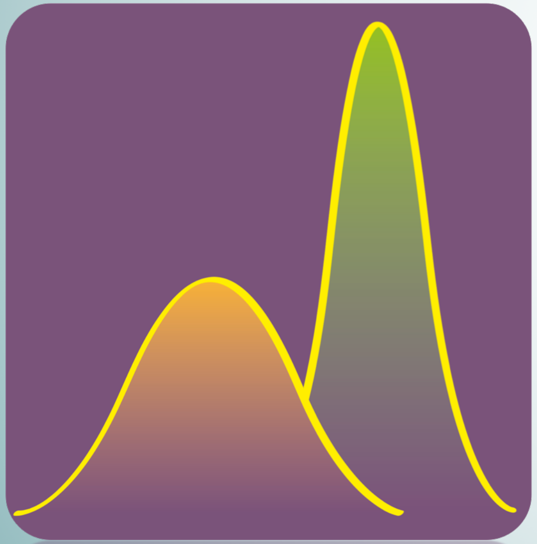
Estadística II

TEMA 5. ESTIMACIÓN POR INTERVALOS DE CONFIANZA



Marta Guijarro Garvi

David Gutiérrez Sobrao

DEPARTAMENTO DE ECONOMÍA

Tema 5. Estimación por intervalos de confianza

Índice de contenidos

- Concepto de intervalo de confianza.
- Construcción de intervalos de confianza: el método del pivote.
- Intervalos de confianza en el muestreo de una población normal: media y varianza.
- Intervalos de confianza en el muestreo de una población de Bernoulli: proporción.
- Intervalos de confianza en el muestreo de dos poblaciones normales independientes: diferencia de medias y cociente de varianzas.
- Intervalos de confianza en el muestreo de dos poblaciones de Bernoulli independientes: diferencia de proporciones.

Tema 5. Estimación por intervalos de confianza

Resultados de aprendizaje

- Conocer e interpretar los elementos relacionados con la construcción de un intervalo de confianza.
- Conocer la diferencia entre un intervalo de probabilidad y un intervalo de confianza.
- Saber obtener intervalos de confianza para la media y la varianza en el muestreo de una población normal.
- Saber hallar intervalos de confianza para la proporción en el muestreo de una población de Bernoulli.
- Saber calcular intervalos de confianza para comparar medias y varianzas en el muestreo de dos poblaciones normales independientes.
- Saber hallar intervalos de confianza para una proporción y una diferencia de proporciones en el muestreo de poblaciones de Bernoulli.

Concepto de intervalo de confianza

Sea X una población con distribución determinada por $f(x;\theta)$ y $(X_1,\dots,X_n)$ una muestra aleatoria simple de tamaño n de X. El intervalo

$$\left[\hat{\theta}_1(X_1,\dots,X_n),\hat{\theta}_2(X_1,\dots,X_n)\right]$$

es un **intervalo de probabilidad** $1-\alpha$ para el parámetro θ si

$$p\left[\hat{\theta}_1(X_1,\dots,X_n)\le\theta\le\hat{\theta}_2(X_1,\dots,X_n)\right]=1-\alpha.$$

Entonces, para una realización de la muestra, $(x_1,\dots,x_n)$, el intervalo

$$\left[\hat{\theta}_1(x_1,\dots,x_n),\hat{\theta}_2(x_1,\dots,x_n)\right]$$

es un **intervalo de confianza** para el parámetro θ al nivel de confianza $1-\alpha$.

Los niveles de confianza más habituales son 0,9, 0,95 y 0,99.

Construcción de intervalos de confianza: el método del pivote

Sea X una población con distribución determinada por $f(x;\theta)$ y sea $(X_1,\ldots,X_n)$ una muestra aleatoria simple de X. Un **pivote** es un estadístico, $T(X_1,\ldots,X_n;\theta)$, con las condiciones siguientes:

- Es una función de la muestra y del parámetro.
- Su expresión permite despejar el parámetro.
- Tiene una distribución de probabilidades conocida.

