MODELO DE REGRESIÓN CON DOS VARIABLES: PROBLEMA DE ESTIMACIÓN

El objeto de la econometría es:

1. Especificar un modelo de relación entre variables económicas **(TEORÍA ECONÓMICA)**
2. Utilizar información muestral acerca de los valores tomados por dichas variables, con el objeto de cuantificar la magnitud de la dependencia entre ellas **(ESTIMACIÓN)**
3. Evaluar críticamente la validez de las hipótesis propuestas por la teoría económica acerca de las relaciones estimadas **(INFERENCIA ESTADÍSTICA)**
4. Efectuar un ejercicio de seguimiento coyuntural de previsión de las variables analizadas **(POLÍTICA ECONÓMICA)**

La tarea fundamental durante el curso consiste en estimar la FRP con base en la FRM, en definitiva se estimará la FRM para aproximarse a la FRP. Se traduce ésto en estimar los $\boldsymbol{\hat{\beta}'s}$ con los datos a disposición para llegar a los $\boldsymbol{\beta's}$ poblacionales utilizando la inferencia estadística.

La condición básica de los métodos de estimación de los $\hat{\beta}'s$ es que garanticen que éstos sean lo más parecido posible a los parámetros poblacionales.

MÉTODOS DE ESTIMACIÓN MÁS COMUNES:

- Mínimos Cuadrados Ordinarios (MCO)
- Máxima Verosimilitud (MV)

1) MÉTODO DE MÍNIMOS CUADRADOS ORDINARIOS (MCO):

Atribuido a Carl F. Gauss, bajo ciertos supuestos cumple con propiedades estadísticas atractivas que lo han convertido en uno de los más eficientes y populares.

$$FRP \qquad Y_i = \beta_1 + \beta_2 X + \varepsilon_i$$

$$FRM \qquad Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{\varepsilon}_i$$

$$Y_i = \hat{Y}_i + \hat{\varepsilon}_i \Longrightarrow \hat{\varepsilon}_i = Y_i - \hat{Y}_i$$

$$\hat{\varepsilon}_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$$

¿QUÉ ES LO QUE SE BUSCA?

R: HACER LOS RESIDUOS O PERTURBACIONES TAN PEQUEÑOS COMO SEA POSIBLE

En realidad lo que se busca es seleccionar la FRM de tal manera que $\sum(\mathbf{Y_i} - \hat{\mathbf{Y}}_\mathbf{i})$ sea la menor posible. Sin embargo, el criterio anterior aunque es muy atractivo no es bueno porque la suma de los residuos podría llegar a ser cero porque todos reciben el mismo peso.

¿QUÉ HACER ENTONCES?
R: MINIMIZAR LA SUMA DE LOS RESIDUOS AL CUADRADO

$$MINIMIZAR \qquad \sum \hat{\varepsilon}_i^{\,2} = \sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$

Este método da mayor importancia a los residuos más grandes y menos a los más pequeños garantizando que $\sum \hat{\varepsilon}_i^{\,2} \neq 0$ (mínima).

LOS ESTIMADORES OBTENIDOS CON ESTE MÉTODO TIENEN PROPIEDADES ESTADÍSTICAS MUY DESEABLES

De la última ecuación se deriva $\sum \hat{\varepsilon}_i^{\,2} = f(\beta_1, \beta_2)$

MCO escoge $\hat{\beta}_1$ y $\hat{\beta}_2$ de tal manera que para una muestra dada o un conjunto de datos $\sum \hat{\varepsilon}_i{}^2$ sea la más pequeña posible. Proporciona valores estimados únicos de los $\hat{\beta}'s$ de tal forma que la sumatoria de los residuos al cuadrado sea mínima.

Procedimiento:

$\sum \hat{\varepsilon}_i{}^2 = \sum(Y_i - \hat{Y}_i)^2 = \sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$ $\longrightarrow$ Función objetivo

Aplicando la metodología tradicional de optimización para minimizar la función objetivo, se toma la primera derivada parcial respecto a $\hat{\beta}_1$ y $\hat{\beta}_2$ y se igualan a cero.

$$\frac{\partial(\sum \hat{\varepsilon}_i{}^2)}{\partial \hat{\beta}_1} = -2\sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = -2\sum \hat{\varepsilon}_i = 0 \qquad (1)$$

$$\sum Y_i - n\hat{\beta}_1 - \hat{\beta}_2 \sum X_i = 0$$

Ecuaciones normales

$$\frac{\partial(\sum \hat{\varepsilon}_i{}^2)}{\partial \hat{\beta}_2} = -2\sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)(X_i) = -2\sum \hat{\varepsilon}_i X_i = 0 \qquad (2)$$

$$\sum Y_i X_i - \hat{\beta}_1 \sum X_i - \hat{\beta}_2 \sum X_i{}^2 = 0$$

Resolviendo el sistema de ecuaciones:

$$\hat{\beta}_1 = \frac{\sum {X_i}^2 \sum Y_i - \sum X_i \sum X_i Y_i}{n \sum {X_i}^2 - \left(\sum X_i\right)^2} = \bar{Y} - \hat{\beta}_2 \bar{X}$$

$$\hat{\beta}_2 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum {X_i}^2 - \left(\sum X_i\right)^2} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \frac{\sum x_i y_i}{\sum {x_i}^2}$$

