



LICENCIATURA EN ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS

Econometría. Primer Parcial, 2010

TEORÍA

1. A partir de T observaciones se ha estimado el modelo $Y_t = \beta_1 X_t + \beta_1^2 X_t^2 + \beta_2 Y_{t-1} + \varepsilon_t$. Indique, razonando la respuesta, el tamaño ajustado T' y los grados de libertad.
2. Escriba la ecuación auxiliar del test de White para el modelo anterior así como sus grados de libertad.
3. ¿Cuál sería el test adecuado para detectar en este modelo la presencia de autocorrelación de primer orden? Justifique la respuesta.
4. Enumere las fases de construcción de un modelo econométrico indicando muy brevemente el sentido de cada una de ellas.
5. Explique brevemente qué significa que un estimador sea insesgado.
6. Explique el concepto de heterocedasticidad e indique alguna de sus consecuencias.
7. En el modelo $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$ se sabe que $X_2 = 2X_1 - X_3 + \varepsilon$. ¿Qué hipótesis del modelo lineal se incumple? Justifique la respuesta.
8. Explique en qué consiste el método de los Mínimos Cuadrados Generalizados y cuál es su utilidad.

PRÁCTICA

Se han observado cuatro variables, Y, X_1, X_2, X_3 durante el periodo comprendido entre enero de 2007 y junio de 2009 habiéndose estimado a partir de las mismas un modelo para explicar la primera de ellas en función de las restantes con el siguiente resultado:

	S.D. var	Coefficient	Std. Error
C		1.065	0.473
X1	0.415	-2.673	0.489
X2	0.994	3.316	0.233
X3	0.501	-1.150	0.250
S.E. of regression	0.271	S.D. dep. var	2.916

habiéndose llegado a la conclusión de que era necesaria su reformulación por existir indicios de multicolinealidad entre las variables explicativas, lo que condujo a la construcción de un nuevo modelo:

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.735	1.375	0.534	0.598
X1	4.081	0.348	11.740	0.000
X3	-4.425	0.288	-15.387	
R-squared		Mean dependent var		0.203
Adjusted R-squared		S.D. dependent var		
S.E. of regression		Akaike info criterion		2.460
Sum squared resid	16.837	Schwarz criterion		2.600
Log likelihood	-33.904	Hannan-Quinn criter.		2.505
F-statistic		Durbin-Watson stat		2.444
Prob(F-statistic)	0.000			

	AC	Q-Stat	Prob
1	-0.231	1.768	0.184
2	0.062	1.901	0.387
3	-0.240	3.954	0.266
4	-0.047	4.035	0.401
5	0.077	4.264	0.512
6	-0.014	4.273	0.640
7	-0.141	5.108	0.647

del que se conoce además que el coeficiente de determinación para el modelo auxiliar de White es 0.215 así como la matriz

$$(X'X)^{-1} = \begin{pmatrix} & -0.5935 & -0.3872 \\ -0.5935 & & -0.0028 \\ -0.3872 & -0.0028 & \end{pmatrix}$$

- a) (1 punto) ¿Está justificada la afirmación hecha sobre el primer modelo? ¿Ha logrado corregirse el problema en el segundo? Conteste a estas cuestiones argumentando a partir de los correspondientes factores de incremento de la varianza.

Conteste las siguientes cuestiones en relación siempre al modelo segundo y para $\alpha=0.05$ cuando sea necesario:

- b) (1,5 puntos) Complete la información sombreada en la tabla y la faltante en la matriz.
 c) Valide el modelo (0,5 puntos) así como el cumplimiento de las hipótesis de partida del mismo (1,5 puntos).
 d) (1 punto) Contraste la hipótesis de que el modelo presenta autocorrelación de primer orden positiva.
 e) (1 punto) ¿Entre qué límites se encontrará con una confianza del 95% el valor de la variable endógena si es

$$X_1 = 2,52 \quad X_3 = 3,47 \quad ? \quad \left[\hat{y} \mp t_{gl, \alpha/2} \hat{\sigma} \sqrt{1 + x'(XX)^{-1}x} \right]$$

- f) (2 puntos) ¿Es aceptable suponer que cada unidad de incremento de X_1 conlleva un incremento de Y de al menos 4.5 unidades? ¿Entre qué límites se encontraría este incremento con una confianza del 95%?
 g) (1,5 puntos) Ante la sospecha de que pudiera haber cierta heterocedasticidad relacionada con el valor de la variable X_1 se dividió la muestra en 3 partes iguales, estimando modelos auxiliares de regresión para la primera y la tercera, resultando en el primer caso una suma de cuadrados residual 3.352 y 10.312 para el segundo. Indique el valor y los grados de libertad para el estadístico F de Goldfeld-Quandt e indique a qué conclusión se llegaría sabiendo que $\text{Prob}(F)=0,046$.

Estadístico	0,025	0,05
t	2,052	1,703
χ^2		11,07
d_L	1,18	1,28
d_U	1,48	1,57

Nota: Especifique, en todos los contrastes, las hipótesis nula y alternativa, el estadístico empleado y la decisión adoptada.

