Repaso de algunas propiedades

a) $E(\overline{X}) = \mu$

b) $V(\overline{X}) = \dfrac{\sigma^2}{n}$

c) Si $X \sim N(\mu, \sigma^2)$ entonces, cualquiera sea n, $\overline{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$, es decir: $\dfrac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}} \sim N(0,1)$.

d) Si n es grande entonces, cualquiera sea la distribución de X, la v.a. $\dfrac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$ tiene aproximadamente distribución $N(0,1)$.

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i).$$

Pero, tratándose de las componentes de una muestra aleatoria es:

$E(X_i) = E(X) = \mu \quad \forall i = 1,2,\ldots,n$. Luego:

$$E(\overline{X}) = \frac{1}{n} n\mu = \mu.$$

$$V(\overline{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}V\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i),$$

$V(X_i) = V(X) = \sigma^2 \quad \forall i = 1,2,\ldots,n$, tenemos

$$V(\overline{X}) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

Intervalos de confianza para la media poblacional μ

i) Intervalo de confianza (bilateral) para la esperanza poblacional μ de una v.a. $X \sim N(\mu, \sigma^2)$ supuesto σ^2 conocido para un tamaño n de la muestra arbitario (grande o pequeño).

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

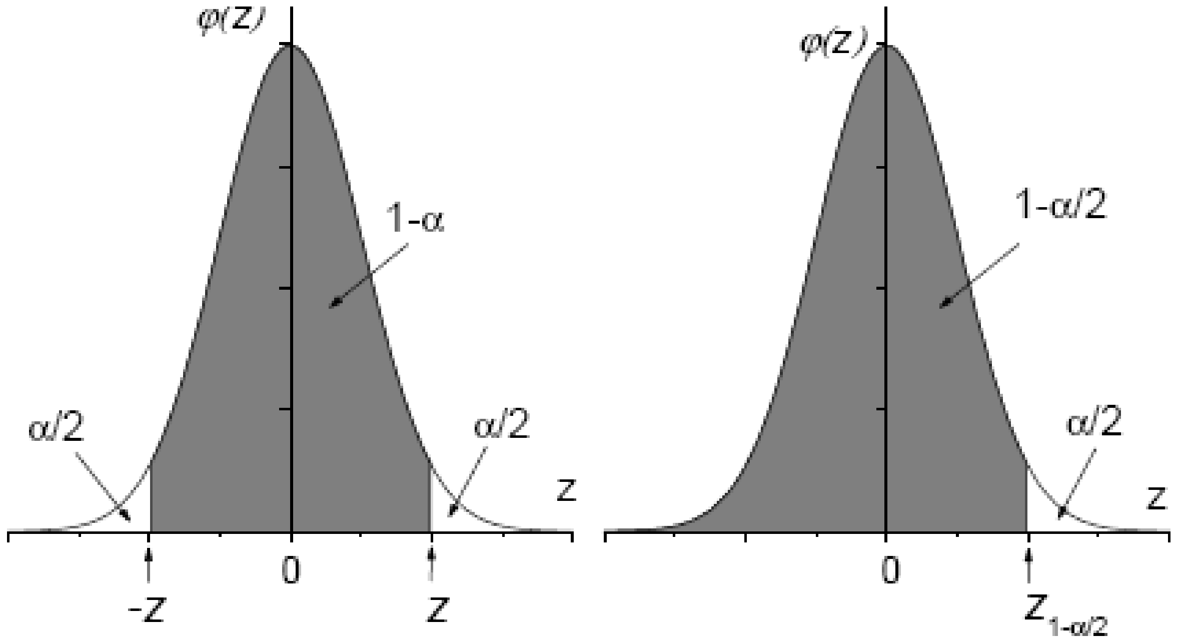
Para construir un intervalo de confianza al nivel de confianza $1\text{-}\alpha$ partiendo del pivote Z, comenzamos por plantear la ecuación probabilística

$$P(-z \leq Z \leq z) = 1\text{-}\alpha,$$

donde la incógnita es el número real z.