Demostración de la simplificación de $\hat{\beta}_2$:

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum {x_i}^2} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \frac{\sum X_i Y_i - \sum X_i \bar{Y} - \sum Y_i \bar{X} + n \bar{X}\,\bar{Y}}{\sum (X_i^2 - 2 X_i \bar{X} + \bar{X}^2)}$$

Se multiplica y se divide por n:

$$\hat{\beta}_2 = \frac{\sum X_i Y_i - n\bar{X}\bar{Y} - n\bar{Y}\bar{X} + n\bar{X}\bar{Y}}{\sum X_i^2 - 2\bar{X}\sum X_i + n\bar{X}^2} = \frac{\sum X_i Y_i - n\bar{X}\bar{Y}}{\sum X_i^2 - 2n\bar{X}^2 + n\bar{X}^2} = \frac{\sum X_i Y_i - n\frac{\sum X_i \sum Y_i}{n \quad n}}{\sum X_i^2 - n\left(\frac{\sum X_i}{n}\right)^2} = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sum X_i^2 - n\left(\frac{\sum X_i}{n}\right)^2} =$$

$$\frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum {X_i}^2 - \left(\sum X_i\right)^2}$$

2) MÉTODO DE MÁXIMA VEROSIMILITUD (MV):

Partiendo del modelo inicial:

$$Y_i = \beta_1 + \beta_2 X + \varepsilon_i$$

Asumiendo $\quad Y_i \sim N(\beta_1 + \beta_2 X;\ \sigma^2)$

La función de densidad de probabilidad conjunta de $Y_1, Y_2, \dots, Y_n$ puede escribirse como:

$$f(Y_1, Y_2, \dots, Y_n | \beta_1 + \beta_2 X;\ \sigma^2) \quad (1)$$

Por la independencia de las Y, la función (1) puede escribirse como:

$$f(Y_1 | \beta_1 + \beta_2 X;\ \sigma^2) f(Y_2 | \beta_1 + \beta_2 X;\ \sigma^2) \dots f(Y_n | \beta_1 + \beta_2 X;\ \sigma^2)$$

Recuerde que la función de densidad de una variable que se distribuye normal viene dada por:

$$f(Y_i) = \frac{1}{\sigma\sqrt{2\pi}} exp\left\{-\frac{1}{2}\frac{(Y_i - \varepsilon)^2}{\sigma^2}\right\} \quad (2)$$

Sustituyendo (2) por cada Y_i en (1):

$$f\left(Y_1, Y_2, \ldots, Y_n \mid \beta_1 + \beta_2 X;\ \sigma^2\right) = \frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n} exp\left\{-\sum \frac{1}{2}\frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2}\right\} \quad (3)$$

La función (3) se llama ***función de verosimilitud*** denotada por:

$$FV\left(\beta_1, \beta_2, \sigma^2\right) = \frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n} exp\left\{-\frac{1}{2}\sum \frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2}\right\} \quad (4)$$

El método de MV consiste en estimar los parámetros desconocidos de tal forma que la probabilidad de observar las Y dadas sea lo más alta posible, por ende se busca maximizar la FV. Para ello se expresa FV en logaritmo:

$$lnFV = -\frac{n}{2} ln\sigma^2 - \frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum \frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2} \quad (5)$$

Diferenciando (5) parcialmente respecto a $\beta_1, \beta_2\ y\ \sigma^2$ se obtiene:

$$\frac{\partial lnFV}{\partial \beta_1} = -\frac{1}{\sigma^2}\sum (Y_i - \beta_1 - \beta_2 X_i)(-1) \quad (6)$$

$$\frac{\partial lnFV}{\partial \beta_2} = -\frac{1}{\sigma^2}\sum(Y_i - \beta_1 - \beta_2 X_i)(-X_i) \qquad (7)$$

$$\frac{\partial lnFV}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum(Y_i - \beta_1 - \beta_2 X_i)^2 \qquad (8)$$

Igualando a cero por CPO de optimización:

$$\frac{1}{\tilde{\sigma}^2}\sum(Y_i - \widetilde{\beta_1} - \widetilde{\beta_2} X_i) = 0 \qquad (9)$$

$$\frac{1}{\tilde{\sigma}^2}\sum(Y_i - \widetilde{\beta_1} - \widetilde{\beta_2} X_i)\,(X_i) = 0 \qquad (10)$$

$$-\frac{n}{2\tilde{\sigma}^2} + \frac{1}{2\tilde{\sigma}^4}\sum(Y_i - \widetilde{\beta_1} - \widetilde{\beta_2} X_i)^2 = 0 \qquad (11)$$

Simplificando 9 y 10, se obtienen las mismas ecuaciones normales que se obtuvieron mediante los estimadores MCO:

$$\sum Y_i - n\hat{\beta}_1 - \hat{\beta}_2 \sum X_i = 0 \qquad (12)$$

$$\sum Y_i X_i - \hat{\beta}_1 \sum X_i - \hat{\beta}_2 \sum X_i{}^2 = 0 \qquad (13)$$

BAJO EL SUPUESTO DE NORMALIDAD LOS ESTIMADORES MCO Y MV SON IGUALES.

