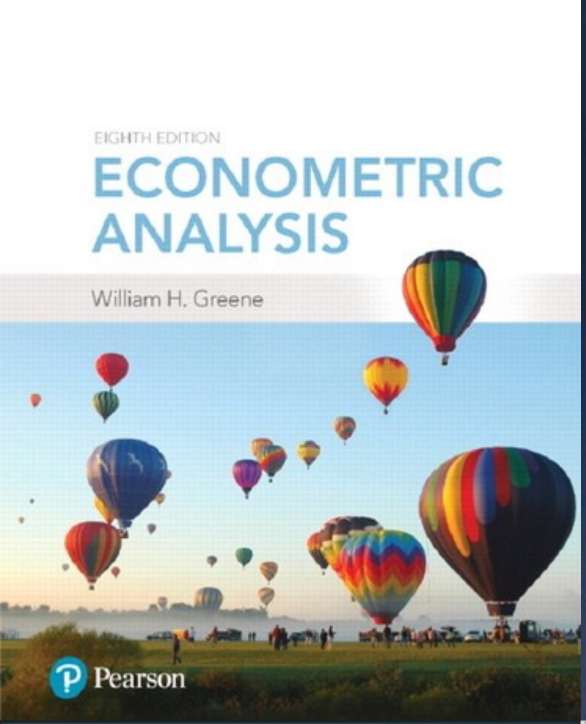
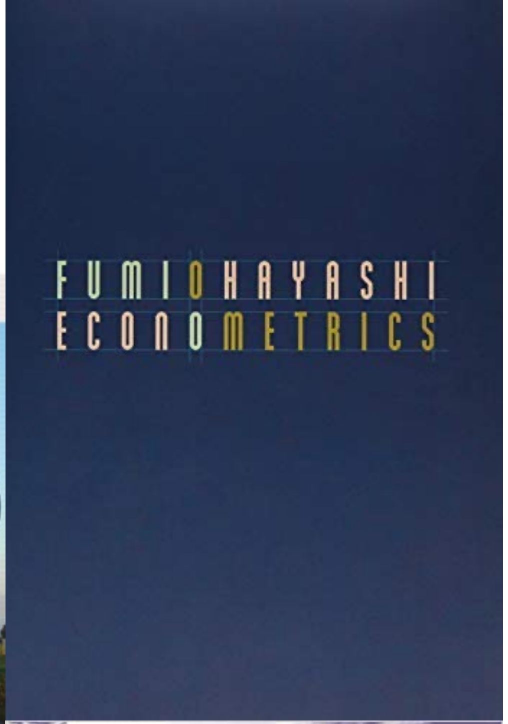
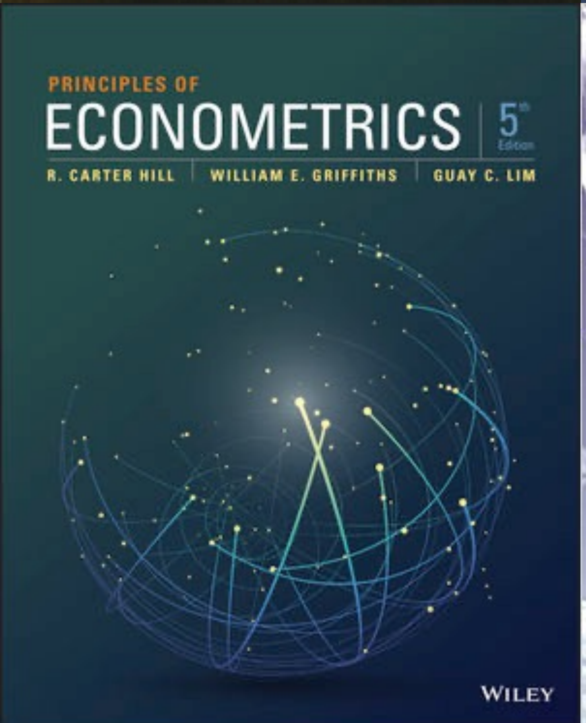
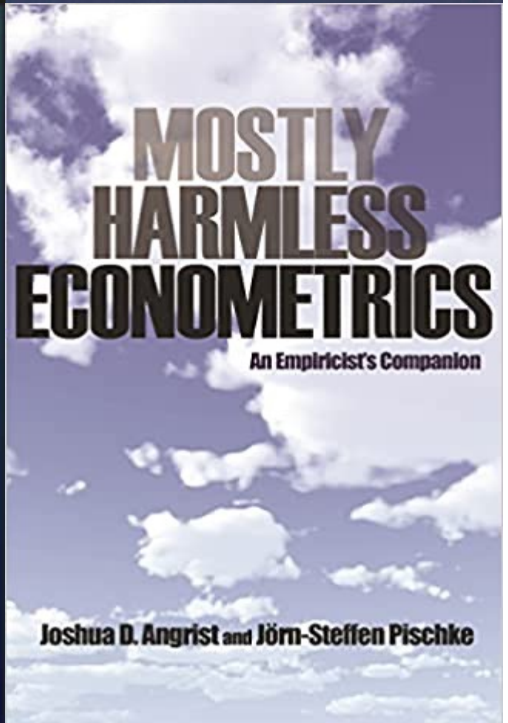
PUCP









