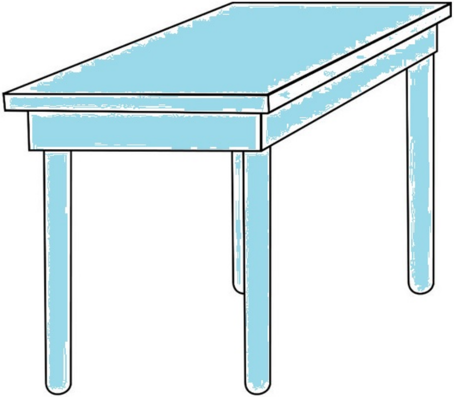
Ciencias de Datos con R: Fundamentos Estadísticos

Intervalos de confianza

Comprar o no comprar la mesa



Intervalos de Confianza - Pesentación

- Usted necesita comprar una mesa que mida TRES metros.
- Más grande, no le entra en la casa.
- Más chica, queda gente sentada en el piso.
- Considere el siguiente conjunto de datos.

$$1.40\ ,\ 0.62\ ,\ 2.40\ ,\ 1.96\ ,\ 0.96\ ,\ 2.16\ ,\ 0.87\ ,\ 2.80\ ,\ 2.31\ ,\ 1.93\ ,$$
$$1.37\ ,\ 0.27\ ,\ 1.30\ ,\ 1.63\ ,\ 0.41\ ,\ 2.78\ ,\ 0.00\ ,\ 0.79\ ,\ 0.83\ ,\ 1.56\ ,$$
$$0.67\ ,\ 1.22\ ,\ 1.84\ ,\ 0.64\ ,\ 1.99\ ,\ 2.93\ ,\ 0.29\ ,\ 1.84\ ,\ 1.58\ ,\ 2.45\ ,$$
$$0.62\ ,\ 1.87\ ,\ 2.80\ ,\ 0.55\ ,\ 1.18\ ,\ 1.03\ ,\ 2.82\ ,\ 2.85\ ,\ 1.22\ ,\ 2.23$$

- Considera usted que la mesa sirve?

Intervalos de Confianza - Ejemplo 2: Calibración

- Objetivo: determinar si un espectrofotómetro está calibrado.
- Ingredientes: gas con 70 ppm de monóxido de carbono.
- Mediciones realizadas con el espectrofotómetro, obteniéndose los siguientes valores:

$$68.67,\ 69.50,\ 70.90,\ 68.64,\ 67.69,\ 71.81,\ 66.06,$$
$$67.35,\ 67.29,\ 73.24$$

- ¿Qué puede decir sobre la calibración del espectrofotómetro?

Intervalos de Confianza - Ejemplo 2: Calibración

- Modelo: $X_i \sim \mathcal{N}(\mu, \sigma_0^2)$.
- Estimador de μ
$$\widehat{\mu}_n = \overline{X}_n$$
- Estimación con los datos:

$$68.67,\ 69.50,\ 70.90,\ 68.64,\ 67.69,\ 71.81,\ 66.06,$$
$$67.35,\ 67.29,\ 73.24$$

- Marta fixit: *Obviamente, μ no tiene por qué valer 69.115* `mean(datos)`
- Vamos a pasar de la estimación puntual a la estimación por intervalo.
- Vamos a dar un **intervalor de valores compatibles con** μ.

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) -

Nueva escuela:
Intervalo de confianza $<-$ Intervalo de compatibilidad

Intervalos de confianza: definición

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

$$\mathbb{P}\left(a(X_1, \ldots, X_n) \ < \ \theta \ < \ b(X_1, \ldots, X_n)\right) = 1 - \alpha \ .$$

Ojo! Necesitamos calcular probabilidades!

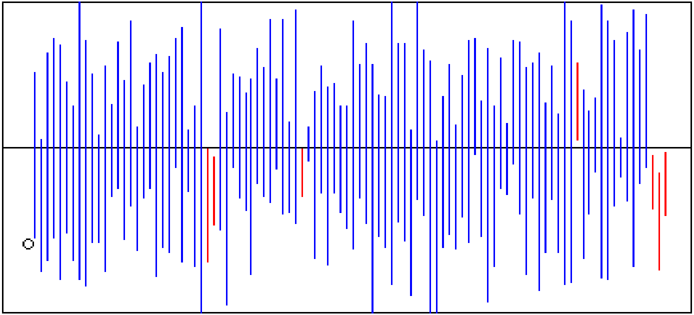
Cada uno con lo suyo.

