Modelo de Regresión Lineal

Mg. Heber J. Baldeón [1]

Grupo Lambda

Marzo 2023

[1] hbaldeon@grupolambda.com.pe

Contenido

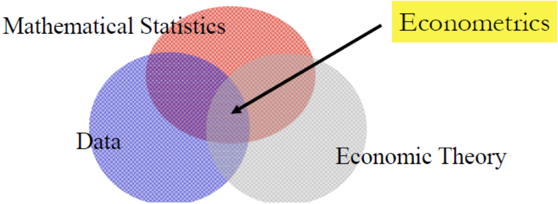
Bibliografía

Textos obligatorios

1. Wooldridge, Jeffrey M. "*Econometric Analysis of Cross Section and Panel Data*" Massachusetts Institute of Technology, 2002.
2. Stock & Watson "*Introduction to Econometrics*" 2nd edition. 2006.
3. Cameron, A. Colin y Pravin K. Trivedi. "*Microeconometrics. Methods and Applications*" Cambridge University Press, 2005.
4. Greene, William H. "*Econometric Analysis*" 6th edition, 2008.
5. Davidson, Russell y James G. Mackinnon. "*Econometric Theory and Methods*".

Econometría

- La econometría es **parte esencial** de la economía, es nuestro distintivo, nuestro valor agregado. No es una disciplina aparte.
- Ragnar Frisch, Econometrica Vol.1 No. 1 (1933) revisited. "*Experience has shown that each of these three view-points, that of* **statistics, economic theory, and mathematics**, *is a necessary, but not by itself a sufficient, condition for a real understanding of the quantitative relations in modern economic life.*
  It is the unification of all three aspects that is powerful. And it is this unification that constitutes econometrics."



Introducción

- Su importancia:
    - Evaluar las teorías económicas
    - Encontrar relaciones relevantes
    - Efecto causal (inferencia causal)
    - Predecir
- La econometría estudia los fenómenos económicos:
    - Fenómenos sociales (heterogeneidad)
    - Fenómenos no medibles o no observables (suerte, preferencias, habilidades, etc)
    - Datos no controlables (no experimentales)
- **Estructura de los datos económicos:**
    - Datos de corte transversal
    - Datos de serie de tiempo
    - Combinación de cortes transversales
    - Datos de panel o longitudinales

Introducción

- La **Estadística** nos brinda una primera aproximación de la relación de los fenómenos económicos (variables aleatorias) utilizando la covarianza y la correlación muestral (medidas de asociatividad lineal).

$$cov(X,Y)=\frac{\sum_{i=1}^{N}(X_i-\bar{X})(Y_i-\bar{Y})}{N-1} \text{ y } corr(X,Y)=\frac{\frac{\sum_{i=1}^{N}(X_i-\bar{X})(Y_i-\bar{Y})}{N-1}}{\sqrt{\frac{\sum_{i=1}^{N}(X_i-\bar{X})^2}{N-1}}\sqrt{\frac{\sum_{i=1}^{N}(Y_i-\bar{Y})^2}{N-1}}}$$

- Pero, la Econometría estudia la **relación** entre la **variable dependiente, explicada, regresada** (y_i), las **variables independientes, explicativa, covariable o regresora** $(\mathbf{x}_i)$ y la **perturbación estocástica o término de error** (u_i), es decir:

$$y_i = f(x_{1i}, x_{2i}, x_{3i}, \ldots, x_{ki}, u_i)$$

Modelo de Regresión Lineal

Si tenemos un **tamaño de muestra de N** y la relación entre la variable dependiente **y** y las **K variables independientes x** (incluyendo a la constante) en forma matricial sería:

$$
\begin{aligned}
y_1 &= \beta_1 + \beta_2 x_{21} + \beta_3 x_{31} + \cdots + \beta_k x_{K1} + u_1 \\
y_2 &= \beta_1 + \beta_2 x_{22} + \beta_3 x_{32} + \cdots + \beta_k x_{K2} + u_2 \\
&\vdots \\
y_i &= \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \cdots + \beta_k x_{Ki} + u_i \\
&\vdots \\
y_N &= \beta_1 + \beta_2 x_{2N} + \beta_3 x_{3N} + \cdots + \beta_k x_{KN} + u_N
\end{aligned}
$$

