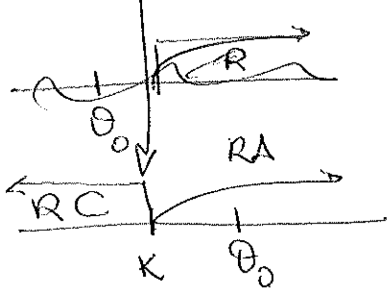
ESTAD_T26. HPTSIS COMPUESTAS y CUMP.
CONTRASTES de SIGNIFICACIÓN, p-valor.
CONTRASTE de RAZÓN de VEROSIMILITUDES.
CONTRASTES sobre $\mu$ y $\sigma^2$ en pobl. normales.
CONTRASTES en pobl. no normales.
MUESTRAS GRANDES

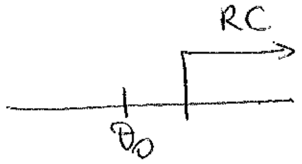
---

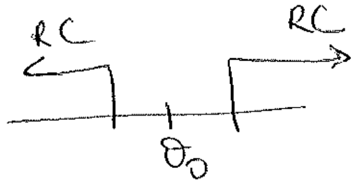
Hipótesis compuestas → $H_0, H_1$ o ambas incluye más de un valor para el parámetro.

Sp. $H_0: \theta = \theta_0$

$\underbrace{H_1: \theta < \theta_0 \qquad H_1: \theta > \theta_0}_{\text{unilat.}} \qquad \underbrace{H_1: \theta \neq \theta_0}_{\text{bilateral}}$




Contraste uniformemente más potente → Contraste de $H_0$ simple frente a $H_1$ alternativa, donde la RC es la mejor RC de tamaño $\alpha$ para contrastar $H_0$ frente a cada una de las hptsis simples que integran $H_1$.

Para los contrastes unilaterales y bilaterales, en general, no está asegurada la existencia de CUMP,

CONTRASTES de SIGNIFICACIÓN

$\left\{ \begin{array}{l} H_0: \theta = \theta_0 \\ H_1: \theta \neq \theta_0 \end{array} \right.$ | se contrasta si existen diferencias significativas entre lo que dice $H_0$ y la evidencia muestral.

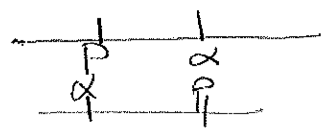
Estadístico: $D = \hat{\theta} - \theta_0$ (cantidad pivotal)

Decisión: si $D \geq d_\alpha$, rechazo $H_0$

$\boxed{\text{P-VALOR}}$ → menor nivel de significación con el que se rechaza $H_0$. Probab. a partir de la cual D cae en RC

p-valor $< \alpha$ ⇒ Rechazo $H_0$

p-valor $\geq \alpha$ ⇒ Acepto $H_0$



CONTRASTE de RAZÓN de VEROSIMILITUD:

Es un contraste de significación, donde la medida de discrepancia que se utiliza es una discrepancia relativa entre verosimilitudes.

$$\begin{cases} H_0: \theta \in \Omega_0 \\ H_1: \theta \in \Omega_1 = \Omega - \Omega_0 \end{cases}$$

Definimos el estadístico razón de verosimilitud:

$$\lambda(x) = \lambda(x_1 \ldots x_n) = \frac{\mathcal{L}_0^*(x,\theta)}{\mathcal{L}^*(x,\theta)} = \frac{\max_{\theta \in \Omega_0} \mathcal{L}(x,\theta)}{\max_{\theta \in \Omega} \mathcal{L}(x,\theta)} \quad / \; 0 \leq \lambda(x) \leq 1$$

Decisión: $\lambda(x) \sim 0 \Rightarrow$ rechazo $H_0$
$\lambda(x) \sim 1 \Rightarrow$ acepto $H_0$

Para poblaciones normales:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$$

$\sigma$ conoc $\rightarrow$ $|\bar{x} - \mu_0| > z_{\alpha/2} \cdot \sigma/\sqrt{n} \Rightarrow$ rechazo $H_0$

$\sigma$ desc $\rightarrow$ $|\bar{x} - \mu_0| > t_{\alpha/2} \cdot S/\sqrt{n-1} \Rightarrow$ rechazo $H_0$

$$\begin{cases} H_0: \mu \leq \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$$

