

9- Test o prueba de hipótesis

9.1 – Introducción

Hasta ahora hemos estudiado el problema de estimar un parámetro desconocido a partir de una muestra aleatoria.

En muchos problemas se requiere tomar una decisión entre aceptar o rechazar una proposición sobre algún parámetro. Esta proposición recibe el nombre de **hipótesis estadística**, y el procedimiento de toma de decisión sobre la hipótesis se conoce como **prueba o test de hipótesis**.

Como se emplean distribuciones de probabilidad para representar poblaciones, también podemos decir que una hipótesis estadística es una proposición sobre la distribución de probabilidad de una variable aleatoria, donde la hipótesis involucra a uno más parámetros de esta distribución.

Por ejemplo, supongamos que cierto tipo de motor de automóvil emite una media de 100 mg de óxidos de nitrógeno (NO_x) por segundo con 100 caballos de fuerza. Se ha propuesto una modificación al diseño del motor para reducir las emisiones de NO_x . El nuevo diseño se producirá si se demuestra que la media de su tasa de emisiones es menor de 100 mg/s. Se construye y se prueba una muestra de 50 motores modificados. La media muestral de emisiones de NO_x es de 92 mg/s, y la desviación estándar muestral es de 21 mg/s.

La variable aleatoria de interés en este caso es X : “tasa de emisión de un motor modificado tomado al azar”.

La preocupación de los fabricantes consiste en que los motores modificados no puedan reducir todas las emisiones; es decir que la media poblacional pudiera ser 100 o mayor que 100.

Entonces, la pregunta es: ¿es factible que esta muestra pueda provenir de una v.a. con media 100 o mayor?

Éste es el tipo de preguntas que las pruebas de hipótesis están diseñadas para responder. Veremos cómo construir una prueba de hipótesis, pero podemos decir que en general se basa en construir a partir de la muestra aleatoria un **estadístico**, y según el valor que tome este **estadístico de prueba** se aceptará o se rechazará la hipótesis.

Se ha observado una muestra con media $\bar{X} = 92$.

Hay dos interpretaciones posibles de esta observación:

1- La media poblacional es realmente mayor o igual que 100, y la media muestral es menor que 100 debido a la **variabilidad** propia de la variable aleatoria \bar{X}

2- La media poblacional es en realidad menor que 100, y la media muestral refleja este hecho.

Estas dos explicaciones tienen nombres: la primera se llama **hipótesis nula**; la segunda es la **hipótesis alternativa**.

En la mayoría de las situaciones la hipótesis nula dice que el efecto que indica la muestra es atribuible solamente a la variación aleatoria del estadístico de prueba.

La hipótesis alternativa establece que el efecto que indica la muestra es verdadero.

Para hacer las cosas más precisas, todo se expresa mediante símbolos. La hipótesis nula se denota por H_0 , la hipótesis alternativa se denota con H_1 . Como es usual la media poblacional se anota μ . Por lo tanto se tiene

$$H_0 : \mu \geq 100 \quad \text{contra} \quad H_1 : \mu < 100 \quad (\text{hipótesis alternativa unilateral})$$

Esencialmente, para realizar una prueba de hipótesis se pone la hipótesis nula en juicio. Se asume que H_0 es verdadera, de la misma manera como se empieza en un juicio bajo el supuesto de que un acusado es inocente. La muestra aleatoria proporciona la evidencia.

Las hipótesis son siempre proposiciones *sobre los parámetros de la población o distribución bajo estudio*, no proposiciones sobre la muestra.

Otros tipos de hipótesis que podrían formularse son

$$H_0 : \mu \leq 100 \quad \text{contra} \quad H_1 : \mu > 100 \quad (\text{hipótesis alternativa unilateral})$$

o

$$H_0 : \mu = 100 \quad \text{contra} \quad H_1 : \mu \neq 100 \quad (\text{hipótesis alternativa bilateral})$$

En el ejemplo tenemos X_1, X_2, \dots, X_{50} muestra aleatoria de la v.a. X definida anteriormente.

Como estamos haciendo una hipótesis sobre la media poblacional es razonable tomar como estadístico de prueba a \bar{X} . El valor observado de la media muestral es $\bar{X} = 92$.

Si el valor de \bar{X} es muy “menor” que 100 entonces se considera que hay evidencia en contra H_0 y se la rechaza, aceptando la hipótesis alternativa.

Si el valor de \bar{X} no es “muy menor” que 100 entonces se considera que no hay evidencia en contra H_0 y se rechaza la hipótesis alternativa.

Ya veremos como construir una **regla de decisión**, supongamos ahora que tenemos la siguiente regla:

$$\begin{cases} \text{se rechaza } H_0 & \text{si } \bar{X} < 95 \\ \text{se acepta } H_0 & \text{si } \bar{X} \geq 95 \end{cases}$$

El intervalo $(95, \infty)$ es la **zona de aceptación**.

La región $(-\infty; 95)$ es la **zona de rechazo o región crítica**.

Mientras que 95 es el **punto crítico**.

Como estamos tomando una decisión basados en el valor de un estadístico podemos cometer dos tipos de errores: rechazar H_0 cuando ésta es verdadera, es decir el estadístico toma valores en la zona de rechazo cuando H_0 es verdadera; o aceptar H_0 cuando ésta es falsa, es decir que el estadístico tome valores en la zona de aceptación cuando H_0 es falsa.

El primero se conoce como **error de tipo I**, y el segundo como **error de tipo II**.

Debido a que la decisión se basa en variables aleatorias es posible asociar probabilidades a los errores de tipo I y II, específicamente anotamos

$$\alpha = P(\text{error de tipo I})$$

$$\beta = P(\text{error de tipo II})$$

A $\alpha = P(\text{error de tipo I})$ se lo conoce como **nivel de significancia del test**.

Para calcular estas probabilidades debemos conocer la distribución del estadístico de prueba en el caso de ser H_0 verdadera, es decir debemos conocer la distribución del estadístico de prueba “bajo H_0 ”.

En el ejemplo anterior la muestra es grande, ya sabemos que por T.C.L. el estadístico

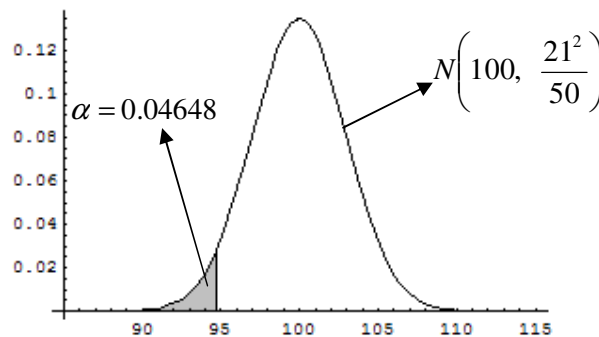
$$Z = \frac{\bar{X} - 100}{s/\sqrt{n}} \approx N(0,1) \text{ si } H_0 \text{ es verdadera, o sea } Z = \frac{\bar{X} - 100}{21/\sqrt{50}} \approx N(0,1)$$

Entonces para calcular α planteamos:

$$\begin{aligned} \alpha &= P(\text{error de tipo I}) = P\left(\text{rechazar } H_0 / H_0 \text{ es V}\right) = P(\bar{X} < 95 / \mu = 100) = \\ &= P\left(\frac{\bar{X} - 100}{21/\sqrt{50}} < \frac{95 - 100}{21/\sqrt{50}}\right) \approx \Phi\left(\frac{95 - 100}{21/\sqrt{50}}\right) = \Phi(-1.6835) = 1 - 0.95352 = 0.04648 \end{aligned}$$

Esto significa que el 4.64% de las muestras aleatorias conducirán al rechazo de la hipótesis $H_0 : \mu \geq 100$ cuando el verdadero μ sea mayor o igual que 100.

En este caso el gráfico de la zona de rechazo es



Del gráfico anterior vemos que podemos reducir α al aumentar la zona de aceptación. Por ejemplo supongamos que ahora la regla de decisión es

$$\begin{cases} \text{se rechaza } H_0 & \text{si } \bar{X} < 93 \\ \text{se acepta } H_0 & \text{si } \bar{X} \geq 93 \end{cases}$$

$$\begin{aligned} \text{Entonces } \alpha &= P(\text{error de tipo I}) = P\left(\text{rechazar } H_0 / H_0 \text{ es V}\right) = P(\bar{X} < 93 / \mu = 100) = \\ &= P\left(\frac{\bar{X} - 100}{21/\sqrt{50}} < \frac{93 - 100}{21/\sqrt{50}}\right) \approx \Phi\left(\frac{93 - 100}{21/\sqrt{50}}\right) = \Phi(-2.357) = 1 - 0.99061 = 0.00939 \end{aligned}$$

También se puede reducir α aumentando el tamaño de la muestra. Supongamos que $n = 85$, entonces

$$\begin{aligned}\alpha &= P(\text{error de tipo I}) = P(\text{rechazar } H_0 / H_0 \text{ es V}) = P(\bar{X} < 95 / \mu = 100) = \\ &= P\left(\frac{\bar{X} - 100}{21/\sqrt{85}} < \frac{95 - 100}{21/\sqrt{85}}\right) \approx \Phi\left(\frac{95 - 100}{21/\sqrt{85}}\right) = \Phi(-2.195) = 1 - 0.98574 = 0.01426\end{aligned}$$

También es importante examinar la probabilidad de cometer error de tipo II, esto es

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

Pero en este caso para llegar a un valor numérico necesitamos tener una alternativa específica pues en nuestro ejemplo:

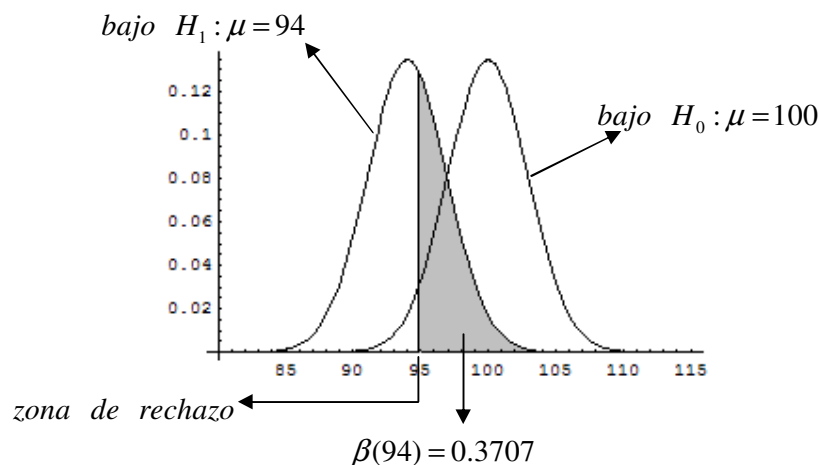
$$\begin{aligned}\beta &= P(\text{error de tipo II}) = P(\text{aceptar } H_0 / H_0 \text{ es falsa}) = P(\bar{X} \geq 95 / \mu \neq 100) = \\ &= P\left(\frac{\bar{X} - \mu}{21/\sqrt{50}} \geq \frac{95 - \mu}{21/\sqrt{50}}\right) = 1 - \Phi\left(\frac{95 - \mu}{21/\sqrt{50}}\right) = \beta(\mu)\end{aligned}$$

Donde anotamos con μ a la verdadera media poblacional **desconocida**.

