\alpha/2$ en la cola de la derecha de la distribución y un área $\alpha/2$ en la cola izquierda. En la tabla de la distribución Normal Estándar se puede ver que los valores de Z que separan $\alpha/2$ = 0.025 en cada una de las colas son -1.96 y 1.96.



Establecida la región crítica el próximo paso es, a partir de los datos muestrales, tomar una decisión y determinar, si es posible, el tipo de error que se ha cometido.

El **procedimiento estándar en la prueba de hipótesis** puede resumirse en los siguientes pasos, una vez identificado el parámetro de interés:

1. Plantear las hipótesis nula y alternativa.
2. Seleccionar un nivel de significación α.
3. Determinar el estadístico de prueba apropiado y su distribución de probabilidad bajo H_0.
4. Determinar la región crítica.
5. Calcular a partir de los datos de la muestra el valor del estadístico utilizado en la prueba.
6. Decidir si se rechaza o no H_0.
7. Concluir interpretando los resultados en términos de la variable bajo estudio e indicando, si es posible, el error cometido en la decisión tomada.

La prueba de hipótesis para la media poblacional con varianza conocida puede resumirse de la siguiente manera:

Supuestos: La población X de interés tiene 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 supone que la población está distribuida normalmente o que se aplican las condiciones del teorema central del límite.
Se desea probar H_0: $\mu = \mu_0$ frente a una de las siguientes hipótesis alternativas posibles:

1. H_1: $\mu \neq \mu_0$

2. H_1: $\mu > \mu_0$
3. H_1: $\mu < \mu_0$.

El estadístico de prueba es $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$ que tiene una distribución normal estándar bajo H_0.

La región crítica o de rechazo asociada a cada hipótesis alternativa es:

1. $\{ z \,/\, z < z_{\alpha/2} \;\text{ o }\; z > z_{1-\alpha/2} \}$
2. $\{ z \,/\, z > z_{1-\alpha} \}$
3. $\{ z \,/\, z < z_{\alpha} \}$

Ejemplo:
Un hipermercado ha cambiado recientemente sus máquinas registradoras, lo que hace suponer que el tiempo promedio de facturación por cliente puede haber variado respecto al tiempo promedio de facturación con las máquinas anteriores, que era de 10 minutos. Para probar la validez de esta suposición, se seleccionó una muestra aleatoria de 16 clientes que arrojó un tiempo promedio de facturación por cliente de 8.5 minutos. Es posible suponer que el tiempo de facturación por cliente es una v.a. distribuida normalmente con un desvío estándar de 2.5 minutos.
A un nivel de significación del 5%, ¿se puede concluir que el tiempo promedio de facturación por cliente ha cambiado respecto del valor histórico de 10 minutos?

Sea la v.a. X = “tiempo de facturación por cliente (en minutos)”, $X \sim N(\mu, \sigma^2 = 2.5^2)$.
$n = 16$, $\overline{x} = 8.5$ minutos.

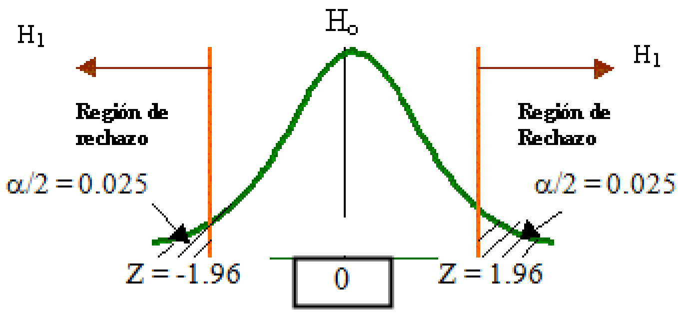
Planteo:

H_0: $\mu = 10$
H_1: $\mu \neq 10$

$\alpha = 0.05$

Estadístico de Prueba bajo H_0 cierta:

$$Z = \frac{\overline{X} - 10}{2.5/\sqrt{16}} \sim N(0, 1)$$



Como $\alpha = 0.05$ entonces $\alpha/2 = 0.025$ y $1-\alpha/2 = 0.975$, por lo tanto $z_{\alpha/2} = -1.96$ y $z_{1-\alpha/2} = 1.96$

Región Crítica:

$$RC = \{ z \,/\, z < z_{\alpha/2} = -1.96 \;\text{ o }\; z > z_{1-\alpha/2} = 1.96 \}$$

El valor del estadístico es:

$$z = \frac{8.5 - 10}{2.5/\sqrt{16}} = \frac{-1.5}{0.625} = -2.4$$

Como $z = -2.4 < -1.96$ entonces rechazo H_0.