Dueño de los datos	estimación con $n = 5$
Alejo	69.252
Gonzalo	66.084
Santiago	69.69
Melanie	71.204
Debora	70.9
Carlos	70.67
Elías	70
Rocio	70.176
Catalina	67.986
Facundo	70.816
Julian	70.126
.	.
.	.

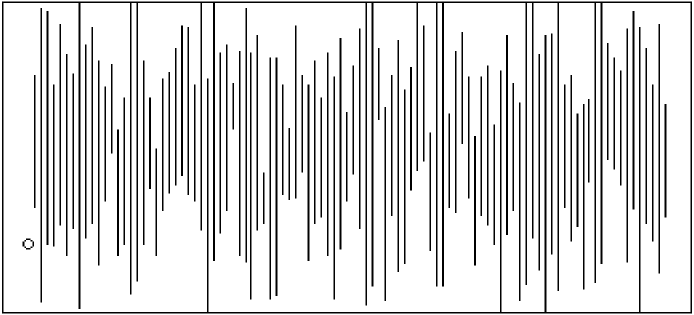
Cada uno con lo suyo.

Dueño de los datos	intervalo de cada uno $n = 5$
Alejo	`(a(datos_alejo),b(datos_alejo))`
Gonzalo	...
Santiago	...
Melanie	...
Debora	...
Carlos	...
Elías	...
Rocio	...
Catalina	...
Facundo	...
Julian	...
.	.
.	.

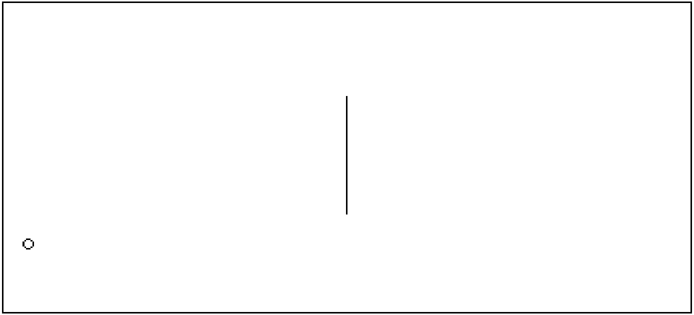
Muchos intervalos y la verdad



Muchos intervalos



Mi intervalo y yo, buena suerte! (confianza)



El mundo normal - Ideal para labo de Física y $n = 5$ datos.

- La distribución χ^2 con k grados de libertad.
- La distribución t- de Student con k grados de libertad.
- La distribución F.

Vamos a la guia. Resolvemos item 2. Mi primera estimación por intervalos - Mundo normal - 10 (minutos)

¿Cómo seguimos?

Elijamos nuestra propia aventura

- Mundo Normal
- Mundo asintótico

Intervalos de Confianza
Modelo normal - 1 Población

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

Nueva notacion para percentiles (o quantiles o...)

Sea z_β con $\mathbb{P}(Z > z_\beta) = \beta$:

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

En particular, utilizaremos mucho $z_{\alpha/2} = \texttt{qnorm(1-alpha/2)}$

- $\alpha = 0.1$, $z_{0.05} =$ `qnorm(1-0.05)` ≈ 1.64
- $\alpha = 0.05$, $z_{0.025} =$ `qnorm(1-0.025)` ≈ 1.96
- $\alpha = 0.01$, $z_{0.005} =$ `qnorm(1-0.005)` ≈ 2.58

$$z_{\alpha/2} = \text{qnorm(1-alpha/2)}$$

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.

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

- ¿Cómo seguimos?

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.

$$\frac{\widehat{\mu}_n - \mu}{\sqrt{\sigma_0^2/n}}\,, \quad \widehat{\mu}_n = \overline{X}_n$$

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

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

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

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

- tenemos que

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

Mundo normal - σ_0^2 conocido

- Equivalentemente, tenemos que

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

y por consiguiente,

$$\left(\widehat{\mu}_n - z_{\alpha/2}\sqrt{\frac{\sigma_0^2}{n}} \quad , \quad \widehat{\mu}_n + z_{\alpha/2}\sqrt{\frac{\sigma_0^2}{n}}\right)$$

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

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

Algunas observaciones

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

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

Para pensar en casa:

- ¿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?

