Solución a la práctica 3 con Eviews

Ejercicio 1. Utilizando datos de 1569 empresas españolas del sector Industria para el año 2014, que se encuentran en el fichero *practica31.wf1*, se quiere explicar el coste de la producción vendida (*prod*) en euros utilizando como variables explicativas el número de empleados a tiempo completo (*emp*) y el inmovilizado material neto (*inm*) en euros. El modelo propuesto es el siguiente:

$$prod_i = \alpha \cdot emp_i^{\beta_1} \cdot inm_i^{\beta_2} \cdot u_i$$

a) **¿Es un modelo lineal en los parámetros? ¿Cómo se puede transformar el modelo para estimarlo por MCO?**

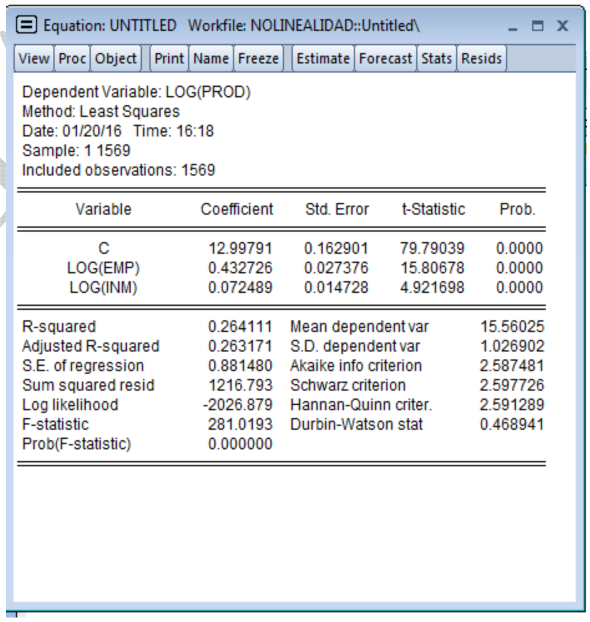
No es un modelo lineal en parámetros por lo que es necesario realizar un proceso de linealización. Aplicando logaritmos al modelo planteado, resulta el siguiente modelo transformado,

$$\log(prod_i) = \beta_0 + \beta_1 \log(emp_i) + \beta_2 \log(inm_i) + \varepsilon_i$$

que es un modelo log-log y a diferencia del modelo original ya es lineal en parámetros.

b) **Estime por MCO el modelo transformado e interprete la estimación de los coeficientes β_1 y β_2.**

Si usamos comandos, tenemos que escribir ***ls log(prod) c log(emp) log(inm)***.



Equation: UNTITLED Workfile: NOLINEALIDAD::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: LOG(PROD)
Method: Least Squares
Date: 01/20/16 Time: 16:18
Sample: 1 1569
Included observations: 1569

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12.99791	0.162901	79.79039	0.0000
LOG(EMP)	0.432726	0.027376	15.80678	0.0000
LOG(INM)	0.072489	0.014728	4.921698	0.0000

R-squared	0.264111	Mean dependent var	15.56025
Adjusted R-squared	0.263171	S.D. dependent var	1.026902
S.E. of regression	0.881480	Akaike info criterion	2.587481
Sum squared resid	1216.793	Schwarz criterion	2.597726
Log likelihood	-2026.879	Hannan-Quinn criter.	2.591289
F-statistic	281.0193	Durbin-Watson stat	0.468941
Prob(F-statistic)	0.000000		

La ecuación del coste de la producción que hemos estimado es:

$$\log(\hat{prod}_i) = 12.998 + 0.433 \log(emp_i) + 0.072 \log(inm_i)$$

Dado que es un modelo log-log, $\hat{\beta}_1$ mide la elasticidad del coste estimado de la producción respecto al trabajo. Es decir, la variación porcentual del coste estimado de la producción ante una variación de un 1% del número de empleados a tiempo completo. Como $\hat{\beta}_1$=0.433, si el número de empleados se incrementa en un 1%, el coste estimado de la producción aumentará aproximadamente en un 0.433%.