Conclusión: Con una probabilidad de error de 0.05, existen evidencias suficientes para afirmar que el tiempo promedio de facturación por cliente con las máquinas nuevas ha variado respecto del tiempo promedio histórico.

Valor P de una prueba de hipótesis

Una manera de presentar los resultados de una prueba de hipótesis es establecer que la hipótesis nula fue o no rechazada con un valor especificado α o nivel de significación. A menudo este planteamiento de las conclusiones resulta inadecuado, ya que no brinda a quien realiza la prueba idea alguna respecto de si el valor del estadístico de prueba calculado está muy próximo al valor crítico que define la región de rechazo o bien se encuentra ubicado dentro de ella. Un segundo método de presentar el resultado de una prueba estadística es mediante el enfoque del valor P. El **valor 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.

Definición:
El **valor P** es el nivel de significación más pequeño que conduce al rechazo de la hipótesis nula H_0. También puede definirse como la probabilidad de error en que se incurriría en caso de Rechazar $H_{0,}$ con los datos que se disponen.

IMPORTANTE: Si el valor P obtenido es menor o igual que valor prefijado como umbral α = 0.05 (**la probabilidad de error máxima que se esta dispuesto a asumir en caso de Rechazar H_0**), entonces se rechaza la hipótesis nula. En caso contrario, No se Rechaza $H_{0.}$ Es decir:

Si P $\leq$ **0.05** entonces Rechaza $H_{0.}$ En este caso se dice que la prueba de Hipótesis es **Significativa.**
Si P > **0.05** entonces No Rechaza H_0. En este caso se dice que la prueba de Hipótesis **no es Significativa.**

Cálculo del valor P para la prueba de hipótesis para la media de una población normal con varianza conocida:

Pruebas unilaterales:
valor P = P($Z \geq z_o$) donde z_o es el valor del estadístico de prueba calculado en los datos de la muestra.
valor P = P($Z \leq z_o$) donde z_o es el valor del estadístico de prueba calculado en los datos de la muestra.

Prueba bilateral:
valor P = P($Z \leq -|z_o|$) + P($Z \geq |z_o|$) = 2 * P($Z \leq -|z_o|$) donde z_o es el valor del estadístico de prueba calculado en los datos de la muestra.

Ejemplo:
En base a los datos del ejemplo anterior el valor P resulta:
valor P = P($Z \leq -|z_o|$) + P($Z \geq |z_o|$) = 2 * P($Z \leq -|-2.4|$) = 2 * 0.0082 = 0.0164.

Interpretación del valor P: Dado que en la prueba de hipótesis anterior resultó un valor P = 0.0164, menor al valor prefijado α = 0.05, entonces **se rechaza** la hipótesis nula, H_0. Luego la conclusión es: con una probabilidad de error 0.0164, existen evidencias suficientes para afirmar que el tiempo promedio de facturación por cliente con las máquinas nuevas ha variado respecto del tiempo promedio histórico.

Prueba de hipótesis para la media de una población normal con varianza desconocida

La población X de interés tiene media μ y varianza σ^2, con σ^2 desconocida. El estadístico de prueba se basa en la media muestral $\overline{X}$. Como X está normalmente distribuida, el estadístico de prueba es $t = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}$ donde s es el desvío muestral.

Se quiere probar H_0: $\mu = \mu_0$ frente a alguna de las hipótesis alternativas planteadas en el caso anterior. Si H_0 es verdadera, el estadístico de prueba $t = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}$ tiene una distribución t de Student con n-1 grados de libertad.

