Econometría I

Regresión Multiple

Carlos Yanes Guerra | Departamento de Economía | 2024-04-08


WELCOME BACK!

😕

Preguntas de la sesion anterior?

Preliminar

La última vez:

1. Estimamos regresión simple **M.C.O**
2. Tenemos los primeros **test** de Parámetros
3. Estuvimos con el tema de retornos `salarios f(Experiencia)`

Formas funcionales

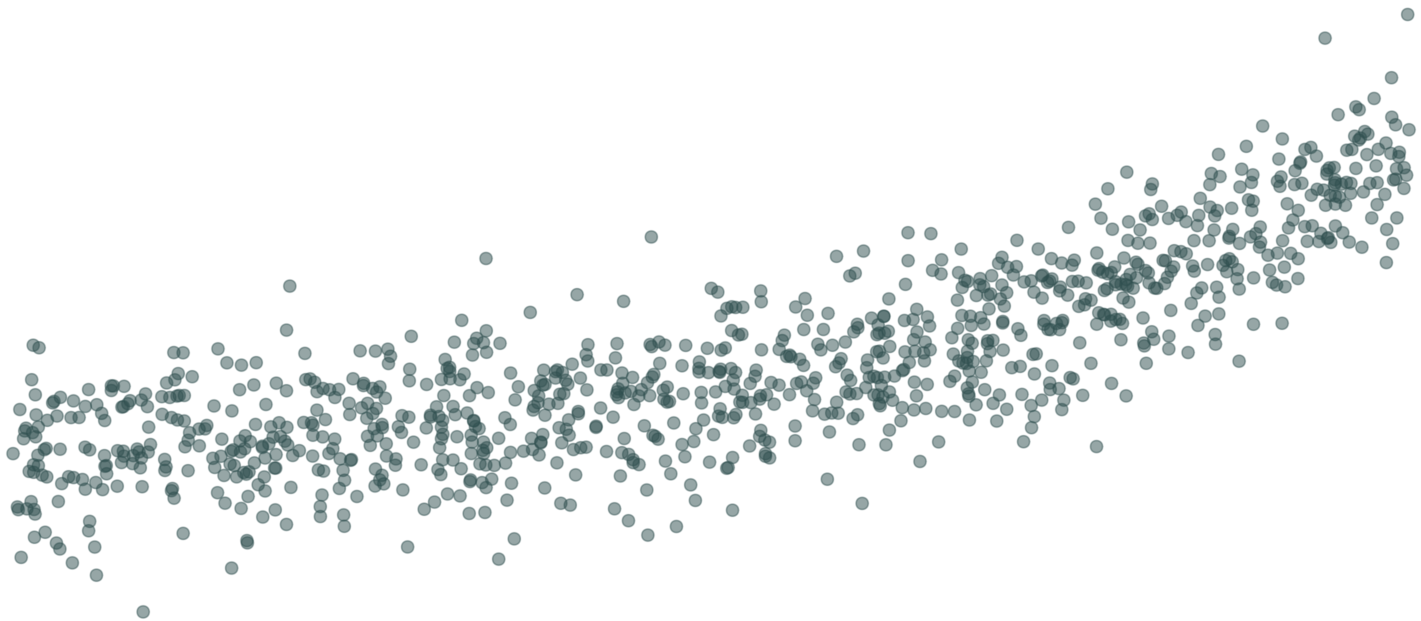
Formas Funcionales

Modelo	Ecuación	β_1	Lectura
N-N	$y = \beta_0 + \beta_1 x$	$\frac{\triangle y}{\triangle x}$	y cambia en β_1 unidades ante un cambio de x
N-L	$y = \beta_0 + \beta_1 Ln(x)$	$\frac{\triangle y}{\triangle x / x}$	y cambia en $\beta_1/100$ unidades ante un cambio del 1% de x
L-N	$Ln(y) = \beta_0 + \beta_1 x$	$\frac{\triangle y / y}{\triangle x}$	y cambia en $\beta_1 * 100\%$ unidades ante un cambio de una unidad de x
L-L	$Ln(y) = \beta_0 + \beta_1 Ln(x)$	$\frac{\triangle y / y}{\triangle x / x}$	Elasticidad: y cambia en $\beta_1\%$ ante un cambio de un 1% de x

