Gestión Aeronáutica: Estadística Teórica
Facultad Ciencias Económicas y Empresariales
Departamento de Economía Aplicada
Profesor: Santiago de la Fuente Fernández

INTERVALOS DE CONFIANZA
CONTRASTES DE HIPÓTESIS

INTERVALOS DE CONFIANZA - CONTRASTES BILATERALES

a) Intervalo de confianza para la media μ de una distribución normal $N(\mu,\sigma)$ de varianza conocida σ^2

$$I_{1-\alpha}(\mu) = \left[\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

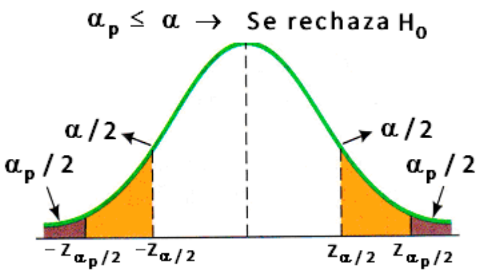
$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu_0 \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \;\rightarrow\; -z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \bar{x}-\mu_0 \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$|\bar{x}-\mu_0| \le z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \;\rightarrow\; \text{Estadístico contraste: } z_{p_{\alpha/2}} = \frac{|\bar{x}-\mu_0|}{\frac{\sigma}{\sqrt{n}}} \le z_{\alpha/2}$$

Contraste bilateral para la media μ de una distribución normal $N(\mu,\sigma)$ de varianza conocida σ^2

$H_0: \mu = \mu_0 \qquad H_a: \mu \ne \mu_0$

Región aceptación: $(-z_{\alpha/2}\,,\, z_{\alpha/2})$



α_p (p-valor) es el error de la primera región crítica de rechazo.

$\alpha = P(\text{Error Tipo I}) = P(\text{Rechazar } H_0 \,/\, H_0 \text{ Cierta})$

Se acepta H_0 si $z_{p_{\alpha/2}} = \dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \le z_{\alpha/2}$ Se rechaza H_0 si $z_{p_{\alpha/2}} = \dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} > z_{\alpha/2}$

SPSS: Calcula el p-valor (Significación bilateral) α_p para un contraste bilateral o de dos colas.

$\alpha_p = 2.P\left[z \ge z_{p_{\alpha/2}}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

La significación bilateral (p-valor) muestra el grado de compatibilidad entre el valor poblacional propuesto y la información muestral disponible. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

b) Intervalo de confianza para la media μ de una distribución normal $N(\mu,\sigma)$ de varianza desconocida σ^2 en muestras grandes $n > 30$

$$I_{1-\alpha}(\mu) = \left[\bar{x} \pm z_{\alpha/2}\frac{s_x}{\sqrt{n}}\right]$$

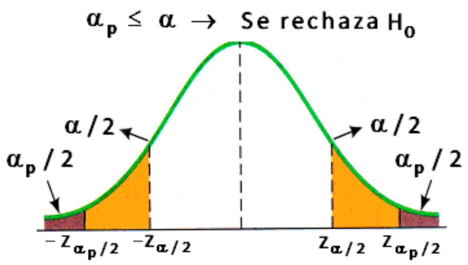
$$\bar{x} - z_{\alpha/2}\frac{s_x}{\sqrt{n}} \le \mu_0 \le \bar{x} + z_{\alpha/2}\frac{s_x}{\sqrt{n}} \;\rightarrow\; -z_{\alpha/2}\frac{s_x}{\sqrt{n}} \le \bar{x}-\mu_0 \le \bar{x} + z_{\alpha/2}\frac{s_x}{\sqrt{n}}$$

$$|\bar{x}-\mu_0| \le z_{\alpha/2}\frac{s_x}{\sqrt{n}} \;\rightarrow\; \text{Estadístico contraste: } z_{p_{\alpha/2}} = \frac{|\bar{x}-\mu_0|}{\frac{s_x}{\sqrt{n}}} \le z_{\alpha/2}$$

▭ Contraste bilateral para la media μ de una distribución normal $N(\mu, \sigma)$ de varianza desconocida σ^2 con muestras grandes $n > 30$

$H_0 : \mu = \mu_0 \qquad H_a : \mu \neq \mu_0$

Región aceptación: $(-z_{\alpha/2}\,,\, z_{\alpha/2})$



α_p (p-valor) es el error de la primera región crítica de rechazo.

$\alpha = P(\text{Error Tipo I}) =$
$= P(\text{Rechazar } H_0 \,/\, H_0 \text{ Cierta})$

Se acepta H_0 si $z_{p_{\alpha/2}} = \dfrac{|\bar{x} - \mu_0|}{s_x / \sqrt{n}} \leq z_{\alpha/2}$ Se rechaza H_0 si $z_{p_{\alpha/2}} = \dfrac{|\bar{x} - \mu_0|}{s_x / \sqrt{n}} > z_{\alpha/2}$

SPSS: Calcula el p-valor (Significación bilateral) α_p para un contraste bilateral o de dos colas.

$\alpha_p = 2.P\left[z \geq z_{p_{\alpha/2}}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

La significación bilateral (p-valor) muestra el grado de compatibilidad entre el valor poblacional propuesto y la información muestral disponible.

Muestras pequeñas $n \leq 30 \quad \rightarrow \quad I_{1-\alpha}(\mu) = \left[\bar{x} \pm t_{\alpha/2,\,(n-1)} \dfrac{s_x}{\sqrt{n}}\right]$

$$\bar{x} - t_{\alpha/2,\,(n-1)} \frac{s_x}{\sqrt{n}} \leq \mu_0 \leq \bar{x} + t_{\alpha/2,\,(n-1)} \frac{s_x}{\sqrt{n}}$$

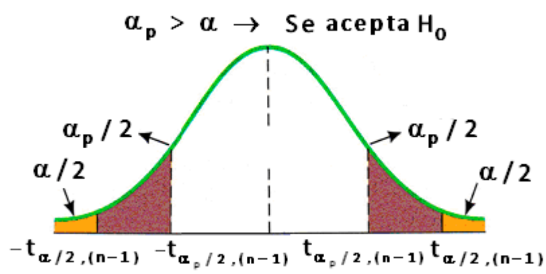
$$-\, t_{\alpha/2,\,(n-1)} \frac{s_x}{\sqrt{n}} \leq \bar{x} - \mu_0 \leq t_{\alpha/2,\,(n-1)} \frac{s_x}{\sqrt{n}} \quad \rightarrow \quad |\bar{x} - \mu_0| \leq t_{\alpha/2,\,(n-1)} \frac{s_x}{\sqrt{n}}$$

Estadístico contraste: $t_{p_{\alpha/2}} = \dfrac{|\bar{x} - \mu_0|}{\dfrac{s_x}{\sqrt{n}}} \leq t_{\alpha/2,\,(n-1)}$

▭ Contraste bilateral para la media μ de una distribución normal $N(\mu, \sigma)$ de varianza desconocida σ^2 en muestras pequeñas $n \leq 30$

$H_0 : \mu = \mu_0 \qquad H_a : \mu \neq \mu_0$

Región aceptación: $(-t_{\alpha/2,\,(n-1)}\,,\, t_{\alpha/2,\,(n-1)})$



Se acepta H_0 si $t_{p_{\alpha/2}} = \dfrac{|\bar{x} - \mu_0|}{s_x / \sqrt{n}} \leq t_{\alpha/2,\,(n-1)}$ Se rechaza H_0 si $t_{p_{\alpha/2}} = \dfrac{|\bar{x} - \mu_0|}{s_x / \sqrt{n}} > t_{\alpha/2,\,(n-1)}$

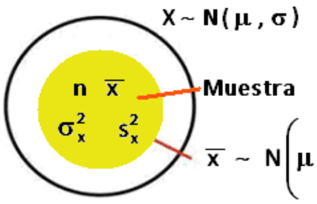
SPSS: Calcula el p-valor (Significación bilateral) α_p para un contraste bilateral o de dos colas.

$\alpha_p = 2.P\left[t_{(n-1)} \geq t_{p_{\alpha/2}}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

La significación bilateral (p-valor) muestra el grado de compatibilidad entre el valor poblacional propuesto y la información muestral disponible. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa

c) Intervalo de confianza para la varianza σ^2 de una distribución normal

$$I_{1-\alpha}(\sigma^2) = \left[\frac{(n-1)\,.\,s_x^2}{\chi^2_{\alpha/2,(n-1)}} \, , \, \frac{(n-1)\,.\,s_x^2}{\chi^2_{1-\alpha/2,(n-1)}} \right]$$



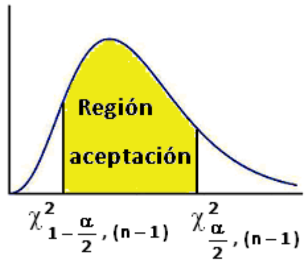
$$\chi^2_{(n-1)} = \frac{\text{Varianza muestral observada}}{\text{Varianza muestral teórica}} = \frac{\sigma_x^2}{\sigma^2/n} = \frac{n\,.\,\sigma_x^2}{\sigma^2} = \frac{(n-1)\,.\,s_x^2}{\sigma^2}$$

$$\bar{x} \sim N\left(\mu \, , \, \frac{\sigma}{\sqrt{n}}\right) \qquad \sigma^2 = \left. \frac{(n-1)\,.\,s_x^2}{\chi^2_{(n-1)}} \right|_{H_0} \quad \rightarrow \quad \chi^2_{p_{\alpha/2}} = \frac{(n-1)\,.\,s_x^2}{\sigma_0^2}$$

🖃 Contraste bilateral para la varianza σ^2 de una distribución normal

$H_0 : \sigma^2 = \sigma_0^2 \qquad H_a : \sigma^2 \neq \sigma_0^2$

Región aceptación: $\left(\chi^2_{1-\alpha/2,(n-1)} \, , \, \chi^2_{\alpha/2,(n-1)}\right)$



Se acepta H_0 si $\chi^2_{p_{\alpha/2}} = \dfrac{(n-1)\,.\,s_x^2}{\sigma_0^2} \in \left[\chi^2_{1-\alpha/2,(n-1)} \, , \, \chi^2_{\alpha/2,(n-1)}\right]$

Se rechaza H_0 si $\chi^2_{p_{\alpha/2}} = \dfrac{(n-1)\,.\,s_x^2}{\sigma_0^2} \notin \left[\chi^2_{1-\alpha/2,(n-1)} \, , \, \chi^2_{\alpha/2,(n-1)}\right]$

En todos los intervalos de confianza $s_x^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$ es la cuasivarianza muestral.

d) Intervalo de confianza para la diferencia de medias de dos distribuciones normales

d.1) Las varianzas poblaciones σ_1^2 y σ_2^2 son conocidas

$$I_{1-\alpha}(\mu_1 - \mu_2) = \left[(\bar{x} - \bar{y}) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right]$$

Cuando el intervalo cubre el 0 no hay diferencia significativa entre las medias poblacionales.

$$(\bar{x} - \bar{y}) - z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq (\mu_1 - \mu_2) \leq (\bar{x} - \bar{y}) + z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$-\, z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq (\bar{x} - \bar{y}) - (\mu_1 - \mu_2) \leq z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\left|(\bar{x}-\bar{y})-(\mu_1-\mu_2)\right| \le z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \quad \rightarrow \quad \frac{\left|(\bar{x}-\bar{y})-(\mu_1-\mu_2)\right|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \le z_{\alpha/2}$$

Estadístico contraste: $z_{p_{\alpha/2}} = \left.\dfrac{\left|(\bar{x}-\bar{y})-(\mu_1-\mu_2)\right|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\right|_{\mu_1-\mu_2=0} = \dfrac{\left|\bar{x}-\bar{y}\right|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \le z_{\alpha/2}$

Contraste bilateral para la diferencia de medias de dos distribuciones normales con varianzas poblaciones σ_1^2 y σ_2^2 conocidas

$$H_0: \mu_1-\mu_2=0 \qquad H_a: \mu_1-\mu_2\neq 0$$

Se acepta H_0 si $z_{p_{\alpha/2}} = \dfrac{\left|\bar{x}-\bar{y}\right|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \le z_{\alpha/2}$ Se rechaza H_0 si $z_{p_{\alpha/2}} = \dfrac{\left|\bar{x}-\bar{y}\right|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} > z_{\alpha/2}$

SPSS: Calcula el p-valor (Significación bilateral) α_p para un contraste bilateral o de dos colas.

$\alpha_p = 2.P\left[z \ge z_{p_{\alpha/2}}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

d.2) Intervalo de confianza para la diferencia de medias de dos distribuciones normales con varianzas poblaciones σ_1^2 y σ_2^2 desconocidas.