Mundo normal - σ_0^2 conocido vs. σ^2 DESCONOCIDO

$$\widehat{\mu}_n = \overline{X}_n$$

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

$$\frac{\widehat{\mu}_n - \mu}{\sqrt{S^2/n}} \sim t_{n-1}$$

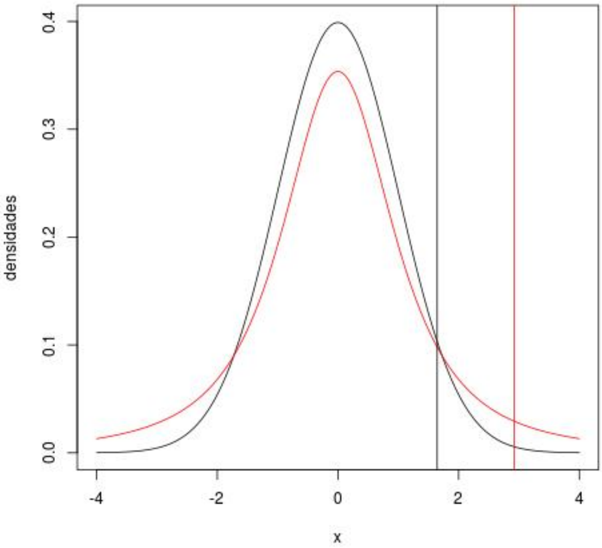
Distribución t-student con $n-1$ grados de libertad.

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

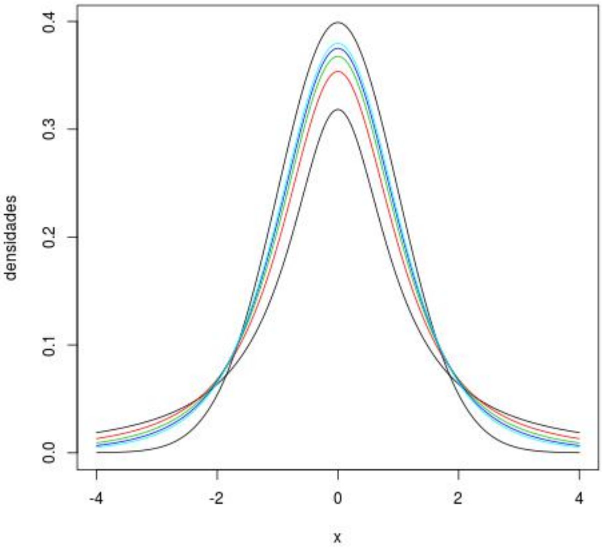
es un intervalo de confianza de nivel $1-\alpha$ para μ bajo el modelo normal, con varianza σ^2 DESconocida.

$t_{n-1,\alpha/2} =$ qt(1-alpha/2,n-1)

Normal vs. t2: mirando percentiles...



Normal y student



Manos a la obra

Recordar que

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

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

$$t_{n-1,\alpha/2} = \text{qt(1-alpha/2,n-1)}$$

Vamos a la guia. Resolvemos item 2. Mi primera estimación por intervalos - Mundo normal junto con 2. en R.

Intervalos de Confianza
Mundo Normal - Dos poblaciones.

Dos Poblaciones: comparación de medias

- Muestra 1: $X_1, \ldots, X_{nx}$ i.i.d, $E(X_i) = \mu_X$ y $V(X_i) = \sigma_X^2$
- Muestra 2: $Y_1, \ldots, Y_{ny}$ i.i.d, $E(Y_i) = \mu_Y$ y $V(X_i) = \sigma_Y^2$

Dos Poblaciones: comparación de medias

- Muestra 1: $X_1, \ldots, X_{nx}$ i.i.d, $E(X_i) = \mu_X$ y $V(X_i) = \sigma_X^2$
- Muestra 2: $Y_1, \ldots, Y_{ny}$ i.i.d, $E(Y_i) = \mu_Y$ y $V(X_i) = \sigma_Y^2$

Nos interesa saber si las medias de ambas poblaciones son o no iguales.

