(1) Tengo una muestra aleatoria tamaño n tomada de una población X con

$$E(X) = \mu \qquad V(X) = \sigma^2$$

Quiero saber cuál es el mejor estimador de μ, este sería $\overline{X}_2$ pues

$$\overline{X}_1 = \frac{1}{n-1} \sum_{i=1}^{n-1} x_i \qquad \overline{X}_2 = \frac{1}{n} \sum_{i=1}^{n} x_i$$

↓ ESTADISTICO usual para estimar σ^2 ↓ ESTADISTICO usual para estimar μ

(2) $X_1, X_2, \dots, X_7$ una muestra aleatoria de una población con media μ y varianza σ^2

Estimadores de μ

$$\hat{\Theta}_1 = \frac{X_1 + X_2 + \dots + X_7}{7} \qquad \hat{\Theta}_2 = \frac{2X_1 - X_6 + X_4}{2} \qquad \hat{\Theta}_3 = \frac{2X_1 - X_7 + X_3}{3}$$

a) Quiero saber si algún estimador de μ es insesgado

sabemos que $\hat{\Theta}$ es insesgado de μ si $E(\hat{\Theta}) = \mu$ (pag 136)

aplicamos $\circledast_1$ propiedad de linealidad de la Esperanza →

• $E(\hat{\Theta}_1) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \underset{n=7}{=} E\left(\frac{1}{7}\sum_{i=1}^{7} X_i\right) = \frac{1}{7} E\left(\sum_{i=1}^{7} X_i\right) = \frac{1}{7}\sum_{i=1}^{7} E(X_i)$

Luego, por ser componentes de una muestra aleatoria vale $E(X_i) = E(X) = \mu$ $\circledast_2$

Luego

$$\frac{1}{7}\underbrace{\sum_{i=1}^{7}}_{7}\underbrace{E(X_i)}_{\mu} = \frac{1}{7}\, 7\mu = \mu$$

Como probamos que $E(\hat{\Theta}_1) = \mu$ $\hat{\Theta}_1$ es insesgado ✓

$\circledast_1$ • $E(\hat{\Theta}_2) = \frac{1}{2} E(2X_1 - X_6 + X_4) = \frac{1}{2}\left[E(2X_1) - E(X_6) + E(X_4)\right] =$

$= \frac{1}{2}\left[2E(X_1) - E(X_6) + E(X_4)\right] \underset{\circledast_2}{=} \frac{1}{2}\left[2\mu - \mu + \mu\right] = \frac{1}{2}\, 2\mu = \mu$

Como probamos que $E(\hat{\Theta}_2) = \mu$ $\hat{\Theta}_2$ es insesgado ✓

• $E(\hat{\Theta}_3) = \frac{1}{3} E(2X_1 - X_7 + X_3) \underset{\circledast_1}{=} \frac{1}{3}\left[2E(X_1) - E(X_7) + E(X_3)\right] \underset{\circledast_2}{=}$

$= \frac{1}{3}(2\mu - \mu + \mu) = \frac{2}{3}\mu \neq \mu$

Como $E(\hat{\Theta}_3) = \frac{2}{3}\mu \neq \mu$, $\hat{\Theta}_3$ NO es insesgado ✗

b) Quiero determinar el error cuadrático medio de los estimadores

- Si el estimador es INSESGADO, el error cuadrático medio será $V(\hat{\theta})$ entonces

$$ECM(\hat{\theta}_1) = V(\hat{\theta}_1) \text{ y } ECM(\hat{\theta}_2) = V(\hat{\theta}_2)$$

(*) $ECM(\hat{\theta}) = V(\hat{\theta})$ (si $\hat{\theta}$ es INSESGADO) pues $b(\hat{\theta}) = 0$

- $ECM(\hat{\theta}_1) = V(\hat{\theta}_1) = V\left(\frac{1}{7}\sum_{i=1}^{7} X_i\right) = \frac{1}{7^2}\sum_{i=1}^{7} \underbrace{V(X_i)}_{=\sigma^2} =$

$$= \frac{1}{7^2} \cdot 7\sigma^2 = \frac{\sigma^2}{7}$$

- $ECM(\hat{\theta}_2) = V(\hat{\theta}_2) = \frac{1}{2^2} V(2X_1 - X_6 + X_4) = \frac{1}{2^2}\left[2^2 V(X_1) + V(X_6) + V(X_4)\right] =$

$$= \frac{1}{2^2}(4\sigma^2 + \sigma^2 + \sigma^2) \quad \frac{6}{4}\sigma^2 = \frac{3\sigma^2}{2}$$

- Si el estimador es SESGADO, $ECM(\hat{\theta}) = V(\hat{\theta}) + (b(\hat{\theta}))^2$

(*) SESGO DE ESTIMADOR

$$b(\hat{\theta}) = E(\hat{\theta}) - \theta$$

(en nuestro caso) $\theta = \mu$

- ① $V(\hat{\theta}_3) = \frac{1}{3^2} V(2X_1 - X_7 + X_3) = \frac{1}{9}\left[4V(X_1) + V(X_7) + V(X_3)\right] =$

$$= \frac{1}{9}(4\sigma^2 + \sigma^2 + \sigma^2) = \frac{6}{9}\sigma^2 = \frac{2}{3}\sigma^2$$

② $(b(\hat{\theta}_3))^2 = (E(\hat{\theta}_3) - \mu)^2 = \left(\frac{2}{3}\mu - \mu\right)^2 = \left(-\frac{1}{3}\mu\right)^2 = \frac{1}{9}\mu^2$

