

Inferencia estadística – Tests de hipótesis

Hasta ahora hemos visto como obtener, a partir de una muestra, un estimador puntual o un intervalo de confianza para un parámetro θ . Frecuentemente el objetivo del estudio es decidir, en base a la información que provee la muestra, entre dos hipótesis relativas a un parámetro.

Ejemplo: Supongamos que el consumo promedio de nafta de los motores utilizados por una empresa automotriz en uno de sus modelos es de 10 litros cada 100 km. Se presenta un proyecto de mejora del motor que produciría una disminución en el consumo pero, por razones de costo, se considera viable el proyecto si la reducción lleva el consumo a un valor menor de 9 litros cada 100 km.

Para estudiar la conveniencia o no de aplicar la mejora a los motores, se aplica esta mejora a una muestra de 25 motores, los cuáles se ponen a funcionar en igualdad de condiciones durante un periodo fijo. El consumo promedio observado es de 8.9 litros cada 100 km. ¿Proveen estos datos evidencia de que vale la pena incorporar la mejora al motor o se deben simplemente al azar?

Supongamos que el consumo de nafta de los motores es una v.a. con distribución normal con varianza igual a 1 y que la muestra es aleatoria, es decir que los 25 consumos son independientes. Es decir, supongamos que X_1, \dots, X_{25} es una m.a., $X_i \sim N(\mu, 1)$. Entonces

$$\bar{X} \sim N\left(\mu, \frac{1}{25}\right) \quad \Leftrightarrow \quad \frac{X - \mu}{\sqrt{1/25}} \sim N(0, 1)$$

Si la media verdadera del consumo en el motor mejorado fuese de 9 litros cada 100 km., ¿cuál es la probabilidad de que una v.a. normal con media 9 y varianza 1/25 tome un valor igual o menor que el observado, 8.9?

$$P(\bar{X} \leq 8.9) = P\left(\frac{\bar{X} - 9}{1/5} \leq \frac{8.9 - 9}{1/5}\right) = \Phi(-0.5) = 0.309 \cong 0.31$$

Esta probabilidad se denomina **p-valor**.

Si el consumo promedio observado hubiese sido $\bar{X} = 8.6$ litros cada 100 km, entonces

$$P(\bar{X} \leq 8.6) = P\left(\frac{\bar{X} - 9}{1/5} \leq \frac{8.6 - 9}{1/5}\right) = \Phi(-2) = 0.023,$$

es decir que, en este último caso, hubiese sido muy poco probable que se observase un valor promedio de 8.6 si la media verdadera es 9.

¿Qué es lo que estamos tratando de decidir? Nuestras hipótesis se refieren a μ , y se

podrían enunciar así:

- i) $\mu = 9$ litros cada 100 km. En este caso no se implementa la mejora a los motores
- ii) $\mu < 9$ litros cada 100 km. En este caso conviene implementar la mejora a los motores

A la primera hipótesis se la denomina **hipótesis nula** y se designa H_0 . Esta hipótesis implica que no hay efecto, es la hipótesis del status quo, o sea del no cambio respecto a la situación inicial. La segunda hipótesis se denomina **hipótesis alternativa** y se designa H_1 . Se la suele llamar la hipótesis del investigador.

Expresadas en términos del parámetro de interés las hipótesis del ejemplo serán

$$H_0: \mu = 9 \quad \text{vs} \quad H_1: \mu < 9$$

Un test es una regla de decisión basada en un **estadístico** o función de la muestra, en este caso \bar{X} , y en una **zona de rechazo**, es decir un conjunto de valores para los cuáles se rechaza la hipótesis nula H_0 .

¿Cómo se elige la zona de rechazo? Observemos que al tomar una decisión en base a una muestra, podemos cometer dos tipos de error.

	No se rechaza H_0	Se rechaza H_0
H_0 es cierta	OK	Error tipo I
H_0 no es cierta	Error tipo II	OK

Debido a la variabilidad muestral, es imposible construir tests en los cuáles estemos absolutamente seguros de tomar la decisión correcta,. Lo que podemos hacer es tratar de mantener bajas las probabilidades de error.

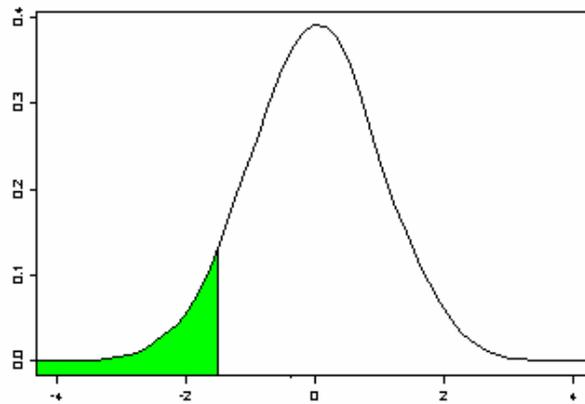
Llamaremos **nivel de significación del test**, y lo designaremos α , a la probabilidad de error tipo I (en realidad a la máxima probabilidad de error tipo I) y designaremos β a la probabilidad de error tipo II.

