PUBLICACIONES DE 4er CURSO

Grado: Economía

Asignatura: ECONOMETRÍA III

Tema 2: Estimación. Ejercicios Resueltos

Grupos:

Profesores: Antonio Aznar y M. Teresa Aparicio

Departamento de ANÁLISIS ECONÓMICO

Curso Académico 2014/15

Facultad de
Economía y Empresa
Universidad Zaragoza

Ejercicios Resueltos

Apartado 2.1

Ejercicio 2.1.1

1) En el marco del modelo $y_t = \alpha + \beta x_t + u_t$ en donde u cumple las hipótesis ideales y también se cumple que $\frac{1}{T}\sum x_t$, $\sum \frac{1}{x_t}$, $\frac{1}{T}\sum \frac{1}{x_t^2}$ convergen a constantes conforme crece el tamaño muestral, se definen los dos estimadores siguientes:

$$\widehat{\beta}_1 = \frac{\sum y_t - T\alpha}{\sum x_t} \qquad \text{y} \qquad \widehat{\beta}_2 = \frac{\sum \frac{y_t}{x_t}}{T}$$

Obtener la media y varianza de ambos estimadores y demostrar que son consistentes.

SOLUCIÓN

Se tiene que

$$\widehat{\beta}_1 = \frac{\sum y_t - T\alpha}{\sum x_t} = \frac{T\alpha + \beta \sum x_t + \sum u_t - T\alpha}{\sum x_t} = \beta + \frac{\sum u_t}{\sum x_t}$$

Por lo que

$$E\widehat{\beta}_1 = \beta \quad por \quad ser \quad E\sum u_t = 0$$

$$Var(\widehat{\beta}_1) = E(\widehat{\beta}_1 - \beta)^2 = E\left(\frac{\sum u_t}{\sum x_t}\right)^2 = \ldots = \frac{T\sigma^2}{\left(\sum x_t\right)^2}$$

El estimador es consistente porque es insesgado y porque la varianza tiende a cero conforme el tamaño muestral crece.

$$LimVar(\widehat{\beta}_1) = Lim\frac{T\sigma^2}{\left(\sum x_t\right)^2} = Lim\frac{\sigma^2}{T\left(\sum x_t / T\right)^2} = 0$$

Con respecto al segundo estimador podemos escribir,

$$\widehat{\beta}_2 = \frac{\sum\frac{y_t}{x_t}}{T} = \frac{\alpha}{T}\sum\frac{1}{x_t} + \beta + \frac{1}{T}\sum\frac{u_t}{x_t}$$

Por lo tanto

$$E(\widehat{\beta}_2) = E\frac{\sum\frac{y_t}{x_t}}{T} = \frac{\alpha}{T}\sum\frac{1}{x_t} + \beta \qquad \text{es sesgado}$$

$$Var(\widehat{\beta}_2) = E(\frac{1}{T}\sum\frac{u_t}{x_t})^2 = \frac{\sigma^2}{T}\frac{1}{T}\sum\frac{1}{x_t^2}$$

Este estimador también es consistente porque aunque es sesgado el sesgo tiende a cero y la varianza también tiende a cero conforme el tamaño muestral crece.

Ejercicio 2.1.2

En el marco del modelo $y_t = \beta x_t + u_t$ en donde u cumple las hipótesis ideales se definen dos estimadores de β: el MCO y el siguiente estimador

$$\widehat{\beta}_1 = \frac{\sum y_t}{\sum x_t}$$

a). Obtener la media y varianza de ambos estimadores e indicar las condiciones que debe cumplir el regresor para que ambos estimadores sean consistentes.

b). Definir los residuos correspondientes a cada estimador y, para la observación t-ésima, derivar la media y varianza del correspondiente residuo. Demostrar que la suma de cuadrados de los residuos MCO es menor que la suma de los cuadrados de los residuos que corresponden a $\widehat{\beta}_1$.

SOLUCIÓN

a). Se tiene que

$$\widehat{\beta}_1 = \frac{\sum y_t}{\sum x_t} = \frac{\beta \sum x_t + \sum u_t}{\sum x_t} = \beta + \frac{\sum u_t}{\sum x_t}$$

Por lo que

$$E\widehat{\beta}_1 = \beta \quad por \quad ser \quad E\sum u_t = 0$$

$$Var(\widehat{\beta}_1) = E(\widehat{\beta}_1 - \beta)^2 = E\left(\frac{\sum u_t}{\sum x_t}\right)^2 = \ldots = \frac{T\sigma^2}{\left(\sum x_t\right)^2}$$

Un estimador es consistente si tanto el sesgo como la varianza tienden a cero conforme el tamaño muestral tiende a infinito. Para este primer estimador la primera condición se cumple por ser insesgado mientras que la varianza tenderá a cero según sea el comportamiento asintótico del sumatorio de los valores del regresor. Por ejemplo, basta que este sumatorio dividido por T tienda a una constante. En este caso tendríamos:

$$LimVar(\widehat{\beta}_1) = Lim\frac{T\sigma^2}{\left(\sum x_t\right)^2} = Lim\frac{\sigma^2}{T\left(\sum x_t / T\right)^2} = 0$$

Con respecto al estimador MCO se tiene que,

$$\hat{\beta} = \frac{\sum x_t y_t}{\sum x_t^2} = \beta + \frac{\sum x_t u_t}{\sum x_t^2}$$

Por lo tanto

$$E\hat{\beta} = \beta \quad por \quad ser \quad E\sum x_t u_t = \sum x_t E u_t = 0$$

$$Var\left(\hat{\beta}\right) = E\left(\frac{\sum x_t u_t}{\sum x_t^2}\right) = \frac{1}{\left(\sum x_t^2\right)^2} E\left(\sum x_t u_t\right)^2 = \frac{\sigma^2}{\sum x_t^2}$$

Este estimador también es consistente porque es insesgado y la varianza tiende a cero ya que el promedio de los cuadrados del regresor tiende a una constante conforme el tamaño muestral crece.

b). El residuo del estimador propuesto es

$$\hat{u}_{1t} = y_t - \widehat{\beta}_1 x_t = \beta x_t + u_t - \beta x_t - \frac{\sum u_t}{\sum x_t} x_t = u_t - \frac{\sum u_t}{\sum x_t} x_t$$

Por lo tanto,

$E\widehat{u}_{1t} = 0$ por lo que es insesgado

$$Var(\widehat{u}_{1t}) = E\left(u_t - x_t \frac{\sum u_t}{\sum x_t} \right)^2 = E\left(u_t^2 + x_t^2 \frac{(\sum u_t)^2}{(\sum x_t)^2} - 2x_t \frac{u_t(\sum u_t)}{\sum x_t} \right)$$

$$= \left(\sigma^2 + x_t^2 \frac{T\sigma^2}{(\sum x_t)^2} - 2x_t \frac{\sigma^2}{\sum x_t} \right)$$