Podemos entonces calcular β para un valor particular de μ , por ejemplo nos puede interesar como se comporta el test cuando la verdadera media es $\mu = 94$, entonces

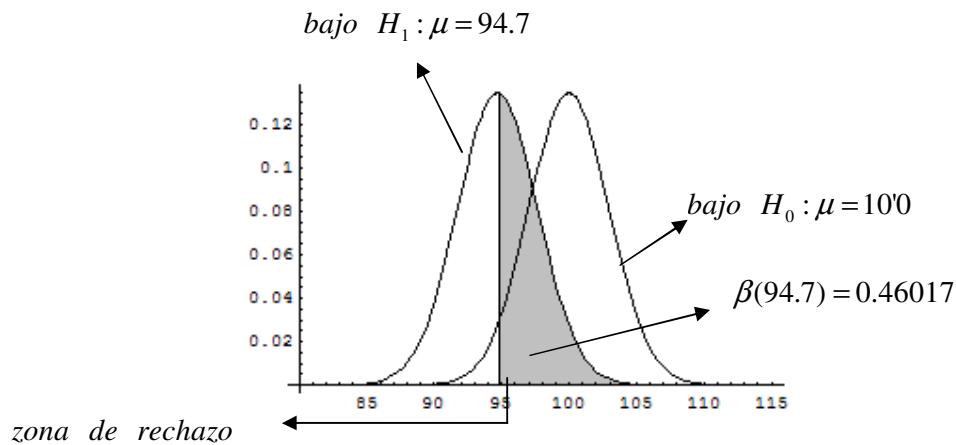
$$\beta(94) = P\left(\frac{\bar{X} - 94}{21/\sqrt{50}} \geq \frac{95 - 94}{21/\sqrt{50}}\right) = 1 - \Phi\left(\frac{95 - 94}{21/\sqrt{50}}\right) = 1 - \Phi(0.3367) = 1 - 0.62930 = 0.3707$$

Gráficamente:



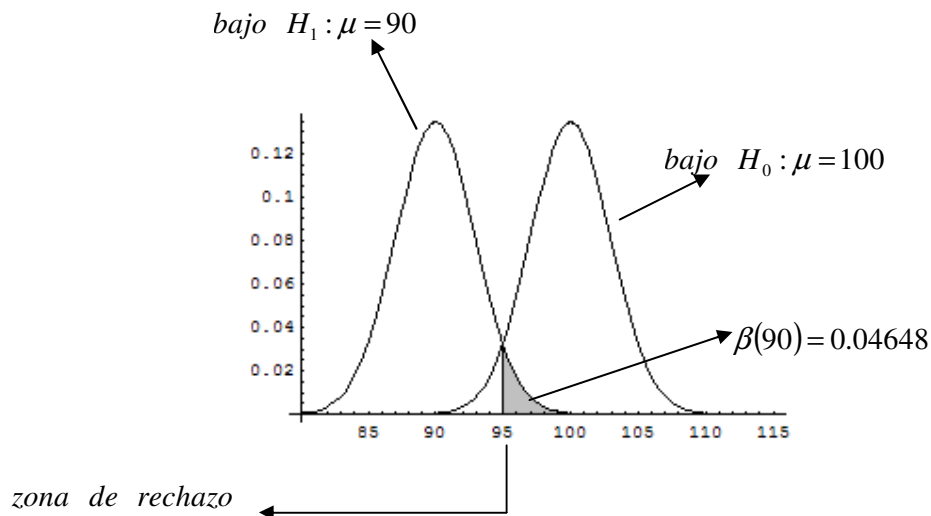
La probabilidad β de cometer error de tipo II **crece** a medida que el valor verdadero de μ se acerca al valor hipotético. Por ejemplo si el verdadero valor de μ fuera 94.7 entonces

$$\beta(94.7) = P\left(\frac{\bar{X} - 94.7}{21/\sqrt{50}} \geq \frac{95 - 94.7}{21/\sqrt{50}}\right) = 1 - \Phi\left(\frac{95 - 94.7}{21/\sqrt{50}}\right) = 1 - \Phi(0.101015) = 1 - 0.53983 = 0.46017$$



Además, la probabilidad β de cometer error de tipo II **disminuye** a medida que el valor verdadero de μ se aleja del valor hipotético. Por ejemplo si el verdadero valor de μ fuera 90 entonces

$$\beta(90) = P\left(\frac{\bar{X} - 90}{21/\sqrt{50}} \geq \frac{95 - 90}{21/\sqrt{50}}\right) = 1 - \Phi\left(\frac{95 - 90}{21/\sqrt{50}}\right) = 1 - \Phi(1.6835) = 1 - 0.95352 = 0.04648$$



También se puede reducir la probabilidad de cometer error de tipo II con el tamaño de la muestra. Por ejemplo si $n = 85$ entonces y $\mu = 94$

$$\beta(94) = P\left(\frac{\bar{X} - 94}{21/\sqrt{85}} \geq \frac{95 - 94}{21/\sqrt{85}}\right) = 1 - \Phi\left(\frac{95 - 94}{21/\sqrt{85}}\right) = 1 - \Phi(0.4390) = 1 - 0.67003 = 0.32997$$

Lo que se ha visto en los ejemplos anteriores se puede generalizar. Podemos recalcar los siguientes puntos importantes:

- 1- El tamaño de la región crítica, y en consecuencia la probabilidad α de cometer error de tipo I, siempre pueden reducirse mediante una selección apropiada de los valores críticos.
- 2- Los errores tipo I y 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 que el tamaño de la muestra no cambie.
- 3- En general, un aumento en el tamaño de la muestra reduce tanto a α como a β , siempre que los valores críticos se mantengan constantes.
- 4- 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 deferencia entre el verdadero valor medio y el propuesto.

En general el investigador controla la probabilidad α del error de tipo I cuando selecciona los valores críticos. Por lo tanto el rechazo de la hipótesis nula de manera errónea se puede fijar de antemano. Eso hace que rechazar la hipótesis nula sea una **conclusión fuerte**.

La probabilidad β de error de tipo II no es constante, sino que depende del valor verdadero del parámetro. También depende β del tamaño de la muestra que se haya seleccionado. Como β está en función del tamaño de la muestra y del valor verdadero del parámetro, la decisión de aceptar la hipótesis nula se la considera una **conclusión débil**, a menos que se sepa que β es aceptablemente pequeño. Por lo tanto **cuando se acepta H_0 en realidad se es incapaz de rechazar H_0 . No se puede rechazar H_0 pues no hay evidencia en contra H_0 .**

Un concepto importante es el siguiente:

La **potencia** de un test es la probabilidad de rechazar la hipótesis nula. La simbolizamos $\pi(\mu)$. Para los **valores de μ tal que la alternativa es verdadera** se tiene

$$\pi(\mu) = P\left(\text{rechazar } H_0 / H_0 \text{ es falsa}\right) = 1 - \beta(\mu)$$

Las pruebas estadísticas se comparan mediante la comparación de sus propiedades de potencia.

La potencia es una medida de la **sensibilidad** del test, donde por sensibilidad se entiende la capacidad de una prueba para detectar diferencias.

En el ejemplo anterior, la sensibilidad de la prueba para detectar la diferencia entre una tasa de emisión media de 100 y otra de 94 es $\pi(94) = 1 - \beta(94) = 1 - 0.3707 = 0.6293$. Es decir si el valor verdadero de la tasa de emisión media es 94, la prueba rechazará de manera correcta H_0 y detectará esta diferencia el 62.93% de las veces. Si el investigador piensa que este valor es bajo entonces el investigador puede aumentar α o el tamaño de la muestra.

9.2 – Prueba de hipótesis sobre la media, varianza conocida

Veamos ahora cómo construir una regla de decisión sobre la media de una población.

Supongamos que la variable aleatoria de interés X tiene una media μ y una varianza σ^2 conocida.

Asumimos que X tiene distribución normal, es decir $X \sim N(\mu, \sigma^2)$.

Nuevamente, como en el ejemplo introductorio, es razonable tomar como estadístico de prueba al promedio muestral \bar{X} . Bajo las suposiciones hechas tenemos que $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

Supongamos que tenemos las hipótesis

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu \neq \mu_0$$

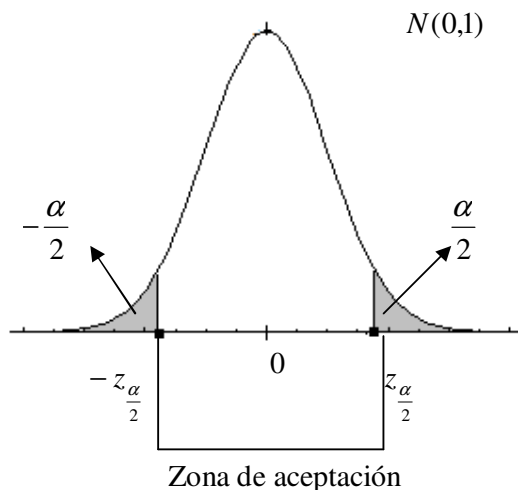
Donde μ_0 es una constante específica. Se toma una muestra aleatoria X_1, X_2, \dots, X_n de la población.

Si $H_0 : \mu = \mu_0$ es verdadera, entonces $\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$, por lo tanto el estadístico

$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ tiene distribución $N(0,1)$ si $H_0 : \mu = \mu_0$ es verdadera

Tomamos a Z como **estadístico de prueba**

Si $H_0 : \mu = \mu_0$ es verdadera entonces $P\left(-z_{\frac{\alpha}{2}} \leq Z \leq z_{\frac{\alpha}{2}}\right) = 1 - \alpha$



Es evidente que una muestra que produce un valor del estadístico de prueba que cae en las colas de la distribución de Z será inusual si $H_0 : \mu = \mu_0$ es verdadera, por lo tanto esto es un indicador que H_0 es falsa.

Entonces la regla de decisión es:

$$\begin{cases} \text{rechazar } H_0 & \text{si } |Z| > z_{\frac{\alpha}{2}} \\ \text{aceptar } H_0 & \text{si } |Z| \leq z_{\frac{\alpha}{2}} \end{cases}$$

Notar que la probabilidad que la estadística de prueba tome un valor que caiga en la zona de rechazo si H_0 es verdadera es igual a α , es decir la probabilidad de cometer error de tipo I es α pues

$$P(\text{error de tipo I}) = P(\text{rechazar } H_0 / H_0 \text{ es V}) = P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| > z_{\frac{\alpha}{2}} / \mu = \mu_0\right) =$$

$$= P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > z_{\frac{\alpha}{2}}\right) + P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -z_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha$$

Ejemplo:

El porcentaje deseado de SiO_2 en cierto tipo de cemento aluminoso es 5.5. Para probar si el verdadero promedio de porcentaje es 5.5 para una planta de producción en particular, se analizaron 16 muestras obtenidas de manera independiente. Supongamos que el porcentaje de SiO_2 en una muestra está normalmente distribuido con $\sigma = 0.3$, y que $\bar{x} = 5.25$.

¿Indica esto de manera concluyente que el verdadero promedio de porcentaje difiere de 5.5?. Utilice $\alpha = 0.01$

Solución:

La v.a. de interés es X : “porcentaje de SiO_2 en cierto tipo de cemento aluminoso”

Asumimos que $X \sim N(\mu, 3^2)$

Podemos plantear las hipótesis

$$H_0 : \mu = 5.5 \quad \text{contra} \quad H_1 : \mu \neq 5.5$$

Tenemos una muestra de tamaño $n = 16$ que dio un promedio muestral $\bar{x} = 5.25$

Como $\alpha = 0.01$ entonces $z_{\frac{\alpha}{2}} = z_{0.005} = 2.575$

Por lo tanto la regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } \left| \frac{\bar{X} - 5.5}{3/\sqrt{16}} \right| > 2.575 \\ \text{aceptar } H_0 & \text{si } \left| \frac{\bar{X} - 5.5}{3/\sqrt{16}} \right| \leq 2.575 \end{cases}$$

El estadístico $\left| \frac{\bar{X} - 5.5}{3/\sqrt{16}} \right|$ toma el valor $z_0 = \left| \frac{5.25 - 5.5}{3/\sqrt{16}} \right| = 0.333333$

Como $z_0 = 0.333333 < 2.575 = z_{\frac{0.01}{2}}$ se acepta H_0

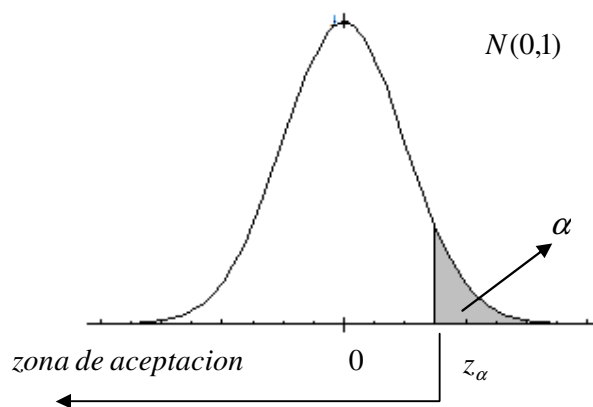
También podemos desarrollar tests o pruebas de hipótesis para el caso de que la hipótesis alternativa es unilateral.

