



## Clase auxiliar # 1

### ¡Recordar!

#### Variables y parámetros:

- Número de observaciones:  $N \in \mathbb{R}$ .
- Número de variables explicativas :  $K \in \mathbb{R}$ .
- Vector de observaciones de la variable dependiente:  $Y = (Y_1, \dots, Y_N) \in \mathbb{R}^{N \times 1}$ .
- Matriz de observaciones de las variables independientes:  $X \in \mathbb{R}^{N \times K}$
- Vector de parámetros poblacionales:  $\beta = (\beta_1, \dots, \beta_N) \in \mathbb{R}^{K \times 1}$
- Vector de errores aleatorios:  $U = (U_1, \dots, U_N) \in \mathbb{R}^{N \times 1}$ . También se suele notar como  $\varepsilon$ .

#### Supuestos de MCO:

1. **Linealidad de parámetros:**  $Y = X\beta + U$ .

Notar que la linealidad es sobre los parámetros, no sobre las variables. Podemos tener perfectamente un modelo de la forma:  $\tanh(Y_i) = \beta_1 \ln(X_{i1}) + \beta_2 \exp(X_{i2}) + \beta_3 X_{i3}^{0,7} + \beta_4 \arctan(X_{i3}) + U_i$ .

2. **Matriz  $X^T X$  invertible:**  $rg(X^T X) = K$ . Se puede verificar.
3. **Errores nulos en esperanza:**  $\mathbb{E}(U) = 0$ . Supuesto aceptable.
4. **Independencia y homocedasticidad:**  $\mathbb{V}(U) = \sigma^2 I_N$ . Se puede testear.
5. **Exogeneidad:**  $\mathbb{E}(X^T U) = 0$ . Supuesto fuerte y problemático. En ciencias sociales siempre hay correlación entre variables que observamos y que no observamos.

#### Formas funcionales:

Forma funcional	Ejemplo de forma funcional	Cambios de Y respecto a X
Lin-lin	$Y_i = \beta_0 + \beta_1 X_i + U_i$	Si X aumenta en una unidad, Y cambia en $\beta_1$ unidades.
Log-log	$\ln(Y_i) = \beta_0 + \beta_1 \ln(X_i) + U_i$	Si X aumenta en un 1%, Y cambia en un 1%
Lin-log	$Y_i = \beta_0 + \beta_1 \ln(X_i) + U_i$	Si X aumenta en un 1%, Y cambia en $\frac{\beta_1}{100}$ unidades.
Log-lin	$\ln(Y_i) = \beta_0 + \beta_1 X_i + U_i$	Si X incrementa en una unidad, Y cambia en un $100 \cdot \beta_1$ %
Cuadrática	$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + U_i$	Si X incrementa en una unidad, Y cambia en $\beta_1 + 2\beta_2 X$ unidades

## 1. Comente

1. **[P1, Control 1, 2017]** Explique si la acumulación de evidencia empírica que no contradice las predicciones de una teoría fortalece a esta última.
2. **[P2, Control 1, 2017]** En un sitio web se afirma que: *Un estudio conjunto de la Universidad de Warwick y Manchester concluye que estar rodeados de amigos mentalmente saludables ayuda a nuestra propia salud mental. Los investigadores explican que las personas que están habitualmente de buen humor a nuestro lado, nos contagian de ese estado de ánimo.*  
Explique qué dificultades podrían haber para probar que salud mental de una persona **es causada** por la de sus amigos.
3. **[P1.c, Control 1, 2016]** Un médico le indica a usted que “la correlación negativa entre número de cigarrillos fumados al día durante el embarazo de una mujer y el peso de su bebé al nacer es evidencia clara del efecto negativo del tabaquismo en la salud del recién nacido. Comente.
4. **[P1.b, Control 1, 2016]** Explique qué son datos de corte transversal, series de tiempo y datos de panel. Dé un ejemplo para cada uno de ellos.

## 2. Formas funcionales (lin-lin, log-log, lin-log, log-lin)

Suponiendo causalidad, demuestre que:

1. En un modelo lin-lin ( $Y_i = \beta_0 + \beta_1 X_i + U_i$ ), si  $X$  aumenta en una unidad,  $Y$  aumenta en  $\beta_1$  unidades.
2. En un modelo log-log ( $\ln Y_i = \beta_0 + \beta_1 \ln X_i + U_i$ ), si  $X$  cambia en un 1 %,  $Y$  cambia en  $\beta_1$  %.
3. En un modelo lin-log ( $Y_i = \beta_0 + \beta_1 \ln X_i + U_i$ ), si  $X$  cambia en un 1 %,  $Y$  cambia en  $\frac{\beta_1}{100}$  unidades.
4. En un modelo log-lin ( $\ln Y_i = \beta_0 + \beta_1 X_i + U_i$ ), si  $X$  cambia en una unidad,  $Y$  cambia en  $100 \cdot \beta_1$  %.
  - a. ¿Qué representan  $\beta_1$  y  $\beta_2$  respectivamente?
  - b. ¿Cómo varía el salario ante un aumento marginal en la escolaridad y la experiencia respectivamente?

## 3. Precios de pasajes aéreos [P1, Control 1, 2015]

Usted tiene un flamante trabajo en la industria de transporte aéreo. Su primera tarea será construir modelos econométricos que expliquen los precios ofrecidos en la industria. Considere que tenemos un modelo para explicar precios efectivos que se cobran por pasajes aéreos de ida y regreso, dentro de Chile en un sitio web. En cada punto, formule un modelo que permita estimar:

1. Efecto porcentual esperado en el precio del día de la semana en que el pasaje se emite y la diferencia porcentual aproximada de dos días de la semana.
2. Efecto porcentual del tiempo de estadía (tiempo entre el viaje de ida y de vuelta), considerando la posible importancia del día de la semana en que se viaje. Por ejemplo: no es lo mismo un viaje viernes-domingo, que uno martes-jueves.

3. La respuesta porcentual del precio a la distancia entre origen y destino, considerando diferencias en dicha respuesta al número de escalas del viaje.
4. Efecto porcentual en el precio del tiempo con anticipación de venta del pasaje, permitiendo un reacción nula para viajes comprados con 6 meses o más de anticipación.

## 4. Estimación computacional de MCO

En este ejercicio analizaremos cómo afecta el gasto per cápita en salud (US\$) a la tasa de mortalidad de menores de 5 años (por cada 1.000) a nivel global. Para lo anterior, ocuparemos datos proporcionados por el Banco Mundial correspondientes al año 2014 y donde cada muestra corresponde a un país. La base de datos se encuentra en U-Cursos bajo el nombre de `mortalidad_gastosalud.xls`.

### 1. **Stata:**

- Importe los datos y realice un gráfico *scatter* (de dispersión) que muestre cómo se comportan conjuntamente ambas variables.
- Regrese la mortalidad en función del gasto per cápita y analice los resultados. Si le parece conveniente, puede realizar transformaciones en las variables con tal de mejorar el ajuste.

### 2. **Matlab:** importe los datos y obtenga el estimador MCO para la misma regresión de la parte anterior. Compare los resultados con los de Stata.

### 3. **Solver de Excel:** proponga un estimador inicial $\hat{\beta}$ y obtenga el estimador MCO para la misma regresión de las partes anteriores. Recordar que queremos minimizar la suma de los residuos al cuadrado. Compare los resultados con los de Stata y Matlab.