Como el estadístico se construye bajo la condición de que H_0 es verdadera, lo que podemos controlar es la probabilidad de error tipo I. Elegiremos la zona de rechazo del test de manera que la probabilidad de error tipo I sea un valor α predeterminado.

Volviendo al ejemplo, sabemos que, si H_0 es cierta,

$$\frac{\bar{X} - 9}{1/5} \sim N(0, 1)$$

Si queremos que el test tenga nivel de significación $\alpha = 0.05$, rechazaríamos H_0 si

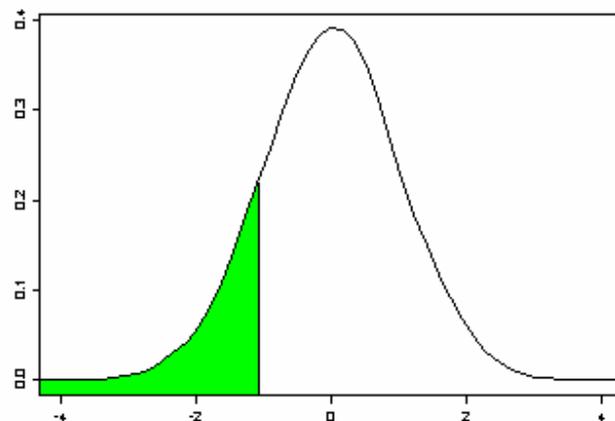


$$\frac{\bar{X} - 9}{1/5} \leq -1.64.$$

Esta es la zona de rechazo del test de nivel 0.05. Si observamos un promedio igual a 8.9, el valor del estadístico es -0.5 y por lo tanto no se rechaza H_0 , mientras que si observamos un promedio igual a 8.6, el valor del estadístico es -2 y se rechaza H_0 .

Si queremos que el test tenga nivel de significación $\alpha = 0.10$, rechazaríamos H_0 si

$$\frac{\bar{X} - 9}{1/5} \leq -1.28$$



Esta es la zona de rechazo del test de nivel 0.10.

Como hemos visto, al seleccionar la región de rechazo controlamos la probabilidad de error tipo I, pero ¿qué ocurre con el error tipo II?

Supongamos que en nuestro ejemplo, observamos un consumo promedio en la muestra de tamaño 25 igual a 8.9 litros cada 100 km y trabajamos con el test de nivel 0.05. En este caso, no rechazamos H_0 (tampoco lo haríamos con el test de nivel 0.10) y por lo tanto, si la mejora en el motor fuese real, podríamos estar cometiendo un error de tipo II.

Por ejemplo, si la modificación en el motor reduce el consumo a 8.5 litros cada 100 km, ¿cuál es la probabilidad de cometer un error tipo II?

$$P_{\mu=8.5} \left(\frac{\bar{X} - 9}{1/5} > -1.64 \right) = P_{\mu=8.5} \left(\bar{X} > -1.64 \cdot \frac{1}{5} + 9 \right) = P_{\mu=8.5} \left(\frac{\bar{X} - 8.5}{1/5} > \frac{-1.64 \cdot \frac{1}{5} + 9 - 8.5}{1/5} \right)$$

$$P_{\mu=8.5} \left(\frac{\bar{X} - 8.5}{1/5} > 0.86 \right) = 1 - \Phi(0.86) = 1 - 0.805 = 0.195$$

Es decir, que la probabilidad de error tipo II para el valor de $\mu = 8.5$ es aproximadamente 0.20.

Definición: La **función de potencia** de un test, $\pi(\mu)$, es la probabilidad de rechazar la hipótesis nula cuando el valor verdadero del parámetro es μ .

Utilizando la función de potencia es posible obtener una expresión general para los dos tipos de errores, pues

$$\pi(\mu) = \begin{cases} \alpha(\mu) & \text{si } \mu \in H_0 \\ 1 - \beta(\mu) & \text{si } \mu \in H_1 \end{cases}$$

donde $\alpha(\mu)$ y $\beta(\mu)$ denota las probabilidades de error tipo I y tipo II respectivamente cuando el verdadero valor del parámetro es μ .

