TEMA 2: *Propiedades de los estimadores MCO*

Econometría I

M. Angeles Carnero

Departamento de Fundamentos del Análisis Económico

Curso 2011-12

Propiedades estadísticas de $\hat{\beta}$

1. Es un estimador lineal. $\hat{\beta}$ es una función lineal de Y al ser X una matriz de constantes (dado el Supuesto 1):
   $\hat{\beta} = (X'X)^{-1} X'Y = WY$

2. Bajo las hipótesis básicas 1 a 4, $\hat{\beta}$ es un estimador insesgado de β, es decir, $E\left(\hat{\beta}\right) = \beta$ ya que

$$\widehat{\beta} = (X'X)^{-1}X'Y = \beta + (X'X)^{-1}X'u$$

   y por tanto

$$E\left[\widehat{\beta}\right] = \beta + (X'X)^{-1}X'E\left[u\right] \underset{\text{puesto que } E[u]=0}{=} \beta$$

   Nótese que el estimador MCO es insesgado con independencia de que se verifique o no el supuesto 5.

3. Bajo las hipótesis básicas del MRL, $Var\left(\hat{\beta}\right) = \sigma^2 \left(X'X\right)^{-1}$ ya que:

- $$\begin{aligned} Var\left[\widehat{\beta}\right] &= E\left\{\left[\widehat{\beta} - E\left[\widehat{\beta}\right]\right]\left[\widehat{\beta} - E\left[\widehat{\beta}\right]\right]'\right\} = E\left\{\left[\widehat{\beta} - \beta\right]\left[\widehat{\beta} - \beta\right]'\right\} \\ &= E\left\{\left[(X'X)^{-1}X'u\right]\left[(X'X)^{-1}X'u\right]'\right\} = \\ &= (X'X)^{-1}X'E(uu')X(X'X)^{-1} = \\ &\underset{\text{Ya que } E(uu')=\sigma^2 I_T}{=} \sigma^2\left(X'X\right)^{-1} \end{aligned}$$

- Teorema de Gauss-Markov:

  Bajo las hipótesis básicas del MRL, el estimador MCO de β es óptimo entre la familia de estimadores lineales e insesgados. Es decir, no es posible encontrar otro estimador de β que siendo lineal e insesgado tenga una varianza menor que el estimador MCO.

Estimación de σ^2 y propiedades estadísticas de $\hat{\sigma}^2$

1. El vector de residuos MCO es $e = Y - \hat{Y} = Y - X\hat{\beta}$. Puede interpretarse como la estimación del vector de errores u.
2. El vector de residuos MCO es una transformación lineal de u: $e = Y - X\hat{\beta} = Y - X\left(X'X\right)^{-1}X'Y = \left[I - X\left(X'X\right)^{-1}X'\right]Y = MY = Mu$ puesto que M es una matriz $M = \left[I - X\left(X'X\right)^{-1}X'\right]$ que cumple las siguientes propiedades:
   1. M es una matriz singular: $|M| = \det(M) = 0$ puesto que $Rg\left(M\right) = Tr\left(M\right) = Tr\left[I_T - X\left(X'X\right)^{-1}X'\right] = Tr\left(I_T\right) - Tr\left[X\left(X'X\right)^{-1}X'\right] = Tr\left(I_T\right) - Tr\left[\left(X'X\right)^{-1}X'X\right] = Tr\left(I_T\right) - Tr\left(I_k\right) = T - k < T$
   2. M es una matriz simétrica: $M = M'$
   3. M es una matriz idempotente: $M = M \cdot M$
   4. M es ortogonal a X: $MX = \left[I - X\left(X'X\right)^{-1}X'\right]X = X - X\left(X'X\right)^{-1}X'X = X - X = 0$

3. $E(e) = 0$ puesto que

$$E[e] = E[Mu] = ME[u] \underset{\text{ya que } E[u]=0}{=} 0$$

4. $Var(e) = \sigma^2 M$ puesto que

$$Var(e) = Var(Mu) = MVar(u)M' = M\sigma^2 IM' = \sigma^2 MM' = \sigma^2 M$$

5. Estimador de σ^2: La varianza de los errores, σ^2, es un valor poblacional junto a β. Es necesario estimarlo para contrastar hipótesis acerca de β o establecer intervalos de confianza. **Intuición:**

$$\sigma^2 = E(u_t^2) \Rightarrow \tilde{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} u_t^2$$

$$\check{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} e_t^2 = \frac{1}{T}e'e$$