MAESTRÍA EN ECONOMÍA
ECONOMETRÍA INTERMEDIA
ECO743 – MÓDULO 2

Sesión 1
Propiedades Finitas MCO y Perturbaciones No Esféricas

Docente: Juan Palomino

PUCP

Índice

1. Supuestos MCO

Supuestos MCO

Supuesto 1: Linealidad

La relación entre la variable dependiente y las variables independientes es lineal. El modelo poblacional es:

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_K x_{iK} + \varepsilon_i \qquad i = 1, \ldots, n$$

Donde β's son parámetros desconocidos a ser estimados, y ε_i son términos de errores observados.

Nota: La propiedad de linealidad es una propiedad de los parámetros, no de las variables.

Supuestos MCO

Supuesto 2: Exogeneidad Estricta

El valor esperado de la perturbación aleatoria debe ser cero para cualquier observación:

$$E[\varepsilon_i|X] = 0 \qquad i = 1,2,\dots,n$$

Para cualquier valor de $x_1, x_2, \dots, x_K$ en la población, el promedio no observable es igual a cero.

Supuestos MCO

Consecuencias de Exogeneidad Estricta:

Definición: Ley de Esperanzas Iteradas

Si $E|y| < \infty$, es decir, la esperanza poblacional de y existe, entonces para cualquier vector x aleatorio:

$$E_X[E(y|x)] = E(y)$$

1. Media Incondicional

La media incondicional del término de error es cero:

$$E(\varepsilon_i) = 0$$

Consecuencias de Exogeneidad Estricta:

2. Regresores son ortogonales al término de error

Si $E(xy)$ de dos variables aleatorias x e y es cero, entonces decimos que x es ortogonal a y. Bajo exogeneidad estricta, los regresores son ortogonales al término de error para todas las observaciones, es decir, no comparten información.

$$E\left(x_{jk}\varepsilon_i\right) = 0 \qquad i, j = 1,2, \dots, n \qquad k = 1, \dots, K$$

o

$$E(x \cdot \varepsilon_i) = \begin{pmatrix} E(x_{j1}\varepsilon_i) \\ E(x_{j2}\varepsilon_i) \\ \vdots \\ E(x_{jK}\varepsilon_i) \end{pmatrix} = \underbrace{0}_{K\times 1} \qquad \forall i, j$$

Supuestos MCO

Consecuencias de Exogeneidad Estricta:

3. Condición de cero correlación

El regresos no está correlacionado con el término de error:

$$\mathrm{Cov}(\varepsilon_i, x_{jk}) = 0$$

4. Media condicional de $\boldsymbol{y_i}$

La media condicional de la variable dependiente es una función lineal de los regresores.

Bajo el supuesto de linealidad y exogeneidad estrictra:

$$E(y_i|X) = x_i'\beta \qquad\qquad i = 1,2,\ldots,n$$

Supuestos MCO

Supuesto 3: Rango Completo

Asume que ninguna de las variables en el modelo es una combinación lineal exacta de las otras.

$$X \; es \; una \; matriz \; n \times K \; con \; rango \; K$$

Significa que X tiene rango de columna completo.

Nota: El rango de una matriz es igual al número de columnas linealmente independientes de la matriz. También se le conoce como **condición de identificación.**

Supuesto 4: Perturbaciones Esféricas

- **Homocedasticidad**

  Asume que todas las unidades tienen el mismo error de varianza

$$E\left[\varepsilon_i^2 \middle| X\right] = \sigma_0^2 > 0 \qquad i = 1,2,\ldots,n$$

- **No Correlación**

  Asume también que no hay correlación entre las observaciones

$$E\left[\varepsilon_i \varepsilon_j \middle| X\right] = 0 \qquad i,j = 1,2,\ldots,n \quad i \neq j$$

Supuestos MCO

Supuesto 5: Regresores no estocásticos

Las observaciones de x_i son fijos en muestras repetidas.