📛 para este tipo de ajustes

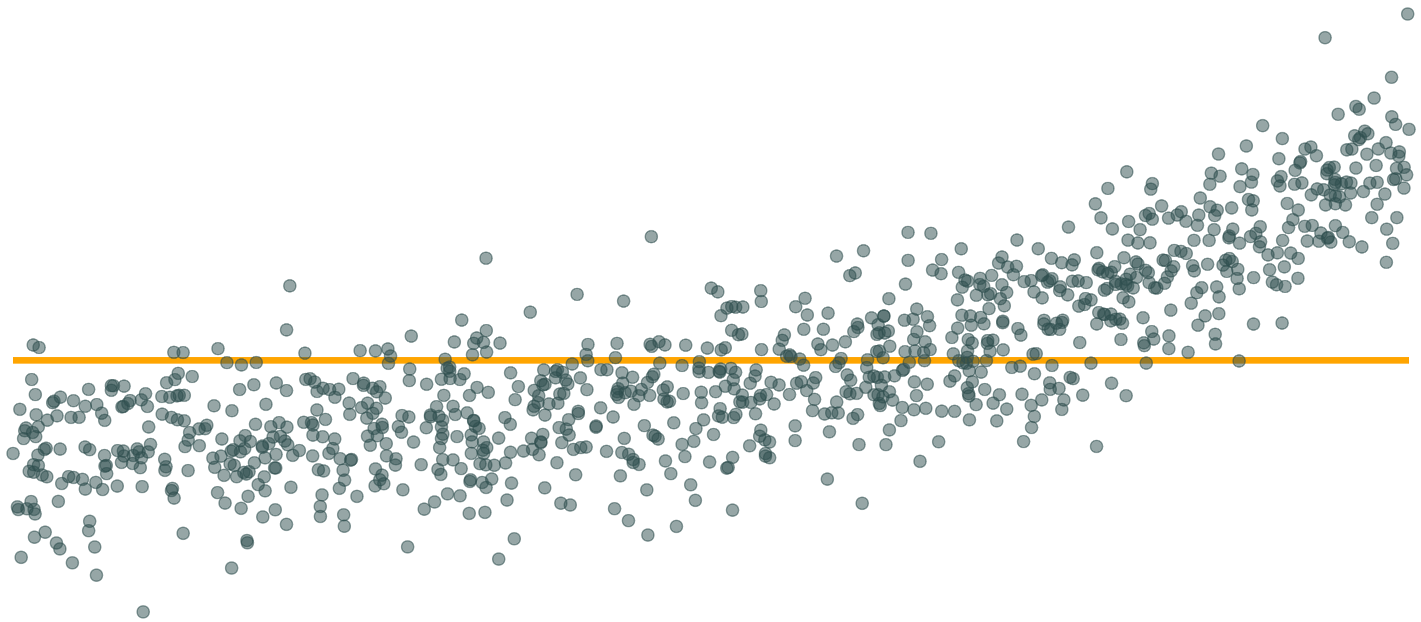
Formas Funcionales

Un gráfico de dispersión



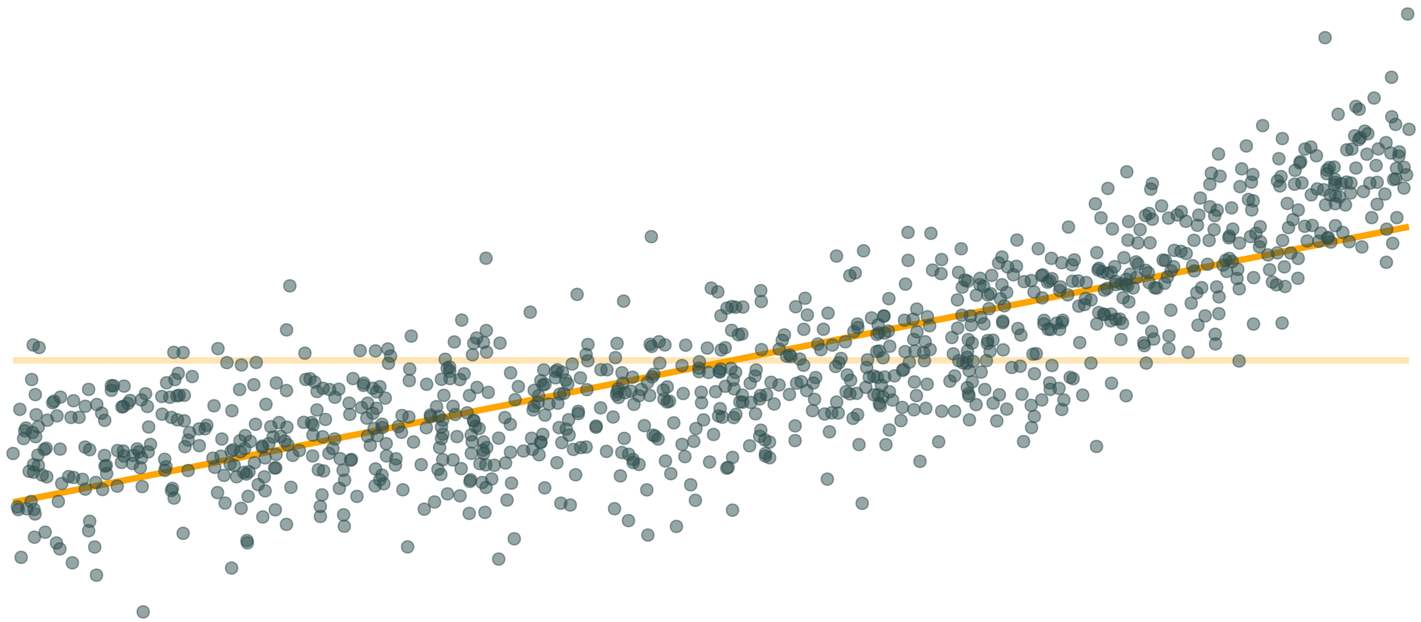
Formas Funcionales

$y_i = \beta_0 + u_i$



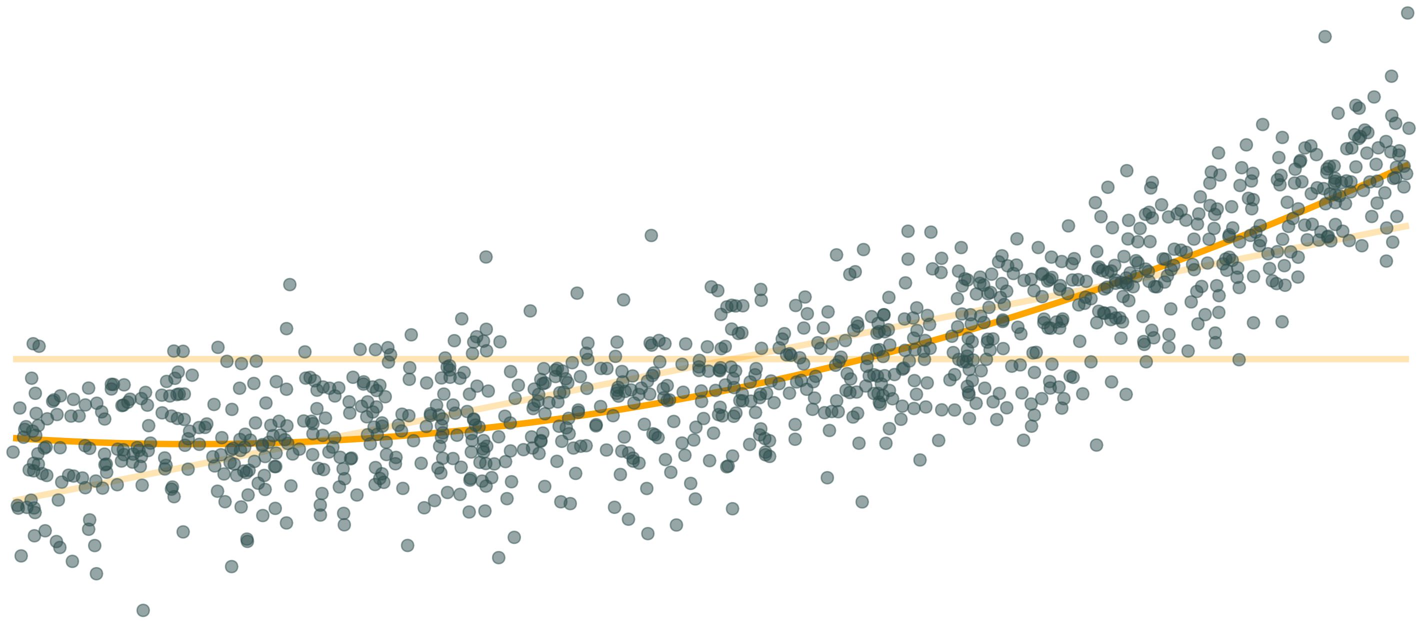
Formas Funcionales

$y_i = \beta_0 + \beta_1 x + u_i$



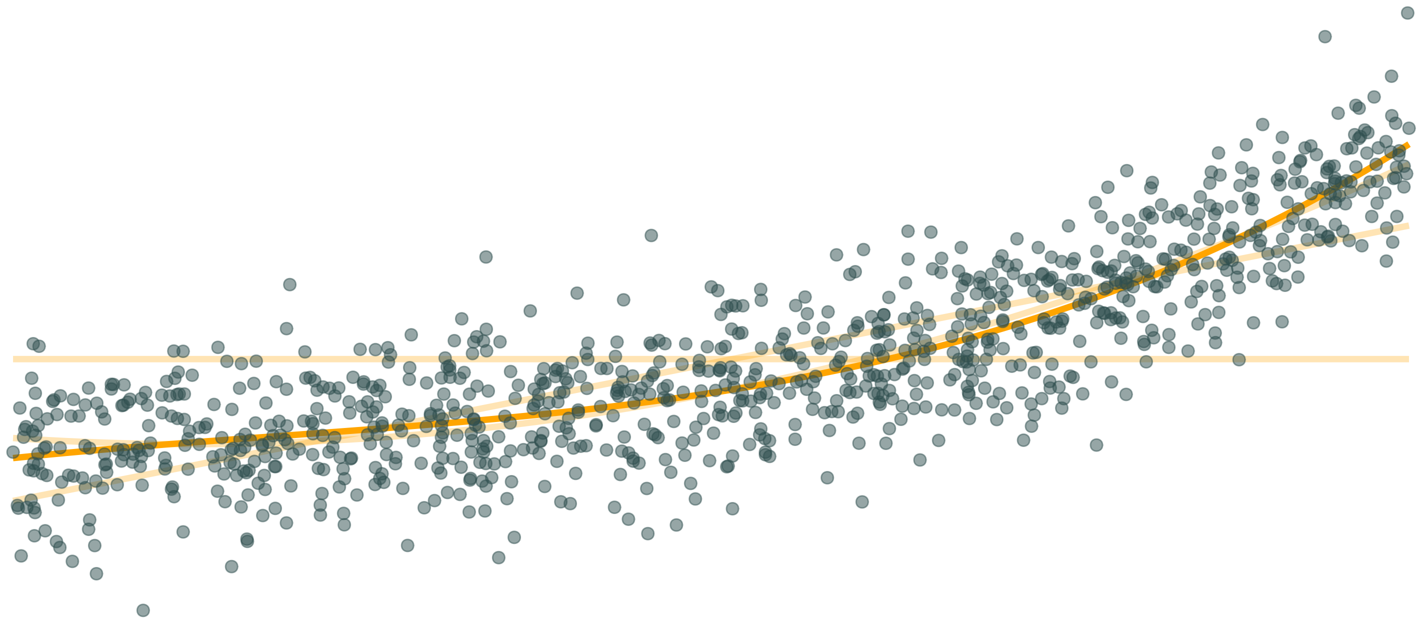
Formas Funcionales

$y_i = \beta_0 + \beta_1 x + \beta_2 x^2 + u_i$



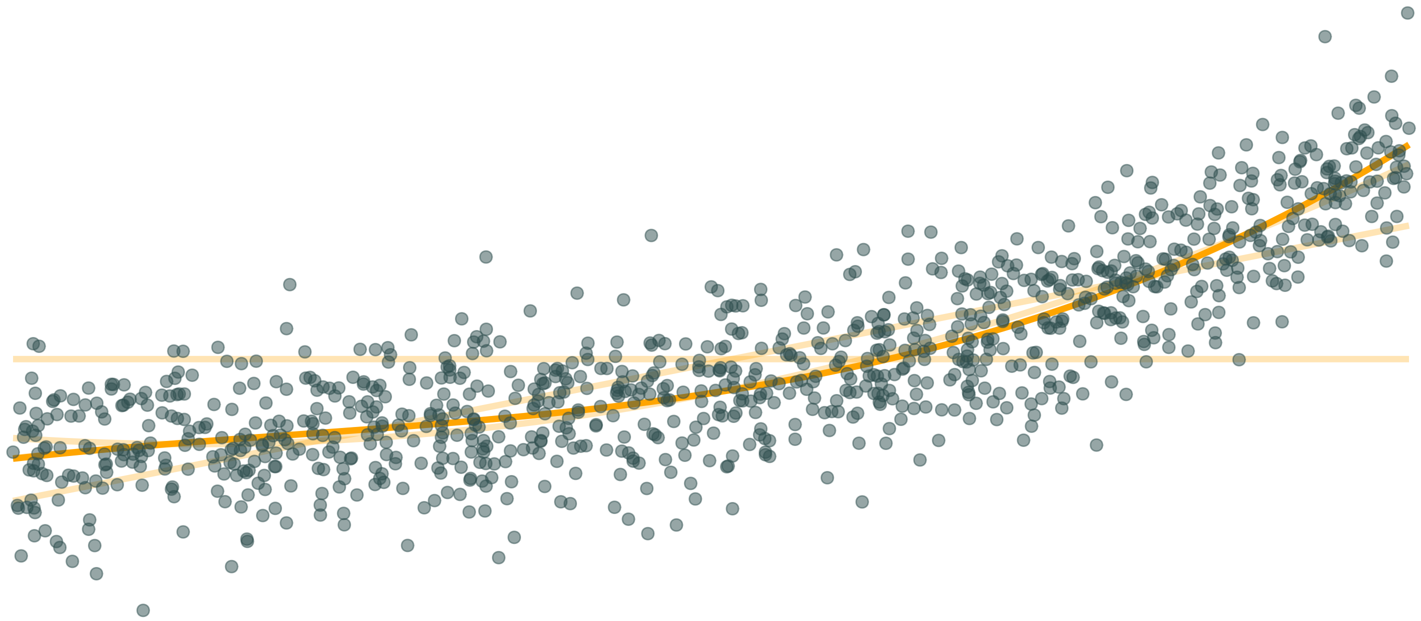
Formas Funcionales

$y_i = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + u_i$



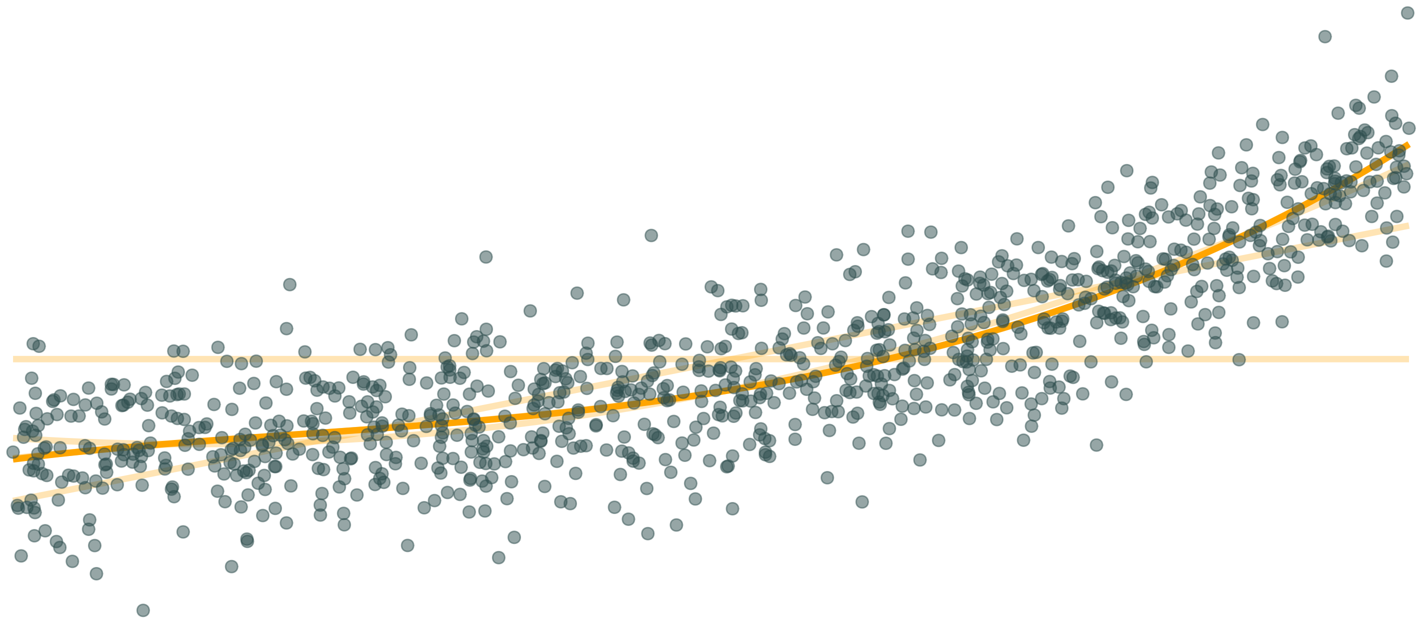
Formas Funcionales

$y_i = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + u_i$



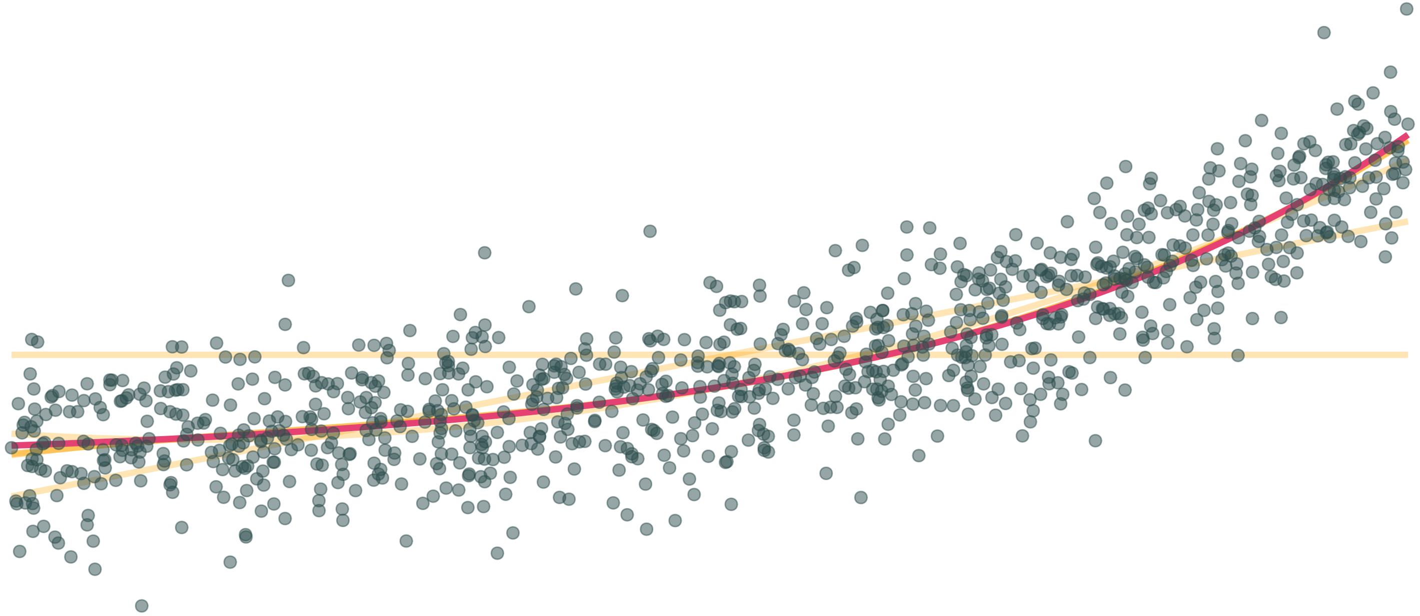
Formas Funcionales

$$y_i = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \beta_5 x^5 + u_i$$



Formas Funcionales

$y_i = 2e^x + u_i$ **siendo este el más real**



Regresión Múltiple

Regresión Múltiple

Cuando se controla (adiciona) mas variables a una regresión, los modelos de regresión lineal se vuelven herramientas mucho mas completas a la hora de estimar **efectos** sobre una variable *objetivo*. Desde luego tienden a ser **mejores** modelos.

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_\rho X_\rho + \mu_i$$

- Los parámetros distintos al **autónomo** serán considerados como parámetros de pendiente.
- Son modelos que se manejan de forma similar a la **regresión simple**.
- El supuesto mas importante:

$$E(u|x_1, x_2, x_3, \cdots, x_k) = 0$$

- Los efectos parciales se miden:

$$\triangle \hat{Y} = \triangle X_1 \hat{\beta}_1 + \triangle X_2 \hat{\beta}_2$$

