

Probabilidades y Estadística (C) – 2019

Intervalos de Confianza – Parte 1 (31/10)

Intervalos de confianza: All of Statistics, Wasserman

- Interval that contains an unknown quantity with a given frequency-

All of Statistics, Wasserman

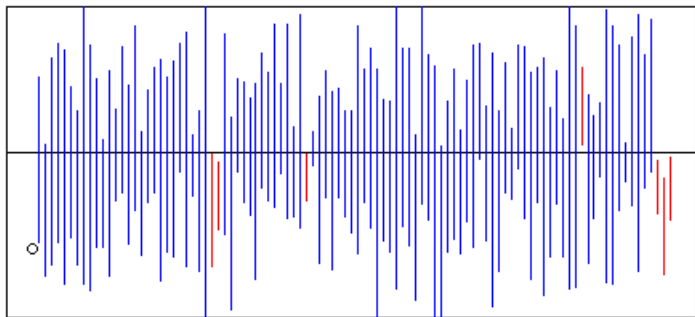
-Intervalo que contiene una cantidad desconocida (parámetro de interés) con cierta frecuencia (nivel) -

Intervalos de confianza: definición

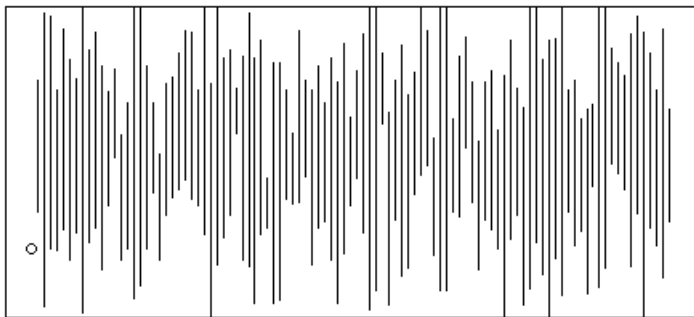
- Diremos que $(a(X_1, \dots, X_n), b(X_1, \dots, X_n))$ es un intervalo de confianza de nivel $1 - \alpha$ para el parámetro θ sii

$$P(a(X_1, \dots, X_n) < \theta < b(X_1, \dots, X_n)) = 1 - \alpha .$$

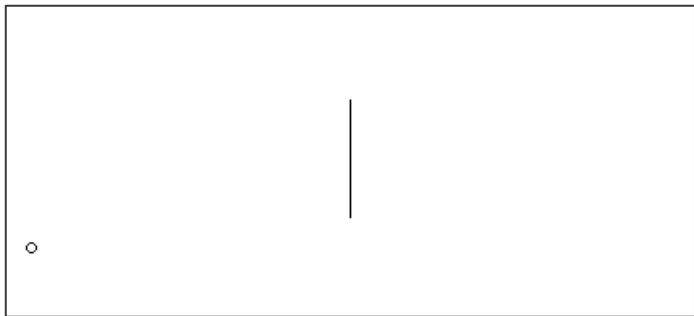
Muchos intervalos y la verdad



Muchos intervalos



Mi intervalo y yo, buena suerte! (confianza)



Ejemplos: Mundo normal

$$X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$$

Mundo normal - σ_0^2 conocido

- Buscamos intervalo de confianza para μ
- Pivot: Cuenta con la muestra y el parámetro de interés, cuya distribución es conocida.

$$\frac{\bar{X}_n - \mu}{\sqrt{\sigma_0^2/n}}$$

- Distribución del *pivot* (pi qué? Pi-vot):

$$\frac{\bar{X}_n - \mu}{\sqrt{\sigma_0^2/n}} \sim Z, Z \sim \mathcal{N}(0, 1)$$

- Sea z_β con $P(Z > z_\beta) = \beta$:

$$z_\beta = \phi^{-1}(1 - \beta)$$

- tenemos que

$$P\left(-z_{\alpha/2} < \frac{\bar{X}_n - \mu}{\sqrt{\sigma_0^2/n}} < z_{\alpha/2}\right) = 1 - \alpha$$

Mundo normal - σ_0^2 conocido

- Equivalentemente, tenemos que

$$P\left(\bar{X}_n - \frac{\sigma_0 z_{\alpha/2}}{\sqrt{n}} < \mu < \bar{X}_n + \frac{\sigma_0 z_{\alpha/2}}{\sqrt{n}}\right) = 1 - \alpha$$

y por consiguiente,

$$\left(\bar{X}_n - \frac{\sigma_0 z_{\alpha/2}}{\sqrt{n}}, \bar{X}_n + \frac{\sigma_0 z_{\alpha/2}}{\sqrt{n}}\right)$$

es un intervalo de confianza de nivel $1 - \alpha$ para μ bajo el modelo normal, con varianza σ_0^2 conocida.

