

Regresión Lineal

Econ. Dax Mancilla

*Universidad Nacional del Callao
Consultoría GEM*

Ago. 8, 2021

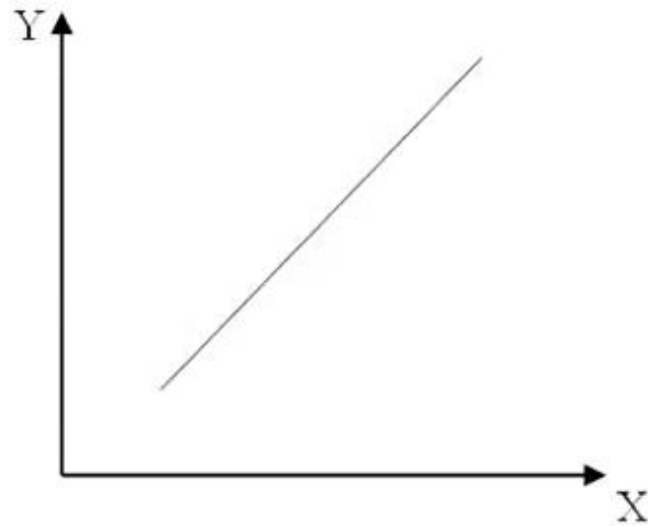
Modelo económico

- Los modelos económicos relacionan variables económicas mediante **relaciones causales** y suelen representarse matemáticamente mediante ecuaciones.
- Los modelos económicos describen una relación exacta o determinística.
- Los modelos económicos permiten realizar **predicciones teóricas**.
- Para comprobar que estas afirmaciones se cumplan en la realidad se necesita trabajar con **datos económicos**.
- Los datos económicos son piezas de información cuantitativa y cualitativa de una realidad concreta.

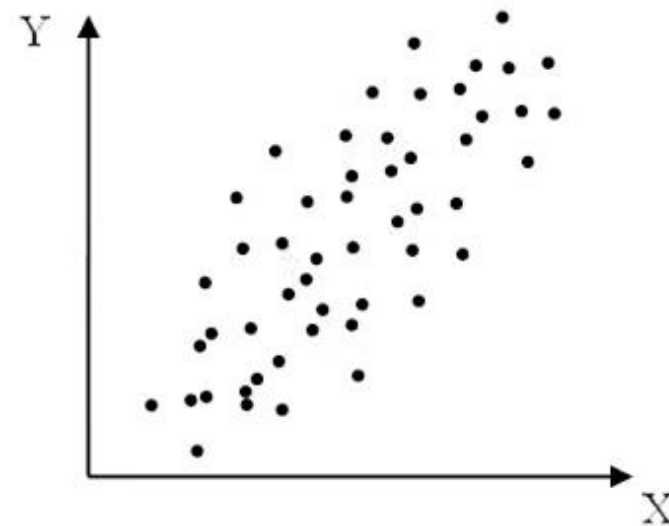
Modelo econométrico

- Para este análisis es necesario emplear técnicas estadísticas que garanticen la rigurosidad de la misma y minimizar los errores a la hora de hacer pronósticos.
- Los modelos econométricos es un modelo estadístico que relaciona dos o más variables.
- Describe una relación inexacta o estadística.

Relación determinística y estadística



Una relación determinística



Una relación estadística

El modelo de Regresión Lineal Clásico (MRLC)

- El objetivo del análisis de regresión es explicar y pronosticar el comportamiento de la variable dependiente a través del comportamiento de la o las variables independiente.

Función de Regresión Poblacional
(FRP):

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

El objetivo es encontrar una contraparte empírica usando la muestra de datos.

Función de Regresión Muestral
(FRM):

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

- \hat{Y}_i es un estimador de $E[Y|X]$
- $\hat{\beta}_0$ y $\hat{\beta}_1$ son estimadores de los parámetros poblacionales β_0 y β_1

