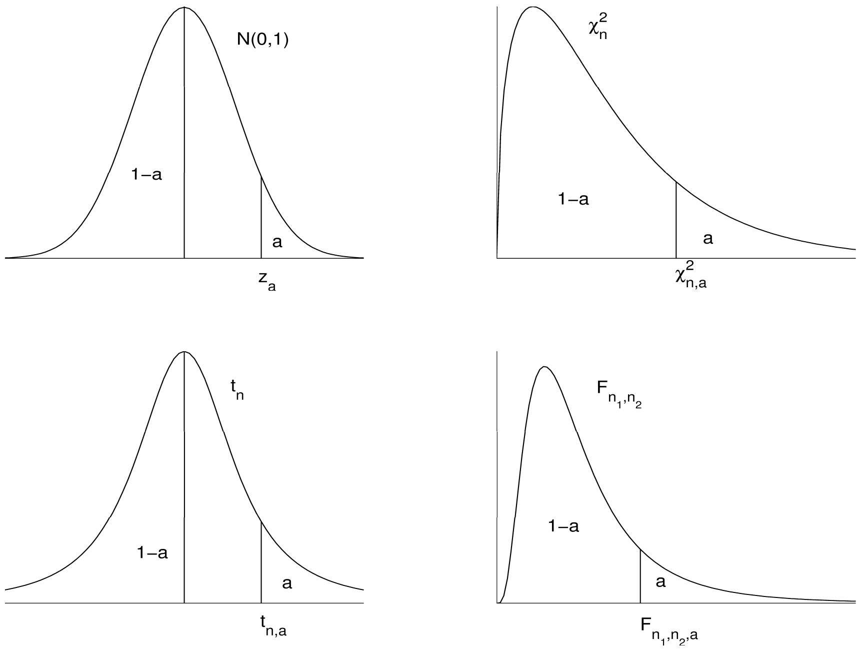
NOTACIÓN:



$(X_1, \ldots, X_n)$ muestra aleatoria de X: $\bar{x} = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} x_i$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

INTERVALOS DE CONFIANZA

1) $X \sim N(\mu, \sigma)$.

* Intervalos de confianza $1 - \alpha$ para μ :

$$I = \left[\bar{x} \mp t_{n-1;\alpha/2} \, \frac{s}{\sqrt{n}} \right]$$

* Intervalo de confianza $1 - \alpha$ para σ^2 :

$$I = \left[\frac{(n-1)s^2}{\chi^2_{n-1;\alpha/2}}, \frac{(n-1)s^2}{\chi^2_{n-1;1-\alpha/2}} \right]$$

2) $X \sim B(1, p)$(muestras grandes).

Intervalo de confianza $1 - \alpha$ para p :

$$I = \left[\bar{x} \mp z_{\alpha/2} \sqrt{\bar{x}(1-\bar{x})/n} \right]$$

3) <u>Intervalo de confianza para la media de una población no necesariamente normal (muestras grandes)</u>

$$I = \left[\bar{x} \mp z_{\alpha/2}\,\frac{s}{\sqrt{n}}\right]$$

4) <u>Dos poblaciones normales independientes.</u>

$X \sim N(\mu_1, \sigma_1)$; $(X_1, \ldots, X_{n_1})$ m. a. de X; se calcula $\bar{x}$ y s_1^2.
$Y \sim N(\mu_2, \sigma_2)$; $(Y_1, \ldots, Y_{n_2})$ m. a. de Y; se calcula $\bar{y}$ y s_2^2.

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

* Intervalos de confianza $1-\alpha$ para $\mu_1 - \mu_2$:

a) σ_1, σ_2 desconocidas, $\sigma_1 = \sigma_2$:

$$I = [\bar{x} - \bar{y} \mp t_{n_1+n_2-2;\alpha/2}\, s_p\, \sqrt{1/n_1 + 1/n_2}]$$

b) σ_1, σ_2 desconocidas, $\sigma_1 \neq \sigma_2$:

$$I = [\bar{x} - \bar{y} \mp t_{f;\alpha/2}\, \sqrt{s_1^2/n_1 + s_2^2/n_2}]$$

donde f es el entero más próximo a

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

* Intervalo de confianza $1-\alpha$ para σ_1^2/σ_2^2 :

$$I = \left[\frac{s_1^2/s_2^2}{F_{n_1-1,n_2-1;\alpha/2}}, \frac{s_1^2}{s_2^2}F_{n_2-1,n_1-1;\alpha/2}\right]$$

5) <u>Comparación de proporciones (muestras grandes e independientes).</u>

$X \sim B(1, p_1)$; $(X_1, \ldots, X_{n_1})$ m. a. de X.
$Y \sim B(1, p_2)$; $(Y_1, \ldots, Y_{n_2})$ m. a. de Y.

