PUBLICACIONES DE 4er CURSO

Grado: Economía

Asignatura: ECONOMETRÍA III

Tema 3: Evaluación. 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 3.1

Ejercicio 3.1.1

a) Definir lo que es la función de potencia de un contraste. Comentar la relación que existe entre esta función y el tamaño del error tipo 2 del contraste. Enuncia, sin demostrar, el contenido del Lema de Neyman-Pearson y comenta la relevancia del mismo para determinar la región crítica de un contraste.
b). Suponer una muestra de tamaño 100 obtenida a partir de una población que sigue una distribución Normal con media desconocida y una varianza igual a la unidad. Se quiere contrastar la hipótesis nula de que la media es igual a 2, para lo cual se propone una región crítica que especifica que, en valor absoluto, la media muestral sea superior a 3. Definir el tamaño del error tipo 1 e indicar el camino a seguir para calcular su valor. Derivar el gradiente y valorarlo utilizando el estimador restringido. Obtener la media y varianza del estimador restringido de la media asociado con la hipótesis nula.

Solución

a). La función de potencia de un contraste es una función tal que nos da, para cada valor del parámetro del que depende la función de probabilidad de la muestra, la probabilidad de rechazar la hipótesis nula. Cuando el parámetro toma el valor especificado bajo la hipótesis alternativa entonces el valor de la función de potencia es la potencia del contraste. El tamaño del error tipo 2 es uno menos la potencia del contraste. El enunciado del Lema de Neyman Pearson puede verse en el Apartado 3.1. La relevancia de este lema es que permite determinar la región crítica del contraste de forma óptima en un marco en el que tanto la hipótesis nula como la alternativa son simples.
b) La región crítica puede escribirse como

$$\left|\bar{x}\right| > 3$$

con la información dada en el enunciado, podemos decir que, bajo la hipótesis nula, se cumple que

$$\bar{x} \sim N(2, \frac{1}{T}) \text{ y como T=100, } \quad \sqrt{100}(\bar{x} - 2) \sim N(0,1)$$

A partir de estos resultados, el tamaño del error tipo 1 se define como

$$\Pr ob\left\{ \left| \bar{x} \right| > 3 / \mu = 2 \right\} = \Pr ob\left\{ \bar{x} > 3 / \mu = 2 \right\} + \Pr ob\left\{ \bar{x} < -3 / \mu = 2 \right\} =$$

$$= \Pr ob\left\{ \frac{\bar{x} - 2}{1/\sqrt{100}} > \frac{3-2}{1/\sqrt{100}} \right\} + \Pr ob\left\{ \frac{\bar{x} - 2}{1/\sqrt{100}} < \frac{-3-2}{1/\sqrt{100}} \right\} \text{ o,}$$

$$= \Pr ob\left\{ N(0,1) > \sqrt{100} \right\} + \Pr ob\left\{ N(0,1) < -5\sqrt{100} \right\} \approx 0$$

El logaritmo de la función de verosimilitud es

$$l(\mu) = -\frac{T}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum (x_t - \mu)^2$$

El gradiente viene dado por

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

Como el estimador restringido es $\tilde{\mu}_R = 2$, el valor del gradiente

será $d(\tilde{\mu}_R) = (\frac{\partial l(\mu)}{\partial \mu})_{\mu=\tilde{\mu}_R} = \sum (x_t - 2)$

Por último, $E\tilde{\mu}_R = 2$ y $Var(\tilde{\mu}_R) = 0$

Ejercicio 3.1.2

Utilizando una muestra de tamaño T obtenida a partir de una población

con media μ y varianza σ^2=1, se trata de contrastar la hipótesis nula

$H_0: \mu = 0$, frente a la alternativa de que la media es diferente de cero. La región crítica que se propone para un nivel de significación del 5% es:

$$\left|\overline{X}\right| \succ 0,5$$

en donde $\overline{x}$ es la media muestral. Calcular el tamaño muestral e indicar cual sería la región crítica si el tamaño muestral fuera T=100. Indicar como se formularía la región crítica en el caso en que se contrastara la hipótesis nula frente a una hipótesis alternativa que dijera que el parámetro toma un valor igual a 2.

Solución

Podemos escribir,

$$\Pr ob(\left|N(0,1)\right| > \frac{0,5}{\sigma / \sqrt{T}}) = 0,05$$

Por las tablas se tiene que

$$\frac{0,5}{1/\sqrt{T}} = 1,96 \Rightarrow T = 16$$

Si el tamaño muestral fuera 100, entonces tendríamos

$$\frac{c}{1/\sqrt{100}} = 1,96 \Rightarrow c = 0,196$$

Si se sabe que el valor bajo la alternativa es 2, entonces la región crítica sería

$$\overline{x} > c$$

determinando c como

$$\frac{c}{1/\sqrt{100}} = 1,65 \Rightarrow c = 0,165$$

Ejercicio 3.1.3

Considerar el siguiente modelo:

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + u_t$$

donde se supone que las variables están medidas en desviaciones respecto a sus medias y la perturbación u cumple las hipótesis ideales.

a). Expresar las varianzas de los estimadores MCO de β en función de las varianzas muestrales de x_1 y x_2 $(S_1^2 \ y \ S_2^2)$, respectivamente y del coeficiente de correlación entre ambas variables (r_{12}).

b).Conociendo que σ^2=1, S_1^2=2, T=8, $\hat{\beta}_1$(MCO)=2,se pide determinar la región crítica para contrastar la hipótesis nula H_0:$\beta_1 = 0$ frente a la alternativa de que es diferente de cero para aquellos casos en que $r_{12}^2 = 0{,}02$ y $r_{12}^2 = 0{,}99$. Utilizar un tamaño del error tipo 1 igual a 0,05.

c). Obtener el tamaño del error tipo 2 para los dos casos contemplados en el apartado anterior suponiendo β_1=1,8.