$\sigma$ conoc $\rightarrow$ $\bar{x} - \mu_0 > z_\alpha \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow$ rechazo $H_0$

$\sigma$ desc $\rightarrow$ $\bar{x} - \mu_0 > t_\alpha \cdot \frac{S}{\sqrt{n-1}} \Rightarrow$ rechazo $H_0$

$$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 \neq \sigma_0^2 \end{cases}$$

$\mu$ conoc $\rightarrow$ $\sum(x_i - \mu)^2 > \sigma_0^2 \cdot \chi^2_{n,\alpha/2}$ ó $<$

$\mu$ desc $\rightarrow$ $\sum(x_i - \bar{x})^2 > \sigma_0^2 \chi^2_{n-1}$

$$\begin{cases} H_0: \sigma^2 \leq \sigma_0^2 \\ H_1: \sigma^2 > \sigma_0^2 \end{cases}$$

$\mu$ conoc $\rightarrow$ $\sum(x_i - \mu)^2 > \sigma_0^2 \chi^2_{n,1-\alpha}$ para un $\alpha$

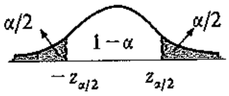
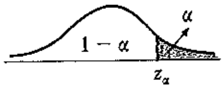
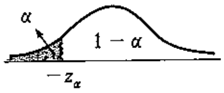
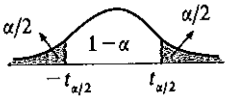
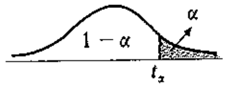
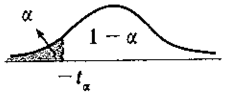
$\mu$ desc $\rightarrow$ $\sum(x_i - \bar{x})^2 > \sigma_0^2 \chi^2_{n-1,1-\alpha}$

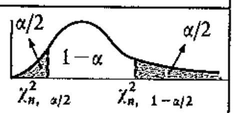
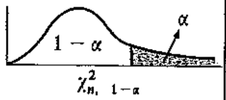
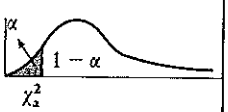
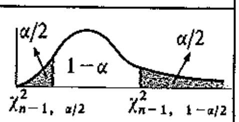
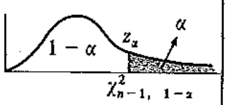
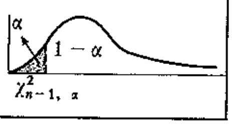
NO NORMALES

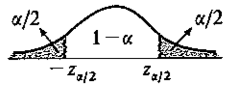
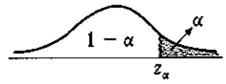
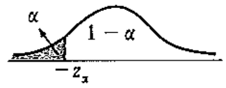
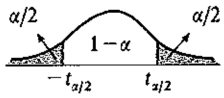
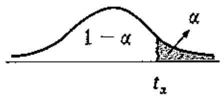
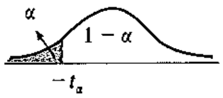
Para muestras sufic. grandes, se puede utilizar TCL

$$S = \frac{\sum(x_i - \bar{x})^2}{n}$$

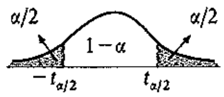
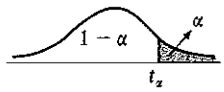
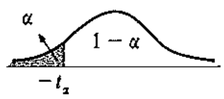
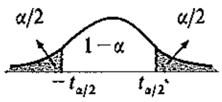
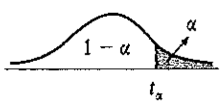
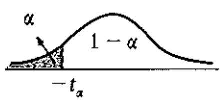
## 6.11. CUADRO RESUMEN DE LOS CONTRASTES DE HIPÓTESIS