Luego, por ① y ②, $ECM(\hat{\theta}_3) = \frac{2}{3}\sigma^2 + \frac{1}{9}\mu^2$

c) Quiero determinar el mejor estimador (el más eficiente)

$$\frac{ECM(\hat{\theta}_1)}{ECM(\hat{\theta}_2)} = \frac{\sigma^2/7}{3\sigma^2/2} = \frac{2}{21} < 1 \rightarrow \hat{\theta}_1 \text{ es más eficiente que } \hat{\theta}_2$$

[$\hat{\theta}_3$ no es insesgado, por lo tanto no será mejor estimador que $\hat{\theta}_1$ y $\hat{\theta}_2$]

③ Sea $X_1, X_2, \ldots, X_m$, muestra aleatoria tamaño m

a) Quiero mostrar que $\bar{X}^2$ es un estimador sesgado de μ^2

sabemos que $\cdot\ E(\bar{X}) = E\left(\frac{1}{m}\sum_{i=1}^{m} X_i\right) = \frac{1}{m} \cdot m\, E(X_i) = \mu$

$$\cdot\ V(\bar{X}) = V\left(\frac{1}{m-1}\sum_{i=1}^{m} X_i\right) \ \frac{1}{m} V(X_i) = \frac{\sigma^2}{m}$$

si $\bar{X}^2$ es un estimador sesgado de $\mu^2 \rightarrow E(\bar{X}^2) \neq \mu^2$

$E(\bar{X}^2) = V(\bar{X}) + (E(\bar{X}))^2 = \frac{\sigma^2}{m} + \mu^2 \neq \mu^2 \therefore \bar{X}^2$ es sesgado de μ^2

Escaneado con CamScanner

b) Quiero determinar la magnitud del sesgo de este estimador

$$b(\hat{\theta}) = b(\bar{X}^2) = E(\bar{X}^2) - \mu^2 = \frac{\sigma^2}{n} + \mu^2 - \mu^2 = \frac{\sigma^2}{n}$$

c) Como $b(\bar{X}^2) = \frac{\sigma^2}{n}$. $\lim_{n \to \infty} b(\bar{X}^2) = \lim_{n \to \infty} \sigma^2/n = 0$

Al tomar límite cuando $n \to \infty$, vemos que a medida que aumenta n, el sesgo tenderá a 0.
Por lo tanto, el estimador resulta ASINTÓTICAMENTE INSESGADO

(4) Tenemos $X \sim P(\lambda)$, X = "número diario de desconexiones accidentales"

En 5 días observamos $x_i = 2, 5, 3, 3, 7$ sea $\theta = \lambda$ (parámetro desconocido)

a) Quiero el ESTIMADOR DE MÁXIMA VEROSIMILITUD de λ

$$L(2, 5, 3, 3, 7, \lambda) = f(2, \lambda).f(5, \lambda).f(3, \lambda).f(3, \lambda).f(7, \lambda) = \prod_{i=1}^{n} f(x_i, \lambda)$$

$$= \prod_{i=1}^{n} P(X = x_i) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$$

$$= e^{-n\lambda} \lambda^{\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} 1/x_i! \quad (*)$$

$$\ln(L(x_i, \lambda)) = \ln\left(\frac{e^{-n\lambda} \lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}\right) = \quad \rightarrow \text{tomamos } \ln(L)$$

$$= \ln(e^{-n\lambda}) + \ln(\lambda^{\sum_{i=1}^{n} x_i}) - \ln(\prod_{i=1}^{n} x_i!)$$

$$= -n\lambda + \sum_{i=1}^{n} x_i \ln(\lambda) - \underbrace{\sum_{i=1}^{n} \ln(x_i!)}_{cte}$$

(*) . $\prod_{i=1}^{n} e^{-\lambda} = e^{-n\lambda}$

$\underbrace{(e^{-\lambda}).(e^{-\lambda})\dots(e^{-\lambda})}_{n \text{ veces}}$

. $\prod_{i=1}^{n} \lambda^{x_i} = \lambda^{\sum_{i=1}^{n} x_i}$ (sumo exponentes)

Derivamos y maximizamos

$$\frac{d(\ln(L(x_i, \lambda)))}{d\lambda} = -n + \sum_{i=1}^{n} x_i . \frac{1}{\lambda} + 0 = 0$$

$$\Leftrightarrow \sum_{i=1}^{n} x_i . \frac{1}{\lambda} = n$$

$$\underbrace{\sum_{i=1}^{n} x_i/n}_{\text{Estadístico usual}} = \hat{\lambda}$$

$\bar{X} = \hat{\lambda}$ → Estimador de máxima verosimilitud de λ

• Es INSESGADO pues $E(\bar{X}) = \sum_{i=1}^{n} x_i/n = \mu = \hat{\lambda}$

Estadístico usual para estimar μ

- Es CONSISTENTE si $\hat{\lambda} \xrightarrow{P} \lambda$ cuando $n \to \infty$

$$\hat{\lambda} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Como

- $E(\hat{\lambda}) = \lambda$ $\qquad \lim_{n\to\infty} E(\hat{\lambda}) = \lambda$
- $V(\hat{\lambda}) = \frac{\lambda}{n}$ $\qquad \lim_{n\to\infty} V(\hat{\lambda}) = 0$

$\therefore$ El estimador es CONSISTENTE

b) Quiero obtener la estimación de λ a partir de la muestra dada

$$L(2,5,3,3,7,\lambda) = \frac{1}{5}\sum_{i=1}^{5} X_i = \frac{2+5+3+3+7}{5} = \frac{20}{5} = 4$$

resuelto en el punto a)

c) Quiero encontrar el estimador de máxima verosimilitud de la probabilidad de que ocurran más de 3 desconexiones accidentales

