5 Contraste de hipótesis

- Nos plantearemos hipótesis y tendremos que rechazarlas o no a partir de la información que proporciona la muestra.
- La hipótesis a validar se denomina *hipótesis nula* H_0 y la enfrentamos a la *hipótesis alternativa* H_1, dando lugar a un contraste *unilateral* o *bilateral*.
- Si los datos no muestran suficiente evidencia para rechazar la hipótesis nula esta se mantiene como cierta (diremos que no hay suficiente evidencia para rechazarla evitando decir que la aceptamos)
- Por ser las características de la población desconocidas, podemos cometer dos tipos de errores: *error de tipo I* (si decidimos rechazar H_0 cuando es cierta) y *error de tipo II* (si decidimos aceptar H_0 cuando es falsa)
- A la probabilidad de cometer un error de tipo I se le denomina *nivel de significación del contraste* y se denota por $\alpha = P(\text{error tipo I}) = P(\text{rechazar } H_0 | H_0 \text{ cierta})$. Así, la probabilidad de aceptar H_0 cuando es cierta vale $P(\text{aceptar } H_0 | H_0 \text{ cierta}) = 1 - \alpha$
- La probabilidad de cometer un error de tipo II se denota por $\beta = P(\text{error tipo II}) = P(\text{aceptar } H_0 | H_0 \text{ falsa})$. Así, a la probabilidad de rechazar H_0 cuando es falsa se le denomina *potencia del contraste* y vale $P(\text{rechazar } H_0 | H_0 \text{ falsa}) = 1 - \beta$
- El nivel a partir del cual se pasa de aceptar la hipótesis nula H_0 a rechazarla se denomina *nivel crítico del contraste* o *p-valor del contraste*. Así, para cualquier contraste, la regla de decisión es *Rechazar* H_0 *si p-valor* $< \alpha$
- En la práctica fijaremos la probabilidad de cometer un error de tipo I (nivel de significación) $\alpha \in \{0.1, 0.05, 0.01\}$

5.1 Contrastes para una población normal

- Contrastes para la media de una población normal con varianza conocida: Sea $X_1, ...X_n$ una m.a.s de una población normal $\mathcal{N}(\mu, \sigma^2)$ con varianza conocida.

$$\text{Unilateral} \quad \left| \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{array} \right. \quad \Longrightarrow^{(\star)} \quad \text{Rechazar } H_0 \text{ si } \frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} > z_\alpha$$

$$\text{Bilateral} \quad \left| \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array} \right. \quad \Longrightarrow^{(\star)} \quad \text{Rechazar } H_0 \text{ si } \frac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}} > z_{\alpha/2}$$

$^{(\star)}$ Sabemos que bajo H_0 (es decir, si H_0 es cierta) $\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$. Así, basta fijar el nivel de significación α para obtener la regla del contraste.

- Contrastes para la media de una población normal con varianza desconocida: Sea $X_1, ...X_n$ una m.a.s de una población normal $\mathcal{N}(\mu, \sigma^2)$ con varianza desconocida.

$$\text{Unilateral} \quad \left| \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{array} \right. \quad \Longrightarrow^{(\star)} \quad \text{Rechazar } H_0 \text{ si } \frac{\overline{x}-\mu_0}{s/\sqrt{n}} > t_{n-1,\alpha}$$

$$\text{Bilateral} \quad \left| \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array} \right. \quad \Longrightarrow^{(\star)} \quad \text{Rechazar } H_0 \text{ si } \frac{|\overline{x}-\mu_0|}{s/\sqrt{n}} > t_{n-1,\alpha/2}$$

$^{(\star)}$ Bajo H_0 se cumple $\frac{\overline{X}-\mu_0}{S/\sqrt{n}} \sim t_{n-1}$

- Contrastes para la varianza de una población normal: Sea $X_1, ...X_n$ una m.a.s de una $\mathcal{N}(\mu, \sigma^2)$.

Unilateral $\left| \begin{array}{l} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 < \sigma_0^2 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{n-1}{\sigma_0^2} s^2 < \mathcal{X}_{n-1,1-\alpha}$

Bilateral $\left| \begin{array}{l} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\{\frac{n-1}{\sigma_0^2} s^2 > \mathcal{X}_{n-1,\alpha/2}$ ó $\frac{n-1}{\sigma_0^2} s^2 < \mathcal{X}_{n-1,1-\alpha/2}\}$

$^{(\star)}$ Bajo H_0 se cumple $\frac{n-1}{\sigma_0^2} S^2 \sim \mathcal{X}_{n-1}^2$

5.2 Contrastes para dos poblaciones normales

- Contrastes para la diferencia de medias de dos poblaciones normales basado en datos pareados.

Unilateral $\left| \begin{array}{l} H_0 : \mu_1 - \mu_2 = d_0 \\ H_1 : \mu_1 - \mu_2 > d_0 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{\overline{d}-d_0}{s_D/\sqrt{n}} > t_{n-1,\alpha}$

Bilateral $\left| \begin{array}{l} H_0 : \mu_1 - \mu_2 = d_0 \\ H_1 : \mu_1 - \mu_2 \neq d_0 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{|\overline{d}-d_0|}{s_D/\sqrt{n}} > t_{n-1,\alpha/2}$

$^{(\star)}$ Bajo H_0 se cumple que, si definimos $D = X - Y$, tenemos $\frac{\overline{D}-(\mu_1-\mu_2)}{S_D/\sqrt{n}} \sim t_{n-1}$

- Contrastes para la diferencia de medias de dos poblaciones normales con varianzas conocidas basado en muestras independientes.