En lo que respecta al residuo MCO podemos escribir,

$$\hat{u}_t = y_t - \hat{\beta} x_t = \beta x_t + u_t - \beta x_t - \frac{\sum' x_t u_t}{\sum x_t^2} x_t = u_t - \frac{\sum x_t u_t}{\sum x_t^2} x_t$$

La esperanza es

$E\hat{u}_t = o$ por lo que es insesgado. Respecto a la varianza,

$$Var(u_t) = E\left(u_t - x_t \frac{\sum x_t u_t}{\sum x_t^2} \right)^2 = E\left(u_t^2 + x_t^2 \frac{(\sum x_t u_t)^2}{(\sum x_t^2)^2} - 2x_t \frac{u_t(\sum x_t u_t)}{\sum x_t^2} \right)$$

$$= \left(\sigma^2 + x_t^2 \frac{\sigma^2}{\sum x_t^2} - 2x_t^2 \frac{\sigma^2}{\sum x_t^2} \right) = \sigma^2 \left(1 - \frac{x_t^2}{\sum x_t^2} \right)$$

Respecto al último punto, basta particularizar el resultado 2.3.3 del Apartado 2.3. Podemos escribir,

$$\hat{u}_{1t} = y_t - \widehat{\beta}_1 x_t = y_t - \hat{\beta} x_t + \hat{\beta} x_t - \widehat{\beta}_1 x_t = \hat{u}_t + (\hat{\beta} - \widehat{\beta}_1) x_t$$

La suma de los cuadrados es,

$$\sum \widehat{u}_{1t}^2 = \sum \hat{u}_t^2 + (\hat{\beta} - \widehat{\beta}_1)^2 \sum x_t^2 + 2(\hat{\beta} - \widehat{\beta}_1) \sum x_t \hat{u}_t$$

El resultado se cumple porque el tercer término de la derecha es cero ya que, por la ecuación normal, se cumple que $\sum x_t \hat{u}_t = 0$.

Apartado 2.2

Ejercicio 2.2.1

Una muestra de T=5 observaciones con los siguientes valores {1,2,4,2,1} se ha obtenido siguiendo un proceso aleatorio a partir de una población con distribución normal

$$f\left(y_t;\theta\right)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{\left(y_t-\mu\right)^2}{2\sigma^2}\right]$$

Con parámetros $\mu\ y\ \sigma^2$.

(a) Suponer que la varianza es conocida e igual a la unidad y que el parámetro desconocido es $\theta=\{\mu\}$.
   (1) Derivar el logaritmo de la función de verosimilitud, el gradiente y la matriz de segundas derivadas.
   (2) Derivar y calcular el estimador MV, $\tilde{\theta}$.
   (3) Calcular $l\left(\tilde{\theta}\right), d\left(\tilde{\theta}\right)\ y\ H\left(\tilde{\theta}\right)$.

(b) Repetir (a) para el caso en que la varianza es desconocida, $\theta=\left\{\sigma^2\right\}$ y la media es conocida igual a 2.

(c) Repetir (a) cuando tanto la media como la varianza son desconocidos, $\theta=\left\{\mu,\sigma^2\right\}$.

Solución.

(a) (1)La función de verosimilitud puede escribirse como

$$L\left(\theta\right)=\left(2\pi\right)^{-\frac{T}{2}}\exp\left\{-\frac{\sum\left(y_t-\mu\right)^2}{2}\right\}$$

El logaritmo es

$$l(\theta) = -\frac{T}{2}\log(2\pi) - \frac{\sum_{1}^{T}(y_t-\mu)^2}{2}$$

El gradiente es

$$d(\theta) = \frac{\partial l(\theta)}{\partial \mu} = \sum(y_t-\mu)$$

En este caso, la matriz de segundas derivadas tiene un solo elemento

$$H(\theta) = -T$$

(2) El estimador MV de μ se obtiene igualando a cero el gradiente

$$\tilde{\mu} = \bar{y} = \frac{1+2+4+2+1}{5} = 2$$

(3) El valor estimado del logaritmo de la función de verosimilitud es

$$l(\tilde{\theta}) = -\frac{5}{2}\log(2\pi) - \frac{\sum(y_t-2)^2}{2} = -7,594$$

El valor del gradiente será cero y el valor de la segunda derivada será el mismo escrito anteriormente porque no depende del parámetro.

(b) (1)El logaritmo de la función de verosimilitud es

$$l(\theta) = -\frac{T}{2}\log(2\pi\sigma^2) - \frac{\sum(y_t-2)^2}{2\sigma^2}$$

y el gradiente toma la forma

$$d(\theta) = \frac{\partial l(\theta)}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}\sum(y_t-2)^2$$

El elemento de segunda derivada es

$$H(\theta) = \frac{\partial^2 l(\theta)}{\partial(\sigma^2)^2} = \frac{T}{2\sigma^4} - \frac{1}{\sigma^6}\sum(y_t-2)^2$$

(2) El estimador MV se obtiene igualando a cero el gradiente, $d\left(\tilde{\theta}\right)=0,$ siendo el estimador

$$\tilde{\sigma}^2=\frac{\sum\left(y_t-2\right)^2}{T}=$$

$$=\frac{\left(1-2\right)^2+\left(2-2\right)^2+\left(4-2\right)^2+\left(2-2\right)^2+\left(1-2\right)^2}{5}=1,2$$

(3) Los valores estimados son:

$$l\left(\tilde{\theta}\right)=-\frac{5}{2}\log\left(2\pi.1,2\right)-\frac{\sum\left(y_t-2\right)^2}{2.1,2}=\text{-7,55}$$

$$d\left(\tilde{\theta}\right)=0$$

$$H\left(\tilde{\theta}\right)=\frac{5}{2(1,2)^2}-\frac{1}{1,2^3}\sum\left(y_t-2\right)^2=\text{-1,736}$$

(c) (1) La función de verosimilitud es

$$L\left(\theta;y_1,y_2....,y_T\right)=\left(2\pi\sigma^2\right)^{-\frac{T}{2}}\exp\left[-\frac{\sum_1^T\left(y_t-\mu\right)^2}{2\sigma^2}\right]$$

Y su logaritmo

$$l\left(\theta\right)=-\frac{T}{2}\log\left(2\pi\sigma^2\right)-\frac{\sum_1^T\left(y_t-\mu\right)^2}{2\sigma^2}$$

El gradiente es un vector de dos elementos

$$d(\theta)=\begin{bmatrix}\dfrac{\partial l(\theta)}{\partial\mu}\\ \dfrac{\partial l(\theta)}{\partial\sigma^2}\end{bmatrix}=\begin{bmatrix}\dfrac{1}{\sigma^2}\sum(y_t-\mu)\\ -\dfrac{T}{2\sigma^2}+\dfrac{1}{2\sigma^4}\sum(y_t-\mu)^2\end{bmatrix}$$