$$P\left(-z \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq z\right) = P\left(-z\frac{\sigma}{\sqrt{n}} \leq \overline{X}-\mu \leq z\frac{\sigma}{\sqrt{n}}\right) = P\left(-\overline{X}-z\frac{\sigma}{\sqrt{n}} \leq -\mu \leq -\overline{X}+z\frac{\sigma}{\sqrt{n}}\right) = 1\text{-}\alpha$$

$$P\left(\overline{X}-z\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X}+z\frac{\sigma}{\sqrt{n}}\right) = 1\text{-}\alpha$$



Al valor de z que verifica esta ecuación se lo suele indicar $z_{1-\frac{\alpha}{2}}$. En consecuencia, el intervalo de confianza bilateral al nivel de significación $1\text{-}\alpha$ queda:

$$\left[\hat{A}_1, \hat{A}_2\right] = \left[\overline{X} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

- $longitud\left(IC(\mu)_{(1-\alpha)}\right) = 2Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$
- $|\overline{X} - \mu| \leq \varepsilon = \ Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ $\Rightarrow$ Error en la estimación

Ejemplo

2. Los siguientes valores corresponden a 10 mediciones del valor de colesterol en un suero, realizadas con un método que tiene una desviación típica de $8,5$.

$$124 \quad 136 \quad 129 \quad 132 \quad 108 \quad 118 \quad 121 \quad 114 \quad 115 \quad 122$$

Se supone que cada medición es una variable aleatoria con distribución normal, cuya media es el verdadero valor. Construya un intervalo de confianza de nivel $0,95$ para el valor de colesterol analizado.

$X_i = nivel\ del\ colesterol\ en\ sangre\ del\ individuo\ i - ésimo$

$\sigma = 8{,}5$ $\bar{X} = 121{,}9$ $Z_{\alpha/2} = 1{,}96$

Se busca en tabla N(0,1)

Intervalo de confianza:

$$[\bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \ ; \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}]$$

$$[121{,}9 - 1{,}96 \frac{8{,}5}{\sqrt{10}} \ ; 121{,}9 + 1{,}96 \frac{8{,}5}{\sqrt{10}}]$$

ii) Intervalo de confianza (bilateral) para la esperanza poblacional μ de una v.a. X con distribución cualquiera supuesto conocida la varianza en el caso de muestra grandes.

$$\left[\hat{A}_1, \hat{A}_2\right] = \left[\overline{X} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right] \quad \text{con} \quad \Phi\left(z_{1-\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2}$$

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \approx N(0,1) \; Por \; TCL$$

$$\Downarrow$$

$$n \geq 30$$

iii) Intervalo de confianza (bilateral) para la esperanza poblacional μ de una v.a. $X \sim N(\mu,\sigma^2)$ supuesto σ^2 desconocido.

$$T = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$$

tiene distribución t de Student con n-1 grados de libertad: $T \sim t_{n-1}$

$$P(-t \leq T \leq t) = 1\text{-}\alpha,$$

$$P\left(-t \leq \frac{\overline{X}-\mu}{S/\sqrt{n}} \leq t\right) = P\left(-t\frac{S}{\sqrt{n}} \leq \overline{X}-\mu \leq t\frac{S}{\sqrt{n}}\right) = P\left(-\overline{X}-t\frac{S}{\sqrt{n}} \leq -\mu \leq -\overline{X}+t\frac{S}{\sqrt{n}}\right) = 1\text{-}\alpha$$

$$P\left(\overline{X}-t\frac{S}{\sqrt{n}} \leq \mu \leq \overline{X}+t\frac{S}{\sqrt{n}}\right) = 1\text{-}\alpha$$

Al valor de t que verifica esta ecuación se lo suele indicar $t_{n-1;1-\frac{\alpha}{2}}$.

$o\ también\ se\ anota\ t_{\frac{\alpha}{2}}(n-1)$ $\quad IC_{1-\alpha}(\mu) = [\overline{X} - t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}};\ \overline{X} + t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}]$

$$longitud\left(IC(\mu)_{(1-\alpha)}\right) = 2t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}$$

$$|\overline{X}-\mu| \leq \varepsilon = t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}} \Rightarrow ERROR\ EN\ LA\ ESTIMACIÓN$$

Distribución t-Student

Teorema Dada un muestra aleatoria $(X_1, X_2, \ldots, X_n)$ con $X_i \sim N(\mu, \sigma^2) \quad \forall i = 1,2,\ldots,n$, entonces la variable aleatoria $\dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$ (t de Student con n-1 grados de libertad)

Definición Sea una variable aleatoria continua X que puede tomar todos los valores reales y cuya función de densidad de probabilidad está dada por

$$f(x) = \frac{\Gamma(n+1)}{\Gamma(n/2)\sqrt{\pi n}} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2} \qquad -\infty < x < \infty, \quad n \in N.$$

Diremos que X tiene una distribución t de Student con n grados de libertad.