ESTIMADORES MÍNIMOS CUADRADOS ORDINARIOS DE UN MODELO DE REGRESIÓN SIMPLE MEDIANTE EL ENFOQUE MATRICIAL

$$Y_i = \beta_1 + \beta_2 X_{i2} + \varepsilon_i$$

$$\underset{n*1}{\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}} = \underset{n*k}{\begin{bmatrix} 1 & X_{12} \\ 1 & X_{22} \\ \vdots & \vdots \\ 1 & X_{n2} \end{bmatrix}} \underset{k*1}{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}} + \underset{n*1}{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

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

<u>Recuerde:</u>

$$\mathrm{E}\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} E(\varepsilon_1) \\ E(\varepsilon_2) \\ \vdots \\ E(\varepsilon_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$E(\varepsilon\varepsilon') = E\begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_n\end{bmatrix}\begin{bmatrix}\varepsilon_1 & \varepsilon_2 & \dots & \varepsilon_n\end{bmatrix}$$

$$E(\varepsilon\varepsilon') = E\begin{bmatrix}\varepsilon_1{}^2 & \varepsilon_1\varepsilon_2 & \dots & \varepsilon_1\varepsilon_n\\ \varepsilon_2\varepsilon_1 & \varepsilon_2{}^2 & \dots & \varepsilon_2\varepsilon_n\\ \vdots & \vdots & \ddots & \vdots\\ \varepsilon_n\varepsilon_1 & \varepsilon_n\varepsilon_2 & \dots & \varepsilon_n{}^2\end{bmatrix}$$

$$E(\varepsilon\varepsilon') = \begin{bmatrix}E(\varepsilon_1{}^2) & E(\varepsilon_1\varepsilon_2) & \dots & E(\varepsilon_1\varepsilon_n)\\ E(\varepsilon_2\varepsilon_1) & E(\varepsilon_2{}^2) & \dots & E(\varepsilon_2\varepsilon_n)\\ \vdots & \vdots & \ddots & \vdots\\ E(\varepsilon_n\varepsilon_1) & E(\varepsilon_n\varepsilon_2) & \dots & E(\varepsilon_n{}^2)\end{bmatrix}$$

$$E(\varepsilon\varepsilon') = \begin{bmatrix}\sigma^2 & 0 & \dots & 0\\ 0 & \sigma^2 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sigma^2\end{bmatrix}$$

$$E(\varepsilon\varepsilon') = \sigma^2\begin{bmatrix}1 & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 1\end{bmatrix}$$

$E(\varepsilon\varepsilon') = \sigma^2 \mathrm{I}.$ Varianza homoscedástica

Para el modelo a estimar:

$$\boldsymbol{Y_i = \widehat{\beta}_1 + \widehat{\beta}_2 X_{i2}} + \hat{\varepsilon}_{\boldsymbol{i}}$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{12} \\ 1 & X_{22} \\ \vdots & \vdots \\ 1 & X_{n2} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} + \begin{bmatrix} \hat{\varepsilon}_1 \\ \hat{\varepsilon}_2 \\ \vdots \\ \hat{\varepsilon}_n \end{bmatrix}$$

n*1 n*k k*1 n*1

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

Lo que se busca es minimizar:

$$\sum \hat{\varepsilon}_{\boldsymbol{i}}{}^{\boldsymbol{2}} = \sum \left(\boldsymbol{Y_i - \widehat{\beta}_1 - \widehat{\beta}_2 X_{2i}}\right)^{\boldsymbol{2}} = \hat{\varepsilon}'\hat{\varepsilon}$$

$$\hat{\varepsilon} = \boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}$$

$$\hat{\varepsilon}'\hat{\varepsilon} = \left(y - X\hat{\beta}\right)'\left(y - X\hat{\beta}\right)$$

$$= y'y - 2\widehat{\beta}'X'y + \widehat{\beta}'X'X\hat{\beta}$$

$$\frac{\partial(\hat{\varepsilon}'\hat{\varepsilon})}{\partial\widehat{\beta}} = -2X'y + 2(X'X)\hat{\beta} = 0 \implies (X'X)\hat{\beta} = X'y$$

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

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

k*1 k*k (k*n)(n*1)

PROPIEDADES NÚMERICAS DE LOS ESTIMADORES MCO:

"Son aquellas que se mantienen como consecuencia del uso de mínimos cuadrados ordinarios sin considerar la forma como se generaron los datos" Davison y MacKinnon (1993).

1) Los estimadores MCO están expresados únicamente en términos de cantidades observables. Son fáciles de calcular.

2) Son estimadores puntuales. Proporciona un valor único del parámetro poblacional relevante.

3) Una vez obtenidos de la información muestral, la recta de regresión muestral puede ser obtenida fácilmente y tiene las siguientes propiedades:

- Pasa a través de las medias muestrales de Y y de X.