Por último, la matriz de segundas derivadas es

$$H(\theta)=\begin{bmatrix}\dfrac{\partial^2 l(\theta)}{\partial\mu^2} & \dfrac{\partial^2 l(\theta)}{\partial\mu\partial\sigma^2}\\ & \dfrac{\partial^2 l(\theta)}{\partial\left(\sigma^2\right)^2}\end{bmatrix}=\begin{bmatrix}-\dfrac{T}{\sigma^2} & -\dfrac{1}{\sigma^4}\sum(y_t-\mu)\\ & \dfrac{T}{2\sigma^4}-\dfrac{1}{\sigma^6}\sum(y_t-\mu)^2\end{bmatrix}$$

(2) Los estimadores MV se obtienen igualando los elementos del gradiente a cero. Estos estimadores son

$$\tilde{\mu}=\frac{1}{T}\sum y_t \qquad \tilde{\sigma}^2=\frac{1}{T}\sum\left(y_t-\bar{y}\right)^2$$

Para obtener las estimaciones MV basta sustituir los cinco valores muestrales en estas expresiones

$$\tilde{\mu}=\frac{1+2+4+1+2}{5}=2$$

$$\tilde{\sigma}^2=\frac{(1-2)^2+(2-2)^2+(4-2)^2+(2-2)^2+(1-2)^2}{5}=1{,}2$$

(3) El logaritmo de la función de verosimilitud evaluado con las estimaciones MV es

$$l\left(\tilde{\theta}\right)=-\frac{T}{2}\log\left(2\pi\tilde{\sigma}^2\right)-\frac{\sum_1^T\left(y_t-\bar{y}\right)^2}{2\tilde{\sigma}^2}=$$

$$-\frac{5}{2}\log\left(2\pi\times1{,}2\right)-\frac{5}{2}=-2{,}55$$

Los elementos del gradiente evaluados con las estimaciones MV son cero porque es la condición necesaria para obtener los estimadores MV.
En lo que respecta a la matriz de segundas derivadas

$$H\left(\tilde{\theta}\right)=\begin{bmatrix} -\frac{5}{1,2} & -\frac{1}{1,2^2}\sum\left(y_t-\overline{y}\right) \\ . & \frac{5}{2(1,2)^2}-\frac{1}{1,2^3}\sum\left(y_t-\overline{y}\right)^2 \end{bmatrix}=\begin{bmatrix} -4,16 & 0 \\ 0 & -1,736 \end{bmatrix}$$

Ejercicio 2.2.2

Se ha obtenido una muestra de tamaño 5, $\left(x_1, x_2 .... x_5\right)$ siguiendo el muestreo aleatorio simple a partir de una población con media μ y varianza conocida, igual a la unidad. Se pide:

a). Escribir la función de verosimilitud de la muestra. Evaluar los elementos del gradiente utilizando los estimadores máximo verosímiles sin restringir y restringidos, suponiendo que la restricción especifica que la media es cero. Derivar el sesgo y la varianza del estimador restringido de la media. Especificar la región crítica que correspondería al contraste de los multiplicadores de Lagrange (LM) si se contrasta como hipótesis nula que la media es cero.

b). Derivar $Var\left(x_i-\overline{x}\right)$ y $E\left[x_i\left(x_j-\overline{x}\right)\right]$ en donde $\overline{x}$ es la media muestral.

SOLUCIÓN. a). La función de verosimilitud es

$$L(\theta)=L(\mu,\sigma^2;x_1,...x_5)=(2\pi\sigma^2)^{-\frac{5}{2}}\exp(-\frac{1}{2\sigma^2}\sum_1^5(x_t-\mu)^2)$$

en donde $\theta^{'}=(\mu,\sigma^2)=(\mu,1)$.

El logaritmo de la función de verosimilitud es

$$l(\theta) = -\frac{5}{2}\log(2\pi) - \frac{1}{2}\sum_{1}^{5}(x_t - \mu)^2$$

El gradiente viene dado por

$$d(\theta) = \sum_{1}^{5}(x_t - \mu)$$

El estimador MV sin restricciones se obtiene igualando el elemento del gradiente a cero, $d(\tilde{\theta}) = 0$,

$$\tilde{\mu} = \overline{x} = \frac{\sum_{1}^{5} x_t}{5}$$

El estimador restringido es

$$\tilde{\mu}_R = 0$$

Ya hemos visto que el gradiente evaluado con los estimadores sin restringir es cero. Si lo evaluamos con los estimadores restringidos el resultado es

$$d(\tilde{\theta}_R) = \sum_{1}^{5} x_t$$

El sesgo y la varianza del estimador restringido son

$$Sesgo(\tilde{\mu}_R) = E\tilde{\mu}_R - \mu = -\mu \quad \text{y} \quad Var(\tilde{\mu}_R) = 0$$

b). Los multiplicadores de Lagrange son

$$LM = d(\tilde{\theta}_R)' I(\tilde{\theta}_R)^{-1} d(\tilde{\theta}_R)$$

La matriz de Información puede escribirse como

$$I(\theta) = T$$

Sustituyendo en la formula de los multiplicadores se obtiene

$$\text{LM}=5\overline{x}\frac{1}{5}5\overline{x} = 5\overline{x}^2$$

Y teniendo en cuenta que, bajo la hipótesis nula, la media muestral sigue una distribución normal con media cero y varianza $\frac{1}{5}$ para muestras grandes podemos decir que LM sigue una distribución chi-cuadrado con un grado de libertad. Por lo tanto, la región crítica será

$$5\overline{x}^2 > \chi^2_{\varepsilon}$$

b) $Var(x_i - \overline{x}) = E(x_i - \overline{x})^2$ por ser $E(x_i - \overline{x}) = \mu - \mu = 0$.
Podemos escribir

$$E(x_i - \overline{x})^2 = Ex_i^2 + E\overline{x}^2 - 2Ex_i\overline{x}$$

Y teniendo en cuenta que

$$Ex_i^2 = \sigma^2 + \mu^2$$

$$E\overline{x}^2 = \frac{\sigma^2}{T} + \mu^2$$

$$Ex_i\overline{x} = \frac{1}{T} E\left(x_1 x_i + \ldots + x_i^2 + \ldots . x_i x_T \right) =$$

$$\frac{1}{T}(\mu^2 + \ldots \sigma^2 + \mu^2 \ldots .. + \mu^2) = \mu^2 + \frac{\sigma^2}{T}$$

Sustituyendo se obtiene

$$E(x_i - \overline{x})^2 = \sigma^2 + \mu^2 - \frac{\sigma^2}{T} - \mu^2 = \sigma^2 - \frac{\sigma^2}{T}$$

Respecto a la segunda expresión

$$E\left[x_i \left(x_j - \overline{x} \right) \right] = Ex_i x_j - Ex_i\overline{x} =$$

$$\mu^2 - \mu^2 - \frac{\sigma^2}{T} = -\frac{\sigma^2}{T}$$

Apartado 2.3

Ejercicio 2.3.1. Tomando los datos del Apartado 2.3 sobre la función de producción se trata de estimar el modelo con las variables en logaritmos. Sea M1 el modelo estimado sin restricciones, M2 el modelo excluyendo la variable capital y el M3 el modelo con rendimientos constantes. Se pide

a). Estimar los tres modelos y comparar los resultados.

b). Escribir la matriz de varianzas y covarianzas de los estimadores de las elasticidades.