$\Downarrow$ (para que sea insesgado)

$$\boxed{\hat{\sigma}^2 = \frac{1}{T-K}\sum_{t=1}^{T} e_t^2 = \frac{1}{T-K}e'e}$$

6. $\hat{\sigma}^2 = \frac{1}{T-k}\sum_{t=1}^{T} e_t^2 = \frac{e'e}{T-k}$ ($T-k$ son los grados de libertad) es un estimador insesgado de σ^2 puesto que

$$E\left(\hat{\sigma}^2\right) = E\left(\frac{e'e}{T-k}\right) = \frac{1}{T-k}E\left(e'e\right) = \frac{\sigma^2\left(T-k\right)}{T-k} = \sigma^2$$

7. Otra expresión de $e'e$:

$$e'e = \left(Y - X\hat{\beta}\right)'\left(Y - X\hat{\beta}\right) = Y'Y - \hat{\beta}'X'Y$$

Matriz de varianzas estimada de $\hat{\beta}$ y errores estándar

Hemos visto que bajo las hipótesis 1 a 5

$$Var\left[\widehat{\beta}\right] = \sigma^2 \left(X'X\right)^{-1}$$

esta matriz es desconocida ya que σ^2 es desconocido. Para saber la fiabilidad de $\widehat{\beta}$ y poder hacer inferencia es importante disponer de un estimador de su varianza. Se define la matriz de varianzas estimada de $\widehat{\beta}$ como

$$\widehat{Var}\left[\widehat{\beta}\right] = \widehat{\sigma}^2 \left(X'X\right)^{-1}$$

En el tema 3 veremos cómo contrastar hipótesis sobre el vector de parámetros β utilizando $\widehat{\beta}$ y $\widehat{Var}\left[\widehat{\beta}\right]$. Nótese que si no se verifica la hipótesis 5, $Var\left[\widehat{\beta}\right] \neq \sigma^2 \left(X'X\right)^{-1}$ y por tanto $\widehat{\sigma}^2 \left(X'X\right)^{-1}$ no sería un estimador apropiado de la varianza de $\widehat{\beta}$.

Se definen los errores estándar como las raices cuadradas de los elementos de la diagonal principal de la matriz $\widehat{Var}\left[\widehat{\beta}\right]$. Es decir

$$SE(\widehat{\beta}_j) = \sqrt{\widehat{\sigma}^2 \left(X'X\right)^{-1}_{jj}} \quad j = 1, .., k$$

donde $\widehat{\beta}_j$ es el elemento j del vector $\widehat{\beta}$ y $\left(X'X\right)^{-1}_{jj}$ es el elemento (j,j) de la matrix $\left(X'X\right)^{-1}$. $SE(\widehat{\beta}_j)$ es un estimador de la desviación típica de $\widehat{\beta}_j$.

- Nota: Si cambiamos las unidades de medida de alguna o algunas de las variables explicativas y/o de la variable dependiente cada uno de los errores estándar variará en la misma proporción que el valor estimado del parámetro correspondiente. Por ejemplo:

$$\begin{aligned} \widehat{\beta}_2^* &= c\widehat{\beta}_2 \\ \widehat{Var}\left[\widehat{\beta}_2^*\right] &= c^2 \widehat{Var}\left[\widehat{\beta}_2\right] \\ &\Downarrow \\ SE\left(\widehat{\beta}_2^*\right) &= cSE\left(\widehat{\beta}_2\right) \end{aligned}$$

Propiedad de la distribución normal multivariante

Si X es un vector $n \times 1$, $X \sim N\left[\mu, \Sigma\right]$, A es una matriz $r \times n$ $(r \leq n)$ no aleatoria y b es un vector $r \times 1$ no aleatorio, entonces:

(i) $AX + b \sim N\left[A\mu + b, A\Sigma A'\right]$

(ii) En particular $\Sigma^{-1/2}\left(X - \mu\right) \sim N\left[0, I_n\right]$

- <u>Definición 1:</u> *Chi-cuadrado con 1 grado de libertad*

  Si $Z \sim N(0,1)$, entonces $Z^2 \sim \chi^2_1$.