Regresión Múltiple

Importante: la regresión lineal permite "ajustar" vía coeficientes $\beta_0, \ldots, \beta_p$ la mejor forma o manera de tratar un problema. Estos coeficientes o parámetros se denominan *marginales*.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1,i} + \hat{\beta}_2 x_{2,i} + \cdots + \hat{\beta}_p x_{p,i} + \varepsilon_i$$

Esto suelen aplicarse en dos escenarios distintos con objetivos bastante diferenciados:

1. **Inferencia Causal** Estimar e interpretar los coeficientes.

2. **Predicción** El enfoque solo es estimar resultados.

Independientemente del objetivo, la forma de "ajustar" (estimar) el modelo es la misma.

Regresión Múltiple

Ajuste de la recta de regresión

Como ocurre con muchos métodos de aprendizaje estadístico, la **regresión** se centra en **minimizar** alguna medida de pérdida/error.

$$e_i = y_i - \hat{y}_i$$

Tenemos entonces que usar para **regresión** lo que es (RSS) *Residual Sum Squares* (siglas en ingles) o la suma de los **residuos al cuadrado** de la regresión.

$$\mathrm{RSS} = e_1^2 + e_2^2 + \cdots + e_n^2 = \sum_{i=1}^{n} e_i^2$$

El MCO escoge el(los) mejores $\hat{\beta}_j$ que minimizan la **RSS**.

Regresión Múltiple

Elección del modelo

Una primera forma para mirar que tanto ajuste tiene un modelo es el R^2, pero, también es bueno mirar *el residuo estándar de la regresión (RSE)*

Residuo estándar de la regresión (RSE)

$$\mathrm{RSE} = \sqrt{\frac{1}{n-p-1}\mathrm{RSS}} = \sqrt{\frac{1}{n-p-1}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}$$

Recuerde que la formula del **R-cuadrado** (**R2**) es:

$$R^2 = \frac{\mathrm{TSS} - \mathrm{RSS}}{\mathrm{TSS}} = 1 - \frac{\mathrm{RSS}}{\mathrm{TSS}} \quad \text{donde} \quad \mathrm{TSS} = \sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2$$

Regresión Múltiple

En la **comparación** de modelos vamos a ver que el R^2 por si solo tiende a sobrestimar la capacidad de los modelos.

$$R^2 = 1 - \frac{\mathrm{RSS}}{\mathrm{TSS}}$$

Al adherir nuevas variables la RSS $\downarrow$ pero en cambio TSS no lo hace. Así, R^2 se incrementa.

Cuando usamos la **RSE**, esta penaliza ligeramente la incorporación de nuevas variables:

$$\mathrm{RSE} = \sqrt{\frac{1}{n-p-1}\mathrm{RSS}}$$

Pero ocurre que **al adicionar una nueva variable:** RSS $\downarrow$ pero p se incrementa. Así, los cambios en el RSE son inciertos.

Regresión Múltiple

Volvemos al problema, si **añadimos** mas variables al modelo, el R^2 *automáticamente* se incrementa.

Solución: Penalizar el número de variables, pero mediante *p.e.*, R^2 Ajustado:

$$\overline{R}^2 = 1 - \frac{\sum_i \left(y_i - \hat{y}_i\right)^2/(n-k-1)}{\sum_i \left(y_i - \overline{y}\right)^2/(n-1)}$$

Nota: El R^2 ajustado no necesariamente esta entre 0 y 1.

Ok… y entonces?

Regresión Múltiple

R2-Ajustado

Entonces el RSE no es la única forma para penalizar la **adición** de nuevas variables…

R2 ajustado es otra forma clásica de solución.

$$R^2\text{Ajustado} = 1 - \frac{\text{RSS}/(n-k-1)}{\text{TSS}/(n-1)}$$

R2 Ajustado ayuda a penalizar esta adición

- RSS siempre decrece cuando se adhiere una nueva variable.
- $\text{RSS}/(n-k-1)$ podría incrementarse o decrecer con una nueva variable.