Solución

1). Teniendo en cuenta que,

$$S_i^2 = \frac{\sum x_i^2}{T} \quad \text{y que} \quad r_{12} = \frac{\sum x_1 x_2}{T S_1 S_2}$$

podemos escribir,

$$Var(\hat{\beta}) = \sigma^2 \begin{pmatrix} \sum x_1^2 & \sum x_1 x_2 \\ \sum x_1 x_2 & \sum x_2^2 \end{pmatrix}^{-1} = \frac{\sigma^2}{T} \begin{pmatrix} \dfrac{1}{S_1^2(1-r_{12}^2)} & \dfrac{-r_{12}}{S_1 S_2 (1-r_{12}^2)} \\ \dfrac{-r_{12}}{S_1 S_2(1-r_{12}^2)} & \dfrac{1}{S_2^2(1-r_{12}^2)} \end{pmatrix}$$

2). Dada la forma que adopta la hipótesis alternativa, la región crítica viene dada por:

$\left|\hat{\beta}_1\right| > c$ en donde c se determina de forma tal que,

$$\Pr ob(\left|N(0{,}1)\right| > \frac{c-0}{\sigma_{\hat{\beta}_1}}) = 0{,}05$$

Por lo que $c = 1{,}96 * \sigma_{\hat{\beta}_1}$

Si $r_{12}^2 = 0{,}02$, entonces $\sigma_{\hat{\beta}_1} = \dfrac{1}{3{,}96} \Rightarrow c = 0{,}5 \Rightarrow$ Rechazo

Notar que: $\hat{\sigma}_{\hat{\beta}_1}^2 = \dfrac{\sigma^2}{T}\dfrac{1}{S_1^2\left(1-r_{12}^2\right)} = \dfrac{1}{8 \times 2 \times 0{,}98} = \dfrac{1}{15{,}68}$

Si $r_{12}^2 = 0{,}99 \Rightarrow c = 4{,}9 \Rightarrow$ **Aceptación**

$$\hat{\sigma}^2_{\hat{\beta}_1} = \frac{\sigma^2}{T}\frac{1}{S_1^2\left(1-r_{12}^2\right)} = \frac{1}{8\times 2\times 0,01} = \frac{1}{0,16}$$

3). El tamaño del error tipo 2 será

$$\delta = 1 - \Pr ob(\left|\hat{\beta}_1\right| > c / \beta = 1.8)$$

Si

$$r_{12}^2 = 0,02 \Rightarrow \delta = 1 - \Pr ob(N(0,1) > \frac{0,5-1.8}{1/3.96}) - \Pr ob(N(0,1) < \frac{-0,5-1.8}{1/3.96}) \cong 0$$

Si

$$r_{12}^2 = 0,99 \Rightarrow \delta = 1 - \Pr ob(N(0,1) > \frac{4.9-1.8}{2.5}) - \Pr ob(N(0,1) < \frac{-4.9-1.8}{2.5}) \cong 0.89$$

Ejercicio 3.1.4

Sea $x_1,.......x_T$ una muestra aleatoria simple obtenida a partir de una población que sigue una distribución normal con media θ, desconocida, y varianza conocida igual a 1. Se va a contrastar la hipótesis nula $H_0{:}\theta = 0$ frente a la alternativa $H_1{:}\theta = 1$ utilizando un procedimiento que minimiza $\varepsilon(A) + b\,\delta(A)$, en donde $\varepsilon(A)$ y $\delta(A)$ son, respectivamente, los tamaños de los errores tipo 1 y tipo 2 de un contraste cualquiera A, y b es una constante positiva.

1). Se especifica $\varepsilon(A)$=0.01 y se pide determinar la región crítica del contraste, b y $\delta(A)$ para tres tamaños muestrales, T=1, 25, 100.

2). Se especifica b=0.16 y se pide determinar la región crítica del contraste, $\varepsilon(A)$ y $\delta(A)$ para tres tamaños muestrales, T=1, 25, 100.

Solución

1). Teniendo en cuenta el Lema de Neyman-Pearson, la región crítica que minimiza la expresión mencionada en el enunciado es,

$$\frac{L_1(x)}{L_0(x)} > \frac{1}{b} = c$$

en donde

$$L_1(x) = (2\pi)^{-\frac{T}{2}}(\sigma^2)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum(x_t-1)^2\right\}$$

sustituyendo, la región crítica resultante es

$$x > c' = \frac{1}{2} + \frac{1}{T}\log c$$

Por ser, bajo la hipótesis nula, $x \approx N(0,\frac{1}{T})$ el tamaño se define como,

$$\varepsilon(A) = \Pr ob(\sqrt{T}x > \sqrt{T}c^{'}) = 0{,}01$$

A partir de una tabla de la normal standard se tiene que:

$$\sqrt{T}c^{'} = 2{,}326 \Rightarrow c^{'} = 2{,}326 \times T^{-\frac{1}{2}}$$

En lo que respecta al tamaño del error tipo 2

$$\delta(A) = \Pr ob(x < c^{'} / H_1) = \Pr ob(N(0{,}1) < \sqrt{T}(c^{'} - 1)$$

Utilizando estos resultados se obtiene

T	$\varepsilon(A)$	$c^{'}$	**b**	$\delta(A)$
1	**0,01**	2,326	**0,16**	$\Pr ob(N(0{,}1)$ $< 1{,}326) = 0{,}91$
25	**0,01**			
100	**0,01**			

Las otras casillas se pueden llenar utilizando los resultados derivados anteriormente.

2). En este caso, la región crítica sería

$$x > \frac{1}{2} + \frac{1}{T}\log 6{,}21 = c^{'}$$

Los tamaños de los errores 1 y 2 ,respectivamente, vienen dados por

$$\varepsilon(A) = \Pr ob(x > c^{'} / H_0) = \Pr ob(N(0{,}1) > \sqrt{T}c^{'})$$

$$\delta(A) = \Pr ob(N(0{,}1) < \sqrt{T}(c^{'} - 1))$$

Los resultados son:

T	b	$c^{'}$	$\varepsilon(A)$	$\delta(A)$
1	0,16	2,3261	$\Pr ob(N(0{,}1) <$ $2{,}3261) = 0{,}01$	$\Pr ob(N(0{,}1) <$ $1{,}3261) = 0{,}91$
25	0,16			
100	0,16			

Las otras casillas se obtienen de forma inmediata.

Apartado 3.2

Ejercicio 3.2.1

Suponer una muestra aleatoria simple de tamaño T, $x_1, x_2, \ldots x_T$, a partir de una población con distribución N(μ, σ^2).