Dos Poblaciones: comparación de medias

- Muestra 1: $X_1, \ldots, X_{nx}$ i.i.d, $E(X_i) = \mu_X$ y $V(X_i) = \sigma_X^2$
- Muestra 2: $Y_1, \ldots, Y_{ny}$ i.i.d, $E(Y_i) = \mu_Y$ y $V(X_i) = \sigma_Y^2$

Nos interesa saber si las medias de ambas poblaciones son o no iguales.

$$\mu_X = \mu_Y ?$$

Dos muestras

Supongamos que tenemos este conjunto de observaciones que sabemos tiene distribucion Normal e igual varianza.

1. Las X‘s 0.44 -1.63 2.59 1.54 0.45 -0.13 -2.76 -1.53

2. Las Y’s 0.06 -0.24 4.65 2.27 3.88 2.35 3.92 -0.73

   Que harián para decidir si $\mu_X = \mu_Y$?

Dos muestras

Supongamos que tenemos este conjunto de observaciones que sabemos tiene distribucion Normal e igual varianza.

1. Las X‘s 0.44 -1.63 2.59 1.54 0.45 -0.13 -2.76 -1.53

2. Las Y’s 0.06 -0.24 4.65 2.27 3.88 2.35 3.92 -0.73

   Que harián para decidir si $\mu_X = \mu_Y$?

   Calculo un intervalo para μ_X obteniendo (-1.60, 1.35) y uno para μ_Y obteniendo (0.27 ,3.76).

Dos muestras

Supongamos que tenemos este conjunto de observaciones que sabemos tiene distribucion Normal e igual varianza.

1. Las X's 0.44 -1.63 2.59 1.54 0.45 -0.13 -2.76 -1.53

2. Las Y's 0.06 -0.24 4.65 2.27 3.88 2.35 3.92 -0.73

   Que harián para decidir si $\mu_X = \mu_Y$?

   Calculo un intervalo para μ_X obteniendo (-1.60, 1.35) y uno para μ_Y obteniendo (0.27 ,3.76).

   Es decir ambos intervalos se intersecan por lo tanto concluyo que las medias son Iguales....

Dos muestras

Supongamos que tenemos este conjunto de observaciones que sabemos tiene distribucion Normal e igual varianza.

1. Las X's 0.44 -1.63 2.59 1.54 0.45 -0.13 -2.76 -1.53

2. Las Y's 0.06 -0.24 4.65 2.27 3.88 2.35 3.92 -0.73

   Que harián para decidir si $\mu_X = \mu_Y$?

   Calculo un intervalo para μ_X obteniendo (-1.60, 1.35) y uno para μ_Y obteniendo (0.27 ,3.76).

   Es decir ambos intervalos se intersecan por lo tanto concluyo que las medias son Iguales....cuando en realidad los datos fueron genreados con $\mu_X = 1$ y $\mu_Y = 2$.

Dos muestras

El modo correcto de hacerlo es contruir un intervalo para $\mu_X - \mu_Y$ en este caso

$$\frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{S_p\sqrt{(1/nx) + (1/ny)}} \sim t_{nx+ny-2}$$

donde $S_p^2 = \frac{(nx-1)S_X^2 + (ny-1)S_Y^2}{nx+ny-2}$.

En el ejemplo queda (-4.22, -0.07)

Intervalos de Confianza
Mundo Asintótico

Intervalos de confianza asintóticos:

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

$$\lim_{n\to\infty} \mathbb{P}\left(a(X_1, \ldots, X_n) \ < \ \theta \ < \ b(X_1, \ldots, X_n)\right) = 1 - \alpha \ .$$

$$\mathbb{P}\left(a(X_1, \ldots, X_n) \ < \ \theta \ < \ b(X_1, \ldots, X_n)\right) \approx 1 - \alpha \ .$$

Ojo! Necesitamos calcular probabilidades aproximadas.
Pediremos ayuda a TCL y sus amigo!

Nueva notacion para percentiles (o c/quantiles o...)

Sea z_β con $\mathbb{P}(Z > z_\beta) = \beta$:

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

En particular, utilizaremos mucho $z_{\alpha/2} =$ `qnorm(1-alpha/2)`

- $\alpha = 0.1$, $z_{0.05} =$ `qnorm(1-0.05)` ≈ 1.64
- $\alpha = 0.05$, $z_{0.025} =$ `qnorm(1-0.025)` ≈ 1.96
- $\alpha = 0.01$, $z_{0.005} =$ `qnorm(1-0.005)` ≈ 2.58

Dibujo!