Unilateral $\left| \begin{array}{l} H_0 : \mu_1 - \mu_2 = d_0 \\ H_1 : \mu_1 - \mu_2 > d_0 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{\overline{x}-\overline{y}-d_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} > z_\alpha$

Bilateral $\left| \begin{array}{l} H_0 : \mu_1 - \mu_2 = d_0 \\ H_1 : \mu_1 - \mu_2 \neq d_0 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{|\overline{x}-\overline{y}-d_0|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} > z_{\alpha/2}$

$^{(\star)}$ Bajo H_0 se cumple que $\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim \mathcal{N}(0,1)$

- Contrastes para la diferencia de medias de dos poblaciones normales con varianzas desconocidas pero iguales basado en muestras independientes.

Unilateral $\left| \begin{array}{l} H_0 : \mu_1 - \mu_2 = d_0 \\ H_1 : \mu_1 - \mu_2 > d_0 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{\overline{x}-\overline{y}-d_0}{\sqrt{s_{12}^2 \frac{n_1+n_2}{n_1 n_2}}} > t_{n_1+n_2-2,\alpha}$

Bilateral $\left| \begin{array}{l} H_0 : \mu_1 - \mu_2 = d_0 \\ H_1 : \mu_1 - \mu_2 \neq d_0 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{|\overline{x}-\overline{y}-d_0|}{\sqrt{s_{12}^2 \frac{n_1+n_2}{n_1 n_2}}} > t_{n_1+n_2-2,\alpha/2}$

$^{(\star)}$ Bajo H_0 se cumple que $\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{S_{12}^2 \frac{n_1 n_2}{n_1 n_2}}} \sim t_{n_1+n_2-2}$ siendo $S_{12}^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

- Contrastes para la diferencia de varianzas de dos poblaciones normales basado en muestras independientes.

Unilateral $\left| \begin{array}{l} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 < \sigma_2^2 \end{array} \right.$ $\Longrightarrow^{(\star)(\star\star)}$ Rechazar H_0 si $\frac{s_1^2}{s_2^2} < F_{n_1-1,n_2-1,1-\alpha}$

Bilateral $\left| \begin{array}{l} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 \neq \sigma_2^2 \end{array} \right.$ $\Longrightarrow^{(\star)(\star\star)}$ Rechazar H_0 si $\{\frac{s_1^2}{s_2^2} < F_{n_1-1,n_2-1,1-\alpha/2}$ ó $\frac{s_1^2}{s_2^2} > F_{n_1-1,n_2-1,\alpha/2}\}$

$^{(\star)}$ Bajo H_0 se cumple que $(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} =)\frac{S_1^2}{S_2^2} \sim F_{n_1-1,n_2-1}$

$^{(\star\star)}$ La distribución $F_{n,m}$ se define como $F_{n,m} = \frac{\frac{\mathcal{X}_n^2}{n}}{\frac{\mathcal{X}_m^2}{m}}$ donde $\mathcal{X}_n^2$ y $\mathcal{X}_m^2$ son distribuciones chi-cuadrado independientes.

5.3 Contrastes para muestras grandes

- Observación: Especificaremos en estas notas breves sólo los contrastes bilaterales.
- Contraste bilateral para una proporción.

Bilateral $\left| \begin{array}{l} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{|\hat{p}-p_0|}{\sqrt{p_0(1-p_0)/n}} > z_{\alpha/2}$

$^{(\star)}$ Bajo H_0 se cumple que $\frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} \sim \mathcal{N}(0,1)$

- Contraste bilateral para la diferencia de dos proporciones.

Bilateral $\left| \begin{array}{l} H_0 : p_1 - p_2 = 0 \\ H_1 : p_1 - p_2 \neq 0 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{|\hat{p}_1-\hat{p}_2|}{\sqrt{q(1-q)/n_1+q(1-q)/n_2}} > z_{\alpha/2}$

$^{(\star)}$ Bajo H_0 se cumple que $\frac{\hat{p}_1-\hat{p}_2}{\sqrt{q(1-q)/n_1+q(1-q)/n_2}} \sim \mathcal{N}(0,1)$, siendo $q = \frac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2}$

- Contraste bilateral para la media de cualquier distribución.

Bilateral $\left| \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{|\overline{x}-\mu_0|}{s/\sqrt{n}} > z_{\alpha/2}$

$^{(\star)}$ Bajo H_0 se cumple que $\frac{\overline{x}-\mu_0}{S/\sqrt{n}} \longrightarrow_{n\to\infty} \mathcal{N}(0,1)$

- Contraste bilateral para un parámetro θ de cualquier distribución.

Bilateral $\left| \begin{array}{l} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{array} \right.$ $\Longrightarrow^{(\star)}$ Rechazar H_0 si $\frac{|\hat{\theta}_{MV}-\theta_0|}{\sqrt{\frac{-1}{L''(\hat{\theta}_{MV})}}} > z_{\alpha/2}$

$^{(\star)}$ Bajo H_0 se cumple que $\hat{\theta}_{MV} \longrightarrow_{n\to\infty} \mathcal{N}(\theta_0, \frac{-1}{L''(\hat{\theta}_{MV})})$