Asumir que los X son fijos quiere decir que, en repetidas muestras de X, los valores obtenidos $x_1, x_2, \ldots, x_K$ van a ser siempre los mismos; es decir, dejan de ser aleatorios.

Bajo este supuesto, ya no es necesario hablar de esperanzas condicionales:

$$E(\varepsilon_i) = 0$$
$$\mathrm{Var}(\varepsilon_i) = \sigma^2$$
$$\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$$

Supuestos MCO

Supuesto 6: Normalidad de los errores

ε_i distribuye normal con media cero y varianza σ^2 condicional a X:

$$\varepsilon_i | X \sim N(0, \sigma^2), \qquad i = 1,2, \ldots, n$$

2. Propiedades Estadísticas de MCO

Insesgadez del Estimador de MCO

Estimador Insesgado

Un estimador $\hat{\theta}$ para θ_0 es insesgado si $E(\hat{\theta}) = \theta_0$

Teorema 1: Insesgadez del Estimador MCO

Bajo el supuesto de Linealidad, Exogeneidad Estricta y Rango Completo:

$$E(\hat{\beta}|X) = \beta_0$$

Entonces, $\hat{\beta}$ es un estimador insesgado de β_0

Insesgadez del Estimador de MCO

Ejemplo:

Tenemos el siguiente modelo:

$$bmi_i = \alpha_0 + \beta_0 income_i + \varepsilon_i$$

Donde β_0 es el efecto del ingreso sobre obesidad y es igual a 1.

El siguiente paso es conseguir una muestra de la población y obtener un estimado de β_0, llamado $\hat{\beta}$.

Imaginen lo siguiente: $B = 50$ muestras de 1000 individuos y para cada muestra estimamos β.

1era muestra → $\hat{\beta}_1 = 0.9$

2da muestra → $\hat{\beta}_1 = 0.94$

3era muestra → $\hat{\beta}_1 = 0.7$

50ava muestra → $\hat{\beta}_1 = 0.5$

En promedio de las 50 muestras, conseguir: $(\frac{1}{B})\sum_{b=1}^{B} \hat{\beta}_b \approx 1$

Insesgadez del Estimador de MCO

Demostración

Tenemos:

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

$$\hat{\beta} = (X'X)^{-1}X'(X\beta_0 + \varepsilon)$$

$$\hat{\beta} = (X'X)^{-1}X'X\beta_0 + (X'X)^{-1}X'\varepsilon$$

$$\hat{\beta} = \beta_0 + (X'X)^{-1}X'\varepsilon$$

$$\hat{\beta} - \beta_0 = A\varepsilon$$

Donde $A = (X'X)^{-1}X'$. Se le conoce como **sampling error**.

Varianza del Estimador MCO

Teorema 2: Varianza del estimador MCO

Bajo el supuesto de Linealidad, Exogeneidad Estricta, Rango Completo y Errores Esféricos:

$$\mathrm{Var}(\hat{\beta}_{OLS}) = \sigma_0^2 (X'X)^{-1}$$

Entonces, $\hat{\beta}$ es un estimador insesgado de β_0

Teorema de Gauss Markov

Teorema 3: Gauss-Markov

Bajo todos los supuestos, el estimador MCO es eficiente en la clase de estimadores lineales insesgados. Es decir, para cualquier estimador insesgado $\tilde{\beta}$ que es lineal en y,

$$\mathrm{Var}(\tilde{\beta}) \geq \mathrm{Var}(\hat{\beta}_{OLS}|X)$$

Este teorema señala que el estimador MCO es eficiente en el sentido que su matriz de varianza-covarianza $Var(\hat{\beta}_{OLS}|X)$ es el más pequeño entre todos los estimadores insesgados. Por esta razón, al estimador MCO se le conoce como Mejor Estimador Lineal Insesgado (MELI) [Best Lineal Unbiased Estimator (BLUE)].