Derivar el contraste de la razón de verosimilitud de la hipótesis nula $H_o: \mu = 0$ frente a la alternativa $H_1: \mu = 1$. Suponiendo que la varianza es igual a la unidad, definir la región crítica del contraste utilizando la media muestral para un nivel de significación igual al 5%

Solución

La función de verosimilitud es

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

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

El logaritmo de la función de verosimilitud es

$$l(\theta) = -\frac{5}{2}\log(2\pi) - \frac{1}{2}\sum_1^5 (x_t - \mu)^2$$

Bajo la hipótesis nula y bajo la alternativa, este logaritmo toma la forma siguiente

$$l(0) = -\frac{5}{2}\log(2\pi) - \frac{1}{2}\sum_1^5 x_t^2$$

$$l(1) = -\frac{5}{2}\log(2\pi) - \frac{1}{2}\sum_1^5 (x_t - 1)^2$$

El contraste de la razón de verosimilitud viene dado por

RV=-2(l(0)-l(1))=2(l(1)-l(0))= $2\sum x_t - T$

La región crítica del contraste es

$RV > c^*$ que es equivalente a $\bar{x} > \frac{c}{2}$ con $c = 1 + \frac{c^*}{T}$

La constante c se determina teniendo en cuenta que

Prob($\bar{x} > \frac{c}{2}$/Hipótesis nula)=0,05 que equivale a

Prob(N(0,1)>1,65)=0,05 por lo que

c/2=1,65 y c=3,3

Ejercicio 3.2.2

Considerar una variable x distribuida como $N(\mu,\sigma^2)$ y suponer que la varianza es conocida. Se obtiene una muestra de tamaño T y se va a contrastar $H_0:\mu = 1$ frente a la alternativa de que es diferente de uno.

a). Derivar el gradiente y la matriz de información de la muestra y evaluar ambos utilizando los estimadores MV con y sin restricciones. Contrastar la hipótesis nula utilizando los contrastes LR y LM. Derivar sus regiones críticas utilizando un estadístico definido en términos de los datos eligiendo previamente el tamaño del error tipo 1.

b). Suponer que el contraste se lleva a cabo utilizando la siguiente región crítica: $LR^{\bullet} \leq$ c para c=.5 y c=1.Derivar las regiones críticas correspondientes en términos de estadísticos definidos a partir de las observaciones muestrales. Calcular aproximadamente el nivel de significación implícito. Suponer que se adopta la siguiente región crítica: $|\bar{x}| \geq 2$. Derivar aproximadamente, el tamaño del error tipo 1 que le corresponde en el caso en que σ^2 = 4 y T=16. Derivar también el tamaño del error tipo 2 para μ = 1,5.Obtener ambos tamaños en el caso en que T=36.

Solución

1). El logaritmo de la función de verosimilitud es

$$l(\mu) = -\frac{T}{2}(\log 2\pi + \log\sigma^2) - \frac{\sum(x_t-\mu)^2}{2\sigma^2}$$

El gradiente, la matriz de segundas derivadas y la matriz de información son, respectivamente

$$d(\mu) = \frac{\sum(x_t-\mu)}{\sigma^2}$$

$$H(\mu) = -\frac{T}{\sigma^2} \quad ; \quad I(\mu) = \frac{T}{\sigma^2}$$

El estimador MV restringido es $\hat{\mu}_R = 1$ mientras que el estimador sin restricciones viene dado por $\hat{\mu} = x$. Sustituyendo estos estimadores en el gradiente y en la matriz de información se llega al resultado.

El estadístico LR puede escribirse como

$$LR = \frac{1}{\sigma^2}(-\sum(x_t-\bar{x})^2 + \sum(x_t-1)^2) = \frac{1}{\sigma^2}T(\bar{x}-1)^2$$

Por ser, bajo la hipótesis nula, $x \underset{H_0}{\overset{d}{\approx}} N(1,\frac{\sigma}{\sqrt{T}})$ la región crítica del contraste es

$$LR > \chi^2_{\varepsilon}(1)$$

habiendo elegido previamente el nivel de significación. El estadístico LM toma la misma forma que el LR por lo que, en este caso, la región crítica de los dos contrastes es la misma.

b). La regla de decisión es equivalente a

$$-2\log LR^{*} > -2\log c$$

o,equivalentemente, se rechaza cuando

$$\frac{T(x-1)^2}{\sigma^2} > -2\log c$$

Si c=1, entonces la desigualdad se cumple siempre por lo que no hay una verdadera regla de decisión. Para que haya una verdadera regla de decisión c tiene que ser menor que la unidad. Si c=.5 entonces la regla de decisión es

$$\frac{T(x-1)^2}{\sigma^2} > -2\log c = 1{,}3862$$

Por cumplirse que $P(\chi^2(1) > 1{,}3862) \approx 0{,}25$, el nivel de significación es, aproximadamente, el 25%.

El tamaño del error tipo 1 viene dado por

$$\varepsilon = \Pr ob\left[|x| > 2 / H_0\right] = \ldots = \Pr ob\left[N(0{,}1) > 2\right] + \Pr ob\left[N(0{,}1) < -6\right] \approx 0{,}0228$$

Respecto al tamaño del error tipo 2,

$$\delta = 1 - \Pr ob\left[|x| > 2 / H_1 : \mu = 1{,}5\right] = \ldots = 1 - \Pr ob\left(N(0{,}1) > 1\right) - \Pr ob(N(0{,}1) < -7) = 0{,}8413$$

Los cálculos para T=36 siguen las mismas pautas.

Ejercicio 3.2.3

Utilizando los datos de la función de producción Cobb-Douglas recogidos en el Apartado 2.3 se trata de contrastar la hipótesis nula asociada con los rendimientos constantes. Es decir, se trata de contrastar la hipótesis nula

$$H_0 : \beta_2 + \beta_3 = 1$$

Se trata de llevar a cabo el contraste utilizando los procedimientos F, LR, Wald y LM. Se pueden utilizar los cálculos presentados en el Ejercicio 2.3.1.

Solución

Según los resultados presentados en el Ejercicio 2.3.1 la suma de cuadrados de los residuos MCO sin y con restricciones son

$$\sum \hat{u}_t^2 = 0{,}0700 \qquad \sum \hat{u}_{Rt}^2 = 0{,}0705$$

Como T=24, los estimadores MV de la varianza de la perturbación serán

$$\tilde{\sigma}^2 = \frac{0,0700}{24} = 0,002916 \quad \tilde{\sigma}_R^2 = \frac{0,0705}{24} = 0,002937$$

El estadístico del contraste F es

$$F = \frac{\left(\sum \hat{u}_{Rt}^2 - \sum \hat{u}_t^2\right)/1}{\hat{\sigma}^2} = \frac{0,0705 - 0,0700}{0,0033} = 0,1575$$

Bajo la hipótesis nula este estadístico sigue una distribución F con 1 grado de libertad en el denominador y 21 grados en el denominador. Para un nivel del 5% el punto crítico es $F_{0,05}(1,21) = 4,32$. Por lo tanto, la región crítica será

F>4,32

En este caso no se rechaza la hipótesis nula.

El estadístico del contraste LR es

$$LR = T\log\frac{\tilde{\sigma}_R^2}{\tilde{\sigma}^2} = 24\log\frac{0,002937}{0,002916} = 0,1720$$

Bajo la hipótesis nula este estadístico sigue una distribución χ^2 con un grado de libertad siendo el punto crítico al 5% igual a 3,84, por lo que la región crítica será: LR>3,84. La hipótesis nula de rendimientos constantes se no rechaza.

