1 Test para la media de una población normal

Sean $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ una muestra aleatoria (v.a.i.i.d. variables aleatorias independientes idénticamente distribuidas). Se quiere testear

$$H_0 : \mu = \mu_0 \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : \mu \neq \mu_0 \qquad H_1 : \mu > \mu_0 \qquad H_1 : \mu < \mu_0$$

1.1 Caso varianza conocida

Estadístico del test

$$\frac{\overline{X} - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} \underset{\substack{\uparrow \\ \text{bajo } H_0}}{\sim} N(0,1)$$

El Statistix no hace este test.

1.2 Caso varianza desconocida

Estadístico del test

$$\frac{\overline{X} - \mu_0}{\sqrt{\frac{s^2}{n}}} \underset{\substack{\uparrow \\ \text{bajo } H_0}}{\sim} t_{n-1}$$

donde

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 \tag{1}$$

es la varianza muestral de las observaciones. Instrucciones para el Statistix: Statistics→One, Two, Multi-sample tests→One-Sample T-Test.

2 Test para la varianza de una población normal

Sean $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ una muestra aleatoria (v.a.i.i.d. variables aleatorias independientes idénticamente distribuidas). Se quiere testear

$$H_0 : \sigma = \sigma_0 \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : \sigma \neq \sigma_0 \qquad H_1 : \sigma > \sigma_0 \qquad H_1 : \sigma < \sigma_0$$

y, la media poblacional μ es desconocida. Estadístico del test

$$(n-1)\frac{s^2}{\sigma_0^2} \underset{\substack{\uparrow \\ \text{bajo } H_0}}{\sim} \chi^2_{n-1}$$

donde s^2 está definido en (1). El Statistix no hace este test.

3 Test asintótico (o aproximado) para la media de una población

Sean $X_1, \ldots, X_n \sim F$ una muestra aleatoria donde F es una distribución no necesariamente normal y n es un número suficientemente grande. Sea $\mu = E(X_1)$. Se quiere testear

$$H_0 : \mu = \mu_0 \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : \mu \neq \mu_0 \qquad H_1 : \mu > \mu_0 \qquad H_1 : \mu < \mu_0$$

Estadístico del test

$$\frac{\overline{X} - \mu_0}{\sqrt{\frac{s^2}{n}}} \underset{\substack{\uparrow \\ \text{bajo } H_0}}{\overset{\text{aprox}}{\sim}} N(0,1)$$

donde s^2 es la varianza muestral definida en (1).

4 Test asintótico para una proporción poblacional

Sean $X_1, \ldots, X_n \sim Bi(1,p)$ una muestra aleatoria con n suficientemente grande. Se quiere testear

$$H_0 : p = p_0 \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : p \neq p_0 \qquad H_1 : p > p_0 \qquad H_1 : p < p_0$$

Estadístico del test

$$\frac{\widehat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \underset{\substack{\uparrow \\ \text{bajo } H_0}}{\overset{\text{aprox}}{\sim}} N(0,1)$$

donde $\widehat{p} = \overline{X}$ es la proporción de éxitos en la muestra. Instrucciones en el Statistix: Statistics→One, Two, Multi-sample tests→Proportion Test.

5 Test para comparar medias de dos poblaciones normales independientes

Sean $X_1, \ldots, X_{n_1} \sim N(\mu_1, \sigma_1^2)$, e $Y_1, \ldots, Y_{n_2} \sim N(\mu_2, \sigma_2^2)$ dos muestras aleatorias independientes entre sí. Se quiere testear

$$H_0 : \mu_1 - \mu_2 = \delta \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : \mu_1 - \mu_2 \neq \delta \qquad H_1 : \mu_1 - \mu_2 > \delta \qquad H_1 : \mu_1 - \mu_2 < \delta$$

5.1 Caso varianzas conocidas

Estadístico del test

$$\frac{\overline{X} - \overline{Y} - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \underset{\substack{\uparrow \\ \text{bajo } H_0}}{\sim} N(0,1)$$