Tipos de hipótesis a testear:

Hipótesis unilaterales:

$$H_0: \theta = \theta_0 \text{ (ó } \theta \leq \theta_0) \quad \text{vs} \quad H_1: \theta > \theta_0$$

$$H_0: \theta = \theta_0 \text{ (ó } \theta \geq \theta_0) \quad \text{vs} \quad H_1: \theta < \theta_0$$

Hipótesis bilaterales:

$$H_0: \theta = \theta_0 \quad \text{vs} \quad H_1: \theta \neq \theta_0$$

La forma de la región de rechazo dependerá de la hipótesis alternativa a testear. Así, en el ejemplo presentado anteriormente, la zona de rechazo consiste en un intervalo de valores en la cola izquierda de la distribución porque la hipótesis alternativa es de la forma $\mu < \mu_0$.

Tests de hipótesis de nivel α para los parámetros de la distribución normal:

Sea X_1, X_2, \dots, X_n una m.a. de una distribución $N(\mu, \sigma^2)$.

Tests para la media cuando la varianza es conocida: Supongamos que $\sigma^2 = \sigma_0^2$ es conocida y consideremos las siguientes hipótesis

- a) $H_0: \mu = \mu_0$ (ó $\mu \leq \mu_0$) vs $H_1: \mu > \mu_0$
- b) $H_0: \mu = \mu_0$ (ó $\mu \geq \mu_0$) vs $H_1: \mu < \mu_0$
- c) $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$

Estadístico del test: $T = \sqrt{n} \frac{\bar{X} - \mu_0}{\sigma_0}$. Bajo $H_0: \mu = \mu_0$, $T \sim N(0,1)$.

Región de rechazo: Como dijimos, la zona de rechazo depende de la hipótesis alternativa. Estará dada, en cada caso, por

- a) $T \geq z_\alpha$
- b) $T \leq -z_\alpha$
- c) $|T| \geq z_{\alpha/2}$

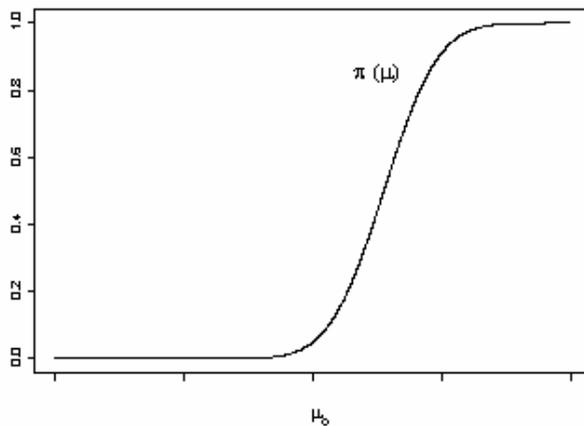
Observemos que, así como la forma de la región de rechazo depende de la alternativa, su tamaño depende del nivel. Por ejemplo, consideremos el caso c). Como la alternativa es $\mu \neq \mu_0$, la forma de la región es $|T| \geq K$, pero como la probabilidad de rechazar H_0 siendo cierta, o sea la probabilidad de Error tipo I, debe ser α ,

$$P_{\mu_0} \left(\left| \sqrt{n} \frac{\bar{X} - \mu_0}{\sigma_0} \right| \geq K \right) = \alpha \Leftrightarrow 1 - P_{\mu_0} \left(\left| \sqrt{n} \frac{\bar{X} - \mu_0}{\sigma_0} \right| < K \right) = 1 - P_{\mu_0} \left(-K < \sqrt{n} \frac{\bar{X} - \mu_0}{\sigma_0} < K \right) = \alpha$$

$$\Leftrightarrow 1 - \Phi(K) + \Phi(-K) = \alpha \Leftrightarrow 2(1 - \Phi(K)) = \alpha \Leftrightarrow \Phi(K) = 1 - \frac{\alpha}{2} \Leftrightarrow K = z_{\alpha/2}$$

Función de potencia: La notación P_μ , como ya hemos visto, indicará la probabilidad cuando el valor verdadero del parámetro es μ . Hallaremos la función de potencia para cada uno de los tests planteados.

$$\begin{aligned} \text{a) } \pi(\mu) &= P_\mu \left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha \right) = P_\mu \left(\frac{\bar{X} - \mu + \mu - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha \right) \\ &= P_\mu \left(\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} \geq z_\alpha + \frac{\mu_0 - \mu}{\sigma_0/\sqrt{n}} \right) = 1 - \Phi \left(z_\alpha + \frac{\mu_0 - \mu}{\sigma_0/\sqrt{n}} \right) \end{aligned}$$