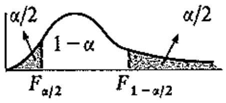
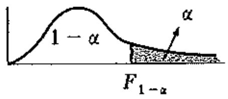
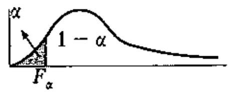
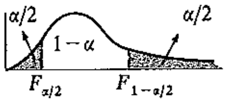
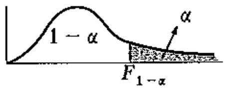
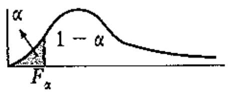
| Población | Hipótesis nula $H_0$ | Hipótesis alternativa $H_1$ | Tamaño de muestra | Estadístico | Distribución | Región crítica | |
|---|---|---|---|---|---|---|---|
| $N(\mu, \sigma)$ $\sigma$ conocida | $\mu = \mu_0$ | $\mu \neq \mu_0$ | $n$ | $z_{\text{exp}} = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ | $N(0, 1)$ | $z_{\text{exp}} < -z_{\alpha/2}$ y $z_{\text{exp}} > z_{\alpha/2}$ |  |
| | $\mu \leqslant \mu_0$ | $\mu > \mu_0$ | | | | $z_{\text{exp}} > z_\alpha$ |  |
| | $\mu \geqslant \mu_0$ | $\mu < \mu_0$ | | | | $z_{\text{exp}} < -z_\alpha$ |  |
| $N(\mu, \sigma)$ $\sigma$ desconocida | $\mu = \mu_0$ | $\mu \neq \mu_0$ | $n$ | $t_{\text{exp}} = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ <br> $s^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ | $t_{n-1}$ | $t_{\text{exp}} < -t_{\alpha/2}$ y $t_{\text{exp}} > t_{\alpha/2}$ |  |
| | $\mu \leqslant \mu_0$ | $\mu > \mu_0$ | | | | $t_{\text{exp}} > t_\alpha$ |  |
| | $\mu \geqslant \mu_0$ | $\mu < \mu_0$ | | | | $t_{\text{exp}} < -t_\alpha$ |  |

| Población | Hipótesis nula $H_0$ | Hipótesis alternativa $H_1$ | Tamaño de muestra | Estadístico | Distribución | Región crítica | |
|---|---|---|---|---|---|---|---|
| $N(\mu, \sigma)$ $\mu$ conocida | $\sigma^2 = \sigma_0^2$ | $\sigma^2 \neq \sigma_0$ | $n$ | $\chi^2_{\text{exp}} = \dfrac{\sum_{i=1}^{n} (x_i - \mu)^2}{\sigma_0^2}$ | $\chi^2_n$ | $\chi^2_{\text{exp}} < \chi^2_{n,\alpha/2}$ y $\chi^2_{\text{exp}} > \chi^2_{n,1-\alpha/2}$ |  |
| | $\sigma^2 \leqslant \sigma_0^2$ | $\sigma^2 > \sigma_0^2$ | | | | $\chi^2_{\text{exp}} > \chi^2_{n,1-\alpha}$ |  |
| | $\sigma^2 \geqslant \sigma_0^2$ | $\sigma^2 < \sigma_0^2$ | | | | $\chi^2_{\text{exp}} < \chi^2_{n,\alpha}$ |  |
| $N(\mu, \sigma)$ $\mu$ desconocida | $\sigma^2 = \sigma_0^2$ | $\sigma^2 \neq \sigma_0$ | $n$ | $\chi^2_{\text{exp}} = \dfrac{(n-1)s^2}{\sigma_0^2}$ <br> $s^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ | $\chi^2_{n-1}$ | $\chi^2_{\text{exp}} < \chi^2_{n-1,\alpha/2}$ y $\chi^2_{\text{exp}} > \chi^2_{n-1,\ 1-\alpha/2}$ |  |
| | $\sigma^2 \leqslant \sigma_0^2$ | $\sigma^2 > \sigma_0^2$ | | | | $\chi^2_{\text{exp}} > \chi^2_{n-1,\ 1-\alpha}$ |  |
| | $\sigma^2 \geqslant \sigma_0^2$ | $\sigma^2 < \sigma_0^2$ | | | | $\chi^2_{\text{exp}} < \chi^2_{n-1,\ \alpha}$ |  |

