ECONOMETRÍA

EJERCICIO 2 - RESUELTO

EN ESTE EJERCICIO SE ILUSTRAN LAS FÓRMULAS Y LOS CÁLCULOS ASOCIADOS CON LA ESTIMACIÓN POR MÍNIMOS CUADRADOS ORDINARIOS DE UN MODELO DE REGRESIÓN LINEAL SIMPLE. ES MUY IMPORTANTE ESTUDIARLO CON ATENCIÓN Y COMPROBAR Y REPRODUCIR A MANO TODOS LOS CÁLCULOS NUMÉRICOS.

ENUNCIADO - PARTE I

A partir de la fórmula general $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, obtener fórmulas específicas para las estimaciones MCO de los parámetros en los dos modelos siguientes:

$$Y = \beta_1 + U. \qquad \text{[M1]}$$

$$Y = \beta_1 + \beta_2 X_2 + U. \qquad \text{[M2]}$$

SOLUCIÓN

En el modelo [M1] ($K = 1$):

$$\boldsymbol{\beta} = \beta_1;\ \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix};\ \mathbf{X} = \mathbf{i} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix};\ \mathbf{X}'\mathbf{X} = N;\ (\mathbf{X}'\mathbf{X})^{-1} = \tfrac{1}{N};\ \mathbf{X}'\mathbf{y} = \textstyle\sum y_i. \qquad \text{[M1.1]}$$

$$\hat{\boldsymbol{\beta}} = \hat{\beta}_1 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \tfrac{1}{N}\textstyle\sum y_i = \bar{y} \ \text{(media muestral de } \mathbf{y}\text{)}. \qquad \text{[M1.2]}$$

En el modelo [M2] (RLS, $K = 2$):

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix};\ \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix};\ \mathbf{X} = [\mathbf{i}, \mathbf{x}_2] = \begin{bmatrix} 1 & x_{12} \\ 1 & x_{22} \\ \vdots & \vdots \\ 1 & x_{N2} \end{bmatrix};$$

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} N & \sum x_{i2} \\ \sum x_{i2} & \sum x_{i2}^2 \end{bmatrix} = \begin{bmatrix} N & N\bar{x}_2 \\ N\bar{x}_2 & \sum x_{i2}^2 \end{bmatrix};\ \mathbf{X}'\mathbf{y} = \begin{bmatrix} \sum y_i \\ \sum x_{i2} y_i \end{bmatrix} = \begin{bmatrix} N\bar{y} \\ \sum x_{i2} y_i \end{bmatrix}. \qquad \text{[M2.1]}$$

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{N(\sum x_{i2}^2 - N\bar{x}_2^2)} \begin{bmatrix} \sum x_{i2}^2 & -N\bar{x}_2 \\ -N\bar{x}_2 & N \end{bmatrix} = \frac{1}{\sum(x_{i2} - \bar{x}_2)^2} \begin{bmatrix} \frac{1}{N}\sum x_{i2}^2 & -\bar{x}_2 \\ -\bar{x}_2 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{N} + \frac{\bar{x}_2^2}{\sum(x_{i2} - \bar{x}_2)^2} & -\frac{\bar{x}_2}{\sum(x_{i2} - \bar{x}_2)^2} \\ -\frac{\bar{x}_2}{\sum(x_{i2} - \bar{x}_2)^2} & \frac{1}{\sum(x_{i2} - \bar{x}_2)^2} \end{bmatrix}. \qquad \text{[M2.2]}$$

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{bmatrix} \frac{1}{N} + \frac{\bar{x}_2^2}{\sum(x_{i2}-\bar{x}_2)^2} & -\frac{\bar{x}_2}{\sum(x_{i2}-\bar{x}_2)^2} \\ -\frac{\bar{x}_2}{\sum(x_{i2}-\bar{x}_2)^2} & \frac{1}{\sum(x_{i2}-\bar{x}_2)^2} \end{bmatrix} \begin{bmatrix} N\bar{y} \\ \sum x_{i2}y_i \end{bmatrix} =$$