Intervalo de confianza $1-\alpha$ para $p_1 - p_2$:

$$I = \left[\bar{x} - \bar{y} \mp z_{\alpha/2}\sqrt{\frac{\bar{x}(1-\bar{x})}{n_1} + \frac{\bar{y}(1-\bar{y})}{n_2}}\right]$$

CONTRASTES DE HIPÓTESIS

NOTACION:
α= nivel de significación del contraste.
n= tamaño de la muestra.
H_0= hipótesis nula.
R= región crítica o de rechazo de H_0.

1) $X \sim N(\mu, \sigma)$.

$H_0 : \mu = \mu_0$ (σ desconocida); $\quad R = \left\{ |\bar{x} - \mu_0| > t_{n-1;\alpha/2} \frac{s}{\sqrt{n}} \right\}$

$H_0 : \mu \leq \mu_0$ (σ desconocida); $\quad R = \left\{ \bar{x} - \mu_0 > t_{n-1;\alpha} \frac{s}{\sqrt{n}} \right\}$

$H_0 : \mu \geq \mu_0$ (σ desconocida); $\quad R = \left\{ \bar{x} - \mu_0 < t_{n-1;1-\alpha} \frac{s}{\sqrt{n}} \right\}$ $(t_{n-1;1-\alpha} = -t_{n-1;\alpha})$

$H_0 : \sigma = \sigma_0$; $\quad R = \left\{ \frac{n-1}{\sigma_0^2} s^2 \notin \left[\chi^2_{n-1;1-\alpha/2} \ , \ \chi^2_{n-1;\alpha/2} \right] \right\}$

$H_0 : \sigma \leq \sigma_0$; $\quad R = \left\{ \frac{n-1}{\sigma_0^2} s^2 > \chi^2_{n-1;\alpha} \right\}$

$H_0 : \sigma \geq \sigma_0$; $\quad R = \left\{ \frac{n-1}{\sigma_0^2} s^2 < \chi^2_{n-1;1-\alpha} \right\}$

2) $X \sim B(1, p)$ (muestras grandes)

$H_0 : p = p_0$; $\quad R = \left\{ |\bar{x} - p_0| > z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}} \right\}$

$H_0 : p \leq p_0$; $\quad R = \left\{ \bar{x} - p_0 > z_{\alpha} \sqrt{\frac{p_0(1-p_0)}{n}} \right\}$

$H_0 : p \geq p_0$; $\quad R = \left\{ \bar{x} - p_0 < z_{1-\alpha} \sqrt{\frac{p_0(1-p_0)}{n}} \right\}$ $(z_{1-\alpha} = -z_{\alpha})$

3) Contrastes para la media de una población no necesariamente normal (muestras grandes)

$H_0 : \mu = \mu_0$ (σ desconocida); $\quad R = \left\{ |\bar{x} - \mu_0| > z_{\alpha/2} \frac{s}{\sqrt{n}} \right\}$

$H_0 : \mu \leq \mu_0$ (σ desconocida); $\quad R = \left\{ \bar{x} - \mu_0 > z_{\alpha} \frac{s}{\sqrt{n}} \right\}$

$H_0 : \mu \geq \mu_0$ (σ desconocida); $\quad R = \left\{ \bar{x} - \mu_0 < z_{1-\alpha} \frac{s}{\sqrt{n}} \right\}$ $(z_{1-\alpha} = -z_{\alpha})$

4) Dos poblaciones normales independientes.

$X \sim N(\mu_1, \sigma_1)$; $(X_1, \ldots, X_{n_1})$ m. a. de X; se calcula $\bar{x}$ y s_1^2.
$Y \sim N(\mu_2, \sigma_2)$; $(Y_1, \ldots, Y_{n_2})$ m. a. de Y; se calcula $\bar{y}$ y s_2^2.

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

$H_0 : \mu_1 = \mu_2$ $(\sigma_1 = \sigma_2)$; $\quad R = \left\{ |\bar{x} - \bar{y}| > t_{n_1+n_2-2;\alpha/2} \ s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right\}$

$H_0 : \mu_1 = \mu_2$ $(\sigma_1 \neq \sigma_2)$; $\quad R = \left\{ |\bar{x} - \bar{y}| > t_{f;\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right\}$

$H_0 : \mu_1 \leq \mu_2$ $(\sigma_1 = \sigma_2)$; $\quad R = \left\{ \bar{x} - \bar{y} > t_{n_1+n_2-2;\alpha} \ s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right\}$

