

# Contrastes de hipótesis más frecuentes

## 1. Contrastes con una muestra

---

### 1.1. Contrastes sobre la media de una distribución $N(\mu, \sigma)$ .

$H_0$	$H_1$	Tamaño de muestra	Estadístico de contraste	Región de rechazo de $H_0$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\frac{(z_{\alpha/2} + z_{\beta})^2 s^2}{\Delta^2}$	$t_{exp} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$ t_{exp}  > t_{n-1, \alpha/2}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\frac{(z_{\alpha} + z_{\beta})^2 s^2}{\Delta^2}$		$t_{exp} > t_{n-1, \alpha}$
$\mu \geq \mu_0$	$\mu < \mu_0$			$t_{exp} < -t_{n-1, \alpha}$

En R :

$H_0$	$H_1$	$x$ =vector que contiene los datos. Regla de decisión: Rechazar $H_0$ si $p$ -valor $< \alpha$
$\mu = \mu_0$	$\mu \neq \mu_0$	<code>t.test(x, mu=mu0)</code>
$\mu \leq \mu_0$	$\mu > \mu_0$	<code>t.test(x, mu=mu0, alternative="greater")</code>
$\mu \geq \mu_0$	$\mu < \mu_0$	<code>t.test(x, mu=mu0, alternative="less")</code>

## 1.2. Contrastes sobre la varianza $\sigma^2$ de una distribución $N(\mu, \sigma)$ .

$H_0$	$H_1$	Tamaño de muestra	Estadístico de contraste	Región de rechazo de $H_0$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$		$c_{exp} = \frac{(n-1)S^2}{\sigma_0^2}$	$c_{exp} < \chi_{n-1, 1-\alpha/2}^2$ ó $c_{exp} > \chi_{n-1, \alpha/2}^2$
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$			$c_{exp} > \chi_{n-1, \alpha}^2$
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$			$c_{exp} < \chi_{n-1, 1-\alpha}^2$

En R : antes de realizar el test, ejecutar `require(TeachingDemos)`

$H_0$	$H_1$	$x$ =vector que contiene los datos. Regla de decisión: Rechazar $H_0$ si $p$ -valor $< \alpha$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	<code>sigma.test(x, sigma=sigma0)</code>
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	<code>sigma.test(sigma, sigma=sigma0, alternative="greater")</code>
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$	<code>sigma.test(sigma, sigma=sigma0, alternative="less")</code>

**Nota:** el paquete `TeachingDemos` no se instala por defecto con R , sino que debe descargarse e instalarse (`Paquetes`→`Instalar Paquetes` →`Madrid (Spain)`→`TeachingDemos`)

### 1.3. Contrastes sobre la proporción $\pi$ de una distribución $B(n, \pi)$ .

	$H_0$	$H_1$	Tamaño de muestra	Estadístico de contraste	Región de rechazo de $H_0$
Contraste exacto	$\pi = \pi_0$	$\pi \neq \pi_0$		$N_E = \text{"número de éxitos en } n \text{ pruebas"}$	$N_E < b_{n,1-\alpha/2}(\pi_0)$ ó $N_E > b_{n,\alpha/2}(\pi_0)$
	$\pi \leq \pi_0$	$\pi > \pi_0$			$N_E > b_{n,\alpha}(\pi_0)$
	$\pi \geq \pi_0$	$\pi < \pi_0$			$N_E < b_{n,1-\alpha}(\pi_0)$
Contraste aproximado ( $n \geq 30$ ) $N_E > 5$ $n - N_E > 5$	$\pi = \pi_0$	$\pi \neq \pi_0$	$\left(\frac{z_{\alpha/2}\sqrt{p_0q_0} + z_{\beta}\sqrt{p_1q_1}}{\Delta}\right)^2$	$z_{exp} = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$	$ z_{exp}  > z_{\alpha/2}$
	$\pi \leq \pi_0$	$\pi > \pi_0$	$\left(\frac{z_{\alpha}\sqrt{p_0q_0} + z_{2\beta}\sqrt{p_1q_1}}{\Delta}\right)^2$		$z_{exp} > z_{\alpha}$
	$\pi \geq \pi_0$	$\pi < \pi_0$			$z_{exp} < -z_{\alpha}$

donde, siendo  $X \approx B(n, \pi_0)$ , se define:

$$b_{n,\alpha}(\pi_0) = \min\{k \mid P(X > k) \leq \alpha\} = \min\{k \mid P(X \leq k) \geq 1 - \alpha\} = \text{qbinom}(1 - \alpha, n, \pi_0)$$

