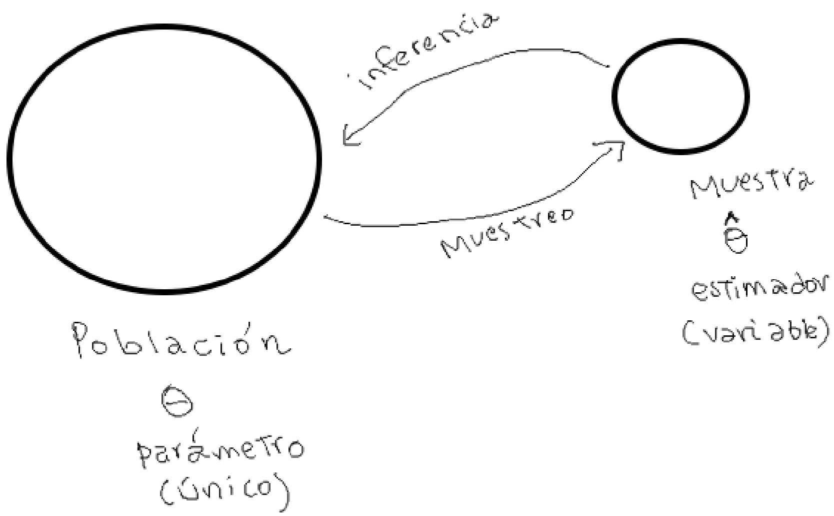
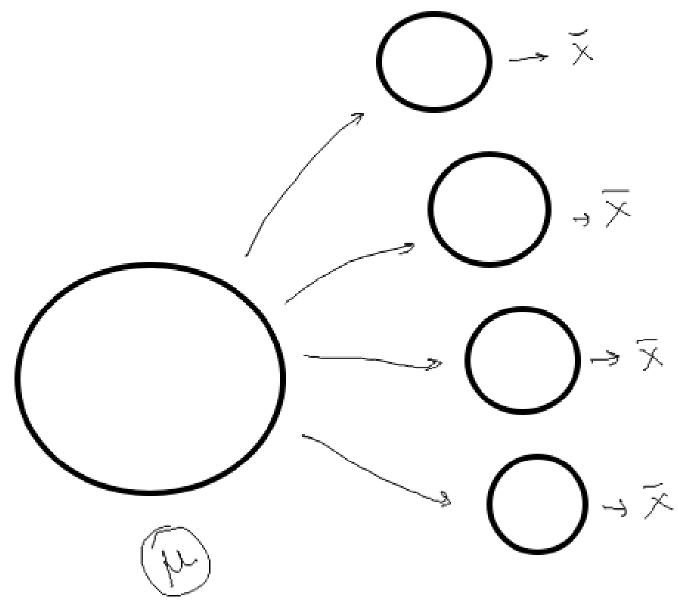
Inferencia estadística



cálculos → datos muestrales

conclusión → poblacional

- ↓ intervalos de confianza
- ↘ prueba de hipótesis
  - paramétricas
  - no paramétricas



Estimación intervalar de la media

Cuando la varianza poblacional es conocida (caso teórico)

$$\left(\underbrace{\bar{X} - \overset{\text{valor de } N(0,1)}{Z_{(1-\alpha/2)}}\frac{\overset{\text{desv. est. poblacional}}{\sigma}}{\sqrt{n}}}_{a} \leq \mu \leq \underbrace{\bar{X} + Z_{(1-\alpha/2)}\frac{\sigma}{\sqrt{n}}}_{b}\right)$$

promedio muestral

Tamaño muestral

intervalo simétrico:

$\bar{X} \pm$ margen de error

$$Z_{(1-\alpha/2)}\frac{\sigma}{\sqrt{n}}$$

$\uparrow n$ margen de error $\downarrow$

$\uparrow$ varianza margen de error $\uparrow$

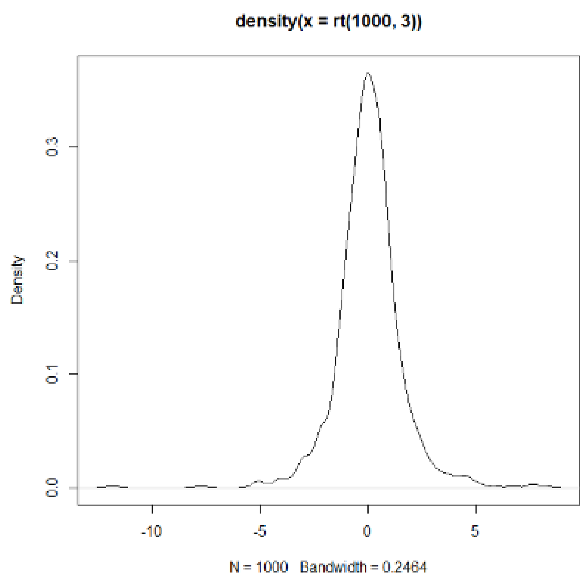
* Cuando la varianza poblacional es desconocida (caso realista)