Supongamos las hipótesis

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu > \mu_0$$

En este caso la región crítica debe colocarse en la cola superior de la distribución normal estándar y el rechazo de H_0 se hará cuando el valor calculado de z_0 sea muy grande, esto es la regla de decisión será

$$\left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha \\ \text{aceptar } H_0 \text{ si } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_\alpha \end{array} \right.$$

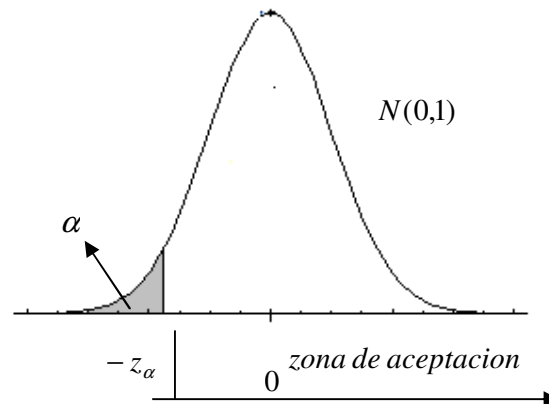


De manera similar para las hipótesis

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu < \mu_0$$

se calcula el valor del estadístico de prueba z_0 y se rechaza H_0 si el valor de z_0 es muy pequeño, es decir la regla de decisión será

$$\left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -z_\alpha \\ \text{aceptar } H_0 \text{ si } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq -z_\alpha \end{array} \right.$$



Ejemplo:

Se sabe que la duración, en horas, de un foco de 75 watts tiene una distribución aproximadamente normal, con una desviación estándar de $\sigma = 25$ horas. Se toma una muestra aleatoria de 20 focos, la cual resulta tener una duración promedio de $\bar{x} = 1040$ horas

¿Existe evidencia que apoye la afirmación de que la duración promedio del foco es mayor que 1000 horas?. Utilice $\alpha = 0.05$.

Solución:

La v.a. de interés es X: “duración en horas de un foco tomado al azar”

Asumimos $X \sim N(\mu, 25^2)$

Podemos plantear las hipótesis

$$H_0 : \mu = 1000 \quad \text{contra} \quad H_1 : \mu > 1000$$

Tenemos una muestra de tamaño $n = 20$ que dio un promedio muestral $\bar{x} = 1040$

Como $\alpha = 0.05$ entonces $z_\alpha = z_{0.05} = 1.645$

$$\text{Por lo tanto la regla de decisión es } \begin{cases} \text{rechazar } H_0 & \text{si } \frac{\bar{X} - 1000}{25/\sqrt{20}} > 1.645 \\ \text{aceptar } H_0 & \text{si } \frac{\bar{X} - 1000}{25/\sqrt{20}} \leq 1.645 \end{cases}$$

$$\text{El estadístico toma el valor } Z = \frac{\bar{X} - 1000}{25/\sqrt{20}} \text{ toma el valor } z_0 = \frac{1040 - 1000}{25/\sqrt{20}} = 7.1554$$

Como $z_0 = 7.1554 > 1.645 = z_{0.05}$ se rechaza H_0

P- valor

Hasta ahora se dieron los resultados de una prueba de hipótesis estableciendo si la hipótesis nula fue o no rechazada con un valor especificado de α o nivel de significancia.

A menudo este planteamiento resulta inadecuado, ya que no proporciona ninguna idea sobre si el valor calculado del estadístico está apenas en la región de rechazo o bien ubicado dentro de ella. Además, esta forma de establecer los resultados impone a otros usuarios el nivel de significancia predeterminado.

Para evitar estas dificultades, se adopta el enfoque del **p-valor**. El **valor p** o **p-valor** 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 es verdadera. Es así como el p-valor da mucha información sobre el peso de la evidencia contra H_0 , de modo que el investigador pueda llegar a una conclusión para cualquier nivel de significancia especificado.

La definición formal del p-valor es la siguiente:

El **valor p** es el nivel de significancia más pequeño que conduce al rechazo de la hipótesis nula H_0

Para las pruebas de distribuciones normales presentadas hasta el momento, es sencillo calcular el p-valor.

Si z_0 es el valor calculado del estadístico de prueba Z , entonces el p-valor es

a) si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$

$$p\text{-valor} = P(|Z| > |z_0|) = 1 - P(|Z| < |z_0|) = 1 - [\Phi(|z_0|) - \Phi(-|z_0|)] = 1 - [2\Phi(|z_0|) - 1] = 2[1 - \Phi(|z_0|)]$$

b) si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu > \mu_0$

$$p\text{-valor} = P(Z > z_0) = 1 - P(Z \leq z_0) = 1 - \Phi(z_0)$$

c) si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu < \mu_0$

$$p\text{-valor} = P(Z < z_0) = \Phi(z_0)$$

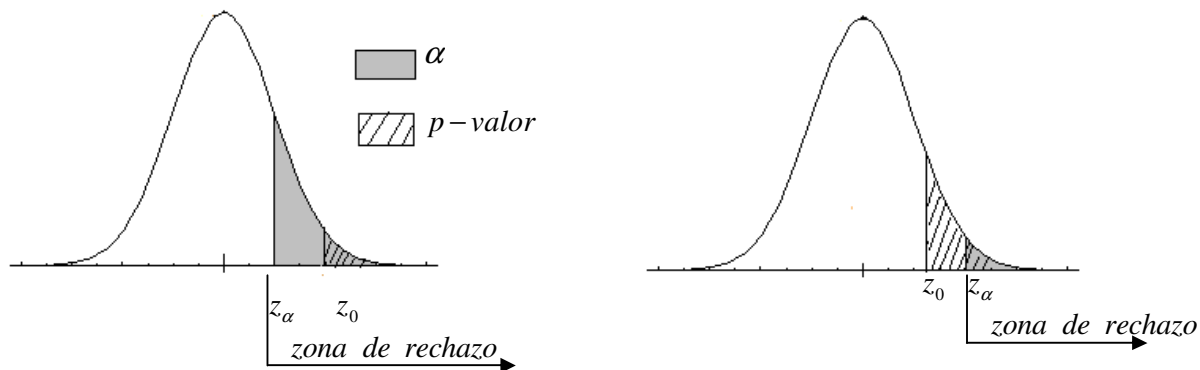
Un p-valor muy chico significa mucha evidencia en contra de H_0 ; un p-valor alto significa que no hay evidencia en contra H_0

Notar que:

Si $\alpha < p\text{-valor}$ entonces se acepta H_0 con nivel de significancia α

Si $\alpha > p\text{-valor}$ entonces se rechaza H_0 con nivel de significancia α

Esto se ilustra en las siguientes figuras:



Ejemplos:

1- En el ejemplo anteúltimo referido al porcentaje deseado de SiO_2 en cierto tipo de cemento aluminoso las hipótesis eran: $H_0 : \mu = 5.5$ contra $H_1 : \mu \neq 5.5$; y el estadístico de prueba tomó el valor

$z_0 = 0.333333 < 2.575 = \frac{z_{0.01}}{2}$; por lo tanto se aceptaba H_0 .

En esta caso $p\text{-valor} = P(|Z| > |z_0|) = 2[1 - \Phi(|z_0|)] = 2[1 - \Phi(0.33333)] = 2[1 - 0.62930] = 0.7414$

Como el p-valor es muy alto no hay evidencia en contra H_0 . Se necesitaría tomar un valor de α mayor a 0.7414 para rechazar H_0 .

2- En el último ejemplo, sobre la duración, en horas, de un foco de 75 watts, las hipótesis eran

$H_0 : \mu = 1000$ contra $H_1 : \mu > 1000$; y el estadístico Z tomó el valor $z_0 = 7.1554 > 1.645 = z_{0.05}$; por lo tanto se rechazaba H_0 .

En este caso

$$p\text{-valor} = P(Z > z_0) = 1 - \Phi(z_0) = 1 - \Phi(7.1554) \approx 0$$

Como el p-valor es casi cero hay mucha evidencia en contra de H_0 . Prácticamente para ningún valor de α se acepta H_0 .

Error de tipo II y selección del tamaño de la muestra

En la prueba de hipótesis el investigador selecciona directamente la probabilidad del error de tipo I. Sin embargo, la probabilidad β de cometer error de tipo II depende del tamaño de la muestra y del valor verdadero del parámetro desconocido.

Supongamos las hipótesis

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu \neq \mu_0$$

Entonces si anotamos con μ al valor verdadero del parámetro

$$\beta = P(\text{aceptar } H_0 / H_0 \text{ es falsa}) = P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \leq z_{\frac{\alpha}{2}} \middle/ \mu \neq \mu_0\right)$$

Como la hipótesis nula es falsa, entonces $\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ **no tiene distribución** $N(0,1)$

Por lo tanto hacemos lo siguiente:

$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu + \mu - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}} ; \quad \text{y ahora como } \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ pues se estandarizó a}$$

\bar{X} con el verdadero μ , entonces

$$\begin{aligned} \beta &= P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \leq z_{\frac{\alpha}{2}} \middle/ \mu \neq \mu_0\right) = P\left(-z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}} \middle/ \mu \neq \mu_0\right) = \\ &= P\left(-z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}} \middle/ \mu \neq \mu_0\right) = P\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) = \end{aligned}$$

$$= \Phi\left(z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) = \Phi\left(z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right)$$

En consecuencia

Si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$, entonces

$$\beta(\mu) = \Phi\left(z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right)$$

Para un valor específico de μ y un valor de α dado, podemos preguntarnos qué tamaño de muestra se necesita para que β sea menor que un valor dado en particular β_0 .

Por ejemplo si $\mu - \mu_0 > 0$ entonces podemos aproximar $\Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) \approx 0$, y planteamos que

$\beta(\mu) = \Phi\left(z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) < \beta_0$. Buscamos en la tabla de la $N(0,1)$ para qué z se cumple que $\Phi(z) = \beta_0$, lo anotamos $-z_{\beta_0}$, y entonces podemos escribir

$$z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n} < -z_{\beta_0} \Rightarrow z_{\frac{\alpha}{2}} + z_{\beta_0} < \frac{(\mu - \mu_0)}{\sigma}\sqrt{n} \Rightarrow n > \frac{\left(z_{\frac{\alpha}{2}} + z_{\beta_0}\right)^2 \sigma^2}{(\mu - \mu_0)^2}$$

En el caso de ser $\mu - \mu_0 < 0$ entonces podemos aproximar $\Phi\left(z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) \approx 1$, y planteamos que

$$\beta(\mu) = 1 - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) < \beta_0. \text{ Es decir } 1 - \beta_0 < \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right)$$

Buscamos en la tabla de la $N(0,1)$ para qué z se cumple que $\Phi(z) = 1 - \beta_0$, lo anotamos z_{β_0} , y entonces podemos escribir

$$-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n} > z_{\beta_0} \Rightarrow z_{\frac{\alpha}{2}} + z_{\beta_0} < -\frac{(\mu - \mu_0)}{\sigma}\sqrt{n} \underset{\mu - \mu_0 < 0}{\Rightarrow} n > \frac{\left(z_{\frac{\alpha}{2}} + z_{\beta_0}\right)^2 \sigma^2}{(\mu - \mu_0)^2}$$

En consecuencia queda la misma fórmula que la anterior

Por lo tanto

Si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$, entonces

$$n > \frac{\left(z_{\frac{\alpha}{2}} + z_{\beta_0}\right)^2 \sigma^2}{(\mu - \mu_0)^2}$$

En forma análoga se puede probar que si las hipótesis son

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu > \mu_0$$

Entonces

$$\begin{aligned} \beta &= P(\text{aceptar } H_0 / H_0 \text{ es falsa}) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_\alpha \mid \mu \neq \mu_0\right) = \\ &= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}} \leq z_\alpha \mid \mu \neq \mu_0\right) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_\alpha - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) = \Phi\left(z_\alpha - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) = \Phi\left(z_\alpha - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n}\right) \end{aligned}$$

Entonces

Si las hipótesis son : $H_0 : \mu = \mu_0$ contra $H_1 : \mu > \mu_0$ entonces

$$\beta(\mu) = \Phi\left(z_\alpha - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n}\right)$$

Y si tenemos las hipótesis $H_0 : \mu = \mu_0$ contra $H_1 : \mu < \mu_0$

$$\begin{aligned} \beta &= P(\text{aceptar } H_0 / H_0 \text{ es falsa}) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq -z_\alpha \mid \mu \neq \mu_0\right) = \\ &= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}} \geq -z_\alpha \mid \mu \neq \mu_0\right) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq -z_\alpha - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) = 1 - \Phi\left(-z_\alpha - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n}\right) \end{aligned}$$

Entonces

Si las hipótesis son : $H_0 : \mu = \mu_0$ contra $H_1 : \mu < \mu_0$ entonces

$$\beta(\mu) = 1 - \Phi\left(-z_\alpha - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n}\right)$$

Y además con una deducción análoga al caso de alternativa bilateral:

Si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu > \mu_0$, (o $H_1 : \mu < \mu_0$) entonces

$$n > \frac{(z_\alpha + z_{\beta_0})^2 \sigma^2}{(\mu - \mu_0)^2}$$

Ejemplos:

1- En el ejemplo referido al porcentaje deseado de SiO_2 en cierto tipo de cemento aluminoso las hipótesis eran: $H_0: \mu = 5.5$ contra $H_1: \mu \neq 5.5$; y el estadístico de prueba tomó el valor $z_0 = 0.333333 < 2.575 = z_{\frac{0.01}{2}}$; por lo tanto se aceptaba H_0 . Teníamos $n = 16$ y $\sigma = 3$

Si el verdadero promedio de porcentaje es $\mu = 5.6$ y se realiza una prueba de nivel $\alpha = 0.01$ con base en $n = 16$, ¿cuál es la probabilidad de detectar esta desviación?