5.2 Caso varianzas desconocidas pero iguales

Estadístico del test

$$\frac{\overline{X}-\overline{Y}-\delta}{\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \underset{\text{bajo } H_0}{\underset{\uparrow}{\sim}} t_{n_1+n_2-2}$$

donde

$$\begin{aligned} s_p^2 &= \frac{s_1^2\left(n_1-1\right)+s_2^2\left(n_2-1\right)}{n_1+n_2-2} \\ &= \frac{1}{n_1+n_2-2}\left(\sum_{i=1}^{n_1}(X_i-\overline{X})^2+\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2\right) \end{aligned}$$

y s_1^2 y s_2^2 son las varianzas muestrales de las *X's* y las *Y's* respectivamente, es decir

$$s_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_i-\overline{X})^2 \tag{2}$$

$$s_2^2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2 \tag{3}$$

5.3 Caso varianzas desconocidas y no necesariamente iguales

En este caso tenemos el test de Welch que tiene nivel aproximado. Estadístico del test

$$\frac{\overline{X}-\overline{Y}-\delta}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \underset{\text{bajo } H_0}{\underset{\uparrow}{\overset{\text{aprox}}{\sim}}} t_K$$

donde

$$K=\left[\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}\right]$$

y s_1^2 y s_2^2 están definidos en (2) y (3) respectivamente.

Estos tres tests y el que sigue para comarar varianzas los realiza el Statistix bajo la instrucción: Statistics→One, Two, Multi-sample tests→Two Sample T-Test.

6 Test para comparar varianzas de dos poblaciones normales independientes

Sean $X_1,\ldots,X_{n_1} \sim N\left(\mu_1,\sigma_1^2\right)$, e $Y_1,\ldots,Y_{n_2} \sim N\left(\mu_2,\sigma_2^2\right)$ dos muestras aleatorias independientes entre sí. Se quiere testear

$$\begin{aligned} H_0 &: \sigma_1^2=\sigma_2^2 \\ H_1 &: \sigma_1^2\neq\sigma_2^2 \end{aligned}$$

Estadístico del test

$$\frac{s_1^2}{s_2^2} \underset{\text{bajo } H_0}{\underset{\uparrow}{\sim}} F_{n_1-1,n_2-1}$$

donde s_1^2 y s_2^2 están definidos en (2) y (3) respectivamente.

7 Test asintótico (o aproximado) para comparar las medias de dos poblaciones

Sean $X_1, \ldots, X_{n_1}$ v.a.i.i.d. y sean $\mu_1 = E(X_1)$ y $\sigma_1^2 = V(X_1)$, e $Y_1, \ldots, Y_{n_2}$ v.a.i.i.d. independientes de las anteriores, y sean $\mu_2 = E(Y_1)$ y $\sigma_2^2 = V(Y_1)$, donde n_1 y n_2 son números suficientemente grandes. Se quiere testear

$$H_0 : \mu_1 - \mu_2 = \delta \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : \mu_1 - \mu_2 \neq \delta \qquad H_1 : \mu_1 - \mu_2 > \delta \qquad H_1 : \mu_1 - \mu_2 < \delta$$

Estadístico del test

$$\frac{\overline{X} - \overline{Y} - \delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \underset{\underset{\text{bajo } H_0}{\uparrow}}{\overset{\text{aprox}}{\sim}} N(0,1)$$

y s_1^2 y s_2^2 son las varianzas muestrales de las X's y las Y's respectivamente, y están dadas por (2) y (3).

8 Test asintótico (o aproximado) para comparar dos proporciones poblacionales

Sean $X_1, \ldots, X_{n_1} \sim Bi(1, p_1)$ v.a.i.i.d. e $Y_1, \ldots, Y_{n_2} \sim Bi(1, p_2)$ v.a.i.i.d. independiente de las anteriores, con n_1 y n_2 suficientemente grandes. Se quiere testear

$$H_0 : p_1 - p_2 = \delta \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : p_1 - p_2 \neq \delta \qquad H_1 : p_1 - p_2 > \delta \qquad H_1 : p_1 - p_2 < \delta$$

Estadístico del test

$$\frac{\widehat{p}_1 - \widehat{p}_2 - \delta}{\sqrt{\frac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}}} \underset{\underset{\text{bajo } H_0}{\uparrow}}{\overset{\text{aprox}}{\sim}} N(0,1)$$

donde $\widehat{p}_1 = \overline{X}$ y $\widehat{p}_2 = \overline{Y}$ son, respectivamente, las proporciones de éxitos en la primer y segunda muestra. Instrucciones en el Statistix: Statistics→One, Two, Multi-sample tests→Proportion Test→Two sample test

