Intervalo de Confianza para Varianza y Media

La construcción del intervalo de confianza para una varianza se basa en el hecho de que, para una muestra de tamaño n, se tiene:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

A partir de esto, por construcción, se satisface:

$$P[\chi^2_{n-1,\alpha/2} < \chi^2 < \chi^2_{n-1,1-\alpha/2}] = 1-\alpha$$

Sustituyendo la primera expresión

$$P[\chi^2_{n-1,\alpha/2} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{n-1,1-\alpha/2}] = 1-\alpha$$

Mediante un proceso algebraico, se puede demostrar que esto equivale a

$$P\left[\frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}}\right] = 1-\alpha$$

Intervalo de confianza σ^2
Sea S^2 el estimador de σ^2 en una muestra de tamaño n. El intervalo de confianza para σ^2 al nivel $1-\alpha$ es $$IC_{\sigma^2} = \left[\frac{(n-1)S^2}{X^2_{n-1,1-\alpha/2}}, \frac{(n-1)S^2}{X^2_{n-1,\alpha/2}}\right]$$

Ejemplo 1. *Se tomó una muestra aleatoria de* n = 114 *estudiantes de sexo masculino de nivel licenciatura, a partir de la cual se estimó un peso (en kg.) medio de* $\bar{x}$ = 75.0588 *y una varianza de* s^2 = 75.0588.

El intervalo de confianza para la varianza al nivel $1 - \alpha = 0.90$ *es*

$$IC_{\sigma^2} = \left[\frac{(113)(160.5513)}{138.8114}, \frac{(113)(160.5513)}{89.4605}\right]$$
$$= (130.6975, 202.7967)$$

donde:

$$X^2_{113,0.05} = 89.4604 \text{ y } x^2_{113,0.95} = 138.8114$$

El intervalo de confianza para σ *es*

$$IC_\sigma = \left(\sqrt{130.6975}, \sqrt{202.7967}\right) = (11.4323, 14.2407)$$

Intervalo de confianza para el cociente de varianza

Si $X_1 \sim \chi^2_m$ y $X_2 \sim \chi^2_n$, se puede demostrar que

$$\frac{X_1/m}{X_2/n} \sim F_{m,n}$$

Además, se sabe que:

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Entonces, si se tienen dos muestras aleatorias de tamaños n_1 y n_2 de poblaciones con varianzas σ_1^2 y σ_2^2 respectivamente, se sigue:

$$\frac{\dfrac{(n_2-1)S_2^2}{(n_2-1){\sigma_2}^2}}{\dfrac{(n_1-1)S_1^2}{(n_1-1)\sigma_1^2}} = \frac{\dfrac{S_2^2}{\sigma_2^2}}{\dfrac{S_1^2}{{\sigma_1}^2}} = \frac{\sigma_1^2 S_2^2}{{\sigma_2}^2 S_1^2} \sim F_{n_2-1,n_1-1}$$

Por tanto

$$P\left[F_{n_2-1,n_1-1,\alpha/2} < \frac{\sigma_1^2 S_2^2}{{\sigma_2}^2 S_1^2} < F_{n_2-1,n_1-1,1-\alpha/2}\right] = 1-\alpha$$

Mediante un proceso análogo a la construcción del intervalo de confianza para la varianza, se llega al siguiente resultado:

Intervalo de confianza para el cociente σ_1^2/σ_2^2
Sea S_1^2 y S_2^2 el estimadores de σ_1^2 y σ_2^2 en m muestras de tamaño n_1 y n_2 respectivamente. El intervalo de confianza para el cociente de varianzas al nivel $1-\alpha$ es $$IC_{\sigma_1^2/\sigma_2^2} = \left[\frac{S_1^2}{S_2^2 F_{n_1-1,n_2-1,1-\alpha/2}}, \frac{S_1^2}{S_2^2 F_{n_1-1,n_2-1,1-\alpha/2}}\right]$$

Ejemplo 2. *Se tomaron muestras aleatorias de* n_1 = 40 *estudiantes de sexo masculino de la facultad de ingeniería y* n_2 = 34 *de la facultad de sistemas nivel licenciatura. Se estimaron las varianzas del peso (en kg.), donde se obtuvo* s_1^2 = 95.26923 *y* s_2^2 = 214.9748 *respectivamente.*

El intervalo de confianza para la varianza al nivel $1-\alpha=0.90$ para el cociente de varianza $\frac{\sigma_1^2}{\sigma_2^2}$ es:

$$\begin{aligned} IC_{\sigma_1^2/\sigma_2^2} &= \left[\frac{95.26923}{(214.9748)(1.7596)}, \frac{95.26923}{(214.9748)(0.5769)}\right] \\ &= (0.2518, 0.7682) \end{aligned}$$

donde $F_{39,33,0.05}$ = 0.9616 *y* $F_{39,33,0.95}$ = 1.0466.

El intervalo de confianza no contiene el valor de 1, lo que estadísticamente supondría igualdad de varianzas.