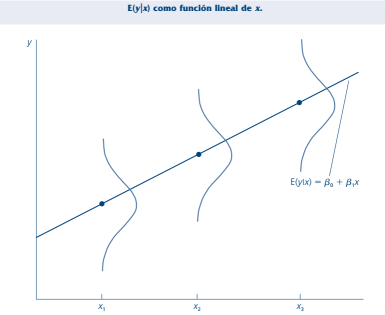
$$
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & x_{21} & \cdots & x_{K1} \\ 1 & x_{22} & \cdots & x_{K2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{2N} & \cdots & x_{KN} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}
$$

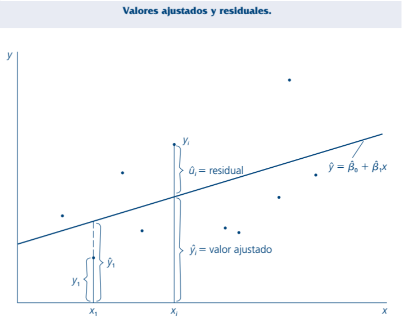
$$
\mathbf{Y}_{N\times 1} = \mathbf{X}_{N\times K}\beta_{K\times 1} + \mathbf{u}_{N\times 1}
$$

Intuición del Modelo de Regresión Lineal

Para el caso particular bivariado, el ML tiene un componente determinístico o teórico y un componente estocástico.

$$y_i = \beta_1 + \beta_2 x_{2i} + u_i$$





Supuestos del Modelo de Regresión Lineal

1. Primer momento: Esperanza condicional $E(u_i/x) = 0$ para $i = 1, \ldots, N$
   $\Rightarrow E[u/x] = 0$ y $Cov(u, x) = 0$ o $E(u\mathbf{x}) = 0$ (**insesgadez**)
   - x_{ki} son fijas (no estocásticas) en muestreo repetido
   - x_{ki} estocasticas pero independientes de la perturbaciones $E[f(x_{ik})g(u_i)] = E[f(x_{ik})]E[g(u_i)]$ o $E[u_i/x_{ik}] = E[u_i]$
   - x_{ki} estocásticas pero sin correlación $E([x_{ik} - E(x_{ik})][u_i - E(u_i)]) = Cov(x_{ik}, u_i) = 0$
2. Segundo momento: Perturbaciones esféricas $Var(u_i/x) = \sigma^2\mathbf{I}_N$ (**eficiencia**)
   - Homoscedasticidad: $var(u_i) = E[(u_i - E(u_i))^2] = E(u_i)^2 = \sigma_u^2$ para $i = 1, \ldots, N$
   - Ausencia de correlación entre perturbaciones: $cov(u_i, u_j) = E[u_i - E(u_i)][u_j - E(u_j)] = E(u_i u_j) = 0$ para $\forall i \neq j$
3. Perturbaciones normales: $u_i/X \sim iidN(0, \sigma_u^2\mathbf{I}_N)$
4. Rango completo: La matriz $\mathbf{x}$ tiene k columnas linealmente independientes (**colinealidad**)
5. Correcta especificación:
   - **Linealidad en parámetros.** Parámetros constantes o alrededor de una constante (sin quiebre estructural),
   - **Problemas de endogeneidad** por variables omitidas, variables irrelevantes, errores de medición, simultaneidad.

Modelo de Regresión Lineal

Los supuestos clásicos, en términos matriciales, se representa así:

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

$$var(\mathbf{u}) = E[(u - E(u))(u - E(u))'] = E[uu']$$

$$var(\mathbf{u}) = \begin{bmatrix} E(u_1^2) & E(u_1 u_2) & \cdots & E(u_1 u_N) \\ E(u_2 u_1) & E(u_2^2) & \cdots & E(u_2 u_N) \\ \vdots & \vdots & \ddots & \vdots \\ E(u_N u_1) & E(u_N u_2) & \cdots & E(u_N^2) \end{bmatrix} = \begin{bmatrix} \sigma_u^2 & 0 & \cdots & 0 \\ 0 & \sigma_u^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_u^2 \end{bmatrix}$$

$$var(\mathbf{u}) = \sigma_u^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \sigma_u^2 I_N$$

Estimador de Mínimos Cuadrados Ordinarios

- El objetivo es estimar los β desconocidos, a partir de datos disponibles, para luego, validar teorías y realizar pronósticos. Para ello, definimos:

  $\hat{Y} = E[Y/X] = X\hat{\beta}$

  Residuo muestral: $\hat{u} = Y - \hat{Y} = Y - X\hat{\beta}$

- El método más frecuentemente utilizado es el de Mínimos Cuadrados Ordinarios (MCO):

$$SCR(\hat{\beta}) = \sum_{i=1}^{N} \hat{u}^2 = \hat{u}\hat{u}' = (Y - X\hat{\beta})(Y - X\hat{\beta})'$$

$$\hat{u}\hat{u}' = YY' - \hat{\beta}'X'Y - Y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} = YY' - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta}$$

$$CPO : \frac{\partial(\hat{u}\hat{u}')}{\partial\hat{\beta}} = 0 \Rightarrow -X'Y + (X'X)\hat{\beta}_{MCO} = 0$$

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

$$Var(\hat{\beta}_{MCO}) = \hat{\sigma}_u^2 \left(X'X\right)^{-1}$$

Estimador de Mínimos Cuadrados Ordinarios

- El objetivo es estimar la varianza de los residuos muestrales
  Residuo muestral: $\hat{u} = Y - \hat{Y} = Y - X\hat{\beta}$
  El estimador $\hat{\sigma}_u^2 = \frac{\hat{u}'\hat{u}}{N-K}$ es insesgado respecto de σ_u^2
- $\hat{u} = Y - \hat{Y} = Y - X\hat{\beta} = Y - X(X'X)^{-1}X'Y = [I_N - X(X'X)^{-1}X']Y = MY$
  $\hat{u} = M(X\beta + u) = MX\beta + Mu = [I_N - X(X'X)^{-1}X']X\beta + Mu = [X - X(X'X)^{-1}X'X]\beta + Mu$
  $\hat{u} = Mu$
- $tr(M) = tr(I_N - X(X'X)^{-1}X') = tr(I_N) - tr(X(X'X)^{-1}X') = N - tr(I_K) = N - K$
- $\mathbb{E}(\hat{\sigma}_u^2) = \mathbb{E}(\frac{u'u}{N-K}) = \mathbb{E}(\frac{u'M'Mu}{N-K}) = \mathbb{E}(\frac{u'Mu}{N-K}) = \mathbb{E}(\frac{tr(u'Mu)}{N-K}) = \mathbb{E}(\frac{tr(u'uM)}{N-K})$
  $\mathbb{E}(\hat{\sigma}_u^2) = \frac{tr(\mathbb{E}(u'uM))}{N-K} = \frac{tr(\sigma_u^2 IM)}{N-K} = \frac{\sigma_u^2 tr(IM)}{N-K} = \frac{\sigma_u^2 tr(M)}{N-K}$
  $\mathbb{E}(\hat{\sigma}_u^2) = \sigma_u^2$

Propiedades del Estimador MCO para muestras finitas

Propiedades de los estimadores

1. Insesgadez: $E(\hat{\beta}) = \beta$, $E[\hat{\sigma}_u^2] = \sigma_u^2$.
2. Eficiencia: $Var(\hat{\beta}_{MCO}) = \hat{\sigma}_\mu^2\,(X'X)^{-1}$

   Dados 2 estimadores $(\hat{\beta}_1, \hat{\beta}_2)$, si $Var(\hat{\beta}_1) < Var(\hat{\beta}_2)$ entonces $\hat{\beta}_1$ es más eficiente.
3. Error cuadrático medio(MSE): $MSE = E[(\hat{\beta} - \beta)^2] = Var(\hat{\beta}) + (Sesgo)^2$

Teorema de Gauss-Markov

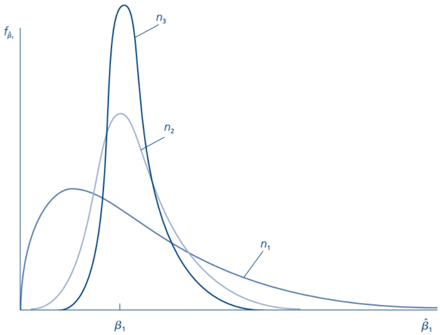
De un grupo de estimadores lineales e insesgados para los parámetros, dentro de los cuales figuran los estimadores de MCO, estoss estimadores MCO son los que tienen menor varianza, es decir, son los más eficientes. Por ello, los estimadores de MCO son los Mejores Estimadores Lineales e Insesgados o MELI.

