Departamento de Fundamentos del Análisis Económico.
Universidad de Alicante. Curso 2011/12

ECONOMETRÍA I

Hoja de problemas del Tema 5

1.- La siguiente ecuación describe el precio medio de las viviendas de un barrio (*price*) en función del nivel de contaminación (*nox*) y del número medio de habitaciones de las viviendas del barrio (*rooms*):

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

a) ¿Qué miden β_1 y β_2? ¿Cuáles son sus signos esperados?

b) ¿Por qué *nox* y *rooms* pueden estar negativamente correlacionadas? Si están negativamente correlacionadas, la regresión simple de $\log(price)$ sobre $\log(nox)$ ¿produce un estimador de β_1 con sesgo positivo o negativo?.

c) Utilizando una muestra de 506 barrios del área de Boston (Estados Unidos) se han obtenido los siguientes resultados

$$\log(\widehat{price}) = 11.71 - 1.043 \log(nox)$$
$$R^2 = 0.264$$

$$\log(\widehat{price}) = 9.23 - 0.718 \log(nox) + 0.306 rooms$$
$$R^2 = 0.514$$

¿Es la relación entre los estimadores de regresión simple y múltiple de la elasticidad del precio de la vivienda respecto al nivel de contaminación la prevista según su respuesta del apartado **b)**?

2.- Un transportista quiere relacionar los kilómetros recorridos por su camión (K_i) con los litros repostados (L_i), para lo que utiliza un modelo de regresión lineal simple $K_i = \beta_1 + \beta_2 L_i + u_i$. Supondremos que este modelo satisface todas las hipótesis del MRL con errores normales. Disponemos de 7 datos, sobre los que sabemos:

$$\sum_{i=1}^{7} L_i = 1384; \quad \sum_{i=1}^{7} K_i = 2605; \quad \sum_{i=1}^{7} L_i^2 = 296792;$$
$$\sum_{i=1}^{7} L_i K_i = 558354; \quad \sum_{i=1}^{7} K_i^2 = 1050531.$$

a) Obtenga el intervalo de confianza al 95% para el número de kilómetros que en media recorrá el camión cuando reposte 140 litros.

b) Obtenga el intervalo de predicción al 95% para el número de kilómetros que recorrerá el camión cuando reposta 140 litros; explique qué significa este intervalo y por qué es lógico que sea mayor que el obtenido en el apartado anterior.

c) El transportista conjetura que los kilómetros que cabe esperar que recorra el camión cuando reposta 100 litros son exactamente 100 menos que los kilómetros que cabe esperar que recorra el camión cuando reposta 200 litros. Contraste esta hipótesis.

3.- Considere el siguiente modelo para el precio de la vivienda

$$\log(price) = \beta_1 + \beta_2 sqrtft + \beta_3 bdrms + u$$

donde *price* es el precio de la vivienda en miles de dólares, $sqrtft$ es la superficie de la vivienda en cientos de pies cuadrados y *bdrms* es el número de dormitorios. En base a una muestra de 88 viviendas se han obtenido los siguientes resultados

$$\widehat{\log(price)} = 4.77 + \underset{(0.0043)}{0.0379} sqrtft + \underset{(0.0296)}{0.0289} bdrms$$

Los números entre paréntesis son los errores estándar.

a) Estime el cambio porcentual en el precio de la vivienda si se añade un dormitorio adicional de 150 pies cuadrados a la vivienda.

b) Sabiendo que la covarianza estimada entre $\widehat{\beta}_2$ y $\widehat{\beta}_3$ es -0.000068 calcule un intervalo de confianza al 95% para el cambio porcentual estimado en **a)** y contraste si es estadísticamente significativo.

c) Si en lugar del modelo anterior, se hubiese estimado el modelo de regresión simple

$$\log(price) = \beta_1 + \beta_2 sqrft_t + u_t$$

¿esperaríamos que la estimación obtenida para β_2 en el primer modelo fuese mayor o menor que la obtenida al estimar el modelo de regresión simple?

4.- Analice si las siguientes afirmaciones son verdaderas o falsas. Si son verdaderas, demuéstrelas; si son falsas, justifique por qué.

a) Considere el modelo lineal

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 X_t^2 + u_t$$

que verifica los supuestos básicos del MRL. Si $\sum_{t=1}^{T} X_t = \sum_{t=1}^{T} X_t^3 = 0$ y omitimos X_t^2, aunque β_3 sea distinto de cero, obtendremos un estimador insesgado de β_2.

b) En un modelo $Y_t = X_t'\beta + u_t$ que cumple todas las hipótesis del MRL queremos predecir el valor de Y_{T+1} sabiendo que $X_{T+1} = X_T$. Si se utiliza como predictor $\widetilde{Y}_{T+1} = Y_T$ entonces el error cuadrático medio de predicción es $ECMP(\widetilde{Y}_{T+1}) = E[(\widetilde{Y}_{T+1} - Y_{T+1})^2] = 2\sigma^2$.

c) En el modelo $Y_t = \beta_1 + \beta_2 X_t + u_t,\ t = 1, ..., T$, que satisface las hipótesis básicas del MRL con errores normales, la longitud del intervalo de predicción para Y_{T+1} coincidirá con la longitud del intervalo de confianza para el valor esperado de la variable Y_{T+1}.

5.- Un economista pretende predecir la cantidad de dinero que recaudará el estado el próximo año a través del impuesto sobre la renta de las personas físicas, $IRPF$. Para ello recoge datos para los años 1986 a 2007 y estima el modelo:

$$\log(IRPF_t) = \beta_1 + \beta_2 paro_t + \beta_3 saldo\_veg_t + u_t$$

donde $IRPF_t$ es el dinero recaudado por el estado español en impuestos durante el año t, $paro_t$ es la tasa de paro (en tanto por ciento) en el año t y $saldo\_veg_t$ es el saldo vegetativo durante el año t, es decir número de nacimientos menos número de muertes, por cada mil habitantes. Los resultados de la estimación aparecen en la tabla adjunta. Si para el próximo año s, se estima una tasa de paro del 20% y un saldo vegetativo igual a 1.13 , ¿qué valor esperaremos obtener para la variable $\log(IRPF_s)$? Obtener un intervalo de predicción al 95% para $\log(IRPF_s)$ sabiendo que $x_s' \ \widehat{var(\widehat{\beta})} \ x_s = 0.0067$, donde $\widehat{\beta} = (\widehat{\beta}_1 \ \widehat{\beta}_2 \ \widehat{\beta}_3)'$ y $x_s = (1, 20, 1.13)'$

Dependent Variable: LIRPF
Method: Least Squares
Date: 11/20/09 Time: 11:48
Sample: 1986 2007
Included observations: 22

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	18.29972	0.187467	97.61588	0.0000
PARO	-0.059176	0.011510	-5.141201	0.0001
SALDO_VEG	-0.220300	0.050696	-4.345520	0.0003

R-squared	0.702087	Mean dependent var	17.12533
Adjusted R-squared	0.670728	S.D. dependent var	0.372465
S.E. of regression	0.213729	Akaike info criterion	-0.122093
Sum squared resid	0.867921	Schwarz criterion	0.026686
Log likelihood	4.343021	F-statistic	22.38853
Durbin-Watson stat	0.820210	Prob(F-statistic)	0.000010