Covarianza entre $\widehat{\beta}$ y $\hat{\varepsilon}$

Teorema 4: Covarianza entre $\widehat{\beta}$ y $\hat{\varepsilon}$

Bajo todos los supuestos de Linealidad, Exogeneidad Estricta, Rango Completo, y Errores Esféricos:

$$Cov\left(\hat{\beta}, \hat{\varepsilon} \middle| X\right) = 0$$

Donde $\hat{\varepsilon} \equiv y - X\widehat{\beta}$

Covarianza entre $\widehat{\beta}$ y $\hat{\varepsilon}$

Demostración

$$
\begin{aligned}
Cov(\hat{\beta}, \hat{\varepsilon}|X) &= E\{[(\hat{\beta} - E(\hat{\beta}|X))][\hat{\varepsilon} - E(\hat{\varepsilon}|X)]'|X\} \\
&= E\{[(\hat{\beta} - \beta_0|X)][M\varepsilon - E(M\varepsilon|X)]'|X\} \\
&= E[A\varepsilon\varepsilon' M|X] \\
&= AE[\varepsilon\varepsilon'|X]M' \\
&= AE[\sigma_0^2 I_n]M \\
&= \sigma_0^2 (X'X)^{-1}X'M \\
&= \sigma_0^2 (X'X)^{-1}\underbrace{(MX)'}_{0} \\
&= 0
\end{aligned}
$$

Teorema 5: Insesgadez de $\hat{\sigma}^2$

Bajo todos los supuestos de Linealidad, Exogeneidad Estricta, Rango Completo, y Errores Esféricos:

$$\mathrm{E}(\hat{\sigma}^2|X) = \sigma_0^2$$

Siempre que $n > K$, de modo que $\hat{\sigma}^2$ está bien definido.

Insesgadez de $\widehat{\sigma}^2$

Demostración

Ya que $\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{(n-K)}$, la prueba equivale a mostrar que $E(\hat{\varepsilon}'\hat{\varepsilon}|X) = (n-K)\sigma_0^2$. Sabemos que $\hat{\varepsilon}'\hat{\varepsilon} = \varepsilon M \varepsilon$, donde M es la matriz aniquiladora. La demostración consiste en probar dos propiedades: (1) $E(\varepsilon M \varepsilon|X) = \sigma_0^2 \cdot traza(M)$, y (2) $traza(M) = n - K$.

1. Probar que$E(\varepsilon M \varepsilon|X) = \sigma_0^2 \cdot traza(M)$: ya que $\varepsilon' M \varepsilon = \sum_{i=1}^{n} \sum_{j=1}^{n} m_{ij}\, \epsilon_i \epsilon_j$, tenemos:

$$E(\varepsilon M \varepsilon|X) = \sum_{i=1}^{n} \sum_{j=1}^{n} m_{ij} E\left(\epsilon_i \epsilon_j \middle| X\right)$$

Ya que m_{ij}'s son funciones de X

$$E(\varepsilon M \varepsilon|X) = \sum_{j=1}^{n} m_{ii}\, \sigma^2$$

Ya que $E\left(\varepsilon_i \varepsilon_j \middle| \mathrm{X}\right) = 0$ para $i \neq j$ por supuesto 4

$$E(\varepsilon M \varepsilon|X) = \sigma^2 \sum_{j=1}^{n} m_{ii} = \sigma^2 \cdot \mathrm{traza(M)}$$

Insesgadez de $\hat{\sigma}^2$

Demostración

2. Probar que $traza(M) = n - K$

$$traza(M) = traza(I_n - P)$$ Ya que $P = X(X'X)^{-1}X'$

$$traza(M) = traza(I_n) - traza(P)$$ El operador de la traza es lineal

$$traza(M) = n - traza(P)$$

y

$$traza(P) = traza[X(X'X)^{-1}X']$$ Ya que $P = X(X'X)^{-1}X'$

$$traza(P) = traza[(X'X)^{-1}X'X]$$ Ya que $traza(AB) = traza(BA)$

$$traza(P) = traza[I_K]$$

$$traza(P) = K$$

Por lo tanto, $traza(M) = n - K$

3. Perturbaciones Esféricas

Supuesto: Perturbaciones Esféricas

❖ **Homocedasticidad**

Asume que todas las unidades tienen el mismo error de varianza

$$E\left[\varepsilon_i^2 \middle| X\right] = \sigma_0^2 > 0 \qquad i = 1,2,\ldots,n$$

❖ **No Correlación**

Asume también que no hay correlación entre las observaciones

$$E\left[\varepsilon_i \varepsilon_j \middle| X\right] = 0 \qquad i,j = 1,2,\ldots,n \qquad i \neq j$$

Perturbaciones Esféricas

Este supuesto señala que la varianza de todos los términos de errores es constante.

