Econometría

Diplomado Banco Central de Honduras

Instituto de Economía

Pontificia Universidad Católica de Chile

Juan Ignacio Urquiza – Junio 2022

¿Qué es la econometría?

- Es una disciplina que emplea la teoría económica y los métodos estadísticos para estudiar relaciones económicas, contrastar distintas teorías, pronosticar variables, y evaluar políticas en forma cuantitativa, utilizando bases de datos.
- Combina elementos de la teoría económica, las matemáticas y la estadística.
- Es una aplicación del método científico en economía.

Función de Esperanza Condicional

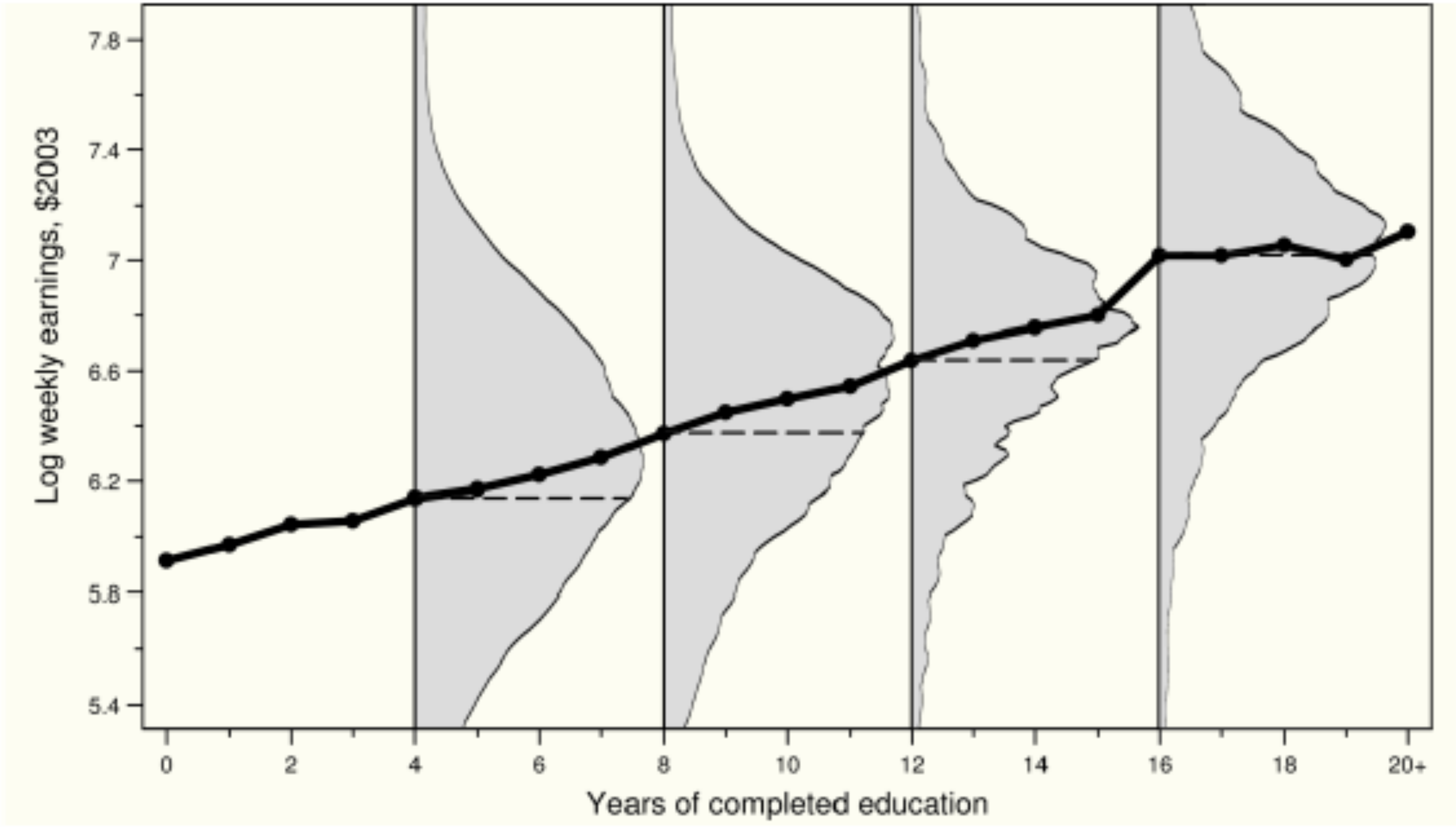
- Se define a la función de esperanza condicional (FEC) para una variable Y_i como su valor esperado cuando X_i está fijo:

$$FEC\colon E(Y_i|X_i)$$

- Dado que X_i es una variable aleatoria, entonces la FEC es también aleatoria.
- También se puede definir la FEC para un valor particular de X_i. Por ejemplo:

$$E(Y_i|X_i = x)$$

Ejemplo


Log weekly earnings, $2003
7.8
7.4
7
6.6
6.2
5.8
5.4
0
2
4
6
8
10
12
14
16
18
20+
Years of completed education

Ley de Esperanzas Iteradas

- La LEI es un complemento importante, que permite obtener el valor esperado de Y_i a partir del promedio de todas las FEC.

$$E(Y_i) = E\left[E\left(Y_i \middle| X_i\right)\right]$$

- Por lo tanto, nos permite descomponer a cualquier variable aleatoria tal que:

$$Y_i = E\left(Y_i \middle| X_i\right) + \varepsilon_i$$

donde $E\left(\varepsilon_i \middle| X_i\right) = 0$ y ε_i es ortogonal a cualquier función de X_i.

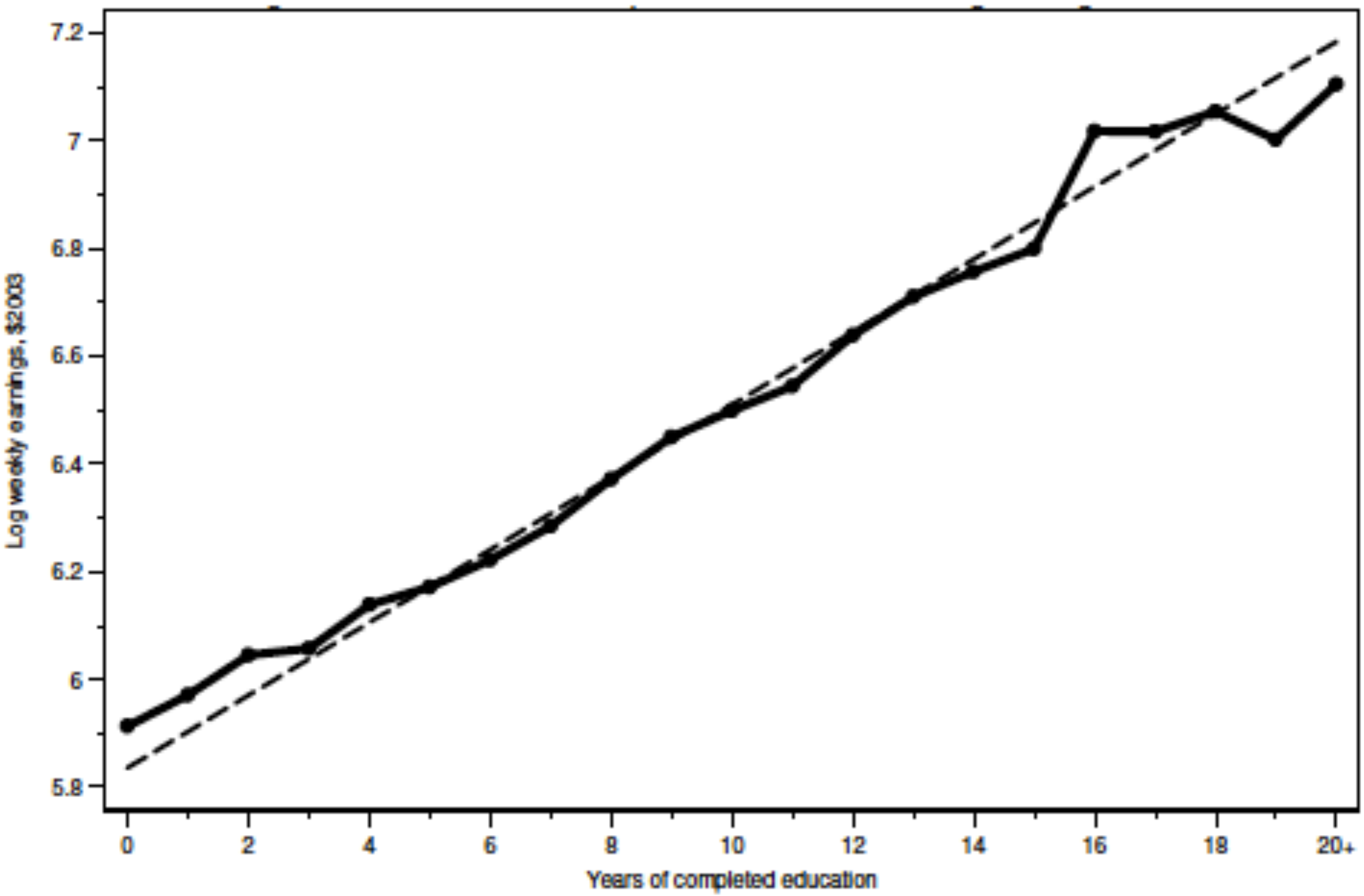
Descomposición de la FEC

- Esto implica que siempre podemos representar a una variable aleatoria Y_i como la suma de su esperanza condicional a X_i más algo que no está relacionado a X_i.
- Además, se puede demostrar que la FEC es el mejor predictor de Y_i dado X_i – menor error cuadrático medio.
- Por lo tanto, la FEC es una buena manera de caracterizar la relación entre X_i e Y_i.

Función de Regresión Poblacional

- ¿Cuál es la relación entre la FCE y la función de regresión poblacional (FRP)?
- Podemos pensar en la FRP como un acercamiento a la FEC.