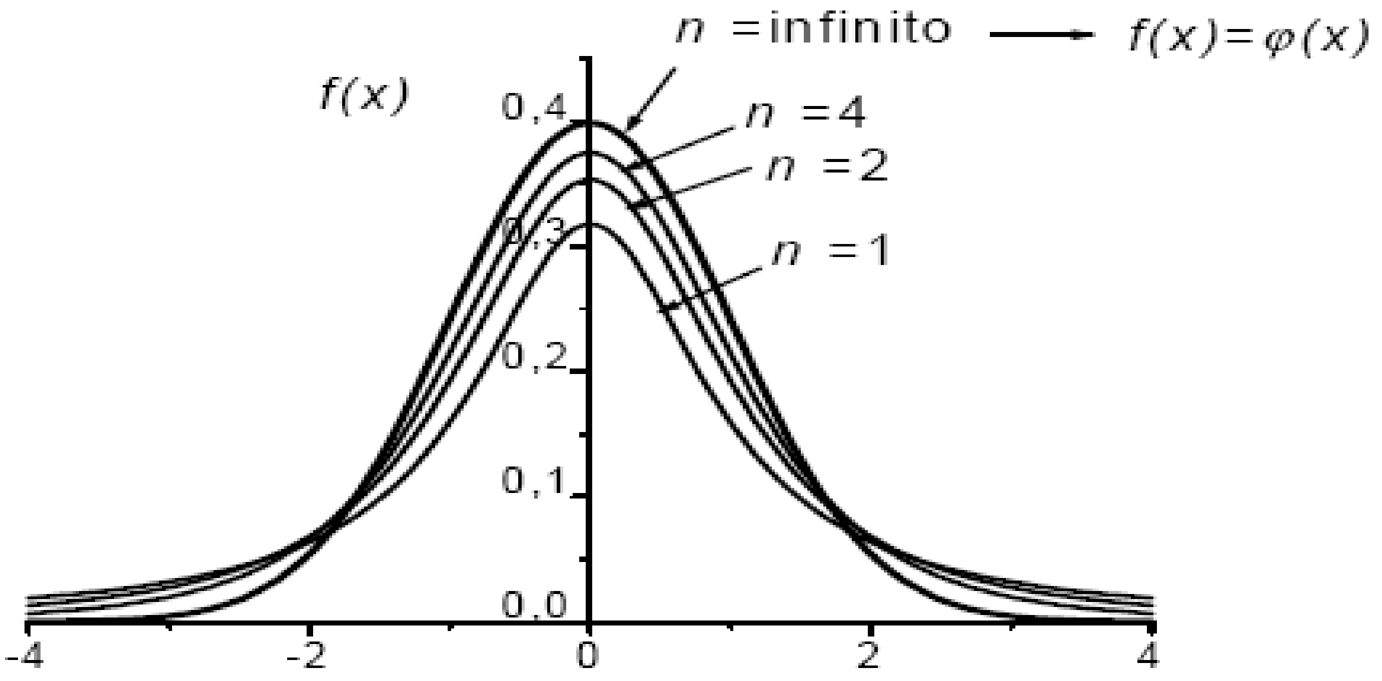
Notación. Si X es una variable aleatoria que tiene una distribución t de student con parámetro n, es decir con n grados de libertad, lo indicaremos $X \sim t_n$.

Propiedades de la distribución t de Student

La distribución es simétrica con respecto al eje $x = 0$.
Es muy semejante a la normal standarizada. De hecho se verifica:

$$\lim_{n\to\infty} f(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \text{ es decir } \quad t_n \underset{n\to\infty}{\longrightarrow} N(0,1).$$

En la figura hemos graficado la distribución t para distintos valores del parámetro n.