Solución

a) Los resultados pueden verse en la siguiente tabla

Concepto	M1	M2	M3
$\hat{\beta}_1\left(\hat{\sigma}_{\hat{\beta}_1}\right)$ $(t-ratio)$	-0,16(0,431) (-0,378)	-1,314 (0,372) (-3,52)	0,015(0,019) (0,792)
$\hat{\beta}_2\left(\hat{\sigma}_{\hat{\beta}_2}\right)$ $(t-ratio)$	0,232(0,063) (3,68)	0	0,253(0,040) (6,32)
$\hat{\beta}_3\left(\hat{\sigma}_{\hat{\beta}_3}\right)$ $(t-ratio)$	0,805(0,144) (5,58)	1,288(0,075) (17,16)	0,747(0,040) (18,67)
Residuos	$\sum\hat{u}_t^2$ =0,0700	$\sum\hat{u}_{Rt}^2$ =0,1153	$\sum\hat{u}_{Rt}^{*2}$ =0,0705
$R^2(\bar{R}^2)$	0,957(0,953)	0,93(0,927)	0,938(0,935)
$\hat{\sigma}^2$	0,0033	0,0052	0,0032
DW	1,52	0,758	1,61

Para estimar el M2 basta tener en cuenta que $\beta_2 = 0$, por lo que se estima el siguiente modelo

$$y_t = \beta_1 + \beta_3 x_{3t} + u_{Rt} \qquad (1)$$

En lo que respecta al modelo M3, la restricción es

$$\beta_2 + \beta_3 = 1 \quad o, bien \quad \beta_2 = 1 - \beta_3$$

Sustituyendo en el modelo original se obtiene

$$y_t - x_{2t} = \beta_1 + \beta_3(x_{3t} - x_{2t}) + u_{Rt}^* \Rightarrow y_t^* = \beta_1 + \beta_3 x_{3t}^* + u_{Rt}^* \qquad (2)$$

El estimador restringido, $\hat{\beta}^*_{3R}$, se obtiene a partir de este modelo y, a continuación, $\hat{\beta}^*_{2r} = 1 - \hat{\beta}^*_{3R}$.
Prestando atención a la suma de cuadrados de los residuos claramente se ve que las dos restricciones distorsionan los datos aunque de forma diferente porque la restricción asociada con los rendimientos constantes apenas varía esa suma mientras que la variación cuando se quita el capital es mucho mayor. Este resultado también se pone de manifiesto en los valores tomados por los coeficientes de determinación y por las estimaciones de la varianza de la perturbación. De todas formas, hay que destacar la ganancia que se obtiene al reducir la colinealidad con dichas restricciones. Hay que tener en cuenta que la correlación lineal entre las dos variables, trabajo y capital, es 0,90. Esto queda reflejado en los valores de las estimaciones de las desviaciones típicas de los estimadores. Por ejemplo, la estimación de la desviación típica de la elasticidad del trabajo pasa de 0,144 en el modelo sin restricciones a 0,075 en el modelo sin la variable capital y a 0,040 en el modelo con rendimientos constantes. Es importante destacar que el modelo M2 parece indicar problemas de autocorrelación por el valor que toma el Durbin-Watson.
b). La matriz de varianzas y covarianzas de los estimadores del modelo sin restricciones puede escribirse como

$$Var\left(\hat{\beta}\right) = \sigma^2 \left(X'X\right)^{-1} = \sigma^2 \begin{bmatrix} T & \sum x_{2t} & \sum x_{3t} \\ . & \sum x_{2t}^2 & \sum x_{2t}x_{3t} \\ . & . & \sum x_{3t}^2 \end{bmatrix}^{-1}$$

Y su estimador y estimación,

$$Var\left(\hat{\beta}\right)=\hat{\sigma}_1^2\left(X'X\right)^{-1}=0,0033\left(X'X\right)^{-1}=\begin{bmatrix} 0,185 & a_{12} & a_{13} \\ & 0,004 & a_{23} \\ & & 0,020 \end{bmatrix}$$

Notar que solo se informa sobre las varianzas. Sobre las covarianzas destacar que todos los elementos fuera de la diagonal principal son diferentes de cero.

En lo que respecta al modelo escrito en (1), primero tener en cuenta que

$$Var\begin{bmatrix} \hat{\beta}_{1R} \\ \hat{\beta}_{3R} \end{bmatrix}=\hat{\sigma}_2^2\begin{bmatrix} T & \sum x_{3t} \\ & \sum x_{3t}^2 \end{bmatrix}^{-1}=0,0052\begin{bmatrix} T & \sum x_{3t} \\ & \sum x_{3t}^2 \end{bmatrix}^{-1}=$$

$$=\begin{bmatrix} 0,138 & b_{12} \\ & 0,0056 \end{bmatrix}$$

Por lo tanto,

$$Var\begin{bmatrix} \hat{\beta}_{1R} \\ \hat{\beta}_{2R} \\ \hat{\beta}_{3R} \end{bmatrix}=\begin{bmatrix} 0,138 & 0 & b_{12} \\ & 0 & 0 \\ & & 0,0056 \end{bmatrix}$$

Se trata de una matriz singular.

En lo que respecta al modelo con la restricción de rendimientos constantes escrito en (2) tenemos que,

$$Var\begin{bmatrix} \hat{\beta}_{1R}^{*} \\ \hat{\beta}_{3R}^{*} \end{bmatrix}=\hat{\sigma}_3^2\begin{bmatrix} T & \sum x_{3t}^{*} \\ & \sum x_{3t}^{*2} \end{bmatrix}=\begin{bmatrix} 0,0003 & c_{12} \\ & 0,0016 \end{bmatrix}$$

Por lo tanto,

$$Var\begin{bmatrix} \hat{\beta}^*_{1R} \\ \hat{\beta}^*_{2R} \\ \hat{\beta}^*_{3R} \end{bmatrix} = \begin{bmatrix} 0,0003 & c_{12} & c_{12} \\ & 0,0016 & 0,0016 \\ & & 0,0016 \end{bmatrix}$$

También en este caso la matriz de varianzas y covarianzas es singular.

Ejercicio 2.3.2. Considerar los dos modelos siguientes:

M1: $y_t = \beta_1 x_{1t} + u_{1t}$

M2: $y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + u_{2t}$

Para estimar β_1 se definen dos estimadores MCO, uno a partir de M1 y otro a partir de M2. suponiendo que genera los datos M2, se pide,

a). Escribir la forma que toman ambos estimadores y derivar el sesgo y la varianza de ambos. Hallar el cociente de las dos varianzas y comentar de qué depende el resultado.

b). Escribir los estimadores MCO de las varianzas de los dos estimadores de β_1 y derivar los sesgos de los mismos.