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X} \implies \bar{Y} = \hat{\beta}_1 + \hat{\beta}_2\bar{X}$$

- El valor promedio de Y estimado es igual al valor promedio de Y observado:

$$
\begin{aligned}
\hat{Y}_i &= \hat{\beta}_1 + \hat{\beta}_2 X_i \\
&= \left(\bar{Y} - \hat{\beta}_2 \bar{X}\right) + \hat{\beta}_2 X_i \\
&= \bar{Y} + \hat{\beta}_2 (X_i - \bar{X}) \\
\bar{\hat{Y}}_i &= \bar{Y}
\end{aligned}
$$

- El valor medio de los residuos estimados es cero. $\sum \hat{\varepsilon}_i = 0$

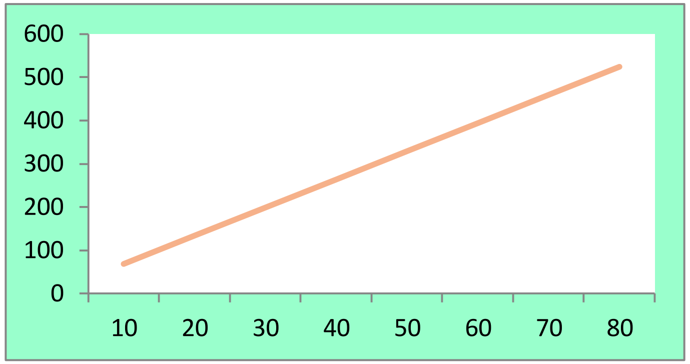
- Los $\hat{u}_i$ no están correlacionados con los valores predichos de Y_i. $\sum \hat{y}_i \hat{\varepsilon}_i = 0$

- Los $\hat{u}_i$ no están correlacionados con los valores fijos de X_i. $\sum X_i \hat{\varepsilon}_i = 0$

SUPUESTOS BÁSICOS DEL MODELO CLÁSICO DE REGRESIÓN LINEAL (MCRL)[1]

1) *El modelo de regresión es lineal en parámetros*[2].

$$Y_i = \beta_1 + \beta_2 X + \varepsilon_i$$



2) *Los valores de las X son fijos en repetidas muestras.* Técnicamente X se supone no estocástica[3]. Por tal razón el análisis de regresión es un análisis de regresión condicional. **CONDICIONADO A LOS VALORES DEL (LOS) REGRESOR(ES).**

[1] Raíz de la mayor parte de la teoría econométrica, es un modelo clásico formulado por Gauss en 1821. Sin embargo existen otros modelos de regresión que no cumplen los supuestos gaussianos.
[2] Existen otros modelos que son intrínsecamente lineales los cuales se verán más adelante en el capítulo 6 y 7.
[3] En algunos casos X puede ser aleatoria, sin embargo el proceso mediante el cual se generan los datos de X no guarda relación alguna con u.

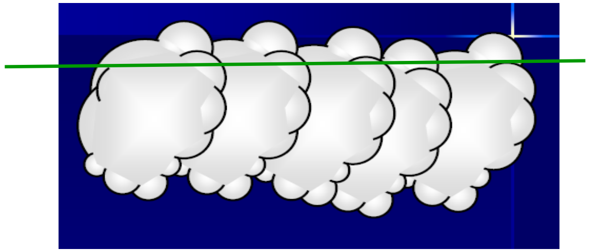
3) ***El valor medio de las perturbaciones u_i es cero.***

$$E(\varepsilon_i|X_i) = 0$$
$$\Downarrow$$
$$E(Y_i|X_i) = \beta_1 + \beta_2 X_i$$

4) Homoscedásticidad o igual varianza de u_i. Dado un valor de X, la varianza de u_i es la misma para todas las observaciones. Lo que significa que las varianza condicionales de u_i son idénticas:

$$Var(\varepsilon_i|X_i) = E[\varepsilon_i - E(\varepsilon_i)|X_i]^2$$

$$= E[(\varepsilon_i{}^2)|X_i]$$
$$= \sigma^2$$



Implica esto también que las varianzas condicionales de Y_i también son homoscedásticas[4]:

$$Var(Y_i|X_i) = \sigma^2$$

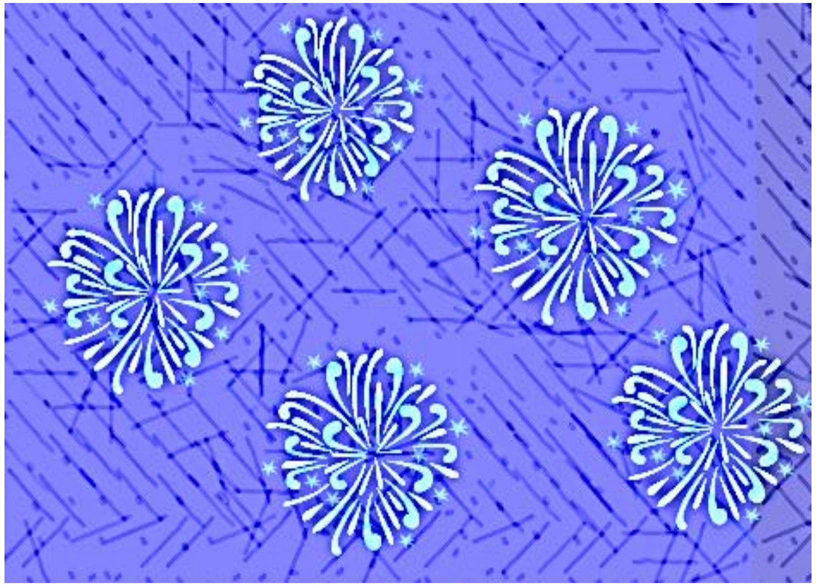
[4] La varianza incondicional de Y es $\sigma_y{}^2$. Ver Apéndice A de D. Gujarati 4ta edición.