Las respectivas regiones críticas son:

1. $\left\{ t_{n-1} \,/\, t_{n-1} < t_{n-1,\ \alpha/2} \ \text{ o } \ t_{n-1} > t_{n-1,\ 1-\alpha/2} \right\}$

2. $\left\{ t_{n-1} \,/\, t_{n-1} > t_{n-1,\ 1-\alpha} \right\}$

3. $\left\{ t_{n-1} \,/\, t_{n-1} < t_{n-1,\ \alpha} \right\}$

Ejemplo:
El Centro de Atención Telefónica del banco Business debe tardar a lo sumo 4 minutos en resolver cualquier trámite de un usuario, para cumplir con lo establecido por el Departamento de Atención al Cliente. Una muestra de 25 llamados arrojó una duración promedio de 4.5 minutos, con una desviación estándar de 1.4 minutos.
Si se supone que la duración de una llamada telefónica para resolver cualquier trámite es una v.a. con distribución normal, a un nivel de significación del 1% ¿el Centro de atención al Cliente cumple con lo fijado por el Departamento del Atención al Cliente?

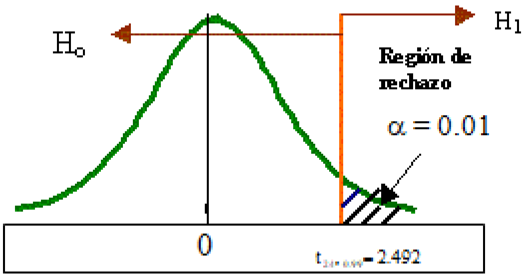
Sea la v.a. X = “duración de una llamada telefónica para resolver cualquier trámite”, $X \sim N(\mu, \sigma^2)$.
σ^2 es desconocido.
$n = 25$, $\overline{x} = 4.5$ minutos y s = 1.4 minutos.

Planteo:

H_0: $\mu \leq 4$ (es equivalente a $\mu = 4$)
H_1: $\mu > 4$

$\alpha = 0.01$

Estadístico de Prueba bajo H_0 cierta



$$t = \frac{\overline{X} - 4}{1.4/\sqrt{25}} \sim t_{n-1}$$

Como $\alpha = 0.01$ entonces $1-\alpha = 0.99$, por lo tanto $t_{24,\ 0.99} = 2.492$

Región Crítica:
$RC = \left\{ t_{24} \,/\, t_{24} > t_{24,\ 0.99} = 2.492 \right\}$

El valor del estadístico es:

$$t = \frac{4{,}5 - 4}{1.4/\sqrt{25}} = \frac{0.5}{0.28} = 1.79$$

Como $t = 1.79 < 2.492$ entonces No rechazo H_0.

Conclusión: Con un nivel de significación de 0.01, no existen evidencias suficientes para afirmar que la duración promedio de una llamada telefónica para resolver cualquier trámite sea superior a los 4 minutos. El Centro de atención al Cliente cumple con lo fijado por el Departamento del Atención al Cliente.

Cálculo del valor P para la prueba de hipótesis para la media de una población normal con varianza desconocida:

Pruebas unilaterales:
valor P = $P(t_{n-1} \geq t_o)$ donde t_o es el valor del estadístico de prueba calculado en los datos de la muestra.
valor P = $P(t_{n-1} \leq t_o)$ donde t_o es el valor del estadístico de prueba calculado en los datos de la muestra.

Prueba bilateral:
valor P = $P(t_{n-1} \leq -|t_o|) + P(t_{n-1} \geq |t_o|) = 2 * P(t_{n-1} \leq -|t_o|)$ donde t_o es el valor del estadístico de prueba calculado en los datos de la muestra.

Ejemplo:
En base a los datos del ejemplo anterior el valor P resulta:
valor P = $P(t_{n-1} \geq t_o) = P(t_{25-1} \geq 1.79) = 0.04557$. Por lo tanto, dado que el valor P = $0.04557 < 0.05$, entonces **se rechaza** la hipótesis nula H_0.

Conclusión: Con una probabilidad de error de 0.04557, existen evidencias suficientes para afirmar que la duración promedio de una llamada telefónica para resolver cualquier trámite es superior a los 4 minutos. El Centro de atención al Cliente no cumple con lo fijado por el Departamento del Atención al Cliente.

