

ECONOMETRÍA. GADE

Prácticas

Tema 5

Ejercicios resueltos

1. Utilizando una muestra de 25 observaciones anuales se estima el siguiente modelo de demanda:

$$D_t = \beta_1 + \beta_2 Y_t + \beta_3 PR_t + u_t$$

Utilizando sólo las 10 primeras observaciones se obtiene la siguiente ecuación estimada ¹:

$$\hat{D}_t = \begin{matrix} 80,50 \\ (86,17) \end{matrix} + \begin{matrix} 0,93Y_t \\ (1,06) \end{matrix} - \begin{matrix} 0,87PR_t \\ (1,9) \end{matrix}$$

donde se ha obtenido $SCR_1 = 125,7$. Del mismo modo, y utilizando las 10 últimas observaciones, se obtiene la siguiente ecuación estimada

$$\hat{D}_t = \begin{matrix} 20,61 \\ (221,44) \end{matrix} + \begin{matrix} 0,53Y_t \\ (0,29) \end{matrix} - \begin{matrix} 0,105PR_t \\ (2,41) \end{matrix}$$

con $SCR_2 = 498,94$. Detectar mediante el test de Goldfeld y Quandt la existencia o no de heteroscedasticidad.

Procediendo de la forma indicada se tiene que la hipótesis nula del contraste es

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$$

En esta situación el estadístico de contraste es:

$$F_{exp} = \frac{SCR_2}{SCR_1} = \frac{498,94}{125,7} = 3,9693$$

y se verifica que $F_{exp} \sim F_{\frac{n-c}{2}-k, \frac{n-c}{2}-k} = F_{\frac{25-5}{2}-3, \frac{25-5}{2}-3} = F_{7,7}$, donde se ha tenido en cuenta que el número de observaciones que han sido omitidas son $c = 5$ y el número de parámetros a estimar $k = 3$.

Como $F_{exp} > 3,79$ entonces se rechaza la hipótesis nula a un nivel de significación del 5 %. Por tanto, concluimos que existe heteroscedasticidad.

2. A partir de los siguientes datos se pide que se realice el contraste de Breusch Pagan. Para ello, se realizan los siguientes pasos:

¹Los valores que aparecen entre paréntesis corresponden a las desviaciones típicas estimadas de cada uno de los parámetros que describen el modelo.

y_i	x_i	\hat{y}_i	e_i	g_i	\hat{g}_i
55	80	60.31	-5.31	0.3582	0.0625
65	100	73.07	-8.07	0.8275	0.2638
70	85	63.5	6.5	0.5368	0.1128
80	110	79.45	0.55	0.0038	0.3644
79	120	85.82	-6.82	0.5910	0.4650
84	115	82.64	1.36	0.0235	0.4147
98	130	92.2	5.8	0.4274	0.5656
95	140	98.58	-3.58	0.1628	0.6663
90	125	89.01	0.99	0.0125	0.5153
75	90	66.69	8.31	0.8774	0.1631
74	105	76.26	-2.26	0.0649	0.3141
110	160	111.34	-1.34	0.0228	0.8675
113	150	104.96	8.04	0.8213	0.7669
125	165	114.52	10.48	1.3955	0.9178
108	145	101.77	6.23	0.4931	0.7166
115	180	124.09	-9.09	1.0498	1.0688
140	225	152.79	-12.79	2.0784	1.5216
120	200	136.85	-16.85	3.6074	1.2700
145	240	162.36	-17.36	3.8291	1.6725
130	185	127.28	2.72	0.0940	1.1191
152	220	149.6	2.4	0.0732	1.4713
144	210	143.23	0.77	0.0075	1.3706
175	245	165.55	9.45	1.1346	1.7228
180	260	175.11	4.89	0.3038	1.8738
135	190	130.47	4.53	0.2607	1.1694
140	205	140.04	-0.04	0.0000	1.3203
178	265	178.3	-0.3	0.0011	1.9241
191	270	181.49	9.51	1.1491	1.9744
137	230	155.98	-18.98	4.5771	1.5719
189	250	168.74	20.26	5.2153	1.7731

