# TEMA 4. INFERENCIA ESTADÍSTICA CON DOS MUESTRAS

## 1. INFERENCIA SOBRE LA DIFERENCIA DE MEDIAS EN POBLACIONES NORMALES CON MUESTRAS INDEPENDIENTES



### 1.1. VARIANZAS CONOCIDAS

· Tenemos:

- $X_1, X_2, \ldots, X_n$ que es una muestra de $X \sim N(\mu_x, \sigma_x^2)$
- $Y_1, Y_2, \ldots, Y_n$ que es una muestra de $Y \sim N\left(\mu_Y, \sigma_y^2\right)$
- $X_i$ es independiente de $Y_j$ para cualquier i, j.

· Entonces: $\overline{X} - \overline{Y} \sim N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}\right)$

· Tipificando obtenemos: $\frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}}} \sim N(0,1)$

· Con lo anterior podemos deducir que el **intervalo de confianza de nivel 100(1-α)% para la diferencia de medias $\mu_X - \mu_Y$** es:

$$\left(\overline{x} - \overline{y} - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}, \overline{x} - \overline{y} + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}\right)$$

· **Contrate de hipótesis** para diferencia de medias con varianza conocida:

- Estadístico de contraste:

$$T = \frac{\overline{X} - \overline{Y} - d_0}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} \sim N(0,1) \text{ bajo } H_0$$

- Hipótesis:

a) $\begin{matrix} H_0: \mu_X - \mu_Y = d_0 \\ H_1: \mu_X - \mu_Y \neq d_0 \end{matrix}$ b) $\begin{matrix} H_0: \mu_X - \mu_Y = d_0 \\ H_1: \mu_X - \mu_Y > d_0 \end{matrix}$ c) $\begin{matrix} H_0: \mu_X - \mu_Y = d_0 \\ H_1: \mu_X - \mu_Y < d_0 \end{matrix}$

- Región crítica:

a) $C = \left(-\infty, -z_{\frac{\alpha}{2}}\right) \cup \left(z_{\frac{\alpha}{2}}, \infty\right)$

b) $C = (z_\alpha, \infty)$

c) $C = (-\infty, -z_\alpha)$

- Cálculo del p-valor:

a) $p - \text{valor} = 2P(Z > |t|)$

b) P-valor=P(Z>t)

c) P-valor=P(Z<t)

## 1.2. VARIANZAS DESCONOCIDAS PERO IGUALES

· Tenemos:

- $X_1, X_2, \ldots, X_n$ que es una muestra de $X \sim N(\mu_x, \sigma_X^2)$
- $Y_1, Y_2, \ldots, Y_n$ que es una muestra de $Y \sim N(\mu_Y, \sigma_y^2)$
- $X_i$ es independiente de $Y_j$ para cualquier i, j.
- $\sigma_X^2 = \sigma_Y^2 = \sigma_2$

· Entonces:

1) $\frac{\bar{X}-\bar{Y}-(\mu_X-\mu_Y)}{\sqrt{\sigma^2\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}} \sim N(0,1)$
2) $\frac{n_X+n_Y-2}{\sigma^2} S_p^2 \sim \chi^2_{n_x+n_Y-2}$ siendo $S_p^2 \equiv \frac{(n_X-1)S_X^2+(n_Y-1)S_Y^2}{n_x+n_Y-2}$
3) $\bar{X}-\bar{Y}$ y $S_p^2$ son v.a. independientes
4) $\frac{\bar{X}-\bar{Y}-(\mu_X-\mu_Y)}{\sqrt{S_p^2\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}} \sim t_{n_{X+n_Y-2}}$

· **Intervalo de confianza de nivel 100(1-α)% para la diferencia de medias de dos poblaciones normales independientes de varianzas desconocidas pero iguales:**

$$\left(\bar{x}-\bar{y}-t_{n_{X+n_Y-2;\frac{\alpha}{2}}}\sqrt{S_p^2\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}, \bar{x}-\bar{y}+t_{n_{X+n_Y-2;\frac{\alpha}{2}}}\sqrt{S_p^2\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}\right)$$

· **Contrate de hipótesis** para diferencia de medias con varianzas desconocidas pero iguales:

- Estadístico de contraste:

$$T = \frac{\bar{X}-\bar{Y}-d_0}{\sqrt{S_p^2\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}} \sim t_{n_{X+n_Y-2}} \text{ bajo } H_0$$

- Hipótesis (igual que antes):

a) $H_0: \mu_x - \mu_Y = d_0$, $H_1: \mu_x - \mu_Y \neq d_0$
b) $H_0: \mu_x - \mu_Y = d_0$, $H_1: \mu_x - \mu_Y > d_0$
c) $H_0: \mu_x - \mu_Y = d_0$, $H_1: \mu_x - \mu_Y < d_0$

- Región crítica:

a) $C = \left(-\infty, -t_{n_{X+n_Y-2;\frac{\alpha}{2}}}\right) \cup \left(t_{n_{X+n_Y-2;\frac{\alpha}{2}}}, \infty\right)$
b) $C = \left(t_{n_{X+n_Y-2;\alpha}}, \infty\right)$
c) $C = \left(-\infty, -t_{n_{X+n_Y-2;\alpha}}\right)$

- Cálculo del p-valor:

a) p-valor=2P($t_{n_{X+n_Y-2}} > |t|$)
b) p-valor=P($t_{n_{X+n_Y-2}}$>t)
c) p-valor=P($t_{n_{X+n_Y-2}}$<t)



C/ Francia, 1 - 03690 S. Vicente del Raspeig +34 965 063 320 www.accademuniversidad.es info@accademuniversidad.es WhatsApp 692 349 023