9 Test para comparar medias de dos muestras apareadas

Sean $(X_1, Y_1), \ldots, (X_n, Y_n)$ una muestra aleatoria de los resultados de dos mediciones realizadas sobre la misma unidad experimental (o que puede considerarse la misma unidad experimental). Supongamos que las diferencias $D_i = X_i - Y_i \sim N(\mu_D, \sigma_D^2)$ independientes entre sí. Se quiere testear

$$H_0 : \mu_D = \delta \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : \mu_D \neq \delta \qquad H_1 : \mu_D > \delta \qquad H_1 : \mu_D < \delta$$

Estadístico del test

$$\frac{\overline{X} - \overline{Y} - \delta}{\sqrt{\frac{s_D^2}{n}}} \underset{\underset{\text{bajo } H_0}{\uparrow}}{\sim} t_{n-1}$$

donde s_D es el desvío estándar de las D_i, es decir

$$s_D^2 = \frac{1}{n-1} \sum_{i=1}^{n} (D_i - \overline{D})^2.$$

Instrucción para el Statistix: Statistics→One, Two, Multi-sample tests→Paired T-Test.

10 Tests no paramétricos para una y dos muestras

10.1 Test de Mann-Whitney-Wilcoxon para dos muestras independientes

Sean $X_1, \ldots, X_{n_1} \sim F_1$, e $Y_1, \ldots, Y_{n_2} \sim F_2$ dos muestras aleatorias independientes entre sí, donde F_1 y F_2 son dos distribuciones continuas no necesariamente normales. Se quiere testear

$$\begin{aligned} H_0 &: F_1 = F_2 \\ H_1 &: F_1 \neq F_2. \end{aligned}$$

Para construir el estadístico se juntan ambas muestras y se ranquean las observaciones, a esos números se los llaman rangos de las observaciones. El estadístico del test es la suma de los rangos del grupo correspondiente al grupo con menor cantidad de observaciones. Puede calcularse (con la tabla apropiada) el p-valor exacto de este test. El Statistix calcula el p-valor exacto cuando las muestras son pequeñas y siempre calcula el p-valor aproximado: Statistics→One, two and multi-sample tests→Wilcoxon Rank sum test.

10.2 Test del signo para la mediana de una población

Sea $X_1, \ldots, X_n \sim F$ una muestra aleatoria donde F es una distribución continua. Sea $\widetilde{\mu}$ la mediana de F. Se quiere testear

$$\begin{aligned} H_0 &: \widetilde{\mu} = m_0 \\ H_1 &: \widetilde{\mu} \neq m_0 \end{aligned}$$

Sean $D_i = X_i - m_0$. El estadístico del test es $U =$ la cantidad de D_i positivos.

$$U \underset{\text{bajo } H_0}{\underset{\uparrow}{\sim}} Bi\left(n-k, \frac{1}{2}\right).$$

donde k es la cantidad de D_i iguales a cero que hay en la muestra. Este test se puede hacer directamente en el Statistix, tener en cuenta que da el p-valor a una cola. Instrucciones del Statistix: Statistics→One, two and multi-sample tests→Sign test.

10.3 Test de rangos signados de Wilcoxon para la mediana de una población

Sea $X_1, \ldots, X_n \sim F$ una muestra aleatoria y F una distribución continua y simétrica. Sea $\widetilde{\mu}$ la mediana de F. Se quiere testear

$$\begin{aligned} H_0 &: \widetilde{\mu} = m_0 \text{ versus} \\ H_1 &: \widetilde{\mu} \neq m_0 \end{aligned}$$

Sean $D_i = X_i - m_0$. Luego se ordenan en orden creciente las observaciones $|D_i|$ y se las ranquea. Sean

$$\begin{aligned} W^+ &= \text{ suma de los rangos de las } D_i \text{ positivas} \\ W^- &= \text{ suma de los rangos de las } D_i \text{ negativas} \end{aligned}$$

El estadístico del test es $U = W^+ - W^-$. Puede calcularse (con la tabla apropiada) el p-valor exacto de este test. El Statistix calcula el p-valor exacto para tests de una cola y el aproximado para el de dos colas. Cuando el tamaño de muestra es suficientemente grande sólo calcula el p-valor aproximado: Statistics→One, two and multi-sample tests→Wilcoxon Signed Rank Test.

Cuando se tiene una muestra apareada $(X_1, Y_1), \ldots, (X_n, Y_n)$ para testear si las medianas de ambas poblaciones coinciden se puede aplicar el test de rangos signados de Wilcoxon o el test del signo sobre las $D_i = X_i - Y_i$.