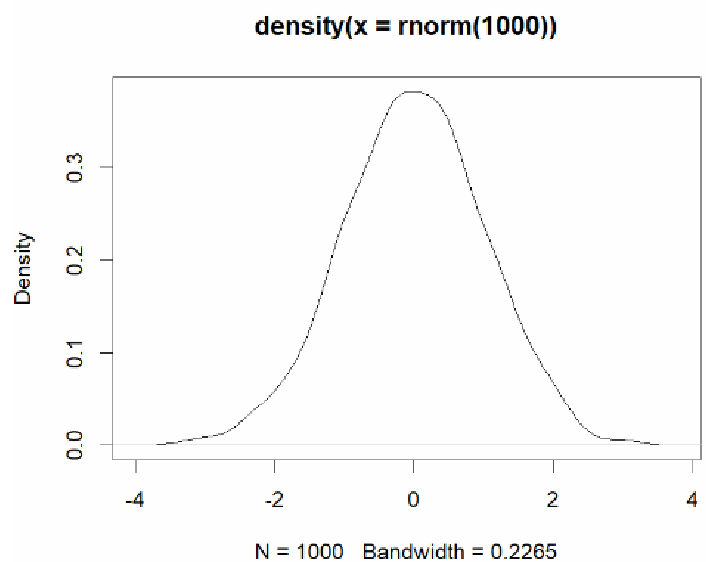
$$\left(\underbrace{\bar{X} - \overset{\text{valor de } t}{t_{(1-\alpha/2;n-1)}}\frac{\overset{\text{desv. est. muestral}}{s}}{\sqrt{n}}}_{a} \leq \mu \leq \underbrace{\bar{X} + t_{(1-\alpha/2;n-1)}\frac{s}{\sqrt{n}}}_{b}\right)$$

promedio muestral

tamaño muestral

intervalo simétrico pero más amplio que el caso teórico






```
> qt(0.975, 2)
[1] 4.302653
> qt(0.975, 15)
[1] 2.13145
> qt(0.975, 30)
[1] 2.042272
> qt(0.975, 80)
[1] 1.990063
> qt(0.975, 200)
[1] 1.971896
> qt(0.975, 500)
[1] 1.96472
> qt(0.975, 1000)
[1] 1.962339
> qnorm(0.975)
[1] 1.959964
```

t es más amplia variable $\Rightarrow$ $IC(\mu)$ con σ^2 desconocida es más amplio

gl $\uparrow$ $\Rightarrow$ $t \to N$

↓ confianza: ↓ amplitud del intervalo

↑ confianza: ↑ amplitud del intervalo

para estar más seguros de que el parámetro estará realmente en el intervalo, optamos por aumentar su amplitud.

Nivel de confianza = $1-\alpha$

4, 7, 3, 6 $\bar{x} = 5$ $n = 4$, $gl = 3$ ⇒ 3 observ. "libres" manteniendo el $\bar{x}$ fijo

5, 5, 4, 6 $\bar{x} = 5$

12, 7, 0, 1 $\bar{x} = 5$

$$\left(\underbrace{\bar{X} - \overbrace{t_{(1-\alpha/2;n-1)}\frac{s}{\sqrt{n}}}^{\text{margen de error}}}_{a} \leq \mu \leq \underbrace{\bar{X} + t_{(1-\alpha/2;n-1)}\boxed{\frac{s}{\sqrt{n}}}}_{b}\right)$$