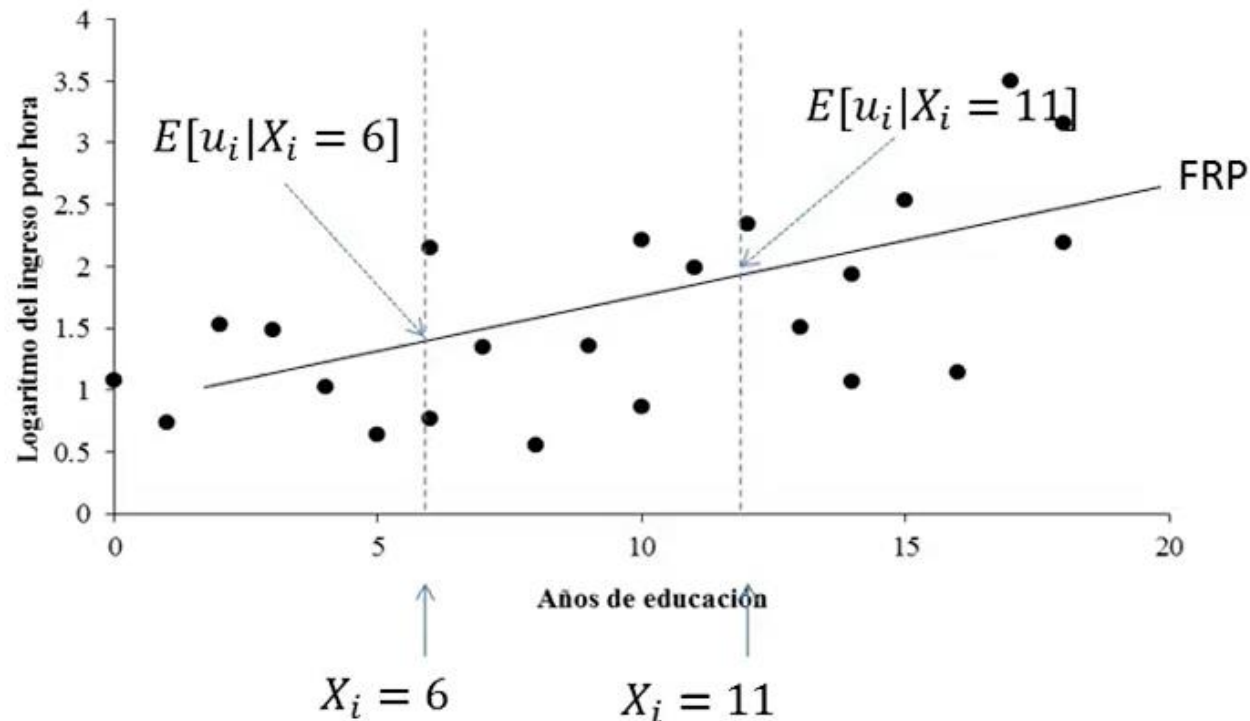
Supuestos del MRLC

Supuesto 1: La relación entre X e Y es una regresión (FRP) lineal en parámetros.

