Universidad Carlos III de Madrid
Estadística II: Introduccion a la Econometría
Examen Final, Convocatoria Ordinaria, Curso 2007-2008.

21 de Junio de 2008

1. (2,5 puntos) Suponga que $\log(wage)$ denota el logaritmo del salario mensual, *educ* los años de formación académica, $\log(exper)$ el logaritmo de los años de experiencia profesional y *age* la edad.

a. (0,5 punto) Utilizando la siguiente ecuación interprete los coeficientes β_1, β_2:

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 \log(\text{exper}) + \beta_3 age + u$$

b. (0,75 puntos) Se estiman los siguientes modelos de regresión por mínimos cuadrados ordinarios:

$$\begin{aligned} \log(\widehat{wage}) &= 6.76 + \underset{(0{,}018)}{0.063} educ + \underset{(0{,}05)}{0.11} \log(\text{exper}) + \underset{se(\widehat{\beta}_3)}{0.0024} \text{age} \qquad (1) \\ R^2 &= 0.245,\ SSR = 3.546,\ n = 49 \end{aligned}$$

* (0,25 puntos) Contraste que la edad no afecta al salario.
* (0,25 puntos) Interprete el coeficiente de determinación en (1).
* (0,25 puntos) ¿Por qué el R^2 en (1) es mayor que en (2)?

b. (0,75 puntos) Suponga ahora que se ha estimado la siguiente ecuación:

$$\log(\widehat{wage}) = \beta_0 + u$$

El estimador de β_0 por mínimos cuadrados ordinarios es $\hat{\beta}_0 = 4,45$, $n = 49$.

* (0,25) Proporcione una expresión para $\hat{\beta}_0$ en términos de los datos.
* (0,25) ¿Qué valor tiene el coeficiente de determinación R^2?
* (0,25) Si la varianza muestral de $\log(\widehat{wage})$ es igual a 5.25, proporcione un estimador de la varianza de $\hat{\beta}_0$.

c. (0,5 puntos) Contraste la hipótesis de que las variables *educ*, $\log(exper)$ y *age* no afectan el salario (son conjuntamente no significativas) al nivel de significatividad del 5%.

2. (2,5 puntos) Suponga que se quiere evaluar cuál es el efecto de una incineradora de basuras sobre los precios de las casas. La construcción de la incineradora comenzó en 1981 y contamos con datos de precios, características y distancia a la incineradora de casas vendidas en 1978 (antes de que se empezase a construir) y en 1982 (después de que entrase en funcionamiento). Se quiere estudiar si las casas que estaban cerca de la incineradora perdieron valor respecto a las que estaban alejadas. Con este fin, empleando datos de 1982 se estima el siguiente modelo:

$$RPRICE = \gamma_0 + \gamma_1 NEARINC + \varepsilon \qquad (1)$$

donde:

$RPRICE$: precio de la casa en términos reales (dólares de 1978)
$NEARINC$: variable dummy que vale 1 si la casa está cerca de la incineradora (a menos de 3 millas) y 0 en caso contrario.
Los resultados de la estimación aparecen en la siguiente tabla:

SALIDA 1
OLS estimates using the 142 observations 180-321
Dependent variable: rprice

Variable	Coefficient	Std. Error	t-statistic	p-value
const	101308	3093.03	32.7535	<0.00001***
nearinc	-30688.3	5827.71	-5.2659	<0.00001***

Mean of dependent variable = 92662.9
Standard deviation of dep. var. = 34070.6
Sum of squared residuals = 1.36614e+011
Standard error of residuals = 31238
Unadjusted R2 = 0.165325
Adjusted R2 = 0.159363
Degrees of freedom = 140
Log-likelihood = -1670.1
Akaike information criterion = 3344.19
Schwarz Bayesian criterion = 3350.1
Hannan-Quinn criterion = 3346.59

a. (0,5 puntos)¿Implican los resultados de la anterior estimación que la causa de los precios más bajos de las casas cercanas a la incineradora es la presencia de la misma? Relacione su respuesta con el signo y la significatividad de los coeficientes estimados en la SALIDA 1.

b. (0,5 puntos) Interprete el coeficiente de la constante en la SALIDA 1. A la luz de los resultados de la SALIDA 1, ¿podemos decir que el precio medio de las viviendas en 1981 es de 70619 dólares?.

c. (0,5 puntos) Se estima de nuevo el Modelo (1), pero en esta ocasión empleando datos de 1978, obteniéndose los siguientes resultados:

SALIDA 2
OLS estimates using the 179 observations 1-179
Dependent variable: rprice

Variable	Coefficient	Std. Error	t-statistic	p-value
const	82517.2	2653.79	31.0941	<0.00000***
nearinc	-18824.4	4744.59	-3.9675	0.00011***

Mean of dependent variable = 76628
Standard deviation of dep. var. = 30626.4
Sum of squared residuals = 1.53324e+011
Standard error of residuals = 29432
Unadjusted R2 = 0.081671
Adjusted R2 = 0.0764827
Degrees of freedom = 177
Log-likelihood = -2094.87
Akaike information criterion = 4193.73
Schwarz Bayesian criterion = 4200.1
Hannan-Quinn criterion = 4196.32