Caso en que la suma $(n_1+n_2) > 30$ con $n_1 \approx n_2$

$$I_{1-\alpha}(\mu_1-\mu_2) = \left[(\bar{x}-\bar{y}) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right]$$

Cuando el intervalo cubre el 0 no hay diferencia significativa entre las medias poblacionales.

$$(\bar{x}-\bar{y}) - z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \le (\mu_1-\mu_2) \le (\bar{x}-\bar{y}) + z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

$$-z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \le (\bar{x}-\bar{y})-(\mu_1-\mu_2) \le z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

$$\left|(\bar{x}-\bar{y})-(\mu_1-\mu_2)\right| \le z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \quad \rightarrow \quad \frac{\left|(\bar{x}-\bar{y})-(\mu_1-\mu_2)\right|}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \le z_{\alpha/2}$$

Estadístico contraste: $$z_{p_{\alpha/2}} = \left.\frac{\left|(\bar{x}-\bar{y})-(\mu_1-\mu_2)\right|}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}\right|_{\mu_1-\mu_2=0} = \frac{\left|\bar{x}-\bar{y}\right|}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}} \leq z_{\alpha/2}$$

Contraste bilateral para la diferencia de medias de dos distribuciones normales con varianzas poblaciones σ_1^2 y σ_2^2 desconocidas, cuando $(n_1+n_2) > 30$ con $n_1 \approx n_2$

$$H_0: \mu_1-\mu_2=0 \qquad H_a: \mu_1-\mu_2 \neq 0$$

Se acepta H_0 si $z_{p_{\alpha/2}} = \dfrac{\left|\bar{x}-\bar{y}\right|}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}} \leq z_{\alpha/2}$ Se rechaza H_0 si $z_{p_{\alpha/2}} = \dfrac{\left|\bar{x}-\bar{y}\right|}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}} > z_{\alpha/2}$

SPSS: Calcula el p-valor (Significación bilateral) α_p para un contraste bilateral o de dos colas.

$$\alpha_p = 2.P\left[z \geq z_{p_{\alpha/2}}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

Caso en que los tamaños muestrales son pequeños $(n_1+n_2) \leq 30$ y las varianzas son desconocidas, pero iguales $\left(\sigma_1^2=\sigma_2^2\right)$

$$I_{1-\alpha}(\mu_1-\mu_2) = \left[(\bar{x}-\bar{y}) \pm t_{\alpha/2,(n_1+n_2-2)} \cdot s_p \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]$$

Cuando el intervalo cubre el 0 no hay diferencia significativa entre las medias poblacionales.

s_p^2 es la media ponderada de las cuasivarianzas muestrales: $s_p^2 = \dfrac{(n_1-1).s_1^2+(n_2-1).s_2^2}{n_1+n_2-2}$

$$(\bar{x}-\bar{y})-t_{\alpha/2,(n_1+n_2-2)} \cdot s_p \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \leq (\mu_1-\mu_2) \leq (\bar{x}-\bar{y})+t_{\alpha/2,(n_1+n_2-2)} \cdot s_p \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$-t_{\alpha/2,(n_1+n_2-2)} \cdot s_p \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \leq (\bar{x}-\bar{y})-(\mu_1-\mu_2) \leq t_{\alpha/2,(n_1+n_2-2)} \cdot s_p \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$\left|(\bar{x}-\bar{y})-(\mu_1-\mu_2)\right| \leq t_{\alpha/2,(n_1+n_2-2)} \cdot s_p \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

Estadístico contraste:

$$t_{p_{\alpha/2}} = \left.\frac{\left|(\bar{x}-\bar{y})-(\mu_1-\mu_2)\right|}{s_p \cdot \sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\right|_{\mu_1-\mu_2=0} = \frac{\left|\bar{x}-\bar{y}\right|}{s_p \cdot \sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \leq t_{\alpha/2,(n_1+n_2-2)}$$

🖃 Contraste bilateral para la diferencia de medias de dos distribuciones normales con varianzas poblaciones σ_1^2 y σ_2^2 desconocidas pero iguales $\left(\sigma_1^2=\sigma_2^2\right)$, con muestras pequeñas $(n_1+n_2)\leq 30$

$$H_0: \mu_1-\mu_2=0 \qquad H_a: \mu_1-\mu_2\neq 0$$

Se acepta H_0 si $\quad t_{p_{\alpha/2}} = \dfrac{|\bar{x}-\bar{y}|}{s_p\cdot\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \leq t_{\alpha/2,(n_1+n_2-2)}$

Se rechaza H_0 si $\quad t_{p_{\alpha/2}} = \dfrac{|\bar{x}-\bar{y}|}{s_p\cdot\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} > t_{\alpha/2,(n_1+n_2-2)}$

SPSS: Calcula el p-valor (Significación bilateral) α_p para un contraste bilateral o de dos colas.

$$\alpha_p = 2.P\left[t_{(n_1+n_2-2)} \geq t_{p_{\alpha/2}}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

🕮 Caso en que los tamaños muestrales son pequeños $(n_1+n_2)\leq 30$ y las varianzas son desconocidas y distintas $\left(\sigma_1^2\neq\sigma_2^2\right)$

$$I_{1-\alpha}(\mu_1-\mu_2) = \left[(\bar{x}-\bar{y}) \pm t_{\alpha/2,f}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right]$$

Cuando el intervalo cubre el 0 no hay diferencia significativa entre las medias poblacionales.

$$\text{f es la aproximación de Welch: } f = \frac{\left(\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1+1}+\dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2+1}} - 2$$

$$(\bar{x}-\bar{y}) - t_{\alpha/2,f}\cdot\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \leq (\mu_1-\mu_2) \leq (\bar{x}-\bar{y}) + t_{\alpha/2,f}\cdot\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

$$-\,t_{\alpha/2,f}\cdot\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \leq (\bar{x}-\bar{y})-(\mu_1-\mu_2) \leq t_{\alpha/2,f}\cdot\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

$$\text{Estadístico contraste: } \quad t_{p_{\alpha/2}} = \left.\frac{|(\bar{x}-\bar{y})-(\mu_1-\mu_2)|}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}\right|_{\mu_1-\mu_2=0} = \frac{|\bar{x}-\bar{y}|}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}} \leq t_{\alpha/2,f}$$

🖮 Contraste bilateral para la diferencia de medias de dos distribuciones normales con varianzas poblaciones σ_1^2 y σ_2^2 desconocidas y distintas, con tamaños muestrales pequeños $(n_1 + n_2) \leq 30$

$$H_0 : \mu_1 - \mu_2 = 0 \qquad H_a : \mu_1 - \mu_2 \neq 0$$

Se acepta H_0 si $t_{p_{\alpha/2}} = \dfrac{|\bar{x} - \bar{y}|}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \leq t_{\alpha/2,f}$ Se rechaza H_0 si $t_{p_{\alpha/2}} = \dfrac{|\bar{x} - \bar{y}|}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} > t_{\alpha/2,f}$

SPSS: Calcula el p-valor (Significación bilateral) α_p para un contraste bilateral o de dos colas.

$\alpha_p = 2.P\left[t_{(n_1+n_2-2)} \geq t_{p_{\alpha/2}}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

e) Intervalo de confianza para la razón de varianzas de dos poblaciones normales

$$I_{1-\alpha}(\sigma_1^2 / \sigma_2^2) = \left[\frac{s_1^2 / s_2^2}{F_{\alpha/2;(n_1-1),(n_2-1)}} \, , \, \frac{s_1^2 / s_2^2}{F_{(1-\alpha/2);(n_1-1),(n_2-1)}}\right]$$

Cuando el intervalo cubre el 1 no hay diferencia significativa entre las varianzas poblacionales.

Hay que considerar la relación: $F_{\alpha;\, n_1, n_2} = \dfrac{1}{F_{(1-\alpha);\, n_2, n_1}}$

Se tiene: $\chi^2_{(n_1-1)} = \dfrac{(n_1-1).\, s^2_{x_1}}{\sigma_1^2}$, $\chi^2_{(n_2-1)} = \dfrac{(n_2-1).\, s^2_{x_2}}{\sigma_2^2}$

$$F_{n_1,n_2} = \frac{\chi^2_{(n_1-1)}/(n_1-1)}{\chi^2_{(n_2-1)}/(n_2-1)} = \frac{\chi^2_{(n_1-1)}/(n_1-1)}{\chi^2_{(n_2-1)}/(n_2-1)} = \frac{\dfrac{(n_1-1).\, s^2_{x_1}}{\sigma_1^2}/(n_1-1)}{\dfrac{(n_2-1).\, s^2_{x_2}}{\sigma_2^2}/(n_2-1)} = \frac{s^2_{x_1}.\, \sigma_2^2}{s^2_{x_2}.\, \sigma_1^2}$$

$$\frac{\sigma_1^2}{\sigma_2^2} = \frac{s^2_{x_1} / s^2_{x_2}}{F_{n_1,n_2}}$$

🖮 Contraste bilateral para la igualdad de varianzas de dos poblaciones normales

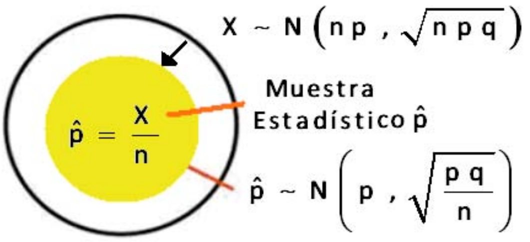
$$H_0 : \sigma_1^2 = \sigma_2^2 \qquad H_a : \sigma_1^2 \neq \sigma_2^2$$

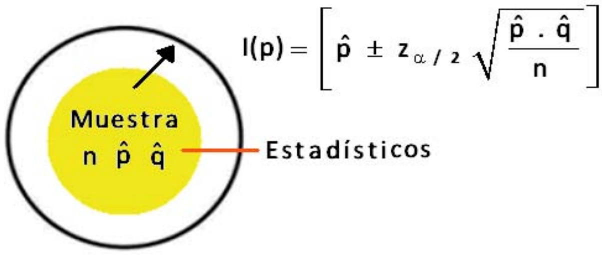
Se acepta H_0 si $F_{p_{\alpha/2}} = \dfrac{s_1^2}{s_2^2} \in \left[F_{(1-\alpha/2);(n_1-1),(n_2-1)} \, , \, F_{\alpha/2;(n_1-1),(n_2-1)}\right]$

Se rechaza H_0 si $F_{p_{\alpha/2}} = \dfrac{s_1^2}{s_2^2} \notin \left[F_{(1-\alpha/2);(n_1-1),(n_2-1)} \, , \, F_{\alpha/2;(n_1-1),(n_2-1)}\right]$

f) Intervalo de confianza para el parámetro p de una distribución binomial de parámetros n, p, B(n, p)

$$I_{1-\alpha}(p) = \left[\hat{p} \;\pm\; z_{\alpha/2}\sqrt{\frac{\hat{p}\,.(1-\hat{p})}{n}}\right]$$





$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\,.(1-\hat{p})}{n}} \le p_0 \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\,.(1-\hat{p})}{n}}$$

$$\left|\hat{p}-p_0\right| \le z_{\alpha/2}\sqrt{\frac{\hat{p}\,.(1-\hat{p})}{n}} \;\rightarrow\; \text{Estadístico contraste: } z_{p_{\alpha/2}} = \frac{\left|\hat{p}-p_0\right|}{\sqrt{\frac{\hat{p}\,.(1-\hat{p})}{n}}} \le z_{\alpha/2}$$

🖃 Contraste bilateral para el parámetro p de una distribución binomial

$$H_0: p = p_0 \qquad H_a: p \neq p_0$$

Se acepta H_0 si $z_{p_{\alpha/2}} = \dfrac{\left|\hat{p}-p_0\right|}{\sqrt{\frac{\hat{p}\,.(1-\hat{p})}{n}}} \le z_{\alpha/2}$ Se rechaza H_0 si $z_{p_{\alpha/2}} = \dfrac{\left|\hat{p}-p_0\right|}{\sqrt{\frac{\hat{p}\,.(1-\hat{p})}{n}}} > z_{\alpha/2}$

SPSS: Calcula el p-valor (Significación bilateral) α_p para un contraste bilateral o de dos colas.

$$\alpha_p = 2.P\left[z \ge z_{p_{\alpha/2}}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

g) Intervalo de confianza para la diferencia de parámetros $(p_1 - p_2)$ de dos distribuciones binomiales.

$$I_{1-\alpha}(p_1 - p_2) = \left[(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}\right]$$

Cuando el intervalo cubre el 0 no hay diferencia significativa entre las proporciones poblacionales.