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

Supuesto 2: La esperanza condicional de u_i dado X_i es igual a cero

$$E[u_i|X_i] = 0$$

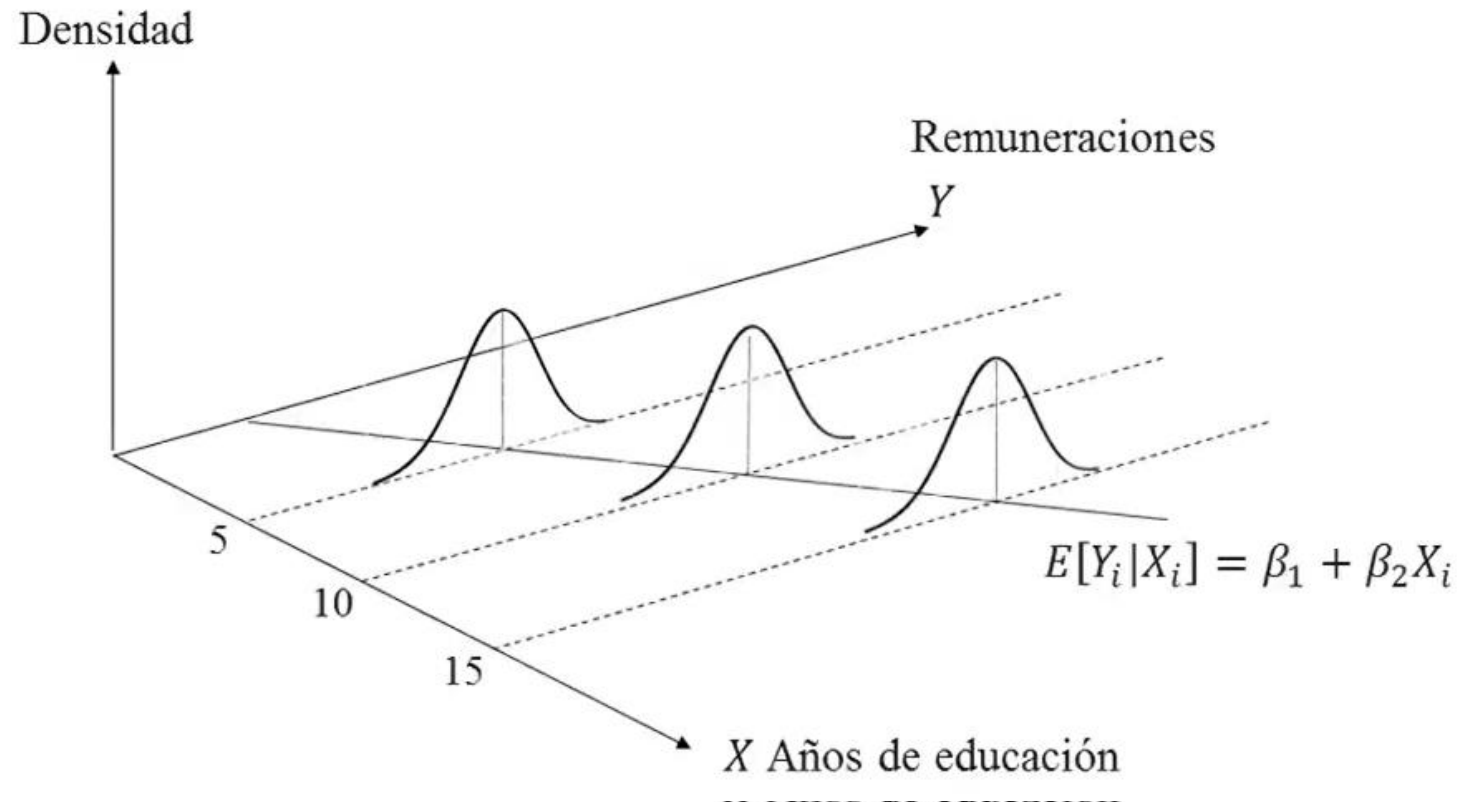


Supuestos del MRLC

Supuesto 3: Las perturbaciones son “esféricas”.

$$\text{var}(u_i|X_i) = \sigma^2 \text{ Homocedasticidad}$$

$$\text{cov}(u_i, u_j|X_i, X_j) = 0 \quad \forall i \neq j \text{ “No autocorrelación”}$$



La dispersión es constante a lo largo de la recta de regresión.