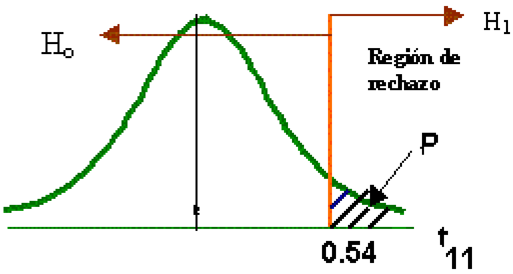
Ejemplo:
Los tiempos de sobrevivencia (en años) de 12 personas que se han sometido a un transplante de corazón son los siguientes:

3.1 0.9 2.8 4.3 0.6 1.4 5.8 9.9 6.3 10.4 0 11.5

Un cardiocirujano afirma que el tiempo de vida promedio de las personas sometidas a transplante de corazón es mayor que 4 años. ¿A qué conclusión se llegará? Suponer que el tiempo de sobrevivencia de una persona es una v.a. Normal. Resolverlo mediante el cálculo del valor P.

Sea la v.a. X = “tiempo de sobrevivencia (en años) de una persona, $X \sim N(\mu, \sigma^2)$, σ^2 es desconocido.
n = 100 propietarios, $\bar{x} = 4.7$años y s = 4.5 años.

Planteo:



H0: $\mu = 4$
H1: $\mu > 4$

Estadístico de Prueba bajo H0 cierta:

$$t_{n-1=11} \cong \frac{\overline{X}-4}{s/\sqrt{12}} \ .$$

El valor del estadístico es: $t = \dfrac{4.7-4}{4.5/\sqrt{12}} = 0.54.$

valor P = P($t_{11} \geq 0.54$) = 0.29998. Por lo tanto, dado que el valor P = 0.29998 > 0.05, entonces **no se rechaza** la hipótesis nula H_0.

Conclusión: Con un nivel de significación de 0.29998, no existen evidencias suficientes para afirmar que el tiempo de vida promedio de las personas sometidas a transplante de corazón es mayor que 4 años.

Observación importante:
Cuando se prueban hipótesis sobre la media μ de una población con σ^2 desconocida, es posible utilizar los procedimientos de prueba vistos anteriormente para la media de una población normal con varianza conocida siempre que el tamaño de la muestra sea grande, por ejemplo $n \geq 30$. La varianza muestral s^2 estará próxima a σ^2 en la mayor parte de las muestras, de modo que es posible sustituir σ por s en los procedimientos de prueba. Estos procedimientos son aproximadamente válidos (como consecuencia del teorema central del límite) sin importar si la población de interés es normal o no. Sin embargo, cuando la muestra es pequeña y σ^2 es desconocida, debe establecerse la distribución subyacente de la población de interés X con la finalidad de obtener un procedimiento de prueba. En muchos casos es razonable suponer que la distribución subyacente es normal.

Cálculo de α y β para la prueba de hipótesis para la media de una población normal con varianza conocida

Considérese los siguientes ejemplos:
1. Un fabricante de barras de acero considera que el proceso de producción funciona correctamente cuando la longitud media de las barras es de 8.6 pulgadas. La longitud de estas barras tiene una distribución normal con un desvío estándar de 0.3 pulgadas. Se desea probar si el proceso de producción funciona correctamente, para ello se plantea la hipótesis H_0: μ = 8.6. Se decide rechazar H_0 toda vez que la longitud media de una muestra de 36 barras seleccionadas al azar sea inferior a 8.52 pulgadas o superior a 8.68 pulgadas. ¿Cuál es la probabilidad de rechazar H_0 siendo H_0 es verdadera?

La región crítica puede expresarse de la siguiente forma:

$$RC = \left\{ \overline{X} \ / \ \overline{X} < 8.52 \ \text{ o } \ \overline{X} > 8.68 \right\}$$

$$\alpha = P(\text{ Rechazar } H_0 / H_0 \text{ verdadera}) = P(\overline{X} < 8.52 / \mu = 8.6) + P(\overline{X} > 8.68 / \mu = 8.6)$$

$$= P(Z < \frac{8.52 - 8.6}{0.3/\sqrt{36}}) + P(Z > \frac{8.68 - 8.6}{0.3/\sqrt{36}}) = P(Z < -1.6) + P(Z > 1.6) = 2 * 0.0548 = 0.1096.$$