Algunas observaciones

$$\left(\bar{X}_n - \frac{\sigma_0 z_{\alpha/2}}{\sqrt{n}} , \bar{X}_n + \frac{\sigma_0 z_{\alpha/2}}{\sqrt{n}} \right)$$

es un intervalo de confianza de nivel $1 - \alpha$ para μ bajo el modelo normal, con varianza σ_0^2 conocida.

- ¿Qué longitud tiene el intervalo?
- ¿Qué ocurre con la longitud del intervalo a medida que n aumenta?
- ¿Qué ocurre con la longitud del intervalo cuando el nivel $1 - \alpha$ aumenta?
- ¿Qué ocurre con la longitud del intervalo si aumenta la varianza σ_0^2 ?

Distribución chi cuadrado: definición

- Sean Z_1, \dots, Z_n i.i.d., $Z_i \sim \mathcal{N}(0, 1)$. Tenemos entonces que

$$Z_i^2 \sim \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) \quad (\text{revisa cambio de variable})$$

$$\text{y por consiguiente} \quad \sum_{i=1}^n Z_i^2 \sim \Gamma\left(\frac{n}{2}, \frac{1}{2}\right).$$

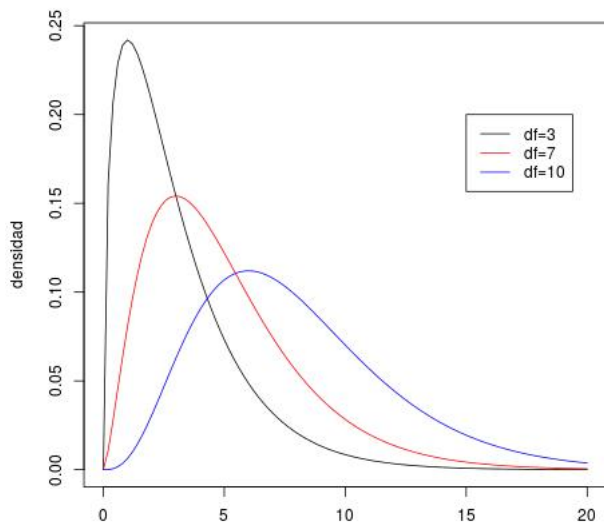
Definición: Llamamos chi cuadrado con n - grados de libertad a la distribución $\Gamma(\frac{n}{2}, \frac{1}{2})$.

- Notación: χ_n^2 : chi cuadrado con n grados de libertad.
- Es decir, utilizamos indistintamente las siguiente notaciones:

$$Z_1^2 + Z_2^2 + \dots + Z_n^2 \sim \chi_n^2, \quad Z_1^2 + Z_2^2 + \dots + Z_n^2 \sim \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$$

- Notar que $\Gamma\left(n, \frac{1}{2}\right) = \chi_{2n}^2$

Densidad chi cuadrado



Caso particular:

- Sean X_1, \dots, X_n i.i.d., $X_n \sim \mathcal{N}(\mu, \sigma^2)$, entonces

$$\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi_n^2$$

- mientras que al reemplazar μ por \bar{X}_n tenemos que (sin demostrar)

$$\frac{\sum_{i=1}^n (X_i - \bar{X}_n)^2}{\sigma^2} \sim \chi_{n-1}^2, \quad \text{es decir } \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

recordemos que
$$S^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

Intervalos de confianza para σ^2 con μ desconocido

- Buscamos intervalo de confianza para σ^2
- Pivot: Cuenta con la muestra y el parámetro de interés, cuya distribución es conocida.

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

- Sea $\chi_{n-1,\beta}$ con $P(X > \chi_{n-1,\beta}) = \beta$ cuando $X \sim \chi_{n-1}$, entonces

$$P\left(\chi_{n-1,1-\alpha/2} < \frac{(n-1)S^2}{\sigma^2} < \chi_{n-1,\alpha/2}\right) = 1 - \alpha$$

y por consiguiente,

$$\left(\frac{(n-1)S^2}{\chi_{n-1,\alpha/2}}, \frac{(n-1)S^2}{\chi_{n-1,1-\alpha/2}}\right)$$

es un intervalo de confianza de nivel $1 - \alpha$ para σ^2 bajo el modelo normal, con μ DESconocida.

Mundo normal - σ^2 desconocido - t de Student

- Buscamos intervalo de confianza para μ con σ^2 desconocido.
- Pivot:

$$\frac{(\bar{X}_n - \mu)}{\sqrt{S^2/n}} \sim t_{n-1} \text{ de Student con } n - 1 \text{ grados de libertad}$$

t_{n-1} : Distribución t - de Student con $n - 1$ grados de libertad

- Hay tabla para t -Student. Sea $t_{n-1,\beta}$ con $P(X > t_{n-1,\beta}) = \beta$ cuando $X \sim t_{n-1}$, entonces