$$= \begin{bmatrix} \bar{y} + \frac{N\bar{x}_2^2\,\bar{y} - \bar{x}_2\sum x_{i2}y_i}{\sum(x_{i2}-\bar{x}_2)^2} \\ \frac{-N\bar{x}_2\,\bar{y} + \sum x_{i2}y_i}{\sum(x_{i2}-\bar{x}_2)^2} \end{bmatrix} = \begin{bmatrix} \bar{y} - \bar{x}_2\frac{\sum x_{i2}y_i - N\bar{x}_2\,\bar{y}}{\sum(x_{i2}-\bar{x}_2)^2} \\ \frac{\sum x_{i2}y_i - N\bar{x}_2\,\bar{y}}{\sum(x_{i2}-\bar{x}_2)^2} \end{bmatrix} = \begin{bmatrix} \bar{y} - \bar{x}_2\hat{\beta}_2 \\ \frac{\sum(x_{i2}-\bar{x}_2)(y_i-\bar{y})}{\sum(x_{i2}-\bar{x}_2)^2} \end{bmatrix}. \qquad [M2.3]$$

ENUNCIADO - PARTE II

Estimar el modelo de regresión lineal simple [M2] con los datos de la tabla siguiente:

$\mathbf{y}$	$\mathbf{x}_2$
84	32
108	40
92	36
110	44
106	48

En las Secciones 1 - 10 de la guía *Introducción al Uso de EViews 4.1* se dan las indicaciones necesarias para obtener con EViews los resultados de algunas de las operaciones que se muestran a continuación.

SOLUCIÓN

En el modelo [M2] (RLS, $K = 2$) con los datos de la tabla anterior ($N = 5$):

$$\underset{2\times1}{\boldsymbol{\beta}} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}; \quad \underset{5\times1}{\mathbf{y}} = \begin{bmatrix} 84 \\ 108 \\ 92 \\ 110 \\ 106 \end{bmatrix}; \quad \underset{5\times2}{\mathbf{X}} = [\,\mathbf{i}, \mathbf{x}_2\,] = \begin{bmatrix} 1 & 32 \\ 1 & 40 \\ 1 & 36 \\ 1 & 44 \\ 1 & 48 \end{bmatrix}. \qquad [1]$$

$$\underset{2\times2}{\mathbf{X}'\mathbf{X}} = \begin{bmatrix} N & \sum_{i=1}^N x_{i2} \\ \sum_{i=1}^N x_{i2} & \sum_{i=1}^N x_{i2}^2 \end{bmatrix} = \begin{bmatrix} 5 & 200 \\ 200 & 8160 \end{bmatrix} \Rightarrow \underset{2\times2}{(\mathbf{X}'\mathbf{X})^{-1}} = \begin{bmatrix} 10.2 & -0.25 \\ -0.25 & 0.00625 \end{bmatrix}. \qquad [2]$$

$$\underset{2\times1}{\mathbf{X}'\mathbf{y}} = \begin{bmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_{i2}y_i \end{bmatrix} = \begin{bmatrix} 500 \\ 20248 \end{bmatrix}. \qquad [3]$$

Estimación MCO de $\boldsymbol{\beta}$:

$$[2] \text{ y } [3] \Rightarrow \underset{2\times1}{\hat{\boldsymbol{\beta}}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \underset{2\times2}{(\mathbf{X}'\mathbf{X})^{-1}} \underset{2\times1}{\mathbf{X}'\mathbf{y}} = \begin{bmatrix} 10.2 & -0.25 \\ -0.25 & 0.00625 \end{bmatrix} \begin{bmatrix} 500 \\ 20248 \end{bmatrix} = \begin{bmatrix} 38 \\ 1.55 \end{bmatrix}. \qquad [4]$$

Observación: Las estimaciones $\hat{\beta}_1$ y $\hat{\beta}_2$ también pueden calcularse en este caso utilizando las fórmulas específicas en [M2.3] para un modelo RLS:

$$\begin{aligned} \hat{\beta}_2 &= \frac{\sum_{i=1}^{N}(x_{i2}-\bar{x}_2)(y_i-\bar{y})}{\sum_{i=1}^{N}(x_{i2}-\bar{x}_2)^2} = \frac{248}{160} = 1.55 \\ &= \frac{\widehat{\text{cov}}[\mathbf{x}_2,\mathbf{y}]}{\widehat{\text{var}}[\mathbf{x}_2]} = \frac{49.6}{32} = 1.55 \\ &= \frac{\widehat{\text{dvt}}[\mathbf{y}]}{\widehat{\text{dvt}}[\mathbf{x}_2]} \times \widehat{\text{corr}}[\mathbf{x}_2,\mathbf{y}] = \frac{\sqrt{104}}{\sqrt{32}} \times \frac{49.6}{\sqrt{32}\times\sqrt{104}} = 1.55, \\ \hat{\beta}_1 &= \bar{y} - \hat{\beta}_2\bar{x}_2 = 100 - 1.55\times 40 = 38, \end{aligned}$$

donde $\bar{y} = \frac{1}{5}\sum_{i=1}^{5} y_i = \frac{1}{5}\times 500 = 100$, $\bar{x}_2 = \frac{1}{5}\sum_{i=1}^{5} x_{i2} = \frac{1}{5}\times 200 = 40$. $\boxed{\text{A}}$

Valores Ajustados y Residuos:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \begin{bmatrix} 1 & 32 \\ 1 & 40 \\ 1 & 36 \\ 1 & 44 \\ 1 & 48 \end{bmatrix} \begin{bmatrix} 38 \\ 1.55 \end{bmatrix} = \begin{bmatrix} 87.6 \\ 100 \\ 93.8 \\ 106.2 \\ 112.4 \end{bmatrix}, \quad \hat{\mathbf{u}} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 84 \\ 108 \\ 92 \\ 110 \\ 106 \end{bmatrix} - \begin{bmatrix} 87.6 \\ 100 \\ 93.8 \\ 106.2 \\ 112.4 \end{bmatrix} = \begin{bmatrix} -3.6 \\ 8 \\ -1.8 \\ 3.8 \\ -6.4 \end{bmatrix}. \qquad [5]$$

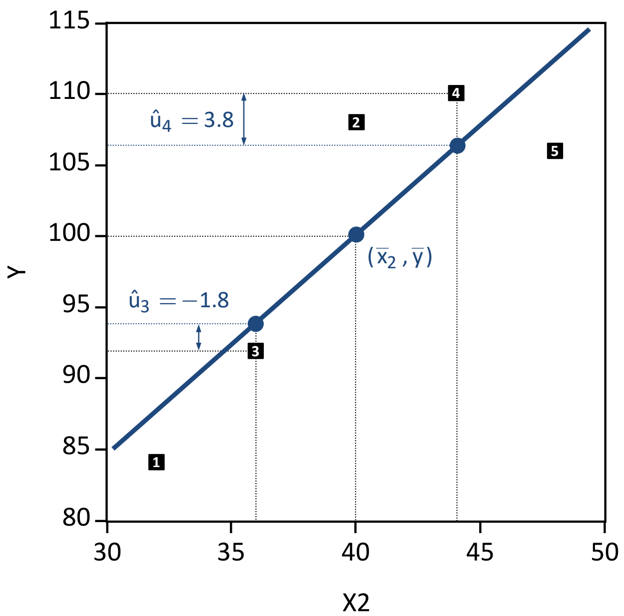
Cada valor ajustado $\hat{y}_i$ puede calcularse como $\hat{y}_i = 38 + 1.55x_{i2}$. Cada residuo $\hat{u}_i$ puede calcularse como $\hat{u}_i = y_i - \hat{y}_i = y_i - (38 + 1.55x_{i2})$ $(i = 1, ..., 5)$.

La ortogonalidad entre variables explicativas y residuos MCO $(\mathbf{X}'\hat{\mathbf{u}} = \mathbf{0})$ implica en este caso que $\sum_{i=1}^{5} \hat{u}_i = 0$ y $\mathbf{x}_2'\hat{\mathbf{u}} = \sum_{i=1}^{5} x_{i2}\hat{u}_i = 0$; la ortogonalidad entre valores ajustados y residuos MCO implica en este caso que $\hat{\mathbf{y}}'\hat{\mathbf{u}} = \sum_{i=1}^{5} \hat{y}_i\hat{u}_i = 0$.