¿Qué valor de n se requiere para satisfacer $\alpha = 0.01$ y $\beta(5.6) = 0.01$?

Solución:

La probabilidad de detectar la desviación es la potencia del test cuando $\mu = 5.6$, es decir

$$\pi(5.6) = P(\text{rechazar } H_0 / H_0 \text{ es falsa}) = 1 - \beta(5.6)$$

Como estamos con hipótesis alternativa bilateral, calculamos

$$\begin{aligned} \beta(5.6) &= \Phi\left(z_{\frac{\alpha}{2}} - \frac{(5.6 - \mu_0)}{\sigma} \sqrt{n}\right) - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(5.6 - \mu_0)}{\sigma} \sqrt{n}\right) = \\ &= \Phi\left(2.575 - \frac{(5.6 - 5.5)}{3} \sqrt{16}\right) - \Phi\left(-2.575 - \frac{(5.6 - 5.5)}{3} \sqrt{16}\right) = \Phi(2.441) - \Phi(-2.708) = \\ &= 0.99266 - (1 - 0.99664) = 0.9893 \quad \Rightarrow \quad \pi(5.6) = 0.0107 \end{aligned}$$

Ahora se quiere hallar n tal que $\beta(5.6) = 0.01$, como el test es bilateral podemos usar directamente la fórmula con $z_{\beta_0} = z_{0.01} = 2.33$

$$n > \frac{\left(z_{\frac{\alpha}{2}} + z_{\beta_0}\right)^2 \sigma^2}{(\mu - \mu_0)^2} = \frac{(2.575 + 2.33)^2 3^2}{(5.6 - 5.5)^2} = 21653.1225 \quad \Rightarrow \quad n \geq 21654$$

2- En el último ejemplo, sobre la duración, en horas, de un foco de 75 watts, las hipótesis eran

$H_0: \mu = 1000$ contra $H_1: \mu > 1000$; y el estadístico Z tomó el valor $z_0 = 7.1554 > 1.645 = z_{0.05}$; por lo tanto se rechazaba H_0 .

En este caso $\sigma = 25$ y $n = 20$

Si la verdadera duración promedio del foco es 1050 horas, ¿cuál es la probabilidad de error de tipo II para la prueba?

¿Qué tamaño de muestra es necesario para asegurar que el error de tipo II no es mayor que 0.10 si la duración promedio verdadera del foco es 1025 hs. ?

Solución:

Como las hipótesis son $H_0: \mu = 1000$ contra $H_1: \mu > 1000$ entonces

$$\beta(\mu) = \Phi\left(z_{\alpha} - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n}\right) = \Phi\left(1.645 - \frac{(1050 - 1000)}{25} \sqrt{20}\right) = \Phi(-7.29927) \neq 0$$

Para hallar n tal que $\beta(1025) \leq 0.1$ aplicamos la fórmula con $z_{\beta_0} = z_{0.1} = 1.285$

$$n > \frac{(z_{\alpha} + z_{\beta_0})^2 \sigma^2}{(\mu - \mu_0)^2} = \frac{(1.645 + 1.285)^2 25^2}{(1025 - 1000)^2} = 8.584 \Rightarrow n \geq 9$$

Relación entre test de hipótesis e intervalos de confianza

Existe una estrecha relación entre la prueba de hipótesis bilateral sobre un parámetro μ y el intervalo de confianza de nivel $1 - \alpha$ para μ .

Específicamente supongamos que tenemos las hipótesis

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu \neq \mu_0$$

La regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } \left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| > z_{\frac{\alpha}{2}} \\ \text{aceptar } H_0 & \text{si } \left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \leq z_{\frac{\alpha}{2}} \end{cases}$$

$$\text{Aceptar } H_0 \text{ si } \left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \leq z_{\frac{\alpha}{2}} \text{ es equivalente a: aceptar } H_0 \text{ si } -z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}}; \text{ y esto es a}$$

su vez equivalente, despejando μ_0 , a:

$$\text{aceptar } H_0 \text{ si } \bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \text{ es decir si } \mu_0 \in \left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$$

Pero resulta que $\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$ es el intervalo de confianza que se construiría para el verdadero parámetro μ de nivel $1 - \alpha$.

Por lo tanto la regla de decisión queda:

$$\begin{cases} \text{rechazar } H_0 & \text{si } \mu_0 \notin \left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] \\ \text{aceptar } H_0 & \text{si } \mu_0 \in \left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] \end{cases}$$

Ejemplo:

En el ejemplo referido al porcentaje deseado de SiO_2 en cierto tipo de cemento aluminoso las hipótesis eran: $H_0 : \mu = 5.5$ contra $H_1 : \mu \neq 5.5$;

y teníamos $n = 16$; $\sigma = 3$; un promedio muestral $\bar{x} = 5.25$

Como $\alpha = 0.01$ entonces $z_{\frac{\alpha}{2}} = z_{0.005} = 2.575$

Construimos un intervalo de confianza de nivel $1 - \alpha = 1 - 0.01 = 0.99$

$$\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] = \left[5.25 - 2.575 \frac{3}{\sqrt{16}}; 5.25 + 2.575 \frac{3}{\sqrt{16}} \right] = [3.31875; 7.18125]$$

Entonces la regla de decisión es:

$$\begin{cases} \text{rechazar } H_0 & \text{si } 5.5 \notin [3.31875; 7.18125] \\ \text{aceptar } H_0 & \text{si } 5.5 \in [3.31875; 7.18125] \end{cases}$$

Como $5.5 \in [3.31875; 7.18125]$, entonces se acepta H_0 .

9.3 – Prueba de hipótesis sobre la media, varianza desconocida para muestras grandes

Hasta ahora se ha desarrollado el procedimiento de test de hipótesis para la hipótesis nula $H_0: \mu = \mu_0$ suponiendo que σ^2 es conocida, pero en la mayoría de las situaciones prácticas σ^2 es desconocida. En general si $n \geq 30$, entonces la varianza muestral S^2 está próxima a σ^2 en la mayor parte de las muestras, de modo que es posible sustituir S^2 por σ^2 . Es decir el estadístico

$$Z = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \approx N(0,1) \quad \text{aproximadamente, si } n \geq 30 \quad \text{si } H_0: \mu = \mu_0$$

Además, si no podemos decir que la muestra aleatoria proviene de una población normal, sea σ^2 conocida o no, por T.C.L. los estadísticos

$$Z = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \approx N(0,1) \quad \text{aproximadamente, si } n \geq 30 \quad \text{si } H_0: \mu = \mu_0$$

Y

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \approx N(0,1) \quad \text{aproximadamente, si } n \geq 30 \quad \text{si } H_0: \mu = \mu_0$$

Las pruebas de hipótesis tendrán entonces un nivel de significancia **aproximadamente de α**

Ejemplo:

Un inspector midió el volumen de llenado de una muestra aleatoria de 100 latas de jugo cuya etiqueta afirmaba que contenían 12 oz. La muestra tenía una media de volumen de 11.98 oz y desviación estándar de 0.19 oz. Sea μ la verdadera media del volumen de llenado para todas las latas de jugo recientemente llenadas con esta máquina. El inspector probará $H_0: \mu = 12$ contra $H_1: \mu \neq 12$

a) Determinar el p-valor

b) ¿Piensa que es factible que la media del volumen de llenado es de 12 oz?

Solución:

La v.a. de interés sería X : “volumen de llenado de una lata tomada al azar”

No se especifica ninguna distribución para X . Anotamos $E(X) = \mu$ y $V(X) = \sigma^2$, ambas desconocidas.

Se toma una muestra de $n = 100$ latas y se obtiene $\bar{x} = 11.98$ y $s = 0.19$

Las hipótesis son $H_0 : \mu = 12$ contra $H_1 : \mu \neq 12$

El estadístico de prueba es

$$Z = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} = \frac{\bar{X} - 12}{S / \sqrt{100}} \text{ y si } H_0 : \mu = 12 \text{ es verdadera entonces } Z \approx N(0,1)$$

$$\text{El estadístico } Z \text{ toma el valor } z_0 = \frac{11.98 - 12}{0.19 / \sqrt{100}} = -1.0526$$

Como la hipótesis alternativa es bilateral entonces

$$p\text{-valor} = P(|Z| > |z_0|) \approx 2[1 - \Phi(1.0526)] = 2[1 - 0.85314] = 0.29372$$

Como el p-valor es mayor que 0.05 se considera que **no hay evidencia en contra de** $H_0 : \mu = 12$

Por lo tanto **es factible que la media del volumen de llenado sea de 12 oz**

9.4 – Prueba de hipótesis sobre la media de una distribución normal, varianza desconocida

Cuando se prueban hipótesis sobre la media μ de una población cuando σ^2 es desconocida es posible utilizar los procedimientos de prueba dados anteriormente siempre y cuando el tamaño de la muestra sea grande ($n \geq 30$). Estos procedimientos son aproximadamente válidos sin importar si la población de interés es normal o no. Pero si la muestra es pequeña y σ^2 es desconocida debe suponerse que la distribución de la variable de interés es normal.

Específicamente, supongamos que la v.a. de interés tiene distribución $N(\mu, \sigma^2)$ donde μ y σ^2 son desconocidas.

Supongamos las hipótesis $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$

Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n de la v.a. X y sean \bar{X} y S^2 la media y la varianza muestrales respectivamente.

El procedimiento se basa en el estadístico

$$T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$$

El cual, **si la hipótesis nula es verdadera**, tiene distribución **Student con $n-1$ grados de libertad**.

Entonces, para un nivel α prefijado, la regla de decisión es

$$\left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } |T| > t_{\frac{\alpha}{2}, n-1} \\ \text{aceptar } H_0 \text{ si } |T| \leq t_{\frac{\alpha}{2}, n-1} \end{array} \right. \text{ es decir } \left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } \left| \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \right| > t_{\frac{\alpha}{2}, n-1} \\ \text{aceptar } H_0 \text{ si } \left| \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \right| \leq t_{\frac{\alpha}{2}, n-1} \end{array} \right.$$

La lógica sigue siendo la misma, si el estadístico de prueba toma un valor inusual, entonces se considera que hay evidencia en contra H_0 y se rechaza la hipótesis nula. Como ahora la distribución del estadístico es Student, nos fijamos si T toma un valor t_0 en las colas de la distribución Student con $n-1$ grados de libertad.

Si la alternativa es $H_1 : \mu > \mu_0$ entonces la regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } T > t_{\alpha, n-1} \\ \text{aceptar } H_0 & \text{si } T \leq t_{\alpha, n-1} \end{cases}$$

Si la alternativa es $H_1 : \mu < \mu_0$ entonces la regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } T < -t_{\alpha, n-1} \\ \text{aceptar } H_0 & \text{si } T \geq -t_{\alpha, n-1} \end{cases}$$

Ejemplo:

Antes de que una sustancia se pueda considerar segura para enterrarse como residuo se deben caracterizar sus propiedades químicas. Se toman 6 muestras de lodo de una planta de tratamiento de agua residual en una región y se les mide el pH obteniéndose una media muestral de 6.68 y una desviación estándar muestral de 0.20. ¿Se puede concluir que la media del pH es menor que 7.0? Utilizar $\alpha = 0.05$ y suponer que la muestra fue tomada de una población normal.