El estadístico del contraste de Wald es

$$W = T\frac{\tilde{\sigma}_R^2 - \tilde{\sigma}^2}{\tilde{\sigma}^2} = 24\frac{0,002937 - 0,002916}{0,002916} = 0,1728$$

La región crítica es la misma que la del LR por lo que la hipótesis nula tampoco se rechaza en este caso.

Por último, el estadístico del contraste LM es

$$LM = T\frac{\tilde{\sigma}_R^2 - \tilde{\sigma}^2}{\tilde{\sigma}_R^2} = 24\frac{0,002937 - 0,002916}{0,002937} = 0,1716$$

También en este caso coincide la región crítica y la hipótesis nula no se rechaza.

Ejercicio 3.2.4

Considerar el siguiente modelo:

$$y_t = \beta + u_t$$

en donde suponemos que u es N (0, σ^2) y β es un parámetro. Suponiendo que σ^2 es conocida y que se dispone de T observaciones se quiere contrastar $H_0 : \beta = 1$.

1). Obtener el gradiente y la matriz de información y evaluarlos utilizando los estimadores MV con y sin restricciones.

2). Utilizando los procedimientos de Wald y de los Multiplicadores de Lagrange contrastar la hipótesis nula especificando sus regiones críticas.

3). Suponer que se especifica la región crítica: $LR^{\bullet} \leq$ c para c=0,5 y 1. Obtener el tamaño del error tipo 1 en ambos casos.

4). Suponer que $\tilde{\beta}$ es el estimador MV. Se especifica la siguiente región crítica $\hat{\beta} \geq 2$. Obtener el tamaño del error tipo 1 suponiendo que σ^2 es igual a 4 y T=16. Para esos mismos valores obtener el tamaño del error tipo 2 para un valor de β=1,5. Indicar cuales serían los tamaños en el caso en que $\sigma^2 = 0{,}64$.

Solución

1). El logaritmo de la función de verosimilitud es

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

A partír de esta expresión, se ve que la respuesta es la misma que en el apartado 1) del ejercicio 3.2.2 siendo el parámetro β en lugar de μ.

2). El estimador Máximo-verosimil de β cumple

$$\hat{\beta} = y \approx N(\beta, \frac{\sigma^2}{T})$$

El estadístico de Wald se define como,

$$W = \frac{(\hat{\beta} - 1)^2}{\sigma^2 / T} = \frac{(y-1)^2}{\sigma^2 / T}$$

A partir de esta expresión se ve que el contraste de Wald toma la misma forma que los contrastes LR y LM derivados en el apartado 2) del ejercicio 3.2.2. Basta extender los resultados de ese apartado.

3). Son aplicables los resultados obtenidos en el mismo apartado del ejercicio 3.2.2.

4). Bajo la hipótesis nula se obtiene que,

$$\frac{\hat{\beta} - 1}{\sigma / \sqrt{T}} \approx N(0,1)$$

El tamaño del error tipo 1 viene dado por

$$\varepsilon = \Pr ob(N(0,1) > \frac{2-1}{2/4} = 2) = 0{,}0228$$

Respecto al tamaño del error tipo 2 para 1,5 se tiene,

$$\delta = 1 - \Pr ob(N(0,1) > \frac{2-1,5}{2/4} = 1) = 1 - 0,1587 = 0,8413$$

Los resultados para $\sigma^2 = 0.64$ se obtienen siguiendo las mismas pautas.

Ejercicio 3.2.5

Considerar el siguiente modelo:

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + u_t \qquad u \approx N(0, \sigma^2 I_T)$$

con:

$$X'X = \frac{1}{3}\begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix} \qquad X'y = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \sigma^2 = 1$$

Se pide:

1). Utilizando los procedimientos LR, W, LM contrastar $H_0 : \beta_2 = 0$.

2). Calcular el estimador previo contraste de β_1 bajo la restricción introducida en 1) y adoptando un nivel de significación del 5%.

3). Comparar los estimadores MCO y restringidos de β_1

Solución

1). El logaritmo de la función de verosimilitud evaluado con los estimadores restringidos y sin restringir es

$$l(\hat{\beta}) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum \hat{u}_t^2$$

$$l(\hat{\beta}_R) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum \hat{u}_{Rt}^2$$

Los elementos del gradiente evaluados con estimadores no restringidos son:

$$d_1(\tilde{\beta}) = (\frac{\partial l(\beta)}{\partial \beta_1})_{\beta = \tilde{\beta}} = \frac{\sum x_{1t}\hat{u}_t}{\sigma^2} = x_1'\hat{u}$$

$$d_2(\tilde{\beta}) = (\frac{\partial l(\beta)}{\partial \beta_2})_{\beta = \hat{\beta}_*} = \frac{\sum x_{2t}\hat{u}_t}{\sigma^2} = x_2'\hat{u}$$

De la misma manera se obtendría $d(\hat{\beta}_R)$.

La matriz de información viene dada por:

$$I(\beta) = X'X$$

El contraste de la razón de verosimilitud es,

$$LR = 2(l(\hat{\beta}) - l(\hat{\beta}_R) = (\hat{u}'_R\hat{u}_R - \hat{u}'\hat{u}) = 4$$

por ser,

$$\hat{u}'_R\hat{u}_R - \hat{u}'\hat{u} = (y - \hat{\beta}_{1R}x_1)'(y - \hat{\beta}_{1R}x_1) - y'My = -\hat{\beta}_{1R}x_1'y + y'X(X'X)^{-1}X'y$$

Por ser LR> $\chi^2(1)$=3,84(tablas)$\Rightarrow$se rechaza la hipótesis nula.

Para aplicar el contraste de Wald, hay que tener en cuenta que $R\hat{\beta} = \hat{\beta}_2$, q=0 y que $RI(\beta)^{-1}R' = 1$

por lo que, W=4, llegandose a la misma conclusión.
En lo que respecta al método de los multiplicadores de Lagrange, tener en cuenta que:

$$d(\hat{\beta}_R) = \begin{pmatrix} x_1'\hat{u}_R \\ x_2'\hat{u}_R \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

Sustituyendo se llega a LM=4, y, también , se concluye rechazando la hipótesis nula.
2). El estimador previo contraste se define como:
$\hat{\beta}_{1PC} = \hat{\beta}_1$ si la hipótesis nula es rechazada
$= \hat{\beta}_{1R}$ si la hipótesis nula no es rechazada.
3).Los dos estimadores se comparan utilizando el Error Cuadrático Medio(ECM).

$$ECM(\hat{\beta}_1) = Sesgo^2 + Varianza = 0 + 4 = 4$$

$$ECM(\hat{\beta}_{1R}) = Sesgo^2 + Varianza = (-2)^2 + 3 = 7$$

Podemos concluir diciendo que es mejor el estimador MCO.