SOLUCIÓN

a). Sean $\hat{\beta}_{1R}$ y $\hat{\beta}_1$ los estimadores de β_1 a partir de M1 y M2, respectivamente, definidos como

$$\hat{\beta}_{1R} = \frac{\sum x_{1t} y_t}{\sum x_{1t}^2}$$

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \left(X'X\right)^{-1} X'y = \begin{pmatrix} \left(x_1' M_2 x_1\right)^{-1} x_1' M_2 y \\ \left(x_2' M_1 x_2\right)^{-1} x_2' M_1 y \end{pmatrix}$$

Como los datos son generados por M2, podemos escribir

$$\hat{\beta}_{1R} = \frac{\sum x_{1t} y_t}{\sum x_{1t}^2} = \frac{\sum x_{1t}\left(\beta_1 x_{1t} + \beta_2 x_{2t} + u_t\right)}{\sum x_{1t}^2} = \beta_1 + \beta_2 \frac{\sum x_{1t} x_{2t}}{\sum x_{1t}^2} + \frac{\sum x_{1t} u_t}{\sum x_{1t}^2}$$

Teniendo en cuenta que: $E\left(\sum x_{1t} u_t\right) =$ …(s.p.a.c.)…=0, se tiene que

$$E\left(\hat{\beta}_{1R}\right) = \beta_1 + \beta_2 \frac{\sum x_{1t} x_{2t}}{\sum x_{1t}^2}; \quad Sesgo\left(\hat{\beta}_{1R}\right) = \beta_2 \frac{\sum x_{1t} x_{2t}}{\sum x_{1t}^2}$$

$$Var(\hat{\beta}_{1R}) = E\left(\frac{\sum x_{1t} u_t}{\sum x_{1t}^2}\right)^2 = \ldots(\text{s.p.a.c.})\ldots = \frac{\sigma^2}{\sum x_{1t}^2}$$

(s.p.a.c.), significa: "siguiendo los procedimientos indicados en los apuntes de clase". Para el estimador de β_1 definido a partir del M2,

$$\hat{\beta}_1 = \left(x_1^{'} M_2 x_1\right)^{-1} x_1^{'} M_2 x_1 \beta_1 + \left(x_1^{'} M_2 x_1\right)^{-1} x_1^{'} M_2 u$$

teniendo en cuenta que: $M_2 = I_T - x_2\left(x_2^{'} x_2\right)^{-1} x_2$ y $M_2 x_2 = 0$. A partir de esta expresión se tiene que

$$E(\hat{\beta}_1) = \beta_1 + \left(x_1^{'} M_2 x_1\right)^{-1} x_1^{'} M_2 E u = \beta_1$$

por ser los regresores no estocásticos. Por lo tanto, $Sesgo\left(\hat{\beta}_1\right) = 0$

$$Var(\hat{\beta}_1) = \ldots(s.p.a.c.)\ldots = \sigma^2\left(x_1^{'} M_2 x_1\right)^{-1}$$

Como puede verse en el Ejercicio 3.1.3, esta varianza puede escribirse como

$$Var(\hat{\beta}_1) = \sigma^2\left(x_1^{'} M_2 x_1\right)^{-1} = \frac{\sigma^2}{\sum x_{1t}^2 (1 - r_{12}^2)}$$

en donde r_{12}^2 es el coeficiente de determinación de los dos regresores. El cociente entre las varianzas de los dos estimadores puede escribirse como:

$$\frac{Var(\hat{\beta}_{1R})}{Var(\hat{\beta}_1)} = 1 - r_{12}^2 < 1$$

Como se ve, depende de la relación lineal entre los dos regresores. Cuanto mayor es la relación menor se hace el cociente.

b). Los estimadores MCO de las varianzas son

$$Va\hat{r}\left(\hat{\beta}_{1R}\right)=\frac{\hat{\sigma}_R^2}{\sum x_{1t}^2} \text{ con } \hat{\sigma}_R^2=\frac{\sum \hat{u}_{Rt}^2}{T-1} \text{ y } \hat{u}_{Rt}=y_t-\hat{\beta}_{1R}x_{1t}$$

$$Va\hat{r}(\hat{\beta}_1)=\frac{\hat{\sigma}^2}{\sum x_{1t}^2(1-r_{12}^2)} \text{ con } \hat{\sigma}^2=\frac{\sum \hat{u}_t^2}{T-2} \text{ y } \hat{u}_t=y_t-\hat{\beta}_1x_{1t}-\hat{\beta}_2x_{2t}$$

Hay que tener en cuenta que

$$\hat{u}_{Rt}=\beta_1x_{1t}+\beta_2x_{2t}+u_t-\beta_1x_{1t}-x_{1t}\frac{\sum x_{1t}x_{2t}}{\sum x_{1t}^2}\beta_2-x_{1t}\frac{\sum x_{1t}u_t}{\sum x_{1t}^2}$$

$$=\beta_2\left(x_{2t}-\hat{x}_{2t}\right)+u_t-\hat{u}_t$$

en donde $\hat{x}_{2t}=x_{1t}\frac{\sum x_{1t}x_{2t}}{\sum x_{1t}^2}$ y $\hat{u}_t=x_{1t}\frac{\sum x_{1t}u_t}{\sum x_{1t}^2}$

Por lo tanto, podemos escribir

$$\sum \hat{u}_{Rt}^2=\beta_2^2\sum\left(x_{2t}-\hat{x}_{2t}\right)^2+\sum u_t^2+\sum \hat{u}_t^2+2\beta_2\sum u_t\left(x_{2t}-\hat{x}_{2t}\right)$$
$$-2\beta_2\sum \hat{u}_t\left(x_{2t}-\hat{x}_{2t}\right)-2\sum u_t\hat{u}_t$$

La esperanza del primer sumando en la derecha, es el mismo término por no ser estocástico. Respecto al segundo se tiene que

$$E(\sum u_t^2)=...(s.p.a.c.)...=T\sigma^2$$

La esperanza del tercer término es

$$E(\sum \hat{u}_t^2)=E\sum x_{1t}^2\frac{(\sum x_{1t}u_t)^2}{(\sum x_{1t}^2)^2}=\ldots(\text{s.p.a.c.})\ldots=\sigma^2$$

Las esperanzas de los dos siguientes términos son cero

$$E(2\beta_2\sum u_t\left(x_{2t}-\hat{x}_{2t}\right))=2\beta_2\sum\left(x_{2t}-\hat{x}_{2t}\right)Eu_t=0$$
$$E(2\beta_2\sum \hat{u}_t\left(x_{2t}-\hat{x}_{2t}\right))=2\beta_2\sum\left(x_{2t}-\hat{x}_{2t}\right)E\hat{u}_t=0$$

La esperanza del último término es

$$E(2\sum u_t\hat{u}_t)=2\frac{1}{\sum x_{1t}^2}E\left(\sum x_{1t}u_t\right)^2=2\sigma^2$$

Combinando todos estos resultados se llega a

$$E\hat{\sigma}_R^2=E\frac{\sum \hat{u}_{Rt}^2}{T-1}=\frac{1}{T-1}\beta_2^2\sum\left(x_{2t}-\hat{x}_{2t}\right)^2+\sigma^2$$

por lo que

$$EVa\hat{r}\left(\hat{\beta}_{1R}\right)=E\frac{\hat{\sigma}_R^2}{\sum x_{1t}^2}=\frac{\beta_2^2\sum\left(x_{1t}-\hat{x}_{1t}\right)^2}{\left(T-1\right)\sum x_{1t}^2}+\frac{\sigma^2}{\sum x_{1t}^2}$$

Por lo tanto, el sesgo es igual al primer término de la derecha.