Solución:

La v.a. de interés es X : “pH de una muestra de lodo tomada al azar”

Asumimos que X tiene distribución $N(\mu, \sigma^2)$

Las hipótesis serían $H_0 : \mu = 7.0$ contra $H_1 : \mu < 7.0$

El estadístico de prueba es $T = \frac{\bar{X} - 7.0}{S/\sqrt{6}}$ y toma el valor $t_0 = \frac{6.68 - 7.0}{0.20/\sqrt{6}} = -3.919$

Buscamos en la tabla de la distribución Student $t_{\alpha, n-1} = t_{0.05, 5} = 2.015$

Entonces como $t_0 = -3.919 < -t_{\alpha, n-1} = -t_{0.05, 5} = -2.015$ se rechaza H_0 , por lo tanto hay evidencia que $\mu < 7.0$

P-valor de un test t

En este caso el cálculo del p-valor se realiza considerando:

Si t_0 es el valor calculado del estadístico de prueba T , entonces el p-valor es

a) las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$

$$p\text{-valor} = P(|T| > |t_0|) = 1 - P(|T| \leq |t_0|) = 2(1 - P(T \leq t_0))$$

b) las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu > \mu_0$

$$p\text{-valor} = P(T > t_0) = 1 - P(T \leq t_0)$$

c) las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu < \mu_0$

$$p\text{-valor} = P(T \leq t_0)$$

Para calcular el p-valor en una prueba t nos encontramos con la dificultad que las tablas de la Student no son completas, por lo tanto en algunas ocasiones se deberá **acotar** el p-valor

En el ejemplo anterior para calcular el p-valor de la prueba como es un test con alternativa unilateral

$$p\text{-valor} = P(T \leq t_0) = P(T \leq -3.919)$$

Buscamos en la tabla de la distribución Student la fila donde figuran $\nu = 5$ grados de libertad y vemos que el valor 3.919 no está tabulado.

Pero $3.365 < 3.919 < 4.032$, y $P(T_5 > 3.365) = 0.01$ y $P(T_5 > 4.032) = 0.005$

Por lo tanto $0.005 < P(T_5 > 3.919) < 0.01$, es decir

$0.005 < p\text{-valor} = P(T_5 < -3.919) < 0.01$

Podemos deducir que existe evidencia de que la media del pH es menor que 0.7

9.5 – Prueba de hipótesis sobre la diferencia de dos medias, varianzas conocidas

Supongamos que tenemos dos variables aleatorias *independientes* normalmente distribuidas:

$$\begin{cases} X_1 \sim N(\mu_1, \sigma_1^2) \\ X_2 \sim N(\mu_2, \sigma_2^2) \end{cases} \text{ y suponemos que las varianzas } \sigma_1^2 \text{ y } \sigma_2^2 \text{ son conocidas.}$$

Sean además

$(X_{11}, X_{12}, \dots, X_{1n_1})$ una muestra aleatoria de tamaño n_1 de X_1

$(X_{21}, X_{22}, \dots, X_{2n_2})$ una muestra aleatoria de tamaño n_2 de X_2 .

El interés recae en probar que $\mu_1 - \mu_2 = \Delta_0$ donde Δ_0 es un valor fijado, por ejemplo si $\Delta_0 = 0$ entonces se querrá probar que $\mu_1 - \mu_2 = 0$ es decir que las medias son iguales.

Ya sabemos que bajo las suposiciones anteriores

$$\begin{cases} \bar{X}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right) \\ \bar{X}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right) \end{cases}$$

Y además

$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).$$

Por lo tanto

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1), \text{ es decir, tiene distribución normal estandarizada.}$$

Si consideramos las hipótesis

$$H_0 : \mu_1 - \mu_2 = \Delta_0 \quad \text{contra} \quad H_1 : \mu_1 - \mu_2 \neq \Delta_0$$

Entonces usamos como estadístico de prueba a $Z = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

$$Y \quad Z = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1) \quad \text{si} \quad H_0 : \mu_1 - \mu_2 = \Delta_0 \quad \text{es verdadera}$$

Por lo tanto la regla de decisión será

$$\begin{cases} \text{rechazar } H_0 & \text{si } |Z| > z_{\frac{\alpha}{2}} \\ \text{aceptar } H_0 & \text{si } |Z| \leq z_{\frac{\alpha}{2}} \end{cases} \quad \text{donde} \quad Z = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$\text{Si } H_1 : \mu_1 - \mu_2 > \Delta_0 \text{ entonces la regla de decisión es } \begin{cases} \text{rechazar } H_0 & \text{si } Z > z_{\alpha} \\ \text{aceptar } H_0 & \text{si } Z \leq z_{\alpha} \end{cases}$$

$$\text{Si } H_1 : \mu_1 - \mu_2 < \Delta_0 \text{ entonces la regla de decisión es } \begin{cases} \text{rechazar } H_0 & \text{si } Z < -z_{\alpha} \\ \text{aceptar } H_0 & \text{si } Z \geq -z_{\alpha} \end{cases}$$

Ejemplos:

1- Un diseñador de productos está interesado en reducir el tiempo de secado de una pintura tapaporos. Se prueban dos fórmulas de pintura. La fórmula 1 tiene el contenido químico estándar, y la fórmula 2 tiene un nuevo ingrediente secante que debe reducir el tiempo de secado. De la experiencia se sabe que la desviación estándar del tiempo de secado es 8 minutos, y esta variabilidad no debe verse afectada por la adición del nuevo ingrediente. Se pintan 10 especímenes con la fórmula 1 y otros 10 con la fórmula 2. los tiempos promedio de secado muestrales fueron $\bar{x}_1 = 121$ minutos y $\bar{x}_2 = 112$ minutos respectivamente.

¿A qué conclusiones debe llegar el diseñador del producto sobre la eficacia del nuevo ingrediente utilizando $\alpha = 0.05$?

Solución:

Aquí las hipótesis son $H_0 : \mu_1 - \mu_2 = 0$ contra $H_1 : \mu_1 - \mu_2 > 0$

$$\text{El estadístico de prueba es } Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{8^2}{10} + \frac{8^2}{10}}} \quad \text{y toma el valor} \quad z_0 = \frac{121 - 112}{\sqrt{\frac{8^2}{10} + \frac{8^2}{10}}} = 2.52$$

Buscamos en la tabla de la normal estándar $z_{\alpha} = z_{0.05} = 1.645$

Como $z_0 = 2.52 > z_{\alpha} = z_{0.05} = 1.645$ se rechaza H_0 al nivel 0.05 y se concluye que el nuevo ingrediente disminuye el tiempo de secado.

El cálculo del p-valor y la deducción de β la probabilidad de cometer error de tipo II se obtienen de manera análoga a los casos anteriores. Por ejemplo para la alternativa bilateral la expresión para β es la siguiente donde anotamos $\mu_1 - \mu_2 - \Delta = \Delta_0 - \Delta = \delta$

$$\beta = P(\text{aceptar } H_0 / H_0 \text{ es falsa}) = \Phi \left(\frac{z_{\alpha/2} - \frac{\delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}}{\right) - \Phi \left(-\frac{z_{\alpha/2} - \frac{\delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}}{\right)$$

En el ejemplo anterior el $p\text{-valor} = P(Z > z_0) = P(Z > 2.52) = 1 - \Phi(2.52) = 0.0059$

También es posible obtener fórmulas para el tamaño de la muestra necesario para obtener una β específica para una diferencia dada en las medias $\mu_1 - \mu_2 - \Delta = \Delta_0 - \Delta = \delta$ y α . Si asumimos que $n_1 = n_2 = n$ entonces

<p>Para $H_1 : \mu_1 - \mu_2 \neq \Delta_0$ es $n > \frac{\left(z_{\alpha/2} + z_{\beta_0}\right)^2 (\sigma_1^2 + \sigma_2^2)}{\delta^2}$</p> <p>Para $H_1 : \mu_1 - \mu_2 > \Delta_0$ o $H_1 : \mu_1 - \mu_2 < \Delta_0$ es $n > \frac{\left(z_{\alpha} + z_{\beta_0}\right)^2 (\sigma_1^2 + \sigma_2^2)}{\delta^2}$</p>
--

9.6 – Prueba de hipótesis sobre la diferencia de dos medias, varianzas desconocidas

Caso 1: $\sigma_1^2 \neq \sigma_2^2$

Supongamos que tenemos dos variables aleatorias **independientes** normalmente distribuidas:

$$\begin{cases} X_1 \sim N(\mu_1, \sigma_1^2) \\ X_2 \sim N(\mu_2, \sigma_2^2) \end{cases} \text{ y las varianzas } \sigma_1^2 \text{ y } \sigma_2^2 \text{ son } \textbf{desconocidas} .$$

y además

$(X_{11}, X_{12}, \dots, X_{1n_1})$ es una muestra aleatoria de tamaño n_1 de X_1
 $(X_{21}, X_{22}, \dots, X_{2n_2})$ es una muestra aleatoria de tamaño n_2 de X_2 .

Si las muestras aleatorias se toma de una distribución normal, donde σ_1 y σ_2 son **desconocidos**, $n_1 \geq 30$ y $n_2 \geq 30$, entonces se puede probar que al reemplazar σ_1 por S_1 y σ_2 por S_2 , el estadístico

$$\frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \approx N(0,1) \quad \text{aproximadamente}$$

Por lo tanto si anotamos $Z = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$ valen las reglas de decisión vistas en la sección anterior,

con la diferencia que el nivel de significancia del test será **aproximadamente** $1 - \alpha$

Si ahora n_1 o n_2 **no son mayores que 30**, entonces

$$T^* = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

tiene distribución aproximadamente Student con ν grados de libertad bajo la hipótesis $H_0 : \mu_1 - \mu_2 = \Delta_0$ donde

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}} \quad \text{si } \nu \text{ no es entero, se toma el entero más próximo a } \nu$$

Por lo tanto, si las hipótesis son

$H_0 : \mu_1 - \mu_2 = \Delta_0$ contra $H_1 : \mu_1 - \mu_2 \neq \Delta_0$ entonces la regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } |T^*| > t_{\frac{\alpha}{2}, \nu} \\ \text{aceptar } H_0 & \text{si } |T^*| \leq t_{\frac{\alpha}{2}, \nu} \end{cases}$$

Si $H_1 : \mu_1 - \mu_2 > \Delta_0$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } T^* > t_{\alpha, \nu} \\ \text{aceptar } H_0 & \text{si } T^* \leq t_{\alpha, \nu} \end{cases}$

Si $H_1 : \mu_1 - \mu_2 < \Delta_0$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } T^* < -t_{\alpha, \nu} \\ \text{aceptar } H_0 & \text{si } T^* \geq -t_{\alpha, \nu} \end{cases}$

Ejemplo:

Un fabricante de monitores prueba dos diseños de microcircuitos para determinar si producen un flujo de corriente equivalente. El departamento de ingeniería ha obtenido los datos siguientes:

<i>Diseño 1</i>	$n_1 = 15$	$\bar{x}_1 = 24.2$	$s_1^2 = 10$
<i>Diseño 2</i>	$n_2 = 10$	$\bar{x}_2 = 23.9$	$s_2^2 = 20$

Con $\alpha = 0.10$ se desea determinar si existe alguna diferencia significativa en el flujo de corriente medio entre los dos diseños, donde se supone que las poblaciones son normales.

Solución:

Las variables aleatorias de interés son

X_1 : “flujo de corriente en diseño 1”

X_2 : “flujo de corriente en diseño 2”

Asumimos que $X_1 \sim N(\mu_1, \sigma_1^2)$ y $X_2 \sim N(\mu_2, \sigma_2^2)$ donde los parámetros son desconocidos

Las hipótesis serían $H_0 : \mu_1 - \mu_2 = 0$ contra $H_1 : \mu_1 - \mu_2 \neq 0$

El estadístico de prueba es

$$|T^*| = \left| \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{15} + \frac{S_2^2}{10}}} \right| \text{ que en este caso toma el valor } t_0^* = \frac{24.2 - 23.9}{\sqrt{\frac{10}{15} + \frac{20}{10}}} = 0.18$$

Debemos buscar en la tabla de la distribución Student $t_{\frac{\alpha}{2}, \nu} = t_{\frac{0.10}{2}, \nu}$ entonces calculamos

$$\nu = \frac{\frac{(S_1^2/n_1 + S_2^2/n_2)^2}{(S_1^2/n_1)^2 + (S_2^2/n_2)^2}}{\frac{n_1 - 1}{15 - 1} + \frac{n_2 - 1}{10 - 1}} = \frac{(10/15 + 20/10)^2}{(10/15)^2 + (20/10)^2} = 14.9333 \Rightarrow \nu = 15$$

Por lo tanto $t_{\frac{\alpha}{2}, \nu} = t_{0.05, 15} = 1.753$

Como $t_0^* = 0.18 < t_{0.05, 15} = 1.753$ entonces se acepta $H_0 : \mu_1 - \mu_2 = 0$

No hay evidencia fuerte que las medias de los dos flujos de corriente sean diferentes.