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \le (p_1 - p_2) \le (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$-z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \le (\hat{p}_1 - \hat{p}_2) - (p_1 - p_2) \le z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$\text{Contraste: } z_{p_{\alpha/2}} = \left.\frac{\left|(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)\right|}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}\right|_{p_1 - p_2 = 0} = \frac{\left|\hat{p}_1 - \hat{p}_2\right|}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \le z_{\alpha/2}$$

🖮 Contraste bilateral para la igualdad de los parámetros de dos distribuciones binomiales $B_1(n_1, p_1)$ y $B_2(n_2, p_2)$

$$H_0 : p_1 = p_2 \qquad H_a : p_1 \neq p_2$$

Se acepta H_0 si $$z_{p_{\alpha/2}} = \frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \leq z_{\alpha/2}$$

Se rechaza H_0 si $$z_{p_{\alpha/2}} = \frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} > z_{\alpha/2}$$

SPSS: Calcula el p-valor (Significación bilateral) α_p para un contraste bilateral o de dos colas.

$$\alpha_p = 2.P\left[z \geq z_{p_{\alpha/2}}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

h) Intervalo de confianza para el parámetro λ de una distribución de Poisson

$$I_{1-\alpha}(\lambda) = \left[\hat{\lambda} \pm z_{\alpha/2}\sqrt{\frac{\hat{\lambda}}{n}}\right]$$

🖮 Contraste bilateral para el parámetro λ de una distribución de Poisson, muestras grandes

$$H_0 : \lambda = \lambda_0 \qquad H_a : \lambda \neq \lambda_0$$

Se acepta H_0 si $z_{p_{\alpha/2}} = \dfrac{|\hat{\lambda} - \lambda_0|}{\sqrt{\dfrac{\hat{\lambda}}{n}}} \leq z_{\alpha/2}$ Se acepta H_0 si $z_{p_{\alpha/2}} = \dfrac{|\hat{\lambda} - \lambda_0|}{\sqrt{\dfrac{\hat{\lambda}}{n}}} > z_{\alpha/2}$

i) Intervalo de confianza para la diferencia de datos apareados

i.1) Para muestras grandes $n > 30$: $I = \left[\overline{d} \pm z_{\alpha/2}\dfrac{s_d}{\sqrt{n}}\right]$

$$d_i = x_i - y_i \qquad \overline{d} = \frac{1}{n}\sum_{i=1}^{n} d_i \qquad s_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \overline{d})^2$$

🖮 Contraste bilateral de igualdad de medias de datos apareados para muestras grandes $n > 30$

$$H_0 : \mu_1 - \mu_2 = 0 \Leftrightarrow d = 0 \qquad H_a : \mu_1 - \mu_2 \neq 0 \Leftrightarrow d \neq 0$$

Se acepta H_0 si $z_{p_{\alpha/2}} = \dfrac{|\overline{d}|}{s_d / \sqrt{n}} \leq z_{\alpha/2}$ Se acepta H_0 si $z_{p_{\alpha/2}} = \dfrac{|\overline{d}|}{s_d / \sqrt{n}} > z_{\alpha/2}$

i.2) Para muestras pequeñas $n \leq 30$: $I = \left[\bar{d} \pm t_{\alpha/2,(n-1)} \dfrac{s_d}{\sqrt{n}} \right]$

🖃 Contraste bilateral de igualdad de medias de datos apareados para muestras pequeñas $n \leq 30$

$$H_0 : \mu_1 - \mu_2 = 0 \iff d = 0 \qquad H_a : \mu_1 - \mu_2 \neq 0 \iff d \neq 0$$

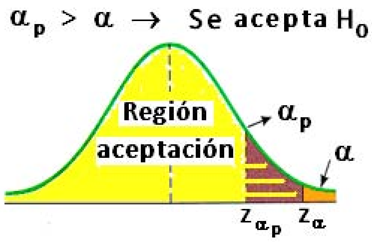
Se acepta H_0 si $t_{p_{\alpha/2}} = \dfrac{|\bar{d}|}{s_d / \sqrt{n}} \leq t_{\alpha/2,(n-1)}$ Se acepta H_0 si $t_{p_{\alpha/2}} = \dfrac{|\bar{d}|}{s_d / \sqrt{n}} > t_{\alpha/2,(n-1)}$

a) Contraste unilateral para la media μ de una población normal con varianza σ^2 conocida.

- Contraste unilateral a la derecha (cola a la derecha)

$H_0 : \mu \leq \mu_0 \qquad H_a : \mu > \mu_0$

Región aceptación $(-\infty \,,\, z_\alpha)$



α_p (p-valor) es el error de la primera región crítica de rechazo.

$\alpha = P(\text{Error Tipo I}) =$
$= P(\text{Rechazar } H_0 \,/\, H_0 \text{ Cierta})$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \leq z_\alpha$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} > z_\alpha$

Intervalo de confianza: $I(\mu) = \left(-\infty \,,\, \mu_0 + z_\alpha \dfrac{\sigma}{\sqrt{n}} \right)$

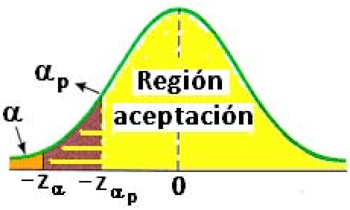
SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la derecha hay que dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = \left(\frac{Sig}{2}\right)_{SPSS} = P\left[z \geq z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

- Contraste unilateral a la izquierda (cola a la izquierda)

$H_0 : \mu \geq \mu_0 \qquad H_a : \mu < \mu_0$

Región aceptación $(-z_\alpha \,,\, \infty)$



α_p (p-valor) es el error de la primera región crítica de rechazo.

$\alpha = P(\text{Error Tipo I}) =$
$= P(\text{Rechazar } H_0 \,/\, H_0 \text{ Cierta})$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \geq -z_\alpha$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} < -z_\alpha$

Intervalo de confianza: $I(\mu) = \left(\mu_0 - z_\alpha \dfrac{\sigma}{\sqrt{n}} \,,\, \infty \right)$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la izquierda hay que restar a 1 el resultado de dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = 1 - \left(\frac{Sig}{2}\right)_{SPSS} = P\left[z \leq -z_{\alpha_p}\right] = P\left[z \geq z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

b) Contraste unilateral para la media μ de una población normal con varianza σ^2 desconocida en muestras grandes $n > 30$.

🖃 Contraste unilateral a la derecha (cola a la derecha)

$H_0 : \mu \le \mu_0$ $\qquad H_a : \mu > \mu_0$

Región aceptación $(-\infty \,,\, z_\alpha)$

α_p (p-valor) es el error de la primera región crítica de rechazo.

$\alpha = P(\text{Error Tipo I}) =$
$= P(\text{Rechazar } H_0 \,/\, H_0 \text{ Cierta})$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \mu_0}{s_x / \sqrt{n}} \le z_\alpha$ $\qquad$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \mu_0}{s_x / \sqrt{n}} > z_\alpha$

Intervalo de confianza: $I(\mu) = \left(-\infty \,,\, \mu_0 + z_\alpha \dfrac{s_x}{\sqrt{n}}\right)$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la derecha hay que dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = \left(\frac{\text{Sig}}{2}\right)_{\text{SPSS}} = P\left[z \ge z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

🖃 Contraste unilateral a la izquierda (cola a la izquierda)

$H_0 : \mu \ge \mu_0$ $\qquad H_a : \mu < \mu_0$

Región aceptación $(-z_\alpha \,,\, \infty)$

α_p (p-valor) es el error de la primera región crítica de rechazo.

$\alpha = P(\text{Error Tipo I}) =$
$= P(\text{Rechazar } H_0 \,/\, H_0 \text{ Cierta})$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \mu_0}{s_x / \sqrt{n}} \ge -z_\alpha$ $\qquad$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \mu_0}{s_x / \sqrt{n}} < -z_\alpha$

Intervalo de confianza: $I(\mu) = \left(\mu_0 - z_\alpha \dfrac{s_x}{\sqrt{n}} \,,\, \infty\right)$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la izquierda hay que restar a 1 el resultado de dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

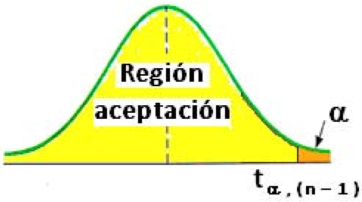
$$\alpha_p = 1 - \left(\frac{\text{Sig}}{2}\right)_{\text{SPSS}} = P\left[z \le -z_{\alpha_p}\right] = P\left[z \ge z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

c) Contraste unilateral para la media μ de una población normal con varianza σ^2 desconocida en muestras pequeñas $n \leq 30$.

Contraste unilateral a la derecha (cola a la derecha)

$H_0: \mu \leq \mu_0$ $\quad$ $H_a: \mu > \mu_0$

Región aceptación $(-\infty \, , \, t_{\alpha\,,\,(n-1)})$



α_p (p-valor) es el error de la primera región crítica de rechazo.

$\alpha = P(\text{Error Tipo I}) =$
$= P(\text{Rechazar } H_0 \, / \, H_0 \text{ Cierta})$

Se acepta H_0 si $t_{p_\alpha} = \dfrac{\bar{x} - \mu_0}{s_x / \sqrt{n}} \leq t_{\alpha\,,\,(n-1)}$ $\quad$ Se rechaza H_0 si $t_{p_\alpha} = \dfrac{\bar{x} - \mu_0}{s_x / \sqrt{n}} > t_{\alpha\,,\,(n-1)}$

Intervalo de confianza: $I(\mu) = \left(-\infty \, , \, \mu_0 + t_{\alpha\,,\,(n-1)} \dfrac{s_x}{\sqrt{n}} \right)$

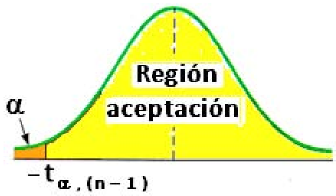
SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la derecha hay que dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = \left(\frac{Sig}{2} \right)_{SPSS} = P\left[t_{(n-1)} \geq t_{\alpha_p} \right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

Contraste unilateral a la izquierda (cola a la izquierda)

$H_0: \mu \geq \mu_0$ $\quad$ $H_a: \mu < \mu_0$

Región aceptación $(-t_{\alpha\,,\,(n-1)} \, , \, \infty)$



α_p (p-valor) es el error de la primera región crítica de rechazo.

$\alpha = P(\text{Error Tipo I}) =$
$= P(\text{Rechazar } H_0 \, / \, H_0 \text{ Cierta})$

Se acepta H_0 si $t_{p_\alpha} = \dfrac{\bar{x} - \mu_0}{s_x / \sqrt{n}} \geq -t_{\alpha\,,\,(n-1)}$ $\quad$ Se rechaza H_0 si $t_{p_\alpha} = \dfrac{\bar{x} - \mu_0}{s_x / \sqrt{n}} < -t_{\alpha\,,\,(n-1)}$

Intervalo de confianza: $I(\mu) = \left(\mu_0 - t_{\alpha\,,\,(n-1)} \dfrac{s_x}{\sqrt{n}} \, , \, \infty \right)$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la izquierda hay que restar a 1 el resultado de dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = 1 - \left(\frac{Sig}{2} \right)_{SPSS} = P\left[t_{(n-1)} \leq -t_{\alpha_p} \right] = P\left[t_{(n-1)} \geq t_{\alpha_p} \right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

d) Contraste unilateral para el parámetro p de una distribución binomial

▭ Contraste unilateral a la derecha (cola a la derecha)

$H_0: p \leq p_0 \qquad H_a: p > p_0$

Región aceptación $(-z_\alpha\ ,\ \infty)$

α_p (p-valor) es el error de la primera región crítica de rechazo.

$\alpha = P(\text{Error Tipo I}) =$
$= P(\text{Rechazar } H_0 \,/\, H_0 \text{ Cierta})$

Se acepta H_0 si $z_{p_\alpha} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{\hat{p}\,.\,(1-\hat{p})}{n}}} \leq z_\alpha$ Se rechaza H_0 si $z_{p_\alpha} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{\hat{p}\,.\,(1-\hat{p})}{n}}} > z_\alpha$

Intervalo de confianza: $I(p) = \left(-\infty\ ,\ p_0 + z_\alpha \sqrt{\dfrac{\hat{p}\,.\,\hat{q}}{n}}\right)$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la derecha hay que dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = \left(\frac{Sig}{2}\right)_{SPSS} = P\left[z \geq z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

▭ Contraste unilateral a la izquierda (cola a la izquierda)

$H_0: p \geq p_0 \qquad H_a: p < p_0$

Región aceptación $(-z_\alpha\ ,\ \infty)$

α_p (p-valor) es el error de la primera región crítica de rechazo.