Propiedades del Estimador MCO para muestras grandes

Propiedades de los estimadores

1. Consistencia: Si el estimador $\hat{\beta}$ converge en probabilidad β, entonces el estimador es consistente (Ley de Grandes Números).
2. Normalidad asintótica: Teorema del Límite Central
3. Eficiencia asintótica

Distribuciones de muestreo de $\hat{\beta}_1$ para muestras de tamaños $n_1 < n_2 < n_3$.



Bondad de ajuste del estimador MCO

¿Cuán lejos se encuentra la recta estimada con respecto a los datos?

$$y = \hat{y}_i + \hat{u}_i \Longrightarrow \underbrace{\sum_{i=1}^{N}(y_i - \bar{y})^2}_{SCT} = \underbrace{\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2}_{SCE} + \underbrace{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}_{SCR}$$

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

- El R^2 mide el porcentaje de variación total observada de la variable endógena explicada **linealmente** por la variación de las variables independientes del modelo estimado.
- Si el modelo poblacional tiene intercepto y se estima sin intercepto, la SCE y SCR puede crecer mucho (valores mayores que 1 o negativos).
- No tiene sentido comparar el R^2 de dos muestras diferentes, dos periodos diferentes, dos países diferentes, de dos modelos con diferente número de regresoras.
- **Regresiones espúreas:** Relación estadísticamente significativa (R2 alto) pero sin sentido (errores presentan persistencia, no estacionarios).

Otras medidas de Bondad de ajuste

1. **Criterio de información de Theil o R^2 ajustado**

$$\bar{R}^2 = 1 - \frac{\frac{SCR}{(N-K)}}{\frac{SCT}{(N-1)}} = 1 - \frac{(N-1)}{(N-K)}(1 - R^2)$$

2. **Criterio de información de Akaike o AIC**

$$AIC = \frac{-2}{N} * (lnL - K) \rightarrow ln(\frac{\varepsilon'\varepsilon}{N}) + \frac{2K}{N}$$

3. **Criterio de información de Bayes - Schwarz o BIC**

$$BIC = -(\frac{2}{N} * lnL - \frac{ln(N)}{N} * K) \rightarrow ln(\frac{\varepsilon'\varepsilon}{N}) + \frac{K}{N} ln(N)$$

4. **Criterio de información de Amemiya**

$$\frac{\varepsilon'\varepsilon}{N} + (1 + \frac{K}{N})$$

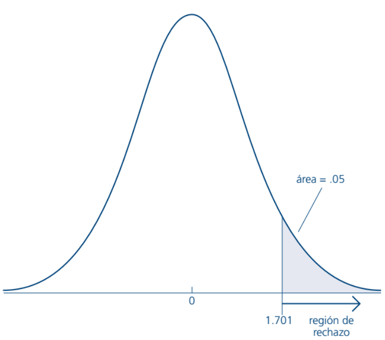
Pruebas de hipótesis del Modelo de Regresión Lineal

Prueba de significación individual:

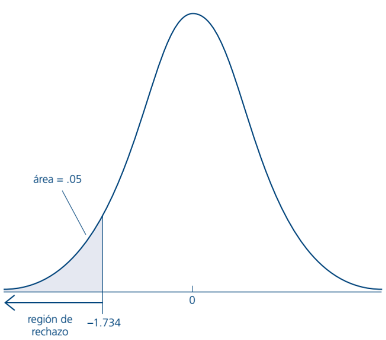
Para $H_0 : c'\hat{\beta} - r = 0$, el estadístico sería: $z = \dfrac{c'\hat{\beta} - r}{\sqrt{\sigma_\mu c'(X'X)^{-1}c}}$.

Pero no tenemos σ_μ, sino $\hat{\sigma}_\mu$. Así, el estadístico sería: $t = \dfrac{c'\hat{\beta} - r}{\sqrt{\hat{\sigma}_\mu c'(X'X)^{-1}c}}$

Regla de rechazo de 5% para la alternativa H_1: $\beta_j > 0$ con 28 *gl*.



Regla de rechazo de 5% para la alternativa H_1: $\beta_j < 0$ con 18 *gl*.

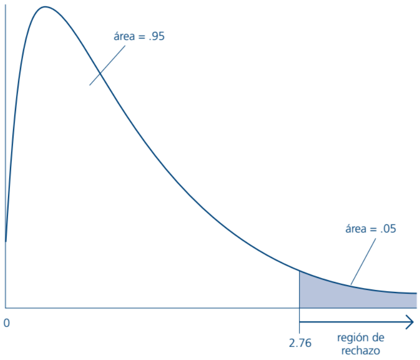


Pruebas de hipótesis del Modelo de Regresión Lineal

Prueba de significación global:

Para $H_0 : \beta_2 = \beta_3 = \beta_4 = \ldots \beta_K = 0$, el estadístico sería: $F = \dfrac{\frac{SCE}{(K-1)}}{\frac{SCR}{(N-K)}}$.

