

Modelos linealizables.

1.- Tras 10 años de funcionamiento, una empresa del sector de las telecomunicaciones, quiere estudiar el beneficio obtenido en dicho periodo en función al número de clientes que utilizan sus servicios. Los datos quedan recogidos en la siguiente tabla:

Beneficios	120	127	148	169	183	293	207	215	220	222
Nº Clientes	22	25	27	30	34	38	40	45	50	59

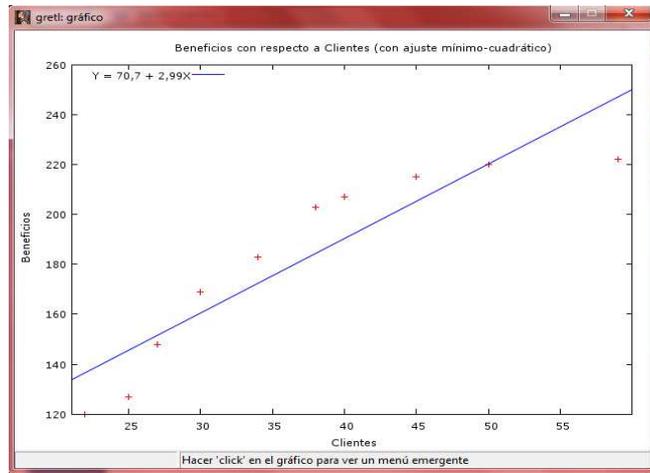
Se pide:

- a) Realice la representación gráfica correspondiente a los datos recogidos en la tabla.
- b) ¿Sería adecuado realizar un ajuste lineal?
- c) Obtenga la estimación de los parámetros de un modelo lineal, modelo semi-logarítmico y un modelo recíproco.
- d) Utilizando la información relacionada con el coeficiente de determinación, ¿Cuál de los tres modelos es más adecuado para describir el comportamiento de los beneficios de la empresa con respecto al número de clientes que tiene?
- e) ¿Qué beneficios esperamos que tenga la empresa cuando su número de clientes ascienda a 100? (nota: utilice el modelo que usted cree más adecuado)

Solución:

- a) Representación de los datos y el modelo lineal. Para obtener la representación gráfica de la variable beneficios en función al número de clientes se realizaran los siguientes pasos:
 - 1) En el menú que encontramos en la parte superior del programa seleccionamos Ver
 - 2) Gráficos
 - 3) Gráficos X-Y (scatter)
 - 4) Seleccionamos como variable dependiente (variable del eje Y) los Beneficios y como variable independiente (variable del eje X) El número de Clientes.
 - 5) Por último, seleccionamos aceptar y obtenemos en gráfico que aparece a continuación. La expresión lineal que aparece

en el gráfico es la correspondiente al modelo lineal estimado por gretl.



b) Utilizando la información que nos aporta el gráfico presentado en el apartado anterior, podemos observar que hay ciertos valores que no quedan ajustados, de una forma adecuada, por el modelo que nos ofrece el software gretl. Aún así, para poder concluir de una forma más correcta a la pregunta planteada, será adecuado recurrir, por ejemplo, al coeficiente de determinación.

c) **Ajuste lineal.** Con el objeto de obtener la estimación del modelo lineal, y para ir adaptándonos a los comandos que presenta Gretl, se utilizará a continuación la opción de Mínimos Cuadrados Ordinarios (M.C.O.). Para ello recurrimos a la opción MODELO en el menú superior de Gretl. A continuación seleccionamos MINIMOS CUADRADOS ORDINARIOS. Una vez que hemos seleccionado el método correcto para realizar la estimación es necesario incluir que variable será la independiente y cuál será la variable dependiente. Nótese que en el cuadro de las variables independientes aparece por defecto siempre const. Esta variable hace referencia a la constante del modelo. Una vez completados todos los pasos se obtiene la siguiente ventana, en la cual se pueden ver los valores de cada uno de los parámetros estimados así como algunas medidas estadísticas, de bondad,...

Teniendo en cuenta los resultados podemos decir que la expresión correspondiente al modelo lineal es $\hat{y}_t = 70.6951 + 2.99203x_t$, donde y_t son los beneficios y x_t son el número de clientes.