Ejercicio 3.2.6
En el modelo con k regresores:

$$y = X\beta + u$$

se va a contrastar la hipótesis nula de no autocorrelación frente a la alternativa de un proceso autoregresivo de segundo orden . Para cinco observaciones, el vector de residuos MCO es:

$$\hat{u}' = (5 \;\; 0 \;\; 1 \;\; -3 \;\; -3\,)$$

A continuación, se hacen las regresiones de este vector,retardado uno y dos periodos, sobre los k regresores del modelo resultando los siguientes vectores de residuos MCO :

$$\hat{u}_1' = (1 \;\; 0 \;\; -1 \;\; 0 \;\; 0\,)$$

$$\hat{u}_2' = (\,0 \;\; 2 \;\; 0 \;\; -1 \;\; -1\,)$$

1). Calcular el valor que toma el estadístico de Durbin-Watson e indicar como se utilizaría para contrastar la hipótesis nula. Comentar la distribución de probabilidad que sigue este estadístico para todo tamaño muestral.
2). En el marco del contraste de los Multiplicadores de Lagrange, escribir el logaritmo de la función de verosimilitud, indicar como se obtendría el gradiente evaluado con los estimadores sin restringir y demostrar la forma que adopta $\tilde{\theta}_R$.
3). Para el contraste de los Multiplicadores de Lagrange formular la región crítica y, utilizando la información comentada anteriormente, concluir aceptando o rechazando la hipótesis nula . Recordar que, para

ε=0.05, el punto crítico para una variable χ^2 con dos grados de libertad es 6.

Solución

1). El estadístico D-W puede escribirse como:

$$d = \frac{\sum(\hat{u}_t - \hat{u}_{t-1})^2}{\sum \hat{u}_{t-1}^2} = \frac{67}{35} = 1{,}9$$

El estadístico d sigue una distribución, bajo la hipótesis nula, que depende de las observaciones de los regresores. Por lo tanto, los puntos críticos de esa distribución no están disponibles de forma general, aunque si se han encontrado dos distribuciones, d_L y d_U, cuya distribución no depende de los regresores. Son los puntos criticos de estas distribuciones los que se utilizan en los contrastes. Atendiendo a estos puntos críticos concluimos aceptando la hipótesis nula de no autocorrelación de orden 1.

2). El proceso autoregresivo de segundo orden que se contempla como alternativa puede escribirse como.

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t$$

El logaritmo de la función de verosimilitud es

$$l(\theta) = -\frac{T}{2}\log(2\pi) - \frac{T}{2}\log\sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}\sum\varepsilon_t^2$$

en donde $\theta' = (\beta', \rho_1, \rho_2, \sigma_\varepsilon^2)$

Para obtener el gradiente se deriva la primera derivada del logaritmo con respecto a cada uno de los parámetros y, a continuación, se sustituyen los parámetros por sus estimaciones MV sin restringir. Respecto a la estimación restringida se tiene

$$\hat{\theta}_R = \begin{pmatrix} \hat{\beta}_R \\ \hat{\rho}_{1R} \\ \hat{\rho}_{2R} \\ \hat{\sigma}_{\varepsilon R}^2 \end{pmatrix} = \begin{pmatrix} (X'X)^{-1}X'y \\ 0 \\ 0 \\ \frac{\hat{u}'\hat{u}}{T} \end{pmatrix}$$

3). El estadístico LM toma la forma siguiente:

$$LM = \frac{1}{\hat{\sigma}_R^2}\hat{u}'\hat{U}_{-2}\left[\hat{U}_{-2}'M\hat{U}_{-2}\right]^{-1}\hat{U}_{-2}'\hat{u}$$

Para calcular este estadístico hay que tener en cuenta que los términos de la matriz que hay entre corchetes son la suma de los productos cruzados de los residuos de la regresión de $\hat{u}$ sobre los regresores X del modelo original.
Sustituyendo se obtiene,

$$LM = \frac{1}{8,8}\begin{pmatrix} 6 & 2 \end{pmatrix}\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{6} \end{pmatrix}\begin{pmatrix} 6 \\ 2 \end{pmatrix} = 2,12$$

Justificar que, con este valor, la hipótesis nula no se rechaza.

Ejercicio 3.2.7
Considerar el modelo:

$$y = X\beta + u$$

en donde X tiene k regresores.La esperanza de u es 0 y se quiere contrastar la hipótesis nula de homoscedasticidad frente a:

$$\sigma_t^2 = (\alpha_1 + \alpha_2 z_t)^2$$

utilizando el contraste de los Multiplicadores de Lagrange:

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

1). Escribir el logaritmo de la función de verosimilitud, indicar como obtendría el gradiente y derivar la forma concreta que adoptan los k primeros elementos del gradiente.
2). Determinar los elementos del vector θ y escribir la forma que toma el vector $\tilde{\theta}_R$ en función de los datos. Evaluar los elementos del gradiente utilizando los estimadores MV con y sin restricciones, respectivamente.
3). Indicar como calcularía el estadístico LM utilizando una regresión auxiliar y establecer la región crítica del contraste utilizando este estadístico.

Solución
1). El logaritmo de la función de verosimilitud es:

$$l(\theta) = -\frac{T}{2}\log 2\pi - \frac{1}{2}\sum \log \sigma_t^2 - \frac{1}{2}\sum \frac{u_t^2}{\sigma_t^2}$$

El gradiente es un vector con k+2 parámetros que se define como se ha indicado en el Capítulo 3.
Los k primeros elementos toman la forma siguiente:

$$d_\beta(\theta) = \frac{\partial l(\theta)}{\partial \beta} = \sum \frac{x_t' u_t}{\sigma_t^2}$$

2). El vector de parámetros es $\theta' = (\beta', \alpha_1, \alpha_2)$ y sus estimaciones

$$\hat{\theta}_R = \begin{pmatrix} \hat{\beta}_R \\ \hat{\alpha}_{1R} \\ \hat{\alpha}_{2R} \end{pmatrix} = \begin{pmatrix} (X'X)^{-1}X'y \\ \left(\frac{\hat{u}'\hat{u}}{T}\right)^{\frac{1}{2}} \\ 0 \end{pmatrix}$$