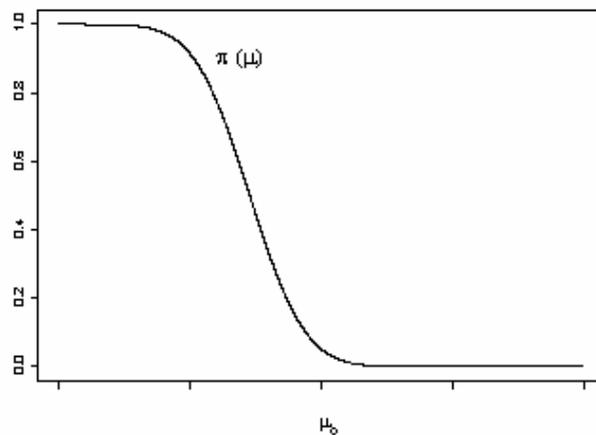


Observemos que esta función es creciente y $\pi(\mu_0) = \alpha$, entonces, si $\mu < \mu_0$, $\pi(\mu) < \alpha$. Por esta razón el test también es de nivel α para las hipótesis

$$H_0: \mu \leq \mu_0 \quad \text{vs} \quad H_1: \mu > \mu_0$$

en el sentido de que la probabilidad de error tipo I es a lo sumo α .

$$\begin{aligned}
 \text{b) } \pi(\mu) &= P_{\mu} \left(\frac{\bar{X} - \mu_0}{\sigma_0 / \sqrt{n}} \leq -z_{\alpha} \right) = P_{\mu} \left(\frac{\bar{X} - \mu + \mu - \mu_0}{\sigma_0 / \sqrt{n}} \leq -z_{\alpha} \right) \\
 &= P_{\mu} \left(\frac{\bar{X} - \mu}{\sigma_0 / \sqrt{n}} \leq -z_{\alpha} + \frac{\mu_0 - \mu}{\sigma_0 / \sqrt{n}} \right) = \Phi \left(-z_{\alpha} + \frac{\mu_0 - \mu}{\sigma_0 / \sqrt{n}} \right)
 \end{aligned}$$



Observemos que esta función es decreciente y $\pi(\mu_0) = \alpha$, entonces, si $\mu > \mu_0$, $\pi(\mu) < \alpha$. Por esta razón el test también es de nivel α para las hipótesis

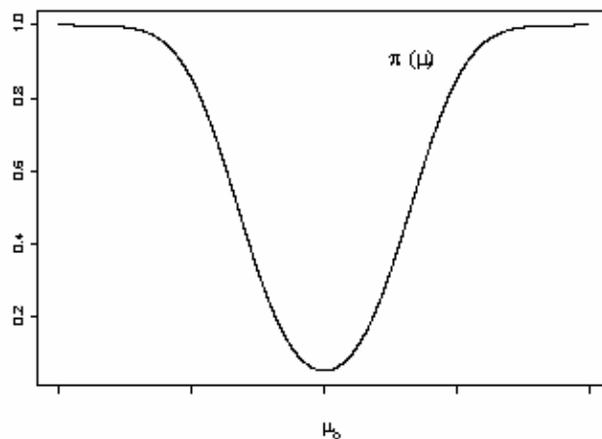
$$H_0: \mu \geq \mu_0 \quad \text{vs} \quad H_1: \mu > \mu_0$$

en el sentido de que la probabilidad de error tipo I es a lo sumo α .

$$\begin{aligned}
 \text{c) } \pi(\mu) &= P_{\mu} \left(\left| \frac{\bar{X} - \mu_0}{\sigma_0 / \sqrt{n}} \right| \geq z_{\alpha/2} \right) = 1 - P_{\mu} \left(\left| \frac{\bar{X} - \mu_0}{\sigma_0 / \sqrt{n}} \right| < z_{\alpha/2} \right) \\
 &= 1 - P_{\mu} \left(-z_{\alpha/2} < \frac{\bar{X} - \mu + \mu - \mu_0}{\sigma_0 / \sqrt{n}} < z_{\alpha/2} \right)
 \end{aligned}$$

$$= 1 - P_{\mu} \left(-z_{\alpha/2} + \frac{\mu_o - \mu}{\sigma_o / \sqrt{n}} < \frac{\bar{X} - \mu}{\sigma_o / \sqrt{n}} < z_{\alpha/2} + \frac{\mu_o - \mu}{\sigma_o / \sqrt{n}} \right)$$

$$= 1 - \Phi \left(z_{\alpha/2} + \frac{\mu_o - \mu}{\sigma_o / \sqrt{n}} \right) + \Phi \left(-z_{\alpha/2} + \frac{\mu_o - \mu}{\sigma_o / \sqrt{n}} \right)$$



Observemos que esta función decrece hasta μ_o donde $\pi(\mu_o) = \alpha$ y crece a partir de allí.