Esto es

$$P(X \geq 3) = 1 - P(X \leq 2) \qquad \overline{X} \sim P(\hat{\lambda}), \quad \hat{\lambda} = 4$$

$$= 1 - 0.23810$$

App

$$= 0.7619$$

(4) $\sigma = 0.04$ cm (Distribución normal)

tamaño muestral $n = 25$

media muestral $\overline{X} = 0.51$

a) Queremos detectar cojinetes indeseables. Suponemos diámetro deseado

$$\mu_0 = 0.5$$

INTERVALOS DE CONFIANZA

- INTERVALO DE CONFIANZA PARA LA MEDIA DE UNA DISTRIBUCIÓN NORMAL VARIANZA CONOCIDA → MÉTODO DEL PIVOTE

EJEMPLO Un ingeniero civil analiza la resistencia a la compresión del concreto. La resistencia está distribuida aproximadamente de manera normal, con varianza $1000\ (psi)^2$. Al tomar una muestra aleatoria de 12 especímenes, se tiene que $\bar{X} = 3250$ psi

a) Construya un intervalo de confianza del 95% para la resistencia a la compresión promedio

b) Construya un intervalo de confianza del 99% para la resistencia a la compresión promedio (Compare ancho de intervalos a) y b)).

DATOS → v.a. X_i: "Resistencia a la compresión del concreto en un espécimen i"

muestra de $n = 12$ especímenes y $X \sim N(\mu, \sigma^2)$ para $i = 1, \ldots 12$

con $\sigma^2 = 1000\ (psi)^2$

a) Queremos I.C. para μ de nivel 95%. Por lo tanto $\alpha = 0.05$

Como $(X_1, X_2, \ldots X_{12})$ es una muestra aleatoria de tamaño $n = 12$ de una v.a. X, donde $X \sim N(\mu, \sigma^2)$, $\sigma^2 = 1000$ (conocido) el intervalo de confianza para μ de nivel $1-\alpha$ está dado por

$$*_1 \quad \left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \ \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

Longitud del intervalo

$$L = 2 z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

A MENOR LONGITUD, MAYOR PRECISIÓN EN LA ESTIMACIÓN

- $z_{\alpha/2} = z_{\frac{0.05}{2}} = z_{0.025} \underset{\text{tabla}}{=} 1,96$
- $\sigma = \sqrt{1000} = \sqrt{1000}$
- $\sqrt{n} = \sqrt{12}$
- $\bar{X} = 3250$

→ Reemplazamos todos los datos y queda

$$\left[3250 - 1,96 \cdot \frac{\sqrt{1000}}{\sqrt{12}}, \ 3250 + 1,96 \frac{\sqrt{1000}}{\sqrt{12}}\right] = [3232.10773, 3267.89227]$$

Intervalo de confianza del 95%

b) Queremos un intervalo de confianza del 99%

Por lo tanto $1 - \alpha = 0.99 \to \alpha = 0.01$ Tenemos

- $z_{\alpha/2} = z_{\frac{0.01}{2}} = z_{0,005} \underset{\text{tabla}}{=} 2,58$
- $\sigma = \sqrt{1000}$
- $\sqrt{n} = \sqrt{12}$
- $\bar{X} = 3250$

Reemplazamos

$$\left[3250 - 2.58 \frac{\sqrt{1000}}{\sqrt{12}}, 3250 + 2,58 \frac{\sqrt{1000}}{\sqrt{12}}\right] = [3226,44, 3273.55]$$

EJEMPLO ¿Qué tamaño n de muestra se necesita para que el intervalo tenga nivel de confianza 95% y longitud la mitad de la longitud del intervalo hallado en a)?

Longitud del intervalo en a) $\rightarrow L = 2 z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 2 \cdot 1{,}96 \frac{\sqrt{1000}}{\sqrt{12}} = 35.78454$

Quiero saber n tq el nivel de confianza se mantenga en 95% pero $L_2 = L/2 = 17.89227$. Planteamos

$$L_2 = 2 z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq 17.89227$$

$$2 \cdot 1{,}96 \frac{\sqrt{1000}}{\sqrt{n}} \leq 17.89227$$

$$\left(\frac{2 \cdot 1{,}96 \sqrt{1000}}{17.89227}\right)^2 \leq n \rightarrow n \geq 48$$

∴ Hay que tomar por lo menos 48 especímenes para que el intervalo tenga la longitud pedida.

Precisión del estimador $\rightarrow \frac{L}{2} = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

EJEMPLO Se estima que el tiempo de reacción a un estímulo de cierto dispositivo electrónico está distribuido normalmente con desviación estándar de 0.05 segundos. ¿Cuál es el número de mediciones temporales que deberá hacerse para que la confianza de que el error de la estimación de la esperanza no exceda 0.01 sea del 95%?

X = "tiempo de reacción a un estímulo"

$X \sim N(\mu, \sigma^2)$ con $\sigma = 0.05$ seg

Queremos calcular el número de mediciones temporales n tq

$$\frac{L}{2} = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq 0.01 \quad , \text{ con } \alpha = 0.05 \quad (\text{pues } 1-\alpha = 0.95 \rightarrow 95\%)$$

$$\Leftrightarrow z_{\frac{0.05}{2}} \, 0.05 \leq 0.01 \cdot \sqrt{n}$$

$$\left(\frac{1{,}96 \cdot 0.05}{0.01}\right)^2 \leq n \rightarrow n \geq 96{,}04$$

∴ Hay que tomar al menos 97 mediciones temporales

* En el caso de que la población de la que se extrae la muestra no sea normal pero $n \geq 30$, el nivel de confianza del intervalo es $\simeq 1-\alpha$.
Pero para muestras pequeñas tomadas de poblaciones que NO SON NORMALES, no se puede garantizar que el nivel de confianza sea $1-\alpha$ (usando la ECUACIÓN $*_1$) *

EJEMPLO: supongamos que X representa la duración de una pieza de equipo y que se probaron 100 de esas piezas dando una duración promedio de 501,2 horas. Se sabe que la desviación estándar poblacional es $\sigma = 4$ h. Se desea tener un intervalo del 95% de confianza para la esperanza poblacional $E(X) = \mu$.

