

# Estimadores y su distribución

## Contents

1	Población y muestra	1
2	Estimadores puntuales de los parámetros del modelo	1
3	Distribución en el muestreo de los estimadores puntuales	2
3.1	Distribución de los datos . . . . .	2
3.2	Distribución del estimador de $\beta$ . . . . .	3
3.3	Distribución del estimador de $\sigma^2$ . . . . .	4

## 1 Población y muestra

Queremos estudiar la relación entre la variable  $y$  y los regresores  $x_1, x_2, \dots, x_k$ . Y queremos obtener **conclusiones generales**. Para ello suponemos que dicha relación viene dada por el siguiente modelo estadístico:

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki} + u_i, \quad u_i \sim N(0, \sigma^2) \quad i = 1, 2, \dots, N$$

- El modelo nos permite obtener conclusiones generales. Dichas conclusiones son válidas en el conjunto de tamaño  $N$ , conocida como **población**.
- Los parámetros del modelo  $\beta_0, \beta_1, \dots, \beta_k, \sigma^2$  son desconocidos y no se pueden calcular porque no conocemos toda la población.
- Para trabajar con dicho modelo tenemos un conjunto de datos de  $y_i, x_{1i}, \dots, x_{ki}$ ,  $i = 1, 2, \dots, n$ , donde  $n < N$ . A estos datos se le conoce como **muestra**.

## 2 Estimadores puntuales de los parámetros del modelo

El modelo de la población tiene  $(k+2)$  parámetros:  $\beta_0, \beta_1, \dots, \beta_k, \sigma^2$ . Lo primero que haremos es estimar un valor para dichos parámetros.

Se denomina **estimador de un parámetro** a cualquier expresión que asigna un valor a dicho parámetro. Y se denomina **estimación** al valor asignado.

Por ejemplo, podríamos utilizar los siguientes estimadores (se indica poniendo el símbolo  $\hat{\square}$ ):

$$\hat{\beta}_k = \bar{x}_k, \quad \hat{\sigma} = \text{Var}(y)$$

En cualquier caso, las estimaciones se tienen que calcular con los datos disponibles, la muestra. Sin embargo, unos estimadores tienen mejores propiedades que otros. Nosotros vamos a utilizar como estimadores los valores calculados con mínimos cuadrados:

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \dots \\ \hat{\beta}_k \end{bmatrix} = (X^T X)^{-1} (X^T y)$$

El estimador de  $\sigma$  se define como (no es un resultado de mínimos cuadrados):

$$\hat{\sigma} = \sqrt{\frac{\sum e_i^2}{n - k - 1}}$$

donde  $k$  es el número de regresores.

### 3 Distribución en el muestreo de los estimadores puntuales

El hecho de considerar un modelo con variables aleatorias hace que los estimadores sean variables aleatorias:

$$u_i \sim Normal \Rightarrow y_i \sim Normal$$

Como

$$\hat{\beta} = (X^T X)^{-1} X^T y \Rightarrow \hat{\beta} \sim Normal$$

Según el modelo,  $y_i$  son variables aleatorias. Y como los estimadores se calculan a partir de  $y_i$ , también son variables aleatorias. Por tanto tienen distribución de probabilidad. En los siguientes apartados se calcula dicha distribución, que se conoce como distribución en el muestreo.

#### 3.1 Distribución de los datos

Como hemos visto en el apartado anterior, el modelo general es:

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki} + u_i, \quad u_i \sim N(0, \sigma^2), \quad i = 1, \dots, n$$

En forma matricial

$$Y = X\beta + U, \quad U \sim N(0, \sigma^2 I)$$

Como  $U \sim Normal$ , debido a las propiedades de la distribución normal,  $Y \sim Normal$ . Concretamente:

$$Y \sim N(X\beta, \sigma^2 I)$$

ya que:

$$E[Y] = E[X\beta + U] = X\beta$$

$$Var[Y] = Var[X\beta + U] = \sigma^2 I$$

## 3.2 Distribución del estimador de $\beta$

### 3.2.1 Con matrices de datos

La ecuación de los estimadores es

$$\hat{\beta} = (X^T X)^{-1} (X^T y)$$

Si consideramos variables aleatorias en lugar de datos:

$$\hat{\beta} = (X^T X)^{-1} (X^T Y)$$

obtenemos la distribución de los estimadores:

$$\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$$

ya que

$$Y \sim Normal \Rightarrow \hat{\beta} \sim Normal$$

$$E[\hat{\beta}] = E[(X^T X)^{-1} X^T Y] = (X^T X)^{-1} X^T E[Y] = (X^T X)^{-1} X^T X \beta = \beta$$

$$Var[\hat{\beta}] = Var[(X^T X)^{-1} X^T Y] = (X^T X)^{-1} X^T Var[Y] ((X^T X)^{-1} X^T)^T = \sigma^2 (X^T X)^{-1}$$