Tamaño de muestra requerido para obtener una probabilidad de error tipo II dada para un valor $\mu = \mu_1$ (fijo) en la alternativa: Recordemos que el error de tipo II se define como “aceptar la hipótesis nula H_o cuando es falsa”. Buscamos el valor de n para que la probabilidad de error tipo II sea menor que β cuando $\mu = \mu_1$ es un valor fijo en H_1 .

$$\text{a) } P_{\mu_1} \left(\frac{\bar{X} - \mu_o}{\sigma_o / \sqrt{n}} < z_{\alpha} \right) \leq \beta \Leftrightarrow 1 - \pi(\mu_1) \leq \beta \Leftrightarrow \pi(\mu_1) \geq 1 - \beta$$

$$\Leftrightarrow 1 - \Phi \left(z_{\alpha} + \frac{\mu_o - \mu_1}{\sigma_o / \sqrt{n}} \right) \geq 1 - \beta \Leftrightarrow \Phi \left(z_{\alpha} + \frac{\mu_o - \mu_1}{\sigma_o / \sqrt{n}} \right) \leq \beta \Leftrightarrow z_{\alpha} + \frac{\mu_o - \mu_1}{\sigma_o / \sqrt{n}} \leq z_{1-\beta}$$

Observemos que en este caso la alternativa es $H_1: \mu > \mu_0$, por lo tanto, $\mu_0 - \mu_1 < 0$ y se obtiene

$$n \geq \left(\frac{(z_\alpha - z_{1-\beta})\sigma_o}{\mu_1 - \mu_0} \right)^2 = \left(\frac{(z_\alpha + z_\beta)\sigma_o}{\mu_1 - \mu_0} \right)^2$$

$$b) P_{\mu_1} \left(\frac{\bar{X} - \mu_0}{\sigma_o / \sqrt{n}} > -z_\alpha \right) \leq \beta \Leftrightarrow 1 - \pi(\mu_1) \leq \beta \Leftrightarrow \pi(\mu_1) \geq 1 - \beta$$

$$\Leftrightarrow \Phi \left(-z_\alpha + \frac{\mu_0 - \mu_1}{\sigma_o / \sqrt{n}} \right) \geq 1 - \beta \Leftrightarrow -z_\alpha + \frac{\mu_0 - \mu_1}{\sigma_o / \sqrt{n}} \geq z_\beta$$

Observemos que en este caso la alternativa es $H_1: \mu < \mu_0$, por lo tanto, $\mu_0 - \mu_1 > 0$ y se obtiene

$$n \geq \left(\frac{(z_\alpha + z_\beta)\sigma_o}{\mu_0 - \mu_1} \right)^2$$

c) Para el caso bilateral, el cálculo del tamaño de muestra se hace en forma aproximada, despreciando la más pequeña de las dos probabilidades.

Tests para la media cuando la varianza es desconocida: Supongamos ahora que la varianza es desconocida y consideremos las mismas hipótesis sobre μ .

- | | | |
|---|----|-----------------------|
| a) $H_0: \mu = \mu_0$ (ó $\mu \leq \mu_0$) | vs | $H_1: \mu > \mu_0$ |
| b) $H_0: \mu = \mu_0$ (ó $\mu \geq \mu_0$) | vs | $H_1: \mu < \mu_0$ |
| c) $H_0: \mu = \mu_0$ | vs | $H_1: \mu \neq \mu_0$ |

Estadístico del test: $T = \sqrt{n} \frac{\bar{X} - \mu_0}{S}$. Bajo $H_0: \mu = \mu_0$, $T \sim t_{n-1}$

Región de rechazo: Como siempre la forma de la zona de rechazo depende de la hipótesis alternativa. Estará dada, en cada caso, por

- $T \geq t_{n-1, \alpha}$
- $T \leq -t_{n-1, \alpha}$

c) $|T| \geq t_{n-1, \alpha/2}$

El tamaño de la zona de rechazo depende del nivel. Por ejemplo, consideremos el caso a). Como la alternativa es $\mu > \mu_0$, la forma de la región es $T \geq K$, pero como la probabilidad de rechazar H_0 siendo cierta, o sea la probabilidad de Error tipo I, debe ser α ,

$$P_{\mu_0} \left(\sqrt{n} \frac{\bar{X} - \mu_0}{S} \geq K \right) = \alpha \Leftrightarrow 1 - P_{\mu_0} \left(\sqrt{n} \frac{\bar{X} - \mu_0}{S} \leq K \right) = \alpha$$

$$\Leftrightarrow 1 - F_T(K) = \alpha \Leftrightarrow F_T(K) = 1 - \alpha \Leftrightarrow K = t_{n-1, \alpha}$$

donde F_T designa la función de distribución de una v.a. t con $n-1$ grados de libertad.

Función de potencia y cálculo del tamaño de muestra para obtener una probabilidad de error tipo II dada: La función de potencia de este test es complicada porque la distribución del estadístico cuando $\mu \neq \mu_0$ es una distribución t no central. Aunque hay tablas y gráficos que permiten obtener probabilidades para una distribución de este tipo, no los estudiaremos en este curso. Por la misma razón, no calcularemos tamaño de muestra para obtener una probabilidad de error tipo II dada para una alternativa fija.