  Nota: $E\left(Z^2\right) = 1, Var\left(Z^2\right) = 2$

- <u>Definición 2:</u> *Chi-cuadrado con n grados de libertad*

  Si $Z_1, Z_2, ..., Z_n$ son n variables aleatorias **independientes e idénticamente distribuidas** (iid) como $N\left[0,1\right]$, entonces $\sum_{i=1}^{n} Z_i^2 \sim \chi^2_n$.

- <u>Definición 3:</u> *t de Student con n grados de libertad*

  Si Z y X son variables aleatorias independientes, $Z \sim N\left[0,1\right]$ y $X \sim \chi^2_n$, entonces

$$\frac{Z}{\sqrt{\frac{X}{n}}} \sim t_n$$

- Definición 4: *F de Snedecor con n y m grados de libertad*

  Si X_1 y X_2 son variables aleatorias independientes $X_1 \sim \chi^2_n$ y $X_2 \sim \chi^2_m$ entonces
  $$\frac{\frac{X_1}{n}}{\frac{X_2}{m}} \sim F_{n,m}$$

- Teorema 1: *Suma de chi-cuadrados*

  Sean X_1 y X_2 dos variables aleatorias independientes con distribución $X_1 \sim \chi^2_{n_1}$ y $X_2 \sim \chi^2_{n_2}$, entonces $X_1 + X_2 \sim \chi^2_{n_1+n_2}$

- Teorema 2:

  Si $Z_1, Z_2, ..., Z_n$ son n variables aleatorias independientes e idénticamente distribuidas (iid) como $N\left[0, \sigma^2\right]$, entonces $\sum_{i=1}^{n} \left(\frac{Z_i}{\sigma}\right)^2 \sim \chi^2_n$.

- Teorema 3: *Distribución de formas cuadráticas de matrices idempotentes en vectores normales estandarizados.*

  Sea $X \sim N(0, I_n)$ de dimensión $(n \times 1)$, A una matriz simétrica e idempotente de dimensión $(n \times n)$, entonces

  $$X'AX \sim \chi^2_J$$

  donde J=$rg(A) = tr(A)$

- Teorema 4: *Independencia de dos formas cuadráticas con matrices idempotentes en un mismo vector normal estandarizado.*

  Sea $X \sim N(0, I_n)$ y A y B dos matrices idempotentes de dimensión $(n \times n)$ tales que $AB = 0$, entonces las dos formas cuadráticas $X'AX$ y $X'BX$ son independientes.

- **EJEMPLO:** Sea $X \sim N(0, I_n)$ y A y B dos matrices $n \times n$ idempotentes de rango n_A y n_B, respectivamente. Utilizando el Teorema 3

$$\begin{aligned} X'AX &\sim \chi^2_{n_A} \\ X'BX &\sim \chi^2_{n_B} \end{aligned}$$

Si $AB = 0$, utilizando el Teorema 4, $X'AX$ y $X'BX$ son independientes y por tanto

$$\frac{(X'AX)\,/n_A}{(X'BX)\,/n_B} \sim F_{n_A, n_B}$$

- Teorema 5: *Independencia de una forma lineal y una forma cuadrática idempotente de un vector normal estandarizado.*

  Sea $X \sim N(0, I_n)$ y sea L una matriz $r \times n$ y A una matriz $n \times n$ idempotente tales que $LA = 0$, entonces la función lineal LX y la forma cuadrática $X'AX$ son independientes.

- **EJEMPLO:** Sea $X \sim N(0, I_n)$, A una matriz $n \times n$ idempotente de rango n_A y L un vector $n \times 1$ tal que $L'L = 1$. Como $X \sim N(0, I_n) \Rightarrow L'X \sim N(0, L'L) = N(0, 1)$ y $X'AX \sim \chi^2_{n_A}$. Si $L'A = 0$, utilizando el Teorema 5, $L'X$ y $X'AX$ son independientes y por tanto

$$\frac{L'X}{\sqrt{(X'AX)/n_A}} \sim t_{n_A}$$

- Teorema 6: *Distribución de formas cuadráticas de matrices de rango completo en vectores normales.*