5) ***No existe autocorrelación entre las perturbaciones (INDEPENDENCIA).*** No hay correlacion serial entre residuos. Si las perturbaciones siguen patrones sistemáticos, hay **CORRELACIÓN SERIAL O AUTOCORRELACION NO SE QUIERE.**

$$Cov(\varepsilon_i, \varepsilon_j | X_i, X_j) = E\{[\varepsilon_i - E(\varepsilon_i)]|X_i\}\{[\varepsilon_j - E(\varepsilon_j)]|X_j\}$$

$$= E(\varepsilon_i|X_i)(\varepsilon_j|X_j)$$

$$= 0$$



LOS SUPUESTOS 4 Y 5 SE DENOMINAN PERTURBACIONES ESFÉRICAS

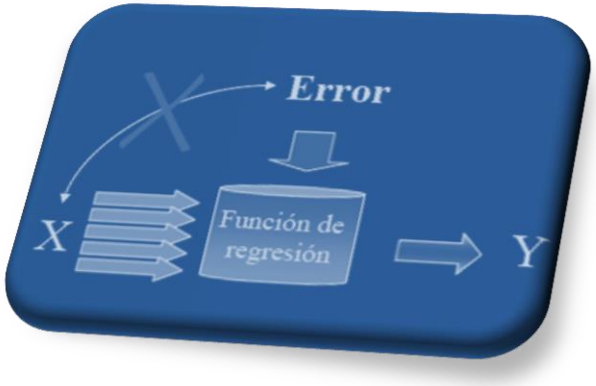
6) ***La covarianza entre u_i y X_i es cero.*** Tanto X como todos los factores no incluidos en el modelo representado por los residuos tienen efectos separados sobre Y.

$$Cov(\varepsilon_i, X_i) = E[\varepsilon_i - E(\varepsilon_i)][X_i - E(X_i)]$$

$$= E[\varepsilon_i(X_i - E(X_i)]$$

$$= E(\varepsilon_i X_i) - E(X_i)E(\varepsilon_i)$$

$$= E(\varepsilon_i X_i) \; = \; 0$$



7) ***El número de observaciones n debe ser mayor que el número de parámetros a estimar k.*** Conocido también como el supuesto de identificabilidad

n>k

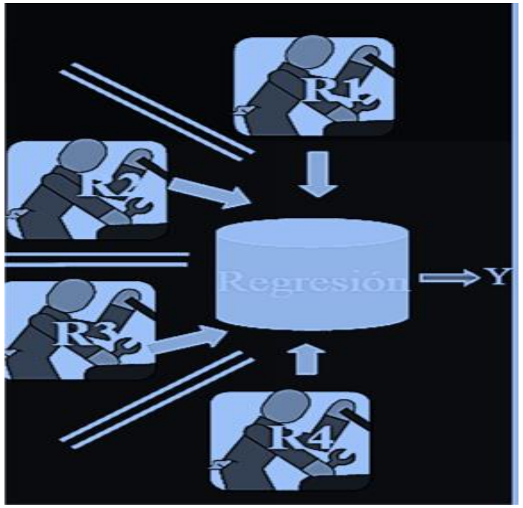
8) ***Variabilidad en los valores de X.*** Los valores que tome la variable X no deben ser iguales en una muestra dada.

$$Var(X) \neq 0$$

9) El modelo de regresión está correctamente especificado. No debe existir sesgo de especificación o error de especificación.

Variables + Forma funcional + Datos Adecuados

10) No hay multicolinealidad perfecta entre variables explicativas.



LOS SUPUESTOS SE RELACIONAN CON LA FRP Y NO CON LA FRM

PRECISIÓN DE LOS ESTIMADORES MCO

No se olvide que los MCO son función de datos muestrales. Sin embargo, es posible que los datos cambien entre distintas muestras por ello se requiere de una ***medida de precisión o de confiabilidad***. En estadística la precisión de un valor estimado se mide a través de su error estándar (desviación estándar)

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2} \qquad ee(\hat{\beta}_2) = \frac{\sigma}{\sqrt{\sum x_i^2}}$$

$$Var(\hat{\beta}_1) = \frac{\sum X_i^2}{n\sum x_i^2}\sigma^2 \qquad ee(\hat{\beta}_1) = \sqrt{\frac{\sum X_i^2}{n\sum x_i^2}}\sigma$$

$$\hat{\sigma}^2 = \frac{\sum \hat{\varepsilon}_i^2}{n-k} \qquad \hat{\sigma} = \sqrt{\frac{\sum \hat{\varepsilon}_i^2}{n-k}}$$

Dado que $\hat{\beta}_1$ y $\hat{\beta}_2$ son ESTIMADORES, éstos no sólo varían de una muestra a otra, sino que también en una muestra dada es probable que dependan entre sí, por tanto:

$$Cov(\hat{\beta}_1, \hat{\beta}_2) = -\bar{X}Var(\hat{\beta}_2)$$

POR LO ANTERIOR SI EL COEFICIENTE β_2 ESTÁ SOBREESTIMADO EL COEFICIENTE β_1 ESTARÁ SUBESTIMADO

PROPIEDADES ESTADISTICAS[5] DE LOS ESTIMADORES MCO

1) *Linealidad:* el estimador es función lineal de la variable aleatoria Y.