La evaluación de los elementos del gradiente es como se indica en el texto.
3). La expresión es
$LM = \frac{1}{2}RSS$ en donde RSS es la varianza explicada de la regresión de
$g_t = \frac{\hat{u}_t^2}{\hat{\sigma}^2}$ sobre una constante y la variable z. La región crítica sería:

$$\frac{1}{2}RSS > \chi^2_{\varepsilon}(1).$$

Ejercicio 3.2.8

Para el modelo:

$$y_t = \phi y_{t-1} + \beta x_t + u_t \qquad |\phi| \prec 1$$

Se va a contrastar la hipótesis nula de no autocorrelación frente a la alternativa de un proceso autoregresivo de segundo orden, utilizando el procedimiento de los Multiplicadores de Lagrange. Se pide:
1). El número de elementos del gradiente indicando cuales de ellos seran cero en dos situaciones: sustituyendo los parámetros por los estimadores restringidos y cuando se utilizan los estimadores sin restringir.
2). Escribir la forma que toma el estadístico de los Multiplicadores de Lagrange indicando como se obtiene cada uno de sus elementos en función de las observaciones muestrales.Escribir la forma que adopta la región crítica indicando los tamaños del error tipo 1 y tipo2 que se cometen cuando se utiliza este contraste.
3).Demostrar que el contraste de la hipótesis nula de no autocorrelación utilizando los Multiplicadores de Lagrange equivale a contrastar la hipótesis nula : $H_0 : \delta = 0$ en el modelo:

$$y_t = \phi y_{t-1} + \beta x_t + \delta z_t + u_t$$

Especificar la forma que debe adoptar z_t.

Solución
1). El número de elementos es igual al número de parámetros, en nuestro caso 5.(Explicar porque). Si se evalúa con los estimadores MV sin restringir, todos los elementos del gradiente serán cero. Si se utilizan los estimadores restringidos, entonces tres serán cero y los otros dos diferentes

de cero.(Indicar cuales serán cero y cuales no dando las razones para este resultado).
2). El estadístico para el contraste es

$$LM = \frac{1}{\hat{\sigma}^2}\hat{u}'\hat{U}_{-p}\left[\hat{U}'_{-P}\hat{U}_{-p} - \hat{U}'_{-P}X(X'X)^{-1}X'\hat{U}_{-p}\right]^{-1}\hat{U}'_{-P}\hat{u}$$

Todos los símbolos se pueden escribir a partir de los datos. Por ejemplo,

$$\hat{U}_{-p} = \begin{pmatrix} 0 & 0 \\ \hat{u}_1 & 0 \\ \hat{u}_2 & \hat{u}_1 \\ \vdots & \vdots \\ \hat{u}_{T-1} & \hat{u}_{T-2} \end{pmatrix}$$

y así sucesivamente. Este contraste tiene una justificación asintótica y, en este marco, se sabe que el nivel de significación es el adoptado previamente y que el tamaño del error tipo 2 tiende a cero conforme la muestra crece. Pero en muestras pequeñas es difícil concretar los tamaños de los dos errores.
3). Basta hacer $Z = \hat{U}_{-p}$ y luego estimar δ como:
$\hat{\delta} = (Z'MZ)^{-1}Z'My$. A continuación, se forma la expresión

$$\hat{\delta}'\left(Var(\hat{\delta})\right)^{-1}\hat{\delta}$$

Se demuestra fácilmente que esta expresión coincide con el estadístico LM definido anteriormente.

Ejercicio 3.2.9
Suponer que para el modelo:

$$y_t = \beta t + u_t \qquad t=1, 2, 3, 4, 5.$$

el vector de residuos MCO es el siguiente:

$$\hat{u}' = (\ 2, -4, 3, 1, -2\ )$$

1). Se pide contrastar la hipótesis nula de no autocorrelación frente a la alternativa de un proceso autorregresivo de orden dos utilizando el procedimiento de los Multiplicadores de Lagrange. Para el cálculo de este estadístico utilizar la siguiente aproximación:

$$[\hat{U}'_{-P}\hat{U}_{-p} - \hat{U}'_{-p}X(X'X)^{-1}X'\hat{U}_{-p}] = \begin{pmatrix} 29 & -18 \\ -18 & 29 \end{pmatrix}$$

Definir $\hat{U}_{-p}$ y $\hat{U}'_{-P}X$ utilizando los datos del problema.
2). Utilizando el mismo contraste, contrastar la hipótesis nula de homoscedasticidad frente a una alternativa del siguiente tipo: $\sigma_t^2 = (\alpha_0 + \alpha_1 t)^2$
3). Utilizando el mismo procedimiento contrastar la hipótesis nula de normalidad.

Solución

1). El estadístico LM puede escribirse como,

$$LM = \frac{1}{\hat{\sigma}_R^2}\hat{u}'\hat{U}_{-2}\left[\hat{U}_{-2}'M\hat{U}_{-2}\right]^{-1}\hat{U}_{-2}'\hat{u}$$

Para calcular este estadístico hay que tener en cuenta que

$$\hat{\sigma}_R^2 = \frac{\hat{u}'\hat{u}}{T} = 6{,}8$$

$$\hat{u}'\hat{U}_{-2} = \left(\sum \hat{u}_t\hat{u}_{t-1} \quad \sum \hat{u}_t\hat{u}_{t-2}\right) = \begin{pmatrix}-19 & -4\end{pmatrix}$$

$$\hat{U}_{-2}'X = \begin{pmatrix}0 & 2 & -4 & 3 & 1\\ 0 & 0 & 2 & -4 & 3\end{pmatrix}\begin{pmatrix}1\\2\\3\\4\\5\end{pmatrix} = \begin{pmatrix}9\\5\end{pmatrix}$$

Sustituyendo se obtiene,

$$LM = \frac{1}{6{,}8}\begin{pmatrix}-19 & -4\end{pmatrix}\begin{pmatrix}29 & -18\\ -18 & 29\end{pmatrix}^{-1}\begin{pmatrix}-19\\ -4\end{pmatrix} = 3{,}86$$

La región crítica del contraste es

$LM > \chi^2_{0,05} = 6$ (Explicar esta región crítica)

por lo que no se rechaza la hipótesis nula.

2).El contraste LM para cotrastar homoscedasticidad es

$$LM = \frac{1}{2}f'Z\left(Z'Z\right)^{-1}Z'f$$

Teniendo en cuenta que,

$$f_t = \frac{\hat{u}_t^2}{\hat{\sigma}^2} - 1$$

se llega a

$$f = \begin{pmatrix}-0{,}411\\ 1{,}352\\ 0{,}323\\ -0{,}852\\ -0{,}411\end{pmatrix}; \quad \left(Z'Z\right)^{-1} = \begin{pmatrix}5 & 15\\ 15 & 55\end{pmatrix}^{-1} = \begin{pmatrix}1.1 & -.3\\ -.3 & .1\end{pmatrix}$$

Por otra parte se tiene también que

$$f'Z = \begin{pmatrix}0 & -2.2\end{pmatrix}$$

Sustituyendo todos estos valores en la expresión del contraste se obtiene:

$$LM = \frac{1}{2}\begin{pmatrix}0 & -2.2\end{pmatrix}\begin{pmatrix}1.1 & -.3\\ -.3 & .1\end{pmatrix}\begin{pmatrix}0\\ -2.2\end{pmatrix} = .242$$

Por ser $\chi^2_{0,05} = 3$, la hipótesis nula de homoscedasticidad no sería rechazada.