  Sea X un vector $n \times 1$, $X \sim N[\mu, \Sigma]$, entonces:
  (i) $\Sigma^{-1/2}(X-\mu) \sim N[0, I_n]$
  (ii) $(X-\mu)' \Sigma^{-1} (X-\mu) \sim \chi^2_n$

Propiedades de los estimadores MCO con errores normales

Con la hipótesis adicional de normalidad de los errores, se puede calcular la distribución exacta de $\hat{\beta}$ y $\widehat{\sigma}^2$. Nótese que la media y la varianza de $\hat{\beta}$ se obtuvieron previamente sin necesidad de imponer esta hipótesis aunque obviamente la distribución, sin hacer este supuesto, es desconocida.

Si $u \sim N(0, \sigma^2 I_T)$ y dado que $\widehat{\beta} = \beta + (X'X)^{-1}X'u$, entonces

$$\underset{k\times 1}{\widehat{\beta}} \sim N(\underset{k\times 1}{\beta}, \sigma^2 \underset{k\times k}{(X'X)^{-1}})$$

La distribución marginal de cada elemento del vector $\widehat{\beta}$ es también normal:

$$\widehat{\beta}_i \sim N(\beta_i, \sigma^2 (X'X)^{-1}_{ii}) \text{ para } i = 1, ..., k$$

donde $\widehat{\beta}_i$ es el elemento $(i, 1)$ del vector $\widehat{\beta}$, β_i es el elemento $(i, 1)$ del vector β y $(X'X)^{-1}_{ii}$ es el elemento (i, i) de la matriz $(X'X)^{-1}$.
Del mismo modo se puede comprobar que bajo la hipótesis adicional de normalidad se tiene que:

$$\begin{aligned} Y &= X\beta + u \sim N(X\beta, \sigma^2 I_T) \\ \hat{Y} &= X\widehat{\beta} \sim N(X\beta, \sigma^2 X(X'X)^{-1}X') \end{aligned}$$

- *Distribución de $\widehat{\sigma}^2$ bajo el supuesto de normalidad de los errores u*

  Si $u \sim N(0, \sigma^2 I_T)$, entonces $\frac{\widehat{\sigma}^2 (T-k)}{\sigma^2} \sim \chi^2_{(T-k)}$

Demostración:

Dado que $\widehat{\sigma}^2 = \frac{e'e}{(T-k)}$, queremos demostrar que $\frac{e'e}{\sigma^2} \sim \chi^2_{(T-k)}$. Sabemos que

$$\frac{e'e}{\sigma^2} = \frac{u'Mu}{\sigma^2} = \left(\frac{u}{\sigma}\right)' M \left(\frac{u}{\sigma}\right) \text{ y } \left(\frac{u}{\sigma}\right) \sim N(0, I_T).$$

y M es una matriz idempotente de rango $T-k$, entonces por el Teorema 3, $\frac{e'e}{\sigma^2} \sim \chi^2_{(T-k)}$

- *Independencia de $\hat{\beta}$ y $\widehat{\sigma}^2$ bajo el supuesto de normalidad de los errores u*

  Si $u \sim N(0, \sigma^2 I_T)$, entonces $\hat{\beta}$ y $\widehat{\sigma}^2$ son independientes entre sí.

Demostración:

Nótese que:

$\frac{(\hat{\beta}-\beta)}{\sigma} = (X'X)^{-1}X'\left(\frac{u}{\sigma}\right)$ $\rightarrow$forma lineal en $\left(\frac{u}{\sigma}\right)$

$\frac{\widehat{\sigma}^2(T-k)}{\sigma^2} = \left(\frac{u}{\sigma}\right)' M \left(\frac{u}{\sigma}\right)$ $\rightarrow$forma cuadrática de M y en $\left(\frac{u}{\sigma}\right)$

Entonces, $\widehat{\sigma}^2$ y $\hat{\beta}$ independientes$\Longleftrightarrow$ $\frac{\widehat{\sigma}^2(T-k)}{\sigma^2}$ y $\frac{(\hat{\beta}-\beta)}{\sigma}$ independientes $\underset{\text{Teorema 5}}{\Leftarrow}$ $(X'X)^{-1}X'M = 0$