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X} \qquad\qquad \hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$$

Demostración 1: Se dice que un estimador $\hat{\theta}$ se define como un estimador lineal de θ, si puede definirse como una función lineal de las observaciones muestrales. Ejemplo:

La media muestral $\bar{X}$, es un estimador lineal porque es una función lineal de los valores muestrales de X.

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}(X_1 + X_2 + X_3 + \cdots + X_n)$$

Asumiendo que $y_i = \beta_2 x_i + \varepsilon_i$[6]. Recuerde que $\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \sum w_i y_i$ donde $\sum w_i = \frac{\sum x_i}{\sum x_i^2}$

Observe que $\hat{\beta}_2 = \sum w_i y_i$, que es lineal respecto a y_i

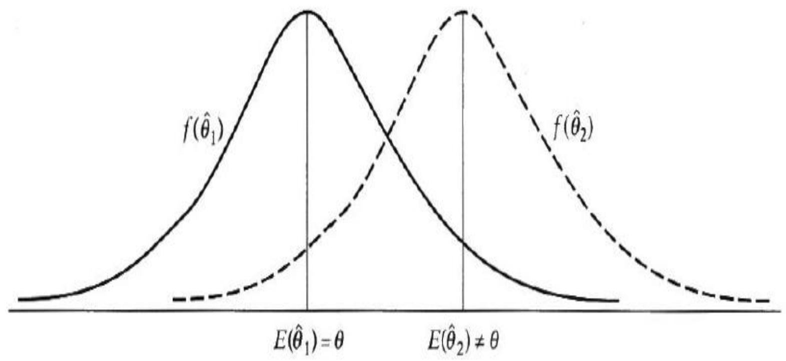
[5] Son aquellas que se mantienen sólo bajo ciertos supuestos según la forma como se generen los datos. Son propiedades de muestras finitas. **VER APÉNDICE A, D. GUJARATI 4ta EDICIÓN**
[6] Se elimina β_1 por simplicidad; las VA están expresadas en desviaciones medias.

2) ***Insesgadez[7]:*** Si un estimador es insesgado, entonces su distribución de probabilidad tiene un valor esperado igual al parámetro que se supone estimará. En este caso, $E(\hat{\beta}_i) = \beta_i$.

Demostración 2[8]: Un estimador, $\hat{\theta}$ de θ, es un estimador insesgado si $\boldsymbol{E(\hat{\theta}) = \theta}$**,** para todos los valores posibles de θ.

$$E(\hat{\beta}_2) = \beta_2$$

$$\begin{aligned} E(\hat{\beta}_2) &= E(\sum w_i y_i) = \sum w_i E(y_i) = \sum w_i E(\beta_2 x_i + \varepsilon_i) \\ &= \beta_2 \sum w_i x_i + \sum \cancel{w_i E}(\varepsilon_i) \\ &= \cancel{\beta_2 \frac{\sum x_i}{\sum x_i^2}} \sum x_i = \beta_2 \end{aligned}$$



Usando algebra lineal:

$$\boldsymbol{E(\hat{\beta}) = E[\beta + (X'X)^{-1}X'\varepsilon] = \beta + (X'X)^{-1}X'E(\varepsilon)}$$
$$\boldsymbol{= \beta}$$

Fuente: **Gujarati**, D. (2003). Pp. 871

[7] La insesgadez no significa que la estimación que se obtuvo de una muestra particular sea igual al parámetro, o incluso muy cercana a éste.
[8] Si W es un estimador sesgado de θ, su sesgo se define como: $\boldsymbol{E(\hat{\theta}) - \theta}$

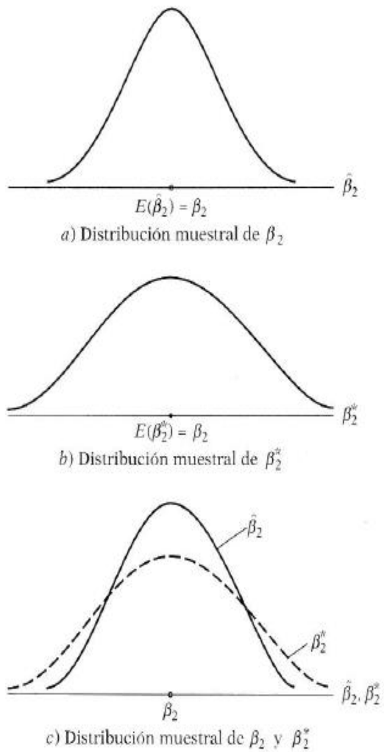
3) *Eficiencia:* el estimador tiene varianza mínima en la clase de todos los estimadores lineales e insesgados.