Para ver esto:

$$Var(\varepsilon_i|X) = E\left(\varepsilon_i^2|X\right) - [E(\varepsilon_i|X)]^2 \qquad \textbf{(Por varianza condicional)}$$

$$Var(\varepsilon_i|X) = E\left(\varepsilon_i^2|X\right) \qquad \textbf{(Por exogeneidad estricta)}$$

Similarmente, la covarianza de la distribución conjunta de $(\varepsilon_i\varepsilon_j)$ condicional a X es cero.

$$Cov\left(\varepsilon_i\varepsilon_j\middle|X\right) = 0 \qquad (i, j = 1,2, \dots, n \qquad i \neq j)$$

$$E\left(\varepsilon_i\varepsilon_j\middle|X\right) - E(\varepsilon_i|X)E(\varepsilon_j|X) = 0$$

$$E\left(\varepsilon_i\varepsilon_j\middle|X\right) = 0$$

Perturbaciones Esféricas

En notación matricial:

$$E(\varepsilon\varepsilon'|X) = \begin{pmatrix} E(\varepsilon_1\varepsilon_1|X) & E(\varepsilon_1\varepsilon_2|X) & \cdots & E(\varepsilon_1\varepsilon_n|X) \\ E(\varepsilon_2\varepsilon_1|X) & E(\varepsilon_2\varepsilon_2|X) & \cdots & E(\varepsilon_2\varepsilon_n|X) \\ \vdots & \vdots & \ddots & \vdots \\ E(\varepsilon_n\varepsilon_1|X) & E(\varepsilon_n\varepsilon_2|X) & \cdots & E(\varepsilon_n\varepsilon_n|X) \end{pmatrix}$$

$$E(\varepsilon\varepsilon'|X) = \begin{pmatrix} \sigma_0^2 & 0 & \cdots & 0 \\ 0 & \sigma_0^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_0^2 \end{pmatrix} = \sigma_0^2 I_n$$

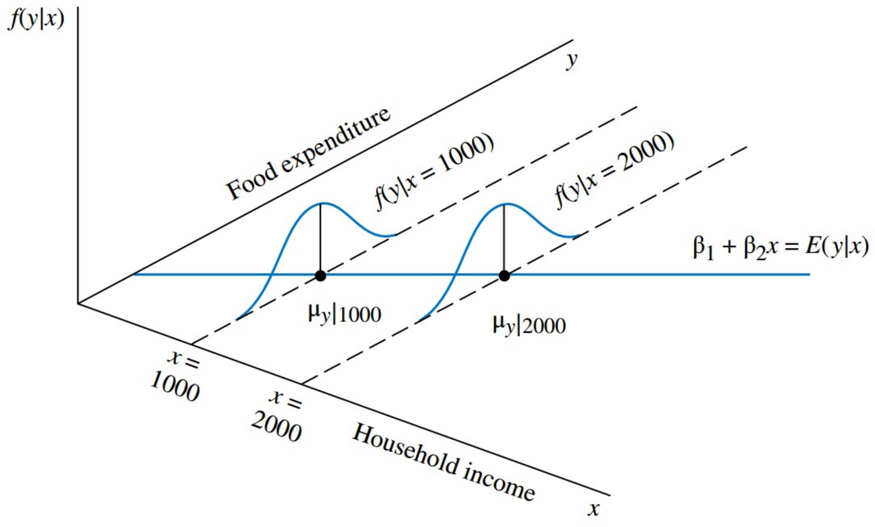
Los términos de errores que cumplen los supuestos de homocedasticidad y no autocorrelación son llamados "**perturbaciones o errores esféricas**".

Perturbaciones Esféricas

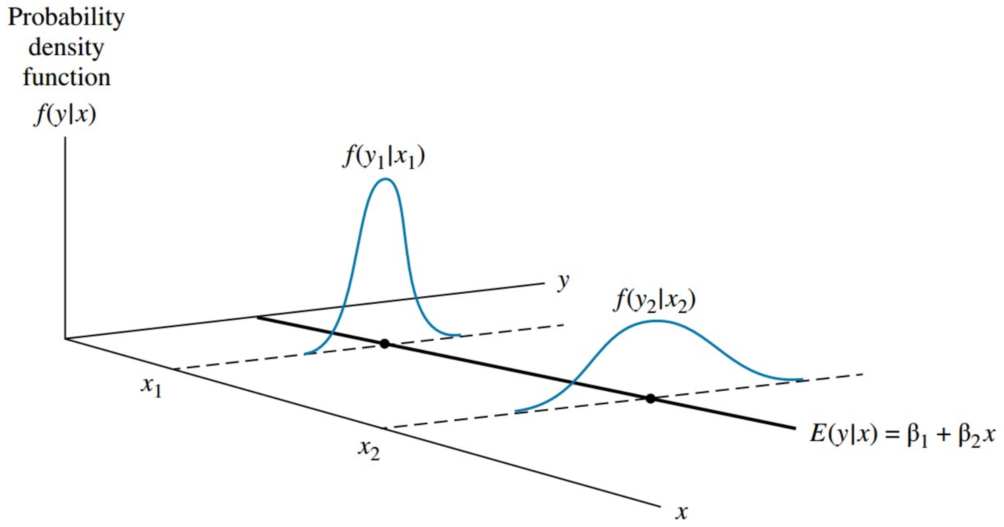
- ¿Qué consecuencias tiene el no cumplimiento de este supuesto?
- ¿Cómo afecta las propiedades estadísticas del estimador de MCO?