$$\hat{\beta}_1 \approx \frac{\Delta pr\hat{o}d \ / \ pr\hat{o}d}{\Delta emp \ / \ emp} = \frac{100 \cdot \Delta pr\hat{o}d \ / \ pr\hat{o}d}{100 \cdot \Delta emp \ / \ emp} = \frac{\% \Delta pr\hat{o}d}{\% \Delta emp}$$

Análogamente, la elasticidad del coste estimado de la producción respecto del inmovilizado neto es $\hat{\beta}_2 = 0.072$, de manera que si el inmovilizado neto aumenta un 1%, el coste estimado de la producción aumentará aproximadamente 0.072%.

$$\hat{\beta}_2 \approx \frac{\Delta pr\hat{o}d \ / \ pr\hat{o}d}{\Delta inm \ / \ inm} = \frac{100 \cdot \Delta pr\hat{o}d \ / \ pr\hat{o}d}{100 \cdot \Delta inm \ / \ inm} = \frac{\% \Delta pr\hat{o}d}{\% \Delta inm}$$

c) ¿Cuál es el valor estimado del logaritmo del coste de la producción vendida para una empresa que tiene 80 empleados a tiempo completo y un inmovilizado material neto de 9 millones de euros? ¿Cuál es el coste estimado de la producción vendida para dicha empresa?

El valor estimado del logaritmo del coste de la producción vendida es:

$$\log(pr\hat{o}d) = 12.998 + 0.433 \cdot \log(80) + 0.072 \cdot \log(9000000) = 16.048$$

Respecto al valor estimado para el coste de la producción, se puede calcular el valor *aproximado* para la estimación de la serie en nivel como la exponencial del del valor estimado para la serie en logaritmos:

$$pr\hat{o}d \approx \exp(\log(pr\hat{o}d)) = \exp(16.048) = 9323046.398 \text{ euros}$$

d) ¿Qué porcentaje de la variación muestral del logaritmo del coste de la producción es explicado por la función de regresión muestral?

Este porcentaje viene dado por el coeficiente de determinación. En este caso, el 26.41%.

Ejercicio 2. El siguiente modelo de regresión relaciona el logaritmo del gasto en consumo de los hogares (*lgasto*), el logaritmo de la renta del hogar (*lrenta*), el número de miembros del hogar (*miembros*) y la edad del sustentador principal (*edad*):

$$lgasto_i = \beta_0 + \beta_1 lrenta_i + \beta_2 miembros_i + \beta_3 edad_i + \beta_4 edad_i^2 + \varepsilon_i \, .$$

Usando datos de la *Encuesta de Presupuestos Familiares* del año 2013 (base 2006) se han obtenido datos de una muestra de 539 hogares que se encuentran en el fichero *practica32.wf1*. Responda a las siguientes cuestiones:

a) Interprete los efectos parciales o efectos *ceteris paribus* estimados de las variables explicativas.

La regresión muestral se obtiene escribiendo:

ls lgasto c lrenta miembros edad edad^2.

Equation: UNTITLED Workfile: GASTO::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: LGASTO
Method: Least Squares
Date: 01/21/16 Time: 13:10
Sample: 1 539
Included observations: 539

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.717565	0.389379	12.11560	0.0000
LRENTA	0.492854	0.037604	13.10626	0.0000
MIEMBROS	0.078985	0.016839	4.690463	0.0000
EDAD	0.022529	0.009375	2.403069	0.0166
EDAD*EDAD	-0.000247	9.02E-05	-2.739711	0.0064

R-squared	0.418319	Mean dependent var	10.30319
Adjusted R-squared	0.413962	S.D. dependent var	0.547630
S.E. of regression	0.419228	Akaike info criterion	1.108429
Sum squared resid	93.85159	Schwarz criterion	1.148222
Log likelihood	-293.7216	Hannan-Quinn criter.	1.123993
F-statistic	96.00742	Durbin-Watson stat	1.842364
Prob(F-statistic)	0.000000		

Por tanto, la ecuación del gasto en consumo de los hogares que hemos estimado es:

$$lga\hat{s}to_i = 4.718 + 0.493lrenta_i + 0.079miembros_i + 0.0225edad_i - 0.0002edad_i^2$$