Como un modelo RLS tiene **término constante**, $\sum_{i=1}^{5} y_i = \sum_{i=1}^{5} \hat{y}_i = 500$, por lo que $\bar{y} = \bar{\hat{y}} = 100$. Además, $\bar{y} = \hat{\beta}_1 + \hat{\beta}_2\bar{x}_2 \Rightarrow 100 = 38 + 1.55\times 40$ [la recta de regresión estimada pasa por el punto $(\bar{x}_2, \bar{y}) = (40, 100)$].

La relación $\mathbf{y}'\mathbf{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \hat{\mathbf{u}}'\hat{\mathbf{u}}$ puede comprobarse en este caso teniendo en cuenta que $\mathbf{y}'\mathbf{y} = \sum_{i=1}^{5} y_i^2 = 50520$, $\hat{\mathbf{y}}'\hat{\mathbf{y}} = \sum_{i=1}^{5} \hat{y}_i^2 = 50384.4$, y $\hat{\mathbf{u}}'\hat{\mathbf{u}} = \sum_{i=1}^{5} \hat{u}_i^2 = 135.6$.

Datos - Modelo Estimado - Algunos Residuos - Medias Muestrales



B

Coeficientes de Determinación y Sumas de Cuadrados:

$$R^2_{\text{NC}} = \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}}}{\mathbf{y}'\mathbf{y}} = \frac{50384.4}{50520} = 1 - \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{\mathbf{y}'\mathbf{y}} = 1 - \frac{135.6}{50520} = 0.9973. \qquad [6]$$

$$R^2 = \widehat{\text{corr}}[\hat{\mathbf{y}}, \mathbf{y}]^2 = \frac{\widehat{\text{cov}}[\hat{\mathbf{y}}, \mathbf{y}]^2}{\widehat{\text{var}}[\hat{\mathbf{y}}] \times \widehat{\text{var}}[\mathbf{y}]} = \frac{76.88^2}{76.88 \times 104} = 0.7392. \qquad [7]$$

$$\text{SCT} = \textstyle\sum_{i=1}^{N} (y_i - \bar{y})^2 = \mathbf{y}'\mathbf{y} - N\bar{y}^2 = 50520 - 5 \times 100^2 = 520\,. \qquad [8]$$

$$\text{SCE} = \textstyle\sum_{i=1}^{N} (\hat{y}_i - \bar{\hat{y}})^2 = \hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{\hat{y}}^2 = 50384.4 - 5 \times 100^2 = 384.4. \qquad [9]$$

$$\begin{aligned} \text{SCR} &= \textstyle\sum_{i=1}^{N} \hat{u}_i^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}} = 135.6 \\ &= \mathbf{y}'\mathbf{y} - \hat{\mathbf{y}}'\hat{\mathbf{y}} = 50520 - 50384.4 = 135.6 \qquad [10] \\ &= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} = 50520 - [38, 1.55]\begin{bmatrix} 500 \\ 20248 \end{bmatrix} = 135.6. \end{aligned}$$

Adicionalmente, la presencia del **término constante** en un modelo RLS implica que

SCT = SCE + SCR [520 = 384.4 + 135.6],

$$R^2 = \frac{\widehat{\text{var}}[\hat{\mathbf{y}}]}{\widehat{\text{var}}[\mathbf{y}]} = \frac{76.88}{104} = \frac{\text{SCE}}{\text{SCT}} = \frac{384.4}{520} = 1 - \frac{\text{SCR}}{\text{SCT}} = 1 - \frac{135.6}{520} = 0.7392\,,$$

$$\bar{R}^2 = 1 - \frac{\frac{\text{SCR}}{N-K}}{\frac{\text{SCT}}{N-1}} = 1 - \frac{\frac{135.6}{5-2}}{\frac{520}{5-1}} = 1 - \frac{(N-1)(1-R^2)}{N-K} = 1 - \frac{(5-1)(1-0.7392)}{5-2} = 0.6523.$$

