Tema 8: Estimación de Parámetros

Síntesis de fórmulas y funciones en R Studio

ESTIMACIÓN DE LA MEDIA POBLACIONAL μ

Estimación de la media μ: $\overline{x} \pm E$

Varianza de la población	*Tamaño de muestra*	
	n≥30	n<30
σ^2 conocida	$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$	
σ^2 desconocida	$E = z_{\alpha/2}\frac{s}{\sqrt{n}}$	$E = t_{\alpha/2,\nu}\frac{s}{\sqrt{n}}$ con $\nu = n-1$ *Población aprox. normal*

EN R STUDIO: se puede hacer directamente con la **función t.test(…)$conf.int** siempre que contemos con las observaciones de la muestra cargadas en un vector x.

```
t.test(x,conf.level = 0.99)$conf.int
```

ESTIMACIÓN DE LA DIFERENCIA DE MEDIAS DE DOS POBLACIONES $\mu_1 - \mu_2$

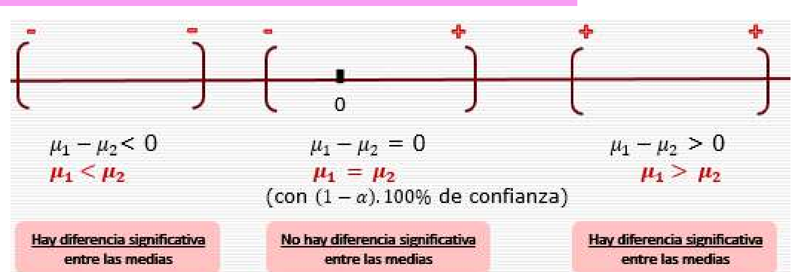
Estimación de la diferencia de medias $\mu_1 - \mu_2$: $(\overline{x_1} - \overline{x_2}) \pm E$

Varianzas de las poblaciones	*Tamaños muestrales*	
	$n_1 \geq 30$ y $n_2 \geq 30$	$n_1 < 30$ y $n_2 < 30$
σ_1^2 y σ_2^2 conocidas	$E = z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$	
σ_1^2 y σ_2^2 desconocidas	$E = z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$	Var. desconocidas IGUALES y NORMALIDAD $E = t_{\alpha/2,\nu}\, s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ con $\nu = n_1+n_2-2$ $s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ Var. desconocidas DIFERENTES y NORMALIDAD $E = t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$ con $\nu = \frac{\left(s_1^2/n_1+s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1}+\frac{\left(s_2^2/n_2\right)^2}{n_2-1}}$

EN R STUDIO: se puede hacer directamente con la **función t.test(…)$conf.int** siempre que contemos con las observaciones de las muestras cargadas en dos vectores x1 y x2, por ejemplo. Se fija "var.equal=TRUE" si las varianzas poblacionales son IGUALES o "var.equal=FALSE" si son diferentes.

```
t.test(x1,x2, var.equal=FALSE, conf.level = 0.90)$conf.int
```

INTERPRETACIÓN



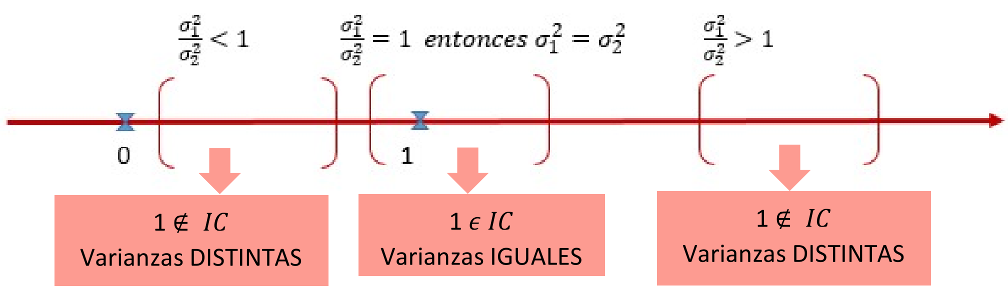
ESTIMACIÓN DEL COCIENTE DE *VARIANZAS* $\frac{\sigma_1^2}{\sigma_2^2}$ DE DOS POBLACIONES

Estimación por intervalo del $(1-\alpha).100\%$ de confianza: $\frac{s_1^2}{s_2^2}\frac{1}{f_{\frac{\alpha}{2}}(\nu_1,\nu_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}\, f_{\frac{\alpha}{2}}(\nu_2,\nu_1)$