El método del pivote permite construir un intervalo de confianza para el parámetro al nivel de confianza $1-\alpha$. Para ello:

1. Se hallan los valores a y b tales que $p\left[a \le T(X_1,\ldots,X_n;\theta) \le b\right] = 1-\alpha$.
2. Se despeja el parámetro y se obtiene el intervalo de probabilidad $1-\alpha$.
3. Para una realización de la muestra, se calcula el correspondiente intervalo de confianza.

Intervalos de confianza en el muestreo de una población normal: media y varianza

En la siguiente tabla se recogen los intervalos de confianza para μ y σ^2 en el muestreo de una población normal, así como el pivote y el planteamiento del procedimiento en cada caso.

Parámetro	Pivote y distribución	Planteamiento del método del pivote	Obtención de los correspondientes valores	Intervalo de confianza
Media (varianza conocida)	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$	$p\left[-z_{\alpha/2} \le \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le z_{\alpha/2}\right]=1-\alpha$	$p\left[Z \le z_{\alpha/2}\right]=1-\dfrac{\alpha}{2}$	$\bar{x} \pm z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$
Media	$\dfrac{\bar{X}-\mu}{S/\sqrt{n-1}} \rightarrow t_{n-1}$	$p\left[-t_{\alpha/2} \le \dfrac{\bar{X}-\mu}{S/\sqrt{n-1}} \le t_{\alpha/2}\right]=1-\alpha$	$p\left[t_{n-1} \le t_{\alpha/2}\right]=1-\dfrac{\alpha}{2}$	$\bar{x} \pm t_{\alpha/2} \cdot \dfrac{s}{\sqrt{n-1}}$
Varianza	$\dfrac{n \cdot S^2}{\sigma^2} \rightarrow \chi^2_{n-1}$	$p\left[a \le \dfrac{n \cdot S^2}{\sigma^2} \le b\right]=1-\alpha$	$p\left[\chi^2_{n-1} \le a\right]=p\left[\chi^2_{n-1} \ge b\right]=\dfrac{\alpha}{2}$	$\left[\dfrac{n \cdot s^2}{b}, \dfrac{n \cdot s^2}{a}\right]$

Intervalos de confianza en el muestreo de una población normal: media y varianza

Ejemplo 5.1. Sea X una población con distribución $N(\mu, 0,5)$. Si $n = 40$, $\bar{x} = 38$ y 1-α=0,95, entonces,

$$p\left[Z \le 1,96\right] = 0,975$$

y el intervalo

$$38 \pm 1,96 \cdot \frac{0,5}{\sqrt{40}} \equiv 38 \pm 0,155$$

es un intervalo de confianza para μ al 95 %.

Ejemplo 5.2. Sea X una población con distribución $N(\mu, \sigma)$. Si $n = 51$, $s = 3$ y 1-α=0,95, entonces,

$$p\left[\chi_{51-1} \le 32,4\right] = p\left[\chi_{51-1} \ge 71,4\right] = 0,025$$

y el intervalo

$$\left[\frac{51 \cdot 3^2}{71,4}, \frac{51 \cdot 3^2}{32,4}\right] \equiv \left[6,429, 14,167\right]$$

es un intervalo de confianza para σ^2 al 95 %.

Intervalos de confianza en el muestreo de una población de Bernoulli: proporción

En la siguiente tabla se recoge el intervalo de confianza para p en el muestreo de una población de Bernoulli, así como el pivote y el planteamiento del procedimiento.