Intervalo de confianza para una media

Sea $\{X_1\}_{i=1}^{n}$ una muestra aleatoria de la $VA\ \ X\ \sim D(\mu,\sigma)$. A partir del Teorema Central del Límite se cumple que, para muestras grandes:

$$Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Si no se conoce σ se utiliza el estimador $S = \hat{\sigma}$ por lo que:

$$T = \frac{\bar{x}-\mu}{S/\sqrt{n}} \sim t_{n-1}$$

Por tanto

$$P\left[t_{n-1,\alpha/2} < \frac{\bar{x}-\mu}{S/\sqrt{n}} < t_{n-1,1-\alpha/2}\right] = 1-\alpha$$

Lo que equivale a

$$P\left[\bar{x} - t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}} < \bar{x} < t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}}\right] = 1-\alpha$$

Intervalo de confianza para la media μ

Sea $\{X_i\}_{i=1}^{n}$ una muestra aleatoria con estimadores $\bar{x}$ y S para media y desviación estándar respectivamente. Un intervalo de confianza para μ al nivel de confianza $1-\alpha$ es

$$IC_\mu = \left[\bar{x} - t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}}, \bar{x} + t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}}\right]$$

O bien

$$IC_\mu = \bar{x} \pm t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}}$$

Fuente: Valdés, J. R. (2019). Estadística. Notas de clase.

Ejemplo 3. *Tomando en cuenta que* $n = 114$, $\bar{x} = 75.0588$ *y* $s^2 = 75.0588$, *el intervalo de confianza para* μ *al nivel* $1 - \alpha = 0.95$ *es*

$$\begin{aligned} IC_\mu &= \bar{x} \pm t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}} \\ &= 75.0588 \pm t_{113,0.975}\frac{12.6709}{\sqrt{114}} \\ &= 75.0588 \pm (1.9812)(1.1867) \\ &= 75.0588 \pm 2.3511 \\ &= (72.7076, 77.4099) \end{aligned}$$

Aquí el error estándar de $\bar{x}$ *es*

$$s_{\bar{x}} = \frac{12.6709}{\sqrt{114}} = 1.1867$$

y el margen de error estimado

$$\hat{e} = (t_{113,0.975})(s_{\bar{x}}) = 2.3511$$

Intervalo de confianza para diferencia de medias

Al igual que la diferencia de proporciones es factible obtener un intervalo de confianza para diferencia de medias estimada depende de si las poblaciones a comparar tienen o no igual varianza.

Si se asume que las varianzas de las poblaciones son iguales $\sigma_1^2 = \sigma_2^2$ entonces el error estándar de la diferencia es:

$$S_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \qquad (4.1)$$

Para el caso de varianza diferentes $\sigma_1^2 \neq \sigma_2^2$ se tiene

$$S_{\bar{x}_1-\bar{x}_2} = \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \qquad (4.2)$$

Donde

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

Si no se tiene información previa se consideran las varianzas distintas. En la construcción del intervalo de confianza para la diferencia de medias $(\mu_1 - \mu_2)$ se parte de que:

$$t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{S_{\bar{x}_1-\bar{x}_2}} \sim t_{n_1+n_2-2}$$

Donde $S_{\bar{x}1} - \bar{x}2$ podrá tomar alguna de las formas de las ecuaciones (4.1) o (4.2).

Al igual que la construcción del $IC_{\hat{p}1} - \hat{p}2$ a partir de que:

$$P\left[t_{n_1+n_2-2,\alpha/2} < t < t_{n_1+n_2-2,1-\alpha/2}\right] = 1-\alpha$$

se tiene el siguiente resultado:

Intervalo de confianza para la diferencia de medias $\mu_1 - \mu_2$

Sean $\{X_{1,i}\}_{i=1}^{n_1}$ y $\{X_{2,i}\}_{i=1}^{n_2}$ muestras aleatorias de dos poblaciones con estimadores $\bar{x}_1$, s_1^2, $\bar{x}_2$ y s_2^2 para medias y varianzas respectivamente. Un intervalo de confianza para la diferencia $\mu_1 - \mu_2$ al nivel de confianza $1 - \alpha$ es

$$IC_{\mu_1-\mu_2} = (\bar{x}_1 - \bar{x}_2) \pm t_{n_1+n_2-2,\alpha/2} S_{\bar{x}_1-\bar{x}_2}$$

donde $S_{\bar{x}_1-\bar{x}_2}$ toma alguna de las formas de las ecuaciones (4.1) o (4.1).

Fuente: Valdés, J. R. (2019). Estadística. Notas de clase.

Ejemplo 4. *En referencia al ejemplo* 2 *se estimó* $\bar{x}_1 = 79.75$ *para ingeniería y* $\bar{x}_2 = 75.4029$ *para sistemas. El intervalo de confianza para* $\mu_1 - \mu_2$ *tomando varianzas distintas, al nivel* $1 - \alpha = 0.95$ *se obtiene como sigue:*

$$\begin{aligned} s_{\bar{x}_1-\bar{x}_1} &= \sqrt{\frac{95.2692}{40} + \frac{214.9748}{34}} \\ &= 2.9503 \end{aligned}$$

que corresponde al error estándar de $\bar{x}_1 - \bar{x}_2$. *Además* $t_{72,0.975} = 1.9935$.
Entonces,

$$\begin{aligned} IC_{\mu_1-\mu_2} &= (79.75 - 75.4029) \pm (1.9935)(2.9503) \\ &= 4.3471 \pm 5.8814 \\ &= (-1.5343, 10.2285) \end{aligned}$$

El margen de error estimado es:

$$\hat{e} = 5.8814$$

Dado que el intervalo de confianza no contiene al cero, estadísticamente no se observa diferencia significativa del peso medio entre ambas facultades.

Referencias:

Hoel, P. G. (1984). Elementary Statistics. John Wiley & Sons.

Valdés, J. R. (2019). Estadística. Notas de clase.