Demostración 3: Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son dos estimadores insesgados de θ, $\hat{\theta}_1$ es eficiente en relación a $\hat{\theta}_2$ cuando $Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$ para todo θ.

Sea otro estimador $\widetilde{\beta_2} = \sum d_i y_i$, el cual también es lineal e insesgado si: $\sum d_i x_i = 1$

$$Var\left(\widetilde{\beta_2}\right) = Var\left(\sum d_i y_i\right) = \sum {d_i}^2 Var(y_i) = \sum {d_i}^2 \sigma^2$$

Se busca encontrar un d_i , que minimice $Var\left(\widetilde{\beta_2}\right)$, sujeto a $\sum d_i x_i = 1$



Gujarati, D. 2003 PP 77

$$Min \sum {d_i}^2 - \lambda\left(\sum d_i x_i - 1\right)$$

$$\frac{\partial \sum {d_i}^2}{\partial d_i} = 2\sum d_i - \lambda \sum x_i = 0$$

$\sum d_i = \frac{\lambda}{2} \sum x_i$ (1) Multiplicando a ambos lados por x_i, tenemos:

$$\sum d_i x_i = \frac{\lambda}{2} \sum {x_i}^2 \quad \rightarrow \quad 1 = \frac{\lambda}{2} \sum {x_i}^2$$

$\lambda = \frac{2}{\sum {x_i}^2}$ (2) Sustituyendo (2) en (1)

$$\sum d_i = \frac{\frac{2}{\sum {x_i}^2}}{2} \sum x_i = \frac{\sum x_i}{\sum {x_i}^2} = \sum w_i$$

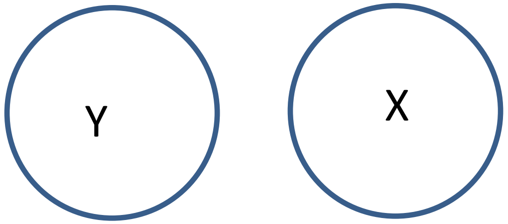
Teorema de Gauss-Markov: Dados los supuestos básicos del modelo clásico de regresión lineal, los estimadores mínimos cuadrados ordinarios, dentro de la clase de estimadores lineales insesgados, tienen varianza mínima, es decir, son **MELI.**

COEFICIENTE DE DETERMINACIÓN: BONDAD DE AJUSTE

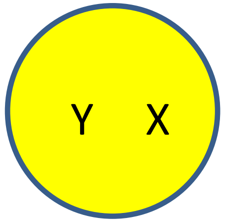
El coeficiente de determinación R^2 dice que tan bien se ajusta la recta de regresión muestral a los datos.

Propiedades:

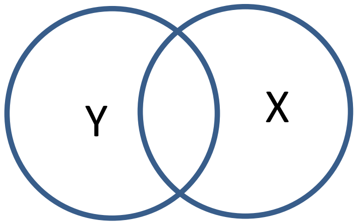
- Es una cantidad no negativa.
- Se encuentra entre $0 < R^2 < 1$



$R^2 = 0$



$\boldsymbol{R^2 = 1}$



$0 < R^2 < 1$

Descomposición de varianza:

1) ***Suma de cuadrados totales (SCT):*** variación total de los valores reales de Y con respecto a su media muestral.

$$\sum \mathrm{y_i}^2 = \sum (\mathrm{Y_i} - \overline{\mathrm{Y}})^2$$

2) ***Suma de cuadrados de la regresión (SCR):*** variación de los valores Y estimados alrededor de su media $\left(\hat{Y} = \bar{Y}\right)$. Es la variación de Y explicada por el modelo:

$$\sum \hat{\mathrm{y}}_{\mathrm{i}}^2 = \sum \left(\hat{\mathrm{Y}}_{\mathrm{i}} - \overline{\mathrm{Y}}\right)^2$$

3) ***Suma de cuadrados de los residuos (SCE):*** variación no explicada por el modelo:

$$\sum \hat{\varepsilon}_{\mathrm{i}}^2$$

SCT = SCR + SCE

Dividiendo entre SCT

$$\frac{SCT}{SCT} = \frac{SCR}{SCT} + \frac{SCE}{SCT}$$

$$1 = \frac{SCR}{SCT} + \frac{SCE}{SCT}$$

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

EL R^2 MIDE LA PROPORCION O EL PORCENTAJE DE LA VARIACIÓN TOTAL EN Y EXPLICADA POR EL MODELO DE REGRESIÓN

El R^2 está altamente relacionado con el coeficiente de correlación ($\hat{\rho}$).