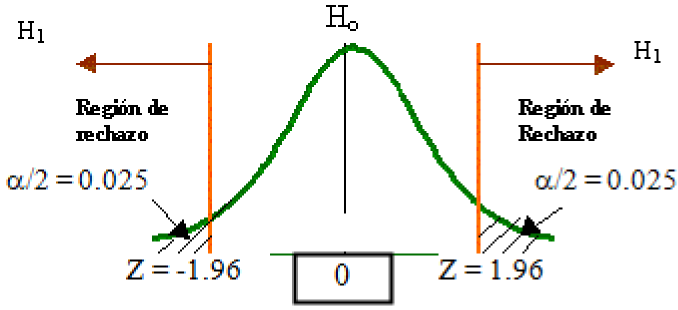
Por lo tanto $\alpha = 0.1096$.

2. Una firma que fabrica arandelas desea averiguar si su proceso de producción funciona correctamente, es decir, si el número de arandelas colocadas en cada caja tiene una media $\mu = 1000$ con un desvío estándar $\sigma = 37.5$. Se supone que la cantidad de arandelas colocadas en cada caja tiene una distribución aproximadamente normal. Para verificar el funcionamiento del proceso de producción se tomó una muestra aleatoria de 9 cajas obteniéndose una cantidad promedio de arandelas de 985 por caja.

   a) A un nivel de significación del 5%, ¿el proceso de producción funciona correctamente?
   b) ¿cuál es la probabilidad de cometer error tipo II suponiendo que en realidad la cantidad promedio de arandelas por caja es de 970?

a) Sea la v.a. X = “cantidad de arandelas por caja” , $X \sim N(\mu, \sigma^2 = 37.5^2)$.
$n = 9$, $\overline{x} = 985$

Planteo:



H_0: $\mu = 1000$
H_1: $\mu \neq 1000$

$\alpha = 0.05$

Estadístico de Prueba bajo H_0 cierta:

$$Z = \frac{\overline{X} - 1000}{37.5/\sqrt{9}} \sim N(0, 1)$$

Como $\alpha = 0.05$ entonces $\alpha/2 = 0.025$ y $1 - \alpha/2 = 0.975$, por lo tanto $z_{\alpha/2} = -1.96$ y $z_{1-\alpha/2} = 1.96$

Región Crítica:
$RC = \left\{ z \,/\, z < z_{\alpha/2} = -1.96 \text{ o } z > z_{1-\alpha/2} = 1.96 \right\}$. Por lo tanto, la **región crítica expresada en términos de la variable** $\overline{X}$ se obtiene de la siguiente manera:

$$\overline{x}_1 = -1.96 * \frac{37.5}{\sqrt{9}} + 1000 = 975.5 \quad \text{y} \quad \overline{x}_2 = 1.96 * \frac{37.5}{\sqrt{9}} + 1000 = 1024.5$$

Luego la región crítica es:

$$RC = \left\{ \overline{X} \,/\, \overline{X} < 975.5 \text{ o } \overline{X} > 1024.5 \right\}$$

Como $975.5 < \overline{x} = 985 < 1024.5$ entonces no rechazo H_0.

Conclusión: Con un nivel de significación de 0.05, no existen evidencias suficientes para afirmar que la cantidad promedio de arandelas por caja difiere de 1000.

b) Para responder a este inciso es necesario calcular β para el valor de $\mu = 970$.

$\beta = P(\text{ No Rechazar } H_0 / H_0 \text{ falsa}) = P(975.5 < \overline{X} < 1024.5 / \mu = 970) =$

$= P\left(\frac{975.5 - 970}{37.5/\sqrt{9}} < Z < \frac{1024.5 - 970}{37.5/\sqrt{9}}\right) = P(0.44 < Z < 4.36) \cong 1.0000 - 0.6700 = 0.3300.$

Luego la probabilidad de cometer error tipo II suponiendo que en realidad la cantidad promedio de arandelas por caja es de 970 es $\beta \cong 0.33$.
De aquí se deduce que la potencia de la prueba es $\pi = 1 - \beta \cong 1 - 0.33 = 0.67$

Prueba de hipótesis para la varianza de una población normal

Se desea probar la hipótesis nula de que la varianza de una población normal σ^2 es igual a un valor específico, por ejemplo σ_0^2. Es decir, dada una muestra aleatoria de *n* observaciones de esta población se quiere probar H_0: $\sigma^2 = \sigma_0^2$ frente a una de las siguientes hipótesis alternativas posibles:

1. H_1: $\sigma^2 \neq \sigma_0^2$
2. H_1: $\sigma^2 > \sigma_0^2$
3. H_1: $\sigma^2 < \sigma_0^2$.

El estadístico de prueba es $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ donde S^2 es la varianza muestral. Si H_0 es verdadera, el estadístico de prueba tiene una distribución chi-cuadrada con n-1 grados de libertad.