3). El contraste LM es

$$LM = \frac{Tg_1^2}{6} + \frac{Tg_2^2}{24} = 0{,}5455$$

La hipótesis nula se aceptaría porque................

Tener en cuenta que

$$g_1 = \frac{m_3}{m_2^{\frac{3}{2}}} = -0{,}4161$$

$$g_2 = \frac{m_4}{m_2^2} - 3 = -1{,}39$$

con, $m_i = \dfrac{\sum \hat{u}_t^i}{5}$

Interpretar g_1 y g_2.

Apartado 3.3

Ejercicio 3.3.1

Suponer el modelo:

$$\begin{aligned} c_t &= \beta y_t + u_t \\ y_t &= c_t + a_t \end{aligned} \qquad (1)\text{-}(2)$$

en donde c_t, y_t y a_t son, respectivamente, consumo, renta y gasto autónomo; u_t es un ruido blanco. Se supone que a_t no es aleatorio.

1). Obtener la forma reducida del modelo, las medias, varianzas y covarianza de c_t e y_t y la media y varianza de c_t dada y_t.

2). Estimar β aplicando MCO en (1). Derivar las propiedades de este estimador tanto en muestras pequeñas como asintóticamente y determinar el signo del sesgo en caso de que exista.

3). Estimar β utilizando el estimador de los mínimos cuadrados en dos etapas y derivar las propiedades de este estimador.

Solución

1). La forma reducida es,

$$y_t = \frac{a_t}{1-\beta} + \frac{u_t}{1-\beta}$$

$$c_t = \frac{\beta a_t}{1-\beta} + \frac{u_t}{1-\beta}$$

Las medias y varianzas son:

$$\begin{pmatrix} y_t \\ c_t \end{pmatrix} \approx N\left[\begin{pmatrix} \dfrac{a_t}{1-\beta} \\ \dfrac{\beta a_t}{1-\beta} \end{pmatrix} \begin{pmatrix} \dfrac{\sigma^2}{(1-\beta)^2} & \dfrac{\sigma^2}{(1-\beta)^2} \\ \dfrac{\sigma^2}{(1-\beta)^2} & \dfrac{\sigma^2}{(1-\beta)^2} \end{pmatrix}\right]$$

Par la distribución condicional vamos a considerar el siguiente resultado

Si $$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \approx N\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix}\right]$$

Entonces la distribución del primer subvector, dado el segundo, viene dada por

$$x_1 / x_2 \approx N\left[\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}\right]$$

Aplicándolo a nuestro caso se obtiene

$$c_t / y_t \approx N(y_t - a_t, 0)$$

2). En este marco, los momentos muestrales convergen a los momentos poblacionales. Además, los momentos muestrales entre a_t y cualquiera de las variables aleatorias convergen a cero. Por ejemplo,

$$T^{-1}\sum a_t u_t \xrightarrow{p} 0$$

Así, a partir de la primera relación de la forma reducida tenemos

$$T^{-1}\sum y_t u_t \xrightarrow{p} \frac{\sigma_u^2}{1-\beta}$$

Por lo tanto, las variables y_t y u_t no son asintóticamente independientes. De la misma manera podemos escribir

$$T^{-1}\sum y_t^2 \xrightarrow{p} \frac{\sigma_a^2 + \sigma_u^2}{(1-\beta)^2}$$

Utilizando estas expresiones la convergencia del estimador MCO de β es la siguiente

$$\hat{\beta} = \beta + \frac{\sum y_t u_t}{\sum y_t^2} \xrightarrow{p} \beta + \frac{(1-\beta)\sigma_u^2}{\sigma_a^2 + \sigma_u^2}$$

Se puede derivar el signo del sesgo y su importancia en cuantía.

3). El estimador en dos etapas toma la forma siguiente

$$\beta^* = \frac{\sum c_t y_t^*}{\sum y_t^{*2}}$$

en donde,

$$y_t^* = \hat{\delta} a_t = \frac{\sum y_t a_t}{\sum a_t^2} a_t$$

Se ve que

$$\hat{\delta} \xrightarrow{p} \frac{1}{1-\beta}$$

La convergencia del estimador bietápico es

$$\beta * = \frac{\hat{\delta}\sum c_t a_t}{\hat{\delta}^2 \sum a_t^2} \xrightarrow{p} \beta$$

Apartado 3.4

Ejercicio 3.4.1

Considerar los dos modelos siguientes:

$$\text{M1: } y_t = \beta_1 x_{1t} + u_{1t}$$
$$\text{M2: } y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + u_{2t}$$

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 discriminar entre los dos modelos utilizando el estimador máximo verosímil de la varianza de ambos estimadores. Determinar la región crítica y obtener el factor de parsimonia correspondiente a la regla de decisión.

SOLUCIÓN

La hipótesis nula se rechaza si

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

o, equivalentemente, cuando

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

(Ver el Ejercicio 3.1.3 para derivar la varianza del segundo modelo)
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, Región Crítica, 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)= $\frac{1}{1 - r_{12}^2}$.

Ejercicio 3.4.2

Considerar los dos modelos siguientes:

$$\text{M1: } y_t = \beta_1 x_{1t} + u_{1t}$$
$$\text{M2: } y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + u_{2t}$$

Sea X la matriz T$\times$2 de las observaciones de las dos variables. Se dispone de la siguiente información:

$$\bar{y}=4; \quad y'y=200; \quad y'X(X'X)^{-1}X'y=190;$$

$$X'y=\begin{pmatrix}40\\20\end{pmatrix}; \quad \sum_{1}^{10} x_{1t}^2=10; \quad \text{T=10}$$

Se pide:

a). Para los criterios $\bar{R}^2, AIC$ calcular sus valores para el modelo amplio y derivar y calcular el factor de parsimonia y el punto critico implícito para esos dos criterios y el de los Multiplicadores de Lagrange.

b). Escribir la región crítica que corresponde a cada criterio y utilizando la información numérica aportada indicar si rechazaría la hipótesis nula.