Intervalo de confianza asintótico para $\mu = \mathbb{E}(X)$

- Por el TCL, sabemos que

$$\frac{(\bar{X}_n - \mu)}{\sqrt{\mathbb{V}(\bar{X}_n)}} = \frac{(\bar{X}_n - \mu)}{\mathrm{se}(\bar{X}_n)} = \frac{(\bar{X}_n - \mu)}{\sigma/\sqrt{n}} \overset{(a)}{\sim} \mathcal{N}(0,1)$$

- Además, tenemos que $\sigma/S \to 1$ y por consiguiente, vale que

$$\frac{\sigma}{S}\frac{(\overline{X}_n - \mu)}{\sigma/\sqrt{n}} = \frac{(\overline{X}_n - \mu)}{S/\sqrt{n}} \overset{(a)}{\sim} \mathcal{N}(0,1)$$

- Entonces

$$\lim_{n\to\infty} \mathbb{P}\left(-z_{\alpha/2} \leq \frac{(\bar{X}_n - \mu)}{S/\sqrt{n}} \leq z_{\alpha/2}\right) = 1 - \alpha$$

- y por consiguiente,

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

es un intervalo de confianza de nivel asintótico $1 - \alpha$ para $\mu = \mathbb{E}(X)$.

Manos a la obra

Sabiendo que

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

es un intervalo de confianza de nivel asintótico $1 - \alpha$ para $\mu = \mathbb{E}(X)$

resolvamos item 3. Intervo asintótico para $\mu = \mathbb{E}(X)$ de la guía de actividades.

$X_i \sim B(1,p)$: Intervalo de confianza asintótico para p

- Por el TCL, sabemos que

$$\frac{\left(\overline{X}_n - p\right)}{\sqrt{\frac{p(1-p)}{n}}} = \frac{\sqrt{n}\left(\overline{X}_n - p\right)}{\sqrt{p(1-p)}} \sim^a \mathcal{N}(0,1)$$

- Además, tenemos que $\widehat{p} = \overline{X}_n \to p$ y por consiguiente, vale que

$$\frac{\sqrt{n}\,(\widehat{p}-p)}{\sqrt{\widehat{p}(1-\widehat{p})}} = \frac{\sqrt{p(1-p)}}{\sqrt{\widehat{p}(1-\widehat{p})}} \frac{\sqrt{n}\,(\widehat{p}-p)}{\sqrt{p(1-p)}} = \frac{\sqrt{n}\,(\widehat{p}-p)}{\sqrt{\widehat{p}(1-\widehat{p})}} \sim^a \mathcal{N}(0,1)$$

- y por consiguiente,

$$\left(\widehat{p} - \frac{\sqrt{\widehat{p}(1-\widehat{p})}\, z_{\alpha/2}}{\sqrt{n}} \quad , \quad \widehat{p} + \frac{\sqrt{\widehat{p}(1-\widehat{p})}\, z_{\alpha/2}}{\sqrt{n}}\right)$$

es un intervalo de confianza de nivel asintótico $1-\alpha$ para p.

Estimadores Asintóticamente Normales

- $\widehat{\theta}_n$ se dice asintóticamente normal (a.n) sii

$$\frac{\widehat{\theta}_n - \theta}{\mathsf{se}} \overset{(a)}{\sim} \mathcal{N}(0,1),$$

donde se $=$ se$(\widehat{\theta}_n)$ denota el desvío estandar del estimador $\widehat{\theta}_n$.

- Ejemplo: $\widehat{\mu}_n = \bar{X}_n$ es a.n., por TCL.

$$\frac{\bar{X}_n - \mu}{\mathsf{se}} \overset{(a)}{\sim} \mathcal{N}(0,1),$$

siendo

$$\mathsf{se} = \mathsf{se}(\bar{X}_n) = \sqrt{\mathbb{V}(\bar{X}_n)} = \sqrt{\frac{\sigma^2}{n}}$$

Intervalos basados en estimadores a.n.