Aunque no conocemos la distribución de X, como la muestra es de

$n = 100 \geq 30$ usamos

$$\left[\overline{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} , \overline{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

- $\overline{X} = 501,2$
- $\sigma = 4$
- $n = 100$
- $z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = 1,96$ ← tabla

$$\left[501,2 - 1,96 \cdot \frac{4}{\sqrt{100}} , 501,2 + 1,96 \frac{4}{\sqrt{100}}\right]$$

$[500,416, 501.984]$

- INTERVALO DE CONFIANZA PARA LA MEDIA DE UNA DISTRIBUCIÓN NORMAL, VARIANZA DESCONOCIDA

↳ STUDENT CON PARÁMETRO $n-1$

Si $(X_1, X_2, \dots, X_n)$ es una muestra aleatoria tamaño n de una v.a. X, donde $X \sim N(\mu, \sigma^2)$, σ^2 desconocido, un intervalo de confianza para μ de nivel $1-\alpha$ es

$$\left[\overline{X} - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} , \overline{X} + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right]$$

EJEMPLO: Se hicieron 10 mediciones sobre la resistencia de cierto tipo de alambre que dieron valores $x_1, x_2, \dots, x_{10}$ tq $\bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 10.48$ ohms y $S = \sqrt{\frac{1}{9}\sum_{i=1}^{10}(x_i - \bar{x})^2} = 1.36$ ohms. Supenga $X \sim N(\mu, \sigma^2)$. Se desea obtener un intervalo de confianza para la esperanza poblacional μ al 90%

Tenemos μ al 90% → $1 - \alpha = 0,9$ → $\alpha = 0.1$

$$\frac{\alpha}{2} = \frac{\frac{1}{10}}{\frac{2}{1}} = \frac{1}{20}$$

Datos
- $\overline{X} = 10.48$
- $t_{\frac{\alpha}{2}, n-1} = t_{0.05, 9} = 1,833$ ← tabla
- $S = 1,36$
- $\sqrt{n} = \sqrt{10}$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Luego

$$\left[\overline{X} - t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}} , \overline{X} + t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}}\right] = \left[10,48 - 1,833 \cdot \frac{1,36}{\sqrt{10}} , 10,48 + 1,833 \frac{1,36}{\sqrt{10}}\right]$$

$$= [9,69 , 11,27]$$

• INTERVALO DE CONFIANZA PARA LA DIFERENCIA DE DOS MEDIAS. VARIANZAS CONOCIDAS

Si X_1 y X_2 son dos variables aleatorias INDEPENDIENTES normalmente distribuidas $X_1 \sim N(\mu_1, \sigma_1^2)$
$X_2 \sim N(\mu_2, \sigma_2^2)$
y suponemos que las varianzas σ_1^2 σ_2^2 son conocidas. Un intervalo de confianza para la diferencia $\mu_1 - \mu_2$ de nivel $1-\alpha$ es

$$\circledast_1 \left[\bar{X}_1 - \bar{X}_2 - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \, , \, \bar{X}_1 - \bar{X}_2 + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right]$$

EJEMPLO se utilizan dos máquinas para llenar botellas de plástico con detergente para máquinas lavaplatos. Se sabe que las desviaciones estándar de volumen de llenado son $\sigma_1 = 0,1$ onzas de líquido y $\sigma_2 = 0,15$ onzas de líquido. Se toman dos muestras aleatorias $n_1 = 12$ botellas y $n_2 = 10$ botellas. Los volúmenes promedio de llenado son $\bar{X}_1 = 30,87$ onzas de líquido y $\bar{X}_2 = 30,68$ onzas de líquido.
Asumiendo que ambas muestras vienen de distribuciones normales, construya un intervalo de confianza de nivel 90% para la diferencia entre las medias de volumen de llenado.

Datos
- $\sigma_1 = 0,1$
- $\sigma_2 = 0,15$
- $n_1 = 12$
- $n_2 = 10$
- $\bar{X}_1 = 30,87$
- $\bar{X}_2 = 30,68$

$\bar{X}_1 - \bar{X}_2 = 0,19$

Quiero IC al 90% $\to 1-\alpha = 0,9 \to \alpha = 0,1$. Sabiendo que $X \sim N(\mu, \sigma_1^2)$ y $X_2 \sim N(\mu_2, \sigma_2^2)$ usamos

$$\left[\bar{X}_1 - \bar{X}_2 - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \, , \, \bar{X}_1 - \bar{X}_2 + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right] \quad \left(z_{\frac{\alpha}{2}} = z_{\frac{0,1}{2}} = z_{0,05} \underset{\text{tabla}}{=} 1,65 \right.$$

$$\left[0,19 - 1,65 \cdot \sqrt{\frac{(0,1)^2}{12} + \frac{(0,15)^2}{10}} \, , \, 0,19 + 1.65 \cdot \sqrt{\frac{(0,1)^2}{12} + \frac{(0,15)^2}{10}} \right]$$