$H_0 : \mu_1 \leq \mu_2$ $(\sigma_1 \neq \sigma_2)$; $\quad R = \left\{ \bar{x} - \bar{y} > t_{f;\alpha} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right\}$

$$H_0: \mu_1 \geq \mu_2 \ (\sigma_1 = \sigma_2); \qquad R = \left\{ \bar{x} - \bar{y} < t_{n_1+n_2-2;1-\alpha} \ s_p \sqrt{\tfrac{1}{n_1} + \tfrac{1}{n_2}} \right\}$$

$$H_0: \mu_1 \geq \mu_2 \ (\sigma_1 \neq \sigma_2); \qquad R = \left\{ \bar{x} - \bar{y} < t_{f;1-\alpha} \sqrt{\tfrac{s_1^2}{n_1} + \tfrac{s_2^2}{n_2}} \right\}$$

$$H_0: \sigma_1 = \sigma_2; \qquad R = \left\{ s_1^2/s_2^2 \notin \left[F_{n_1-1;n_2-1;1-\alpha/2} \ , \ F_{n_1-1;n_2-1;\alpha/2} \right] \right\}$$

$$H_0: \sigma_1 \leq \sigma_2; \qquad R = \left\{ s_1^2/s_2^2 > F_{n_1-1;n_2-1;\alpha} \right\}$$

$$H_0: \sigma_1 \geq \sigma_2; \qquad R = \left\{ s_1^2/s_2^2 < F_{n_1-1;n_2-1;1-\alpha} \right\}$$

$$\text{donde } f = \text{entero más próximo a } \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$$

5) <u>Comparación de proporciones (muestras grandes e independientes).</u>

$X \sim B(1, p_1)$; $(X_1, \ldots X_{n_1})$ m. a. de X.
$Y \sim B(1, p_2)$; $(Y_1, \ldots Y_{n_2})$ m. a. de Y.

$$H_0: p_1 = p_2; \qquad R = \left\{ |\bar{x} - \bar{y}| > z_{\alpha/2} \sqrt{\bar{p}(1-\bar{p}) \left(\tfrac{1}{n_1} + \tfrac{1}{n_2} \right)} \right\}$$

$$H_0: p_1 \leq p_2; \qquad R = \left\{ \bar{x} - \bar{y} > z_{\alpha} \sqrt{\bar{p}(1-\bar{p}) \left(\tfrac{1}{n_1} + \tfrac{1}{n_2} \right)} \right\}$$

$$H_0: p_1 \geq p_2; \qquad R = \left\{ \bar{x} - \bar{y} < z_{1-\alpha} \sqrt{\bar{p}(1-\bar{p}) \left(\tfrac{1}{n_1} + \tfrac{1}{n_2} \right)} \right\}$$

$$\text{donde } \bar{p} = \frac{\sum x_i + \sum y_i}{n_1 + n_2} = \frac{n_1 \bar{x} + n_2 \bar{y}}{n_1 + n_2}$$

REGRESIÓN LINEAL SIMPLE

Modelo: $y_i = a + bx_i + e_i,\ i = 1, \ldots, n$

Estadísticos básicos, notaciones:

$v_x = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2,\ v_y = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2,\ cov_{x,y} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$

Estimadores de a y b:

$\hat{b} = \frac{cov_{x,y}}{v_x} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2},\ \hat{a} = \bar{y} - \hat{b}\bar{x}$

Estimación de la varianza residual ($\sigma^2 = V(e)$) y del coeficiente de correlación ρ:

$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i-\hat{a}-\hat{b}x_i)^2}{n-2}$

$\hat{\rho} = \frac{cov_{x,y}}{\sqrt{v_x v_y}}$

Contrastes sobre el coeficiente de regresión: Se supone que las e_i son vaiid $N(0, \sigma)$.

$H_0 : b = 0; \qquad R = \{|t| > t_{n-2;\alpha/2}\}$

$H_0 : b \geq 0; \qquad R = \{t < t_{n-2;1-\alpha}\}$

$H_0 : b \leq 0; \qquad R = \{t > t_{n-2;\alpha}\}$

siendo $t = \frac{\hat{b}}{\hat{\sigma}/\sqrt{(\sum(x_i-\bar{x})^2}}$

Intervalo de predicción para la observación y_0 cuando $x = x_0$.

Para un nivel de confianza $1 - \alpha$:

$$I = \left[\hat{y}_0 \pm t_{n-2;\alpha/2}\ \hat{\sigma}\ \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}}\right]$$

siendo $\hat{y}_0 = \hat{a} + \hat{b}x_0$.

Intervalo de confianza para la media de y_0 cuando $x = x_0$.

Para un nivel de confianza $1 - \alpha$:

$$I = \left[\hat{y}_0 \pm t_{n-2;\alpha/2}\ \hat{\sigma}\ \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}}\right]$$