Parámetro	Pivote y distribución	Planteamiento del método del pivote	Obtención de los correspondientes valores	Intervalo de confianza
Proporción	$\dfrac{\hat{p}-p}{\sqrt{\dfrac{p\cdot(1-p)}{n}}} \xrightarrow[\text{aprox.}]{} N(0,1)$	$p\left[-z_{\alpha/2} \le \dfrac{\hat{p}-p}{\sqrt{\dfrac{\hat{p}\cdot(1-\hat{p})}{n}}} \le z_{\alpha/2}\right]=1-\alpha$	$p\left[Z \le z_{\alpha/2}\right]=1-\dfrac{\alpha}{2}$	$\hat{p} \pm z_{\alpha/2}\cdot\sqrt{\dfrac{\hat{p}\cdot(1-\hat{p})}{n}}$

Ejemplo 5.3. Sea X una población con distribución $b(p)$. Si n = 100, $\hat{p}=0,4$ y 1-α=0,95, entonces,

$$p\left[Z\leq 1,96\right]=0,975$$

y el intervalo

$$0,4\pm 1,96\cdot\sqrt{\frac{0,4\cdot(1-0,4)}{100}}\equiv 0,4\pm 0,096$$

es un intervalo de confianza para p al 95 %.

Intervalos de confianza para el muestreo de dos poblaciones normales independientes: diferencia de medias

En la siguiente tabla se recogen los intervalos de confianza para $\mu_1 - \mu_2$ en el muestreo de dos poblaciones normales independientes, así como el pivote y el planteamiento del procedimiento en cada caso.

Pivote y distribución	**Planteamiento del método del pivote**	**Obtención de los correspondientes valores**	**Intervalo de confianza**
	Desviaciones típicas conocidas		
$\dfrac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} \to N(0,1)$	$p\left[-z_{\alpha/2} \le \dfrac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} \le z_{\alpha/2}\right] = 1 - \alpha$	$p\left[Z \le z_{\alpha/2}\right] = 1 - \dfrac{\alpha}{2}$	$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2} \cdot \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$
	Desviaciones típicas desconocidas e iguales		
$\dfrac{\sqrt{\dfrac{n_1 \cdot n_2}{n_1 + n_2}}\left(\bar{X}_1 - \bar{X}_2\right) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{n_1 \cdot S_1^2 + n_2 \cdot S_2^2}{n_1 + n_2 - 2}}} \to t_{n_1 + n_2 - 2}$	$p\left[-t_{\alpha/2} \le \dfrac{\sqrt{\dfrac{n_1 \cdot n_2}{n_1 + n_2}}\left[\left(\bar{X}_1 - \bar{X}_2\right) - (\mu_1 - \mu_2)\right]}{\sqrt{\dfrac{n_1 \cdot S_1^2 + n_2 \cdot S_2^2}{n_1 + n_2 - 2}}} \le t_{\alpha/2}\right] = 1 - \alpha$	$p\left[t_{n_1+n_2-2} \le t_{\alpha/2}\right] = 1 - \dfrac{\alpha}{2}$	$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} \cdot \dfrac{\sqrt{\dfrac{n_1 \cdot s_1^2 + n_2 \cdot s_2^2}{n_1 + n_2 - 2}}}{\sqrt{\dfrac{n_1 \cdot n_2}{n_1 + n_2}}}$

Intervalos de confianza para el muestreo de dos poblaciones normales independientes: diferencia de medias

Ejemplo 5.4. Sean X_1 y X_2, poblaciones independientes con distribuciones $N(\mu_1, \sigma_1)$ y $N(\mu_2, \sigma_2)$. Si $n_1 = 40$ y $n_2 = 90$, $\bar{x}_1 = 25$, $\bar{x}_2 = 42$, $s_1^2=2$, $s_2^2=3$ y $1-\alpha=0{,}95$, entonces,

$$p\left[t_{40+90-2} \leq 1{,}96\right] = 0{,}975$$

y el intervalo

$$25-42 \pm 1{,}96 \cdot \frac{\sqrt{\dfrac{40\cdot 2+90\cdot 3}{40+90-2}}}{\sqrt{\dfrac{40\cdot 90}{40+90}}} \equiv -17 \pm 0{,}616$$

es un intervalo de confianza para $\mu_1 - \mu_2$ al 95 %.

Intervalos de confianza para el muestreo de dos poblaciones normales independientes: cociente de varianzas

En la siguiente tabla se recoge el intervalo de confianza para σ_1^2/σ_2^2 en el muestreo de dos poblaciones normales independientes, así como el pivote y el planteamiento del procedimiento.