En el cálculo del tamaño de muestra,  $p_0$  es un valor inicial para  $\pi_0$  (obtenido por muestreo piloto o de la bibliografía), y:

$$p_1 = \begin{cases} p_0 + \Delta & \text{si } p_0 \leq \frac{1}{2} \\ p_0 - \Delta & \text{si } p_0 > \frac{1}{2} \end{cases}, \quad q_0 = 1 - p_0, \quad q_1 = 1 - p_1$$

**En R :**

$H_0$	$H_1$	$x = \text{Número de éxitos}, \quad n = \text{número de pruebas}$ Regla de decisión: Rechazar $H_0$ si $p\text{-valor} < \alpha$
$\pi = \pi_0$	$\pi \neq \pi_0$	<code>binom.test(x,n,p=p0)</code>
$\pi \leq \pi_0$	$\pi > \pi_0$	<code>binom.test(x,n,p=p0,alternative="greater")</code>
$\pi \geq \pi_0$	$\pi < \pi_0$	<code>binom.test(x,n,p=p0,alternative="less")</code>

## 2. Contrastes con dos muestras.

---

### 2.1. Diferencia de medias de variables con distribuciones $N(\mu_1, \sigma_1)$ y $N(\mu_2, \sigma_2)$ .

#### 2.1.1. Muestras Independientes:

Varianzas distintas ( $\sigma_1^2 \neq \sigma_2^2$ )				
$H_0$	$H_1$	Tamaño de muestra	Estadístico de contraste	Región de rechazo de $H_0$
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$\frac{(z_{\alpha/2} + z_{\beta})^2 (s_1^2 + s_2^2)}{\Delta^2}$	$t_{exp} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$ t_{exp}  > t_{n, \alpha/2}$
$\mu_1 \leq \mu_2$	$\mu_1 > \mu_2$	$\frac{(z_{\alpha} + z_{\beta})^2 (s_1^2 + s_2^2)}{\Delta^2}$		$t_{exp} > t_{n, \alpha}$
$\mu_1 \geq \mu_2$	$\mu_1 < \mu_2$			$t_{exp} < -t_{n, \alpha}$

Varianzas iguales ( $\sigma_1^2 = \sigma_2^2$ ): Sólo hacer este contraste si hay razones fundadas para suponer que las varianzas de ambas variables son iguales.				
$H_0$	$H_1$	Tamaño de muestra	Estadístico de contraste	Región de rechazo de $H_0$
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$\frac{(z_{\alpha/2} + z_{\beta})^2 (2s^2)}{\Delta^2}$	$t_{exp} = \frac{\bar{X}_1 - \bar{X}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$ t_{exp}  > t_{n_1 + n_2 - 2, \alpha/2}$
$\mu_1 \leq \mu_2$	$\mu_1 > \mu_2$	$\frac{(z_{\alpha} + z_{\beta})^2 (2s^2)}{\Delta^2}$		$t_{exp} > t_{n_1 + n_2 - 2, \alpha}$
$\mu_1 \geq \mu_2$	$\mu_1 < \mu_2$			$t_{exp} < -t_{n_1 + n_2 - 2, \alpha}$

siendo:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}, \quad n = \text{REDONDEO} \left[ \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^2}{n_1}\right)^2 \frac{1}{n_1 - 1} + \left(\frac{s_2^2}{n_2}\right)^2 \frac{1}{n_2 - 1}} \right]$$

En R :

$H_0$	$H_1$	$x1$ =vector que contiene los datos de la primera variable. $x2$ =vector que contiene los datos de la segunda variable. Regla de decisión: Rechazar $H_0$ si $p\text{-valor} < \alpha$
$\mu = \mu_0$	$\mu \neq \mu_0$	<code>t.test(x1,x2)</code>
$\mu \leq \mu_0$	$\mu > \mu_0$	<code>t.test(x1,x2,alternative="greater")</code>
$\mu \geq \mu_0$	$\mu < \mu_0$	<code>t.test(x1,x2,alternative="less")</code>

NOTAS:

- En caso de que haya razones para pensar que las varianzas de ambas variables son iguales, añadir dentro del comando `t.test` la opción `var.equal=TRUE`.
- Si los datos se leen desde un archivo en el que se ha especificado una única variable  $x$  en dos muestras, definidas por los valores de la variable `grupo` (que puede tomar, por ejemplo, los valores 1 y 2), entonces puede utilizarse la sintaxis:

`t.test(x ~ grupo)`

especificando la alternativa si el test es unilateral.