Valor crítico correspondiente a 5% y región de rechazo en la distribución $F_{3,60}$.



Validación de los Supuestos Clásicos

1. **Parámetros cambiantes**
   - Test basado en residuos recursivos - Cusum
     El residuos recursivos, $e_t = Y_t - x'\hat{\beta}_{t-1}$, se normaliza y $w_t = \frac{e_t}{\sqrt{1+x_t'(x_{t-1}'x_{t-1})^{-1}x_t}} \sim N(0, \sigma_u^2)$ y se define el estadístico $W_t = \sum_{j=k+1}^{t} w_j/\hat{\sigma}, t = k+1, \ldots, n$. Donde σ es la desv. est. residual del modelo estimado con n datos. Bajo la $H_0 : W_t = 0$, ausencia de cambio estructural, el estadístico debería fluctuar sobre el valor 0.
   - Test basado en residuos recursivos - Cusum 2
     En caso el estadístico $S_t = \frac{\sum_{r=k}^{t} w_r^2}{\sum_{r=k}^{n} w_r^2}$ permanezca dentro de las bandas de confianza de $E(S_t) = \frac{t-k}{n-k}$, no habría quiebre estructural.
   - Test estructural: Test de Chow (más potente)
     Separar la muestra en 2 submuestras para comparar los errores de los modelos de las partes con el modelo total. Si $\sum_{t=1}^{n_1} e_{t1}^2 + \sum_{t=1}^{n_1} e_{t2}^2 \approx \sum_{t=1}^{n_1} e_t^2$ no existirá cambio estructural.

$$F = \frac{[e'e - (e_1'e_1 + e_2'e_2)]/k}{(e_1'e_1 + e_2'e_2)/n - 2k} \sim F(k, n-2k)$$

Solución:

- Incorporar variables dicotómicas
- Mejorar la especificación del modelo

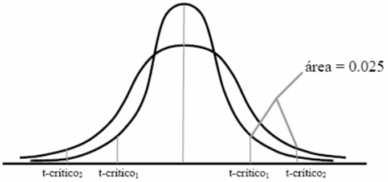
Validación de los Supuestos Clásicos

2. Multicolinealidad

El $rang(x) = k$ garantiza que $(X'X)^{-1}$ sea invertible, por que el estimador MCO tenga solución.

<u>Consecuencias de la multicolinealidad:</u>

- Pruebas t sesgadas a la baja. Nos induciría al error tipo 1 (no rechazar H_0, cuando es falsa)



- t poco significativas y R^2 alto
- Sensibilidad de los parámetros a pequeños cambios en la muestra (pérdida de datos).

Validación de los Supuestos Clásicos

Detección de la multicolinealidad:

- **Prueba del tamaño relativo de los autovalores**

  Calcular los autovalores de $X'X$ y examinar si algún autovalor es cercano a 0.

  $I = \sqrt{\frac{\lambda_{MAX}}{\lambda_{MIX}}}$. Regla heurística: mayor a 20 o 25 implica alta colinealidad.

- **Factor de inflación de la Varianza** (VIF):

  $\hat{var}(\hat{\beta}) = s^2(X'X)^{-1} \rightarrow \hat{var}(\hat{\beta}) = \dfrac{s^2}{(N-1)\hat{var}(x_j)} \dfrac{1}{1-R_j^2}$ donde R_j^2 resulta de la regresión alternativa

  $x_j = g(x_1, \ldots, x_{j-1}, x, \ldots, x_N)$ y $VIF = \dfrac{1}{1-R_j^2}$

  Si el $VIF > 10$ hay una alta multicolinealidad. El VIF promedio debe ser cercano a 1.

- **Test de Farrar y Glauber:** Utilizan 3 test estadísticos para verificar la colinealidad. i) Prueba χ^2, con $H_0 : X_i$ ortogonales entre si; ii) Prueba F, con $H_0 : R^2_{max} = 0$; y iii) utilizando la matriz de correlación, la prueba t, con $H_0 : r_{max} = 0$.

  Solución:

  Agregar más datos o eliminar el regresor que genera la mayor colinealidad.