$\alpha = P(\text{Error Tipo I}) =$
$= P(\text{Rechazar } H_0 \,/\, H_0 \text{ Cierta})$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{\hat{p}\,.\,(1-\hat{p})}{n}}} \geq -z_\alpha$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{\hat{p}\,.\,(1-\hat{p})}{n}}} < -z_\alpha$

Intervalo de confianza: $I(p) = \left(p_0 - z_\alpha \sqrt{\dfrac{\hat{p}\,.\,\hat{q}}{n}}\ ,\ \infty\right)$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la izquierda hay que restar a 1 el resultado de dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

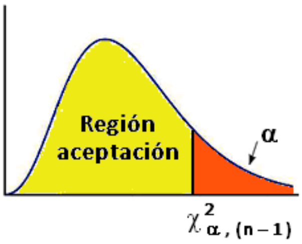
$$\alpha_p = 1 - \left(\frac{Sig}{2}\right)_{SPSS} = P\left[z \leq -z_{\alpha_p}\right] = P\left[z \geq z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

e) Contraste unilateral para la varianza σ^2 de una distribución normal

- Contraste unilateral a la derecha (cola a la derecha)

$$H_0 : \sigma^2 \leq \sigma_0^2 \qquad H_a : \sigma^2 > \sigma_0^2$$

Región aceptación: $\leq \chi^2_{\alpha, (n-1)}$



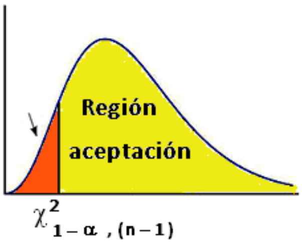
Se acepta H_0 si $\chi^2_{\alpha_p} = \dfrac{(n-1).\, s_x^2}{\sigma_0^2} \leq \chi^2_{\alpha, (n-1)}$

Se rechaza H_0 si $\chi^2_{\alpha_p} = \dfrac{(n-1).\, s_x^2}{\sigma_0^2} > \chi^2_{\alpha, (n-1)}$

- Contraste unilateral a la izquierda (cola a la izquierda)

$$H_0 : \sigma^2 \geq \sigma_0^2 \qquad H_a : \sigma^2 < \sigma_0^2$$

Región aceptación: $\geq \chi^2_{1-\alpha, (n-1)}$



Se acepta H_0 si $\chi^2_{\alpha_p} = \dfrac{(n-1).\, s_x^2}{\sigma_0^2} \geq \chi^2_{1-\alpha, (n-1)}$

Se rechaza H_0 si $\chi^2_{\alpha_p} = \dfrac{(n-1).\, s_x^2}{\sigma_0^2} < \chi^2_{1-\alpha, (n-1)}$

f) Contraste unilateral para la diferencia de medias de dos distribuciones normales con varianzas poblaciones σ_1^2 y σ_2^2 conocidas

- Contraste unilateral a la derecha (cola a la derecha)

$$H_0 : (\mu_1 - \mu_2) \leq 0 \qquad H_a : (\mu_1 - \mu_2) > 0$$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\overline{x} - \overline{y}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \leq z_\alpha$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\overline{x} - \overline{y}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} > z_\alpha$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la derecha hay que dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = \left(\frac{Sig}{2}\right)_{SPSS} = P\left[z \geq z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

♣ Contraste unilateral a la derecha (cola a la derecha, constante)

$$H_0 : (\mu_1 - \mu_2) \leq k \qquad H_a : (\mu_1 - \mu_2) > k$$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \bar{y} - k}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \leq z_\alpha$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \bar{y} - k}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} > z_\alpha$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la derecha hay que dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = \left(\frac{Sig}{2}\right)_{SPSS} = P\left[z \geq z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

▭ Contraste unilateral a la izquierda (cola a la izquierda)

$$H_0 : (\mu_1 - \mu_2) \geq 0 \qquad H_a : (\mu_1 - \mu_2) < 0$$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \geq -z_\alpha$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} < -z_\alpha$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la izquierda hay que restar a 1 el resultado de dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = 1 - \left(\frac{Sig}{2}\right)_{SPSS} = P\left[z \leq -z_{\alpha_p}\right] = P\left[z \geq z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

♣ Contraste unilateral a la izquierda (cola a la izquierda, constante)

$$H_0 : (\mu_1 - \mu_2) \geq k \qquad H_a : (\mu_1 - \mu_2) < k$$

Se acepta H_0 si $z_{p_\alpha} = \dfrac{\bar{x} - \bar{y} - k}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \geq -z_\alpha$ Se rechaza H_0 si $z_{p_\alpha} = \dfrac{\bar{x} - \bar{y} - k}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} < -z_\alpha$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la izquierda hay que restar a 1 el resultado de dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = 1 - \left(\frac{Sig}{2}\right)_{SPSS} = P\left[z \leq -z_{\alpha_p}\right] = P\left[z \geq z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

g) Contraste unilateral para la diferencia de medias de dos distribuciones normales con varianzas poblaciones σ_1^2 y σ_2^2 desconocidas, cuando $(n_1 + n_2) > 30$ con $n_1 \approx n_2$

🖮 Contraste unilateral a la derecha (cola a la derecha)

$$H_0 : (\mu_1 - \mu_2) \leq 0 \qquad H_a : (\mu_1 - \mu_2) > 0$$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \leq z_\alpha$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} > z_\alpha$

SPSS: Para calcular el nivel de significación α_p de un contraste unilateral a la derecha hay que dividir entre 2 el contraste bilateral de salida de SPSS. Es el nivel de significación más pequeño posible para el que todavía se aceptaría la hipótesis alternativa.

$$\alpha_p = \left(\frac{Sig}{2}\right)_{SPSS} = P\left[z \geq z_{\alpha_p}\right] \quad \text{cuando } \alpha_p > \alpha \text{ se acepta } H_0$$

♣ Contraste unilateral a la derecha (cola a la derecha, constante)

$$H_0 : (\mu_1 - \mu_2) \leq k \qquad H_a : (\mu_1 - \mu_2) > k$$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \bar{y} - k}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \leq z_\alpha$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \bar{y} - k}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} > z_\alpha$

SPSS: $\alpha_p = \left(\dfrac{Sig}{2}\right)_{SPSS} = P\left[z \geq z_{\alpha_p}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

🖮 Contraste unilateral a la izquierda (cola a la izquierda)

$$H_0 : (\mu_1 - \mu_2) \geq 0 \qquad H_a : (\mu_1 - \mu_2) < 0$$

Se acepta H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \geq -z_\alpha$ Se rechaza H_0 si $z_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} < -z_\alpha$

SPSS: $\alpha_p = 1 - \left(\dfrac{Sig}{2}\right)_{SPSS} = P\left[z \leq -z_{\alpha_p}\right] = P\left[z \geq z_{\alpha_p}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

♣ Contraste unilateral a la izquierda (cola a la izquierda, constante)

$$H_0 : (\mu_1 - \mu_2) \geq k \qquad H_a : (\mu_1 - \mu_2) < k$$

Se acepta H_0 si $z_{p_\alpha} = \dfrac{\bar{x} - \bar{y} - k}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \geq -z_\alpha$ Se rechaza H_0 si $z_{p_\alpha} = \dfrac{\bar{x} - \bar{y} - k}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} < -z_\alpha$

SPSS: $\alpha_p = 1 - \left(\dfrac{Sig}{2}\right)_{SPSS} = P\left[z \leq -z_{\alpha_p}\right] = P\left[z \geq z_{\alpha_p}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

h) Contraste unilateral para la diferencia de medias de dos distribuciones normales con varianzas poblaciones desconocidas e iguales $\sigma_1^2 = \sigma_2^2$, en muestras pequeñas $(n_1 + n_2) \leq 30$

Contraste unilateral a la derecha (cola a la derecha)

$$H_0: (\mu_1 - \mu_2) \leq 0 \qquad H_a: (\mu_1 - \mu_2) > 0$$

Se acepta H_0 si $t_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{s_p \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \leq t_{\alpha,(n_1+n_2-2)}$

Se rechaza H_0 si $t_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{s_p \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} > t_{\alpha,(n_1+n_2-2)}$

Media ponderada de cuasivarianzas:

$$s_p^2 = \frac{(n_1 - 1).s_1^2 + (n_2 - 1).s_2^2}{n_1 + n_2 - 2}$$

Se rechaza H_0 si $t_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{s_p \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} < -t_{\alpha,(n_1+n_2-2)}$

SPSS: $\alpha_p = \left(\dfrac{\text{Sig}}{2}\right)_{SPSS} = P\left[t_{(n_1+n_2-2)} \geq t_{\alpha_p}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

Contraste unilateral a la izquierda (cola a la izquierda)

$$H_0: (\mu_1 - \mu_2) \geq 0 \qquad H_a: (\mu_1 - \mu_2) < 0$$

Se acepta H_0 si $t_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{s_p \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \geq -t_{\alpha,(n_1+n_2-2)}$

Media ponderada de cuasivarianzas:

$$s_p^2 = \frac{(n_1 - 1).s_1^2 + (n_2 - 1).s_2^2}{n_1 + n_2 - 2}$$

Se rechaza H_0 si $t_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{s_p \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} < -t_{\alpha,(n_1+n_2-2)}$

SPSS: $\alpha_p = \left(\dfrac{\text{Sig}}{2}\right)_{SPSS} = P\left[t_{(n_1+n_2-2)} \leq -t_{\alpha_p}\right] = P\left[t_{(n_1+n_2-2)} \geq t_{\alpha_p}\right]$ cuando $\alpha_p > \alpha$ se acepta

i) Contraste unilateral para la diferencia de medias de dos distribuciones normales con varianzas poblaciones desconocidas y distintas $\sigma_1^2 \neq \sigma_2^2$, en muestras pequeñas $(n_1 + n_2) \leq 30$

Contraste unilateral a la derecha (cola a la derecha)

$$H_0: (\mu_1 - \mu_2) \leq 0 \qquad H_a: (\mu_1 - \mu_2) > 0$$

Se acepta H_0 si $t_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \leq t_{\alpha,f}$ Se rechaza H_0 si $t_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} > t_{\alpha,f}$

SPSS: $\alpha_p = \left(\dfrac{\text{Sig}}{2}\right)_{SPSS} = P\left[t_f \geq t_{\alpha_p}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

f es la aproximación de Welch: $f = \dfrac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1 + 1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2 + 1}} - 2$

🖃 Contraste unilateral a la izquierda (cola a la izquierda)

$$H_0 : (\mu_1 - \mu_2) \geq 0 \qquad H_a : (\mu_1 - \mu_2) < 0$$

Se acepta H_0 si $t_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \geq -t_{\alpha,\,f}$ Se rechaza H_0 si $t_{\alpha_p} = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} < -t_{\alpha,\,f}$

SPSS: $\alpha_p = \left(\dfrac{\text{Sig}}{2}\right)_{\text{SPSS}} = P\left[t_f \leq -t_{\alpha_p}\right] = P\left[t_f \geq t_{\alpha_p}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

j) Contraste unilateral para la diferencia de parámetros de dos distribuciones binomiales $B_1(n_1\,,\,p_1)$ y $B_2(n_2\,,\,p_2)$ en muestras grandes.

🖃 Contraste unilateral a la derecha (cola a la derecha)

$$H_0 : (p_1 - p_2) \leq 0 \qquad H_a : (p_1 - p_2) > 0$$

Se acepta H_0 si $z_p = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\hat{p}_1\,(1 - \hat{p}_1)}{n_1} + \dfrac{\hat{p}_2\,(1 - \hat{p}_2)}{n_2}}} \leq z_\alpha$

Se rechaza H_0 si $z_p = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\hat{p}_1\,(1 - \hat{p}_1)}{n_1} + \dfrac{\hat{p}_2\,(1 - \hat{p}_2)}{n_2}}} > z_\alpha$

SPSS: $\alpha_p = \left(\dfrac{\text{Sig}}{2}\right)_{\text{SPSS}} = P\left[z \geq z_{\alpha_p}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

🖃 Contraste unilateral a la izquierda (cola a la izquierda)

$$H_0 : (p_1 - p_2) \geq 0 \qquad H_a : (p_1 - p_2) < 0$$

Se acepta H_0 si $z_p = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\hat{p}_1\,(1 - \hat{p}_1)}{n_1} + \dfrac{\hat{p}_2\,(1 - \hat{p}_2)}{n_2}}} \geq -z_\alpha$

Se rechaza H_0 si $z_p = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\hat{p}_1\,(1 - \hat{p}_1)}{n_1} + \dfrac{\hat{p}_2\,(1 - \hat{p}_2)}{n_2}}} < -z_\alpha$

SPSS: $\alpha_p = 1 - \left(\dfrac{\text{Sig}}{2}\right)_{\text{SPSS}} = P\left[z \leq -z_{\alpha_p}\right] = P\left[z \geq z_{\alpha_p}\right]$ cuando $\alpha_p > \alpha$ se acepta H_0

k) Contraste unilateral para la diferencia de varianzas de dos poblaciones normales.