Respecto al p-valor, cuando se utilizan tablas sólo es posible obtener una cota, ya que las tablas proveen solamente algunos valores críticos de la distribución t .

Tests para la varianza cuando la media es desconocida: Las hipótesis a testear son

- a) $H_0: \sigma^2 = \sigma_0^2$ (ó $\sigma^2 \leq \sigma_0^2$) vs $H_1: \sigma^2 > \sigma_0^2$
 b) $H_0: \sigma^2 = \sigma_0^2$ (ó $\sigma^2 \geq \sigma_0^2$) vs $H_1: \sigma^2 < \sigma_0^2$
 c) $H_0: \sigma^2 = \sigma_0^2$ vs $H_1: \sigma^2 \neq \sigma_0^2$

Estadístico del test: $U = \frac{(n-1)S^2}{\sigma_0^2}$. Bajo $H_0: \sigma^2 = \sigma_0^2$, $U \sim \chi_{n-1}^2$.

Región de rechazo: Como siempre la forma de la zona de rechazo depende de la hipótesis alternativa. En este caso, estará dada por

- a) $U \geq \chi_{n-1, \alpha}^2$
 b) $U \leq \chi_{n-1, 1-\alpha}^2$
 c) $U \geq \chi_{n-1, \alpha/2}^2$ ó $U \leq \chi_{n-1, 1-\alpha/2}^2$

El tamaño de la zona de rechazo depende del nivel. Por ejemplo, consideremos el caso b). Como la alternativa es $\sigma^2 < \sigma_0^2$, la forma de la región es $U \leq K$, pero como la

probabilidad de rechazar H_0 siendo cierta ($P(\text{Error tipo I})$) debe ser α ,

$$P_{\sigma_0^2} \left(\frac{(n-1)S^2}{\sigma_0^2} \leq K \right) = \alpha \Leftrightarrow K = \chi_{n-1, \alpha}^2$$

Función de potencia: La obtendremos sólo para el caso b). Los otros dos casos se obtienen en forma similar.

$$\pi(\sigma_1^2) = P_{\sigma_1^2} \left(\frac{(n-1)S^2}{\sigma_0^2} \leq \chi_{n-1, 1-\alpha}^2 \right) = P_{\sigma_1^2} \left(\frac{(n-1)S^2}{\sigma_1^2} \leq \frac{\sigma_0^2}{\sigma_1^2} \chi_{n-1, 1-\alpha}^2 \right) = F_{\chi_{n-1}^2} \left(\frac{\sigma_0^2}{\sigma_1^2} \chi_{n-1, 1-\alpha}^2 \right)$$

donde $F_{\chi_{n-1}^2}$ indica la función de distribución chi-cuadrado con $n-1$ grados de libertad.

Utilizando tablas sólo es posible obtener una cota para la potencia ya que las tablas proveen solamente algunos valores críticos de la distribución χ^2 .

Por la misma razón, al calcular el p-valor utilizando tablas, sólo es posible obtener una cota.

Ejercicio: ¿Qué estadístico utilizaría en el caso en que la media μ fuese conocida?. ¿Cuál es la distribución de dicho estadístico? ¿Cómo se modifican las regiones de rechazo y la función de potencia de los tests?

Ejemplo: Se toman 25 determinaciones de la temperatura en cierto sector de un reactor, obteniéndose

$$\bar{x} = 243^\circ C \quad \text{y} \quad s = 2.8^\circ C$$

Interesa saber, a nivel 0.05

- si existe evidencia para decidir que la temperatura media en ese sector del reactor es menor que $250^\circ C$.
- si existe evidencia para decidir que la varianza de la temperatura en ese sector del reactor es mayor que $(2^\circ C)^2$.

a) Las hipótesis a testear son

$$H_0: \mu = 250 \text{ (ó } \mu \geq 250) \quad \text{vs} \quad H_1: \mu < 250$$

El estadístico del test será $T = \sqrt{n} \frac{\bar{X} - 250}{S}$,

y la región de rechazo estará dada por los valores de T tales que

$$T = \sqrt{n} \frac{\bar{X} - 250}{S} \leq -t_{n-1, 0.05}$$

En nuestro caso, $n = 25$ y por lo tanto $-t_{24, 0.05} = -1.71$. Como el valor observado de T es -12.5 , se rechaza H_0 , es decir hay evidencia de que la temperatura media del reactor es menor que $250^\circ C$.

b) Las hipótesis a testear son

$$H_0: \sigma^2 = 4 \text{ (ó } \sigma^2 \leq 4) \quad \text{vs} \quad H_1: \sigma^2 > 4$$

El estadístico del test será $U = \frac{(n-1)S^2}{\sigma^2}$,

y la región de rechazo estará dada por los valores de U tales que

$$U = \frac{(n-1)S^2}{4} \geq \chi_{n-1, 0.05}^2$$

En nuestro caso, $n = 25$ y por lo tanto $\chi_{24, 0.05}^2 = 36.42$. Como el valor observado de U es 47.04 , se rechaza H_0 . Es decir, hay evidencia de que la varianza de la temperatura del reactor es mayor que $(2^\circ C)^2$.