Pivote y distribución	**Planteamiento del método del pivote**	**Obtención del valor correspondiente**	**Intervalo de confianza**
$\dfrac{\dfrac{S_{c_1}^2}{\sigma_1^2}}{\dfrac{S_{c_2}^2}{\sigma_2^2}} \to F_{n_1-1,n_2-1}$	$p\left[a \leq \dfrac{\dfrac{S_{c_1}^2}{\sigma_1^2}}{\dfrac{S_{c_2}^2}{\sigma_2^2}} \leq b\right] = 1-\alpha$	$p\left[F_{n_1-1,n_2-1} \leq a\right] = p\left[F_{n_1-1,n_2-1} \geq b\right] = \dfrac{\alpha}{2}$	$\left[\dfrac{s_{c_1}^2}{b \cdot s_{c_2}^2}, \dfrac{s_{c_1}^2}{a \cdot s_{c_2}^2}\right]$

Intervalos de confianza para el muestreo de dos poblaciones normales independientes: cociente de varianzas

Ejemplo 5.5. Sean X_1 y X_2, poblaciones independientes con distribuciones $N(\mu_1,\sigma_1)$ y $N(\mu_2,\sigma_2)$. Si $n_1 = 25$ y $n_2 = 31$, $s_{c_1}^2$ =5, $s_{c_2}^2$ =8 y 1-α=0,9 , entonces,

$$p\left[F_{25-1,31-1} \leq 0,51\right] = p\left[F_{25-1,31-1} \geq 1,89\right] = 0,05$$

y el intervalo

$$\left[\frac{5}{1,89\cdot 8},\frac{5}{0,51\cdot 8}\right] \equiv \left[0,331, 1,225\right]$$

es un intervalo de confianza para σ_1^2/σ_2^2 al 95 %.

Intervalos de confianza para el muestreo de dos poblaciones de Bernoulli independientes: diferencia de proporciones

En la siguiente tabla se recoge el intervalo de confianza para $p_1 - p_2$ en el muestreo de dos poblaciones de Bernoulli independientes, así como el pivote y el planteamiento del procedimiento.

Pivote y distribución	**Planteamiento del método del pivote**	**Obtención del valor correspondiente**	**Intervalo de confianza**
$\dfrac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\dfrac{p_1 \cdot (1 - p_1)}{n_1} + \dfrac{p_2 \cdot (1 - p_2)}{n_2}}} \xrightarrow[\text{aprox.}]{} N(0,1)$	$p\left[-z_{\alpha/2} \leq \dfrac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\dfrac{\hat{p}_1 \cdot (1 - \hat{p}_1)}{n_1} + \dfrac{\hat{p}_2 \cdot (1 - \hat{p}_2)}{n_2}}} \leq z_{\alpha/2}\right] = 1 - \alpha$	$p\left[Z \leq z_{\alpha/2}\right] = 1 - \dfrac{\alpha}{2}$	$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \cdot \sqrt{\dfrac{\hat{p}_1 \cdot (1 - \hat{p}_1)}{n_1} + \dfrac{\hat{p}_2 \cdot (1 - \hat{p}_2)}{n_2}}$

Intervalos de confianza para el muestreo de dos poblaciones de Bernoulli independientes: diferencia de proporciones

Ejemplo 5.6. Sean X_1 y X_2, poblaciones independientes con distribuciones $b(p_1)$ y $b(p_2)$. Si $n_1 = 100$, $n_2 = 200$, $\hat{p}_1 = 0,4$, $\hat{p}_2 = 0,3$ y $1\text{-}\alpha=0,95$ entonces,

$$p\left[Z \leq 1,96\right] = 0,975$$

y el intervalo

$$0,4 - 0,3 \pm 1,96 \cdot \sqrt{\frac{0,4 \cdot (1-0,4)}{100} + \frac{0,3 \cdot (1-0,3)}{200}} \equiv 0,1 \pm 0,059$$

es un intervalo de confianza para $p_1 - p_2$ al 95 %.