Si calculamos el p-valor

$$p\text{-valor} = P(|T^*| > t_0^*) = P(|T^*| > 0.18) > 0.40$$

Caso 2: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Supongamos que tenemos dos variables aleatorias **independientes** normalmente distribuidas:

$$\begin{cases} X_1 \sim N(\mu_1, \sigma_1^2) \\ X_2 \sim N(\mu_2, \sigma_2^2) \end{cases} \text{ y las varianzas } \sigma_1^2 \text{ y } \sigma_2^2 \text{ son } \textbf{desconocidas} \text{ pero } \textbf{iguales}.$$

y además

$(X_{11}, X_{12}, \dots, X_{1n_1})$ es una muestra aleatoria de tamaño n_1 de X_1

$(X_{21}, X_{22}, \dots, X_{2n_2})$ es una muestra aleatoria de tamaño n_2 de X_2 .

Sean \bar{X}_1 y \bar{X}_2 las medias muestrales y S_1^2 y S_2^2 las varianzas muestrales. Como S_1^2 y S_2^2 son los estimadores de la varianza común σ^2 , entonces construimos un **estimador combinado** de σ^2 . Este estimador es

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Se puede comprobar que es un estimador insesgado de σ^2 .

Ya vimos que se puede probar que el estadístico

$$T = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{tiene distribución Student con } n_1 + n_2 - 2 \text{ grados de libertad}$$

Por lo tanto, si las hipótesis son

$H_0 : \mu_1 - \mu_2 = \Delta_0$ contra $H_1 : \mu_1 - \mu_2 \neq \Delta_0$ entonces la regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } |T| > t_{\frac{\alpha}{2}, n_1 + n_2 - 2} \\ \text{aceptar } H_0 & \text{si } |T| \leq t_{\frac{\alpha}{2}, n_1 + n_2 - 2} \end{cases}$$

Si $H_1 : \mu_1 - \mu_2 > \Delta_0$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } T > t_{\alpha, n_1 + n_2 - 2} \\ \text{aceptar } H_0 & \text{si } T \leq t_{\alpha, n_1 + n_2 - 2} \end{cases}$

Si $H_1 : \mu_1 - \mu_2 < \Delta_0$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } T < -t_{\alpha, n_1 + n_2 - 2} \\ \text{aceptar } H_0 & \text{si } T \geq -t_{\alpha, n_1 + n_2 - 2} \end{cases}$

Ejemplo:

Se tienen las mediciones del nivel de hierro en la sangre de dos muestras de niños: un grupo de niños sanos y el otro padece fibrosis quística. Los datos obtenidos se dan en la siguiente tabla:

sanos	$n_1 = 9$	$\bar{x}_1 = 18.9$	$s_1^2 = 5.9^2$
enfermos	$n_2 = 13$	$\bar{x}_2 = 11.9$	$s_2^2 = 6.3^2$

Podemos asumir que las muestras provienen de poblaciones normales independientes con iguales varianzas.

Es de interés saber si las dos medias del nivel de hierro en sangre son iguales o distintas. Utilizar $\alpha = 0.05$

Solución:

Las variables de interés son

X_1 : “nivel de hierro en sangre de un niño sano tomado al azar”

X_2 : “nivel de hierro en sangre de un niño con fibrosis quística tomado al azar”

Asumimos que $X_1 \sim N(\mu_1, \sigma^2)$ y $X_2 \sim N(\mu_2, \sigma^2)$

Consideramos las hipótesis

$$H_0 : \mu_1 - \mu_2 = 0 \quad \text{contra} \quad H_1 : \mu_1 - \mu_2 \neq 0$$

Para calcular el valor del estadístico de prueba, primero calculamos

$$S_p = \sqrt{S_p^2} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{(9 - 1)5.9^2 + (13 - 1)6.3^2}{9 + 13 - 2}} = 6.14$$

$$\text{El estadístico de prueba es } |T| = \left| \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{9} + \frac{1}{13}}} \right| \text{ y toma el valor } t_0 = \left| \frac{18.9 - 11.9}{6.14 \sqrt{\frac{1}{9} + \frac{1}{13}}} \right| = 2.63$$

Buscamos en la tabla de la distribución Student $t_{\frac{\alpha}{2}, n_1 + n_2 - 2} = t_{0.025, 20} = 2.086$

Como $t_0 = 2.63 > t_{0.025, 20} = 2.086$ entonces se rechaza $H_0 : \mu_1 - \mu_2 = 0$

Si calculamos el p-valor de la prueba

$$p - \text{valor} = 2(1 - P(T < t_0)) = 2(1 - P(T < 2.63)) = 2P(T > 2.63)$$

Vemos de la tabla de la Student que $t_{0.01, 20} = 2.528$ y $t_{0.005, 20} = 2.845$ por lo tanto

$$2 \times 0.005 < p - \text{valor} = 2P(T > 2.63) < 2 \times 0.01 \quad \text{es decir} \quad 0.01 < p - \text{valor} < 0.02$$

9.7 – Prueba de hipótesis sobre la diferencia de dos medias para datos de a pares

Ya se vio el caso, cuando se habló de intervalos de confianza para una diferencia de medias, de datos dados de a pares, es decir $(X_{11}, X_{21}); (X_{12}, X_{22}); \dots; (X_{1n}, X_{2n})$.

Las variables aleatorias X_1 y X_2 tienen medias μ_1 y μ_2 respectivamente.

Consideramos $D_j = X_{1j} - X_{2j}$ con $j = 1, 2, \dots, n$.

Entonces

$$E(D_j) = E(X_{1j} - X_{2j}) = E(X_{1j}) - E(X_{2j}) = \mu_1 - \mu_2$$

y

$$V(D_j) = V(X_{1j} - X_{2j}) = V(X_{1j}) + V(X_{2j}) - 2Cov(X_{1j}, X_{2j}) = \sigma_1^2 + \sigma_2^2 - 2Cov(X_1, X_2)$$

$$\text{Estimamos } E(D_j) = \mu_1 - \mu_2 \text{ con } \bar{D} = \frac{1}{n} \sum_{j=1}^n D_j = \frac{1}{n} \sum_{j=1}^n (X_{1j} - X_{2j}) = \bar{X}_1 - \bar{X}_2$$

$$\text{En lugar de tratar de estimar la covarianza, estimamos la } V(D_j) \text{ con } S_D = \frac{1}{n-1} \sum_{j=1}^n (D_j - \bar{D})^2$$

Anotamos $\mu_D = \mu_1 - \mu_2$ y $\sigma_D^2 = V(D_j)$

Asumimos que $D_j \sim N(\mu_D, \sigma_D^2)$ con $j = 1, 2, \dots, n$

Las variables aleatorias en pares diferentes son independientes, no lo son dentro de un mismo par.

Para construir una regla de decisión nuevamente, consideramos el estadístico

$$T = \frac{\bar{D} - \mu_D}{S_D / \sqrt{n}} \quad \text{con distribución } t_{n-1}$$

Si tenemos las hipótesis

$$H_0 : \mu_1 - \mu_2 = \Delta_0 \quad \text{contra} \quad H_1 : \mu_1 - \mu_2 \neq \Delta_0$$

Entonces el estadístico de prueba es

$$T = \frac{\bar{D} - \Delta_0}{S_D / \sqrt{n}} \quad \text{y tiene distribución } t_{n-1} \quad \text{si } H_0 : \mu_1 - \mu_2 = \Delta_0 \quad \text{es verdadera}$$

$$\text{Por lo tanto, la regla de decisión es } \begin{cases} \text{rechazar } H_0 & \text{si } |T| > t_{\frac{\alpha}{2}, n-1} \\ \text{aceptar } H_0 & \text{si } |T| \leq t_{\frac{\alpha}{2}, n-1} \end{cases} \quad \text{donde } T = \frac{\bar{D} - \Delta_0}{S_D / \sqrt{n}}$$

$$\text{Si } H_1 : \mu_1 - \mu_2 > \Delta_0 \quad \text{entonces la regla de decisión es } \begin{cases} \text{rechazar } H_0 & \text{si } T > t_{\alpha, n-1} \\ \text{aceptar } H_0 & \text{si } T \leq t_{\alpha, n-1} \end{cases}$$

$$\text{Si } H_1 : \mu_1 - \mu_2 < \Delta_0 \quad \text{entonces la regla de decisión es } \begin{cases} \text{rechazar } H_0 & \text{si } T < -t_{\alpha, n-1} \\ \text{aceptar } H_0 & \text{si } T \geq -t_{\alpha, n-1} \end{cases}$$

Ejemplo:

Se comparan dos microprocesadores en una muestra de 6 códigos de puntos de referencia para determinar si hay una diferencia en la rapidez. Los tiempos en segundos utilizados para cada procesador en cada código están dados en la siguiente tabla:

	Código					
	1	2	3	4	5	6
Procesador A	27.2	18.1	27.2	19.7	24.5	22.1
Procesador B	24.1	19.3	26.8	20.1	27.6	29.8

¿Puede concluir que las medias de la rapidez de ambos procesadores son diferentes con nivel de significancia 0.05?

Solución:

Las variables aleatorias de interés son

X_1 : “rapidez del procesador A en un código tomado al azar”

X_2 : “rapidez del procesador B en un código tomado al azar”

Como ambas variables se miden sobre un mismo código no podemos asumir que son independientes.

Las hipótesis son $H_0 : \mu_1 - \mu_2 = 0$ contra $H_1 : \mu_1 - \mu_2 \neq 0$

Necesitamos la muestra de las diferencias D_j :

3.1, -1.2; 0.4; -0.4; -3.1; -7.7

De esta muestra obtenemos $\bar{d} = -1.483333$ y $s_D = 3.66246$

Además $\alpha = 0.05 \rightarrow t_{\frac{\alpha}{2}, n-1} = t_{0.025, 5} = 2.571$

El estadístico de prueba es $|T| = \left| \frac{\bar{D}}{S_D / \sqrt{6}} \right|$ y toma el valor $t_0 = \left| \frac{-1.483333}{3.66246 / \sqrt{6}} \right| = 0.99206$

Como $t_0 = 0.99206 < t_{\frac{\alpha}{2}, n-1} = t_{0.025, 5} = 2.571$ entonces se acepta la hipótesis nula. No hay evidencia de que las medias de la rapidez de ambos procesadores sean diferentes.

9.8 – Tests de hipótesis sobre la varianza

Supongamos que se desea probar la hipótesis de que la varianza de una población normal es igual a un valor específico, por ejemplo σ_0^2 .

Sea (X_1, X_2, \dots, X_n) una muestra aleatoria de tamaño n de una v.a. X , donde $X \sim N(\mu, \sigma^2)$.