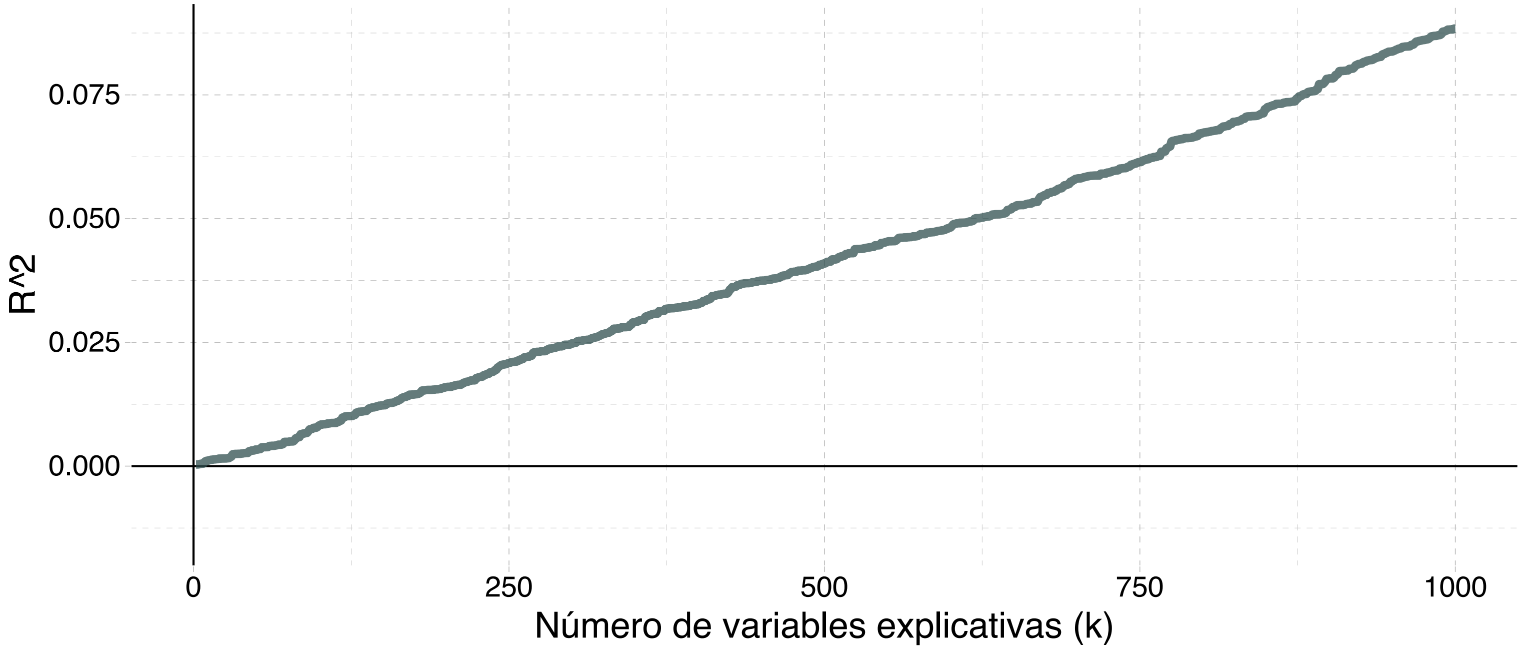
Regresión Múltiple

	(1)	(2)	(3)	(4)
(Intercept)	-0.424	-0.193	-0.207 *	-0.168
	(0.590)	(0.147)	(0.091)	(0.091)
x1	1.253 ***	1.088 ***	1.004 ***	0.985 ***
	(0.317)	(0.079)	(0.049)	(0.049)
x2		4.754 ***	4.933 ***	4.921 ***
		(0.123)	(0.078)	(0.076)
x3			2.083 ***	2.038 ***
			(0.166)	(0.164)
x4				0.083 *
				(0.038)
Observaciones	100	100	100	100
R2	0.138	0.947	0.980	0.981

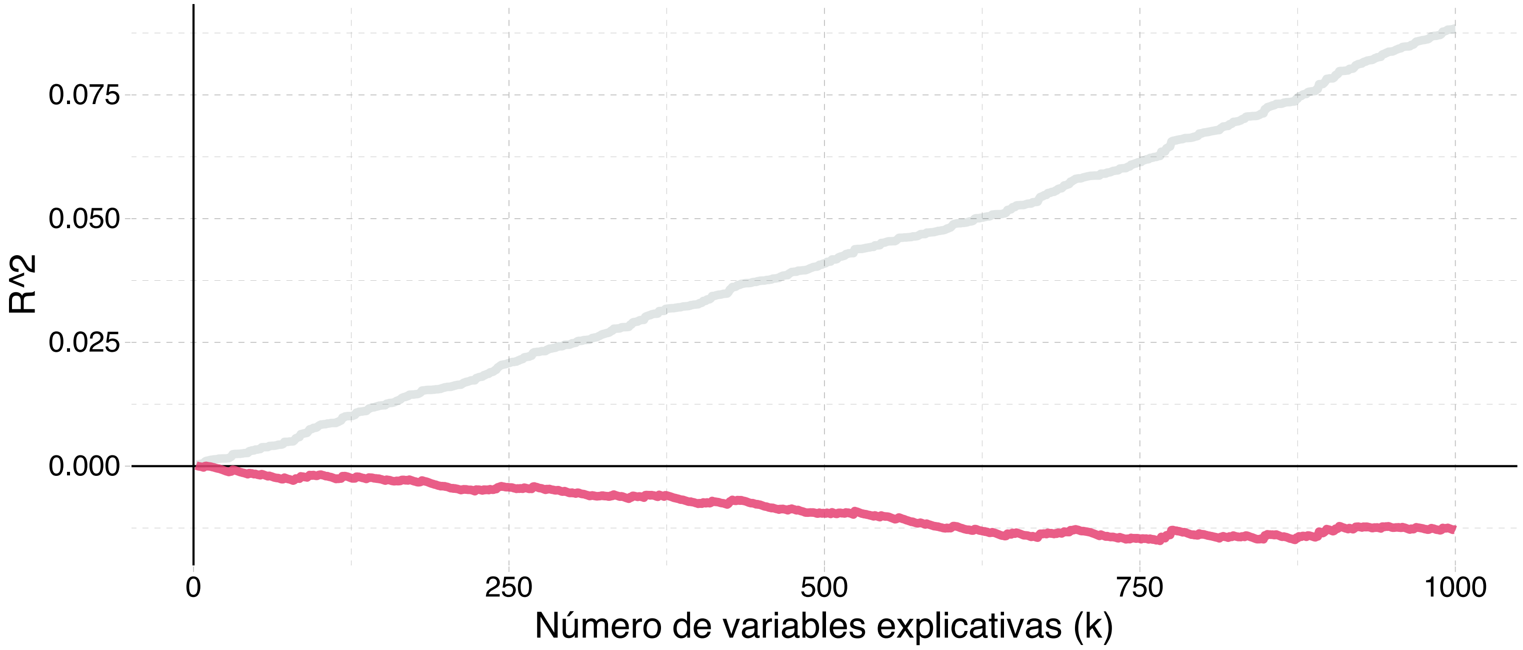
Regresión Múltiple con R2 Ajustado

	(1)	(2)	(3)	(4)
(Intercept)	-0.424	-0.193	-0.207 *	-0.168
	(0.590)	(0.147)	(0.091)	(0.091)
x1	1.253 ***	1.088 ***	1.004 ***	0.985 ***
	(0.317)	(0.079)	(0.049)	(0.049)
x2		4.754 ***	4.933 ***	4.921 ***
		(0.123)	(0.078)	(0.076)
x3			2.083 ***	2.038 ***
			(0.166)	(0.164)
x4				0.083 *
				(0.038)
Observaciones	100	100	100	100
R2 Ajustado	0.129	0.946	0.979	0.980