C

Observación: Los residuos, la SCE y el R^2 también pueden calcularse en este caso utilizando las siguientes fórmulas específicas para un modelo RLS:

$$\hat{\mathbf{u}} = \tilde{\mathbf{y}} - \tilde{\mathbf{x}}_2\hat{\beta}_2 = \begin{bmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ \vdots \\ y_N - \bar{y} \end{bmatrix} - \begin{bmatrix} x_{12} - \bar{x}_2 \\ x_{22} - \bar{x}_2 \\ \vdots \\ x_{N2} - \bar{x}_2 \end{bmatrix} \times \hat{\beta}_2 = \begin{bmatrix} -16 \\ 8 \\ -8 \\ 10 \\ 6 \end{bmatrix} - \begin{bmatrix} -8 \\ 0 \\ -4 \\ 4 \\ 8 \end{bmatrix} \times 1.55 = \begin{bmatrix} -3.6 \\ 8 \\ -1.8 \\ 3.8 \\ -6.4 \end{bmatrix},$$

$$\text{SCE} = \hat{\beta}_2\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}} = \hat{\beta}_2 \times \textstyle\sum_{i=1}^{N}(x_{i2} - \bar{x}_2)(y_i - \bar{y}) = 1.55 \times 248 = 384.4\,,$$

$$R^2 = \widehat{\text{corr}}[\mathbf{x}_2\,,\mathbf{y}]^2 = \frac{\widehat{\text{cov}}[\mathbf{x}_2\,,\mathbf{y}]^2}{\widehat{\text{var}}[\mathbf{x}_2]\times\widehat{\text{var}}[\mathbf{y}]} = \frac{49.6^2}{32\times 104} = 0.7392\,.$$

D

Estimación de σ^2 y $\text{Var}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}]$:

$$\hat{\sigma}^2 = \frac{\text{SCR}}{N-K} = \frac{135.6}{5-2} = 45.2. \qquad [11]$$

$$\widehat{\text{Var}}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}] = \hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 10.2\hat{\sigma}^2 & -0.25\hat{\sigma}^2 \\ -0.25\hat{\sigma}^2 & 0.00625\hat{\sigma}^2 \end{bmatrix} =$$

$$= \begin{bmatrix} 461.04 & -11.3 \\ -11.3 & 0.2825 \end{bmatrix} = \begin{bmatrix} \widehat{\text{Var}}[\hat{\beta}_{1\mathbf{W}}] & \widehat{\text{Cov}}[\hat{\beta}_{1\mathbf{W}}\,,\hat{\beta}_{2\mathbf{W}}] \\ \widehat{\text{Cov}}[\hat{\beta}_{2\mathbf{W}}\,,\hat{\beta}_{1\mathbf{W}}] & \widehat{\text{Var}}[\hat{\beta}_{2\mathbf{W}}] \end{bmatrix}. \qquad [12]$$

Observación: Las varianzas y la covarianza estimadas anteriores también pueden calcularse en este caso utilizando las fórmulas específicas para el modelo RLS que se obtienen para los elementos de $\widehat{\text{Var}}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}] = \hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}$ a partir de [M2.2]:

$$\begin{aligned}
\widehat{\text{Var}}[\hat{\beta}_{1\mathbf{W}}] &= \hat{\sigma}^2\left(\frac{1}{N} + \frac{\bar{x}_2^2}{\sum_{i=1}^{N}(x_{i2} - \bar{x}_2)^2}\right) = 45.2 \times \left(\frac{1}{5} + \frac{40^2}{160}\right) = 461.04, \\
\widehat{\text{Var}}[\hat{\beta}_{2\mathbf{W}}] &= \hat{\sigma}^2\left(\frac{1}{\sum_{i=1}^{N}(x_{i2} - \bar{x}_2)^2}\right) = 45.2 \times \left(\frac{1}{160}\right) = 0.2825, \\
\widehat{\text{Cov}}[\hat{\beta}_{1\mathbf{W}}\,,\hat{\beta}_{2\mathbf{W}}] &= \hat{\sigma}^2\left(-\frac{\bar{x}_2}{\sum_{i=1}^{N}(x_{i2} - \bar{x}_2)^2}\right) = 45.2 \times \left(-\frac{40}{160}\right) = -11.3.
\end{aligned}$$

Resumen del Modelo Estimado:

E

$$\begin{aligned}
\hat{Y} \;=\; &\underset{(21.47)}{38.00} \;+\; \underset{(0.53)}{1.55}X_2\,, \\
N = 5,\; R^2 &= 0.7392,\; \bar{R}^2 = 0.6523,\; \hat{\sigma}^2 = 45.2.
\end{aligned}$$