### 2.1.2. Muestras Emparejadas:

$H_0$	$H_1$	Tamaño de muestra	Estadístico de contraste	Región de rechazo de $H_0$
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$\frac{(z_{\alpha/2} + z_{\beta})^2 s_D^2}{\Delta^2}$	$t_{exp} = \frac{\bar{X}_1 - \bar{X}_2}{S_D / \sqrt{n}}$	$ t_{exp}  > t_{n-1, \alpha/2}$
$\mu_1 \leq \mu_2$	$\mu_1 > \mu_2$	$\frac{(z_{\alpha} + z_{\beta})^2 s_D^2}{\Delta^2}$		$t_{exp} > t_{n-1, \alpha}$
$\mu_1 \geq \mu_2$	$\mu_1 < \mu_2$			$t_{exp} < -t_{n-1, \alpha}$

siendo:

$$S_D = \sqrt{\frac{\sum_{i=1}^n ((X_{1i} - X_{2i}) - (\bar{X}_1 - \bar{X}_2))^2}{n-1}} = \sqrt{S_1^2 + S_2^2 - 2S_{12}} = \sqrt{S_1^2 + S_2^2 - 2rS_1S_2}$$

**En R :**

$H_0$	$H_1$	$x1$ =vector que contiene los datos de la primera variable. $x2$ =vector que contiene los datos de la segunda variable. <b>Regla de decisión:</b> Rechazar $H_0$ si $p\text{-valor} < \alpha$
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	<code>t.test(x1,x2,paired=T)</code>
$\mu_1 \leq \mu_2$	$\mu_1 > \mu_2$	<code>t.test(x1,x2,paired=T,alternative="greater")</code>
$\mu_1 \geq \mu_2$	$\mu_1 < \mu_2$	<code>t.test(x1,x2,paired=T,alternative="less")</code>

## 2.2. Diferencia de tendencia central entre distribuciones no normales.

Cuando las distribuciones de las variables  $X_1$  y  $X_2$  que se comparan no son normales (y las muestras son pequeñas), se puede utilizar el test de Wilcoxon-Mann-Whitney. Este test asigna *rangos* a los valores de ambas variables y contrasta si ambas tienen la misma tendencia central, o una de ellas tiende a dar valores más altos que la otra.

**En R :**

Test de Wilcoxon-Mann-Whitney		
$H_0$	$H_1$	$x1$ =vector que contiene los datos de la primera variable. $x2$ =vector que contiene los datos de la segunda variable. Regla de decisión: Rechazar $H_0$ si $p\text{-valor} < \alpha$
$\mathcal{F}_1 = \mathcal{F}_2$	$\mathcal{F}_1 \neq \mathcal{F}_2$	<code>wilcox.test(x1,x2,paired=F)</code>
$\mathcal{F}_1 \leq \mathcal{F}_2$	$\mathcal{F}_1 > \mathcal{F}_2$	<code>wilcox.test(x1,x2,paired=F,alternative="greater")</code>
$\mathcal{F}_1 \geq \mathcal{F}_2$	$\mathcal{F}_1 < \mathcal{F}_2$	<code>wilcox.test(x1,x2,paired=F,alternative="less")</code>

**Notas:**

- Si los datos son independientes, utilizar `paired=F`, y si están emparejados `paired=T`.
- Si los datos se leen desde un archivo en el que se ha especificado una única variable  $x$  en dos muestras, definidas por los valores de la variable *grupo* (que puede tomar, por ejemplo, los valores 1 y 2), entonces puede utilizarse la sintaxis:

`wilcox.test(x ~ grupo)`

especificando la alternativa si el test es unilateral y, en su caso, el posible emparejamiento.

### 2.3. Cociente de varianzas $\sigma_1^2/\sigma_2^2$ de variables con distribución normal (muestras independientes).

$H_0$	$H_1$	Tamaño de muestra	Estadístico de contraste	Región de rechazo de $H_0$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 = \sigma_2^2$	$1 + 4 \left( \frac{t_{\alpha} + t_{2\beta}}{\ln \phi} \right)^2$	$F_{exp} = \frac{S_1^2}{S_2^2}$	$F_{exp} < F_{n_1-1, n_2-1, 1-\alpha/2}$ ó $F_{exp} > F_{n_1-1, n_2-1, \alpha/2}$
$\sigma_1^2 \leq \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	$1 + 4 \left( \frac{t_{2\alpha} + t_{4\beta}}{\ln \phi} \right)^2$		$F_{exp} > F_{n_1-1, n_2-1, \alpha}$
$\sigma_1^2 \geq \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$			$F_{exp} < F_{n_1-1, n_2-1, 1-\alpha}$

El tamaño de muestra es aproximado, siendo  $\phi = \frac{\sigma_1^2}{\sigma_2^2}$  el primer valor del cociente de varianzas para el que interesa que el contraste tenga una potencia  $1 - \beta$ .

