Intervalos de Confianza - Parte 2
Modelo normal - 1 Población

Mundo normal

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

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

Dibujo!

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 $P(Z > z_\beta) = \beta$:

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

- 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.

- ¿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

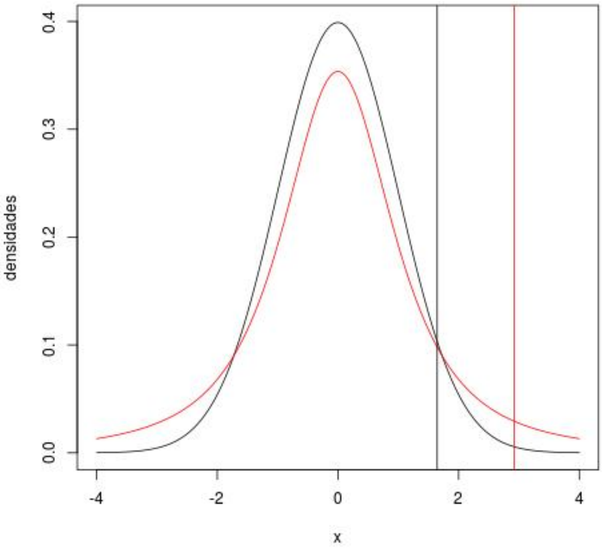
$$\widehat{\mu}_n = \overline{X}_n$$

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

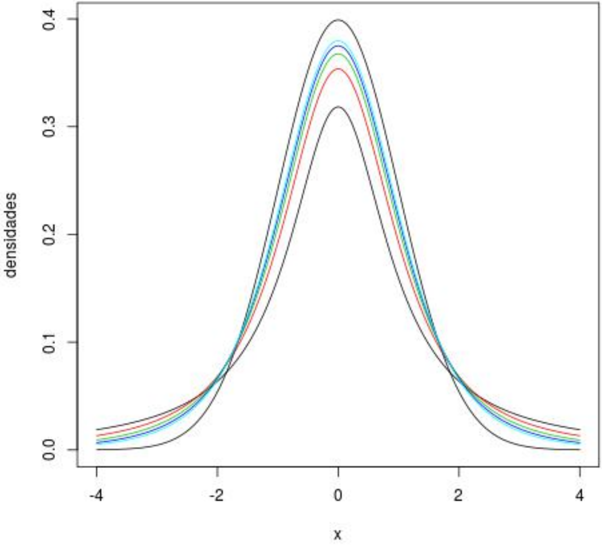
$$\frac{\widehat{\mu}_n - \mu}{\sqrt{S^2/n}} \sim t_{n-1}$$

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

Normal vs. t2: mirando percentiles...



Normal y student



Algunos percentiles útiles

Distribución t de Student con n grados de libertad.
Se tabula $t_{n,\alpha}$, tal que $P(X \geq t_{n,\alpha}) = \alpha$, con $X \in t_n$.

	α					
n	0'250	0'1	0'05	0'025	0'01	0'005
1	1'0000	3'0777	6'3137	12'706	31'821	63'656
2	0'8165	1'8856	2'9200	4'3027	6'9645	9'9250
3	0'7649	1'6377	2'3534	3'1824	4'5407	5'8408
4	0'7407	1'5332	2'1318	2'7765	3'7469	4'6041
5	0'7267	1'4759	2'0150	2'5706	3'3649	4'0321
6	0'7176	1'4398	1'9432	2'4469	3'1427	3'7074
7	0'7111	1'4149	1'8946	2'3646	2'9979	3'4995
8	0'7064	1'3968	1'8595	2'3060	2'8965	3'3554
9	0'7027	1'3830	1'8331	2'2622	2'8214	3'2498
10	0'6998	1'3722	1'8125	2'2281	2'7638	3'1693

Mundo normal - σ^2 desconocido

- Buscamos intervalo de confianza para μ con σ^2 desconocido.
- Pivot:
$$\frac{\widehat{\mu}_n - \mu}{\sqrt{S^2/n}} \sim t_{n-1} \ , \quad \widehat{\mu}_n = \overline{X}_n \ .$$
- Sea $t_{n-1,\beta}$ con $\mathbb{P}(X > t_{n-1,\beta}) = \beta$ cuando $X \sim t_{n-1}$, entonces
$$\mathbb{P}\left(-t_{n-1,\alpha/2} \ < \ \frac{\widehat{\mu}_n - \mu}{\sqrt{S^2/n}} \ < \ t_{n-1,\alpha/2}\right) = 1 - \alpha$$

