

Tema 9: Pruebas de Hipótesis

Síntesis de fórmulas y funciones en R Studio

CRITERIOS GENERALES PARA DECIDIR

	Decisión	Conclusión	Error que puede cometerse
$p - value < \alpha$	Rechazar H_0	Hay evidencia muestral suficiente para afirmar H_1 .	Tipo I
$p - value > \alpha$	No rechazar H_0	No hay evidencia muestral suficiente para afirmar H_1 .	Tipo II

↑ La decisión la expresamos en función de H_0 ↑ La conclusión la expresamos en función de H_1

PRUEBA DE HIPÓTESIS PARA LA MEDIA POBLACIONAL μ **PLANTEO**

$$H_0: \mu = \mu_0 \quad \text{vs}$$

$$H_1: \mu < \mu_0 \quad \text{o} \quad H_1: \mu > \mu_0 \quad \text{o} \quad H_1: \mu \neq \mu_0$$

ESTADÍSTICOS DE PRUEBA

VARIANZA POBLACIONAL	$n \geq 30$	$n < 30$
σ^2 CONOCIDA	$z_{calc} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$	$z_{calc} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$
σ^2 DESCONOCIDA	$z_{calc} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$	$t_{calc} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$ CON $v = n - 1$ <small>*POBLACIÓN NORMAL*</small>

Hipótesis nula	Hipótesis alternativa	P-value	Cálculo de P-value en R Studio
$H_0: \mu = \mu_0$	$H_1: \mu < \mu_0$	$P(Z \leq z_{calc}) \quad \text{o} \quad P(T \leq t_{calc})$	<code>pnorm(zcalc, lower.tail=TRUE)</code> ó <code>pt(tcalc, df=n-1, lower.tail=TRUE)</code>
	$H_1: \mu > \mu_0$	$P(Z \geq z_{calc}) \quad \text{o} \quad P(T \geq t_{calc})$	<code>pnorm(zcalc, lower.tail=FALSE)</code> ó <code>pt(tcalc, df=n-1, lower.tail=FALSE)</code>
	$H_1: \mu \neq \mu_0$	$2P(Z \geq z_{calc}) \quad \text{o} \quad 2P(T \geq t_{calc})$	<code>2*pnorm(abs(zcalc), lower.tail=FALSE)</code> ó <code>2*pt(abs(tcalc), df=n-1, lower.tail=FALSE)</code>



EN R STUDIO: se puede hacer directamente con la función **t.test** siempre que contemos con las observaciones de la muestra cargadas en un vector x. Decidimos en función del p-value.

```
t.test(x, #variable
       alternative = "two.sided", # "two.sided"(dist),"greater">>,"less"><
       mu=10) #valor de mu_0
```

PRUEBA DE HIPÓTESIS PARA LA DIFERENCIA DE MEDIAS DE DOS POBLACIONES $\mu_1 - \mu_2$

PLANTEO $H_0: \mu_1 - \mu_2 = d_0$ vs $H_1: \mu_1 - \mu_2 < d_0$ ó $H_1: \mu_1 - \mu_2 > d_0$ ó $H_1: \mu_1 - \mu_2 \neq d_0$

ESTADÍSTICOS DE PRUEBA

VARIANZAS POBLACIONALES	$n_1 \geq 30$ y $n_2 \geq 30$	$n_1 < 30$ y $n_2 < 30$
σ_1^2 y σ_2^2 CONOCIDAS		$z_{calc} = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$
σ_1^2 y σ_2^2 DESCONOCIDAS	$z_{calc} = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	<p>VAR. DESCONOCIDAS IGUALES Y NORMALIDAD</p> $t_{calc} = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ <p>CON $\nu = n_1 + n_2 - 2$; $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$</p> <p>VAR. DESCONOCIDAS DIFERENTES Y NORMALIDAD</p> $t_{calc} = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \text{ CON } \nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1} + \frac{\left(s_2^2/n_2\right)^2}{n_2-1}}$

EN R STUDIO: se puede hacer directamente con la **función t.test(...)** 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. Decidimos en función del p-value.

```
t.test(x1, x2, # variables
       var.equal=TRUE, # ¿varianzas poblacionales iguales?
       alternative = "greater",# "two.sided"(dist),"greater"(>),"less"(<)
       mu=0) #valor de d_0
```

PRUEBA DE HIPÓTESIS PARA COCIENTE DE VARIANZAS $\frac{\sigma_1^2}{\sigma_2^2}$ DE DOS POBLACIONES

PLANTEO $H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$ vs $H_1: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$



EN R STUDIO: se puede hacer directamente con la **función var.test** siempre que contemos con las observaciones de las muestras cargadas en dos vectores x1 y x2, por ejemplo. Decidimos en función del p-value.

```
var.test(x1, x2, alternative = "two.sided")
```

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