Regresión Múltiple



Regresión Múltiple

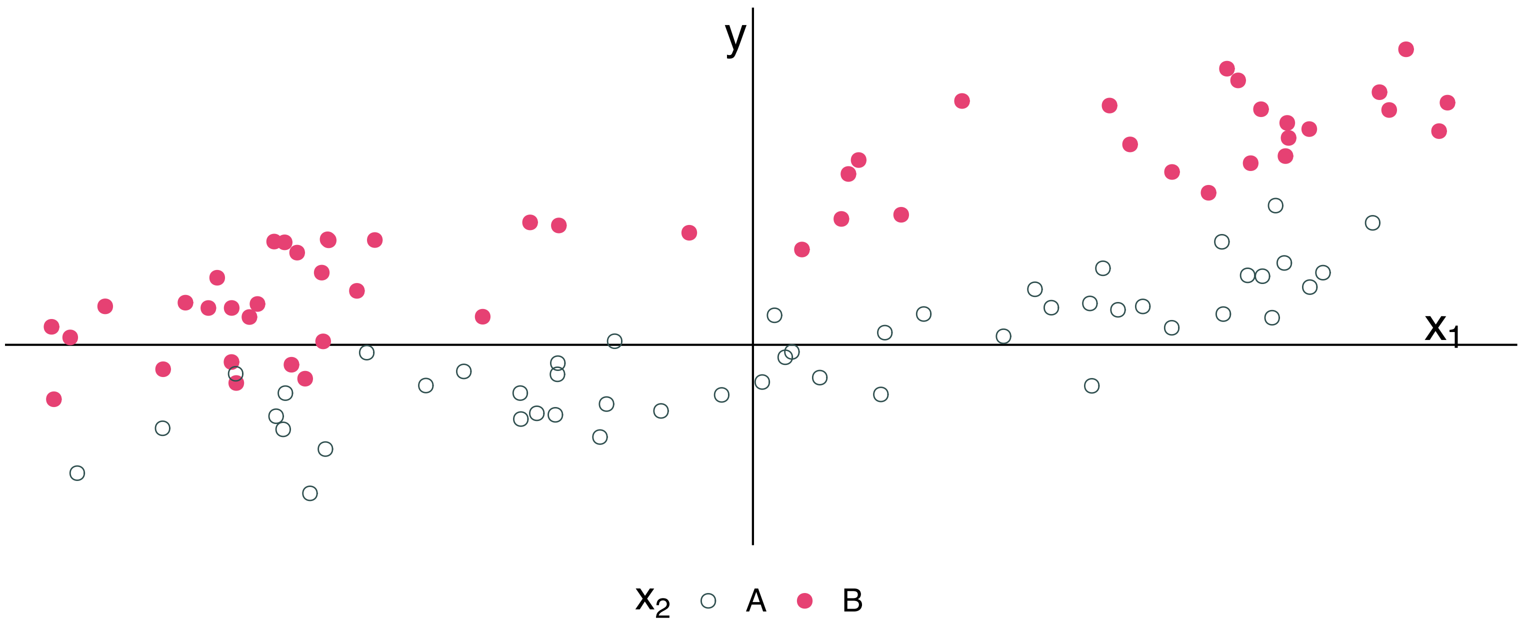


La solución: R^2 Ajustado

Regresión con variables cualitativas y cuantitativas

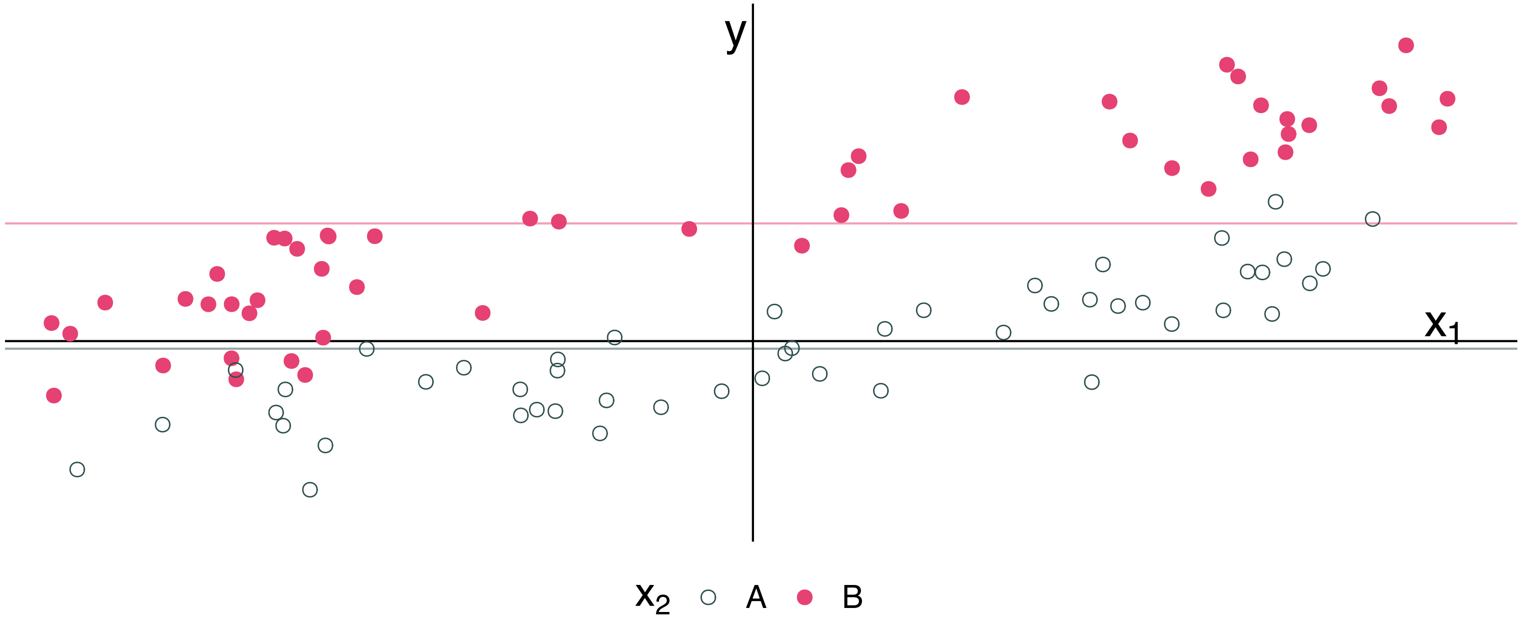
Regresion Múltiple con variable continua y discreta

$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$ donde x_1 es continua y x_2 es categórica



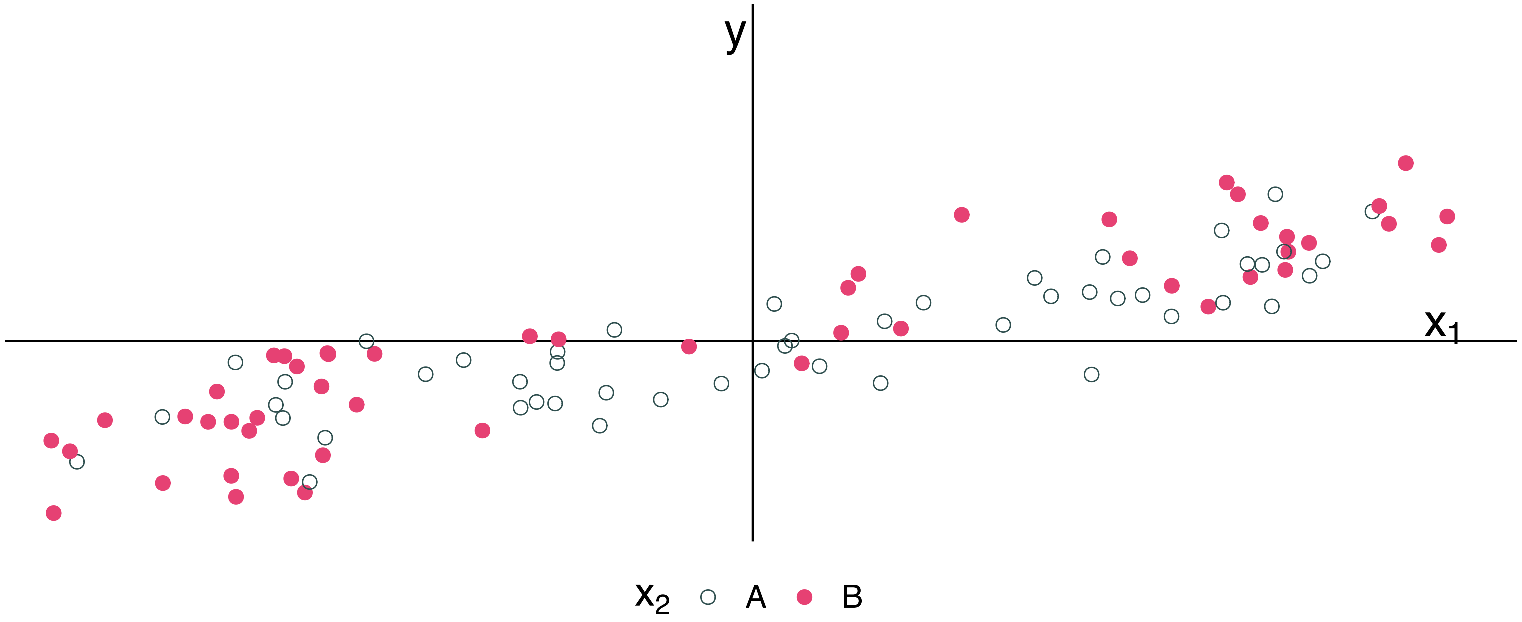
Regresion Múltiple con variable continua y discreta

El intercepto y la variable categórica x_2 controla la media por grupos.