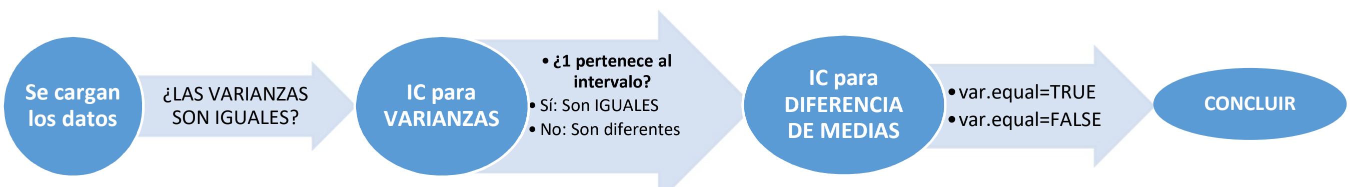
EN R STUDIO: se puede hacer directamente con la **función var.test(…)$conf.int** siempre que contemos con las observaciones de las muestras cargadas en dos vectores x1 y x2, por ejemplo.

```
var.test(x1, x2, conf.level = 0.94)$conf.int
```

INTERPRETACIÓN



OBSERVACIÓN: si se tiene que hacer un IC para diferencia de medias y no se sabe nada acerca de las varianzas poblacionales, los pasos a seguir son:




```
var.test(x1,x2, conf.level = 0.90)$conf.int
```

```
t.test(x1,x2, var.equal=…, conf.level = 0.90)$conf.int
```

ESTIMACIÓN DE UNA VARIANZA σ^2 DE UNA POBLACIÓN

Intervalo de confianza para σ^2

Si s^2 es la varianza de una muestra aleatoria de tamaño n de una población normal, un intervalo de confianza del $100(1-\alpha)\%$ para σ^2 es

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}},$$

donde $\chi^2_{\alpha/2}$ y $\chi^2_{1-\alpha/2}$ son valores χ^2 con $v = n-1$ grados de libertad, que dejan áreas de $\alpha/2$ y $1-\alpha/2$, respectivamente, a la derecha.

Un intervalo de confianza aproximado a $100(1-\alpha)\%$ para σ se obtiene tomando la raíz cuadrada de cada extremo del intervalo para σ^2.

EN R STUDIO: Aplicamos la fórmula ingresando el ALFA/2 que corresponde.

```
(n-1)*var(x)/qchisq(0.025,df=4,lower.tail=FALSE)

(n-1)*var(x)/qchisq(0.025,df=4,lower.tail=TRUE)
```

ESTIMACIÓN DE LA PROPORCIÓN DE ÉXITOS POBLACIONAL P

$$\hat{p} \pm z_{\frac{\alpha}{2}} . \sqrt{\frac{\hat{p}.\hat{q}}{n}} \qquad \text{siendo } \hat{p} = \frac{x}{n} \qquad \hat{q} = 1 - \hat{p}$$

EN R STUDIO: se puede hacer directamente con la **función prop.test(…)$conf.int** para lo cual necesitamos x (nro. éxitos en muestra) y n (tamaño de la muestra).

```
prop.test(x=, n=, conf.level=0.95)
```

ESTIMACIÓN DE LA DIFERENCIA DE PROPORCIONES DE DOS POBLACIONES $p_1 - p_2$

Intervalo de confianza para $p_1 - p_2$ de una muestra grande

Si $\hat{p}_1$ y $\hat{p}_2$ son las proporciones de éxitos en muestras aleatorias de tamaños n_1 y n_2, respectivamente, $\hat{q}_1 = 1 - \hat{p}_1$ y $\hat{q}_2 = 1 - \hat{p}_2$, un intervalo de confianza aproximado del $100(1 - \alpha)\%$ para la diferencia de dos parámetros binomiales $p_1 - p_2$ es dado por

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}},$$

donde $z_{\alpha/2}$ es el valor z que deja una área de $\alpha/2$ a la derecha.

EN R STUDIO: se puede hacer directamente con la **función prop.test(…)$conf.int** para lo cual necesitamos x1 (nro. éxitos en muestra 1), x2 (nro. éxitos en muestra 2), n1 (tamaño de muestra 1) y n2 (tamaño de muestra 2).

```
prop.test(x=c(x1, x2), n=c(n1, n2), conf.level=0.95)
```