

ECONOMETRÍA. GADE

Prácticas

Tema 4

Ejercicios resueltos

1. Proponga un modelo para explicar la recaudación anual del cine (Y) en el periodo 1985-2015, en función del número de entradas vendidas los días de diario (X_1), el número de entradas vendidas los días festivos (X_2) y el número de entradas totales vendidas (X_3). ¿Considera que puede existir colinealidad en este modelo?. ¿De que tipo?

RESOLUCIÓN:

El modelo se puede formular como $Y_t = \beta_1 + \beta_2 X_{1t} + \beta_3 X_{2t} + \beta_4 X_{3t} + u_t$. Este modelo tendrá colinealidad perfecta ya que el número de entradas totales vendidas (X_3) se calcula como la suma del número de entradas vendidas los días de diario (X_1) y el número de entradas vendidas los días festivos (X_2).

2. Se tienen datos sobre la evolución de las ventas (Y) del grupo Inditex, el número de tiendas (X_1), número de países (X_2), número de marcas (X_3) y número de empleados (X_4). Se pide:
 - a) Analice la posible existencia de colinealidad a partir de la siguiente matriz de correlaciones:

Y	X_1	X_2	X_3	X_4	
1	0,9945	0,9376	0,8253	0,9971	Y
	1	0,9529	0,8500	0,9977	X_1
		1	0,9606	0,9551	X_2
			1	0,8530	X_3
				1	X_4

RESOLUCIÓN:

A la vista de los coeficientes de correlación se detecta que:

- La variable X_1 esta fuertemente relacionada con las variables X_2 y X_4 ya que los coeficientes de correlación son respectivamente 0,9529 y 0,9977. Tambien esta relacionada con la variable X_3 ya que el coeficiente de correlación es igual 0,85.
- La variable X_2 esta muy relacionada con las variables X_3 y X_4 ya que los coeficientes de correlación son 0,9606 y 0,9551, respectivamente.
- La variable X_3 esta relacionada con la variable X_4 ya que el coeficiente de correlación es igual a 0,8530.

Todas las variables independientes estan relacionadas entre sí, por lo que se puede decir que existe problemas de colinealidad.

- b) Analice la posible existencia de colinealidad a partir de las siguientes regresiones auxiliares:

RESOLUCIÓN:

$$\hat{X}_1 = \begin{matrix} 108+ & 4,83X_2 & -43,9X_3+ & 0,0471X_4 \\ (299) & (15,0) & (129) & (0,00445) \end{matrix} \quad R^2 = 0,995$$

$$\hat{X}_2 = \begin{matrix} -14,8+ & 0,00133X_1 & +8,09X_3+ & 0,000213X_4 \\ (3,32) & (0,00413) & (0,708) & (0,000202) \end{matrix} \quad R^2 = 0,99$$

$$\hat{X}_3 = \begin{matrix} 2,11 & -0,000164X_1 & +0,110X_2 & -1,91 \times 10^{-5}X_4 \\ (0,240) & (0,000482) & (0,00963) & (2,39 \times 10^{-5}) \end{matrix} \quad R^2 = 0,97$$

$$\hat{X}_4 = \begin{matrix} 2,52 \times 10^3+ & 18,6X_1 & +305X_2 & -2,02 \times 10^3X_3 \\ (5,93 \times 10^{-5}) & (1,76) & (290) & (2,52 \times 10^3) \end{matrix} \quad R^2 = 0,996$$

A la vista de los datos aportados, se puede calcular el FIV como $FIV = \frac{1}{1-R^2}$ para cada una de las variables explicativas obteniéndose los siguientes resultados:

$$FIV(X_1) = \frac{1}{1-0,995} = 200$$

$$FIV(X_2) = \frac{1}{1-0,99} = 100$$

$$FIV(X_3) = \frac{1}{1-0,97} = 33,33$$

$$FIV(X_4) = \frac{1}{1-0,996} = 250$$

Todos los valores son mayores que 10, por lo que se concluye que existe colinealidad en el modelo.