Recuerde que el coeficiente de correlación:

- Puede tener signo positivo o negativo.
- Se encuentra en el intervalo $[-1; 1]$
- Es simétrico.
- Es independiente del origen y de la escala.
- No implica relación causa-efecto

$$\hat{\rho} = \pm\sqrt{R^2}$$

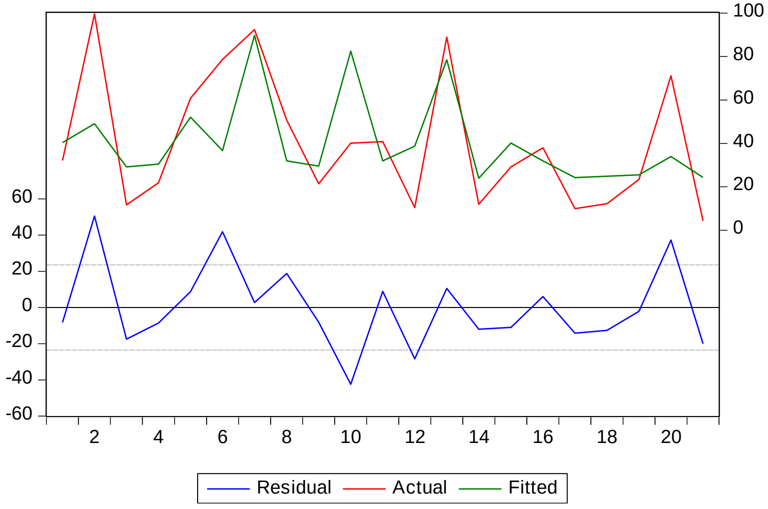
Ejemplo 1: Impacto de los gastos en publicidad. Wall Street Journal 1984.

$$\boldsymbol{Impresiones = \hat{\beta}_1 + \hat{\beta}_2 Gasto + \hat{\varepsilon}_i}$$

Dependent Variable: IMPRESIONES
Method: Least Squares
Sample: 1 21
Included observations: 21

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	22.16269	7.089479	3.126139	0.0056
GASTO	0.363174	0.097120	3.739425	0.0014

R-squared	0.423951	Mean dependent var	40.46667
Adjusted R-squared	0.393633	S.D. dependent var	30.18061
S.E. of regression	23.50152	Akaike info criterion	9.242400
Sum squared resid	10494.11	Schwarz criterion	9.341878
Log likelihood	-95.04520	Hannan-Quinn criter.	9.263989
F-statistic	13.98330	Durbin-Watson stat	2.371839
Prob(F-statistic)	0.001389		



Ejemplo 2: Relación tasa de interés e inflación en Venezuela 1982-2005

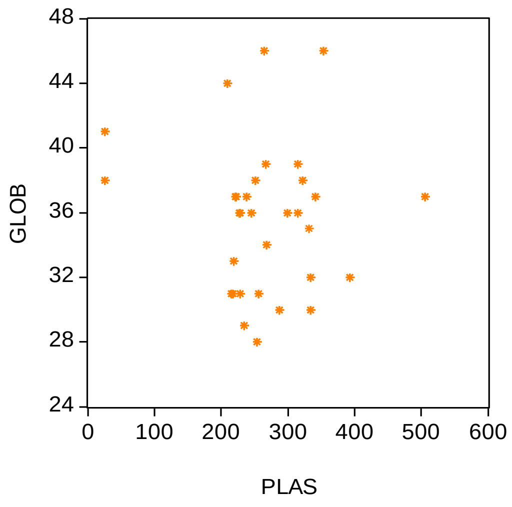
Dependent Variable: TI
Method: Least Squares
Sample: 1982 2005
Included observations: 24

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	13.93345	3.566080	3.907217	0.0008
INF	0.219672	0.089295	2.460058	0.0222

R-squared	0.215739	Mean dependent var	21.03333
Adjusted R-squared	0.180091	S.D. dependent var	11.33273
S.E. of regression	10.26165	Akaike info criterion	7.574361
Sum squared resid	2316.634	Schwarz criterion	7.672532
Log likelihood	-88.89233	Hannan-Quinn criter.	7.600405
F-statistic	6.051885	Durbin-Watson stat	0.970288
Prob(F-statistic)	0.022218		

$$\widehat{TI} = 13.93345 + 0.219672INF$$

Ejemplo 3: Nivel de plasma fibrinógeno (gm/lt) en función de la gammaglobilina (gm/lt).



Dependent Variable: GLOB
Method: Least Squares
Sample: 1 32
Included observations: 32

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	36.80305	2.605308	14.12618	0.0000
PLAS	-0.004340	0.009357	-0.463790	0.6461

R-squared	0.007119	Mean dependent var	35.65625
Adjusted R-squared	-0.025977	S.D. dependent var	4.583346
S.E. of regression	4.642495	Akaike info criterion	5.968842
Sum squared resid	646.5827	Schwarz criterion	6.060451
Log likelihood	-93.50148	Hannan-Quinn criter.	5.999208
F-statistic	0.215101	Durbin-Watson stat	1.694466
Prob(F-statistic)	0.646146		

$$\widehat{Glob} = 322.7400 - 1.640386 Plas$$

Lecturas obligatorias:

- ✓ Gujarati, D. y Porter, D. (2010). Econometría. 5ta. Edición McGraw Hill. Capítulo 3.
- ✓ Novales, A. (1993). Econometría. 2da. Edición. McGraw Hill. Interamericana. Capítulo 3.
- ✓ Wooldridge, J. (2010). Introducción a la Econometría. Un enfoque moderno. Capítulo 2.