🖮 Contraste unilateral a la derecha (cola a la derecha)

$$H_0 : (\sigma_1^2 - \sigma_2^2) \leq 0 \qquad H_a : (\sigma_1^2 - \sigma_2^2) > 0$$

Se acepta H_0 si $F_{p_\alpha} = \dfrac{s_1^2}{s_2^2} \leq F_{\alpha;\,(n_1-1),(n_2-1)}$ Se rechaza H_0 si $F_{p_\alpha} = \dfrac{s_1^2}{s_2^2} > F_{\alpha;\,(n_1-1),(n_2-1)}$

🖮 Contraste unilateral a la izquierda (cola a la izquierda)

$$H_0 : (\sigma_1^2 - \sigma_2^2) \geq 0 \qquad H_a : (\sigma_1^2 - \sigma_2^2) < 0$$

Se acepta H_0 si $F_{p_\alpha} = \dfrac{s_1^2}{s_2^2} \geq F_{\alpha;\,(n_1-1),(n_2-1)}$ Se rechaza H_0 si $F_{p_\alpha} = \dfrac{s_1^2}{s_2^2} < F_{\alpha;\,(n_1-1),(n_2-1)}$

HIPÓTESIS SOBRE LA MEDIA DE UNA POBLACIÓN: REGIÓN DE RECHAZO

$X \approx N(\mu, \sigma)$ σ^2 conocida $z_\alpha = -z_{1-\alpha}$			
$H_0: \mu = \mu_0$	$H_1: \mu_1 > \mu_0$	$R = \left\{ (\bar{x} - \mu_0) > z_\alpha \dfrac{\sigma}{\sqrt{n}} \right\}$	unilateral simple derecha
$H_0: \mu = \mu_0$	$H_1: \mu_1 < \mu_0$	$R = \left\{ (\bar{x} - \mu_0) < z_{1-\alpha} \dfrac{\sigma}{\sqrt{n}} \right\}$	unilateral simple izquierda
$H_0: \mu = \mu_0$	$H_1: \mu > \mu_0$	$R = \left\{ (\bar{x} - \mu_0) > z_\alpha \dfrac{\sigma}{\sqrt{n}} \right\}$	unilateral compuesta derecha
$H_0: \mu = \mu_0$	$H_1: \mu < \mu_0$	$R = \left\{ (\bar{x} - \mu_0) < z_{1-\alpha} \dfrac{\sigma}{\sqrt{n}} \right\}$	unilateral compuesta izquierda
$H_0: \mu = \mu_0$	$H_1: \mu \neq \mu_0$	$R = \left\{ \lvert \bar{x} - \mu_0 \rvert > z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}} \right\}$	bilateral compuesta
$H_0: \mu \leq \mu_0$	$H_1: \mu > \mu_0$	$R = \left\{ (\bar{x} - \mu_0) > z_\alpha \dfrac{\sigma}{\sqrt{n}} \right\}$	unilateral compuesta derecha
$H_0: \mu \geq \mu_0$	$H_1: \mu < \mu_0$	$R = \left\{ (\bar{x} - \mu_0) < z_{1-\alpha} \dfrac{\sigma}{\sqrt{n}} \right\}$	unilateral compuesta

$X \approx N(\mu, \sigma)$ σ^2 desconocida $n.\sigma_x^2 = (n-1).\, s_x^2 \rightarrow \dfrac{s_x}{\sqrt{n}} = \dfrac{\sigma_x}{\sqrt{n-1}}$ $t_{\alpha, n} = -t_{1-\alpha, n}$			
$H_0: \mu = \mu_0$	$H_1: \mu_1 > \mu_0$	$R = \left\{ (\bar{x} - \mu_0) > t_{\alpha, (n-1)} \dfrac{s_x}{\sqrt{n}} \right\}$	unilateral simple derecha
$H_0: \mu = \mu_0$	$H_1: \mu_1 < \mu_0$	$R = \left\{ (\bar{x} - \mu_0) < t_{1-\alpha, (n-1)} \dfrac{s_x}{\sqrt{n}} \right\}$	unilateral simple izquierda
$H_0: \mu = \mu_0$	$H_1: \mu > \mu_0$	$R = \left\{ (\bar{x} - \mu_0) > t_{\alpha, (n-1)} \dfrac{s_x}{\sqrt{n}} \right\}$	unilateral compuesta derecha
$H_0: \mu = \mu_0$	$H_1: \mu \neq \mu_0$	$R = \left\{ \lvert \bar{x} - \mu_0 \rvert > t_{\alpha/2, (n-1)} \dfrac{s_x}{\sqrt{n}} \right\}$	bilateral compuesta
$H_0: \mu \leq \mu_0$	$H_1: \mu > \mu_0$	$R = \left\{ (\bar{x} - \mu_0) > t_{\alpha, (n-1)} \dfrac{s_x}{\sqrt{n}} \right\}$	unilateral compuesta derecha
$H_0: \mu \leq \mu_0$	$H_1: \mu < \mu_0$	$R = \left\{ (\bar{x} - \mu_0) < t_{1-\alpha, (n-1)} \dfrac{s_x}{\sqrt{n}} \right\}$	unilateral compuesta izquierda

HIPÓTESIS SOBRE LA IGUALDAD DE MEDIAS: REGIÓN DE RECHAZO

Igualdad de medias: $X \approx N(\mu_1, \sigma_1)$, $Y \approx N(\mu_2, \sigma_2)$	
$H_0: \mu_1 = \mu_2$ ((σ_1, σ_2 conocidas)	$R = \left\{ \lvert \bar{x} - \bar{y} \rvert > z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right\}$
$H_0: \mu_1 - \mu_2 = k$ (σ_1, σ_2 conocidas)	$R = \left\{ \lvert \bar{x} - \bar{y} - k \rvert > z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right\}$
$H_0: \mu_1 = \mu_2$ σ_1, σ_2 desconocidas pero iguales	$R = \left\{ \lvert \bar{x} - \bar{y} \rvert > t_{\alpha/2, (n_1 + n_2 - 2)}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right\}$
$H_0: \mu_1 = \mu_2$ σ_1, σ_2 desconocidas y distintas	$R = \left\{ \lvert \bar{x} - \bar{y} \rvert > t_{\alpha/2, f}\, s_p \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right\}$
$H_0: \mu_1 \leq \mu_2$ ((σ_1, σ_2 conocidas)	$R = \left\{ (\bar{x} - \bar{y}) > z_{\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right\}$
$H_0: (\mu_1 - \mu_2) \leq k$ (σ_1, σ_2 conocidas)	$R = \left\{ (\bar{x} - \bar{y} - k) > z_{\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right\}$
$H_0: \mu_1 \leq \mu_2$ ($\sigma_1 = \sigma_2$)	$R = \left\{ (\bar{x} - \bar{y}) > t_{\alpha, (n_1 + n_2 - 2)}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right\}$
$H_0: \mu_1 \leq \mu_2$ ($\sigma_1 \neq \sigma_2$)	$R = \left\{ (\bar{x} - \bar{y}) > t_{\alpha, f} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right\}$
$H_0: \mu_1 \geq \mu_2$ (σ_1, σ_2 conocidas)	$R = \left\{ (\bar{x} - \bar{y}) > z_{1-\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right\}$
$H_0: (\mu_1 - \mu_2) \geq k$ (σ_1, σ_2 conocidas)	$R = \left\{ (\bar{x} - \bar{y} - k) < -z_{\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right\}$
$H_0: \mu_1 \geq \mu_2$ ($\sigma_1 = \sigma_2$)	$R = \left\{ (\bar{x} - \bar{y}) > t_{1-\alpha, (n_1 + n_2 - 2)}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right\}$

HIPÓTESIS SOBRE LA IGUALDAD DE MEDIAS: REGIÓN DE RECHAZO

Igualdad de medias: $X \approx N(\mu_1, \sigma_1)$, $Y \approx N(\mu_2, \sigma_2)$

$H_0: \mu_1 \geq \mu_2 \quad (\sigma_1 = \sigma_2)$

$$R = \left\{ (\bar{x} - \bar{y}) > t_{1-\alpha, f}\; s_p \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right\}$$

HIPÓTESIS SOBRE LA IGUALDAD DE VARIANZAS: REGIÓN DE RECHAZO

$H_0: \sigma_1 = \sigma_2$	$R = \left\{ \frac{s_1^2}{s_2^2} \notin \left[F_{1-\alpha/2, (n_1-1), (n_2-1)}, F_{\alpha/2, (n_1-1), (n_2-1)} \right] \right\}$
$H_0: \sigma_1 \leq \sigma_2$	$R = \left\{ \frac{s_1^2}{s_2^2} > F_{\alpha, (n_1-1), (n_2-1)} \right\}$
$H_0: \sigma_1 \geq \sigma_2$	$R = \left\{ \frac{s_1^2}{s_2^2} < F_{1-\alpha, (n_1-1), (n_2-1)} \right\}$


Ejercicios de
Intervalos de Confianza

1. El peso (en gramos) de las cajas de cereales de una determinada marca sigue una distribución $N(\mu, 5)$. Se han tomado los pesos de 16 cajas seleccionadas aleatoriamente, y los resultados obtenidos han sido:

506, 508, 499, 503, 504, 510, 497, 512, 514, 505, 493, 496, 506, 502, 509, 496.

a) Obtener los intervalos de confianza del 90%, 95% y 99% para la media poblacional.

b) Determinar cuál sería el tamaño muestral necesario para conseguir, con un 95% de confianza, un intervalo de longitud igual a 2 gramos.

c) Suponiendo ahora que σ es desconocida, calcular los intervalos de confianza para la media al 90%, 95% y 99%.

Solución:

a) Se trata de construir un intervalo de confianza para la media poblacional μ de varianza conocida $\sigma^2 = 25$. El intervalo de confianza de nivel $(1-\alpha)$ viene dado por:

$$I_{1-\alpha}(\mu) = \Big[\overbrace{\bar{x}}^{\substack{\text{media}\\\text{muestral}}} \pm \overbrace{z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}^{\substack{\text{error}\\\text{estimación}}}\Big] \qquad \begin{cases} L_{1-\alpha} = 2\,.\,z_{\alpha/2}\,.\,\dfrac{\sigma}{\sqrt{n}} \;\to\; n = \left(\dfrac{2\,.\,z_{\alpha/2}\,.\,\sigma}{L_{1-\alpha}}\right)^2 \\ L = \text{longitud o amplitud} \qquad \text{Error estimación } \in = z_{z_{\alpha/2}}\dfrac{\sigma}{\sqrt{n}} \end{cases}$$

$$\bar{x} = \frac{1}{16}\sum_{i=1}^{16} x_i = 503,75$$

$$\begin{cases} 1-\alpha = 0,90 & \alpha = 0,10 & \alpha/2 = 0,05 & z_{0,05} = 1,645 \\ 1-\alpha = 0,95 & \alpha = 0,05 & \alpha/2 = 0,025 & z_{0,025} = 1,96 \\ 1-\alpha = 0,99 & \alpha = 0,01 & \alpha/2 = 0,005 & z_{0,005} = 2,575 \end{cases}$$

A medida que el nivel de confianza es mayor, aumenta longitud del intervalo.

$$P\left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = 1-\alpha$$

Los intervalos de confianza solicitados serán:

$$I_{0,90}(\mu) = \left[503,75 \pm 1,645\frac{5}{\sqrt{16}}\right] = \left[503,75 - 1,645\frac{5}{\sqrt{16}}\,,\, 503,75 + 1,645\frac{5}{\sqrt{16}}\right] = \left[501,69\,,\,505,81\right]$$

$$I_{0,95}(\mu) = \left[503,75 \pm 1,96\frac{5}{\sqrt{16}}\right] = \left[503,75 - 1,96\frac{5}{\sqrt{16}}\,,\, 503,75 + 1,96\frac{5}{\sqrt{16}}\right] = \left[501,30\,,\,506,20\right]$$

$$I_{0,99}(\mu) = \left[503,75 \pm 2,575\frac{5}{\sqrt{16}}\right] = \left[503,75 - 2,575\frac{5}{\sqrt{16}}\,,\, 503,75 + 2,575\frac{5}{\sqrt{16}}\right] = \left[500,53\,,\,506,97\right]$$

Longitud de cada uno de los intervalos de confianza:

$$L_{0,90}(\mu) = 505,81 - 501,69 = 4,12$$
$$L_{0,95}(\mu) = 506,20 - 501,30 = 4,9$$
$$L_{0,99}(\mu) = 506,97 - 500,53 = 6,44$$

b) La amplitud o longitud vendrá dado por la fórmula: $I_{1-\alpha}(\mu) = \left[\overline{x} \pm z_{\alpha/2}\dfrac{s_x}{\sqrt{n}}\right]$

$$\begin{pmatrix}\text{Amplitud o}\\ \text{Longitud}\end{pmatrix} = \left(\overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) - \left(\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 2\,.\,z_{\alpha/2}\,.\,\frac{\sigma}{\sqrt{n}} \;\rightarrow\; n = \left(\frac{2\,.\,z_{\alpha/2}\,.\,\sigma}{\text{Amplitud}}\right)^2$$

siendo $n = \left(\dfrac{2 \times 1{,}96 \times 5}{2}\right)^2 \approx 96$ cajas de cereales

c) Se trata de construir un intervalo de confianza para la media poblacional μ de varianza poblacional desconocida, con muestras pequeñas ($n \leq 30$).

El intervalo de confianza de nivel $(1-\alpha)$, viene dado por:

$$I_{1-\alpha}(\mu) = \left[\overline{x} \pm t_{\alpha/2,(n-1)}\frac{s_x}{\sqrt{n}}\right] \quad \begin{cases} 1-\alpha = 0{,}90 & \alpha = 0{,}10 & t_{0{,}05,15} = 1{,}753 \\ 1-\alpha = 0{,}95 & \alpha = 0{,}05 & t_{0{,}025,15} = 2{,}131 \\ 1-\alpha = 0{,}99 & \alpha = 0{,}01 & t_{0{,}005,15} = 2{,}947 \end{cases}$$

Cuasivarianza muestral: $s_x^2 = \dfrac{1}{15}\displaystyle\sum_{i=1}^{16}(x_i - \overline{x})^2 = 36{,}037 \;\rightarrow\; s_x \approx 6$ cuasidesviación típica

Los intervalos de confianza solicitados serán:

$$I_{0,90}(\mu) = \left[503{,}75 \pm 1{,}753\frac{6}{\sqrt{16}}\right] = \left[501{,}12\,,\ 506{,}38\right]$$

$$I_{0,95}(\mu) = \left[503{,}75 \pm 2{,}131\frac{6}{\sqrt{16}}\right] = \left[500{,}55\,,\ 506{,}95\right]$$

$$I_{0,99}(\mu) = \left[503{,}75 \pm 2{,}947\frac{6}{\sqrt{16}}\right] = \left[499{,}33\,,\ 508{,}17\right]$$

Señalar que a mayor nivel de confianza $(1-\alpha)$ mayor es la amplitud del intervalo, y, en consecuencia, los intervalos de confianza son mayores.

