# 1 Test para comparar medias 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

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

## 1.1 Caso varianzas conocidas

Estadístico del test

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

## 1.2 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

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

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

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

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

## 1.3 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 (1) y (2) respectivamente.

## 2 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

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2 \qquad H_1 : \sigma_1^2 > \sigma_2^2 \qquad H_1 : \sigma_1^2 < \sigma_2^2$$

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 (1) y (2) respectivamente.

## 3 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 $D_i = X_i - Y_i \sim N\left(\mu_D, \sigma_D^2\right)$ son 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

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

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

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

## 4 Test asintótico (o aproximado) para comparar 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. e 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

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

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 (1) y (2).

# 5 Test asintótico (o aproximado) para comparar dos proporciones

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. e independientes 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

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

donde $\widehat{p}_1 = \overline{X}$ y $\widehat{p}_2 = \overline{Y}$ son, respectivamente, las proporciones de éxitos en la primer y segunda muestra.

# 6 Tests no paramétricos para una y dos muestras

## 6.1 Test del signo para la mediana de una población

Sea $X_1, \ldots, X_n \sim F$ una muestra aleatoria donde $F$ es una función de distribución continua. Sea $\theta$ la mediana de $F$. Se quiere testear

$$H_0 : \theta = \theta_0 \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : \theta \neq \theta_0 \qquad H_1 : \theta > \theta_0 \qquad H_1 : \theta < \theta_0$$

Sean $D_i = X_i - \theta_0$. El estadístico del test es $U =$ la cantidad de $D_i$ positivos.

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

donde $k$ es la cantidad de $D_i$ iguales a cero que hay en la muestra.

## 6.2 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 $\theta$ la mediana de $F$. Se quiere testear

$$H_0 : \theta = \theta_0 \text{ versus alguna de las alternativas siguientes}$$
$$H_1 : \theta \neq \theta_0 \qquad H_1 : \theta > \theta_0 \qquad H_1 : \theta < \theta_0$$

Sean $D_i = X_i - \theta_0$. Luego se ordenan en orden creciente las observaciones $|D_i|$ y se las ranquea.
El estadístico del test es

$$W = \text{ suma de los rangos de las } D_i \text{ positivas.}$$

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 $D_i = X_i - Y_i$.

## 6.3 Test de Mann-Whitney-Wilcoxon para dos muestras independientes

### 6.3.1 Modelo 1:

Sean $X_1, \ldots, X_{n_1} \sim F$, e $Y_1, \ldots, Y_{n_2} \sim G$ dos muestras aleatorias independientes entre sí, donde $F$ y $G$ son dos ditribuciones continuas que satisfacen que $G(x) = F(x+c)$. Sean $\theta_X$ la mediana de $F$ y $\theta_Y$ la mediana de G. Se quiere testear

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

### 6.3.2 Modelo 2:

Sean $X_1, \ldots, X_{n_1} \sim F$, e $Y_1, \ldots, Y_{n_2} \sim G$ dos muestras aleatorias independientes entre sí, donde $F$ y $G$ son dos distribuciones continuas cualesquiera. Se quiere testear

$$\begin{aligned} &H_0 : F(x) = G(x) \ \forall x \text{ vs.} \\ &H_1 : \exists x \ / \ F(x) \neq G(x) \end{aligned}$$

### 6.3.3 Estadístico del test

Para construirlo 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 con menor cantidad de observaciones.

## 6.4 Test de la Mediana para dos muestras independientes

Sean $X_1, \ldots, X_{n_1} \sim F$, e $Y_1, \ldots, Y_{n_2} \sim G$ dos muestras aleatorias independientes entre sí, donde $F$ y $G$ son dos distribuciones continuas cualesquiera. Sean $\theta_X$ la mediana de $F$ y $\theta_Y$ la mediana de G. Se quiere testear

$$\begin{aligned} &H_0 : \theta_X = \theta_Y \text{ vs.} \\ &H_1 : \theta_X \neq \theta_Y \end{aligned}$$

Si $T$ es el estadístico correspondiente, entonces

$$T \underset{\substack{\uparrow \\ \text{bajo } H_0}}{\sim} \chi^2_1.$$