$[0,0984 \, , \, 0.2816]$

- Si conozco σ_1, σ_2 y $n_1 = n_2$ puede determinarse el tamaño requerido de la muestra de manera tal que la longitud del intervalo sea menor que ℓ

$$L = 2 z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}} \leq \ell$$

- Si σ_1 y σ_2 son desconocidos y $n_1, n_2 \geq 30$, puedo reemplazar $\sigma_x \to S_x$ y construir el intervalo para $\mu_1 - \mu_2$ usando $\circledast_1$, su nivel es $\simeq 1-\alpha$
- Si las muestras no son normales pero $n_x \geq 30$ el nivel de confianza del intervalo $\circledast_1$ es $\simeq 1-\alpha$

EJEMPLO De una muestra de 150 lámparas del fabricante A se obtuvo una vida media de 1400hs y una desviación estándar típica de 120 hs. Mientras que de una muestra de 100 lámparas del fabricante B se obtuvo una vida media de 1200hs y una desviación típica de 80hs. Hallar los límites de confianza del 95% para la diferencia de los vidas medios de las poblaciones A y B.

X_1 = " duración (horas) de una lámpara A"

X_2 = " duración (horas) de una lámpara B"

$n_1 = 150$ $\qquad \bar{X}_1 = 1400$ $\qquad \sigma_1 = 120$ $\qquad \bar{X}_1 - \bar{X}_2 = 200$

$n_2 = 100$ $\qquad \bar{X}_2 = 1200$ $\qquad \sigma_2 = 80$

quiero hallar los límites de confianza del 95% para $\mu_1 - \mu_2$.

Aunque no se las distribuciones, como $n_1 = 150 \geq 30$ y $n_2 = 100 \geq 30$, puedo usar ⊛.

Además $95\% \to 1-\alpha = 0{,}95 \to \alpha = 0{,}05$

$$z_{\frac{\alpha}{2}} = z_{\frac{0{,}05}{2}} = z_{0{,}025} = 1{,}96$$

↑ tabla

luego

$$\left[\bar{X}_1 - \bar{X}_2 - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\ ,\ \bar{X}_1 - \bar{X}_2 + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

$$\left[200 - 1{,}96\sqrt{\frac{120^2}{150} + \frac{80^2}{100}}\ ,\ 200 + 1{,}96\sqrt{\frac{120^2}{150} + \frac{80^2}{100}}\right]$$

$[175{,}2077 ,\ 224{,}7923]$

• INTERVALO DE CONFIANZA PARA LA DIFERENCIA DE DOS MEDIAS, VARIANZAS DESCONOCIDAS

Si X_1 y X_2 son dos variables aleatorias independientes normalmente distribuidas $X_1 \sim N(\mu_1, \sigma^2)$, $X_2 \sim N(\mu_2, \sigma^2)$ y las varianzas σ^2 son desconocidas e IGUALES

Un intervalo de confianza para la diferencia $\mu_1 - \mu_2$ de nivel $1-\alpha$ es

$$\left[\bar{X}_1 - \bar{X}_2 - t_{\frac{\alpha}{2}, n_1+n_2-2}\, S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\ ,\ \bar{X}_1 - \bar{X}_2 + t_{\frac{\alpha}{2}, n_1+n_2-2}\, S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]$$

EJEMPLO se piensa que la concentración del ingrediente activo de un detergente es afectada por el tipo de catalizador utilizado en el proceso de fabricación. Se realizan 10 observaciones con cada catalizador y se obtienen los datos

Catalizador 1 = 57.9, 66.2, 65.4, 65.4, 65.2, 62.6, 67.6, 63.7, 67.2, 71.0
Catalizador 2 = 66.4, 71.7, 70.3, 69.3, 64.8, 69.6, 68.6, 69.4, 65.3, 68.8

a) Encuentre un intervalo de confianza del 95% para la diferencia entre los medios de las concentraciones activos para los dos catalizadores. Asumir que ambas muestras fueron extraídos de poblaciones normales con varianzas iguales

Sean las v.a

X_1 = "concentración del ingrediente activo con catalizador 1".

X_2 = "concentración del ingrediente activo con catalizador 2".

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

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

$\sigma_1 = \sigma_2 = \sigma$

$\overline{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i = 65.22$ $\quad n_1 = n_2 = 10$

* Usamos el estadístico usual para estimar μ

$\overline{X}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} x_i = 68.42$ $\quad \overline{X}_1 - \overline{X}_2 = -3,2$

$$S_1 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_i-\overline{x}_1)^2 = \frac{1}{9}\sum_{i=1}^{10}(x_i-65.22)^2 = 3.444$$

$$S_2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2}(x_i-\overline{x}_2)^2 = \frac{1}{9}\sum_{i=1}^{10}(x_i-68.42)^2 = 2.224$$

Calculamos

$$S_p^2 = \frac{(9 \cdot 3,444^2) + (9 \cdot 2.224^2)}{18} = 8.4036 \rightarrow S_p = \sqrt{8.4036} = 2.899$$

$95\% \rightarrow 1-\alpha = 0,95 \rightarrow \alpha = 0,05$ $\quad t_{\frac{\alpha}{2}, n_1+n_2-2} = t_{0.025,18} = 2,101$ (tabla)

Luego el intervalo de confianza será

$$\left[\overline{x}_1 - \overline{x}_2 - t_{\frac{\alpha}{2}, n_1+n_2-2}\, S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\ ,\ \overline{x}_1 - \overline{x}_2 + t_{\frac{\alpha}{2}, n_1+n_2-2}\, S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]$$

$$\left[-3,2 - 2,101 \cdot 2,899\sqrt{\frac{1}{5}}\ ,\ -3,2 + 2,101 \cdot 2,899\sqrt{\frac{1}{5}}\right]$$

$[-5,9239\ ,\ -0,4761]$

TEST DE HIPOTESIS

- PRUEBA DE HIPÓTESIS SOBRE LA MEDIA, VARIANZA CONOCIDA.