Tomamos como estimador puntual de σ^2 a $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

Luego a partir de este estimador puntual construimos el estadístico $X = \frac{(n-1)S^2}{\sigma^2}$

Este estadístico contiene al parámetro desconocido a estimar σ^2 y ya sabemos que tiene una distribución llamada **ji-cuadrado con n-1 grados de libertad**

Supongamos las hipótesis

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{contra} \quad H_1 : \sigma^2 \neq \sigma_0^2$$

Tomamos como estadístico de prueba a

$$X = \frac{(n-1)S^2}{\sigma_0^2} \quad \text{y} \quad \text{si} \quad H_0 : \sigma^2 = \sigma_0^2 \quad \text{es verdadera, entonces} \quad X = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi_{n-1}^2$$

Nuevamente, el razonamiento es: si el estadístico X que bajo $H_0 : \sigma^2 = \sigma_0^2$ tiene distribución χ_{n-1}^2 toma un valor “inusual”, se considera que hay evidencia en contra H_0

Recordar que la distribución χ_{n-1}^2 es asimétrica. Entonces la regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } X > \chi_{\frac{\alpha}{2}, n-1}^2 \quad \text{ó} \quad X < \chi_{1-\frac{\alpha}{2}, n-1}^2 \\ \text{aceptar } H_0 & \text{si } \chi_{1-\frac{\alpha}{2}, n-1}^2 \leq X \leq \chi_{\frac{\alpha}{2}, n-1}^2 \end{cases} \quad \text{donde} \quad X = \frac{(n-1)S^2}{\sigma_0^2}$$

Si $H_1: \sigma^2 > \sigma_0^2$ entonces la regla de decisión es
$$\begin{cases} \text{rechazar } H_0 & \text{si } X > \chi_{\alpha, n-1}^2 \\ \text{aceptar } H_0 & \text{si } X \leq \chi_{\alpha, n-1}^2 \end{cases}$$

Si $H_1: \sigma^2 < \sigma_0^2$ entonces la regla de decisión es
$$\begin{cases} \text{rechazar } H_0 & \text{si } X < \chi_{1-\alpha, n-1}^2 \\ \text{aceptar } H_0 & \text{si } X \geq \chi_{1-\alpha, n-1}^2 \end{cases}$$

Para calcular el p-valor, si el estadístico X tomó el valor x_0 , y teniendo en cuenta que no hay simetría en la distribución ji-cuadrado, hacemos:

Si $H_1: \sigma^2 > \sigma_0^2$ entonces $p\text{-valor} = P(X > x_0)$

Si $H_1: \sigma^2 < \sigma_0^2$ entonces $p\text{-valor} = P(X < x_0)$

Si $H_1: \sigma^2 \neq \sigma_0^2$ entonces $p\text{-valor} = 2 \min \left(P(X < x_0), P(X > x_0) \right)$

Ejemplo:

Consideremos nuevamente el ejemplo visto en la sección de intervalos de confianza para la varianza sobre la máquina de llenado de botellas. Al tomar una muestra aleatoria de 20 botellas se obtiene una varianza muestral para el volumen de llenado de $s^2 = 0.0153 \text{ oz}^2$.

Si la varianza del volumen de llenado es mayor que 0.01 oz^2 , entonces existe una proporción inaceptable de botellas que serán llenadas con una cantidad menor de líquido. ¿Existe evidencia en los datos muestrales que sugiera que el fabricante tiene un problema con el llenado de las botellas? Utilice $\alpha = 0.05$

Solución:

La variable de interés es X : “volumen de llenado de una botella tomada al azar”

Asumimos $X \sim N(\mu, \sigma^2)$

Los datos son $s^2 = 0.0153$ de una muestra de tamaño $n = 20$

Las hipótesis son $H_0: \sigma^2 = 0.01$ contra $H_1: \sigma^2 > 0.01$

$$\alpha = 0.05 \rightarrow \chi_{\alpha, n-1}^2 = \chi_{0.05, 19}^2 = 30.14$$

El estadístico de prueba es $X = \frac{(n-1)S^2}{\sigma_0^2} = \frac{19 \times S^2}{0.01}$ y toma el valor

$$x_0 = \frac{19 \times S^2}{0.01} = \frac{19 \times 0.0153}{0.01} = 29.07$$

Como $x_0 = 29.07 < \chi_{0.05, 19}^2 = 30.14$ entonces no hay evidencia fuerte de que la varianza del volumen de llenado sea mayor que 0.01

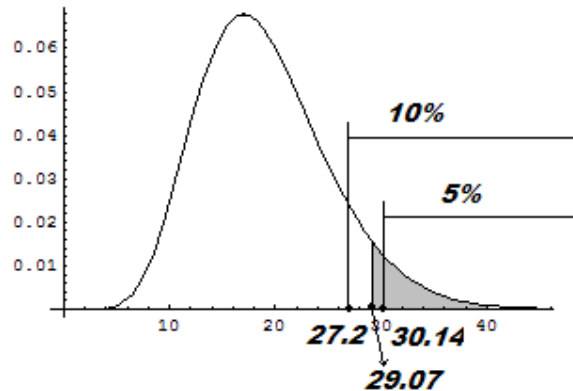
Para calcular el p-valor

$$p\text{-valor} = P(X > x_0) = P(X > 29.07)$$

Buscamos en la tabla de la distribución ji-cuadrado y vemos que en la fila con $\nu = 19$ no figura 29.07, pero $27.20 < 29.07 < 30.14$, y además

$$\begin{cases} P(X > 27.20) = 0.10 \\ P(X > 30.14) = 0.05 \end{cases} \Rightarrow 0.05 < p\text{-valor} < 0.10$$

En la figura siguiente se ilustra la situación



9.9 – Tests de hipótesis sobre la igualdad de dos varianzas

Supongamos que tenemos interés en dos poblaciones normales independientes, donde las medias y las varianzas de la población son desconocidas. Se desea probar la hipótesis sobre la igualdad de las dos varianzas, específicamente:

Supongamos que tenemos dos variables aleatorias *independientes* normalmente distribuidas:

$$\begin{cases} X_1 \sim N(\mu_1, \sigma_1^2) \\ X_2 \sim N(\mu_2, \sigma_2^2) \end{cases} \text{ y } \mu_1; \mu_2; \sigma_1^2 \text{ y } \sigma_2^2 \text{ son desconocidos}$$

y además

$(X_{11}, X_{12}, \dots, X_{1n_1})$ es una muestra aleatoria de tamaño n_1 de X_1

$(X_{21}, X_{22}, \dots, X_{2n_2})$ es una muestra aleatoria de tamaño n_2 de X_2 .

Sean S_1^2 y S_2^2 las varianzas muestrales, S_1^2 y S_2^2 son los estimadores de σ_1^2 y σ_2^2 respectivamente.

Consideramos el estadístico

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}$$

Notar que F contiene al parámetro de interés $\frac{\sigma_1^2}{\sigma_2^2}$, pues $F = \frac{S_1^2 \times \sigma_2^2}{S_2^2 \times \sigma_1^2}$

Sabemos que F tiene una distribución llamada Fisher con $n_1 - 1$ y $n_2 - 1$ grados de libertad.

Sean las hipótesis $H_0 : \sigma_1^2 = \sigma_2^2$ contra $H_1 : \sigma_1^2 \neq \sigma_2^2$

Tomamos como estadístico de prueba a $F = \frac{S_1^2}{S_2^2}$

Vemos que $F = \frac{S_1^2}{S_2^2} \sim F_{n_1-1, n_2-1}$ si $H_0 : \sigma_1^2 = \sigma_2^2$ es verdadera

Recordando que la distribución Fisher es asimétrica, la regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } F > f_{\frac{\alpha}{2}, n_1-1, n_2-1}^2 \text{ ó } F < f_{1-\frac{\alpha}{2}, n_1-1, n_2-1}^2 \\ \text{aceptar } H_0 & \text{si } f_{1-\frac{\alpha}{2}, n_1-1, n_2-1}^2 \leq F \leq f_{\frac{\alpha}{2}, n_1-1, n_2-1}^2 \end{cases}$$

Si $H_1 : \sigma_1^2 > \sigma_2^2$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } F > f_{\alpha, n_1-1, n_2-1}^2 \\ \text{aceptar } H_0 & \text{si } F \leq f_{\alpha, n_1-1, n_2-1}^2 \end{cases}$

Si $H_1 : \sigma_1^2 < \sigma_2^2$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } F < f_{1-\alpha, n_1-1, n_2-1}^2 \\ \text{aceptar } H_0 & \text{si } F \geq f_{1-\alpha, n_1-1, n_2-1}^2 \end{cases}$

Para calcular el p-valor, si el estadístico F tomó el valor f_0 , y teniendo en cuenta que no hay simetría en la distribución Fisher, hacemos:

Si $H_1 : \sigma_1^2 > \sigma_2^2$ entonces $p\text{-valor} = P(F > f_0)$

Si $H_1 : \sigma_1^2 < \sigma_2^2$ entonces $p\text{-valor} = P(F < f_0)$

Si $H_1 : \sigma_1^2 \neq \sigma_2^2$ entonces $p\text{-valor} = 2 \min \left(P(F < f_0), P(F > f_0) \right)$

Ejemplo:

En una serie de experimentos para determinar la tasa de absorción de ciertos pesticidas en la piel se aplicaron cantidades medidas de dos pesticidas a algunos especímenes de piel. Después de un tiempo se midieron las cantidades absorbidas (en μg). Para el pesticida A la varianza de las cantidades absorbidas en 6 muestras fue de 2.3; mientras que para el B la varianza de las cantidades absorbidas en 10 especímenes fue de 0.6. Suponga que para cada pesticida las cantidades absorbidas constituyen una muestra aleatoria de una población normal. ¿Se puede concluir que la varianza en la cantidad absorbida es mayor para el pesticida A que para el B? Utilizar $\alpha = 0.05$

Solución:

Las variables aleatorias de interés son

X_1 : “cantidad absorbida de pesticida A en un espécimen de piel tomado al azar”

X_2 : “cantidad absorbida de pesticida B en un espécimen de piel tomado al azar”

Asumimos que $X_1 \sim N(\mu_1, \sigma_1^2)$ y $X_2 \sim N(\mu_2, \sigma_2^2)$

Las hipótesis son $H_0 : \sigma_1^2 = \sigma_2^2$ contra $H_1 : \sigma_1^2 < \sigma_2^2$

Los datos son $s_1^2 = 2.3$ y $s_2^2 = 0.6$

$n_1 = 6$; $n_2 = 10$

El estadístico de prueba es $F = \frac{S_1^2}{S_2^2}$ y toma el valor $f_0 = \frac{2.3}{0.6} = 3.83$

Buscamos en la tabla de la distribución Fisher $f_{0.05,5,9} = 3.48$

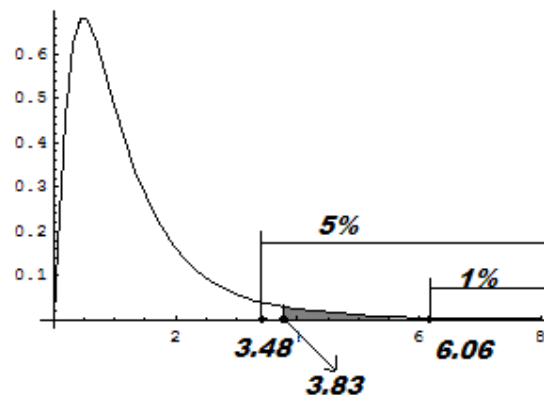
Como $f_0 = \frac{2.3}{0.6} = 3.83 > 3.48 = f_{0.05,5,9}$ se rechaza $H_0 : \sigma_1^2 = \sigma_2^2$

Para saber cuánta evidencia hay contra la hipótesis nula, calculamos el p-valor

De la tabla de la Fisher vemos que $f_{0.05,5,9} = 3.48 < 3.83 < f_{0.01,5,9} = 6.06$

Por lo tanto $0.01 < p\text{-valor} < 0.05$

En la figura siguiente se ilustra la situación



9.10 – Tests de hipótesis sobre una proporción

En muchos problemas se tiene interés en una variable aleatoria que sigue una distribución binomial. Por ejemplo, un proceso de producción que fabrica artículos que son clasificados como aceptables o defectuosos. Lo más usual es modelar la ocurrencia de artículos defectuosos con la distribución binomial, donde el parámetro binomial p representa la proporción de artículos defectuosos producidos. En consecuencia, muchos problemas de decisión incluyen una prueba de hipótesis con respecto a p . Consideremos las hipótesis

$$H_0 : p = p_0 \quad \text{contra} \quad H_1 : p \neq p_0$$

Supongamos que consideramos una muestra aleatoria (X_1, X_2, \dots, X_n) de tamaño n , donde X_i tiene una distribución binomial con parámetros 1 y p : $X_i \sim B(1, p)$.

Ya sabemos que $X = X_1 + X_2 + \dots + X_n$, es una v.a. cuya distribución es binomial con parámetros n

y p : $X \sim B(n, p)$. De acuerdo con esto, la variable aleatoria \hat{P} definida: $\hat{P} = \frac{X}{n}$ representa la proporción

de individuos de la muestra que verifican la propiedad de interés.

Además

$$E(\hat{P}) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = \frac{1}{n} np = p, \text{ y } V(\hat{P}) = V\left(\frac{X}{n}\right) = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}$$

Consideramos el estadístico de prueba

$$Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Si $H_0 : p = p_0$ es verdadera entonces $Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \approx N(0,1)$ aproximadamente por T.C.L.