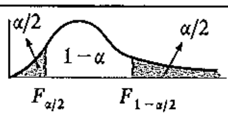
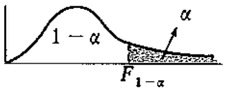
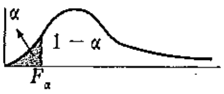
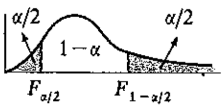
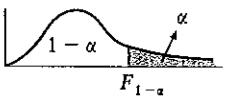
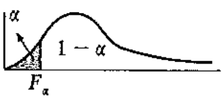
| Población | Hipótesis nula $H_0$ | Hipótesis alternativa $H_1$ | Tamaño de muestra | Estadístico | Distribución | Región crítica |
|---|---|---|---|---|---|---|
| $N(\mu_x, \sigma_x)$<br>$N(\mu_y, \sigma_y)$<br>$\sigma_x, \sigma_y$<br>conocidas | $\mu_x-\mu_y=d_0$<br>$\mu_x-\mu_y\leqslant d_0$<br>$\mu_x-\mu_y\geqslant d_0$ | $\mu_x-\mu_y\neq d_0$<br>$\mu_x-\mu_y>d_0$<br>$\mu_x-\mu_y<d_0$ | $n_x, n_y$ | $z_{\exp}=\dfrac{\bar{x}-\bar{y}-d_0}{\sqrt{\dfrac{\sigma_x^2}{n_x}+\dfrac{\sigma_y^2}{n_y}}}$ | $N(0, 1)$ | $z_{\exp}<-z_{\alpha/2}$ y $z_{\exp}>z_{\alpha/2}$  <br>$z_{\exp}>z_\alpha$  <br>$z_{\exp}<-z_\alpha$   |
| $N(\mu_x, \sigma_x)$<br>$N(\mu_y, \sigma_y)$<br>$\sigma_x, \sigma_y$<br>desconocidas<br>$\sigma_x^2=\sigma_y^2$ | $\mu_x-\mu_y=d_0$<br>$\mu_x-\mu_y\leqslant d_0$<br>$\mu_x-\mu_y\geqslant d_0$ | $\mu_x-\mu_y\neq d_0$<br>$\mu_x-\mu_y>d_0$<br>$\mu_x-\mu_y<d_0$ | $n_x, n_y$ | $t_{\exp}=\dfrac{\bar{x}-\bar{y}-d_0}{s'\sqrt{\dfrac{1}{n_x}+\dfrac{1}{n_y}}}$<br>$s'^2=\dfrac{(n_x-1)s_x^2+(n_y-1)s_y^2}{n_x+n_y-2}$ | $t_{n_x+n_y-2}$ | $t_{\exp}<-t_{\alpha/2}$ y $t_{\exp}>t_{\alpha/2}$  <br>$t_{\exp}>t_\alpha$  <br>$t_{\exp}<-t_\alpha$   |



| Población | Hipótesis nula $H_0$ | Hipótesis alternativa $H_1$ | Tamaño de muestra | Estadístico | Distribución | Región crítica |
|---|---|---|---|---|---|---|
| $N(\mu_x, \sigma_x)$<br>$N(\mu_y, \sigma_y)$<br>$\sigma_x, \sigma_y$<br>desconocidas<br>$\sigma_x^2\neq\sigma_y^2$ | $\mu_x-\mu_y=d_0$<br>$\mu_x-\mu_y\leqslant d_0$<br>$\mu_x-\mu_y\geqslant d_0$ | $\mu_x-\mu_y\neq d_0$<br>$\mu_x-\mu_y>d_0$<br>$\mu_x-\mu_y<d_0$ | $n_x, n_y$ | $t_{\exp}=\dfrac{\bar{x}-\bar{y}-d_0}{\sqrt{\dfrac{s_x^2}{n_x}+\dfrac{s_y^2}{n_y}}}$<br>$s_x^2=\dfrac{1}{n_x-1}\sum_{i=1}^{n_x}(x_i-\bar{x})^2$<br>$s_y^2=\dfrac{1}{n_y-1}\sum_{i=1}^{n_y}(y_i-\bar{y})^2$ | $t_\nu$<br>$\nu=\dfrac{\left(\dfrac{s_x^2}{n_x}+\dfrac{s_y^2}{n_y}\right)^2}{\dfrac{\left(\dfrac{s_x^2}{n_x}\right)^2}{n_x-1}+\dfrac{\left(\dfrac{s_y^2}{n_y}\right)^2}{n_y-1}}$ | $t_{\exp}<-t_{\alpha/2}$ y $t_{\exp}>t_{\alpha/2}$  <br>$t_{\exp}>t_\alpha$  <br>$t_{\exp}<-t_\alpha$   |
| $N(\mu_x, \sigma_x)$<br>$N(\mu_y, \sigma_y)$<br>muestras apareadas<br>$(X_i, Y_i)$ | $\mu_x-\mu_y=d_0$<br>$\mu_x-\mu_y\leqslant d_0$<br>$\mu_x-\mu_y\geqslant d_0$ | $\mu_x-\mu_y\neq d_0$<br>$\mu_x-\mu_y>d_0$<br>$\mu_x-\mu_y<d_0$ | $n_x, n_y$ | $t_{\exp}=\dfrac{\bar{d}-d_0}{s_d/\sqrt{n}}$<br>$\bar{d}=\dfrac{1}{n}\sum_{i=1}^{n}d_i$<br>$s_d^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(d_i-\bar{d})^2$<br>$d_i=x_i-y_i$ | $t_{n-1}$ | $t_{\exp}<-t_{\alpha/2}$ y $t_{\exp}>t_{\alpha/2}$  <br>$t_{\exp}>t_\alpha$  <br>$t_{\exp}<-t_\alpha$   |