- Sea $\widehat{\theta}_n$ asintóticamente normal

$$\frac{\widehat{\theta}_n - \theta}{\mathsf{se}} = \frac{\widehat{\theta}_n - \theta}{\mathsf{se}(\widehat{\theta}_n)} \overset{(a)}{\sim} \mathcal{N}(0,1) \ ,$$

- Sea $\widehat{\mathsf{se}}$ tal que $\frac{\mathsf{se}}{\widehat{\mathsf{se}}} \to 1$, entonces

$$\frac{\widehat{\theta}_n - \theta}{\widehat{\mathsf{se}}} \overset{(a)}{\sim} \mathcal{N}(0,1) \ ,$$

Llegamos así a que

$$\left(\widehat{\theta}_n - z_{\alpha/2}\,\widehat{\mathsf{se}} \quad , \quad \widehat{\theta}_n + z_{\alpha/2}\,\widehat{\mathsf{se}}\right)$$

es un intervalo de confianza de nivel asintótico $1-\alpha$ para θ.

Intervalo - Error de Estimación

- Intervalo de confianza de nivel asintótico $1-\alpha$ para θ

$$\left(\widehat{\theta}_n - z_{\alpha/2}\,\widehat{\mathsf{se}} \quad , \quad \widehat{\theta}_n + z_{\alpha/2}\,\widehat{\mathsf{se}}\right)$$

- Otra manera de informar

$$\widehat{\theta}_n(\pm\widehat{\mathsf{se}})$$

En general

Si

$$\frac{\widehat{\theta}_n - \theta}{\widehat{\textsf{mongo}}} \overset{(a)}{\sim} \mathcal{N}(0,1) \ ,$$

entonces,

$$\left(\widehat{\theta}_n - z_{\alpha/2}\ \widehat{\textsf{mongo}} \quad , \quad \widehat{\theta}_n + z_{\alpha/2}\ \widehat{\textsf{mongo}}\right)$$

es un intervalo de confianza de nivel asintótico $1-\alpha$ para θ.

Intervalo de confianza para la mediana?

- Distribución de la mediana muestral: asintóticamente normal

$$\frac{\mathsf{med}(X_1,\ldots,X_n)-\mathsf{med}(X)}{\mathsf{se}} \approx \mathcal{N}(0,1) \quad n \text{ grande}$$

- Desvio del Estimador:

$$\mathsf{se}=\mathsf{se}(\mathsf{med}(X_1,\ldots,X_n))=\sqrt{\mathbb{V}_F\{\mathsf{med}(X_1,\ldots,X_n)\}}=???$$

- $\widehat{\mathsf{se}}=??$
- Bootstrap! $\widehat{\mathsf{se}}_{boot}$

$$\text{Intervalo de confianza :}$$

$$\left(\mathsf{med}(X_1,\ldots,X_n)-z_{\alpha/2}\widehat{\mathsf{se}}_{boot}\ ,\ \mathsf{med}(X_1,\ldots,X_n)+z_{\alpha/2}\widehat{\mathsf{se}}_{boot}\right)$$

Intevalos Bootstrap Normal

- $\widehat{\theta}_n$ asintóticamente normal si

$$\frac{\widehat{\theta}_n - \theta}{\mathsf{se}} \approx \mathcal{N}(0,1)$$

  con $\mathsf{se} = \mathsf{se}(\widehat{\theta}_n)$

- Sea $\widehat{\mathsf{se}}_{\text{boot}}$ el estimador bootstrap de $\mathsf{se}(\widehat{\theta}_n)$

intervalo boot normal nivel $1-\alpha$:

$$\left(\widehat{\theta}_n - z_{\alpha/2}\,\widehat{\mathsf{se}}_{\text{boot}}\;,\;\widehat{\theta}_n + z_{\alpha/2}\,\widehat{\mathsf{se}}_{\text{boot}}\right)$$

Intevalos Bootstrap Percentil.