Dada la muestra, la estimación del efecto *ceteris paribus* del logaritmo de la renta sobre el logaritmo del consumo de los hogares es $\hat{\beta}_1 = 0.493$. Como estas variables están en logaritmos, $\hat{\beta}_1$ se interpreta como la elasticidad del gasto en consumo de los hogares respecto de su renta. Esto es, si la renta de un hogar aumenta 1%, el gasto esperado en consumo de dicho hogar aumentará aproximadamente en 0.493%.

$$\hat{\beta}_1 \approx \frac{\Delta ga\hat{s}to / ga\hat{s}to}{\Delta renta / renta} = \frac{100 \cdot \Delta ga\hat{s}to / ga\hat{s}to}{100 \cdot \Delta renta / renta} = \frac{\%\Delta ga\hat{s}to}{\%\Delta renta}$$

En cuanto a $\hat{\beta}_2$,

$$\hat{\beta}_2 \approx \frac{\Delta ga\hat{s}to / ga\hat{s}to}{\Delta miembros}$$

Es decir, $\hat{\beta}_2$ representa la semielasticidad del gasto estimado en consumo de los hogares respecto del número de miembros. Multiplicando por 100 para expresar la variación del gasto estimado en consumo de los hogares en porcentaje, se tiene que

$$100 \cdot \hat{\beta}_2 \approx \frac{100 \cdot \Delta g\hat{as}to \,/\, g\hat{as}to}{\Delta miembros} \Rightarrow 100 \cdot \hat{\beta}_2 \approx \frac{\%\Delta g\hat{as}to}{\Delta miembros}$$

A partir de los resultados de la estimación, $100 \cdot \hat{\beta}_2 = 7.9$, y se interpreta como que, *ceteris paribus*, un hogar con un miembro más que otro gasta en promedio un 7.9% más, aproximadamente.

El efecto *ceteris paribus* de la *edad* sobre el logaritmo del gasto estimado en consumo de los hogares es aproximadamente:

$$\frac{\Delta lg\hat{as}to}{\Delta edad} \approx \hat{\beta}_3 + 2\hat{\beta}_4 edad_i = 0.0225 - 2 \cdot 0.0002 \cdot edad_i$$

Dado que el coeficiente estimado de *edad* es positivo y el de *edad*^{2} es negativo, el efecto marginal de *edad* es decreciente, y positivo únicamente hasta cierto valor de *edad*. En concreto, cuando aumenta la edad del sustentador principal en un año, el gasto esperado en consumo varía aproximadamente en:

$$\%\Delta g\hat{as}to \approx 100(0.0225 - 2 \cdot 0.0002 \cdot edad_i)$$

Nótese que $\hat{\beta}_3$ y $\hat{\beta}_4$ no tienen interpretación por separado.

b) Para un hogar cuyo sustentador principal tiene 25 años, ¿cuál es la variación esperada en el gasto en consumo si cumple un año más? ¿Y para un hogar que tiene un sustentador principal de 40 años?

Para un hogar cuyo sustentador principal tiene 25 años, la variación aproximada del logaritmo del gasto esperado en consumo por cumplir un año más es de $0.0225 - 2 \cdot 0.0002 \cdot 25 = 0.0125$. Esto supone un incremento del gasto en consumo esperado del 1.25% aproximadamente.

De la misma manera, para un hogar que tiene un sustentador principal de 40 años, un año adicional le reportará un incremento aproximado del gasto esperado en consumo de un 0.65% aproximadamente ($(0.0225 - 2 \cdot 0.0002 \cdot 40) \cdot 100 = 0.65$)