En lo que respecta al estimador obtenido a partir del modelo M2, hay que tener en cuenta que

$$E\hat{\sigma}^2 = E\frac{\sum \hat{u}_t^2}{T-2} = \sigma^2$$

por estar el modelo bien especificado. Por lo tanto,

$$EVa\hat{r}(\hat{\beta}_1) = E\frac{\hat{\sigma}^2}{\sum x_{1t}^2(1-r_{12}^2)} = \frac{\sigma^2}{\sum x_{1t}^2(1-r_{12}^2)}$$

se trata de un estimador insesgado.

Ejercicio 2.3.3

Considerar los dos modelos siguientes:

M1: $y_t = \beta_1 x_{1t} + u_{1t}$

M2: $y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + u_{2t}$

Se dispone de la siguiente información

$$\left(X'X\right) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad X'y = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad y'y = 10 \quad T = 10$$

a). Estimar β_1 con MCO, utilizando ambos modelos. Comparar las propiedades (media y varianza) de los dos estimadores suponiendo que genera los datos el modelo restringido. Estimar insesgadamente la varianza de los dos estimadores. Obtener las propiedades (media y varianza) del estimador restringido de β_2.

b). Se va a discriminar entre los dos modelos utilizando el estimador máximo verosímil de la varianza de ambos estimadores. Escribir la forma que adoptaría la regla de discriminación si lo escribimos como un contraste F derivando el punto crítico implícito. Aplicar esa regla y determinar el modelo que sería rechazado en este caso. Comparar este punto crítico con el que corresponde al criterio SBIC.

Solución

a). Sean $\hat{\beta}_{1R}$ y $\hat{\beta}_1$ los estimadores de β_1 a partir de M1 y M2, respectivamente, definidos como

$$\hat{\beta}_{1R} = \frac{\sum x_{1t} y_t}{\sum x_{1t}^2} = \frac{2}{2} = 1$$

$$\hat{\beta}=\begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix}=\left(X^{'}X\right)^{-1}X^{'}y=\frac{1}{3}\begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix}\begin{pmatrix}2\\ 1\end{pmatrix}=\begin{pmatrix}1\\ 0\end{pmatrix}$$

Se ve como los dos estimadores de β_1 coinciden.

Como los datos son generados por M1, podemos escribir

$$\hat{\beta}_{1R}=\frac{\sum x_{1t}y_t}{\sum x_{1t}^2}=\frac{\sum x_{1t}\left(\beta_1 x_{1t}+u_{1t}\right)}{\sum x_{1t}^2}=\beta_1+\frac{\sum x_{1t}u_{1t}}{\sum x_{1t}^2}$$

Teniendo en cuenta que: $E\left(\sum x_{1t}u_t\right)=\ldots$(s.p.a.c.)…=0, se tiene que

$$E\left(\hat{\beta}_{1R}\right)=\beta_1;\quad Sesgo\left(\hat{\beta}_{1R}\right)=0$$

$$Var(\hat{\beta}_{1R})=E\left(\frac{\sum x_{1t}u_{1t}}{\sum x_{1t}^2}\right)^2=\ldots\text{(s.p.a.c.)}\ldots=\frac{\sigma^2}{\sum x_{1t}^2}$$

(s.p.a.c.), significa: “siguiendo los procedimientos indicados en los apuntes de clase”. Para el estimador de β_1 definido a partir del M2,

$$\hat{\beta}_1=\left(x_1^{'}M_2x_1\right)^{-1}x_1^{'}M_2x_1\beta_1+\left(x_1^{'}M_2x_1\right)^{-1}x_1^{'}M_2u$$

teniendo en cuenta que: $M_2=I_T-x_2\left(x_2^{'}x_2\right)^{-1}x_2$ y $M_2x_2=0$. A partir de esta expresión se tiene que

$$E(\hat{\beta}_1)=\beta_1+\left(x_1^{'}M_2x_1\right)^{-1}x_1^{'}M_2Eu=\beta_1$$

por ser los regresores no estocásticos. Por lo tanto, $Sesgo\left(\hat{\beta}_1\right)=0$

$$Var(\hat{\beta}_1)=...(s.p.a.c.)...=\sigma^2\left(x_1^{'}M_2x_1\right)^{-1}$$

Como puede verse en el Ejercicio 3.1.3, esta varianza puede escribirse como

$$Var(\hat{\beta}_1)=\sigma^2\left(x_1^{'}M_2x_1\right)^{-1}=\frac{\sigma^2}{\sum x_{1t}^2(1-r_{12}^2)}$$

en donde r_{12}^2 es el coeficiente de determinación de los dos regresores. El cociente entre las varianzas de los dos estimadores puede escribirse como:

$$\frac{Var(\hat{\beta}_{1R})}{Var(\hat{\beta}_1)}=1-r_{12}^2<1$$

Como se ve, depende de la relación lineal entre los dos regresores. Cuanto mayor es la relación menor se hace el cociente.

Para estimar insesgadamente la varianza de ambos estimadores, hay que sustituir σ^2 por su estimador MCO definido en el modelo amplio. Es decir, sustituido por

$$\hat{\sigma}^2 = \frac{\hat{u}'\hat{u}}{T-2} = \frac{8}{8} = 1$$

en donde,

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

$$= 10 - 2 = 8$$

Los estimadores insesgados son

$$Var\hat{r}(\hat{\beta}_1) = \frac{\hat{\sigma}^2}{\sum x_{1t}^2(1-r_{12}^2)} \quad \text{y} \quad Var\hat{r}(\hat{\beta}_{1R}) = \frac{\hat{\sigma}^2}{\sum x_{1t}^2},$$

y las correspondientes estimaciones, 0,66 y 0,5, respectivamente. Teniendo en cuenta que

$$r_{12}^2 = \frac{\left(\sum x_{1t}x_{2t}\right)^2}{\sum x_{1t}^2 \sum x_{2t}^2} = \frac{1}{4} = 0,25$$

Respecto al estimador restringido de β_2 se tiene que

$$E\hat{\beta}_{2R} = 0 \quad \text{y} \quad Var(\hat{\beta}_{2R}) = 0$$

b). La hipótesis nula se rechaza si

$$Va\tilde{r}\left(\hat{\beta}_{1R}\right) > Va\tilde{r}\left(\hat{\beta}_1\right)$$

o, equivalentemente, cuando

$$\frac{\tilde{\sigma}_R^2}{\sum x_{1t}^2} > \frac{\tilde{\sigma}^2}{\sum x_{1t}^2(1-r_{12}^2)}$$

en donde ahora

$$\tilde{\sigma}_R^2 = \frac{\sum \hat{u}_{Rt}^2}{T} \quad \text{y} \quad \tilde{\sigma}^2 = \frac{\sum \hat{u}_t^2}{T}$$