Ejemplo

La estatura de los individuos de una cierta región es una variable aleatoria con distribución de probabilidad normal. Se extrae una muestra aleatoria simple con los siguientes resultados

170, 164, 165, 169, 165, 168, 165, 162 , 166

Se desea calcular un intervalo de confianza del 99% para la estatura media de la población.

$$\bar{X} = 166$$

$$n = 9$$

$$S = 2{,}549$$

$$t_{\frac{\alpha}{2}}(n-1) = t_{0{,}005}(8) = 3{,}3554$$

Se busca en tabla T-Student

$$[\bar{X} - t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}} ; \bar{X} + t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}]$$

$$[166 - 3{,}3554\frac{2{,}549}{\sqrt{9}} ; 166 + 3{,}3554\frac{2{,}549}{\sqrt{9}}]$$

$$X_i = estatura\ del\ individuo\ i - \acute{e}simo$$

iv) Intervalo de confianza (bilateral) para la esperanza poblacional μ de una v.a. X con distribución cualquiera supuesto σ^2 desconocido y tamaño de la muestra grande.

Puesto que para n grande $S \to \sigma$, y teniendo en cuenta el *teorema del límite central*, vemos que

$$Z = \frac{\overline{X} - \mu}{S / \sqrt{n}} \to \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

$$\left[\hat{A}_1, \hat{A}_2\right] = \left[\overline{X} - z_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \overline{X} + z_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right] \quad \text{con} \quad \Phi\left(z_{1-\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2}$$

Ejemplo

Una muestra de 56 muestras de algodón produjo un porcentaje de alargamiento promedio muestral de 8.17 y una desviación estándar de 1.42 . Calcule un intervalo de confianza de nivel 0,90 para el porcentaje de alargamiento promedio verdadero .

$$\overline{X} = 8{,}17$$
$$n = 56$$
$$S = 1{,}42$$
$$z_{\frac{\alpha}{2}} = z_{0{,}05} = 1{,}645$$
$$[\overline{X} - z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \, ; \overline{X} + z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}]$$
$$[8{,}17 - 1{,}645 \frac{1{,}42}{\sqrt{56}} \, ; 8{,}17 + 1{,}645 \frac{1{,}42}{\sqrt{56}}]$$

Xi=porcentaje de alargamiento de la fibra i-ésima

Intervalo de confianza para una proporción

Si $\hat{p}$ es la proporción de individuos de una muestra aleatoria de tamaño grande, que presentan una característica concreta, el estadístico pivote para la construcción del intervalo es

$$\frac{\hat{P}-P}{\sqrt{\frac{\hat{P}(1-\hat{P}}{n}}} \approx N(0.1)$$

y el intervalo de confianza para la proporción p de la población

$$\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

$$IC_{1-\alpha}(P) = [\widehat{P} - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{P}(1-\widehat{P})}{n}}\ ; \widehat{P} + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{P}(1-\widehat{P})}{n}}]$$

$\Rightarrow$ Intervalo de confianza para la proporción

$$long(IC_{1-\alpha}(P)) = 2z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{P}(1-\widehat{P})}{n}}$$

$$\varepsilon = z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{P}(1-\widehat{P})}{n}}$$

Ejemplo

Se recibe_un gran lote de artículos provenientes de un fabricante que asegura que la proporción de artículos defectuosos de su producción es del 1,5%. Al seleccionar una muestra de 200 artículos de la producción e inspeccionarlos, se descubren que hay 8 defectuosos

- a) Obtener un intervalo de confianza del 99% para la proporción de artículos defectuosos en el proceso de fabricación
- b) ¿Qué se puede afirmar acerca de la afirmación del fabricante?

a) $X = número\ de\ artículos\ defectuosos\ entre\ los\ 200\ artículos$

$\widehat{P} = 8/200$ $n = 200$ $z_{\frac{\alpha}{2}} = z_{0,005} = 2,575$

$$[\widehat{P} - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{P}\left(1-\widehat{P}\right)}{n}}\ ; \widehat{P} + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{P}(1-\widehat{P})}{n}}]$$

$$\left[\frac{8}{200} - 2,575\sqrt{\frac{\frac{8}{200}\left(1-\frac{8}{200}\right)}{200}}\ ;\ \frac{8}{200} + 2,575\sqrt{\frac{\frac{8}{200}\,(1-\frac{8}{200})}{200}}\right]$$

$$[0,0043; 0,075]$$

b)

$0,015 \in [0,0043; 0,075]$ la afirmación del fabricante no puede ser rechazada