Tests de hipótesis de nivel aproximado (o asintótico) α para la media de una distribución cualquiera:

Sea X_1, X_2, \dots, X_n una m.a. de una distribución con media μ y varianza $\sigma^2 < \infty$. Aplicando el Teorema Central del Límite, sabemos que

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$$

Además, utilizando la propiedad enunciada al construir intervalos de confianza de nivel asintótico $(1 - \alpha)$ para la media de una distribución cualquiera,

$$\left. \begin{array}{l} \sqrt{n} \frac{\bar{X} - \mu}{\sigma} \xrightarrow{d} N(0,1) \\ \frac{\sigma}{S} \xrightarrow{p} 1 \end{array} \right\} \Rightarrow \sqrt{n} \frac{\bar{X} - \mu}{S} \xrightarrow{d} N(0,1)$$

Por lo tanto, si n es suficientemente grande,

$$\sqrt{n} \frac{\bar{X} - \mu}{S} \stackrel{(a)}{\sim} N(0,1)$$

Supongamos que se desea testear a nivel aproximado α alguna de las hipótesis siguientes:

- | | | |
|---|----|-----------------------|
| a) $H_0: \mu = \mu_0$ (ó $\mu \leq \mu_0$) | vs | $H_1: \mu > \mu_0$ |
| b) $H_0: \mu = \mu_0$ (ó $\mu \geq \mu_0$) | vs | $H_1: \mu < \mu_0$ |
| c) $H_0: \mu = \mu_0$ | vs | $H_1: \mu \neq \mu_0$ |

y que n es suficientemente grande. Utilizando como estadístico $T = \sqrt{n} \frac{\bar{X} - \mu_0}{s}$, las siguientes regiones de rechazo proveen tests de nivel aproximado α para cada una de las hipótesis:

- a) $T \geq z_{\alpha}$
- b) $T \leq -z_{\alpha}$
- c) $|T| \geq z_{\alpha/2}$

Función de potencia aproximada: Un estimador de la función de potencia puede obtenerse reemplazando el estadístico S por su valor observado s , o sea:

$$\begin{aligned} \pi(\mu) &= P_{\mu} \left(\left| \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \right| \geq z_{\alpha/2} \right) = 1 - P_{\mu} \left(\left| \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \right| \leq z_{\alpha/2} \right) \\ &= 1 - P_{\mu} \left(-z_{\alpha/2} \leq \frac{\bar{X} - \mu + \mu - \mu_0}{s/\sqrt{n}} \leq z_{\alpha/2} \right) \\ &= 1 - P_{\mu} \left(-z_{\alpha/2} + \frac{\mu_0 - \mu}{s/\sqrt{n}} \leq \frac{\bar{X} - \mu}{s/\sqrt{n}} \leq z_{\alpha/2} + \frac{\mu_0 - \mu}{s/\sqrt{n}} \right) \end{aligned}$$

$$\cong 1 - \Phi \left(z_{\alpha/2} + \frac{\mu_o - \mu}{s/\sqrt{n}} \right) + \Phi \left(-z_{\alpha/2} + \frac{\mu_o - \mu}{s/\sqrt{n}} \right)$$

En forma similar, se obtiene la función de potencia aproximada en los otros dos casos.

Ejemplo: En algunos casos, la varianza y la media dependen del mismo parámetro y no es necesario reemplazar σ por un estimador. Por ejemplo sea X_1, X_2, \dots, X_n una m.a. de una distribución de Poisson de parámetro λ . Entonces, si n es suficientemente grande,

$$\frac{\bar{X} - \lambda}{\sqrt{\lambda/n}} \stackrel{(a)}{\sim} N(0,1)$$

Supongamos que se desea testear a nivel aproximado α

$$H_0: \lambda = \lambda_o \quad \text{vs} \quad H_1: \lambda > \lambda_o$$

Entonces, bajo H_0 ,

$$\frac{\bar{X} - \lambda_o}{\sqrt{\lambda_o/n}} \stackrel{(a)}{\sim} N(0,1)$$

y, el test con región de rechazo

$$\frac{\bar{X} - \lambda_o}{\sqrt{\lambda_o/n}} \geq z_\alpha$$

tiene nivel aproximado α .