- Sean $\widehat{\theta}^*_1, \ldots, \widehat{\theta}^*_{Nboot}$ estadisticos bootstrap de su estimador.

$$\text{intervalo boot percentil } 1-\alpha: \quad \left(\widehat{\theta}^*_{\alpha/2} \, , \, \widehat{\theta}^*_{1-\alpha/2}\right)$$

Estimadores Asintóticamente normales y el método Delta

- $\widehat{\theta}_n$ asintóticamente normal:

$$\frac{\widehat{\theta}_n - \theta}{\mathsf{se}} \approx \mathcal{N}(0,1)$$

- f función suave. Entonces, haciendo Taylor tenemos

$$\frac{f(\widehat{\theta}_n) - f(\theta)}{\sqrt{\{f'(\theta)\}^2 \mathsf{se}^2}} \approx \mathcal{N}(0,1)$$

- Luego

$$\frac{f(\widehat{\theta}_n) - f(\theta)}{\sqrt{\{f'(\widehat{\theta})\}^2 \widehat{\mathsf{se}}^2}} \approx \mathcal{N}(0,1)$$

$$\left(f(\widehat{\theta}_n) - z_{\alpha/2}\sqrt{\{f'(\widehat{\theta})\}^2 \widehat{\mathsf{se}}^2}\,,\ f(\widehat{\theta}_n) + z_{\alpha/2}\sqrt{\{f'(\widehat{\theta})\}^2 \widehat{\mathsf{se}}^2} \right)$$

Por si querés más explicaciones

Un gran amigo del TCL: Slutsky (Teo - caso particular)

- Sea $\widehat{\theta}_n$ a.n.. y consideramos $\widehat{\mathsf{se}}$, de forma tal que

$$\frac{\mathsf{se}}{\widehat{\mathsf{se}}} \to 1 \ .$$

- En el ejemplo,

$$\mathsf{se} = \mathsf{se}(\bar{X}_n) = \sqrt{\mathbb{V}(\bar{X}_n)} = \sqrt{\frac{\sigma^2}{n}} \qquad \widehat{\mathsf{se}} = \sqrt{\frac{S^2}{n}}$$

- Teorema: Sea $\widehat{\theta}_n$ a.n., y sea $\widehat{\mathsf{se}}$ tal que $\mathsf{se}/\widehat{\mathsf{se}} \to 1$, entonces

$$\frac{\widehat{\theta}_n - \theta}{\widehat{\mathsf{se}}} \overset{(a)}{\sim} \mathcal{N}(0,1).$$

- En el ejemplo,

$$\frac{\bar{X}_n - \mu}{\sqrt{\frac{S^2}{n}}} \overset{(a)}{\sim} \mathcal{N}(0,1) \ , \quad \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sqrt{S^2}} \overset{(a)}{\sim} \mathcal{N}(0,1) \ .$$

Intervalos de Confianza Asintóticamente Normal

- Sea $\widehat{\theta}_n$ asintóticamente normal

$$\frac{\widehat{\theta}_n - \theta}{\mathsf{se}} = \frac{\widehat{\theta}_n - \theta}{\mathsf{se}(\widehat{\theta}_n)} \overset{(a)}{\sim} \mathcal{N}(0,1)\ ,$$

Intervalos de Confianza Asintóticamente Normal

- Sea $\widehat{\theta}_n$ asintóticamente normal

$$\frac{\widehat{\theta}_n - \theta}{\mathsf{se}} = \frac{\widehat{\theta}_n - \theta}{\mathsf{se}(\widehat{\theta}_n)} \overset{(a)}{\sim} \mathcal{N}(0,1)\ , \quad \mathbb{P}\left(\frac{\widehat{\theta}_n - \theta}{\mathsf{se}(\widehat{\theta}_n)} \leq x\right) \to \phi(x)$$

- Sea $\widehat{\mathsf{se}}$ tal que $\frac{\mathsf{se}(\widehat{\theta}_n)}{\widehat{\mathsf{se}}} \to 1$,
- Tenemos entonces que $\mathbb{P}\left(-z_{\alpha/2} \ \leq \ \frac{\widehat{\theta}_n - \theta}{\widehat{\mathsf{se}}} \ \leq \ z_{\alpha/2}\right) \to 1 - \alpha$ y por consiguiente

$$\mathbb{P}\left(\widehat{\theta}_n - z_{\alpha/2}\,\widehat{\mathsf{se}} \ \leq \ \theta \ \leq \ \widehat{\theta}_n + z_{\alpha/2}\,\widehat{\mathsf{se}}\right) \to 1 - \alpha,$$

Llegamos así a que

$$\left(\widehat{\theta}_n - z_{\alpha/2}\,\widehat{\mathsf{se}} \quad , \quad \widehat{\theta}_n + z_{\alpha/2}\,\widehat{\mathsf{se}}\right)$$

es un intervalo de confianza de nivel asintótico $1 - \alpha$ para θ.