La regla de decisión puede escribirse como

$$\frac{\sum \hat{u}_{Rt}^2}{\sum \hat{u}_t^2} > \frac{1}{1-r_{12}^2}$$

Por lo tanto, el factor de parsimonia es h (T)= $\frac{1}{1-r_{12}^2}$. Rescribiendo la anterior expresión se obtiene

$$F=\frac{\sum \hat{u}_{Rt}^2-\sum \hat{u}_t^2}{\sum \hat{u}_t^2/(T-2)} > \frac{(T-2)r_{12}^2}{1-r_{12}^2}$$

El punto crítico implícito viene dado por el término de la derecha de esta desigualdad. Teniendo en cuenta que

$$r_{12}^2=\frac{(\sum_1^T x_{1t}x_{2t})^2}{\sum x_{1t}^2 \sum x_{2t}^2}=0,25$$

el valor que toma el punto crítico implícito en este ejemplo es 2,7. Por otra parte, el estadístico F toma el valor 0 por ser

$$\sum \hat{u}_{Rt}^2=\sum y_t^2-\hat{\beta}_{1R}\sum x_{1t}y_t=10-2=8$$

Por lo tanto, la hipótesis nula no se rechaza.
El punto crítico implícito correspondiente al criterio SBIC viene dado por

$$\frac{(T-2)\ln T}{T}=\frac{8\times 2,3}{10}=1,84$$

Ejercicio 2.3.4

Suponer que la variable y está relacionada con la variable x según el modelo:

$$y_i=\beta x_i+u_i \quad u_i \sim \text{iid } N(0,\sigma^2) \quad i=1,2\ \ldots\ldots n$$

Un investigador agrupa las observaciones en G grupos y toma el promedio de las dos variables en los grupos considerando el modelo de promedios

$$\bar{y}_g=\beta\bar{x}_g+\bar{u}_g, \qquad \bar{u}_g=\frac{\sum_{j=1}^{n_g} u_{gj}}{n_g}$$

$$n_1+n_2+\ldots\ldots+n_G=n$$

donde $\bar{y}_g, \bar{x}_g, \bar{u}_g$ son los promedios respectivos. Se pide:

Transformar el modelo de promedios de forma tal que el nuevo modelo obtenido sea homoscedástico. Se definen tres estimadores MCO de β; el primero, utilizando el modelo original, el segundo, utilizando el modelo de promedios y el tercero utilizando el modelo transformado. Obtener la media y varianza de los tres y determinar cual de ellos es el mejor estimador.

SOLUCIÓN

El primer estimador es el OLS que podemos escribir como

$$\hat{\beta} = \frac{\sum x_t y_t}{\sum x_t^2} = \beta + \frac{\sum x_t u_t}{\sum x_t^2}$$

La esperanza y la varianza de este estimador son,

$$E\hat{\beta} = \beta + \frac{\sum x_t E u_t}{\sum x_t^2} = \beta \quad \text{(Insesgado)}$$

$$Var(\hat{\beta}) = E\left(\hat{\beta} - \beta\right)^2 = \frac{\sum x_t^2 E u_t^2}{(\sum x_t^2)^2} = \frac{\sigma^2}{\sum x_t^2}$$

El segundo estimador definido a partir del modelo en promedios es,

$$\hat{\beta}_1 = \frac{\sum \bar{x}_g \bar{y}_g}{\sum \bar{x}_g^2} = \beta + \frac{\sum \bar{x}_g \bar{u}_g}{\sum \bar{x}_g^2}$$

La esperanza y la varianza de este estimador son,

$$E\hat{\beta}_1 = \beta + \frac{\sum \bar{x}_g E\bar{u}_g}{\sum \bar{x}_g^2} = \beta \quad \text{(Insesgado)}$$

$$Var(\hat{\beta}_1) = E\left(\hat{\beta}_1 - \beta\right)^2 = \frac{\sum \bar{x}_g^2 E\bar{u}_g^2}{(\sum \bar{x}_g^2)^2} = \frac{\sigma^2 / n_g}{\sum \bar{x}_g^2}$$

Hay que tener en cuenta que

$$E\bar{u}_g = \frac{\sum_{i=1}^{n_g} E u_{gi}}{n_g} = 0$$ porque cada perturbación tiene una esperanza igual a cero. En cuanto a la varianza,

$$Var(\bar{u}_g) = \frac{1}{n_g^2} \sum_{i=1}^{n_g} E u_{gi}^2 = \frac{n_g \sigma^2}{n_g^2} = \frac{\sigma^2}{n_g} \quad \text{(heteroscedastica).}$$

El modelo transformado puede escribirse como,

$$y_g^{\bullet} = \beta x_g^{\bullet} + u_g^{\bullet}$$

donde $y_g^{\bullet} = \sqrt{n_g}\,\bar{y}_g$, $x_g^{\bullet} = \sqrt{n_g}\,\bar{x}_g$ y $u_g^{\bullet} = \sqrt{n_g}\,\bar{u}_g$.

La perturbación de este modelo transformado tiene una esperanza igual a cero y es homoscedastica por cumplirse

$$Var\left(u_g^{\bullet}\right)=\left(\sqrt{n_g}\right)^2 E\bar{u}_g^2=\sigma^2$$

El tercer estimador a partir de este modelo transformado puede escribirse como

$$\hat{\beta}_2=\frac{\sum x_g^{\bullet} y_g^{\bullet}}{\sum x_g^{\bullet 2}}=\beta+\frac{\sum x_g^{\bullet} u_g^{\bullet}}{\sum x_g^{\bullet 2}}$$

La esperanza y la varianza de este estimador son,

$$E\hat{\beta}_2=\beta+\frac{\sum x_g^{\bullet} E u_g^{\bullet}}{\sum x_g^{\bullet 2}}=\beta \quad \text{(Insesgado)}$$

$$Var(\hat{\beta}_2)=E\left(\hat{\beta}_2-\beta\right)^2=\frac{\sum x_g^{\bullet 2} E u_g^{\bullet 2}}{(\sum \bar{x}_g^2)^2}=\frac{\sigma^2}{\sum x_g^{\bullet 2}}=\frac{\sigma^2}{\sum n_g \bar{x}_g^2}$$

Se ve que los tres estimadores son insesgados, por lo que la comparación habrá de basarse en las varianzas. Se demuestra fácilmente que el tercer estimador tiene una varianza menor que el segundo estimador. Respecto a la relación entre el primero y el tercero, hay que tener en cuenta que,

$$\sum_{i=1}^{T} x_i=\sum_{g}^{G}\sum_{j}^{T}\left(x_{gj}-\bar{x}_g+\bar{x}_g\right)^2=\sum_{g}^{G}\sum_{j}^{T}\left(x_{gj}-\bar{x}_g\right)^2$$

$$+\sum_{g}^{G} n_g \bar{x}_g^2+2\sum_{g}^{G}\bar{x}_g\sum_{j}^{T}\left(x_{gj}-\bar{x}_g\right)$$

Teniendo en cuenta que el último término de la derecha es cero, se cumple que

$$\sum_{i=1}^{T} x_i=\sum_{g}^{G}\sum_{j}^{T}\left(x_{gj}-\bar{x}_g+\bar{x}_g\right)^2>\sum_{g}^{G} n_g \bar{x}_g^2$$

Por lo que la varianza del primer estimador será menor.