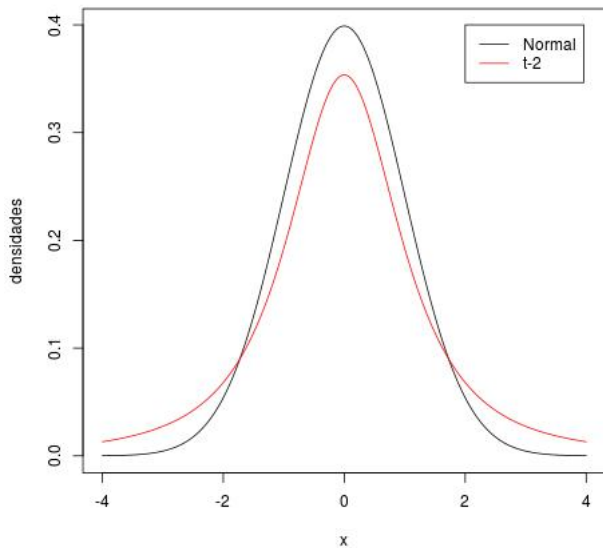
$$P\left(-t_{n-1,\alpha/2} < \frac{(\bar{X}_n - \mu)}{\sqrt{S^2/n}} < t_{n-1,\alpha/2}\right) = 1 - \alpha$$

y por consiguiente,

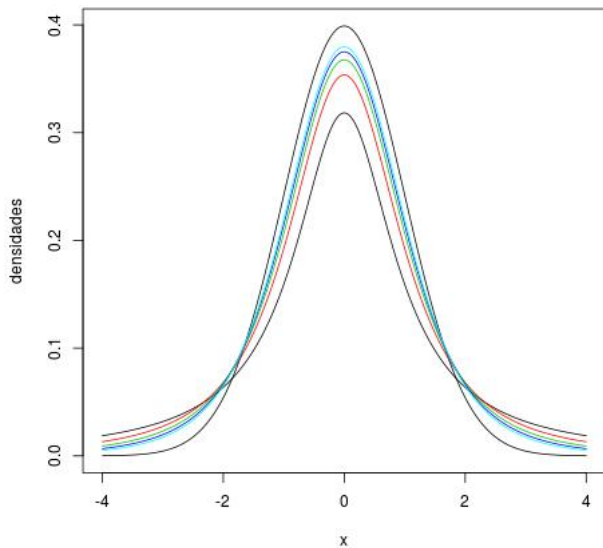
$$\left(\bar{X}_n - \frac{\sqrt{S^2} t_{n-1,\alpha/2}}{\sqrt{n}}, \quad \bar{X}_n + \frac{\sqrt{S^2} t_{n-1,\alpha/2}}{\sqrt{n}}\right)$$

es un intervalo de confianza de nivel $1 - \alpha$ para μ bajo el modelo normal, con varianza σ^2 DESconocida.

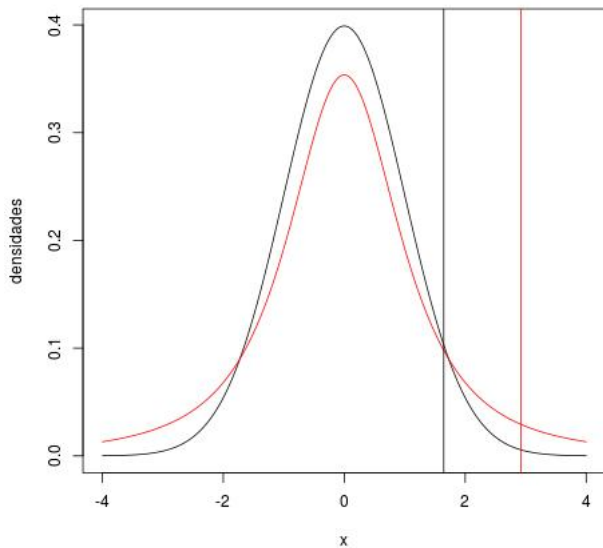
Normal vs. t2



Normal y student



Normal vs. t2: mirando percentiles...



Construcción de intervalos de confianza - PIVOT

- Buscamos $(a(X_1, \dots, X_n), b(X_1, \dots, X_n))$ de forma tal que

$$P(a(X_1, \dots, X_n) < \theta < b(X_1, \dots, X_n)) = 1 - \alpha .$$

- Construcción mediante la utilización de un **Pivot**: función de la muestra y θ cuya distribución es conocida.

$$H(X_1, \dots, X_n, \theta) \sim \text{Distribución } \textit{tabulada}$$

- Encerramos al pivot entre los percentiles de su distribución y despejamos θ .
- Identifique el pivot para cada uno de los intervalos de confianza construídos.

Mundo Exponencial

- $X_i \sim \mathcal{E}(\lambda)$
- $2\lambda X_i \sim \mathcal{E}(1/2) = \Gamma(1, 1/2)$?
- Armamos el *Pivot*

$$2\lambda \sum_{i=1}^n X_i \sim \Gamma\left(n, \frac{1}{2}\right) = \chi_{2n}^2$$

- y ahora?