Supuesto de Homocedasticidad

Varianza del error no observable ε condicionado sobre las variables explicativas es constante.



Homocedasticidad



Heterocedasticidad

4. Perturbaciones No Esféricas

Perturbaciones No Esféricas - Definición

Recordemos que hemos asumido que:

$$Var(\varepsilon|X) = \sigma_0^2 I_n = \begin{pmatrix} \sigma_0^2 & 0 & \cdots & 0 \\ 0 & \sigma_0^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_0^2 \end{pmatrix}$$

Eliminando el supuesto de homocedasticidad:

$$Var(\varepsilon|X) = E(\varepsilon\varepsilon'|X) = \sigma_0^2 \Omega(X)$$

Perturbaciones No Esféricas - Definición

La matriz $\Omega(X)$ es no singular y conocida. Esta matriz puede contener elementos distintos en su diagonal principal y elementos diferentes de cero fuera de su diagonal:

$$\Omega(X) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{12} & \sigma_{2n} & \dots & \sigma_n^2 \end{bmatrix}$$

Es decir, la varianza no es necesariamente constante y la covarianza entre diferentes perturbaciones podría ser diferente de cero.

Consecuencias de relajar el supuesto: Insesgadez

Evaluar insesgadez:

$$E\left(\hat{\beta}_{MCO}\middle|X\right) =?$$

Demostración?

Consecuencias de relajar el supuesto: Insesgadez

Demostración

Tenemos:

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

$$\hat{\beta} = (X'X)^{-1}X'(X\beta_0 + \varepsilon)$$

$$\hat{\beta} = (X'X)^{-1}X'X\beta_0 + (X'X)^{-1}X'\varepsilon$$

$$\hat{\beta} = \beta_0 + (X'X)^{-1}X'\varepsilon$$

Luego:

$$E(\hat{\beta}_{MCO}|X) = \beta + (X'X)^{-1}X'E(\varepsilon|X)$$

$$E(\hat{\beta}_{MCO}|X) = \beta$$

Sigue siendo **insesgado**.

Consecuencias de relajar el supuesto: Varianza

Recordemos que la matriz de varianza-covarianza homocedástica es:

$$Var(\hat{\beta}_{MCO}) = \sigma_0^2 (X'X)^{-1}$$

Demostración?

Consecuencias de relajar el supuesto: Varianza

Demostración: Varianza Homocedástica

Tenemos:

$$Var(\hat{\beta}|X) = E(\hat{\beta} - E[\hat{\beta}]|X)^2 \quad \textbf{(Por varianza condicional)}$$

$$Var(\hat{\beta}|X) = E(\hat{\beta} - \beta_0|X)^2 \quad \textbf{(Por insesgadez)}$$

$$Var(\hat{\beta}|X) = E(A\varepsilon|X)^2 \quad \textbf{(Por sampling error)}$$

$$Var(\hat{\beta}|X) = E(A\varepsilon\varepsilon' A'|X) \quad \textbf{(Forma cuadrática)}$$

$$Var(\hat{\beta}|X) = AE(\varepsilon\varepsilon'|X)A' \quad \textbf{(A es una función de X)}$$

$$Var(\hat{\beta}|X) = A(\sigma_0^2 I_n)A' \quad \textbf{(Por homocedasticidad)}$$

$$Var(\hat{\beta}|X) = \sigma_0^2 AA'$$

$$\boldsymbol{Var(\widehat{\beta}|X) = \sigma_0^2(X'X)^{-1}}$$

Consecuencias de relajar el supuesto: Varianza

Evaluar matriz de varianza-covarianza no homocedástica:

$$Var(\hat{\beta}_{MCO}) = ?$$

Demostración?