#### En R :

$H_0$	$H_1$	$x1$ =vector que contiene los datos de la primera variable. $x2$ =vector que contiene los datos de la segunda variable. Regla de decisión: Rechazar $H_0$ si $p\text{-valor} < \alpha$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 = \sigma_2^2$	<code>var.test(x1,x2)</code>
$\sigma_1^2 \leq \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	<code>var.test(x1,x2,alternative="greater")</code>
$\sigma_1^2 \geq \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$	<code>var.test(x1,x2,alternative="less")</code>

Si los datos se leen desde un archivo en el que se ha especificado una única variable  $x$  en dos muestras, definidas por los valores de la variable `grupo` (que puede tomar, por ejemplo, los valores 1 y 2), entonces puede utilizarse la sintaxis:

`var.test(x ~ grupo)`

especificando la alternativa si el test es unilateral.



## 2.4. Diferencia de proporciones en muestras independientes.

Contraste aproximado si $n_1 \geq 30$ y $n_2 \geq 30$ , $n_1p^*, n_1q^*, n_2p^*, n_2q^* > 5$				
$H_0$	$H_1$	Tamaño de muestra	Estadístico de contraste	Región de rechazo de $H_0$
$\pi_1 = \pi_2$	$\pi_1 \neq \pi_2$	$\frac{1}{2} \left( \frac{z_{\alpha/2} + z_{\beta} \sqrt{1 - \Delta^2}}{\Delta} \right)^2$	$z_{exp} = \frac{p_1 - p_2}{\sqrt{p^*q^* \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$	$ z_{exp}  > z_{\alpha/2}$
$\pi_1 \leq \pi_2$	$\pi_1 > \pi_2$	$\frac{1}{2} \left( \frac{z_{\alpha} + z_{2\beta} \sqrt{1 - \Delta^2}}{\Delta} \right)^2$		$z_{exp} > z_{\alpha}$
$\pi_1 \geq \pi_2$	$\pi_1 < \pi_2$			$z_{exp} < -z_{\alpha}$

### NOTAS:

- $p_1$  y  $p_2$  son las respectivas proporciones muestrales.
- $p^* = \frac{n_1p_1 + n_2p_2}{n_1 + n_2}$ ,  $q^* = 1 - p^*$
- $\Delta$  es la diferencia mínima a detectar entre  $\pi_1$  y  $\pi_2$ .

**En R :** si se dispone de tamaños muestrales grandes:

Contraste aproximado si $n_1 \geq 30$ y $n_2 \geq 30$ , $n_1p^*, n_1q^*, n_2p^*, n_2q^* > 5$		
$H_0$	$H_1$	$(x1, x2)$ = Número de éxitos en cada una de las dos muestras. $(n1, n2)$ = Tamaños de las muestras. Regla de decisión: Rechazar $H_0$ si $p\text{-valor} < \alpha$
$\pi_1 = \pi_2$	$\pi_1 \neq \pi_2$	<code>prop.test(c(x1, x2), c(n1, n2))</code>
$\pi_1 \leq \pi_2$	$\pi_1 > \pi_2$	<code>prop.test(c(x1, x2), c(n1, n2), alternative="greater")</code>
$\pi_1 \geq \pi_2$	$\pi_1 < \pi_2$	<code>prop.test(c(x1, x2), c(n1, n2), alternative="less")</code>

En caso de que se disponga de una variable `result` con dos valores (éxito-fracaso), y otra variable `grupo` que especifica la muestra, puede emplearse la sintaxis:

`prop.test(table(result, grupo), alternative="...")`

especificando en cada caso la alternativa de interés (o no incluyendo el comando `alternative` si el test es bilateral).

**En R : si los tamaños muestrales son pequeños:**

Si no se dan las condiciones anteriores ( $n_1 \geq 30$  y  $n_2 \geq 30$ ,  $n_1p^*$ ,  $n_1q^*$ ,  $n_2p^*$ ,  $n_2q^* > 5$ ), puede emplearse el *test exacto de Fisher*:

<i>Test Exacto de Fisher</i>		
$H_0$	$H_1$	$x$ =Variable con dos valores (éxito y fracaso). $grupo$ =Variable que indica la muestra (1 ó 2). <b>Regla de decisión:</b> Rechazar $H_0$ si $p\text{-valor} < \alpha$
$\pi_1 = \pi_2$	$\pi_1 \neq \pi_2$	<code>fisher.test(table(x,grupo))</code>
$\pi_1 \leq \pi_2$	$\pi_1 > \pi_2$	<code>fisher.test(table(x,grupo),alternative="greater")</code>
$\pi_1 \geq \pi_2$	$\pi_1 < \pi_2$	<code>fisher.test(table(x,grupo),alternative="less")</code>