	Coefficiente	Desv. Típica	Estadístico t	Valor p	
const	70,6951	17,9556	3,937	0,0043	***
Clientes	2,99203	0,464480	6,442	0,0002	***
Media de la vble. dep.	181,4000	D.T. de la vble. dep.	38,57230		
Suma de cuad. residuos	2164,320	D.T. de la regresión	16,44810		
R-cuadrado	0,838368	R-cuadrado corregido	0,818164		
F(1, 8)	41,49508	Valor p (de F)	0,000200		
Log-verosimilitud	-41,07577	Criterio de Akaike	86,15154		
Criterio de Schwarz	86,75671	Crit. de Hannan-Quinn	85,48767		

Ajuste semi-logarítmico. La expresión de un modelo semi-logarítmico queda definida por $y_t = \alpha + \beta \ln x_t + u_t$. Realizando el cambio de variable $\ln x_t = z_t$, la expresión anterior queda reducida a un modelo lineal, $y_t = \alpha + \beta z_t + u_t$. Así pues, lo único que tendríamos que hacer es seguir los mismos pasos que se han aplicado en el caso anterior. En este caso la variable dependiente sigue siendo los Beneficios y la variable independiente es l_Clientes (logaritmo del número de clientes). Luego para realizar el ajuste definimos la variable l_Clientes siguiendo los siguientes pasos:

- 1) Seleccionamos la variable Clientes
- 2) En el menú superior, dentro de la opción AÑADIR, se selecciona LOGARITMOS DE LAS VARIABLES SELECCIONADAS.
- 3) A continuación, en el listado de variables, encontramos una nueva variable llamada l_Clientes.
- 4) Una vez descritas las variables, se realiza el ajuste lineal entre la variable Beneficios y l_Clientes.

	Coefficiente	Desv. Típica	Estadístico t	Valor p	
const	-234,964	42,3768	-5,545	0,0005	***
l_Clientes	116,773	11,8427	9,860	9,43e-06	***
Media de la vble. dep.	181,4000	D.T. de la vble. dep.	38,57230		
Suma de cuad. residuos	1018,039	D.T. de la regresión	11,28073		
R-cuadrado	0,923972	R-cuadrado corregido	0,914469		
F(1, 8)	97,22506	Valor p (de F)	9,43e-06		
Log-verosimilitud	-37,30463	Criterio de Akaike	78,60925		
Criterio de Schwarz	79,21442	Crit. de Hannan-Quinn	77,94538		

Con los resultados que se muestran en la Figura, podemos concluir que la expresión del modelo semi-logarítmico es $\hat{y}_t = -234.964 + 116.773 \ln x_t$.

Ajuste recíproco. En el caso del modelo recíproco, sabemos que la relación que se describe entre la variable dependiente e independiente viene dada por la expresión $y_t = \alpha + \beta \frac{1}{x_t} + u_t$. Considerando el cambio de variable $\frac{1}{x_t} = z_t$, la expresión del modelo recíproco coincide con la correspondiente al modelo lineal. Por tanto, lo único que deberíamos hacer es describir esa nueva variable y volver a utilizar M.C.O.

Pasos para describir la variable z_t :

1. Seleccionamos en el menú superior la opción AÑADIR.
2. Dentro de la opción añadir nos quedamos con DEFINIR NUEVA VARIABLE...
3. Al seleccionar la opción definir nueva variable se abre una nueva ventana en la cual introduciremos el nombre de la nueva variable y la expresión correspondiente al cambio que se desea hacer. Luego, dentro del cuadro escribimos $\text{Inv_Clientes}=1/\text{Clientes}$

Donde Inv_Clientes es la variable que se obtiene de realizar la inversa sobre la variable Clientes .

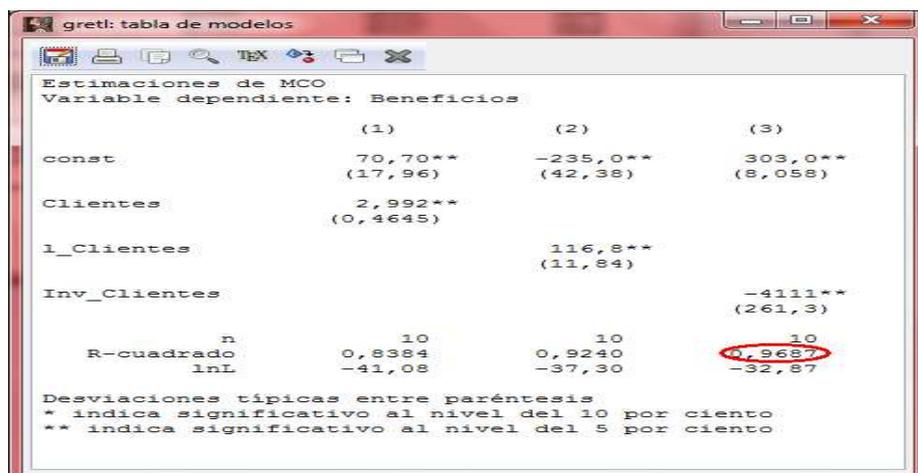
Una vez definida la nueva variable, se realiza el ajuste lineal entre la variable dependiente Beneficios y la variable independiente Inv_Clientes . En este caso, el modelo recíproco queda determinado por $\hat{y}_t = 302.983 - 4111.37 \frac{1}{x_t}$.