Las respectivas regiones críticas son:

1. $\left\{\chi^2{}_{n-1} / \chi^2{}_{n-1} < \chi^2{}_{n-1,\ \alpha/2} \text{ o } \chi^2{}_{n-1} > \chi^2{}_{n-1,\ 1-\alpha/2}\right\}$
2. $\left\{\chi^2{}_{n-1} / \chi^2{}_{n-1} > \chi^2{}_{n-1,\ 1-\alpha}\right\}$
3. $\left\{\chi^2{}_{n-1} / \chi^2{}_{n-1} < \chi^2{}_{n-1,\ \alpha}\right\}$

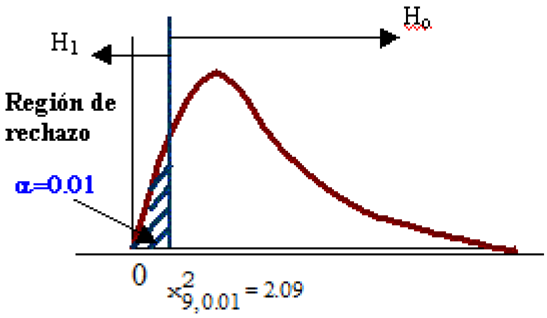
Observación importante:
Si se desea probar una afirmación respecto del desvío estándar σ de una población, se deberá plantear las hipótesis y realizar la prueba en términos de la varianza poblacional σ^2, puesto que el estadístico con distribución conocida $\chi^2{}_{n-1}$, está en función de σ^2.
Ejemplo:
Un supervisor de control de calidad en una enlatadora sabe que la cantidad exacta contenida en cada lata varía, pues hay ciertos factores imposibles de controlar que afectan la cantidad de llenado. La variación de la cantidad de llenado es importante; si es grande, algunas latas contendrán muy poco y otros demasiado. Las agencias reguladoras especifican que el desvío estándar de la cantidad de llenado debe ser menor que 0.1 onzas. El supervisor de control de calidad muestreó $n = 10$ latas y midió la cantidad de llenado en cada una, obteniendo un desvío estándar igual a 0.043. ¿Proporciona esta información prueba suficiente de que el desvío estándar σ de las mediciones de llenado es menor que 0.1 onzas, trabajando a un nivel de significación del 1%? Se puede suponer que la cantidad de llenado sigue una distribución normal.

Sea la v.a. X = "cantidad de llenado por lata", $X \sim N(\mu, \sigma^2)$.
$n = 10$ y s = 0.043 minutos.

Planteo:

H_0: $\sigma^2 = 0.01$ (es equivalente a $\sigma^2 \geq 0.01$)
H_1: $\sigma^2 < 0.01$

$\alpha = 0.01$



Estadístico de prueba bajo H_0:

$$\frac{(n-1)S^2}{\sigma^2} = \frac{(10-1)S^2}{0.01} = \chi^2_{n-1}$$

Región crítica:

RC = $\{ \chi^2_{n-1} \,/\, \chi^2_{n-1} < \chi^2_{n-1,\alpha} \} = \{ \chi^2_9 \,/\, \chi^2_9 < \chi^2_{9,0.01} \} = \{ \chi^2_9 \,/\, \chi^2_9 < 2.09 \}$.

El valor del estadístico es:

$$\chi^2_9 = \frac{9 * (0.043)^2}{0.01} = 1.66$$

Como $\chi^2_9 = 166 \in$ Región crítica, por lo tanto rechazo H_0.

Conclusión: Con una probabilidad de error del 0.01, la varianza de la población de todas las cantidades de llenado es menor que 0.01 (o σ menor a 0.1), por lo tanto la enlatadora está operando dentro de los límites de variabilidad deseados.

Fin Capítulo VII