- Se ha estimado el modelo de regresión $\vec{y} = X\vec{\beta} + \vec{u}$ obteniendo, una vez aplicado el estimador mínimo cuadrático, $\vec{\hat{y}} = 9,29031 + 0,637785X$. Y se han obtenido los residuos realizando la operación $e_i = y_i - \hat{y}_i$.
- A partir de dicha información, se tiene que la Suma de los Cuadrados de los residuos generados es 2361.153. Luego, de forma inmediata se tendría que $\tilde{\sigma}^2 = \frac{SCR}{n} = \frac{2361,153}{30} = 78,7051$.
- Dividiendo los residuos al cuadrado entre $\tilde{\sigma}^2 = 78,7051$ se obtienen los residuos cuadráticos medios, g_i , a partir de los cuales se obtiene la estimación del modelo $g_i = \vec{z}_i^t \vec{\alpha} + u_i$, donde se considera que la variable que es la causante de la heteroscedasticidad es X_i . La estimación obtenida es $\hat{g}_i = -0,742509 + 0,0100626X_i$.
- El modelo estimado nos permite calcular la SCE como $\sum_{i=1}^{30} (\hat{g}_i - \bar{g})^2 = 10,4267$.
- Así pues, el estadístico de contraste tomaría el valor

$$\chi_{exp}^2 = \frac{SCE}{2} = \frac{10,4267}{2} = 5,21335$$

Considerando un nivel de significación del 5%, se tiene que como $\chi_{exp}^2 > \chi_{k-1,1-\alpha}^2 = \chi_{1,0,95}^2 = 3,8414$ se rechaza la hipótesis nula. Así pues, no se puede rechazar el supuesto de heteroscedasticidad.

3. A partir de los datos del ejemplo anterior, realizar el contraste de Glesjer sabiendo que:

y_i	55	65	70	80	79	84	...	178	191	137	189
X_i	80	100	85	110	120	115	...	265	270	230	250
$ e_i $	5.31	8.07	6.5	0.55	6.82	1.36	...	0.3	9.51	18.98	20.26

Se realizan los siguientes pasos:

- Para $h = \pm 1, \pm 2, \pm \frac{1}{2}$ se estima el modelo de regresión en el cual la variable endógena es el valor absoluto de los errores de estimación, $|e_t|$, y la variable exógena es la variable que pensamos que puede ser la causante del problema de heteroscedasticidad, X_i .

$h = 1$	$ \hat{e}_t = 1.1126 + 0.0331 \cdot z_t$ (3.6348) (0.01988)	$R^2 = 0.1188$
$h = -1$	$ \hat{e}_t = 11.0619 - 639.379 \cdot z_t^{-1}$ (1.9876) (282.044)	$R^2 = 0.0814$
$h = 2$	$ \hat{e}_t = 3.6566 + 9.56 \cdot 10^{-5} \cdot z_t^2$ (1.3159) ($3.34 \cdot 10^{-5}$)	$R^2 = \mathbf{0.1233}$
$h = -2$	$ \hat{e}_t = 8.5926 - 35036.3 \cdot z_t^{-2}$ (1.1709) (18621.3)	$R^2 = 0.0575$
$h = 1/2$	$ \hat{e}_t = -3.8895 + 0.8289 \cdot z_t^{1/2}$ (4.0212) (0.3056)	$R^2 = 0.1126$
$h = -1/2$	$ \hat{e}_t = 16.0441 - 115.249 \cdot z_t^{-1/2}$ (3.8293) (47.1948)	$R^2 = 0.0932$

- Utilizando el estadístico de contraste $|t_{exp}| = \left| \frac{\hat{\delta}_1}{\sqrt{\text{var}[\hat{\delta}_1]}} \right|$ (recordemos que los valores que aparecen entre paréntesis son las desviaciones típicas estimadas de cada uno de los parámetros del modelo) se estudia la significatividad del parámetro δ_1 de la regresión $|\vec{e}_t| = \delta_0 + \delta_1 z_t^h + v_t$. Si se rechaza la hipótesis nula entonces estamos ante un problema de heteroscedasticidad.

$h = 1$	$ t_{exp} = 1,6649 \not> t_{28,0,975} = 2,048$	Se mantiene H_0
$h = -1$	$ t_{exp} = 2,2669 > t_{28,0,975} = 2,048$	Se rechaza H_0
$h = 2$	$ t_{exp} = 2,8713 > t_{28,0,975} = 2,048$	Se rechaza H_0
$h = -2$	$ t_{exp} = 1,8815 \not> t_{28,0,975} = 2,048$	Se mantiene H_0
$h = 1/2$	$ t_{exp} = 2,7124 > t_{28,0,975} = 2,048$	Se rechaza H_0
$h = -1/2$	$ t_{exp} = 2,4419 > t_{28,0,975} = 2,048$	Se rechaza H_0

Luego, con las regresiones auxiliares para $h = -1; 2; 1/2; -1/2$ concluimos que existe heteroscedasticidad, siendo el modelo $|\hat{e}_t| = 3,6566 + 9,56 \cdot 10^{-5} \cdot z_t^2$ el mejor ya que es el que presenta el mayor coeficiente de determinación. En esta situación, podemos pensar que $E[u_t^2] = \sigma^2 z_t^2$.