SOLUCIONES

Teoría

- Como hay un retardo de un periodo es $T' = T - 1$. Como se han estimado dos parámetros (β_1, β_2) es $gl = T' - 2 = T - 3$.
- Sería $\sigma_i^2 = \alpha_0 + \alpha_1 X_i + \alpha_2 X_i^2 + \alpha_3 Y_{i-1} + \alpha_4 X_i^4 + \alpha_5 Y_{i-1}^2 + \alpha_6 X_i^3 + \alpha_7 X_i Y_{i-1} + \alpha_8 X_i^2 Y_{i-1} + \eta_i$ con 8 grados de libertad (el número de variables de la ecuación).
- El de Durbin y no el habitual de Durbin-Watson ya que el modelo presenta con explicativa a la variable endógena retardada.
- Especificación: Selección de variables y de modelo así como determinación de hipótesis de partida. Estimación: determinación de estimaciones de los parámetros a partir de la muestra. Validación: Contraste sobre valores paramétricos, modelo e hipótesis de partida.
- Significa que aunque ningún valor concreto de la estimación coincida necesariamente con el del parámetro sí lo hace el valor medio de los mismos.
- La heterocedasticidad se produce cuando la variabilidad de los residuos de un modelo no es constante. Como consecuencia son poco fiables las estimaciones hechas sobre las varianzas de los estimadores de los parámetros del modelo y por tanto los intervalos de confianza o los contrastes que se les asocian.
- Falla la hipótesis de no multicolinealidad, pues una de las variables explicativas puede ser calculada, al menos aproximadamente, por el resto de explicativas.
- Es un método que se emplea para obtener buenos estimadores de los parámetros de un modelo lineal cuando se incumplen las hipótesis de homocedasticidad o no autocorrelación. Consiste en incorporar en el procedimiento de cálculo de los estimadores información sobre las varianzas y covarianzas residuales.

Práctica

- a) Los factores de incremento de la varianza para las tres variables son $F_1 = 16,215$; $F_2 = 21,179$; $F_3 = 6,204$. Puesto que los dos primeros son mayores que 10 son indicio claro de multicolinealidad. En el segundo modelo, una vez estimada la desviación típica residual, $\hat{\sigma} = 0,79$ (ver apartado b), los nuevos factores son $F_1 = 0,966$; $F_3 = 0,969$ por lo que sí se parece haberse corregido el problema.

- b) El modelo tiene $30-3=27$ grados de libertad. Las estimaciones son $\hat{\sigma} = \sqrt{\frac{16,837}{27}} = 0,79$; $\bar{R}^2 = 1 - \left(\frac{0,79}{2,916}\right)^2 = 0,927$; $R^2 = 1 - (1 - 0,927) \frac{27}{29} = 0,932$; $F = \frac{0,932}{1 - 0,932} \frac{27}{2} = 185,029$.

Como $|-15,387| > 1,74 \Rightarrow \text{Prob}(T_3) < \text{Prob}(T_1) \approx 0$. Finalmente, de ser $0,79\sqrt{a_{00}} = 1,375$ se obtiene $a_{00} = 3,029$ y de la misma manera, $a_{11} = 0,194$; $a_{22} = 0,133$.

- c) Para el contraste global, por ser $\text{Prob}(F) = 0 < \alpha$ se llega a la conclusión de que el modelo es lineal. Los contrastes individuales en cada variable verifican $\text{Prob}(T_1), \text{Prob}(T_3) = 0 < \alpha$ por lo que es lineal en ambas. En cuanto a la hipótesis de homocedasticidad es $W = 30 \cdot 0,215 = 6,45 > \chi_{0,05}^2 = 11,07$ por lo que se puede aceptar que el modelo es homocedástico. Finalmente, para la prueba de autocorrelación es preferible el test de Ljung-Box dado que el otro disponible, Durbin-Watson sólo puede detectar autocorrelación de primer orden. Como $\text{Prob}(Q_1), \dots, \text{Prob}(Q_7) > \alpha$ podrá aceptarse que el modelo no presenta indicios de autocorrelación de ningún orden de los contrastables.
- d) Para el contraste de hipótesis nula $\rho_1 > 0$ puesto que $D = 2,444$ y $4 - d_{U;0,05} = 2,43$ se rechazará la hipótesis nula.
- e) La estimación puntual es $\hat{y} = -4,336$. El intervalo es $-4,336 \mp 2,052 \cdot 0,79 \sqrt{1+0,135} = (-6,062; -2,61)$.
- f) Para el contraste de hipótesis nula $\beta_1 > 4,5$ puesto que $T_1 = \frac{4,081 - 4,5}{0,348} = -1,204 > -t_{25;0,05} = -1,703$ se puede aceptar la hipótesis nula. El intervalo es $4,081 \mp 2,052 \cdot 0,348 = (3,367; 4,795)$.
- g) El valor del estadístico de contraste es $F = \frac{10,312}{3,352} = 3,076$ con $10 - 3 = 7$ y 7 grados de libertad. Puesto que $\text{Prob}(F) = 0,046 < \alpha$ habría que suponer que existen indicios de heterocedasticidad ligados a la primera variable.