### 3.2.2 Con matrices de covarianzas

Si en lugar de utilizar las matrices  $(X^T X)^{-1}$  y  $(X^T y)$  utilizamos matrices de varianzas y covarianzas, los estimadores son:

$$\hat{\beta}_a = S_{XX}^{-1} S_{Xy} = \left( \frac{1}{n-1} X_a^T X_a \right)^{-1} \left( \frac{1}{n-1} X_a^T Y_a \right) = (X_a^T X_a)^{-1} (X_a^T Y_a)$$

donde  $\hat{\beta}_a = [\hat{\beta}_1 \ \hat{\beta}_2 \ \cdots \ \hat{\beta}_k]^T$ . Por tanto hay que obtener la distribución de  $Y_a$ . El modelo generador de datos es:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + U_i, \quad i = 1, 2, \dots, n$$

Teniendo en cuenta que:

$$\bar{Y} = \beta_0 + \beta_1 \bar{x}_1 + \beta_2 \bar{x}_2 + \cdots + \beta_k \bar{x}_k$$

y restando ambas ecuaciones se obtiene:

$$Y_i - \bar{Y} = \beta_1 (x_{1i} - \bar{x}_1) + \beta_2 (x_{2i} - \bar{x}_2) + \cdots + \beta_k (x_{ki} - \bar{x}_k) + U_i, \quad i = 1, 2, \dots, n$$

Estas  $n$  ecuaciones se pueden expresar en forma matricial de la misma forma que hicimos antes, obteniendo:

$$Y_a = X_a \beta_a + U, \quad U \sim N(0, \sigma I)$$

Por tanto,

$$Y_a \sim N(X_a \beta_a, \sigma^2 I)$$

Ahora se puede demostrar que

$$\hat{\beta}_a \sim N\left(\beta_a, \frac{\sigma^2}{n-1} S_{XX}^{-1}\right)$$

ya que

$$Y_a \sim Normal \Rightarrow \hat{\beta}_a \sim Normal$$

$$E[\hat{\beta}_a] = E[(X_a^T X_a)^{-1} X_a^T Y_a] = (X_a^T X_a)^{-1} X_a^T E[Y_a] = (X_a^T X_a)^{-1} X_a^T X_a \beta_a = \beta_a$$

$$Var[\hat{\beta}_a] = Var[(X_a^T X_a)^{-1} X_a^T Y_a] = (X_a^T X_a)^{-1} X_a^T Var[Y_a] (X_a^T X_a)^{-1} X_a^T X_a = \sigma^2 (X_a^T X_a)^{-1}$$

Finalmente:

$$S_{XX} = \frac{1}{n-1} X_a^T X_a \Rightarrow (X_a^T X_a)^{-1} = \frac{1}{n-1} S_{XX}^{-1}$$

Faltaría el estimador de  $\beta_0$  que se obtiene con la ecuación

$$\hat{\beta}_0 = \bar{Y} - \bar{x}^T \hat{\beta}_a$$

dónde  $\bar{x} = [\bar{x}_1 \ \bar{x}_2 \ \dots \ \bar{x}_n]^T$ . Se puede demostrar que

$$\hat{\beta}_0 \sim N\left(\beta_0, \sigma^2 \left(\frac{1}{n} + \frac{1}{n-1} \bar{x}^T S_{XX}^{-1} \bar{x}\right)\right)$$

ya que:

$$\bar{Y} \sim N\left(\beta_0 + \bar{x}^T \beta_a, \frac{\sigma^2}{n}\right)$$

### 3.3 Distribución del estimador de $\sigma^2$

El modelo tiene un parámetro más, la varianza de los errores,  $\sigma^2$ . Este parámetro también hay que estimarlo. Se puede demostrar que

$$\frac{\sum e_i^2}{\sigma^2} \rightarrow \chi_{n-k-1}^2$$

donde n es el número de observaciones y k es el número de regresores. Por ello se propone el siguiente estimador

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-k-1}$$

ya que es un estimador insesgado de  $\sigma^2$ . Efectivamente

$$E[\hat{\sigma}^2] = E \left[ \frac{\sum e_i^2}{n - k - 1} \right] = \sigma^2$$

ya que  $E[\chi_n^2] = n$ . Al término  $\sum e_i^2 / (n - k - 1)$  también se lo conoce como **varianza residual** y se representa por  $\hat{s}_R^2$ .

$$\hat{s}_R^2 = \frac{\sum e_i^2}{n - k - 1}$$

A la raíz cuadrada se le conoce como **residual standard error**. El término  $(n - k - 1)$  son los *grados de libertad*. La distribución en el muestreo de la varianza residual es

$$\frac{\sum e_i^2}{\sigma^2} \rightarrow \chi_{n-k-1}^2 \Rightarrow \frac{(n - k - 1)\hat{s}_R^2}{\sigma^2} \rightarrow \chi_{n-k-1}^2$$