Por lo tanto la regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } |Z| > z_{\frac{\alpha}{2}} \\ \text{aceptar } H_0 & \text{si } |Z| \leq z_{\frac{\alpha}{2}} \end{cases} \quad \text{donde } Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Si $H_1 : p > p_0$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } Z > z_{\alpha} \\ \text{aceptar } H_0 & \text{si } Z \leq z_{\alpha} \end{cases}$

Si $H_1 : p < p_0$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } Z < -z_{\alpha} \\ \text{aceptar } H_0 & \text{si } Z \geq -z_{\alpha} \end{cases}$

Observaciones:

1- La prueba descrita anteriormente requiere que la proporción muestral esté normalmente distribuida. Esta suposición estará justificada siempre que $np_0 > 10$ y $n(1-p_0) > 10$, donde p_0 es la proporción poblacional que se especificó en la hipótesis nula.

2- También se podía haber tomado como estadístico de prueba a $Z = \frac{X - np_0}{\sqrt{np_0(1-p_0)}}$ donde $X \sim B(n, p)$

Ejemplo:

Un fabricante de semiconductores produce controladores que se emplean en aplicaciones de motores automovilísticos. El cliente requiere que la fracción de controladores defectuosos en uno de los pasos de manufactura críticos no sea mayor que 0.05, y que el fabricante demuestre esta característica del proceso de fabricación con este nivel de calidad, utilizando $\alpha = 0.05$. El fabricante de semiconductores

toma una muestra aleatoria de 200 dispositivos y encuentra que 4 de ellos son defectuosos. ¿El fabricante puede demostrar al cliente la calidad del proceso?

Solución:

Sea la v.a. X : “número de controladores defectuosos en la muestra”

Entonces $X \sim B(200, p)$ donde p es la proporción de controladores defectuosos en el proceso

Las hipótesis son $H_0 : p = 0.05$ contra $H_1 : p < 0.05$

Como $\alpha = 0.05$ entonces $-z_\alpha = -z_{0.05} = -1.645$

El estadístico de prueba es $Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\hat{P} - 0.05}{\sqrt{\frac{0.05(1-0.05)}{200}}}$ y toma el valor $z_0 = -1.95$

Como $z_0 = -1.95 < -z_\alpha = -z_{0.05} = -1.645$ entonces se rechaza H_0 , y se concluye que la fracción de controladores defectuosos es menor que 0.05.

Calculamos el p-valor

$$p\text{-valor} = P(Z < z_0) = P(Z < -1.95) = \Phi(-1.95) = 0.0256$$

Valor de β y selección del tamaño de la muestra

Podemos obtener expresiones aproximadas para la probabilidad de cometer error de tipo II de manera análoga a las obtenidas para los test para la media

Si $H_1 : p \neq p_0$ entonces

$$\begin{aligned} \beta(p) &= P(\text{aceptar } H_0 / H_0 \text{ es falsa}) \approx \\ &\approx \Phi\left(\frac{p_0 - p + z_{\frac{\alpha}{2}} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) - \Phi\left(\frac{p_0 - p - z_{\frac{\alpha}{2}} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) \end{aligned}$$

Si $H_1 : p < p_0$ entonces

$$\beta(p) = P(\text{aceptar } H_0 / H_0 \text{ es falsa}) \approx 1 - \Phi\left(\frac{p_0 - p - z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

Si $H_1 : p > p_0$ entonces

$$\beta(p) = P(\text{aceptar } H_0 / H_0 \text{ es falsa}) \approx \Phi\left(\frac{p_0 - p + z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

Estas ecuaciones pueden resolverse para encontrar el tamaño aproximado de la muestra n para que con un nivel de significancia de α la probabilidad de cometer error de tipo II sea menor o igual que un valor específico β_0 . Las ecuaciones se deducen como en casos anteriores y son

$$\text{Si } H_1 : p \neq p_0 \text{ entonces } n \geq \left(\frac{z_{\frac{\alpha}{2}} \sqrt{p_0(1-p_0)} + z_{\beta_0} \sqrt{p(1-p)}}{p - p_0} \right)^2$$

$$\text{Si } H_1 : p < p_0 \text{ ó } H_1 : p > p_0 \text{ entonces } n \geq \left(\frac{z_{\alpha} \sqrt{p_0(1-p_0)} + z_{\beta_0} \sqrt{p(1-p)}}{p - p_0} \right)^2$$

Ejemplo:

Volviendo al ejemplo anterior, supongamos que la verdadera proporción de componentes defectuosos en el proceso es $p = 0.03$, ¿cuál es el valor de β si $n = 200$ y $\alpha = 0.05$?

Solución:

Ya que la alternativa es $H_1 : p < p_0$ aplicamos la fórmula

$$\begin{aligned} \beta(p) &= P(\text{aceptar } H_0 / H_0 \text{ es falsa}) \approx 1 - \Phi \left(\frac{p_0 - p - z_{\alpha} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} \right) = \\ &= 1 - \Phi \left(\frac{0.05 - 0.03 - 1.645 \sqrt{\frac{0.05(1-0.05)}{200}}}{\sqrt{\frac{0.03(1-0.03)}{200}}} \right) = 1 - \Phi(-0.44) = 0.67 \end{aligned}$$

Como la probabilidad de aceptar que el proceso tiene la calidad deseada cuando en realidad $p = 0.03$ es bastante alta, podemos preguntar qué tamaño de muestra se necesita para que en el test anterior sea $\beta < 0.1$ si la verdadera proporción de defectuosos es $p = 0.03$. En este caso aplicamos la fórmula donde $z_{\beta_0} = z_{0.1} = 1.28$

$$n \geq \left(\frac{z_{\alpha} \sqrt{p_0(1-p_0)} + z_{\beta_0} \sqrt{p(1-p)}}{p - p_0} \right)^2 = \left(\frac{1.645 \sqrt{0.05(1-0.05)} + 1.28 \sqrt{0.03(1-0.03)}}{0.03 - 0.05} \right)^2 \approx 832$$

La muestra requerida es muy grande, pero la diferencia a detectar $p - p_0 = 0.03 - 0.05$ es bastante pequeña.

9.11 – Tests de hipótesis sobre dos proporciones

Las pruebas de hipótesis sobre diferencia de medias pueden adaptarse al caso donde tenemos dos parámetros binomiales p_1 y p_2 de interés.

Específicamente, supongamos que se toman dos muestras aleatorias

$(X_{11}, X_{12}, \dots, X_{1n_1})$ es una muestra aleatoria de tamaño n_1 de X_1

$(X_{21}, X_{22}, \dots, X_{2n_2})$ es una muestra aleatoria de tamaño n_2 de X_2

Donde $X_1 \sim B(1, p_1)$; $X_2 \sim B(1, p_2)$ y X_1 y X_2 independientes.

Ya sabemos que $\hat{P}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i}$ y $\hat{P}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i}$ son estimadores *insesgados* de p_1 y p_2 respectivamente, con varianzas $V(\hat{P}_1) = \frac{p_1(1-p_1)}{n_1}$ y $V(\hat{P}_2) = \frac{p_2(1-p_2)}{n_2}$

Supongamos las hipótesis

$$H_0 : p_1 - p_2 = 0 \quad \text{contra} \quad H_1 : p_1 - p_2 \neq 0$$

Notar que si la hipótesis nula es verdadera entonces $p_1 = p_2 = p$, donde p es *desconocido*.

El estadístico $Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ tiene distribución *aproximadamente* $N(0,1)$ por T.C.L. si

$H_0 : p_1 - p_2 = 0$ es verdadera. Tomamos como *estimador de p* a $\hat{P} = \frac{\sum_{i=1}^{n_1} X_{1i} + \sum_{i=1}^{n_2} X_{2i}}{n_1 + n_2}$ y lo reemplazamos en Z

Entonces el estadístico de prueba es $Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ que bajo $H_0 : p_1 - p_2 = 0$ se puede probar que tiene distribución aproximadamente $N(0,1)$

Entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } |Z| > z_{\frac{\alpha}{2}} \\ \text{aceptar } H_0 & \text{si } |Z| \leq z_{\frac{\alpha}{2}} \end{cases}$ donde $Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

Si $H_1 : p_1 - p_2 > 0$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } Z > z_{\alpha} \\ \text{aceptar } H_0 & \text{si } Z \leq z_{\alpha} \end{cases}$

Si $H_1 : p_1 - p_2 < 0$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } Z < -z_{\alpha} \\ \text{aceptar } H_0 & \text{si } Z \geq -z_{\alpha} \end{cases}$

Ejemplo:

En una muestra de 100 lotes de un producto químico comprado al distribuidor A, 70 satisfacen una especificación de pureza. En una muestra de 70 lotes comprada al distribuidor B, 61 satisfacen la especificación. ¿Puede concluir que una proporción mayor de los lotes del distribuidor B satisface la especificación?

Solución:

Los parámetros de interés son p_1 y p_2 las verdaderas proporciones de lotes que cumplen las especificaciones de pureza.

Tenemos una muestra aleatoria $(X_{11}, X_{12}, \dots, X_{1n_1})$ de tamaño $n_1 = 100$ donde $\hat{P}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i} = \frac{70}{100} = 0.7$

Y otra muestra $(X_{21}, X_{22}, \dots, X_{2n_2})$ de tamaño $n_2 = 70$ donde $\hat{P}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i} = \frac{61}{70}$

Las hipótesis son $H_0 : p_1 - p_2 = 0$ contra $H_1 : p_1 - p_2 < 0$

El estadístico de prueba es $Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ donde $\hat{P} = \frac{\sum_{i=1}^{n_1} X_{1i} + \sum_{i=1}^{n_2} X_{2i}}{n_1 + n_2}$

En este caso $\hat{P} = \frac{\sum_{i=1}^{n_1} X_{1i} + \sum_{i=1}^{n_2} X_{2i}}{n_1 + n_2} = \frac{70 + 61}{100 + 70} = \frac{131}{170}$

El estadístico toma el valor $z_0 = \frac{\frac{70}{100} - \frac{61}{70}}{\sqrt{\frac{131}{170}\left(1 - \frac{131}{170}\right)\left(\frac{1}{100} + \frac{1}{70}\right)}} = -2.6163$

Para saber cuánta evidencia hay contra $H_0 : p_1 - p_2 = 0$ calculamos el p-valor

$p\text{-valor} = P(Z < z_0) = \Phi(-2.6163) = 0.0045$

Como el p-valor es menor que 0.05, se considera que hay mucha evidencia contra $H_0 : p_1 - p_2 = 0$ y se rechaza la hipótesis nula.

Valor de β

Cuando $H_0 : p_1 - p_2 = 0$ es falsa, la varianza de $\hat{P}_1 - \hat{P}_2$ es

$$V(\hat{P}_1 - \hat{P}_2) = V(\hat{P}_1) + V(\hat{P}_2) = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$$

Anotamos $\sigma_{\hat{P}_1 - \hat{P}_2} = \sqrt{V(\hat{P}_1 - \hat{P}_2)} = \sqrt{V(\hat{P}_1) + V(\hat{P}_2)} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

Entonces

Si $H_1 : p_1 - p_2 \neq 0$

$$\beta \approx \Phi \left(\frac{z_{\alpha/2} \sqrt{\bar{p}\bar{q} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} - (p_1 - p_2)}{\sigma_{\hat{p}_1 - \hat{p}_2}} \right) - \Phi \left(\frac{-z_{\alpha/2} \sqrt{\bar{p}\bar{q} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} - (p_1 - p_2)}{\sigma_{\hat{p}_1 - \hat{p}_2}} \right)$$

Donde $\bar{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$ y $\bar{q} = 1 - \bar{p}$

$$\text{Si } H_1 : p_1 - p_2 > 0 \text{ entonces } \beta \approx \Phi \left(\frac{z_{\epsilon} \sqrt{\bar{p}\bar{q} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} - (p_1 - p_2)}{\sigma_{\hat{p}_1 - \hat{p}_2}} \right)$$

$$\text{Si } H_1 : p_1 - p_2 < 0 \text{ entonces } \beta \approx 1 - \Phi \left(\frac{-z_{\alpha} \sqrt{\bar{p}\bar{q} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} - (p_1 - p_2)}{\sigma_{\hat{p}_1 - \hat{p}_2}} \right)$$

Podemos deducir fórmulas para el tamaño de la muestra, nuevamente asumiendo que $n_1 = n_2 = n$