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 distribuida con $\sigma = 3$ y que $\bar{x} = 5.25$. ¿Indica esto de manera concluyente que el verdadero promedio de porcentaje difiere de 5.5? Utilice $\alpha = 0.01$

sea v.a. X = "porcentaje de SiO_2 en cierto tipo de cemento aluminoso"

$X \sim N(\mu, \sigma^2)$ con $\sigma = 3$

Podemos plantear $H_0 : \mu = 5.5$ contra $H_1 : \mu \neq 5.5$

Además $n = 16$ $\quad \mu_0 = 5.5$

$\bar{x} = 5.25$ (promedio muestral)

si $\alpha = 0.01 \rightarrow z_{\frac{\alpha}{2}} = z_{\frac{0.01}{2}} = z_{0.005} = 2.58$

↑ tabla

∴ la regla de decisión es:

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

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

↓ estadístico de prueba

El estadístico toma el valor $\left| \dfrac{5.25 - 5.5}{3/4} \right| = 0.33 \leq 2.58$

∴ como $|Z| \leq z_{\frac{\alpha}{2}}$, se acepta H_0

EJEMPLO Se sabe que la duración en horas de un foco de 75 watts tiene una distribución aproximadamente normal, con $\sigma = 25$ horas. Se toma una muestra aleatoria de 20 focos, la cual resulta tener una duración promedio de $\bar{x} = 1040$ hs. ¿Existe evidencia que apoye la afirmación de que la duración promedio del foco es mayor que 1000 horas? Utilice $\alpha = 0.05$

sea v.a. X = "duración en horas de un foco de 75 watts"

$X \sim N(\mu, \sigma^2)$ con $\sigma = 25$ $\quad \mu_0 = 1000$

Podemos plantear Hipótesis $H_0 : \mu = 1000$ contra $H_1 : \mu > 1000$

Además $n = 20$ $\quad \bar{x} = 1040$

Si $\alpha = 0.05 \rightarrow z_{\alpha} = z_{0.05} = 1.64$

$\therefore$ La regla de decisión es $\begin{cases} \text{rechazar } H_0 \text{ si } \dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} > z_\alpha \\ \text{aceptar } H_0 \text{ si } \dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \leq z_\alpha \end{cases}$

El estadístico toma el valor $z = \dfrac{1040-1000}{25/\sqrt{20}} = 7.15 > \overbrace{1.64}^{z_\alpha}$

$\therefore$ como $z > z_\alpha$ se rechaza H_0

p- valor = nivel de significancia más pequeño que conduce al rechazo de la hipótesis nula H_0

$\alpha <$ p- valor ✓
$\alpha >$ p- valor ✗

H_1 $\mu \neq \mu_0$

p- valor $= P(|Z| > |z_0|) = 1 - P(|Z| < |z_0|) = 1 - [\phi(|z_0|) - \phi(-|z_0|)] = 1 - [2\phi(|z_0|) - 1] = 2[1-\phi(|z_0|)]$

$H_1 : \mu > \mu_0$

p- valor $= P(Z > z_0) = 1 - P(Z \leq z_0) = 1 - \phi(z_0)$

$H_1 : \mu < \mu_0$

p- valor $= P(Z < z_0) = \phi(z_0)$

EJEMPLO (continuación 1) 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 $\beta(5.6) = 0.01$?

Probabilidad de detectar la desviación

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

Hipótesis alternativa bilateral, calculamos

$$\beta(5.6) = \phi\left(z_{\frac{\alpha}{2}} \frac{(5.6-\mu_0)\sqrt{n}}{\sigma}\right) - \phi\left(-z_{\frac{\alpha}{2}} - \frac{(5.6-\mu_0)\sqrt{n}}{\sigma}\right) = \phi(2.441) - \phi(-2.708)$$

$= 0{,}9893 \longrightarrow \pi(5.6) = 1 - \beta(5.6) = 1 - 0.9893 = 0.0107$

TEST HIPÓTESIS

EJEMPLO pp 179

sea X_i = "porcentaje de SiO_2 en la i-ésima muestra de cemento aluminoso", con $i = 1, 2, \ldots, 16$

<u>Datos</u> $X \sim N(\mu, \sigma^2)$ • $\mu = 5.5$ • $\sigma = 3$ • $\overline{X} = 5.25$ • $n = 16$

$\alpha = 0.01$

• Planteamos hipótesis $H_0: \mu = 5.5$ contra $H_1: \mu \neq 5.5$

Con los datos dados, la regla de decisión es

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

con $|Z| = \left|\dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}\right|$ y $z_{\frac{\alpha}{2}} = z_{0.005} \underset{\text{tabla}}{\overset{\uparrow}{=}} 2.575$

Luego

$$\left|\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}\right| = \left|\frac{5.25 - 5.5}{3/\sqrt{16}}\right| = |-1/3| = \underbrace{1/3}_{z_0} < 2.575$$

∴ como $z_0 = 1/3 < 2.575 = z_{\frac{\alpha}{2}}$, se acepta H_0

EJEMPLO pp 181

Sea X = "duración en horas de un foco de 75 watts"

① <u>Datos</u> • $X \sim N(\mu, \sigma^2)$ • $\sigma = 25$ • $n = 20$ • $\overline{X} = 1040$