Validación de los Supuestos Clásicos

3. Heteroscedasticidad: $H_0 : var(u/x_1, x_2, \ldots, x_k) = E(u^2/x_1, x_2, \ldots, x_k) = E(u^2) = \sigma^2$

- Pruebas gráficas
- Test de Glesjer
- Test de Park
- Ratio de Verosimilitud
- Test de Breusch - Pagan: Se estima un modelo MCO y los $\hat{u}$. Se estima una ecuación alternativa de $\hat{u}^2$ con todos los regresores, para hallar el $R^2_{\hat{u}^2}$

  Y se plantea el estadístico $F = \dfrac{R^2_{\hat{u}^2}/K}{(1 - R^2_{\hat{u}^2})/N - K - 1} \sim F_{(K,N-K-1)}$ o $LM = nR^2_{\hat{u}^2} \sim \chi^2_K$
- Test de White: Se estima un modelo MCO y los $\hat{u}$. Se estima una ecuación alternativa de $\hat{u}^2$ con todos los regresores, regresores al cuadrado y producto cruzado. Se calcula el $R^2_{\hat{u}^2}$, para luego utilizar los estadísticos F y LM.

  Problema: sin son muchos regresores se pierde muchos grados de libertad, por lo que se podría utilizar una regresión alternativa: $\hat{u}^2 = \delta_0 + \delta_1\hat{y} + \delta_2\hat{y}^2 + error$, y se halla el $R^2_{\hat{u}^2}$ para calcular los estadísticos F o LM.

Validación de los Supuestos Clásicos

Solución:

- Mínimos cuadrados ponderados o WLS

$\sigma(z_i) = \sigma$

$$\left\{ \frac{y}{\sigma(z_i)} \right\} = \left\{ \frac{x}{\sigma(z_i)} \right\}' \beta + \varepsilon_i$$

- Mínimos cuadrados generalizados

$$\hat{\beta}_{MCG} = \left(X'\Omega^{-1}X \right)^{-1} X'\Omega^{-1}Y$$

- Mínimos cuadrados generalizados factibles

$$\hat{\beta}_{MCGF} = \left(X'\hat{\Omega}^{-1}X \right)^{-1} X'\hat{\Omega}^{-1}Y$$

$$\hat{V}(\hat{\beta}_{MCGF}) = \left(X'\hat{\Omega}^{-1}X \right)^{-1}$$

Validación de los Supuestos Clásicos

4. Autocorrelación

- Correlogramas
  Autocorrelación simple y parcial significativas (fuera de las bandas de Barlett)
- Test de Durbin Watson
  $H_0 : \rho = 0$

$$DW = \frac{\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n}\hat{u}_t^2}, \leftrightsquigarrow DW \approx 2(1-\hat{\rho})$$

- Test de Breusch-Godfrey (para q rezagos)
  Se realice una regresión MCO, se calcula los residuos $(\hat{u}_t)$ de la regresión, posteriormente, se estima una regresión alternativa $\hat{u}_t = f(x_{t1}, x_{t2}, \ldots, x_{tk}, \hat{u}_{t-1}, \hat{u}_{t-2}, \ldots, \hat{u}_{t-q})$, y se calcula el $R^2_{\hat{u}}$

$$LM = (n-q)R^2_{\hat{u}} \sim \chi^2_q$$

- **Solución**: Mínimos cuadrados generalizados factibles
  - Estimador de Prais - Winsten: $\tilde{y}_t = \beta_0\tilde{x}_{t0} + \beta_1\tilde{x}_{t1} + \cdots + \beta_k\tilde{x}_{tk} + error_t$ donde: $\tilde{x}_{10} = (1-\hat{\rho}^2)^{1/2}$ y $\tilde{x}_{t0} = (1-\hat{\rho})$ para $t \geqq 2$.
  - Estimador de Cochrane - Orcutt: $\tilde{y}_t = \beta_1\tilde{x}_{t1} + \cdots + \beta_k\tilde{x}_{tk} + error_t$
  - Estimador de Newey - West:

Validación de los Supuestos Clásicos

5. **Variables redundantes o irrelevantes**

   Modelo correcto: $Y = X_1\beta_1 + u$

   Modelo estimado: $Y = X_1\beta_1 + X_2\beta_2 + v$

   Antes: $Y = X\beta + u \longmapsto \hat{\beta}_{MCO} = (X'X)^{-1}X'Y$

   Ahora: $Y = X_1\beta_1 + X_2\beta_2 + u$

   Por el Teorema de Frisch-Waugh-Lovell (1933):

$$\hat{\beta}_1 = (X_1'M_2X_1)^{-1}X_1'M_2Y$$

$$\hat{\beta}_2 = (X_2'M_1X_2)^{-1}X_2'M_1Y$$

   Donde: $M_i = I - X_i(X_i'X_i)^{-1}X_i'$

Validación de los Supuestos Clásicos

5. Variables redundantes o irrelevantes

- Impactos en el valor esperado:

  $E[\hat{\beta}_1] = \beta_1 + (X_1' M_2 X_1)^{-1} X_1' M_2 u$

  $V(\hat{\beta}_1 | X_1, X_2) = \sigma^2 (X_1' M_2 X_1)^{-1}$ vs $V(\hat{\beta}_1^* | X_1) = \sigma^2 (X_1' X_1)^{-1}$

  La inclusión de variable irrelevante no produce sesgo.

- Impactos en la Varianza:

  $E(\sigma^2) = E(\frac{\hat{u}' \hat{u}}{N - k_1 - k_2})$

  Pero si aumenta los errores estándar de los parámetros. Esto también es perjudicial puesto que tenderemos a cometer error tipo II con mayor probabilidad (no rechazar la nula cuando esta es falsa)

Validación de los Supuestos Clásicos

6. **Variables omitidas (endogeneidad)**

   Modelo correcto: $Y = X_1\beta_1 + X_2\beta_2 + u$

   Modelo estimado: $Y = X_1\beta_1 + v$

- Impactos en el valor esperado:

  $E[\hat{\beta}_1] = \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2$

  $V(\hat{\beta}_1^*|X_1, X_2) - V(\hat{\beta}_1|X_1)$ es (semi) definida positiva (estimadores más eficientes?)

  La omisión de variables relevantes sesga los parámetros. La dirección del sesgo dependerá del signo del parámetro de la variable omitida y de la covarianza de la variable omitida con las otras variables del modelo.

- Impactos en la Varianza:

  Estimador de la varianza de los errores (σ_ε^2) sezgado hacia arriba.

  Los errores estándar serán menores lo cual es perjudicial para la inferencia. Si se comete error tipo I o II con mayor probabilidad dependerá del tamaño del sesgo y de cuánto se subestimen los errores estándar.

Validación de los Supuestos Clásicos

7. Errores de medida (endogeneidad)

- Error de medida en la variable dependiente:

  $Y^* = \beta_0 + \beta_1 X + \varepsilon$, siendo $v_0 = Y - Y^* \Rightarrow Y = Y^* + v_0$

  Si $E(v_0) = 0 \wedge Cov(X, v_0) = 0$, el MCO es consistente.

  Los estimadores MCO empleando Y son más ineficientes que los obtenidos con Y^* (sin error de medida)

- Error de medida en la variable independiente:

  $Y = \beta_0 + \beta_1 X^* + \varepsilon$, siendo $v_1 = X - X^* \Rightarrow X = X^* + v_1$

  En presencia de errores de medida, tenderemos a **infraestimar** la magnitud (en valor absoluto) de la pendiente de la variable que se mide con error (X).

  Si $V(X^*)$ es grande en relación a $V(v_1)$, la inconsistencia puede llegar a ser despreciable.

Validación de los Supuestos Clásicos

8. **Endogeneidad (Simultaneidad)**

- $E(u_i/x) = 0$ para $i = 1, \ldots, N \Rightarrow E(u\mathbf{x}) = 0$
- Sol: Aplicar MC2E (Variables Instrumentales)

Concepto

Una Variable Instrumental (Z) es tal que esta correlacionada con la variable X pero que esta incorrelacionada con la variable u

$$\text{Z es una VI} \Leftrightarrow E(XZ) \neq 0 \text{ pero } E(Zu) = 0$$

$$y = \alpha + \beta x + u$$

- Como $E(Xu) \neq 0$, no puedo utilizar ni MCO ni MGM. Pero dado que conozco una VI para X, entonces aplico el método de estimación más sencillo, pero incorporándole la variable instrumental (Z).
- Estimador con variables instrumentales:
  Condiciones: $E(Zu) = 0$ y $E(u) = 0$
  $$\hat{\beta}_{VI_Z} = \frac{Cov(Z,y)}{Cov(Z,X)}$$