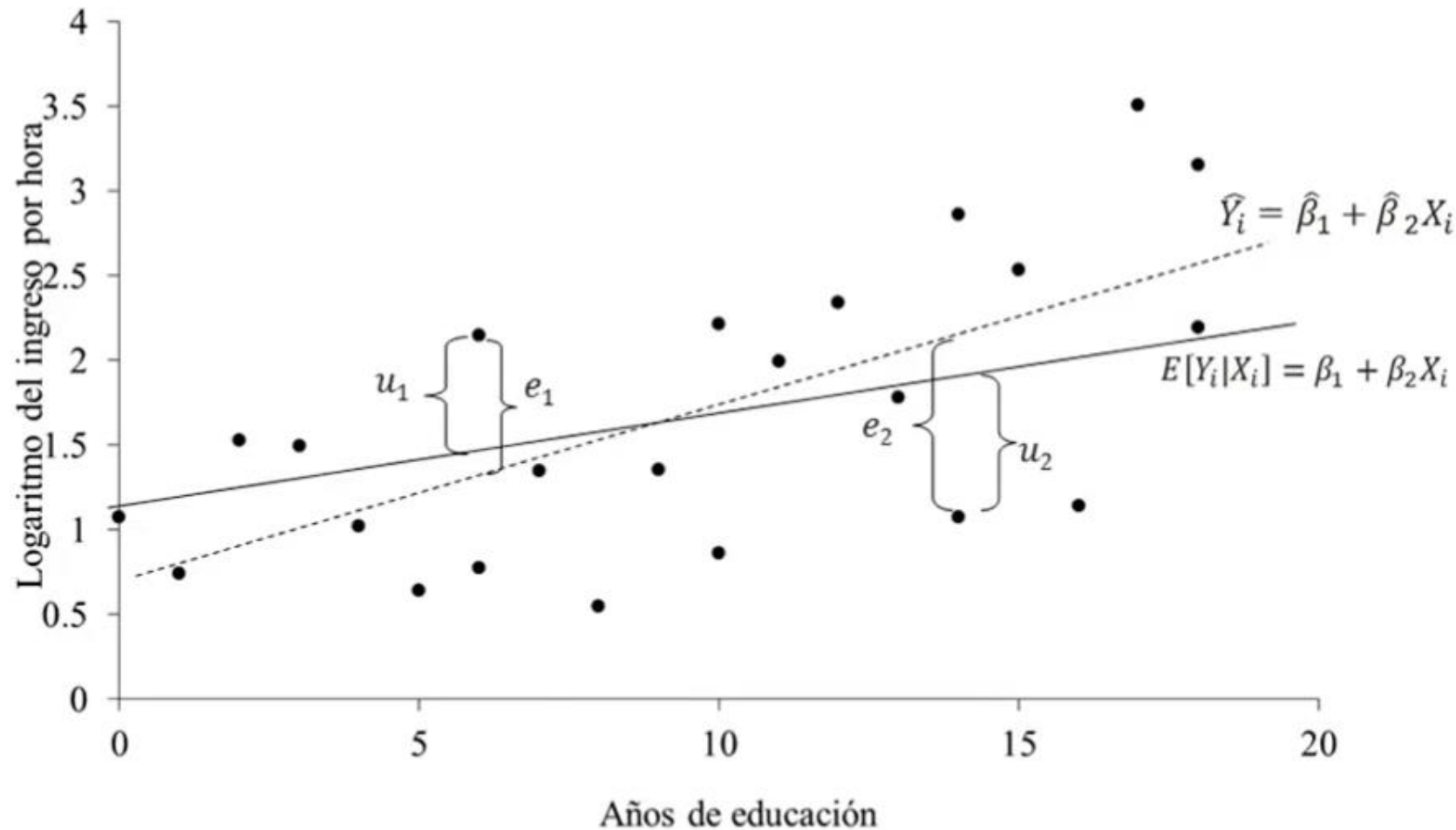
Supuestos del MRLC

Supuesto 4: Los errores siguen una distribución normal

Considerando a los supuestos 2 y 3, este supuesto dice que cada u_i es independiente e idénticamente distribuido de los demás errores , como una normal.

$$u_i \sim N(0, \sigma^2) \quad \forall i$$

Estimación por Mínimos Cuadrado Ordinario (MCO)



- Definimos al residuo como:

$$e_1 = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$$

Estimación por Mínimos Cuadrado Ordinario (MCO)

- Busca trazar una recta que minimice las distancias de las observaciones a la recta, es decir que los residuos en conjunto sean lo más pequeños posibles.

$$SCR = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$$

- El método consiste en escoger los valores de $\hat{\beta}_0$ y $\hat{\beta}_1$, tal que se minimice la SCR.
- La solución para ambos parámetros es la siguiente:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

Regresión Lineal Múltiple

Permite analizar la relación entre dos o más variables a través de ecuaciones. La Representación del modelo de regresión lineal múltiple en su forma matricial es la siguiente:

$$Y = \beta X + \epsilon$$

En donde la solución vendrá dado por la siguiente ecuación:

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

Significancia individual

La significancia individual viene dada por la siguiente hipótesis:

$$H_0: \beta_0 = 0$$

$$H_1: \beta_0 \neq 0$$

El objetivo es evaluar que los coeficientes son diferentes de 0 de forma individual. El estadístico de prueba sigue una distribución t de Student

Significancia global

La significancia global viene dada por la siguiente hipótesis:

$$H_0: \beta_1 = \beta_2 = \beta_k = 0$$

$$H_1: \beta_1 \neq \beta_2 \neq \beta_k \neq 0$$

El objetivo es evaluar que los coeficientes en conjunto son estadísticamente significativos. El estadístico de prueba sigue una distribución F.

Gracias