$\alpha = 0.05$

② • Planteamos hipótesis $H_0: \mu = 1000$ contra $H_1: \mu > 1000$

③ Con los datos dados, la regla de decisión es:

$$\begin{cases} \text{rechazar } H_0 \text{ si } z_0 > z_\alpha \\ \text{aceptar } H_0 \text{ si } z_0 \leq z_\alpha \end{cases}$$

con $z_0 = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$ y $z_\alpha = z_{0.05} \underset{\text{tabla}}{\overset{\uparrow}{=}} 1.645$

④ Luego $\dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{1040 - 1000}{25/\sqrt{20}} = 7,15 > 1.645$

∴ Como $z_0 = 7,15 > 1,645 = z_\alpha$, se rechaza H_0

EJEMPLO pp 182

Siguiendo ejercicio de la pp 179 pero ahora con el criterio del p-valor

$\begin{cases} \text{si } \alpha < p\text{-valor, se acepta } H_0 \text{ con nivel de significancia } \alpha. \\ \text{si } \alpha \geq p\text{-valor se rechaza } H_0 \text{ con nivel de significancia } \alpha \end{cases}$

HIP H_0 $\mu = 5.5$ contra H_0 $\mu \neq 5.5$

$p\text{-valor} = P(|Z| > |z_0|) = 1 - P(|Z| \leq |z_0|) = 2(1 - \phi(|z_0|)) \underset{z_0 = 0,3}{=} (2(1-\phi(0,3)) \ominus$

$\ominus$ $2(1 - (0,6293)) = 0,7414 > 0,05 = \alpha$

tabla acumulada

Como el p-valor $> \alpha = 0,05$ se acepta H_0 con nivel de significancia α.

EJEMPLO pp 183

Siguiendo ejercicio de la pp 181 pero ahora con el criterio del p-valor

$\begin{cases} \text{si } \alpha < p\text{-valor, se acepta } H_0 \text{ con nivel de significancia } \alpha \\ \text{si } \alpha \geq p\text{-valor, se rechaza } H_0 \text{ con nivel de significancia } \alpha. \end{cases}$

HIP H_0 $\mu = 1000$ contra $H_1: \mu > 1000$

$p\text{-valor} = P(Z > z_0) = 1 - P(Z \leq z_0) = 1 - \phi(z_0) \underset{z_0 = 7.15}{=} 1 - \overbrace{\phi(7,15)}^{\simeq 1} \simeq 0 < 0.05$

Como el p-valor $< 0,05 = \alpha$, se rechaza H_0 con nivel de significancia α.

EJEMPLO pp 188

sea X = "Volumen de llenado para una lata de jugo"

Datos $\cdot\ n = 100 > 30$ $\quad \cdot\ \overline{X} = 11.98$ $\quad \cdot\ S = 0,19$ $\quad \cdot\ \mu = 12$

No sabemos la distribución pero como $n = 100 > 30$, por TCL se cumple

$Z \simeq N(0,1)$

Planteamos HIP H_0 $\mu = 12$ contra H_0 $\mu \neq 12$. El estadístico de prueba es

$Z = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$, si $H_0: \mu = 12$ es V $\rightarrow Z \simeq N(0,1)$

El estadístico toma el valor $z_0 = \dfrac{11.98 - 12}{0.19/\sqrt{100}} = -1,05$

Calculamos $p\text{-valor} = P(|Z| > |z_0|) = 2(1 - \phi(|z_0|))$ $2(1 - \phi(1,05)) = 2(1 - 0.8531) \ominus$

$\ominus$ $0,2938 > 0,05$

$\therefore$ Como $p\text{-valor} = 0,2938 > 0,05 = \alpha$, se acepta H_0 $\mu = 12$ con nivel de significancia α

CRITERIO DEL P-VALOR

- $p\text{-valor} > \alpha$, acepta H_0
- $p\text{-valor} < \alpha$, rechaza H_0

(con nivel de significancia α)

EJEMPLO pp 190

Sea X = "ph de una muestra de leche tomada al azar"

① Datos

- $n = 6 < 30$ • $\bar{X} = 6.68$ • $S = 0.2$ • $X \sim N(\mu, \sigma^2)$ • $\alpha = 0.05$

② Planteamos HIP $H_0: \mu = 7.0$ contra $H_1: \mu < 7.0$

③ El estadístico de prueba para estos casos es $T = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$ y toma el valor $t_0 = \dfrac{6.68 - 7}{0.2/\sqrt{6}} = -3{,}92$

④ La regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } t_0 < -t_{\alpha, n-1} \\ \text{aceptar } H_0 & \text{si } t_0 \geq -t_{\alpha, n-1} \end{cases}$$

Calculamos $-t_{\alpha, n-1} = -t_{0{,}05, 5} \underset{\text{tabla}}{=} 2{,}015$

Como $t_0 = -3{,}92 < 2{,}015 = t_{\alpha, n-1}$ se rechaza H_0

Criterio del p-valor. A partir de la HIP

$p\text{-valor} = P(T \leq -3{,}92)$

- Busco en TABLA fila con $V = n-1 = 5$ grados de libertad (si no lo encuentro, acoto)

$3{,}365 < 3{,}92 < 4{,}032$ y $P(T_5 > 3{,}365) = 0{,}01$, $P(T_5 > 4{,}032) = 0{,}005$

$\rightarrow 0{,}005 < p\text{-valor} < 0{,}01 < 0{,}05 = \alpha$

∴ Como $p\text{-valor} < \alpha = 0{,}05$, se rechaza H_0

CRITERIO p-valor
- p-valor > α, acepta H_0
- p-valor < α, rechaza H_0

(con nivel de sign α)

EJEMPLO pp 192

Sea $X_{1,2}$ = "Tiempo de secado de una pintura tipo aparovos."

