# UNIVERSIDAD NACIONAL DE INGENIERIA

Facultad de Ingeniería Económica, Estadistíca y CC.SS.

## Curso: Modelos Lineales

# Modelos Lineal General

Eduardo Marcos Sánchez

# Modelos

- Tenemos el modelo con K variables

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_K X_{Ki} + u_i, \;\; i = 1, \ldots n$$

- Y vimos que usando matrices puede escribirse como

$$Y = X\beta + u$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{K1} \\ 1 & X_{22} & X_{32} & \cdots & X_{K2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{Kn} \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

donde $X$ es $(n \times K)$, $Y$ y $u$ son $(n \times 1)$, y $\beta$ es $(K \times 1)$.

# Definiciones

- Vector de estimadores $\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_K \end{bmatrix}$
- Vector de estimaciones de $Y$ $(n \times 1)$: $\hat{Y} \equiv X\hat{\beta}$
- Vector de residuos o errores de estimación $(n \times 1)$:
  $e \equiv Y - \hat{Y} = Y - X\hat{\beta}$

# Definiciones

- $$X'X = \begin{bmatrix} n & \sum X_{2i} & \sum X_{3i} & \cdots & \sum X_{Ki} \\ \sum X_{2i} & \sum X_{2i}^2 & \sum X_{2i}X_{3i} & \cdots & \sum X_{2i}X_{Ki} \\ \sum X_{3i} & \sum X_{2i}X_{3i} & \sum X_{3i}^2 & \cdots & \sum X_{3i}X_{Ki} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum X_{Ki} & \sum X_{2i}X_{Ki} & \sum X_{3i}X_{Ki} & \cdots & \sum X_{Ki}^2 \end{bmatrix}$$
- La existencia de $(X'X)^{-1}$ se garantiza por el supuesto de no multicolinealidad perfecta.
- Recordemos: si no hay multicolinealidad perfecta $\rho(X) = K$, lo que a su vez implica que $\rho(X'X) = K \Rightarrow \exists (X'X)^{-1}$

# Supuestos

1. $E(u) = \begin{bmatrix} E(u_1) \\ E(u_2) \\ \vdots \\ E(u_n) \end{bmatrix} = \mathbf{0}$

2. $VarCov(u) = \begin{bmatrix} V[u_1] & Cov[u_1, u_2] & \cdots & Cov[u_1, u_n] \\ Cov[u_2, u_1] & V[u_2] & \cdots & Cov[u_2, u_n] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[u_n, u_1] & Cov[u_n, u_2] & \cdots & V[u_n] \end{bmatrix} = \sigma^2 \mathbf{I_n}$

3. $X_{nxK}$ no estocástica con $\rho(X) = K$

# Supuestos

- Insesgadez: $E[\hat{\beta}] = \beta$
- Matriz de varianzas y covarianzas: $VarCov(\hat{\beta}) = \sigma^2(X'X)^{-1}$

$$= \begin{bmatrix} V[\hat{\beta}_1] & Cov[\hat{\beta}_1, \hat{\beta}_2] & \cdots & Cov[\hat{\beta}_1, \hat{\beta}_K] \\ Cov[\hat{\beta}_1, \hat{\beta}_2] & V[\hat{\beta}_2] & \cdots & Cov[\hat{\beta}_2, \hat{\beta}_K] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[\hat{\beta}_1, \hat{\beta}_K] & Cov[\hat{\beta}_2, \hat{\beta}_K] & \cdots & V[\hat{\beta}_K] \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^2 A_{11} & \sigma^2 A_{12} & \cdots & \sigma^2 A_{1K} \\ \sigma^2 A_{21} & \sigma^2 A_{22} & \cdots & \sigma^2 A_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2 A_{K1} & \sigma^2 A_{K2} & \cdots & \sigma^2 A_{KK} \end{bmatrix}$$

donde $A_{kk}$ es el elemento en la fila $k$ y columna $k$ de la matriz $(X'X)^{-1}$

- El estimador de $VarCov(\hat{\beta})$ es $Var\hat{C}ov(\hat{\beta}) = S^2(X'X)^{-1}$ con $S^2 = \frac{e'e}{n-K}$

# Supuestos

**Definición:** matriz semidefinida positiva

Sea $H$ una matriz $m \times m$ y sea $c$ cualquier vector de $m$ constantes. Se dice que $H$ es semidefinida positiva si y sólo si $c'Hc \geq 0 \ \forall c \in \mathbb{R}^m$

- Notar:
  - que $c'Hc$ es $(1 \times 1)$, o sea, un número real.
  - $c'Hc$ es una **forma cuadrática.** Para ver por qué, desarrollar la expresión para $m = 2$.
- **Importante**: cualquier matriz de varianzas y covarianzas es semidefinida positiva.
  - En particular, $VarCov[\hat{\beta}]$ es semidefinida positiva.

# Supuestos

- Entonces, para cualquier vector de constantes $c$ de dimensión $K \times 1$ se cumple que $c' VarCov[\hat{\beta}]c \geq 0$.
- La expresión $c' VarCov[\hat{\beta}]c$ es la varianza de la combinación lineal $c'\hat{\beta}$.
- Veamos el ejempo para $K = 2$:
  - $c'\hat{\beta} = c_1\hat{\beta}_1 + c_2\hat{\beta}_2$
  - $V[c'\hat{\beta}] = c' VarCov[\hat{\beta}]c = c_1^2 V[\hat{\beta}_1] + c_2^2 V[\hat{\beta}_2] + 2c_1c_2 Cov[\hat{\beta}_1, \hat{\beta}_2]$
  - La semidefinición positiva garantiza que esta última expresión es $\geq 0$.
  - Entonces todos los siguientes casos particulares son $\geq 0$:
    - Si $c_1 = 1$ y $c_2 = 0 \Rightarrow V[c'\hat{\beta}] = V[\hat{\beta}_1] \geq 0$
    - Si $c_1 = 0$ y $c_2 = 1 \Rightarrow V[c'\hat{\beta}] = V[\hat{\beta}_2] \geq 0$
    - Si $c_1 = 1$ y $c_2 = 1 \Rightarrow V[c'\hat{\beta}] = V[\hat{\beta}_1] + V[\hat{\beta}_2] + 2Cov[\hat{\beta}_1, \hat{\beta}_2] \geq 0$
    - Si $c_1 = 1$ y $c_2 = -1 \Rightarrow V[c'\hat{\beta}] = V[\hat{\beta}_1] + V[\hat{\beta}_2] - 2Cov[\hat{\beta}_1, \hat{\beta}_2] \geq 0$

# Teorema de Gauss Markov

- Bajo todos los supuestos clásicos, el estimador de MCO es el más eficiente de todos los estimadores lineales e insesgados.
- **Mejor Estimador Lineal e Insesgado (MELI).**
- **Además**: sea $c$ un vector de $K$ constantes arbitrarias $c_1, c_2, ...c_K$, $c'\hat{\beta}$ es el mejor estimador lineal e insesgado de $c'\beta$.
  - Es decir, la combinación lineal de los estimadores MCO es MELI para estimar la combinación lineal de los parámetros.