2. Una muestra aleatoria extraída de una población normal de varianza 100, presenta una media muestral $\overline{x} = 160$. Con una muestra de tamaño 144, se pide:

a) Calcular un intervalo de confianza del 95 por ciento para la media poblacional.

b) Si se quiere tener una confianza del 95 por ciento de que su estimación se encuentra a una distancia de 1,2 cm más o menos de la verdadera media poblacional, ¿cuántas observaciones adicionales deben tomarse?

Solución:

a) $n = 144 \quad \overline{x} = 160 \quad \sigma = 10 \quad 1-\alpha = 0,95 \quad \alpha/2 = 0,025 \quad z_{0,025} = 1,96$

$$I_{0,95}(\mu) = \left[160 - 1,96\,\frac{10}{12}\,,\ 160 + 1,96\,\frac{10}{12}\right] = \left[158,37\,,\ 161,63\right]$$

b) El error absoluto que se quiere cometer es de 1,2, aplicando la fórmula para la determinación de la muestra a un nivel de confianza del 95 por 100, se tiene:

$$\epsilon = z_{\alpha/2}\,\frac{\sigma}{\sqrt{n}} \ \rightarrow\ n = \left(\frac{z_{\alpha/2}\,.\,\sigma}{\epsilon}\right)^2 \ \rightarrow\ n = \left(\frac{1,96\,.\,10}{1,2}\right)^2 \approx 267$$

Se debería tomar una muestra adicional de $267 - 144 = 123$ elementos

3. La afluencia de visitantes al parque de Monfragüe durante un mes, medida a través de una muestra aleatoria durante 10 días elegidos aleatoriamente, han sido los siguientes:

682, 553, 555, 666, 657, 649, 522, 568, 700, 552

Suponiendo que los niveles de afluencia siguen una distribución normal, y que la desviación típica muestral es de 56,99.

a) Se podría afirmar, con un 95 por ciento de confianza, que la afluencia media al parque es de 600 personas al mes.

b) Los adjudicatarios de la explotación al parque, en negociaciones con la Junta de Extremadura, afirmaron que la afluencia media era constante y que la dispersión sería de unas 15 personas. ¿Queda esta afirmación probada con los datos disponibles con un 95% de confianza?

Solución:

a) Se trata intervalo de confianza para la media μ de una distribución normal de varianza poblacional desconocida siendo la muestra pequeña $n \leq 30$

$$I_{1-\alpha}(\mu) = \left[\overline{x} \pm t_{\alpha/2,\,(n-1)}\,\frac{s_x}{\sqrt{n}}\right] \quad \begin{cases} n\,.\,\sigma_x^2 = (n-1)\,.\,s_x^2 \ \rightarrow\ s_x^2 = \dfrac{n\,.\,\sigma_x^2}{n-1} \\ s_x^2 = \dfrac{10\,.\,56,99^2}{9} = 3608,73 \ \rightarrow\ s_x = \sqrt{3608,73} = 60,07 \\ 1-\alpha = 0,95 \quad \alpha/2 = 0,025 \quad t_{\alpha/2,\,(n-1)} = t_{0,025;9} = 2,262 \end{cases}$$

$$\overline{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 610,04 \qquad s_x = 60,07$$

$$I_{0,95}(\mu) = \left[610,04 \pm 2,262\ \frac{60,07}{\sqrt{10}}\right] = \left[610,04 - 2,262\ \frac{60,07}{\sqrt{10}}\ ,\ 610,04 + 2,262\ \frac{60,07}{\sqrt{10}}\right] =$$
$$= \left[567,07\ ,\ 653,01\right]$$

Como $567,07 \leq 600 \leq 653,01$ se puede afirmar que con un 95 por ciento de confianza la afluencia media es de 600 personas al mes.

b) Intervalo de confianza para la varianza σ^2 de una distribución normal:

$$I_{1-\alpha}(\sigma^2) = \left[\frac{(n-1).\ s_x^2}{\chi^2_{\alpha/2,(n-1)}}\ ,\ \frac{(n-1).\ s_x^2}{\chi^2_{1-\alpha/2,(n-1)}}\right] \quad \begin{cases} s_x^2 = 3608,73 \\ 1-\alpha = 0,95 \quad \chi^2_{\alpha/2,(n-1)} = \chi^2_{0,025,9} = 19,023 \\ \chi^2_{1-\alpha/2,(n-1)} = \chi^2_{0,975,9} = 2,70 \end{cases}$$

$$I_{0,95}(\sigma^2) = \left[\frac{9\ .\ 3608,73}{19,023}\ ,\ \frac{9\ .\ 3608,73}{2,70}\right] = \left[1707,33\ ,\ 12029,1\right]$$

$$I_{0,95}(\sigma) = \left[\sqrt{\frac{9\ .\ 3608,73}{19,023}}\ ,\ \sqrt{\frac{9\ .\ 3608,73}{2,70}}\right] = \left[\sqrt{1707,33}\ ,\ \sqrt{12029,1}\right] = \left[41,32\ ,\ 109,68\right]$$

$15 \notin \left[41,32\ ,\ 109,68\right] \rightarrow$ El intervalo de la desviación típica no contiene el valor 15, con lo cual no se puede afirmar con una confianza del 95% que la dispersión de afluencia sea de 15 personas.

4.a) Se desea tomar una muestra estratificada de las personas mayores de edad de un municipio, cuyos estratos son los siguientes intervalos de edades, en años: de 18 a 30, de 31 a 45, de 45 a 60 y mayores de 60. En el primer intervalo hay 7500 personas, en el segundo 8400, en el tercero 5700 y en el cuarto 3000. Calcula el tamaño de la muestra total y su composición, sabiendo que el muestreo se hace con afijación proporcional y se han elegido 375 personas del primer estrato.

b) Dada la población {2, 4, 6}, construye todas las muestras posibles de tamaño 2, que puedan formar mediante muestreo aleatorio simple, y halla la varianza de las medias muestrales de todas las muestras.

Solución:

a) Como la afijación es proporcional, el peso de cada estrato en la muestra es directamente proporcional a los individuos de la población correspondiente.

Si en el primer estrato, formado por 7500 personas se han elegido $n_1 = 375$

En el segundo estrato con 8400 personas se eligen $n_2 = \dfrac{8400 - 375}{7500} = 420$ personas

En el tercer estrato con 5700 personas se eligen $n_3 = \dfrac{5700 - 375}{7500} = 285$ personas

En el cuarto estrato con 3000 personas se eligen $n_4 = \dfrac{3000 - 375}{7500} = 150$ personas

El tamaño total de la muestra es $\displaystyle\sum_{i=1}^{4} n_i = 375 + 420 + 285 + 150 = 1230$ personas

Se elige 1 de cada 20 personas en cada estrato.

b) En el muestreo aleatorio simple se mantiene la probabilidad de extracción en cada caso. Por tanto, las extracciones deben hacerse con reemplazamiento.

El número de muestras de tamaño 2 que pueden obtenerse de la población {2, 4, 6} son 9:

{2, 2} ; {2, 4} ; {2, 6} ; {4, 2} ; {4, 4} ; {4, 6} ; {6, 2} ; {6, 4} y {6, 6}

Muestra	Elementos	Media de la muestras $\overline{x}_i$	$(\overline{x}_i - \overline{x})^2$
m_1	{2, 2}	2	4
m_2	{2, 4}	3	1
m_3	{2, 6}	4	0
m_4	{4, 2}	3	1
m_5	{4, 4}	4	0
m_6	{4, 6}	5	1
m_7	{6, 2}	4	0
m_8	{6, 4}	5	1
m_9	{6, 6}	6	4
		$\overline{x} = \frac{1}{9}\sum_{i=1}^{9} \overline{x}_i = 4$	$\sum_{i=1}^{9} (\overline{x}_i - \overline{x})^2 = 12$

La media y varianza de la población son:

$$\mu = \frac{2+4+6}{3} = 4 \qquad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{3} = \frac{8}{3}$$

Media de las muestras: $\overline{x} = \frac{1}{9}\sum_{i=1}^{9} \overline{x}_i = \frac{36}{9} = 4$

Varianza de las muestras: $\sigma_x^2 = \frac{1}{9}\sum_{i=1}^{9}(\overline{x}_i - \overline{x})^2 = \frac{12}{9} = \frac{4}{3} \quad \rightarrow \quad \sigma_x = \frac{2}{\sqrt{3}}$

Se verifica que $\mu = \overline{x} = 4$; y que $\sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{\frac{8}{3}}}{\sqrt{2}} = \sqrt{\frac{8}{6}} = \sqrt{\frac{4}{3}} = \frac{2}{\sqrt{3}}$

5. En una ciudad se desea estimar la proporción de hogares que reciclan sus envases de plástico. La ciudad está dividida en cuatro barrios (A, B, C y D) con 800, 2000, 1200 y 1000 hogares respectivamente.
Se selecciona mediante muestreo estratificado con afijación proporcional una muestra de 400 hogares.

a) ¿Cuántos hogares de cada uno de los barrios se incluirán en la muestra?

b) Si en el barrio B, 64 hogares de la muestra reciclan, ¿cuál es la estimación de hogares que reciclan en ese barrio?

c) Proporcionar un intervalo de confianza al 95% para la estimación puntual anterior.

Solución:

a) Total de hogares $= 800 + 2000 + 1200 + 1000 = 5000$ Muestra $= 400$

Muestreo estratificado con afijación proporcional: $\dfrac{400}{5000} = (0{,}08) = \dfrac{n_1}{800} = \dfrac{n_2}{2000} = \dfrac{n_3}{1200} = \dfrac{n_4}{1000}$

Muestreo estratificado se aplica el 8% a cada barrio $\begin{cases} \text{barrio A:} & 800 \times 0{,}08 = 64 \text{ hogares} \\ \text{barrio B:} & 2000 \times 0{,}08 = 160 \text{ hogares} \\ \text{barrio C:} & 1200 \times 0{,}08 = 96 \text{ hogares} \\ \text{barrio D:} & 1000 \times 0{,}08 = 80 \text{ hogares} \end{cases}$

b) $\hat{p}_B = \dfrac{64}{160} = 0{,}4$

c) Un intervalo de confianza para la estimación puntual del barrio B.

$n = 160 \quad \hat{p}_B = 0{,}4 \quad \hat{q}_B = 0{,}6 \quad 1 - \alpha = 0{,}95 \quad \alpha/2 = 0{,}025 \quad z_{0,025} = 1{,}96$

$$I_{0,95}(p_B) = \left[\hat{p}_B \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_B \,.\, \hat{q}_B}{n}}\right] = \left[0{,}4 \pm 1{,}96\sqrt{\frac{0{,}4 \,.\, 0{,}6}{160}}\right] = \left[0{,}4 \pm 0{,}076\right] = \left[0{,}324\ ,\ 0{,}476\right]$$

6. Una biblioteca desea estimar el porcentaje de libros infantiles que posee. La biblioteca está compuesta de 4 salas (Norte, Sur, Este y Oeste) con 2500, 2740, 4000 y 6900 libros, respectivamente. Se selecciona mediante muestreo estratificado aleatorio una muestra del 5% de los libros con afijación proporcional.

a) ¿Cuántos libros, de cada una de las salas hay en la muestra?

b) Si en la muestra de la sala Sur hay 30 libros infantiles, ¿Cuál es la estimación de la proporción de libros infantiles en esa sala?

c) Para un nivel de confianza del 90%, obtener el error máximo cometido con la estimación puntual anterior. Justificar las respuestas.

Solución:

a) Con afijación proporcional la muestra es el 5% de cada sala $\begin{cases} \text{Norte:} & 2500 \times 0{,}05 = 125 \text{ libros} \\ \text{Sur:} & 2740 \times 0{,}05 = 137 \text{ libros} \\ \text{Este:} & 4000 \times 0{,}05 = 200 \text{ libros} \\ \text{Oeste:} & 6900 \times 0{,}05 = 345 \text{ libros} \end{cases}$

b) $\hat{p}_{Sur} = \dfrac{30}{137} = 0{,}219$

c) Intervalo de confianza para la estimación puntual: $I_{0,90}(p_{Sur}) = \left[\hat{p}_{Sur} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}_{Sur}\,.\,\hat{q}_{Sur}}{n}}\right]$

$n = 137 \quad \hat{p}_{Sur} = 0,219 \quad \hat{q}_{Sur} = 0,781 \quad 1-\alpha = 0,90 \quad \alpha/2 = 0,05 \quad z_{0,05} = 1,645$

Error de estimación: $\in = z_{\alpha/2}\sqrt{\dfrac{\hat{p}_{Sur}\,.\,\hat{q}_{Sur}}{n}} = 1,645\sqrt{\dfrac{0,219\,.\,0,781}{137}} = 0,0581$

7. Se realiza un estudio sobre el tiempo de reacción de los conductores ante un imprevisto. Se considera una población de 10000 conductores, de los cuales 5000 tienen una antigüedad superior a 10 años, 3000 tienen una antigüedad entre 3 y 10 años y el resto tienen una antigüedad inferior a los 3 años. Se selecciona una muestra de 500 conductores mediante muestreo estratificado con afijación proporcional. Se pide, justificando la respuesta:

a) ¿Cuántos conductores de cada uno de los estratos mencionados anteriormente se incluirán en la muestra?

b) En los conductores con una antigüedad de menos de 3 años que resultan elegidos en la muestra, se observa que el tiempo medio de reacción es de 1,2 segundos. Supuesta que dicha variable tiene una distribución normal de desviación típica 0,3, proporcionar un intervalo de confianza al 95% para el tiempo medio de reacción de estos conductores.

