Profesor: Héctor Bahamonde.
e:hector.bahamonde@uoh.cl
w:www.hectorbahamonde.com
Curso: OLS.

La Mecánica detras del OLS Pensemos en la relación entre *educación* e *ingreso*, "controlando por" *años de experiencia laboral*. Supongamos que tenemos la siguiente base de datos:

Nombre (i)	**Ingreso (Y)**	**Educacion (x1)**	**Experiencia (x2)**
Pedro	3	2	2
Juan	5	7	4
Diego	7	3	6

Hypotesis: "A más educación, más ingreso, para el promedio de años de experiencia laboral". A que se refiere "para el promedio de años promedio"?

I. Cuánto sube mi ingreso si aumento mi educacion?

El modelo de regresión lineal está dado por la siguiente formula:

$$Y_i \;=\; \beta 0 + \beta 1 x1_i \;+\; \beta 2 x2_i + e_i$$

- **Lo que conocemos:** $\boldsymbol{x}$ e y.
- **Lo que no conocemos, pero debemos estimar**: $\boldsymbol{\beta}$ y ϵ.

Volvamos a repensar la formula del modelo de regresión, pero en términos de matrices:

Definamos lo que conocemos:

$$Y \;=\; \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$$

$$X = \begin{bmatrix} 2 & 2 \\ 7 & 4 \\ 3 & 6 \end{bmatrix}$$

...y veamos cómo se ve OLS pero con matrices:

$$\begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}_y \;=\; \beta 0 + \beta 1 \begin{bmatrix} 2 \\ 7 \\ 3 \end{bmatrix}_{x1} \;+\; \beta 2 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}_{x2} + e_i$$

- Es fácil ver que:
- Debemos multiplicar $\beta 1$ por $x1$ y $\beta 2$ por $x2$.
- $\beta 0$, $\beta 1$, $\beta 2$ y ϵ_i son cantidades desconocidas, pero que debemos estimar.
- $\beta 0$, $\beta 1$y $\beta 2$ son escalares constantes, donde $\boldsymbol{\beta} \;=\; \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}_{\boldsymbol{\beta}}$
- El único parámetro que está indexado, es ϵ_i. Otra clase hablaremos de esto. **Adelanto**: Pensemos el caso de "Pedro". Si por ejemplo $\beta 0 \;=\; -3$, $\beta 1 \;=\; 1$ y $\beta 2 \;=\; 2$, tendremos que $y_{\text{Pedro}} - 3 \;+\; 1(2) + 2(2) = 3$. Entonces:

$$y_{\text{Pedro}} \;=\; 3 \;=\; -3 \;+1(2) \;+2(2) \;+\; 0$$

Aquí ϵ_{Pedro}=0. En ese sentido, e_i es la diferencia entre lo que estimamos y lo que observamos. "Filosoficamente", significa otra cosa.

Continuemos.

Formula para sacar β (el efecto de X sobre Y):

$\beta \;=\; \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T y$

Recapitulemos lo observado, y hagamos el cálculo:

$$y = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$$

$$\boldsymbol{X} = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 7 & 4 \\ 1 & 3 & 6 \end{bmatrix}$$

$$\boldsymbol{X}^T = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 7 & 3 \\ 2 & 4 & 6 \end{bmatrix}$$

$$\boldsymbol{X}^T \times \boldsymbol{X} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 7 & 3 \\ 2 & 4 & 6 \end{bmatrix}_{\boldsymbol{X}^T} \times \begin{bmatrix} 1 & 2 & 2 \\ 1 & 7 & 4 \\ 1 & 3 & 6 \end{bmatrix}_{\boldsymbol{X}} = \begin{bmatrix} 3 & 12 & 12 \\ 12 & 62 & 50 \\ 12 & 50 & 56 \end{bmatrix}$$

$$\left(\boldsymbol{X}^T \times \boldsymbol{X}\right)^{-1} = \frac{1}{\left(\boldsymbol{X}^T \boldsymbol{X}\right)} = \frac{1}{\det\left(\boldsymbol{X}^T \boldsymbol{X}\right)} \times \text{Adj}\left(\boldsymbol{X}^T \boldsymbol{X}\right) = \begin{bmatrix} 3 & -0.22 & -0.44 \\ -0.22 & 0.074 & -0.0185 \\ -0.44 & -0.0185 & 0.129 \end{bmatrix}$$

$$\boldsymbol{\beta} = \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T y = \begin{bmatrix} 3 & -0.22 & -0.44 \\ -0.22 & 0.074 & -0.0185 \\ -0.44 & -0.0185 & 0.129 \end{bmatrix} \times \begin{bmatrix} 1 & 1 & 1 \\ 2 & 7 & 3 \\ 2 & 4 & 6 \end{bmatrix} \times \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}_{\boldsymbol{\beta}}$$

Esto quiere decir que $\beta 0 = 1$, $\beta 1 = 0$ y $\beta 2 = 1$. Volvamos a re-escribir nuestra formula:

$$\boldsymbol{\beta} \;=\; \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T y \;=\; \begin{bmatrix}1\\0\\1\end{bmatrix}_{\boldsymbol{\beta}}$$

La formula que teníamos antes: $\begin{bmatrix}3\\5\\7\end{bmatrix}_y \;=\; \beta 0 + \beta 1 \begin{bmatrix}2\\7\\3\end{bmatrix}_{x1} \;+\; \beta 2 \begin{bmatrix}2\\4\\6\end{bmatrix}_{x2} + e_i$

Los resultados que tenemos ahora: $\begin{bmatrix}3\\5\\7\end{bmatrix}_y \;=\; 1 + 0 \begin{bmatrix}2\\7\\3\end{bmatrix}_{x1} \;+\; 1 \begin{bmatrix}2\\4\\6\end{bmatrix}_{x2} + e_i$