En los casos primero y cuarto se concluye que no existe heteroscedasticidad ya que en ambos se mantiene la hipótesis nula.

- Supongamos que deseamos analizar los dividendos de una empresa en función de sus beneficios disponemos de las 20 observaciones de la siguiente tabla: Tras realizar la estimación por MCO se obtiene que:

$$\hat{D}_t = 10'2229 + 0'0638456 \cdot B_t$$

(1'35823) (0'011223)

con $R^2 = 0'642$ y donde entre paréntesis se especifica la desviación típica estimada de cada coeficiente estimado. Tras realizar la estimación por MCO del modelo original (ejemplo anterior)

Empresas	Dividendos	Beneficios
1	13'2	61
2	15	78
3	22'2	158
4	15'2	110
5	16'1	85
6	18'5	150
7	15'5	140
8	15	70
9	20	122
10	15	70
11	21	140
12	16'2	91
13	18'5	105
14	17	115
15	17'5	115
16	22	160
17	18	165
18	23	170
19	17	130
20	17	90

hemos obtenido los residuos a partir de los cuales hemos ajustado las siguientes regresiones auxiliares:

$$|\hat{e}_t| = 2'15499 - 8633'51 \cdot B_t^{-2} \quad R^2 = 0'3617$$

(0'3191) (2703'4)

$$|\hat{e}_t| = 3'1728 - 197'755 \cdot B_t^{-1} \quad R^2 = 0'4163$$

(0'5479) (55'1884)

$$|\hat{e}_t| = 5'212 - 40'6989 \cdot B_t^{-1/2} \quad R^2 = 0'4374$$

(1'0576) (10'8788)

$$|\hat{e}_t| = -2'93457 + 0'39747 \cdot B_t^{1/2} \quad R^2 = 0'4619$$

(1'09025) (0'1011)

$$|\hat{e}_t| = -0'891826 + 0'01888 \cdot B_t \quad R^2 = 0'4646$$

(0'5783) (0'004778)

$$|\hat{e}_t| = 0'138792 + 0'000079 \cdot B_t^2 \quad R^2 = 0'4536$$

(0'3424) (0'0000205)

En todos los casos rechazamos H_0 , por lo que hay heteroscedasticidad. Además, como el mejor modelo es el quinto (mayor coeficiente de determinación), pensamos que $E[u_t^2] = \sigma^2 B_t$.

Gracias al test de Glesjer hemos supuesto que $E[u_t^2] = \sigma^2 \cdot B_t$, es decir, la perturbación aleatoria depende directamente de los beneficios. Por tanto, lo que en teoría se denotaba como w_t corresponde a la variable beneficios, B_t . Luego para transformar los datos habrá que dividir por la raíz cuadrada de dicha variable:

$$D_t^* = \frac{D_t}{\sqrt{B_t}}, \quad cte_t^* = \frac{1}{\sqrt{B_t}}, \quad B_t^* = \frac{B_t}{\sqrt{B_t}} = \sqrt{B_t},$$

con $t = 1, \dots, 20$.

A partir de los datos transformados se obtiene la siguiente estimación por MCO del modelo transformado: $\hat{D}_t^* = 10'2147 \cdot cte_t^* + 0'0639 \cdot B_t^*$ con $R^2 = 0'9931$ (adviértase que los valores numéricos de las nuevas estimaciones no difieren mucho de las originales).

Nota: Se facilita estimación hecha en excel en Prado.

Empresa	B_t^*	D_t^*	cte_t^*
1	7'81025	1'690087	0'1280369
2	8'83176	1'698416	0'1132277
3	12'56981	1'766137	0'07956
4	10'48809	1'449263	0'09535
5	9'21954	1'746290	0'1084652
6	12'24745	1'510519	0'08165
7	11'83216	1'309989	0'08452
8	8'36660	1'792843	0'1195229
9	11'04536	1'810715	0'09054
10	8'36660	1'792843	0'1195229
11	11'83216	1'774824	0'08452
12	9'53939	1'698221	0'1048285
13	10'24695	1'805415	0'09759
14	10'72381	1'585258	0'09325
15	10'72381	1'631883	0'09325
16	12'64911	1'739253	0'07906
17	12'84523	1'401298	0'07785
18	13'03840	1'764019	0'07670
19	11'40175	1'490999	0'08771
20	9'48683	1'791957	0'1054093

Referencias

- [1] García, C.B., Sánchez, J.M. y Salmerón, R. (2017) Econometría básica para la economía y la empresa. Ed. Fleming.