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

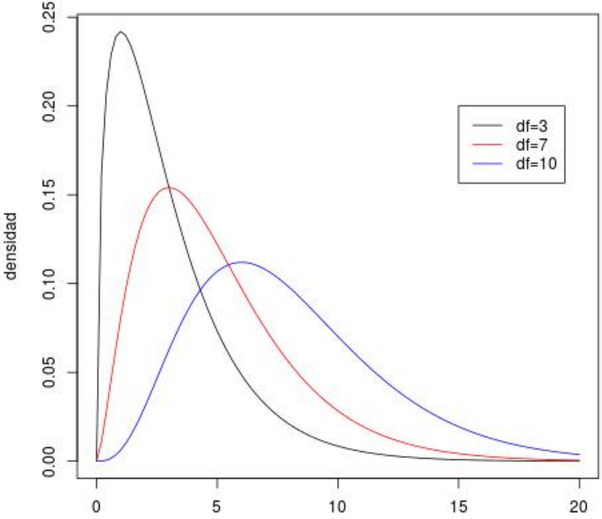
  es un intervalo de confianza de nivel $1-\alpha$ para μ bajo el modelo normal, con varianza σ^2 DESconocida.

Vayamos a la Guía de actividades

Distribución chi cuadrado: definición

- Sean $Z_1, \ldots, Z_n$ i.i.d., $Z_i \sim \mathcal{N}(0,1)$.
  Definición: Llamamos chi cuadrado con n- grados de libertad a la distribución de
  $$Z_1^2 + Z_2^2 + \cdots + Z_n^2$$
- Notación: χ_n^2
- Fonética: chi cuadrado con n grados de libertad.
- Es decir,
  $$Z_1^2 + Z_2^2 + \cdots + Z_n^2 \sim \chi_n^2 \ .$$
- En R: $\star$`chisq` - cuantiles `qchisq(p, df)`

Densidad chi cuadrado (Ojo! no es simétrica en el 0.)



Caso particular:

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

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

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

$$\frac{\sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2}{\sigma^2} \sim \chi^2_{n-1} \,, \quad \text{es decir } \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

recordemos que $\quad S^2 := \dfrac{1}{n-1} \displaystyle\sum_{i=1}^{n} (X_i - \overline{X}_n)^2$

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.

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

- ¿Cómo seguimos?

Distribución χ^2 de Pearson con n grados de libertad.
Se tabula $\chi^2_{n,\alpha}$, tal que $P(X \geq \chi^2_{n,\alpha}) = \alpha$, con $X \in \chi^2_n$.

	α									
n	0'995	0'990	0'975	0'950	0'900	0'100	0'050	0'025	0'010	0'005
1	0'0000	0'0002	0'0010	0'0039	0'0158	2'706	3'841	5'024	6'635	7'879
2	0'0100	0'0201	0'0506	0'1026	0'2107	4'605	5'991	7'378	9'210	10'597
3	0'0717	0'1148	0'2158	0'3518	0'5844	6'251	7'815	9'348	11'345	12'838
4	0'2070	0'2971	0'4844	0'7107	1'0636	7'779	9'488	11'143	13'277	14'860
5	0'4118	0'5543	0'8312	1'1455	1'6103	9'236	11'070	12'832	15'086	16'750
6	0'6757	0'8721	1'2373	1'6354	2'2041	10'645	12'592	14'449	16'812	18'548
7	0'9893	1'2390	1'6899	2'1673	2'8331	12'017	14'067	16'013	18'475	20'278
8	1'3444	1'6465	2'1797	2'7326	3'4895	13'362	15'507	17'535	20'090	21'955
9	1'7349	2'0879	2'7004	3'3251	4'1682	14'684	16'919	19'023	21'666	23'589
10	2'1558	2'5582	3'2470	3'9403	4'8652	15'987	18'307	20'483	23'209	25'188
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

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.

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

- Sea $\chi^2_{n-1,\beta}$ con $\mathbb{P}(X > \chi^2_{n-1,\beta}) = \beta$ cuando $X \sim \chi^2_{n-1}$, entonces

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

y por consiguiente,

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

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

Intervalos de confianza para σ^2 con μ_0 conocido

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

$$\frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma^2} \sim \chi_n^2$$

- Sea $\chi^2_{n,\beta}$ con $\mathbb{P}(X > \chi^2_{n,\beta}) = \beta$ cuando $X \sim \chi^2_n$, entonces

$$\mathbb{P}\left(\chi^2_{n,1-\alpha/2} \ < \ \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma^2} \ < \ \chi^2_{n,\alpha/2}\right) = 1 - \alpha$$

y por consiguiente,

$$\left(\frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\chi^2_{n,\alpha/2}} \quad , \quad \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\chi^2_{n,1-\alpha/2}}\right)$$

es un intervalo de confianza de nivel $1 - \alpha$ para σ^2 bajo el modelo normal, con $\mu = \mu_0$ CONOCIDA.

Construcción de intervalos de confianza - PIVOT

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

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

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

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

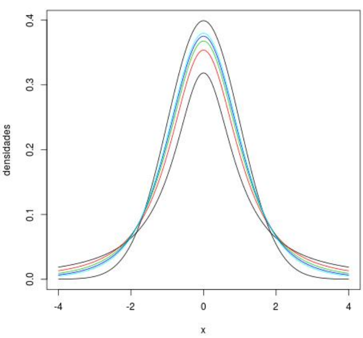
- Encerramos al pivot entre los percentiles de su distribución y despejamos θ.
- Identifique el pivot para cada uno de los intervalos de confianza construídos.