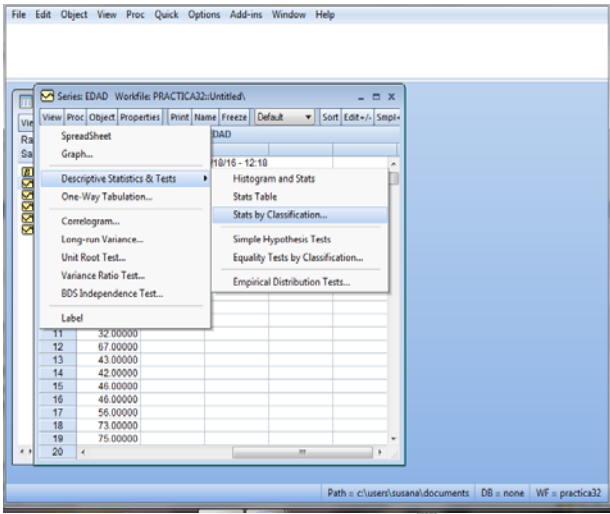
c) ¿A partir de qué edad del sustentador principal, un año más produce una disminución del gasto en consumo de los hogares? ¿Cuántos individuos de la muestra tienen o superan esa edad? Comente los resultados.

Tenemos que encontrar el valor de *edad* para el que el efecto marginal de *edad* sobre *lgasto* sea cero. Este es el punto donde

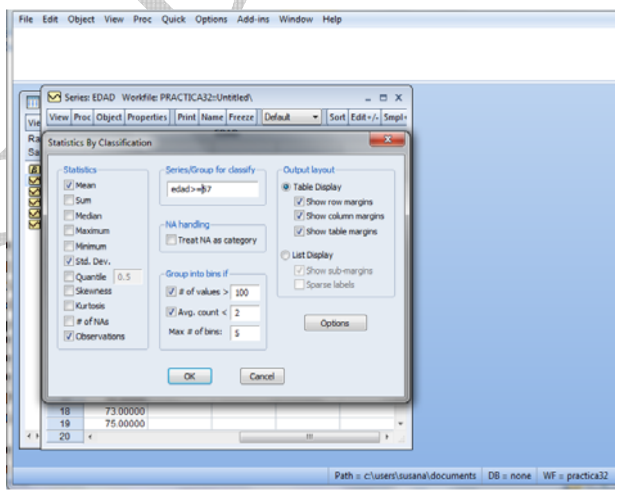
$$0.0225 - 2 \cdot 0.0002 \cdot edad_i = 0$$

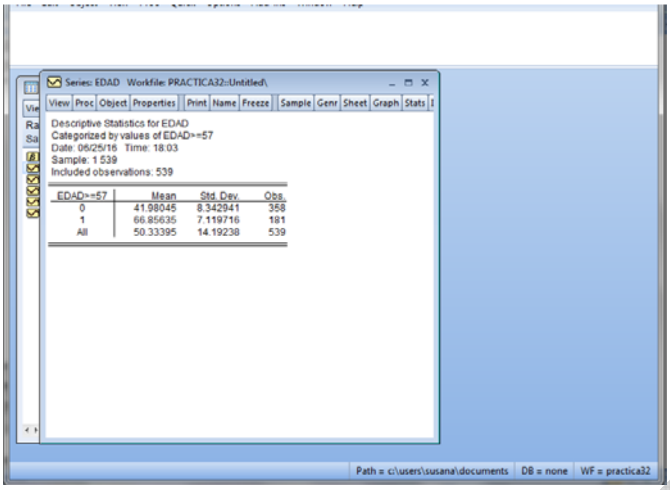
lo que ocurre para *edad*=56.25. Por tanto, si el sustentador principal de un hogar tiene 57 o más años, el efecto parcial de la edad cambia de signo y será negativo.

Para calcular el número de hogares cuyos sustentadores principales tienen 57 o más años, abrimos la variable *edad*, puesto que la información que nos pide el enunciado se refiere a esa variable. A continuación, segmentamos la muestra según la variable *edad* seleccionando ***View/Descriptive statistics/Stats by Classification*:**



En la pantalla que aparece a continuación, escribimos ***edad>=57*** en la casilla ***Series/Group for classify*** que aparece en blanco. Para el resto de las casillas dejamos las opciones que E-Views selecciona por defecto. Para finalizar, pinchamos en ***OK*** y aparecerá la siguiente pantalla:





Hay 181 hogares de 539 que tienen un sustentador principal de 57 o más años , lo que supone aproximadamente un 31% del total. Por tanto, en un 31% de los hogares el efecto parcial de la variable edad es negativo.