- En particular, se puede demostrar que:
  - Si la FEC es lineal, entonces la FEC es la FRP.
  - La FRP es el mejor predictor *lineal* de Y_i dado X_i.
  - La FRP es la mejor aproximación lineal de la FEC.

Ejemplo



7.2
7
6.8
6.6
6.4
6.2
6
5.8
Log weekly earnings, $2003
0
2
4
6
8
10
12
14
16
18
20+
Years of completed education

Modelo de Regresión Lineal Simple

- MRL Simple:

$$\boxed{Y = \beta_0 + \beta_1 X + \varepsilon}$$

  - Y: variable dependiente, explicada, predicha, de respuesta, regresando…
  - X: variable independiente, explicativa, predictora, de control, regresor…
  - β_0 , β_1: parámetros poblacionales desconocidos.
  - ε: término de error o perturbación inobservable.
    - Representa factores distintos de X que afectan a Y.

Ejemplos

- <u>Ecuación de salarios</u>:
  - Y: salario
  - X: años de educación
  - ε: experiencia laboral, capacidad/habilidad innata, años de antigüedad en la empresa…
- <u>Rendimiento y fertilizante</u>:
  - Y: rendimiento de un cultivo
  - X: cantidad de fertilizante
  - ε: calidad de la tierra, lluvia, sol…

Modelo de Regresión Lineal Múltiple

- MRL Múltiple:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K + \varepsilon$$

donde β_0 es la constante o el intercepto y $\beta_{i \neq 0}$ es el parámetro de pendiente asociado a una de las K variables explicativas.