A la luz de los resultados de la SALIDA 2, ¿se puede concluir que la incineradora se construyó en una zona en la que los precios de las casas eran más bajos?

d. (1 punto) Suponga que definimos una variable dummy, $d78$, que toma el valor 1 si la observación corresponde a una casa que fue vendida en el año 1978 y toma el valor 0 si fue vendida en el año 1982. Estimamos el siguiente modelo:$RPRICE = \beta_0 + \beta_1 NEARINC + \alpha_0 d78 + \alpha_1 (NEARINC * d78) + \varepsilon$

* (0,5 puntos) Especifique cómo contrastaría la hipótesis nula de que las casas vendidas en 1978 y en 1982 siguen la misma función de regresión.
* (0,5 puntos) Según el modelo ¿Cuál es el efecto neto sobre el precio de las casas que se encuentran cerca de la incineradora después de su construcción?

3. (2,5 puntos) Para explicar el efecto de los cursos de adiestramiento de los trabajadores sobre la productividad de las empresas manufactureras de Michigan, se considera el siguiente modelo:

$$\log(scrap) = \beta_0 + \beta_1 hrsemp + \beta_2 \log(sales) + \beta_3 \log(employ) + u$$

donde *scrap* denota una tasa de artículos desechados: el número de artículos defectuosos de cada 100 producidos que deben ser desechados, *hrsemp* son las horas anuales de adiestramiento por empleado, *sales* es el total facturado al año (en dólares), y *employ* es el número de trabajadores en la empresa. La ecuación estimada es:

$$\log(\widehat{scrap}) = \underset{(4.91)}{13.72} - \underset{(0.019)}{0.028} hrsemp - \underset{(0.41)}{1.21} \log(sales) + \underset{(0.43)}{1.48} \log(employ)$$

$$n = 30, \quad R^2 = 0.431.$$

Los errores estándar se encuentran en paréntesis.

a. (0,5 puntos) Si los empleados en una empresa son adiestrados 5 horas más al año que en otra empresa con la misma facturación anual y el mismo número de trabajadores, ¿cuál es la diferencia estimada entre las dos empresas en la tasa de artículos desechados *scrap*?

b. (0,5 puntos) Demuestre que el modelo puede reescribirse como

$$\log(scrap) = \beta_0 + \beta_1 hrsemp + \beta_2 \log\left(\frac{sales}{employ}\right) + \theta_3 \log(employ) + u \tag{1}$$

en términos del parámetro $\theta_3 = \beta_2 + \beta_3$. Ayuda: $\log(x/y) = \log(x) - \log(y)$.

c. (1 punto)

* (0,5 puntos) Escriba el modelo en forma reducida bajo la hipótesis nula $H_0 : \beta_1 = -\beta_3$. Escriba la hipótesis nula en términos del parámetro θ_3.
* (0,5 puntos) Explique como contrastar H_0 en la dirección de que $H_1 : \beta_1 \neq -\beta_3$ utilizando el R^2 en el modelo restringido y no restringido.

d. (0,5 puntos) El modelo en su formulación (1) ha sido estimado como sigue:

$$\log(\widehat{scrap}) = \underset{(4.57)}{11.74} - \underset{(0.019)}{0.042} hrsemp - \underset{(0.370)}{0.951} \log(\tfrac{sales}{employ}) + \underset{(0.205)}{0.041} \log(employ)$$

$$n = 43, \quad R^2 = 0.31.$$

Contraste la hipótesis nula $H_0 : \beta_1 = -\beta_3$ en la dirección $H_1 : \beta_1 \neq -\beta_3$.

4. (2,5 puntos)La siguiente ecuación de regresión describe el precio promedio de las viviendas en una localidad en términos de la contaminación *nox* (concentración de óxido nítrico),

$$\log(price) = \alpha_0 + \alpha_1 \log(nox) + u_1.$$

También se considera una ecuación que tiene en cuenta el número de habitaciones en la vivienda (*rooms*):

$$\log(price) = \beta_0 + \beta_1 \log(nox) + \beta_2 rooms + u_2.$$

Suponiendo que los errores u_1 y u_2 no están correlacionados con ninguna variable explicativa

a. (1 punto) Suponiendo que $\beta_2 \neq 0$, en que circunstancia $\beta_1 = \alpha_1$.

b. (1 punto) En base a una muestra aleatoria se estima las siguiente ecuaciónes:

$$\begin{aligned} \log(\widehat{price}) &= \underset{(0.132)}{11.71} - \underset{(0.078)}{1.043} \log(nox). \\ n &= 506, \qquad R^2 = 0.264 \end{aligned} \tag{2}$$

$$\begin{aligned} \log(\widehat{price}) &= \underset{(0.188)}{9.23} - \underset{(0.066)}{0.718} \log(nox) + \underset{(0.019)}{0.306} rooms. \\ n &= 506, \qquad R^2 = 0.514 \end{aligned} \tag{3}$$

Proporcione una relación entre los estimadores de α_1 en (2) y de β_1 en (3).

c. (0,5 puntos) Si la varianza muestral de $\log(nox)$ es $0,75$, ¿cuál es el valor de la covarianza muestral entre $\log(nox)$ y *rooms*?