Consecuencias de relajar el supuesto: Varianza

Demostración: Varianza No Homocedástica

Tenemos:

$$Var(\hat{\beta}_{MCO}|X) = E[(\hat{\beta} - E(\hat{\beta}))\left(\hat{\beta} - E(\hat{\beta})\right)'|X]$$

$$Var(\hat{\beta}_{MCO}|X) = E(\hat{\beta} - E(\hat{\beta})|X)^2$$

$$Var(\hat{\beta}_{MCO}|X) = E(\hat{\beta} - \beta_0|X)^2$$

$$Var(\hat{\beta}|X) = E(A\varepsilon|X)^2 \qquad \textbf{(Por sampling error)}$$

$$Var(\hat{\beta}_{MCO}|X) = E(A\varepsilon\varepsilon'A'|X) \qquad \textbf{(Forma cuadrática)}$$

$$Var(\hat{\beta}_{MCO}|X) = AE(\varepsilon\varepsilon'|X)A' \qquad \textbf{(A es una función de X)}$$

$$Var(\hat{\beta}_{MCO}|X) = (X'X)^{-1}X'E(\varepsilon\varepsilon')X(X'X)^{-1}$$

$$\boldsymbol{Var(\hat{\beta}_{MCO}|X) = (X'X)^{-1}X'\Omega X(X'X)^{-1}}$$

Esta difiere de la matriz de varianza-covarianza homocedástica

Consecuencias de relajar el supuesto

El estimador de σ^2, $\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-K}$, ya no es insesgado. Tarea: **Demostrarlo**.

El Teorema de Gauss Markov no se mantiene, por lo tanto el estimador MCO no son MELI. ¿Por qué?

El estadístico t y F no son válidas. Los estadísticos utilizados en los intervalos de confianza y en los contrastes de hipótesis ya no siguen las distribución señaladas, de modo que carecen de validez.

5. Mínimos Cuadrados Generalizados

Mínimos Cuadrados Generalizados

- Dado los fallos que ocurren en las propiedades de los estimadores MCO, surge la conveniencia de buscar estimadores alternativos que verifiquen mejores propiedades que los de MCO.

- Este es el caso de los estimadores de Mínimos Cuadrados Generalizados (MCG)

- Para construirlos basta observar una propiedad del nuevo modelo, que depende de la descomposición de la matriz de varianza-covarianza de las perturbaciones.

Modelo Transformado

- Dado que Ω por construcción es una matriz simétrica y definida positiva, podemos encontrar una matriz C no singular (cuadrada) y de orden n tal que:

$$\Omega^{-1} = C'C$$

- La inversa de Ω también debe ser una matriz positiva definida y simétrica, y C es una matriz $n \times n$ no singular. Es posible reescribir:

$$\Omega = (C'C)^{-1} = C^{-1}(C')^{-1}$$

$$C\Omega C' = CC^{-1}(C')^{-1}\ C' = I$$

Modelo Transformado

- Crear un nuevo modelo de regresión transformando (y, X, ε) por C como:

$$\tilde{y} \equiv Cy \qquad \tilde{X} \equiv CX \qquad \tilde{\varepsilon} \equiv C\varepsilon$$

- Por supuesto de linealidad para (y, X, ε) implica que $(\tilde{y}, \tilde{X}, \tilde{\varepsilon})$ también satisface linealidad:

$$\tilde{y} = \tilde{X}\beta_0 + \tilde{\varepsilon}$$

- Supongamos que la matriz C, que aún no sabemos como es, cumple su objetivo, teniendo entonces que $E\left(\tilde{\varepsilon}\middle|\tilde{X}\right) = 0$ y $Var\left(\tilde{\varepsilon}\middle|\tilde{X}\right) = \sigma_0^2 I_n$.

Modelo Transformado

Este modelo transformado cumple supuesto de exogeneidad estricta:

$$E\left(\tilde{\varepsilon}\middle|\tilde{X}\right) = ?$$

Demostración: Exogeneidad Estricta

Tenemos:

$$\begin{aligned} E\left(\tilde{\varepsilon}\middle|\tilde{X}\right) &= E(\tilde{\varepsilon}|X) \\ &= E(C\varepsilon|X) \\ &= CE(\varepsilon|X) \\ &= 0 \end{aligned}$$

Modelo Transformado

Supuesto de homocedasticidad se cumple para el modelo transformado:

$$E\left(\tilde{\varepsilon}\tilde{\varepsilon}'\middle|\tilde{X}\right) = ?$$

Demostración: Homocedasticidad

Tenemos:

$$\begin{aligned} E\left(\tilde{\varepsilon}\tilde{\varepsilon}'\middle|\tilde{X}\right) &= E(\tilde{\varepsilon}\tilde{\varepsilon}'|X) \\ &= CE(\varepsilon\varepsilon'|X)C' \\ &= C\sigma_0^2\Omega C' \\ &= \sigma_0^2 C\Omega C' \\ &= \sigma_0^2 I_n \end{aligned}$$