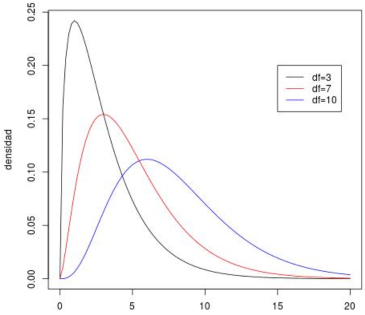
Resumen bajo normalidad: $X_i \sim \mathcal{N}(\mu, \sigma^2)$

Parámetro	Supuesto	Pivot
$\mu = \mathbb{E}(X)$	Normales σ^2 conocido	$\frac{\widehat{\mu}_n - \mu}{\sqrt{\sigma^2/n}} \sim \mathcal{N}(0,1)$
$\mu = \mathbb{E}(X)$	Normales σ^2 DESconocido	$\frac{\widehat{\mu}_n - \mu}{\sqrt{S^2/n}} \sim t_{n-1}$
$\sigma^2 = V(X)$	Normales μ conocido	$\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2_n$
$\sigma^2 = V(X)$	Normales μ DESconocido	$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

$$\widehat{\mu}_n = \overline{X}_n \ , \quad (n-1)S^2 = \sum_{i=1}^n (X_i - \overline{X}_n)^2$$

normal y student – chi cuadrado





Agregamos algo más (formal) sobre la distribución t-student, que vamos a usar luego...

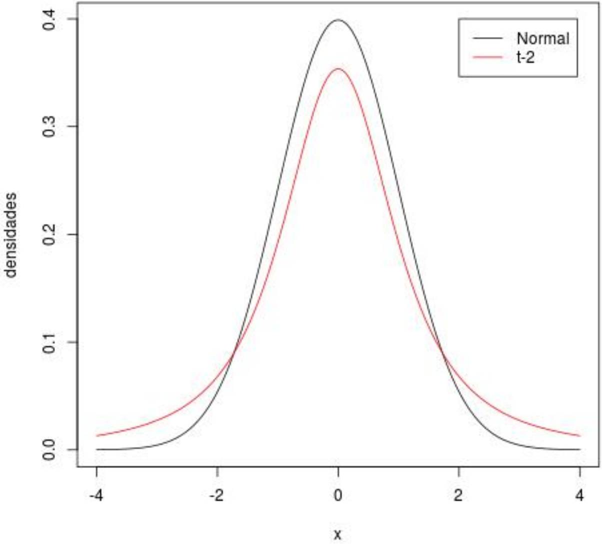
Distribución t de student: definición

- Z, V v.a. independientes,
- $Z \sim \mathcal{N}(0, 1)$ y $V \sim \chi^2_n$.
- Llamamos t de student con n grados de libertad a la distribución de

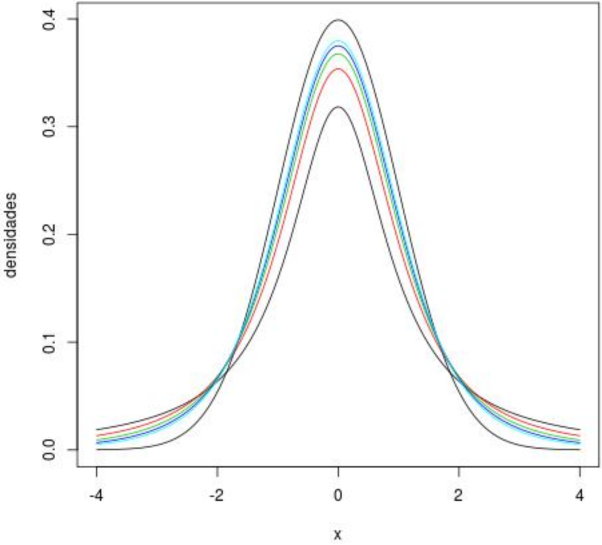
$$T_n = \frac{Z}{\sqrt{V/n}}.$$

En adelante, utilizamos t_n para denotar la t de student con n grados de libertad.

Normal vs. t2



Normal y student



Caso particular:

- $X_1, \ldots, X_n$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$. $\overline{X}_n \sim \mathcal{N}(\mu, \sigma^2/n)$. Entonces

$$Z := \frac{\left(\overline{X}_n - \mu\right)}{\sqrt{\sigma^2/n}} \sim \mathcal{N}(0,1) \quad V := \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

- y además son independientes sin demostrar. Luego,

$$\frac{Z}{\sqrt{V/(n-1)}} = \frac{\left(\overline{X}_n - \mu\right)}{\sqrt{S^2/n}} \sim t_{n-1}$$