Su función de potencia aproximada se obtiene en la forma siguiente:

$$\begin{aligned} \pi(\lambda_1) &= P_{\lambda_1} \left(\frac{\bar{X} - \lambda_o}{\sqrt{\lambda_o/n}} \geq z_\alpha \right) = P_{\lambda_1} \left(\bar{X} \geq z_\alpha \sqrt{\lambda_o/n} + \lambda_o \right) = P_{\lambda_1} \left(\frac{\bar{X} - \lambda_1}{\sqrt{\lambda_1/n}} \geq \frac{z_\alpha \sqrt{\lambda_o/n} + \lambda_o - \lambda_1}{\sqrt{\lambda_1/n}} \right) \\ &\cong 1 - \Phi \left(z_\alpha \sqrt{\frac{\lambda_o}{\lambda_1}} + \frac{\lambda_o - \lambda_1}{\sqrt{\lambda_1/n}} \right) \end{aligned}$$

Test de hipótesis de nivel aproximado (o asintótico) α para una proporción (parámetro p de la distribución binomial): Sea X_1, X_2, \dots, X_n una m.a. de una

distribución $\text{Bi}(1, p)$. Entonces, $X = \sum_{i=1}^n X_i \sim \text{Bi}(n, p)$. Aplicando el Teorema Central del Límite, si n es suficientemente grande,

$$\frac{\bar{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{d} Z \sim N(0,1)$$

siendo \bar{X} la proporción muestral o frecuencia relativa de éxitos.

Un test de nivel aproximado α para las hipótesis:

- a) $H_0: p = p_0$ vs $H_1: p > p_0$
- b) $H_0: p = p_0$ vs $H_1: p < p_0$
- c) $H_0: p = p_0$ vs $H_1: p \neq p_0$

se basa en el estadístico $\frac{\bar{X} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$, el cual, si H_0 es cierta, tiene distribución

aproximada $N(0,1)$. Las regiones de rechazo estarán dadas por

a) $\frac{\bar{X} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \geq z_\alpha$

b) $\frac{\bar{X} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \leq -z_\alpha$

c) $\left| \frac{\bar{X} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \right| \geq z_{\alpha/2}$

Ejercicio: Deducir las funciones de potencia aproximadas en los 3 casos.

Relación entre tests de hipótesis bilaterales e intervalos de confianza:

Introduciremos esta idea a través de un ejemplo. Sea X_1, X_2, \dots, X_n una m.a. de una distribución $N(\mu, \sigma^2)$. Sabemos que el intervalo de confianza para μ de nivel $1 - \alpha$ está dado por

$$\left[\bar{X} - t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}, \bar{X} + t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} \right].$$

Supongamos ahora que deseamos testear a nivel α las siguientes hipótesis:

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

Dado que el intervalo construido contiene con alta probabilidad al valor verdadero de μ , si μ_0 no pertenece al intervalo, ésto nos llevaría a sospechar que la hipótesis nula es falsa.

Es decir, podríamos construir un test de nivel α rechazando H_0 si μ_0 no pertenece al intervalo de confianza, dado que

$$\begin{aligned} P(EI) &= P_{\mu_0} \left(\mu_0 \notin \left[\bar{X} - t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}, \bar{X} + t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} \right] \right) \\ &= 1 - P_{\mu_0} \left(\mu_0 \in \left[\bar{X} - t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}, \bar{X} + t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} \right] \right) = 1 - (1 - \alpha) = \alpha. \end{aligned}$$

Proposición: Sea $IC(X_1, X_2, \dots, X_n)$ un intervalo de confianza de nivel $1 - \alpha$ para un parámetro θ , obtenido a partir de una m.a. X_1, X_2, \dots, X_n . Consideremos el problema de testear las hipótesis

$$H_0: \theta = \theta_0 \quad \text{vs} \quad H_1: \theta \neq \theta_0$$

El test que rechaza H_0 cuando $\theta_0 \notin IC(X_1, X_2, \dots, X_n)$, tiene nivel α .

Ejemplo: Sea X_1, X_2, \dots, X_n una m.a. de una distribución exponencial de parámetro λ .

Recordemos que, usando que $2\lambda \sum_{i=1}^n X_i \sim \chi_{2n}^2$, hemos obtenido el siguiente intervalo de confianza de nivel exacto $1 - \alpha$ para λ

$$IC_{\lambda} = \left[\frac{\chi_{2n,1-\alpha/2}^2}{2 \sum_{i=1}^n X_i}, \frac{\chi_{2n,\alpha/2}^2}{2 \sum_{i=1}^n X_i} \right]$$

Si deseamos testear las hipótesis

$$H_0: \lambda = \lambda_0 \quad \text{vs} \quad H_1: \lambda \neq \lambda_0$$

El test que rechaza H_0 si $\lambda_0 \notin IC_{\lambda}$ tiene nivel α .