error estándar

"desv. estándar" de $\bar{X}$

Estimación intervalar de la varianza

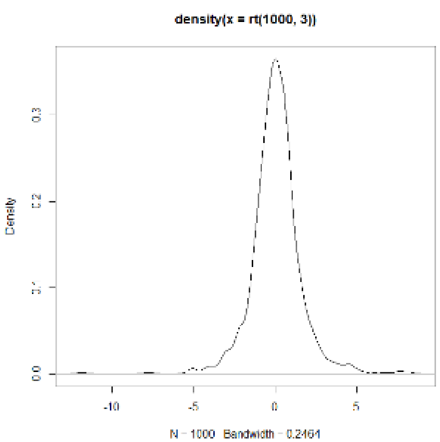
Tamaño muestral >0

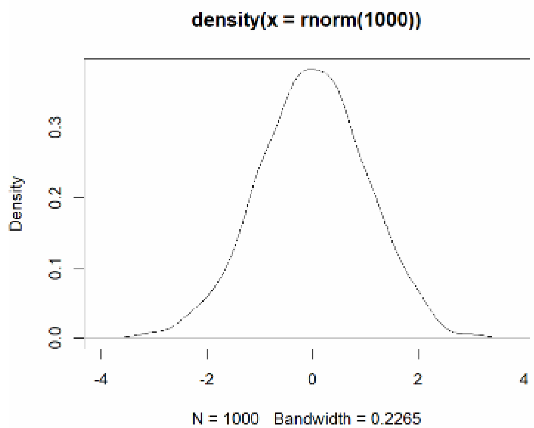
varianza muestral >0

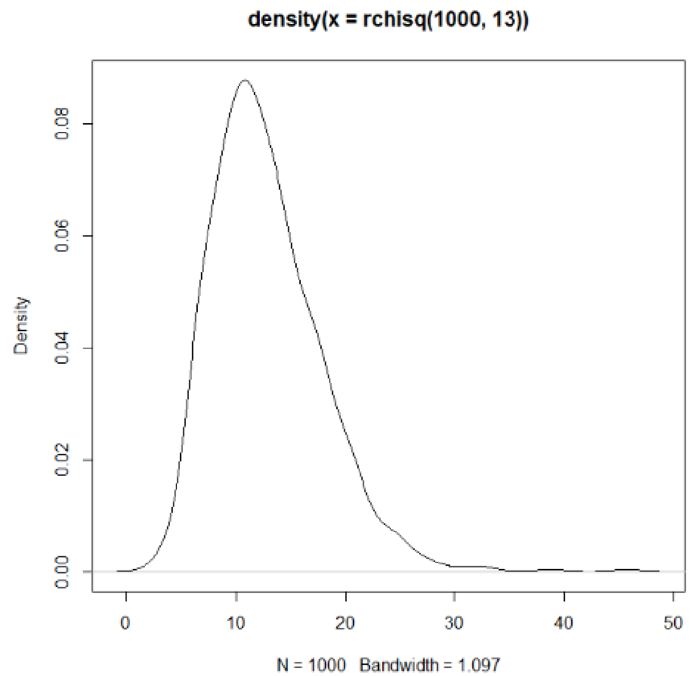
$$\underbrace{\frac{(n-1)s^2}{\chi^2_{(1-\alpha/2;n-1)}}}_{a} \leq \sigma^2 \leq \underbrace{\frac{(n-1)s^2}{\chi^2_{(\alpha/2;n-1)}}}_{b}$$

valor Chi cuadrado >0

intervalo asimétrico y positivo







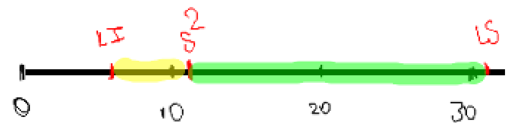
```
> sodio |>
+   varTest(conf.level = 0.95)|>
+   use_series(conf.int)
      LCL       UCL
5.565681 31.972759
attr(,"conf.level")
[1] 0.95
```

$IC(\sigma^2) =$ 5.565681 31.972759 } estimación intervalar

```
> sodio |> var()
[1] 11.09091
```

$= s^2$

estimación puntual



intervalo asimétrico

Estimación intervalar de la proporción

Caso 1: Aproximación Normal

$(n > 30,\ np > 5,\ n(1-p) > 5)$

p : proporción $\in (0,1)$

margen de error

$$\underbrace{p - Z_{(1-\alpha/2)}\sqrt{\frac{p(1-p)}{n}}}_{a} \leq \pi \leq \underbrace{p + Z_{(1-\alpha/2)}\boxed{\sqrt{\frac{p(1-p)}{n}}}}_{b}$$

→ error estándar de p

proporción muestral

Tamaño muestral

intervalo simétrico : $p \pm$ M. error

$\uparrow n$ margen de error $\downarrow$

$N = 10000$

$n = 9000$

$n = 90$

mayor margen de error

¿qué valor de p maximiza $\sqrt{\frac{p(1-p)}{n}}$?

$p = 0.50 \Rightarrow$ desconocimiento o incertidumbre total.