  - Note que hay $K+1$ parámetros poblacionales desconocidos.

- Es importante destacar que el análisis de regresión no se aplica para determinar causalidad sino para ver si dos variables guardan una relación positiva o negativa.

Interpretación

- Los parámetros de la FRP muestran la relación entre X_j e Y, manteniendo todo lo demás constante.
- Es decir, los parámetros de pendiente representan el cambio esperado en Y ante un cambio en X_j, manteniendo todo lo demás constante:

$$\Delta E(Y|X_1, X_2, \ldots, X_K) = \beta_j \Delta X_j$$

- Por lo tanto,

$$\beta_j = \frac{\Delta E(Y|X_1, X_2, \ldots, X_K)}{\Delta X_j}$$

Estimación

- Sea el modelo de regresión lineal (MRL) simple:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

- Nuestro objetivo consiste en estimar los parámetros poblacionales a partir de una muestra aleatoria.
- Para cada observación de la muestra $\{i = 1, 2, \ldots, n\}$ se cumple que:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

donde ε_i es el término de error para la observación *i*.

Estimación

- Al análogo muestral del término de error se lo conoce como *"residual o residuo"*.

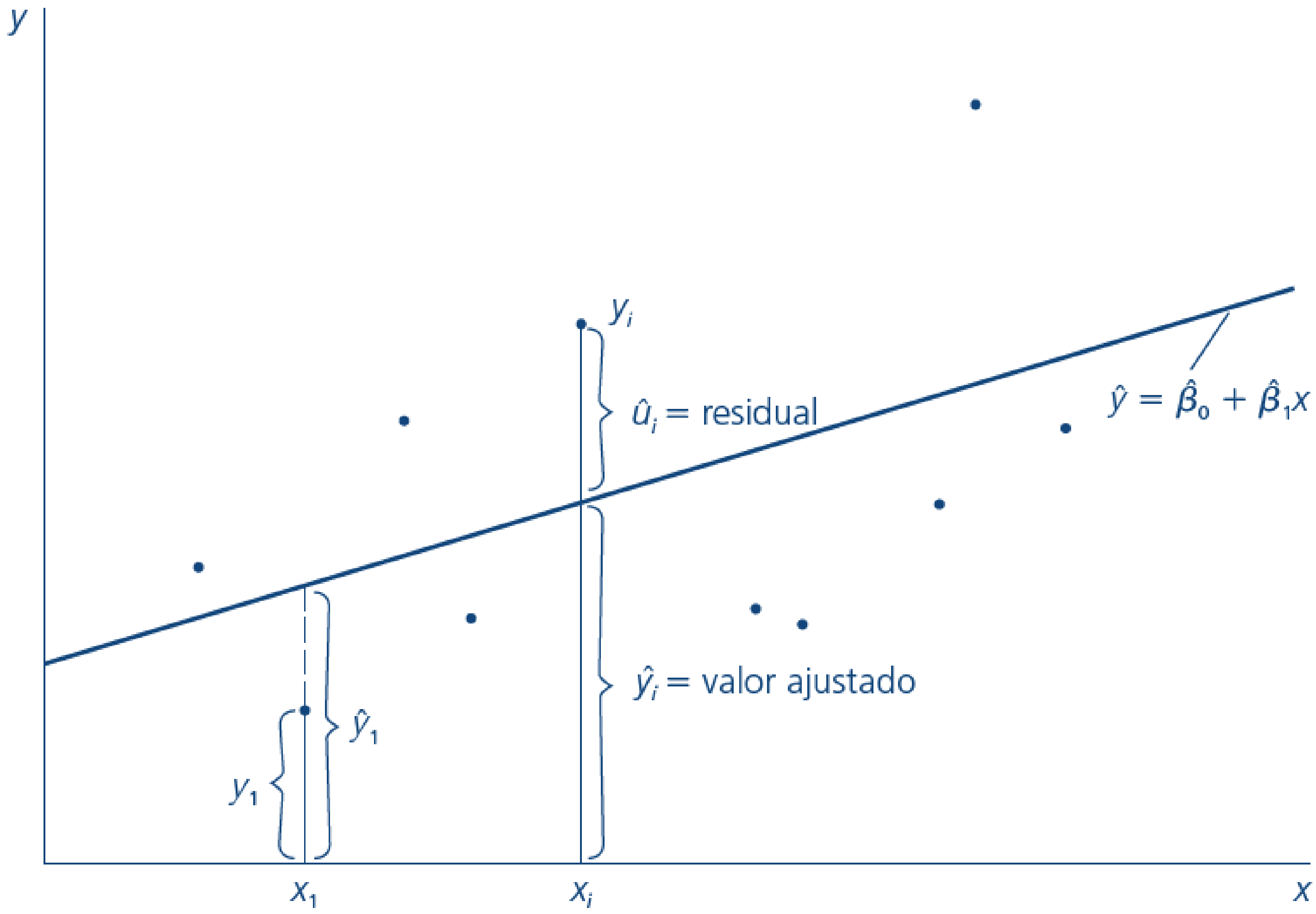
- ***Definición***: llamaremos residuo a la diferencia entre el *"valor observado"* y el *"valor ajustado"*:

$$\widehat{\varepsilon}_i = Y_i - \widehat{Y}_i = Y_i - \left(\widehat{\beta}_0 + \widehat{\beta}_1 X_i\right)$$

- Los estimadores de Mínimos Cuadrados Ordinarios (MCO) son los argumentos que minimizan la suma de los residuos al cuadrado:

$$\sum_{i=1}^{n} \widehat{\varepsilon}_i^2 = \sum_{i=1}^{n} (Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i)^2$$

Valor observado vs. ajustado



Estimación

- Este criterio penaliza más a los residuos más grandes.
- CPO:

$$\sum_i \widehat{\varepsilon}_i = 0 \quad \Rightarrow \quad \widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1 \overline{X}$$

$$\sum_i X_i \widehat{\varepsilon}_i = 0 \quad \Rightarrow \quad \widehat{\beta}_1 = \frac{\sum_i (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_i (X_i - \overline{X})^2}$$

- Implicancia:
  - Si en la muestra X e Y se correlacionan positivamente (negativamente), entonces $\hat{\beta}_1$ es positivo (negativo).

Forma matricial

- Sea el modelo de regresión lineal (MRL) múltiple:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K + \varepsilon$$

- En términos matriciales, se puede representar de la siguiente manera:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

  - En este caso, las letras minúsculas representan vectores y las mayúsculas matrices.

Forma matricial

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \beta_0 \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \beta_1 \begin{bmatrix} x_{1,1} \\ x_{2,1} \\ \vdots \\ x_{n,1} \end{bmatrix} + \beta_2 \begin{bmatrix} x_{1,2} \\ x_{2,2} \\ \vdots \\ x_{n,2} \end{bmatrix} + \cdots + \beta_k \begin{bmatrix} x_{1,k} \\ x_{2,k} \\ \vdots \\ x_{n,k} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{1,1} & x_{1,2} & \cdots & x_{1,k} \\ 1 & x_{2,1} & x_{2,2} & \cdots & x_{2,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & x_{n,2} & \cdots & x_{n,k} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\boxed{\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}}$$

Estimación

- Se debe minimizar la suma de cuadrados residuales.
- ¿Cómo?

$$
\begin{aligned}
\rightarrow \quad SCR &= \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} = \left(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)'\left(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right) \\
&= \boldsymbol{y}'\boldsymbol{y} - \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} + \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} \\
&= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{y}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} + \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}
\end{aligned}
$$

puesto que, tratándose de escalares, $\left(\widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}\right) = \left(\boldsymbol{y}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)$.