Ejercicio 2.3.5

Considerar los dos siguientes modelos.

M1: $y_t=\beta_1 x_{1t}+u_{1t}$ o, matricialmente, $y=X_1\beta_1+u_1$

M2: $y_t=\beta_1 x_{1t}+\beta_2 x_{2t}+u_{2t}$ 0, matricialmente, $y=X\beta+u$

a) Para ambos modelos, definir los vectores de residuos MCO y derivar la media y matriz de varianzas y covarianzas de ambos suponiendo que los datos han sido generados por M2.

b) ¿Puede ser la suma de cuadrados de los residuos del primer modelo inferior a la del segundo si genera los datos el primer modelo?. Justifica la respuesta. ¿Puede ser el estimador MCO de σ^2 del modelo restringido menor que el del modelo sin restringir? ¿Puede ser menor si el estimador es el Máximo Verosímil?. Justifica las respuestas.

Solución

a) a) Los vectores de residuos son, respectivamente,

$$\hat{u}_1 = y - X_1\hat{\beta}_{1R} = y - X_1(X_1^{'}X_1)^{-1}X_1^{'}y = M_1y = M_1X_2\beta_2 + M_1u$$
$$\hat{u}_2 = y - X\hat{\beta} = y - X(X^{'}X)^{-1}Xy = My = Mu$$

La esperanza matemática de ambos vectores es

$$E(\hat{u}_1) = M_1X_2\beta_2 \quad \text{por ser} \quad E(M_1u) = M_1Eu = 0$$
$$E(\hat{u}_2) = 0 \quad \text{por ser} \quad E(Mu) = MEu = 0$$

En lo que respecta a la matriz de varianzas y covarianzas

$$Var(\hat{u}_1) = E(\hat{u}_1 - E\hat{u}_1)(\hat{u}_1 - E\hat{u}_1)^{'} = E(M_1uu^{'}M_1) = M_1E(uu)M_1 = \sigma^2M_1$$
$$Var(\hat{u}_2) = E(\hat{u}_2 - E\hat{u}_2)(\hat{u}_2 - E\hat{u}_2)^{'} = E(Muu^{'}M) = ME(uu)M = \sigma^2M$$

b). Nunca puede ser menor como puede verse en los resultados 2.3.3 y 2.3.25.
El estimador MCO puede ser mayor o menor. No hay una regla general. Para el estimador MV si que hay una regla general porque el denominador es el mismo. En este caso el estimador definido a partir del modelo restringido siempre será mayor ya que la suma de cuadrados del numerador siempre es mayor para este modelo con restricciones.

Apartado 2.4

Ejercicio 2.4.1 (Ajuste Parcial)

Suponer que el valor óptimo de una variable se determina como

$$y_t^* = \alpha + \beta x_t$$

Y que el valor observado de la variable se ajusta al siguiente modelo

$$y_t - y_{t-1} = \gamma\left(y_t^* - y_{t-1}\right) + u_t \qquad 0 < \gamma \leq 1$$

Se pide

1) Derivar un modelo en términos de las variables observables que permita estimar los parámetros presentes en las dos relaciones.
2) Definir el estimador MCO de γ en ese modelo y comentar las propiedades de dicho estimador haciendo hipótesis diferentes respecto a la perturbación aleatoria.

Solución.

1). Sustituyendo la primera relación en la segunda se obtiene

$$y_t - y_{t-1} = \gamma\alpha + \gamma\beta x_t - \gamma y_{t-1} + u_t$$

o, equivalentemente,

$$y_t = \gamma\alpha + \gamma\beta x_t + (1-\gamma) y_{t-1} + u_t$$

Este modelo está en términos de variables observables y los parámetros pueden estimarse.

Los estimadores MCO del modelo anterior son

$$\begin{bmatrix} \gamma\alpha \\ \gamma\beta \\ 1-\gamma \end{bmatrix} = \begin{bmatrix} T & \sum x_t & \sum x_t y_{t-1} \\ & \sum x_t^2 & \sum x_t y_{t-1} \\ & & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y_t \\ \sum x_t y_t \\ \sum y_{t-1} y_t \end{bmatrix}$$

Por lo que el estimador de γ será: $\hat{\gamma} = 1 - \left(1-\gamma\right)$. Las propiedades de este estimador dependerán de las propiedades de la perturbación aleatoria. Si esta perturbación es un ruido blanco el estimador será sesgado pero consistente. Si la perturbación tiene autocorrelación entonces el estimador MCO ni será insesgado ni consistente.

Ejercicio 2.4.2 (Expectativas Adaptables)

Se supone que la demanda de dinero, y_t, depende del tipo de interés esperado a largo plazo, x_t^*, pudiendose escribir

$$y_t = \beta_0 + \beta_1 x_t^* + u_t \qquad (1)$$

Las expectativas no son observables y se propone el siguiente esquema de formación

$$x_t^* - x_{t-1}^* = \gamma\left(x_t - x_{t-1}^*\right) \qquad 0 < \gamma \leq 1 \quad (2)$$

Se pide

1) Derivar un modelo en términos de las variables observables que permita estimar los parámetros presentes en las dos relaciones.
2) Definir el estimador MCO de γ en ese modelo y comentar las propiedades de dicho estimador haciendo hipótesis diferentes respecto a la perturbación aleatoria.

Solución

1). En primer lugar, se escribe (2) como

$$x_t^* = \gamma x_t + (1-\gamma) x_{t-1}^* \qquad (3)$$

A la relación en (1) le restamos la misma relación retardada un periodo y multiplicada por $1-\gamma$ obteniéndose

$$y_t = \gamma\beta_0 + \beta_1\left[x_t^* - (1-\gamma)x_{t-1}^*\right] + (1-\gamma)y_{t-1} + u_t - (1-\gamma)u_{t-1}$$
$$\gamma\beta_0 + \beta_1\gamma x_t + (1-\gamma)y_{t-1} + u_t - (1-\gamma)u_{t-1}$$

A partir de este modelo se pueden estimar los tres parámetros del modelo.

2). Los estimadores MCO son los mismos que en el Ejercicio 2.4.1. En este caso, sabemos que la perturbación tiene autocorrelación por lo que el estimador MCO del parámetro no será insesgado ni consistente.