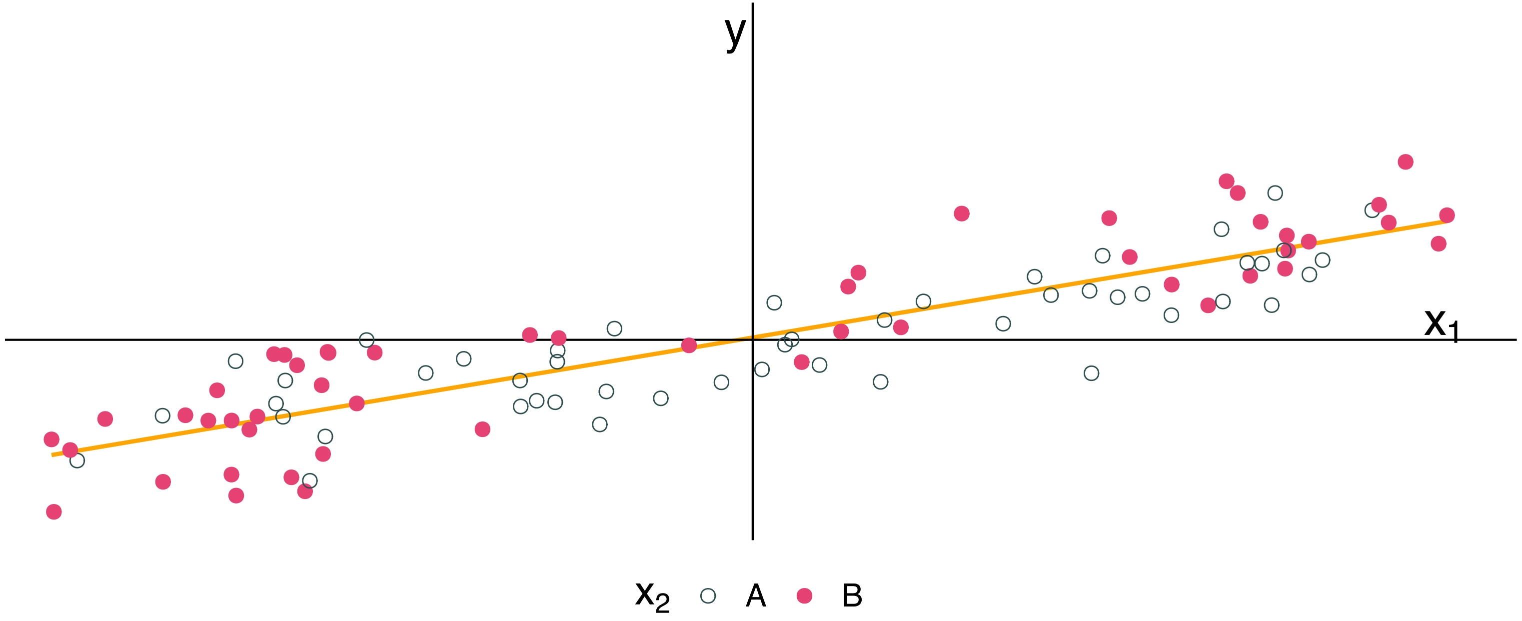
Regresion Múltiple con variable continua y discreta

Cuando es removida x_2



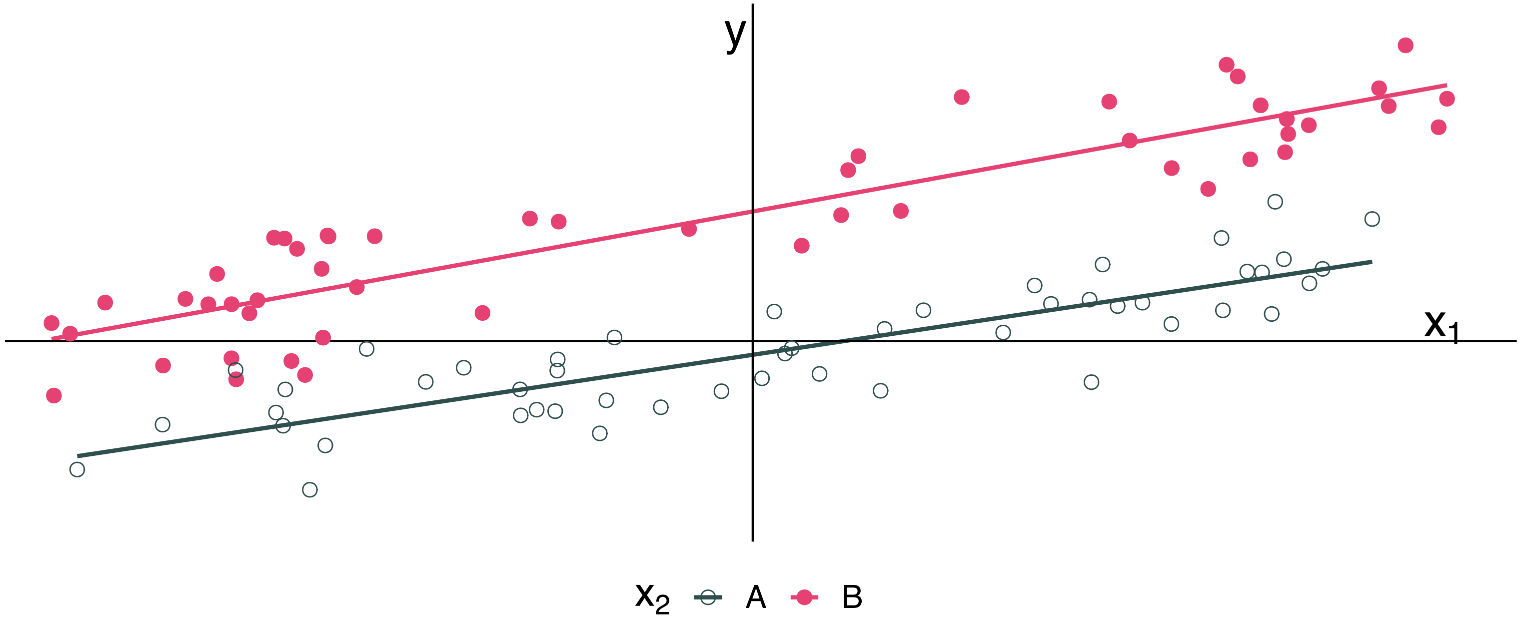
Regresion Múltiple con variable continua y discreta

El parámetro $\hat{\beta}_1$ estima la relación entre y_i y x_1 después de mantener constante a x_2.



Regresion Múltiple

Otra forma de verlo



Regresion Múltiple

Si buscamos un **estimador**

Esto en un modelo de regresión **simple** $y_i = \beta_0 + \beta_1 x_i + u_i$

$$\begin{aligned}
\hat{\beta}_1 = & \\
= & \frac{\sum_i \left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{\sum_i \left(x_i - \overline{x}\right)} \\
= & \frac{\sum_i \left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)/(n-1)}{\sum_i \left(x_i - \overline{x}\right)/(n-1)} \\
= & \frac{\hat{\text{Cov}}(x, y)}{\hat{\text{Var}}(x)}
\end{aligned}$$

Regresion Múltiple

Estimador lineal **simple**

$$\hat{\beta}_1 = \frac{\hat{\text{Cov}}(x, y)}{\hat{\text{Var}}(x)}$$

cuando vamos a la parte de regresión **múltiple**, el estimador **cambia** solo un poco:

$$\hat{\beta}_1 = \frac{\hat{\text{Cov}}(\tilde{x}_1, y)}{\hat{\text{Var}}(\tilde{x}_1)}$$

Donde $\tilde{x}_1$ es el *residuo* de la variable x_1 la variación que queda en x después de controlar las otras variables explicativas.