- Reglas para la diferenciación:

$$\boxed{\frac{\partial \boldsymbol{A}\widehat{\boldsymbol{\beta}}}{\partial \widehat{\boldsymbol{\beta}}} = \boldsymbol{A}'}$$

$$\boxed{\frac{\partial \widehat{\boldsymbol{\beta}}'\boldsymbol{C}\widehat{\boldsymbol{\beta}}}{\partial \widehat{\boldsymbol{\beta}}} = 2\boldsymbol{C}\widehat{\boldsymbol{\beta}}}$$

Estimación

- Sea el MRL múltiple en notación matricial:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$$

- Se debe minimizar la suma de cuadrados de residuales.

$$\begin{aligned}\min_{\widehat{\boldsymbol{\beta}}} SCR &= \min_{\widehat{\boldsymbol{\beta}}} \left(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)' \left(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right) \\ &= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{y}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} + \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}\end{aligned}$$

- CPO:

$$\frac{\partial SCR}{\partial \widehat{\boldsymbol{\beta}}} = \mathbf{0} \quad \rightarrow \quad -2\boldsymbol{X}'\boldsymbol{y} + 2(\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}} = \mathbf{0}$$

Estimación

- Por lo tanto, tenemos que:

$$\boxed{\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}}$$

- El vector $\widehat{\boldsymbol{\beta}}$ recibe el nombre de estimador de MCO.
- Para que esta solución sea efectivamente un mínimo es necesario que:

$$\frac{\partial SCR}{\partial \widehat{\boldsymbol{\beta}} \partial \widehat{\boldsymbol{\beta}}'} = 2(\boldsymbol{X}'\boldsymbol{X})$$

  sea definida positiva.
- Puesto que $\boldsymbol{X}$ es de rango completo (RLM.3), entonces $\widehat{\boldsymbol{\beta}}$ es único y minimiza la suma de cuadrados residuales.

Ejemplo

- Usted dispone de los siguientes datos muestrales:

x	y
0	4
1	6
2	7
3	9
4	11

- $n = 5, k = 1.$
- Estime la regresión simple de Y sobre X por MCO.

Ejemplo

- Caracterización de **X** e **y**:

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \quad , \quad y = \begin{bmatrix} 4 \\ 6 \\ 7 \\ 9 \\ 11 \end{bmatrix}$$

- Estimador de MCO:

$$\widehat{\boldsymbol{\beta}} = (X'X)^{-1}(X'y)$$

Ejemplo

- Cómputo de la matriz $(\boldsymbol{X}'\boldsymbol{X})^{-1}$:

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$

$$X'X = \begin{bmatrix} 5 & 10 \\ 10 & 30 \end{bmatrix}$$

$$(X'X)^{-1} = \begin{bmatrix} 0.6 & -0.2 \\ -0.2 & 0.1 \end{bmatrix}$$

Ejemplo

- Cómputo del vector ***X´*y**:

$$X'y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \\ 7 \\ 9 \\ 11 \end{bmatrix}$$

$$X'y = \begin{bmatrix} 37 \\ 91 \end{bmatrix}$$

- Cálculo de los estimadores:

$$\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \widehat{\beta_0} \\ \widehat{\beta_1} \end{bmatrix} = \begin{bmatrix} 0.6 & -0.2 \\ -0.2 & 0.1 \end{bmatrix} \begin{bmatrix} 37 \\ 91 \end{bmatrix} = \begin{bmatrix} 4 \\ 1.7 \end{bmatrix}$$

Propiedades algebraicas

- Las propiedades algebraicas se derivan de las llamadas ecuaciones normales para MCO.