Datos

- $n_1 = n_2 = n = 10 < 30$ • $\sigma_{12} = 8$ • $\bar{X}_1 = 121$ • $\bar{X}_2 = 112$ • suponemos distr normal
- $\alpha = 0{,}05$

Planteamos HIP $H_0: \mu_1 - \mu_2 = 0$ contra $H_1: \mu_1 - \mu_2 > 0$

El estadístico de prueba para estos casos es $Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - \overset{=0}{\Delta_0}}{\sqrt{\sigma_1^2 + \sigma_2^2}}$ y toma el valor $z_0 = \dfrac{(121-112)-0}{\sqrt{\dfrac{2(8)^2}{10}}} = 2,515$

La REGLA DE DECISIÓN es $\begin{cases} \text{se acepta } H_0 \text{ si } z_0 \leq z_\alpha \\ \text{se rechaza } H_0 \text{ si } z_0 > z_\alpha \end{cases}$

con $\alpha = 0,05 \rightarrow z_\alpha = z_{0,05} = 1,645 < 2,515 = z_0$

$\therefore$ Como $z_0 > z_\alpha$, se rechaza H_0 al nivel $\alpha = 0,05$

EJEMPLO pp 194

Sea X_{12} = "Flujo de corriente en diseño$_{12}$"

DATOS . distribuciones normales, varianzas desconocidas (como no me dice que son =, los asumo ≠)

. $n_1 = 15$. $\bar{X}_1 = 24.2$. $S_1^2 = 10$ $\alpha = 0,1$

. $n_2 = 10$. $\bar{X}_2 = 23.9$. $S_2^2 = 20$

Planteamos HIP $H_0: \mu_1 - \mu_2 = 0$ contra $H_1: \mu_1 - \mu_2 \neq 0$

El ESTADÍSTICO a utilizar en estos casos es

$$|T^*| = \left| \frac{(\bar{X}_1 - \bar{X}_2) - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \right| \text{ que toma el valor } |t_0| = \left| \frac{(24,2-23,9)-0}{\sqrt{\frac{10}{15} + \frac{20}{10}}} \right| = 0,1837$$

La REGLA DE DECISIÓN es $\begin{cases} \text{se acepta } H_0 \text{ si } |t_0| \leq t_{\frac{\alpha}{2},v} \\ \text{se rechaza } H_0 \text{ si } |t_0| > t_{\frac{\alpha}{2},v} \end{cases}$

$$\text{con } v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}} = \frac{\left(\frac{10}{15} + \frac{20}{10}\right)^2}{\frac{(10/15)^2}{14} + \frac{(20/10)^2}{9}} = 14,93 \simeq 15$$

con $\frac{\alpha}{2} = \frac{0,1}{2} = 0,05$

Luego $t_{\frac{\alpha}{2},v} = t_{0,05,15} \underset{\text{tabla}}{=} 1,753 > 0,1837 = |t_0|$

$\therefore$ Como $|t_0| < t_{\frac{\alpha}{2},v}$, se acepta H_0 (no hay evidencia de que los dos medios de flujo sean ≠)

EJEMPLO pp 196

Sea $X_{1,2}$ = "Nivel de hierro en sangre de niño sano (enfermo)"

DATOS dos muestras

- $n_1 = 9$. $\bar{X}_1 = 18,9$. $S_1^2 = (5,9)^2$. $\sigma_1^2 = \sigma_2^2$ y desconocidos
- $n_2 = 13$. $\bar{X}_2 = 11,9$. $S_2^2 = (6,3)^2$. sup $X_1 \sim N(\mu_1, \sigma_1^2)$ y $X_2 \sim N(\mu_2, \sigma_2^2)$

Planteamos HIP $H_0: \mu_1 - \mu_2 = 0$ contra $H_1: \mu_1 - \mu_2 \neq 0$

El ESTADÍSTICO a utilizar en estos casos es $|T| = \left| \dfrac{(\bar{X}_1 - \bar{X}_2) - \Delta_0}{S_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \right|$ que toma

el valor $|t_0| = \left| \dfrac{(18,9 - 11,9) - 0}{6,14 \sqrt{\dfrac{1}{9} + \dfrac{1}{13}}} \right|$. Calculo $S_p = \sqrt{\dfrac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}}$

$|t_0| = 2,629$

$S_p = 6,14$

El CRITERIO DE DECISIÓN es

$$\begin{cases} \text{Rechazar } H_0 \text{ si } |t_0| > t_{\alpha/2,\, n_1+n_2-2} \\ \text{Aceptar } H_0 \text{ si } |t_0| \leq t_{\alpha/2,\, n_1+n_2-2} \end{cases}$$

$$\left. \begin{array}{l} \alpha = 0,05 \rightarrow \dfrac{\alpha}{2} = 0,025 \\ n_1 + n_2 - 2 = 20 \end{array} \right\} t_{0,025,\,20} \underset{\text{tabla}}{=} 2,086$$

Luego, como $|t_0| = 2,629 > 2,086 = t_{\frac{\alpha}{2},\, n_1+n_2-2}$ se rechaza H_0

Si calculamos el p-valor

p-valor $= 2(1 - P(T < 2,629))$

Acotamos por la tabla $t_{0,01,\,20} = 2,528$ y $t_{0,005,\,20} = 2,845$

$\rightarrow 0,005 <$ p-valor $< 0,01 < 0,05$

$\therefore$ p-valor $< 0,05$