Solución:

a) Muestreo estratificado con afijación proporcional: $\dfrac{500}{10000} = (0,05) = \dfrac{n_1}{5000} = \dfrac{n_2}{3000} = \dfrac{n_3}{2000}$

$$\begin{cases} n_1 = 5000 . 0,05 = 250 & \text{conductores de antigüedad superior a diez años} \\ n_2 = 3000 . 0,05 = 150 & \text{conductores de antigüedad entre tres y diez años} \\ n_3 = 2000 . 0,05 = 100 & \text{conductores de antigüedad inferior a tres años} \end{cases}$$

b) Intervalo de confianza para la media μ de una distribución normal $N(\mu, 0,3)$

$n = 100 \quad \overline{x} = 1,2 \quad 1-\alpha = 0,95 \quad \alpha = 0,05 \quad \alpha/2 = 0,025 \quad z_{0,025} = 1,96$

$$I_{0,95}(\mu) = \left[\overline{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = \left[1,2 \pm 1,96\,\frac{0,3}{\sqrt{100}}\right] = \left[1,2 \pm 0,059\right] = \left[1,141\,,1,259\right]$$

8. Una agencia de alquiler de automóviles necesita estimar el número medio de kilómetros diarios que realiza su flota de automóviles. A tal fin, en varios días de la semana toma los recorridos de cien vehículos de su flota y obtiene que la media muestral es de 165 km/día, y la cuasidesviación estándar muestral de 6 km/día. Se pide:

a) Bajo la hipótesis de normalidad de la característica en estudio (número de kilómetros por día), construir un intervalo de confianza para la media de dicha distribución a un nivel de confianza del 95 por 100.

b) Bajo la misma hipótesis de normalidad del apartado anterior, construir un intervalo de confianza del 90 por 100 para la varianza de dicha distribución.

Solución:

a) Se trata de construir un intervalo de confianza para la media de una distribución normal $N(\mu, \sigma)$ de varianza desconocida.

Como el tamaño de la muestra es grande, se tiene que el intervalo de confianza será:

$$I_{1-\alpha}(\mu) = \left[\bar{x} \pm z_{\alpha/2} \frac{s_x}{\sqrt{n}}\right]$$

En este caso: $\bar{x} = 165 \quad s_x = 6 \quad n = 100 \quad z_{0,025} = 1,96$

$$I(\mu) = \left[165 \pm 1,96 \,.\, \frac{6}{\sqrt{100}}\right] = \left[163,82 \ , \ 166,18\right]$$

b) El Intervalo de confianza para la varianza poblacional:

$$I_{1-\alpha}(\sigma^2) = \left[\frac{(n-1).\, s_x^2}{\chi^2_{\alpha/2,(n-1)}} \ , \ \frac{(n-1).\, s_x^2}{\chi^2_{1-\alpha/2,(n-1)}}\right] \qquad \chi^2_{0,05\,,\,99} = 124 \qquad \chi^2_{0,95\,,\,99} = 77$$

con lo cual, $I(\sigma^2) = \left[\dfrac{99 \,.\, 36}{124} \ , \ \dfrac{99 \,.\, 36}{77}\right] = \left[28,74 \ , \ 46,29\right]$

9. El gasto diario en llamadas telefónicas de dos departamentos X e Y de una misma empresa sigue una distribución normal, con gasto medio desconocido en ambos. Sin embargo, se conocen las desviaciones típicas, que son 100 y 110 céntimos de euro para X e Y, respectivamente. La dirección ha observado que una muestra aleatoria de 20 días, el gasto medio diario en llamadas realizadas por el departamento X ha sido de 1100 céntimos, y de 1400 en el departamento Y. Obtener un intervalo de confianza para la diferencia de gastos medios entre ambos departamentos.

Solución:

La variables aleatorias siguen, respectivamente, las distribuciones normales $N(\mu_1, 100)$ y $N(\mu_2, 110)$. El intervalo de confianza para la diferencia de medias $(\mu_1 - \mu_2)$ con varianzas poblacionales conocidas viene dado por la expresión:

$$I_{1-\alpha}(\mu_1 - \mu_2) = \left[(\bar{x} - \bar{y}) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right] \qquad \begin{cases} \sigma_1^2 = 100^2 \quad \sigma_2^2 = 110^2 \\ \bar{x} = 1100 \quad n_1 = 20 \quad \bar{y} = 1400 \quad n_2 = 20 \\ 1-\alpha = 0,90 \quad \alpha/2 = 0,05 \quad z_{\alpha/2} = 1,645 \end{cases}$$

$$I_{0,90}(\mu_1 - \mu_2) = \left[(1100-1400) \pm (1,645)\sqrt{\frac{100^2}{20}+\frac{110^2}{20}}\right] = \left[-354,68\,,-245,32\right]$$

El intervalo de confianza no cubre el 0 por 100, lo que indica que existe diferencia significativa en el gasto de llamadas telefónicas.

Como el intervalo de confianza es negativo, se deduce que el gasto medio en llamadas telefónicas del departamento Y es superior al del departamento X, con una confianza del 90 por ciento.

10. Se selecciona una muestra aleatoria de 600 familias, a las que se pregunta si tienen o no ordenador en casa. Contestaron afirmativamente 240 familias. Obtener un intervalo de confianza al nivel del 95% para la proporción real de familias que poseen ordenador en casa.

Solución:

La característica en estudio es dicotómica, hay que construir un intervalo de confianza para el parámetro p (proporción) de la variable aleatoria binomial asociada al estudio de la característica.

Como el tamaño de la muestra es suficientemente grande $n = 600$ se puede utilizar la aproximación normal.

$$I_{1-\alpha}(p) = \left[\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}.(1-\hat{p})}{n}}\right] \quad \begin{cases} \hat{p} = 240/600 = 0,4 \quad \hat{q} = 1-\hat{p} = 0,6 \quad n = 600 \\ 1-\alpha = 0,95 \quad \alpha = 0,05 \quad \alpha/2 = 0,025 \quad z_{\alpha/2} = z_{0,025} = 1,96 \end{cases}$$

$$I_{0,95}(p) = \left[0,4 \pm 1,96\sqrt{\frac{0,4\,.\,0,6}{600}}\right] = \left[0,36\;,\;0,44\right]$$

Con una confianza del 95% se puede afirmar que las familias poseen ordenador entre el 36% y el 44%.

11. Según los dirigentes del partido A, la intención de voto del partido rival B, en Andalucía, es la misma que la que tiene en Madrid. Se realiza una encuesta a 100 personas en Andalucía de los que 25 mostraron su apoyo al partido B, y a otras 100 personas en Madrid de las que 30 se inclinaron por el partido B.

a) Construir un intervalo de confianza del 90% para la proporción de personas que votarían al partido B en Andalucía

b) ¿A cuántas personas habría que encuestar para obtener un margen de error o error de estimación $\pm$ 2% al nivel de confianza anterior?.

c) Construir un intervalo de confianza al 90% para la diferencia de proporciones en la estimación del voto del partido B en las dos comunidades. ¿Se puede afirmar que los dirigentes del partido A tienen razón?.

Solución:

a) La característica en estudio en ambas comunidades es dicotómica, hay que construir un intervalo de confianza para el parámetro p_1 (proporción) de la variable aleatoria binomial asociada al estudio de la característica en la comunidad de Andalucía.

Como el tamaño de la muestra es suficientemente grande $n_1 = 100$ se puede utilizar la aproximación normal.

$$I_{1-\alpha}(p_1) = \left[\hat{p}_1 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1 \,.\, (1-\hat{p}_1)}{n}}\right] \quad \begin{cases} \hat{p}_1 = 25/100 = 0{,}25 \quad \hat{q}_1 = 1-\hat{p}_1 = 0{,}75 \quad n_1 = 100 \\ 1-\alpha = 0{,}90 \quad \alpha = 0{,}10 \quad \alpha/2 = 0{,}05 \quad z_{\alpha/2} = z_{0,05} = 1{,}645 \end{cases}$$

$$I_{0,90}(p_1) = \left[0{,}25 \pm 1{,}645\sqrt{\frac{0{,}25 \,.\, 0{,}75}{100}}\right] = \left[0{,}179 \,,\, 0{,}321\right]$$

En Andalucía la intención de voto del partido B se encuentra entre el 17,9% y 32,1%, con un nivel de confianza del 90%.

b) La amplitud o longitud vendrá dado por la fórmula:

$$I_{1-\alpha}(p) = \left[\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p} \,.\, (1-\hat{p})}{n}}\right] \quad \rightarrow \quad \epsilon^2 = \left(z_{\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}}\right)^2$$

de donde, $n = \dfrac{(z_{\alpha/2})^2 \,.\, (\hat{p}\,.\,\hat{q})}{\epsilon^2}$

El caso más desfavorable será cuando $\hat{p} = \hat{q} = 0{,}5$

Siendo $\epsilon^2 = (\pm\, 0{,}02)^2 = 0{,}0004 \quad \rightarrow \quad n = \dfrac{(1{,}645)^2 \,.\, (0{,}5 \,.\, 0{,}5)}{0{,}0004} \approx 1691$

c) Se trata de un intervalo de confianza para la diferencia de parámetros poblacionales $(p_1 - p_2)$ de dos distribuciones binomiales, con el tamaño de las muestras suficientemente grandes, $n_1 = n_2 = 100$ para utilizar la aproximación normal.

$$I_{1-\alpha}(p_1 - p_2) = \left[(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\,(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2\,(1-\hat{p}_2)}{n_2}}\right]$$

$$\begin{cases} \hat{p}_1 = 25/100 = 0{,}25 \quad \hat{q}_1 = 1-\hat{p}_1 = 0{,}75 \quad n_1 = 100 \\ \hat{p}_2 = 30/100 = 0{,}3 \quad \hat{q}_2 = 1-\hat{p}_2 = 0{,}70 \quad n_2 = 100 \\ 1-\alpha = 0{,}90 \quad \alpha = 0{,}10 \quad \alpha/2 = 0{,}05 \quad z_{\alpha/2} = z_{0,05} = 1{,}645 \end{cases}$$

$$I_{0,90}(p_1 - p_2) = \left[(0{,}25 - 0{,}3) \pm 1{,}645\sqrt{\frac{0{,}25 \,.\, 0{,}75}{100} + \frac{0{,}3 \,.\, 0{,}70}{100}}\right] = \left[-0{,}153 \,,\, 0{,}053\right]$$

El intervalo de confianza cubre el cero, lo que indica que no existe diferencia significativa entre la intención de voto del partido B en ambas comunidades, con lo cual los dirigentes del partido A tienen razón con una fiabilidad del 90%.

12. Una central de transformación de productos lácteos recibe diariamente leche de dos granjas A y B. Para analizar la calidad de la leche, que sigue una ley normal, se extraen dos muestras al azar y se analiza el contenido en materia grasa, obteniendo los resultados adjuntos en tantos por ciento:

Granja A	$\bar{x}_A = 8,7\%$	$s_A^2 = 1,02\,(\%)^2$	$n_A = 33$
Granja B	$\bar{x}_B = 10,9\%$	$s_B^2 = 1,73\,(\%)^2$	$n_B = 27$

Construir un intervalo de confianza del 95 por 100 para la diferencia del contenido medio en grasa de la leche de ambas granjas.

Solución:

Sea la variable aleatoria $X_A \equiv$ ' Contenido en grasa de la leche de la granja A' que sigue una distribución normal $N(\mu_A, \sigma_A)$. Análogamente, la variable aleatoria $X_B \equiv$ ' Contenido en grasa de la leche de la granja B' que sigue una distribución normal $N(\mu_B, \sigma_B)$.

Se trata de elaborar un intervalo de confianza para la diferencia de medias poblacionales $(\mu_A - \mu_B)$ con varianzas desconocidas y muestras grandes $n_A + n_B = 33 + 27 = 60 > 30$

$$I_{1-\alpha}(\mu_A - \mu_B) = \left[(\bar{x} - \bar{y}) \pm z_{\alpha/2} \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}} \right]$$

$1-\alpha = 0,95 \quad \alpha = 0,05 \quad \alpha/2 = 0,025 \quad z_{0,025} = 1,96$

$$I_{0,95}(\mu_A - \mu_B) = \left[(8,7 - 10,9) \pm 1,96 \sqrt{\frac{1,02}{33} + \frac{1,73}{27}} \right] = \left[-2,804\%\,,\ -1,596\% \right]$$

El intervalo no cubre el 0 % indicando que existe diferencia significativa entre el contenido en grasa de la leche de ambas granjas. Por otra parte, se observa un mayor contenido en grasa en la leche de la granja B.