- Recuerde que:

$$-\boldsymbol{X}'\boldsymbol{y} + (\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}} = \mathbf{0}$$

$$-\boldsymbol{X}'\left(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right) = \mathbf{0}$$

$$-\boldsymbol{X}'\widehat{\boldsymbol{u}} = \mathbf{0} \quad \rightarrow \quad \boldsymbol{X}'\widehat{\boldsymbol{u}} = \mathbf{0}$$

- Esto implica que los residuos suman cero y que son ortogonales a los regresores.

- Además, implica que: $\widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\widehat{\boldsymbol{u}} = \widehat{\boldsymbol{y}}'\widehat{\boldsymbol{u}} = 0.$

Ejemplo

- Curva de regresión muestral: $\hat{y} = \widehat{\beta_0} + \widehat{\beta_1}x$

x	y	$\hat{y}$	$\hat{u} = y - \hat{y}$	$x * \hat{u}$	$\hat{y} * \hat{u}$
0	4	4	0	0	0
1	6	5,7	0,3	0,3	1,71
2	7	7,4	-0,4	-0,8	-2,96
3	9	9,1	-0,1	-0,3	-0,91
4	11	10,8	0,2	0,8	2,16
		SUMA=	0	0	0

Descomposición de la varianza

- Sea SCT la suma de cuadrados totales:

$$SCT = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

- Sea SCE la suma de cuadrados explicados:

$$SCE = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

- Sea SCR la suma de cuadrados residuales:

$$SCR = \sum_{i=1}^{n} \hat{u}_i^2$$

Bondad del ajuste

- Entonces, se puede demostrar que:

$$SCT = SCE + SCR$$

- A partir de esta descomposición, se define el R^2 o coeficiente de determinación como:

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT}$$

- Dado que $0 \leq SCE \leq SCT$, entonces $0 \leq R^2 \leq 1$.

- Corresponde a la proporción de la variación muestral de y que es explicada por la regresión de MCO.

Bondad del ajuste

- Resulta interesante destacar que el R^2 nunca disminuye cuando se agregan variables explicativas al modelo.
- Esto ocurre incluso si las variables que se agregan no son relevantes en el modelo poblacional.
- Se necesita de una medida que ajuste por la cantidad de regresores. Se define entonces al R^2 ajustado como:

$$\bar{R}^2 = 1 - \frac{(n-1)}{(n-k-1)} \times \frac{SCR}{SCT}$$

- El principal atractivo es que penaliza por el número de regresores. Nótese que incluso podría ser negativo.

```
. reg testscr str

      Source |       SS       df       MS              Number of obs =     420
-------------+------------------------------           F(  1,   418) =   22.58
       Model |  7794.11004     1  7794.11004           Prob > F      =  0.0000
    Residual |  144315.484   418  345.252353           R-squared     =  0.0512
-------------+------------------------------           Adj R-squared =  0.0490
       Total |  152109.594   419  363.030056           Root MSE      =  18.581

------------------------------------------------------------------------------
     testscr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -2.279808   .4798256    -4.75   0.000     -3.22298   -1.336637
       _cons |    698.933   9.467491    73.82   0.000     680.3231    717.5428
------------------------------------------------------------------------------
```

```
. reg testscr str el_pct expn_stu

      Source |       SS       df       MS              Number of obs =     420
-------------+------------------------------           F(  3,   416) =  107.45
       Model |  66409.8837     3  22136.6279           Prob > F      =  0.0000
    Residual |  85699.7099   416  206.008918           R-squared     =  0.4366
-------------+------------------------------           Adj R-squared =  0.4325
       Total |  152109.594   419  363.030056           Root MSE      =  14.353

------------------------------------------------------------------------------
     testscr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -.2863992   .4805232    -0.60   0.551    -1.230955     .658157
      el_pct |  -.6560227   .0391059   -16.78   0.000    -.7328924   -.5791529
    expn_stu |   .0038679   .0014121     2.74   0.006     .0010921    .0066437
       _cons |   649.5779   15.20572    42.72   0.000     619.6883    679.4676
------------------------------------------------------------------------------
```