| Población | Hipótesis nula $H_0$ | Hipótesis alternativa $H_1$ | Tamaño de muestra | Estadístico | Distribución | Región crítica |
|---|---|---|---|---|---|---|
| $N(\mu_x, \sigma_x)$<br>$N(\mu_y, \sigma_y)$<br>$\mu_x, \mu_y$<br>conocidas | $\sigma_x^2 = \sigma_y^2$<br>$\sigma_x^2 \leqslant \sigma_y^2$<br>$\sigma_x^2 \geqslant \sigma_y^2$ | $\sigma_x^2 \neq \sigma_y^2$<br>$\sigma_x^2 > \sigma_y^2$<br>$\sigma_x^2 < \sigma_y^2$ | $n_x, n_y$ | $F_{exp} = \dfrac{\frac{1}{n_x}\sum_{i=1}^{n_x}(x_i-\mu_x)^2}{\frac{1}{n_y}\sum_{i=1}^{n_y}(y_i-\mu_y)^2}$ | $F_{n_x, n_y}$ | $F_{exp} < F_{\alpha/2}$ y $F_{exp} > F_{1-\alpha/2}$  <br>$F_{exp} > F_{1-\alpha}$  <br>$F_{exp} < F_{\alpha}$   |
| $N(\mu_x, \sigma_x)$<br>$N(\mu_y, \sigma_y)$<br>$\mu_x, \mu_y$<br>desconocidas | $\sigma_x^2 = \sigma_y^2$<br>$\sigma_x^2 \leqslant \sigma_y^2$<br>$\sigma_x^2 \geqslant \sigma_y^2$ | $\sigma_x^2 \neq \sigma_y^2$<br>$\sigma_x^2 > \sigma_y^2$<br>$\sigma_x^2 < \sigma_y^2$ | $n_x, n_y$ | $F_{exp} = \dfrac{s_x^2}{s_y^2}$<br>$s_x^2 = \dfrac{1}{n_x-1}\sum_{i=1}^{n_x}(x_i-\bar{x})^2$<br>$s_y^2 = \dfrac{1}{n_y-1}\sum_{i=1}^{n_y}(y_i-\bar{y})^2$ | $F_{n_x-1, n_y-1}$ | $F_{exp} < F_{\alpha/2}$ y $F_{exp} > F_{1-\alpha/2}$  <br>$F_{exp} > F_{1-\alpha}$  <br>$F_{exp} < F_{\alpha}$   |



