Intervalos de Confianza

Juan Sosa, PhD

Modelo

$$X_1, X_2, \ldots, X_n \stackrel{\text{iid}}{\sim} \text{Ber}(\pi)$$

Intervalo de confianza (bilaterial)

$$\text{IC}_{100(1-\alpha)\%}\,(\pi) = \hat{\pi} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$$

Nota: Esta aproximación es apropiada para $n \geq 30$, $n\hat{\pi} \geq 5$ y $n(1-\hat{\pi}) \geq 5$.

Tamaño de muestra

$$n = \frac{\left(z_{1-\alpha/2}\right)^2 \pi_0(1-\pi_0)}{ME^2}$$

donde π_0 es la proporción muestral de una de un estudio piloto (por ejemplo) y ME es el margen de error.

Modelo

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \mathsf{N}(\mu, \sigma^2)$$

Intervalo de confianza (bilaterial)

$$\mathsf{IC}_{100(1-\alpha)\%}\,(\mu) = \bar{x} \pm \mathsf{t}_{n-1,1-\alpha/2}\,\frac{s}{\sqrt{n}}$$

Nota: Para $n \geq 30$ se tiene que $\mathsf{t} \approx \mathsf{N}(0,1)$ y la población no tiene que ser Normal.

Tamaño de muestra

$$n = \frac{\left(z_{1-\alpha/2}\right)^2 \sigma_0^2}{ME^2}$$

donde σ_0^2 es la varianza muestral de un estudio piloto (por ejemplo) y ME es el margen de error.

Modelo

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \mathsf{N}(\mu, \sigma^2)$$

Intervalo de confianza (bilateral)

$$\mathsf{IC}_{100(1-\alpha)\%}\left(\sigma^2\right) = \left(\frac{(n-1)s^2}{\chi^2_{n-1,1-\alpha/2}} \,;\, \frac{(n-1)s^2}{\chi^2_{n-1,\alpha/2}} \right)$$

donde χ^2_{n-1} denota la distribución chi cuadrado con $n-1$ grados de libertad.

Modelo

Poblaciones independientes:

$$X_1, X_2, \ldots, X_{n_1} \overset{\text{iid}}{\sim} \text{Ber}(\pi_1) \qquad Y_1, Y_2, \ldots, Y_{n_2} \overset{\text{iid}}{\sim} \text{Ber}(\pi_2)$$

Intervalo de confianza (bilateral)

$$\text{IC}_{100(1-\alpha)\%}(\pi_1 - \pi_2) = \hat{\pi}_1 - \hat{\pi}_2 \pm z_{1-\alpha/2}\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}$$

Nota: Esta aproximación es apropiada para

$$n_1, n_2 \geq 30 \quad n_1\hat{\pi}_1, n_1\hat{\pi}_2 \geq 5 \quad n_1(1-\hat{\pi}_1), n_2(1-\hat{\pi}_2) \geq 5$$

Modelo

Poblaciones independientes:

$$X_1, X_2, \ldots, X_{n_1} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_1, \sigma_1^2) \qquad Y_1, Y_2, \ldots, Y_{n_2} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_2, \sigma_2^2)$$

Intervalo de confianza (bilateral) bajo homogenedidad ($\sigma_1^2 = \sigma_2^2$)

$$\mathsf{IC}_{100(1-\alpha)\%}\,(\mu_1 - \mu_2) = \bar{x_1} - \bar{x}_2 \pm t_{n_1+n_2-2,1-\alpha/2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

donde

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

es la varianza conjugada de las muestras.
Nota: Para $n \geq 30$ se tiene que t $\approx \mathsf{N}(0, 1)$ y la población no tiene que ser Normal.

Modelo

Poblaciones independientes:

$$X_1, X_2, \ldots, X_{n_1} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_1, \sigma_1^2) \qquad Y_1, Y_2, \ldots, Y_{n_2} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_2, \sigma_2^2)$$

Intervalo de confianza (bilateral) bajo heterogenedidad ($\sigma_1^2 \neq \sigma_2^2$)

$$\mathsf{IC}_{100(1-\alpha)\%}\,(\mu_1 - \mu_2) = \bar{x_1} - \bar{x}_2 \pm t_{\nu,1-\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

donde

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

corresponde a los grados de libertad.

Nota: Para $n \geq 30$ se tiene que t $\approx \mathsf{N}(0, 1)$ y la población no tiene que ser Normal.

Modelo

Poblaciones independientes:

$$X_1, X_2, \ldots, X_{n_1} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_1, \sigma_1^2) \qquad Y_1, Y_2, \ldots, Y_{n_2} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_2, \sigma_2^2)$$

Intervalo de confianza (bilateral)

$$\mathsf{IC}_{100(1-\alpha)\%}\left(\frac{\sigma_1^2}{\sigma_2^2}\right) = \left(F_{\alpha/2, n_2-1, n_1-1}\frac{s_1^2}{s_2^2}; F_{1-\alpha/2, n_2-1, n_1-1}\frac{s_1^2}{s_2^2}\right)$$

donde F_{n_2-1, n_1-1} denota la distribución F con $n_2 - 1$ grados de libertad en el numerador y $n_1 - 1$ grados de libertad en el denominador.

Distribución F (Fisher-Snedecor)

La función de densidad de una variable aleatoria F con r_1 grados de libertad del numerador y r_2 grados de libertad del denominador es:

$$f(x) = \frac{(r_1/r_2)^{r_1/2}\Gamma[(r_1+r_2)/2]x^{(r_1/2)-1}}{\Gamma[r_1/2]\Gamma[r_2/2][1+(r_1x/r_2)]^{(r_1+r_2)/2}} \qquad x \geq 0$$

donde $\Gamma(z) = \int_0^\infty t^{z-1}\exp\{-t\}\,dt$ es la función gamma.