13. Un instituto de investigaciones agronómicas siembra, en cinco parcelas diferentes, dos tipos de maíz híbrido. Las producciones en quintales métricos por hectárea son:

	1	2	3	4	5
Híbrido I	90	85	95	76	80
Híbrido II	84	87	90	92	90

a) Construir un intervalo de confianza para el cociente de varianzas con un error de significación de 0,10.

b) Construir un intervalo de confianza del 90% para la diferencia entre las producciones medias.

Solución:

a) Sea la variable aleatoria $X_1 \equiv$ 'Producción de maíz del híbrido I' que sigue una distribución normal $N(\mu_1, \sigma_1)$. Análogamente, la variable aleatoria $X_2 \equiv$ ' Producción de maíz del híbrido II' sigue una distribución normal $N(\mu_2, \sigma_2)$

Al construir un intervalo de confianza para el cociente de varianzas se puede concluir si las varianzas poblacionales desconocidas son o no distintas.

De modo que, si el intervalo de confianza para el cociente de varianzas $(\sigma_1^2 / \sigma_2^2)$ cubre al punto 1 podremos partir de que las varianzas son desconocidas pero iguales.

$$I_{1-\alpha}(\sigma_1^2 / \sigma_2^2) = \left[\frac{s_1^2 / s_2^2}{F_{\alpha/2\,;\,(n_1-1),(n_2-1)}} \,,\, \frac{s_1^2 / s_2^2}{F_{(1-\alpha/2)\,;\,(n_1-1),(n_2-1)}} \right]$$

siendo, $F_{1-\alpha/2\,;\,(n_1-1),(n_2-1)} = \dfrac{1}{F_{\alpha/2\,;\,(n_2-1),(n_1-1)}}$

En el caso,

$n_1 = 5 \quad \bar{x}_1 = 85,20 \quad s_1^2 = 57,7 \quad \alpha/2 = 0,05$

$n_2 = 5 \quad \bar{x}_2 = 88,6 \quad s_2^2 = 9,8 \quad s_1^2 / s_2^2 = 57,7/9,8 = 5,89$

$F_{0,05\,;\,4,4} = 6,3883 \quad F_{0,95\,;\,4,4} = 1/F_{0,05\,;\,4,4} = 1/6,3883 = 0,1565$

$$I_{0,90}(\sigma_1^2 / \sigma_2^2) = \left[\frac{5,89}{6,3883} \,,\, \frac{5,89}{0,1565} \right] = \left[0,92 \,,\, 37,64 \right]$$

El intervalo cubre el uno, y concluimos que las varianzas poblacionales son desconocidas e iguales, con una fiabilidad del 90%.

b) Se trata de un intervalo de confianza para la diferencia de medias poblacionales $(\mu_1 - \mu_2)$ con varianzas poblacionales desconocidas pero iguales, con muestras pequeñas $n_1 + n_2 = 10 < 30$.

$$I_{1-\alpha}(\mu_1 - \mu_2) = \left[(\bar{x} - \bar{y}) \pm t_{\alpha/2,\,(n_1+n_2-2)} \cdot s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right]$$

donde, $s_p^2 \equiv$ media ponderada de las cuasivarianzas muestrales:

$$s_p^2 = \frac{(n_1-1).s_1^2 + (n_2-1).s_2^2}{n_1 + n_2 - 2} \quad \rightarrow \quad s_p^2 = \frac{4\,.57,7 + 4\,.9,8}{5+5-2} = 33,75 \quad s_p = 5,81$$

$n_1 = n_2 = 5 \quad \bar{x}_1 = 85,20 \quad \bar{x}_2 = 88,6 \quad t_{\alpha/2,\,(n_1+n_2-2)} = t_{0,05\,,\,8} = 1,860$

$$I_{0,90}(\mu_1 - \mu_2) = \left[(85,20 - 88,6) \pm 1,860\,.\,5,81\,.\sqrt{\frac{1}{5} + \frac{1}{5}} \right] = \left[-10,23 \,,\, 3,43 \right]$$

El intervalo de confianza cubre el cero, por lo que no existe diferencia significativa entre las producciones medias, con una fiabilidad del 90%.

14. Un fabricante de televisores está desarrollando un nuevo modelo de televisor en color, y para este fin se pueden utilizar dos tipos de esquemas transistorizados. El fabricante selecciona una muestra de esquemas transistorizados del primer tipo de tamaño 16, y otra del segundo tipo de tamaño 13. Los datos muestrales respecto a la vida media de cada esquema son los siguientes:

$$\bar{x}_1 = 1400 \text{ horas} \quad s_1 = 30 \text{ horas} \quad n_1 = 16$$
$$\bar{x}_2 = 1500 \text{ horas} \quad s_2 = 17 \text{ horas} \quad n_2 = 13$$

Construir un intervalo de confianza del 90% para la diferencia de vida media de cada tipo de esquema.

Solución:

Sea la variable aleatoria $X_1 \equiv$ 'Vida media del primer esquema' que sigue una distribución normal $N(\mu_1, \sigma_1)$. Análogamente, la variable aleatoria $X_2 \equiv$ 'Vida media del segundo esquema sigue una distribución normal $N(\mu_2, \sigma_2)$

Hay que construir un intervalo de confianza para la diferencia de medias poblacionales $(\mu_1 - \mu_2)$ con varianzas poblacionales desconocidas, y se desconoce si distintas o no, siendo las muestras pequeñas $n_1 + n_2 = 29 < 30$

Para dilucidar si las varianzas poblacionales desconocidas son o no distintas, se construye primero un intervalo de confianza para el cociente de varianzas $(\sigma_1^2 / \sigma_2^2)$, de modo que si el intervalo cubre al punto 1 se puede partir de que las varianzas son desconocidas pero iguales.

Para construir un intervalo de confianza para el cociente de varianzas se emplea la fórmula:

$$I_{1-\alpha}(\sigma_1^2 / \sigma_2^2) = \left[\frac{s_1^2 / s_2^2}{F_{\alpha/2\,;\,(n_1-1),\,(n_2-1)}} \, , \, \frac{s_1^2 / s_2^2}{F_{(1-\alpha/2)\,;\,(n_1-1),\,(n_2-1)}} \right]$$

siendo, $F_{1-\alpha/2\,;\,(n_1-1),\,(n_2-1)} = \dfrac{1}{F_{\alpha/2\,;\,(n_2-1),\,(n_1-1)}}$

$$\begin{cases} s_1^2 = 30^2 = 900 \quad n_1 = 16 \quad\quad s_2^2 = 17^2 = 289 \quad n_2 = 13 \quad s_1^2 / s_2^2 = 900/289 = 3,114 \\ 1-\alpha = 0,90 \quad\quad \alpha/2 = 0,05 \end{cases}$$

$$F_{0,05\,;\,15\,,\,12} = 2,6169 \qquad F_{0,95\,;\,15\,,\,12} = 1/F_{0,05\,;\,12\,,\,15} = 1/2,4753 = 0,404$$

de donde, $I_{0,90}(\sigma_1^2 / \sigma_2^2) = \left[\dfrac{3,114}{2,6169} \, , \, \dfrac{3,114}{0,404} \right] = \left[1,19 \, , \, 7,71 \right]$

El intervalo no cubre el punto uno, y por tanto las varianzas poblacionales son desconocidas y distintas con una fiabilidad del 90%.

La situación es un intervalo de confianza para la diferencia de medias poblacionales $(\mu_1 - \mu_2)$ con varianzas poblacionales desconocidas y distintas o no, con muestras pequeñas $n_1 + n_2 < 30$

$$I_{1-\alpha}(\mu_1 - \mu_2) = \left[(\bar{x} - \bar{y}) \pm t_{\alpha/2,f} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right]$$

siendo,

$$s_1^2 = 30^2 = 900 \qquad n_1 = 16 \qquad s_1^2 / n_1 = 900/16 = 56{,}25 \qquad \left(s_1^2 / n_1\right)^2 = 3164{,}06$$

$$s_2^2 = 17^2 = 289 \qquad n_2 = 13 \qquad s_2^2 / n_2 = 289/13 = 22{,}23 \qquad \left(s_2^2 / n_2\right)^2 = 494{,}17$$

$$\frac{\left(s_1^2 / n_1\right)^2}{17} = 186{,}12 \qquad \frac{\left(s_2^2 / n_2\right)^2}{14} = 35{,}298 \qquad \left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2 = 6159{,}11$$

f es la aproximación de Welch:

$$f = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1 + 1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2 + 1}} - 2 = \frac{6159{,}11}{186{,}12 + 35{,}298} - 2 = 25{,}82 \approx 26$$

$$\bar{x}_1 = 1400 \text{ horas} \qquad \bar{x}_2 = 1500 \text{ horas} \qquad t_{\alpha/2,\, f} = t_{0,05\,,\,26} = 1{,}706$$

$$I_{0,90}(\mu_1 - \mu_2) = \left[(1400 - 1500) \pm 1{,}706\sqrt{\frac{900}{16} + \frac{289}{13}}\right] = \left[-115{,}113\,,\,-84{,}887\right]$$

El intervalo no cubre el cero, concluyendo que existe diferencia significativa entre la vida media de cada esquema, siendo mayor la vida media del segundo esquema con una fiabilidad del 90%.

15. Un equipo de investigación biológica está interesado en ver si una nueva droga reduce el colesterol en la sangre. Con tal fin toma una muestra de diez pacientes y determina el contenido de colesterol en la sangre antes y después del tratamiento. Los datos muestrales expresados en miligramos por 100 mililitros son los siguientes:

Paciente	1	2	3	4	5	6	7	8	9	10
Antes	217	252	229	200	209	213	215	260	232	216
Después	209	241	230	208	206	211	209	228	224	203

Construir un intervalo de confianza del 95 por 100 para la diferencia del contenido medio de colesterol en la sangre antes y después del tratamiento.

Solución:

Se trata de datos apareados en los que no existe independencia entre las muestras.

El intervalo de confianza al ser la muestra pequeña $n = 10 < 30$: $I = \left[\bar{d} \pm t_{\alpha/2,\,(n-1)} \dfrac{s_d}{\sqrt{n}}\right]$

$$d_i = x_i - y_i \qquad \bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i \qquad s_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$$

donde $\bar{d}$ es la media de las diferencias y s_d la desviación estándar de estas diferencias.

X = 'Antes'	217	252	229	200	209	213	215	260	232	216
Y = 'Después'	209	241	230	208	206	211	209	228	224	203
$d_i = x_i - y_i$	8	11	-1	-8	3	2	6	32	8	13

$n = 10 \quad \overline{d} = 7,40 \quad s_d^2 = 112,1481 \quad s_d = 10,59 \quad t_{\alpha/2,\,(n-1)} = t_{0,025\,,\,9} = 2,262$

$$I_{0,95}(\mu_1 - \mu_2) = \left[7,40 \pm 2,262\,.\,\frac{10,59}{\sqrt{10}}\right] = \left[-0,17\,,\,14,97\right]$$

El intervalo abarca el cero, por lo que no existe diferencia significativa en la diferencia del contenido medio del colesterol antes y después del tratamiento, con una fiabilidad del 95%.

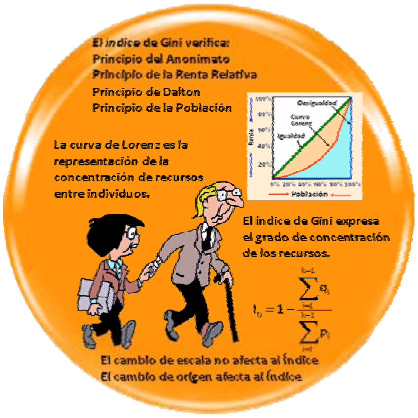

Tratamiento de Datos
Anova de un Factor con Medidas repetidas.
Análisis de la Varianza con varios factores.
Análisis Multivariable


Análisis de la Varianza con varios factores


ANÁLISIS VARIANZA


En cualquier experimento se pueden distinguir dos fases. Una primera, que consiste en la observación y análisis de los hechos que acontecen, y otra segunda, que interpreta y obtiene conclusiones.
La Estadística Descriptiva se encarga de dar una descripción numérica, ordenar y simplificar la información recogida en la primera fase.
La Estadística Teórica, con la información obtenida en el muestreo de la primera fase, pretende generalizar o predecir los resultados para toda la población.


Análisis Multivariante
AF
ACP
El Análisis de Componentes Principales (ACP) tiene como objetivo simplificar la estructura de los datos transformando las variables en análisis en el menor número posible de factores comunes o componentes principales que sean combinaciones lineales de las variables.


El índice de Gini verifica:
Principio del Anonimato
Principio de la Renta Relativa
Principio de Dalton
Principio de la Población
La curva de Lorenz es la representación de la concentración de recursos entre individuos.
El índice de Gini expresa el grado de concentración de los recursos.
El cambio de escala no afecta al índice
El cambio de origen afecta al índice


SPSS
BONDAD DE AJUSTE
PRUEBAS PARAMÉTRICAS
PRUEBAS NO PARAMÉTRICAS


Análisis Multivariante
INVESTIGACIÓN en Cómic
Análisis Factorial
Componentes Principales (ACP)
Análisis Cluster
SPSS

Universidad Autónoma de Madrid

Gestión Aeronáutica: Estadística Teórica
Facultad Ciencias Económicas y Empresariales
Departamento de Economía Aplicada
Profesor: Santiago de la Fuente Fernández