Finalmente, $\tilde{\varepsilon}|\tilde{X}$ es normal porque la distribución de $\tilde{\varepsilon}|\tilde{X}$ es la misma que $\tilde{\varepsilon}|X$ y $\tilde{\varepsilon}$ es una transformación lineal de ε

Estimador MCG

- Estimador de mínimo cuadrados generalizados

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

$$\hat{\beta}_{MCG} = [(CX)'(CX)]^{-1}(CX)'Cy$$

$$\hat{\beta}_{MCG} = (X'C'CX)^{-1}(X'C'Cy)$$

$$\boldsymbol{\hat{\beta}_{MCG} = (X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}y)}$$

Propiedades de MCG

Demostración: Insesgadez de $\widehat{\beta}_{MCG}$

Tomando el valor esperado de $\widehat{\beta}_{MCG}$:

$$E(\hat{\beta}_{MCG}|X) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}E(y|X)$$

$$E(\hat{\beta}_{MCG}|X) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}E(X\beta_0 + \varepsilon|X)$$

$$E(\hat{\beta}_{MCG}|X) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}X\beta_0$$

$$E(\hat{\beta}_{MCG}|X) = \beta_0$$

El estimador sigue siendo insesgado.

Propiedades de MCG

Su varianza condicional es:

$$Var(\hat{\beta}_{MCG}|X) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Var(y|X)\Omega^{-1}X'(X'\Omega^{-1}X)^{-1}$$

$$Var(\hat{\beta}_{MCG}|X) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}(\sigma^2\Omega)\Omega^{-1}X'(X'\Omega^{-1}X)^{-1}$$

$$Var(\hat{\beta}_{MCG}|X) = \sigma^2(X'\Omega^{-1}X)^{-1}$$

El estimador $\hat{\beta}_{MCG}$ es eficiente (mínima varianza).

Propiedades de MCG

❖ **Teorema de Aitken**

El estimador $\hat{\beta}_{MCG}$ es lineal, insesgado y su varianza es incluso menor que la varianza de MCO.

$$Var(\hat{\beta}_{MCG}|X) < Var(\hat{\beta}_{MCO}|X)$$

$$\sigma^2(X'\Omega^{-1}X)^{-1} < \sigma_0^2(X'X)^{-1}$$

Para el modelo de perturbaciones no esféricas, el estimador MCG es el de menor varianza dentro de la clase de estimadores lineales e insesgados.

Propiedades de MCG

- El estimador insesgado de σ^2 es:

$$\hat{\sigma}^2 = \frac{(y - X\hat{\beta}_{MCG})'\Omega^{-1}(y - X\hat{\beta}_{MCG})'}{n-k}$$

- Se verifica que el estadístico:

$$\frac{\hat{\beta}_{MCG} - \beta}{\sqrt{\widehat{Var}(\hat{\beta}_{MCG})}} \sim t(n-k)$$

Inferencia

- Dicho estadístico se puede utilizar para obtener el intervalo de confianza para β a un nivel de confianza $1-\alpha$:

$$\hat{\beta}_{MCG} \pm t_{n-k;1-\frac{\alpha}{2}} \sqrt{\widehat{Var}(\hat{\beta}_{MCG})}$$

- Así como resolver contrastes de hipótesis sobre β
- Si se quiere contrastar:

$$H_0 : R\beta = r$$

$$H_1 : R\beta \neq r$$

- El estadístico de contraste es:

$$F = \frac{\left(R\hat{\beta}_{MCG} - r\right)' [R(X'\Omega^{-1}X)R']^{-1} \left(R\hat{\beta}_{MCG} - r\right) / \#r}{q\hat{\sigma}^2_{MCG}} \sim F(q, n-k)$$

De forma que se rechaza H_0 si $F_{exp} > F_{q,n-k;1-\alpha}$.

6. Mínimos Cuadrados Generalizados Factibles

¿Y si Ω no es conocida?

- En la práctica, la matriz Ω raramente es conocida, por lo que el estimador MCG no se puede calcular.
- La alternativa es sustituir esta matriz por un estimador suyo $\widehat{\Omega}$, obteniendo así el estimador de mínimos cuadrados generalizados factibles (MCGF)

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

- Como consecuencia, las propiedades que satisface el vector de estimadores MCG ya no se mantienen. De esta forma, el estimador MCGF, en general, no es insesgado ni óptimo.

PUCP