Regresion Múltiple

Formalmente tenemos nuestro **Modelo** de **Regresión Múltiple** de la siguiente forma:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u_i$$

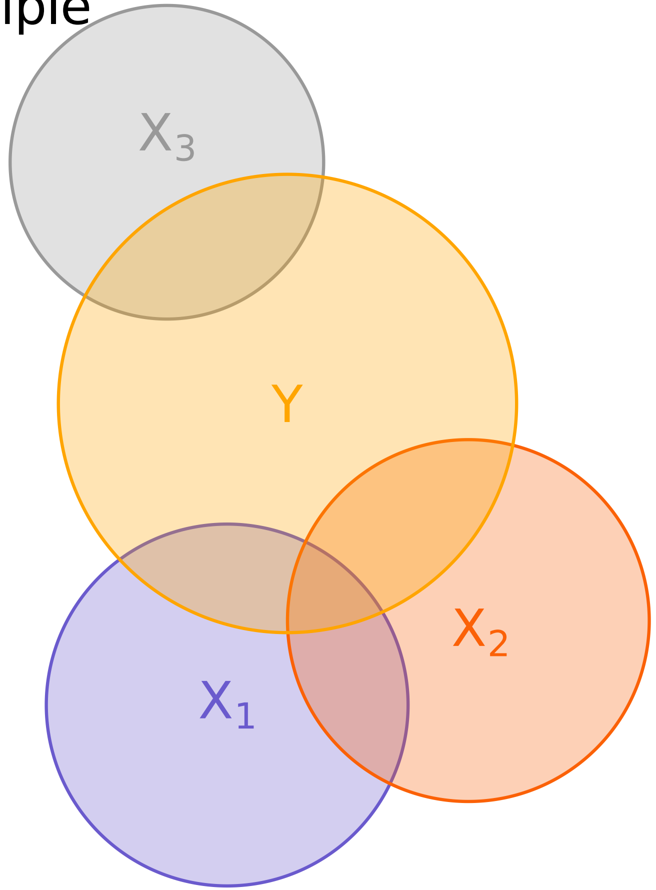
Ya residualizado x_1 (el cual nombramos $\tilde{x}_1$) se obtiene de la regresión de x_1 sobre un intercepto y todas las demás variables explicativas y del final de los residuos, *p.e.*,

$$\hat{x}_{1i} = \hat{\gamma}_0 + \hat{\gamma}_2 \, x_{2i} + \hat{\gamma}_3 \, x_{3i}$$
$$\tilde{x}_{1i} = x_{1i} - \hat{x}_{1i}$$

Lo que nos permite entender mejor el **estimador** de la regresión múltiple

$$\hat{\beta}_1 = \frac{\hat{\text{Cov}}(\tilde{x}_1, \, y)}{\hat{\text{Var}}(\tilde{x}_1)}$$

Regresión Multiple



Otra forma de verlo

Estan las matrices

- Concepto de matrices
- Matriz identidad, nula, vectores
- Operaciones de matrices

Esto tomelo como repaso

Enfoque matricial

Una **matriz** es una colección de números ordenados rectangularmente.

$$[X] = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}_{n*k}$$

La **matriz identidad** es una matriz cuya *diagonal* principal tiene números uno (1).

$$[X] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & 1 & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{n*k}$$

Enfoque matricial

> **Transpuesta** cambiar los elementos de una *fila* por una *columna*

- Se obtiene creando una matriz cuya i-esima **fila** es la misma j-esima **columna**.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \quad A' = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{bmatrix}$$

> **Vectores** estos son lineas de las matrices

- Están los tipo **fila** y los tipo **columna**

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{1k} \\ \cdots & x_{22} & \cdots & \cdots \\ \cdots & x_{32} & \cdots & \cdots \\ \cdots & x_{n2} & \cdots & \cdots \end{bmatrix}$$

Enfoque matricial

Operaciones con matrices: Suma

> **Suma de Matrices** deben tener mismo tamaño y funciona sumando cada uno de los elementos de una matriz con los de la siguiente matriz.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}$$

$$X + A = \begin{bmatrix} x_{11} + a_{11} & x_{12} + a_{12} & \cdots & x_{1k} + a_{1k} \\ x_{21} + a_{21} & x_{22} + a_{22} & \cdots & x_{2k} + a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} + a_{n1} & x_{n2} + a_{n2} & \cdots & x_{nk} + a_{nk} \end{bmatrix}$$

Enfoque matricial

Operaciones con matrices: Multiplicación

Para esto hay que tener en consideración:

- No es necesario que sean cuadradas.
- Deben ser siempre **conformables**.

> Para que una matriz sea *conformable* debe considerarse lo siguiente:

$$A_{m*n} * X_{n*p} = C_{m*p}$$

Debe coincidir o ser de igual tamaño las columnas de la primera matriz con las filas de la siguiente matriz en el orden de la operación.

Enfoque matricial

Operaciones con matrices: Multiplicación

Si tenemos dos vectores $A = (a_1, a_2, \ldots, a_n)$ y $B = (b_1, b_2, \ldots, b_n)$ entonces:

$$a * b = a_1 * b_1 + a_2 * b_2 + \ldots + a_n * b_n$$

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}_{3*4} \quad B = \begin{bmatrix} 3 & 7 & 9 \\ 4 & 9 & 2 \\ 2 & 7 & 1 \\ 4 & 7 & 2 \end{bmatrix}_{4*3} \quad A \times B = \begin{bmatrix} 33 & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}_{3*3}$$

es importante que conozca las propiedades de la *multiplicación*:

- A*I=A
- Ley asociativa $(AB)C = A(BC)$
- Ley distributiva $A(B + C) = AB + AC$
- Ley transpuesta: $(AB)' = B'A'$

De vuelta al estimador ahora matricial de M.C.O

Forma matricial del modelo

☆ Hasta ahora lo que hemos deducido es el estimador de mínimos cuadrados ordinarios MCO para una variable dependiente y una independiente.

☆ Debemos observar ahora como se "deriva" el **estimador** cuando se tiene **más** de una variable **independiente**.

> Debemos recordar que la información que se tiene cuando se estima un **modelo de regresión** tiene la siguiente forma: (Las variables se organizan por columnas).

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{n*1} = \begin{bmatrix} 1 & x_{12} & \cdots & x_{1k} \\ 1 & x_{22} & \vdots & x_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{n2} & \vdots & x_{nk} \end{bmatrix}_{n*k} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}_{k*1} + \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}_{n*1}$$

Forma matricial del modelo

Tenemos por cada observación una **ecuación** que debe ser escrita:

$$
\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \cdots + \beta_k X_{1k} + \mu_1 \\
Y_2 &= \beta_0 + \beta_1 X_{21} + \cdots + \beta_k X_{2k} + \mu_2 \\
\cdots &= \cdots + \cdots + \cdots + \cdots + \cdots \\
Y_n &= \beta_0 + \beta_1 X_{n1} + \cdots + \beta_k X_{nk} + \mu_n
\end{aligned}
$$

Forma matricial del modelo

- Lo cual nos permite tener un sistema así:

$$Y = XB + U$$

$$S = (y - X\beta)'(y - X\beta) = \mu'\mu$$

Donde

$$y'y - y'X\beta - \beta'X'y + \beta'X'X\beta$$

$$\beta'X'y = (\beta'X'y) = y'X\beta$$

$$y'y - 2\beta'X'y + X'X\beta^2$$

Debe derivar B:

$$\frac{\partial S}{\partial \beta} = -2X'y + 2X'X\beta = 0$$

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

Bibliografía

📘 Álvarez, R. A. R., Calvo, J. A. P., Torrado, C. A. M., & Mondragón, J. A. U. (2013). *Fundamentos de econometría intermedia: teoría y aplicaciones*. Universidad de los Andes.

📘 Stock, J. H., Watson, M. W., & Larrión, R. S. (2012). *Introducción a la Econometría*.

📘 Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.

Gracias!

Alguna pregunta adicional?

Carlos Andres Yanes Guerra

cayanes@uninorte.edu.co

keynes37