- c) Analice la posible existencia de colinealidad sabiendo que el máximo autovalor de la matrix $X'X$ es igual a 4.735 y el mínimo es igual a 0.001. RESOLUCIÓN:

Teniendo en cuenta que el número de condición se calcula como:

$$k(X) = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} = \sqrt{\frac{4,735}{0,001}} = 68,81.$$

se observa que toma un valor superior a 30 y por tanto se concluye que existe colinealidad en el modelo.

3. Calcule a partir de las siguientes matrices, la inversa y el determinante y concluya acerca de la relación entre las variables:

$$X'X = \begin{pmatrix} 1 & 0,95 \\ 0,95 & 1 \end{pmatrix}$$

$$X'X = \begin{pmatrix} 1 & 0,995 \\ 0,995 & 1 \end{pmatrix}$$

RESOLUCIÓN:

A partir de los datos anteriores, se obtiene:

$$(X'X)^{-1} = \begin{pmatrix} 10,26 & -9,74 \\ -9,74 & 10,26 \end{pmatrix}; |X'X| = 0,0975$$

$$(X'X)^{-1} = \begin{pmatrix} 100,25 & -99,75 \\ -99,75 & 100,25 \end{pmatrix}; |X'X| = 0,009975$$

4. A partir del fichero de muestra data7-3 de Ramanathan que contiene datos para 14 viviendas sobre el precio de venta de las viviendas (**price**), se pide estimar el precio en función de la superficie en pies cuadrados (**sqft**), el número de dormitorios (**bedrms**) y el número de baños (**baths**). Concluye sobre la posible existencia de colinealidad a partir de:

- a) Los coeficientes de correlación muestral. RESOLUCIÓN:

Coeficientes de correlación, usando las observaciones 1 - 14
valor crítico al 5 % (a dos colas) = 0.5324 para n = 14

price	sqft	bedrms	baths	
1,0000	0,9058	0,3156	0,6696	price
	1,0000	0,4647	0,7873	sqft
		1,0000	0,5323	bedrms
			1,0000	baths

- b) El coeficiente de determinación de las regresiones de cada una de las variables explicativas sobre el resto de las variables independientes del modelo. RESOLUCIÓN:

$$\hat{sqft} = -658 + 74,0 * bedrms + 975 * baths; R^2 = 0,623$$

$$\hat{bedrms} = 2,30 + 0,000103 * sqft + 0,488 * baths; R^2 = 0,289$$

$$\hat{baths} = 0,647 + 0,000532 * sqft + 0,191 * bedrms; R^2 = 0,655$$

- c) A partir del FIV. RESOLUCIÓN:

$$FIV(sqft) = 2,651$$

$$FIV(bedrms) = 1,406$$

$$FIV(baths) = 2,900$$

- d) A partir del NC Se obtiene que el minimo autovalor es igual a 0.008 y el maximo igual a 3.938, por lo tanto el numero de condición es igual:

$$K(X) = \sqrt{\frac{3,938}{0,008}} = 22,422$$

Notese que a partir de los FIV se concluiría que no existe multicolinealidad, mientras que el NC presenta un valor superior a 20 que indicaría multicolinealidad. Si se obtienen los estadísticos principales de las variables se observa que el rango de la variable **baths** es muy pequeño (varia entre 3 y 4) presentando un coeficiente de variación muy pequeño.

Estadísticos principales, usando las observaciones 1 - 14

Variable	Media	Mediana	Mínimo	Máximo
price	317.49	291.50	199.90	505.00
sqft	1910.9	1835.0	1065.0	3000.0
bedrms	3.6429	4.0000	3.0000	4.0000
baths	2.3571	2.2500	1.7500	3.0000
Variable	Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
price	88.498	0.27874	0.65346	-0.52983
sqft	577.76	0.30234	0.48526	-0.67212
bedrms	0.49725	0.13650	-0.59628	-1.6444
baths	0.44629	0.18934	0.33161	-1.3902
Variable	porc. 5 %	porc. 95 %	Rango IQ	Observaciones ausentes
price	indefinido	indefinido	154.50	0
sqft	indefinido	indefinido	832.75	0
bedrms	indefinido	indefinido	1.0000	0
baths	indefinido	indefinido	0.81250	0

5. A partir del fichero de muestra data4-7 de Ramanathan se dispone de una base de datos anuales sobre las tasas de mortalidad por enfermedades coronarias y sus determinantes para el periodo de 1947 a 1980 en E.E.U.U. Se pide:

- a) Especifica el modelo y razona sobre el signo esperado en cada coeficiente.
b) Estima el modelo por MCO. RESOLUCIÓN:

Modelo 1: MCO, usando las observaciones 1947–1980 ($T = 34$)
Variable dependiente: chd

	Coefficiente	Desv. Típica	Estadístico t	valor p
const	226.002	146.830	1.539	0.1363
cal	-69.9825	78.5568	-0.8909	0.3815
unemp	-0.613405	1.58641	-0.3867	0.7023
cig	10.1164	5.07126	1.995	0.0571
edfat	2.80992	1.66827	1.684	0.1046
meat	0.111592	0.243007	0.4592	0.6500
spirits	21.7156	8.45724	2.568	0.0166
beer	-3.46661	1.29774	-2.671	0.0131
wine	-4.56184	16.2472	-0.2808	0.7812
Media de la vble. dep.	354.8147	D.T. de la vble. dep.	14.94605	
Suma de cuad. residuos	1980.298	D.T. de la regresión	8.900109	
R^2	0.731364	R^2 corregido	0.645401	
$F(8, 25)$	8.507846	Valor p (de F)	0.000016	
Log-verosimilitud	-117.3428	Criterio de Akaike	252.6857	
Criterio de Schwarz	266.4229	Hannan-Quinn	257.3705	
$\hat{\rho}$	0.252080	Durbin-Watson	1.333807	

- c) Interpreta los coeficientes del modelo anterior.
- d) Comenta los resultados obtenidos de la estimación en terminos de bondad de ajuste, significatividad y signos de los coeficientes estimados. Razona si te parecen adecuados.
- e) Detecta la posible existencia de multicolinealidad a partir del FIV y del NC.

$$FIV(cal) = 5,786$$

$$FIV(unemp) = 1,967$$

$$FIV(cig) = 11,756$$

$$FIV(edfat) = 25,570$$

$$FIV(meat) = 6,250$$

$$FIV(spirits) = 16,703$$

$$FIV(beer) = 36,218$$

$$FIV(wine) = 49,066$$

Se obtiene que el minimo autovalor es igual a 0,00008 y el maximo igual a 8.644, por lo tanto el numero de condición es igual:

$$K(X) = \sqrt{\frac{8,644}{0,0008}} = 328,84$$

Si centramos las variables originales, el valor de los FIV se mantienen pero el del NC seria:

$$K(X) = \sqrt{\frac{6,502}{0,012}} = 23,63$$

Se habria corregido la multicolinealidad no esencial, pero seguiria existiendo colinealidad esencial.

6. A partir de los datos de [10] actualizados desde Junio de 2008 a abril de 2019 estimar el rendimiento a 52 semanas (c_{52}) en función del rendimiento a 13 y 26 semanas (c_{13} y c_{26} , respectivamente). Analizar la posible existencia de multicolinealidad y estimar mediante el metodo ortogonal. Este ejemplo es desarrollado completamente en [9]. Se especifica el siguiente modelo:

$$\mathbf{c}_{52} = \beta_1 + \beta_2 \mathbf{c}_{13} + \beta_3 \mathbf{c}_{26} + \mathbf{u}, \quad (1)$$

Obteniendose la siguiente matriz de correlación:

$$\begin{matrix} & \mathbf{c}_{52} & \mathbf{c}_{13} & \mathbf{c}_{26} \\ \mathbf{c}_{52} & 1,000 & & \\ \mathbf{c}_{13} & 0,981 & 1,000 & \\ \mathbf{c}_{26} & 0,995 & 0,993 & 1,000 \end{matrix} \left(\begin{matrix} 1,000 \\ 0,981 & 1,000 \\ 0,995 & 0,993 & 1,000 \end{matrix} \right).$$

Al estimarlo por MCO se obtiene:

MCO, usando las observaciones ($T = 131$)
Variable dependiente: r52

	Coefficiente	Desv. Típica	Estadístico t	valor p
const	0.0504730	0.00853647	5.913	0.0000
r13	-0.556909	0.0654995	-8.502	0.0000
r26	1.56218	0.0625815	24.96	0.0000
Media de la vble. dep.	0.696565	D.T. de la vble. dep.		0.781464
Suma de cuad. residuos	0.515394	D.T. de la regresión		0.063455
R^2	0.993508	R^2 corregido		0.993407
$F(2, 128)$	9794.297	Valor p (de F)		9.8e-141
Log-verosimilitud	176.8594	Criterio de Akaike		-347.7189
Criterio de Schwarz	-339.0933	Hannan-Quinn		-344.2139
$\hat{\rho}$	0.352018	Durbin-Watson		1.279550

Notese que el signo del parámetro β_2 no es coherente con la matriz de correlaciones. Si se calcula el FIV se obtiene un valor de 71.516 y para el NC se obtiene un valor igual a 23.233, que indicaría la existencia de multicolinealidad. En relación a la multicolinealidad no esencial se concluye que no existe ya que los coeficientes de variación de las variables no son bajos: ($CV(\mathbf{c}_{13}) = 1,486997$ y $CV(\mathbf{c}_{26}) = 1,280229$). Si se aplica el método de estimación ortogonal, en primer lugar, hay que estimar la siguiente regresión auxiliar:

$$\mathbf{c}_{26} = \alpha_1 + \alpha_2 \mathbf{c}_{13} + \mathbf{v}, \quad (2)$$

y, posteriormente, guardar dichos residuos y usarlos como variable explicativa sustituyendo a la variable c_{26} , de manera que se estima el siguiente modelo:

MCO, usando las observaciones ($T = 131$)
Variable dependiente: r52

	Coeficiente	Desv. Típica	Estadístico t	valor p
const	0.183112	0.00668113	27.41	0.0000
r13	1.06664	0.00774526	137.7	0.0000
e	1.56218	0.0625815	24.96	0.0000
Media de la vble. dep.	0.696565	D.T. de la vble. dep.		0.781464
Suma de cuad. residuos	0.515394	D.T. de la regresión		0.063455
R^2	0.993508	R^2 corregido		0.993407
$F(2, 128)$	9794.297	Valor p (de F)		9.8e-141
Log-verosimilitud	176.8594	Criterio de Akaike		-347.7189
Criterio de Schwarz	-339.0933	Hannan-Quinn		-344.2139
$\hat{\rho}$	0.352018	Durbin-Watson		1.279550

Notese que en este modelo se ha corregido el simbolo negativo del estimador del parámetro de la variable c_{13} . Si se calcula el FIV se obtendrá un valor igual a 1 (el minimo posible ya que ahora las variables explicativas son ortogonales) y el NC tendrá un valor de 1.878.

Referencias

- [1] García, C.B., Sánchez, J.M. y Salmerón, R. (2017) *Econometría básica para la economía y la empresa*. Ed. Fleming.
- [2] García, J., Jiménez, J.F. y Cerrillo, J.R. *Econometría práctica*. Edo. Libreria Universitaria de Almería.
- [3] Johnston, J. (1984) *Métodos de econometría*. Ed. Vicens Vives.
- [4] Pena, B., Estavillo, J., Galindo, E., Leceta, M. y Zamora, M. (1999). *Cien ejercicios de econometría*. Ed. Pirámide.
- [5] Sánchez, C., López, M.M. y García, T. (2015) *Econometría*. Ed. Fleming.
- [6] Matilla García, m., Pérez Pascual, Pedro y Sanz Carnero, B. (2013) *Econometría y predicción*. Mc Graw Hill.
- [7] Olva Maldonado, H. (2009). *Análisis de la función de producción Cobb-Douglas y su aplicación en el sector productivo Mexicano*. Tesis profesional. Universidad Autónoma de Chapingo, México. Dirección web: <http://bit.ly/1pFtQnn><http://bit.ly/1pFtQnn>.
- [8] Wikipedia: Función de producción de Cobb-Douglas. Dirección web: <http://bit.ly/1gfKeFP><http://bit.ly/1gfKeFP>.
- [9] García, C. B., Salmerón, R., García, C., García, J. (2019). *Residualization: justification, properties and application*. Journal of Applied Statistics, 1-21.
- [10] Wooldridge, J. (2008). *Introducción a la econometría. Un enfoque moderno*, 2nd ed., Thomson Paraninfo, Madrid.