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	302,983	8,05796	37,60	2,75e-010 ***
Inv_Clientes	-4111,37	261,261	-15,74	2,66e-07 ***

Media de la vble. dep.	181,4000	D.T. de la vble. dep.	38,57230
Suma de cuad. residuos	419,0360	D.T. de la regresión	7,237368
R-cuadrado	0,968706	R-cuadrado corregido	0,964795
F(1, 8)	247,6420	Valor p (de F)	2,66e-07
Log-verosimilitud	-32,86624	Criterio de Akaike	69,73249
Criterio de Schwarz	70,33766	Crit. de Hannan-Quinn	69,06862

d) Para identificar cual de los tres modelos es el más adecuado, para describir el comportamiento de los beneficios de la empresa con respecto al número de clientes, estudiamos el valor del coeficiente de determinación. Cada una de las salidas obtenidas en el apartado anterior nos muestra el valor R_cuadrado. Alternativamente, se puede construir una tabla, como la que aparece en la siguiente figura, la cual recoge los valores estimados de cada uno de los parámetros así como el valor del coeficiente R_cuadrado. Para obtener esta salida es necesario seguir los siguientes pasos:

1. Cuando se realiza el ajuste de cada uno de los modelos probabilísticos, es necesario guardar SESIÓN COMO ICONO (dentro de la opción archivo de la ventana del modelo estimado)
2. Una vez guardado cada uno de los modelos nos vamos a la pantalla principal de Gretl. En ella en la parte inferior de la pantalla encontramos la opción VISTA DE ICONOS DE SESIÓN. Con esta opción se abre una ventana nueva donde aparece, entre otras cosas, los modelos que hemos estimado y Tabla de Modelos.
3. Situando el puntero del ratón sobre cada uno de los modelos y utilizando el botón derecho del mismo, seleccionamos la opción AÑADIR A LA TABLA DE MODELOS.
4. Una vez añadidos los tres modelos a la tabla de modelos, abrimos la tabla de modelos y obtenemos la ventana que aparece a continuación.



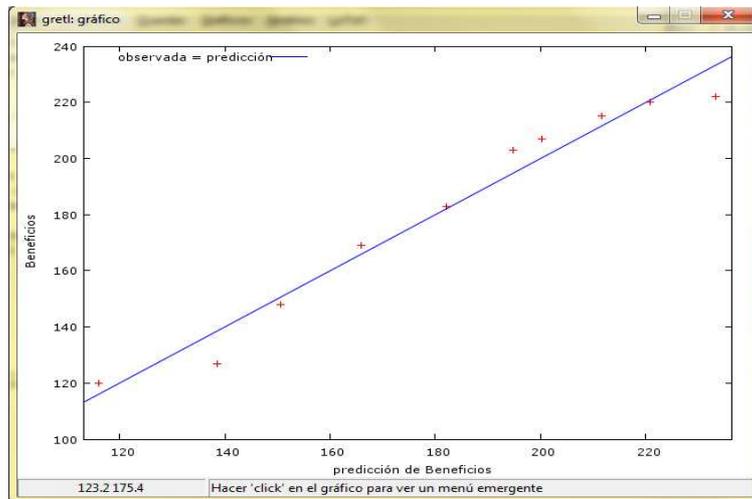
	(1)	(2)	(3)
const	70,70** (17,96)	-235,0** (42,38)	303,0** (8,058)
Clientes	2,992** (0,4645)		
l_Clientes		116,8** (11,84)	
Inv_Clientes			-4111** (261,3)
n	10	10	10
R-cuadrado	0,8384	0,9240	0,9687
lnL	-41,08	-37,30	-32,87

Desviaciones típicas entre paréntesis
* indica significativo al nivel del 10 por ciento
** indica significativo al nivel del 5 por ciento

El modelo que presenta el mayor valor de R-cuadrado corresponde al obtenido con el modelo recíproco. Por tanto, podemos concluir que sería adecuado recurrir a un modelo de la forma

$$\hat{y}_t = 302.983 - 4111.37 \frac{1}{x_t}$$

donde y_t es el beneficio de la empresa y x_t hace referencia a la variable número de clientes. Si representamos gráficamente (en la ventana del modelo estimado gráficos— gráfico de variable estimada y observada—observada vs. estimada), se observa que el modelo recíproco mejora los resultados del modelo lineal representado en el apartado a).



e) Utilizando la estimación obtenida con el modelo recíproco podemos decir que la empresa tendrá unos beneficios de 261.8693 cuando esta tenga 100 clientes.