| Población | Hipótesis nula $H_0$ | Hipótesis alternativa $H_1$ | Tamaño de muestra | Estadístico | Distribución | Región crítica |
|---|---|---|---|---|---|---|
| $N(\mu_x, \sigma_x)$<br>$N(\mu_y, \sigma_y)$<br>$\mu_x$<br>conocida<br>$\mu_y$<br>desconocida | $\sigma_x^2 = \sigma_y^2$<br>$\sigma_x^2 \leqslant \sigma_y^2$<br>$\sigma_x^2 \geqslant \sigma_y^2$ | $\sigma_x^2 \neq \sigma_y^2$<br>$\sigma_x^2 > \sigma_y^2$<br>$\sigma_x^2 < \sigma_y^2$ | $n_x, n_y$ | $F_{exp} = \dfrac{\frac{1}{n_x}\sum_{i=1}^{n_x}(x_i-\mu_x)^2}{s_y^2}$ | $F_{n_x, n_y-1}$ | $F_{exp} < F_{\alpha/2}$ y $F_{exp} > F_{1-\alpha/2}$  <br>$F_{exp} > F_{1-\alpha}$  <br>$F_{exp} < F_{\alpha}$   |
| $N(\mu_x, \sigma_x)$<br>$N(\mu_y, \sigma_y)$<br>$\mu_x$<br>desconocida<br>$\mu_y$<br>conocida | $\sigma_x^2 = \sigma_y^2$<br>$\sigma_x^2 \leqslant \sigma_y^2$<br>$\sigma_x^2 \geqslant \sigma_y^2$ | $\sigma_x^2 \neq \sigma_y^2$<br>$\sigma_x^2 > \sigma_y^2$<br>$\sigma_x^2 < \sigma_y^2$ | $n_x, n_y$ | $F_{exp} = \dfrac{s_x^2}{\frac{1}{n_y}\sum_{i=1}^{n_y}(y_i-\mu_y)^2}$ | $F_{n_x-1, n_y}$ | $F_{exp} < F_{\alpha/2}$ y $F_{exp} > F_{1-\alpha/2}$  <br>$F_{exp} > F_{1-\alpha}$  <br>$F_{exp} < F_{\alpha}$   |

| Población | Hipótesis nula $H_0$ | Hipótesis alternativa $H_1$ | Tamaño de muestra | Estadístico | Distribución | Región crítica |
|---|---|---|---|---|---|---|
| $B(1, p)$ | $p = p_0$<br>$p \leqslant p_0$<br>$p \geqslant p_0$ | $p \neq p_0$<br>$p > p_0$<br>$p < p_0$ | $n$ | $z_{exp} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$ | $N(0, 1)$ | $z_{exp} < z_{\alpha/2}$ y $z_{exp} > z_{\alpha/2}$<br>$z_{exp} > z_\alpha$<br>$z_{exp} < -z_\alpha$ |
| $B(1, p_X)$<br>$B(1, p_Y)$ | $p_X = p_Y$<br>$p_X \leqslant p_Y$<br>$p_X \geqslant p_Y$ | $p_X \neq p_Y$<br>$p_X > p_Y$<br>$p_X < p_Y$ | $n_X$<br>$n_Y$ | $z_{exp} = \dfrac{\hat{p}_x - \hat{p}_y}{\sqrt{\dfrac{n_x + n_y}{n_x n_y} pq}}$<br>$\hat{p} = \dfrac{n_x \hat{p}_x + n_y \hat{p}_y}{n_x + n_y} = \dfrac{x + y}{n_x + n_y}$ | $N(0, 1)$ | $z_{exp} < -z_{\alpha/2}$ y $z_{exp} > z_{\alpha/2}$<br>$z_{exp} > z_\alpha$<br>$z_{exp} < z_\alpha$ |

# Capítulo 7
# CONTRASTES DE BONDAD DE AJUSTE Y TABLAS DE CONTINGENCIA

## 7.1. INTRODUCCIÓN

En los capítulos anteriores hemos estudiado la estimación tanto puntual como por intervalos, y también nos hemos ocupado del contraste de hipótesis paramétrico, estudiando algunos contrastes con cierto detalle.

En este capítulo podíamos haber realizado una amplia introducción sobre los contrastes no paramétricos, pero por razones didácticas hemos preferido dejar esa introducción para el capítulo siguiente, aunque los tests que aquí vamos a estudiar pueden considerarse como no paramétricos, como indicaremos con más detalle también en la introducción del capítulo siguiente.

Hasta ahora hemos utilizado la distribución $\chi^2$ de Pearson para hacer inferencias sobre la varianza poblacional, aquí, sin embargo, utilizaremos este estadístico $\chi^2$ de Pearson para:

1. Contrastar si una supuesta distribución se ajusta a un conjunto de datos. **Contrastes de bondad de ajuste.**
2. Contrastar si existe dependencia entre dos características de la misma población. **Contraste de independencia.**
3. Contrastar si varias muestras proceden de la misma población. **Contraste de homogeneidad.**