SOLUCIÓN

Primero, hay que tener en cuenta que

$$u_1'u_1=y'y-y'x_1(x_1'x_1)^{-1}x_1y=200-160=40$$

$$u_2'u_2=y'y-y'X(X'X)^{-1}X'y=200-190=10$$

$$\sum_{1}^{10}\left(y_t-\bar{y}\right)^2=y'y-T\bar{y}^2=200-10.16=40$$

$$k_1=1 \qquad k_2=2$$

a). El valor de los criterios es

$$\bar{R}_2^2=1-\frac{T-1}{T-k_2}\frac{\hat{u}_2'\hat{u}_2}{\sum\left(y_t-\bar{y}\right)^2}=0,718$$

$$AIC_2=\ln\tilde{\sigma}_2^2+\frac{2\left(k_2+1\right)}{T}=0,6$$

$$LM_2=T\,\frac{\hat{u}_1'\hat{u}_1-\hat{u}_2'\hat{u}_2}{\hat{u}_1'\hat{u}_1'}=7,5$$

Respecto a los factores de parsimonia, decimos que rechazamos la hipótesis nula si

$$\bar{R}_1^2 < \bar{R}_2^2 \Rightarrow 1 - \frac{T-1}{T-k_1}\frac{\hat{u}_1'\hat{u}_1}{\sum\left(y_t-\bar{y}\right)^2} < 1 - \frac{T-1}{T-k_2}\frac{\hat{u}_2'\hat{u}_2}{\sum\left(y_t-\bar{y}\right)^2} \Rightarrow$$

$$\frac{\hat{u}_1'\hat{u}_1}{T-k_1} > \frac{\hat{u}_2'\hat{u}_2}{T-k_2} \Rightarrow \frac{\hat{u}_1'\hat{u}_1}{\hat{u}_2'\hat{u}_{'2}} > 1 + \frac{k_2-k_1}{T-k_2} = h(\bar{R}^2)$$

De la misma manera se pueden obtener los factores de parsimonia para los otros dos criterios. Sustituyendo, se obtiene

$$h(\bar{R}^2) = 1 + \frac{k_2-k_1}{T-k_2} = 1{,}11$$

$$h(AIC) = 1 + \frac{2(k_2-k_1)}{T} = 1{,}2$$

$$h(LM) = \frac{\chi_\varepsilon^2}{T-\chi_\varepsilon^2} = 1{,}62$$

Respecto al punto crítico implicito, se tiene que

$$\bar{R}_1^2 < \bar{R}_2^2 \Rightarrow \frac{\hat{u}_1'\hat{u}_1}{\hat{u}_2'\hat{u}_{'2}} > 1 + \frac{k_2-k_1}{T-k_2} \Rightarrow \frac{\hat{u}_1'\hat{u}_1 - \hat{u}_2'\hat{u}_{'2}}{\hat{u}_2'\hat{u}_{'2}} > \frac{k_2-k_1}{T-k_2} \Rightarrow$$

$$\frac{(\hat{u}_1'\hat{u}_1 - \hat{u}_2'\hat{u}_{'2})/(k_2-k_1)}{\hat{u}_2'\hat{u}_{'2}/(T-k_2)} > 1$$

Por lo tanto,

$$F_{\bar{R}^2} = 1$$

Siguiendo similar proceso para los otros dos criterios se obtiene

$$F_{AIC} \approx 2$$

$$F_{LM} = \frac{T-k_2}{k_2-k_1}\frac{\chi_\varepsilon^2}{T-\chi_\varepsilon^2} = \frac{8.3{,}84}{6{,}16} = 4{,}98$$

b). Por lo que se ha visto al derivar el factor de parsimonia, la hipótesis nula se rechaza si

para $\bar{R}^2$ se cumple que $\dfrac{\hat{u}_1'\hat{u}_1}{\hat{u}_2'\hat{u}_{'2}} > 1 + \dfrac{k_2-k_1}{T-k_2} = 1{,}11$

para AIC se cumple que $\dfrac{\hat{u}_1'\hat{u}_1}{\hat{u}_2'\hat{u}_{'2}} > 1 + \dfrac{2(k_2-k_1)}{T} = 1{,}2$

para LM se cumple que $\dfrac{\hat{u}_1'\hat{u}_1}{\hat{u}_2'\hat{u}_{'2}} > \dfrac{T}{T-\chi_\varepsilon^2} = 1{,}62$

Por ser

$$\frac{\hat{u}_1'\hat{u}_1}{\hat{u}_2'\hat{u}_2} = \frac{40}{10} = 4$$ la hipótesis nula se rechaza con los tres criterios.

Ejercicio 3.4.3

Considerar los dos siguientes modelos.

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

M2: $y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + u_t$ 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). Para seleccionar entre los dos modelos, se utilizan los criterios $\bar{R}^2$, AIC y LM. Escribir la forma que adoptan los estadísticos correspondientes a los tres criterios. Suponiendo que la hipótesis nula viene dada por el modelo restringido escribir, para cada uno de los tres criterios, la región crítica y derivar el correspondiente factor de parsimonia.

SOLUCIÓN

a) Los vectores de residuos son, respectivamente,

M1:

$$\hat{u}_1 = y - X_1\hat{\beta}_{1R} = y - X_1(X_1'X_1)^{-1}X_1'y = M_1y = X_2\beta_2 + M_1u$$

$$\hat{u}_2 = y - X\hat{\beta} = y - X(X'X)^{-1}Xy = My = Mu$$

La esperanza matemática de ambos vectores es

$$E(\hat{u}_1) = X_2\beta_2 \quad \text{por ser} \quad E(M_1u) = M_1Eu = 0$$

$$E(\hat{u}_2) = 0 \quad \text{por ser} \quad E(Mu) = MEu = 0$$

En lo que respecta a la matriz de varianzas y covarianzas

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

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

b) La tres regiones críticas son

$$\bar{R}_1^2 < \bar{R}_2^2 \qquad \bar{R}_i^2 = 1 - \frac{T-1}{T-k_i}\frac{\hat{u}_i'\hat{u}_i}{\sum\left(y_t - \bar{y}\right)^2}$$

$$LM > \chi_\varepsilon^2(k_2 - k_1) \qquad \text{LM=T}\frac{\tilde{\sigma}_1^2 - \tilde{\sigma}_2^2}{\tilde{\sigma}_1^2}$$

$$AIC_1 > AIC_2 \qquad\qquad AIC_i = \ln \tilde{\sigma}_i^2 + \frac{2k_i}{T}$$

Estas tres regiones críticas son casos particulares de la expresión general

$\hat{u}_1^{'}\hat{u}_1 > \hat{u}_2^{'}\hat{u}_2 h(.)$ (Hay que demostrar este resultado para cada criterio). h(.) es el factor de parsimonia

$$h(\bar{R}^2) = \frac{T - k_1}{T - k_2}$$

$$h(LM) = \frac{T}{T - \chi_\varepsilon^2}$$

$$h(AIC) = \exp(\frac{2(k_2 - k_1)}{T}) \simeq 1 + \frac{2(k_2 - k_1)}{T}$$