

ECONOMETRÍA I - CURSO 2016

PRACTICO 5

Variables Ficticias – Inconstancia del vector  $\beta$

EJERCICIO 1

Utilizando los datos correspondientes al año 1985, contenidos en el archivo **cps\_ej1\_practico5.gdt** estime las siguientes regresiones, previo análisis del significado económico de las variables utilizadas en ellos:

$$[1] \text{ wage}_i = \beta_1^I + \beta_2^I \text{ female}_i + \varepsilon_i^I$$

$$[2] \text{ wage}_i = \beta_1^{II} + \beta_2^{II} \text{ female}_i + \beta_3^{II} \text{ educ}_i + \beta_4^{II} \text{ exper}_i + \varepsilon_i^{II}$$

$$[3] \ln\_ \text{ wage}_i = \beta_1^{III} + \beta_2^{III} \text{ female}_i + \beta_3^{III} \text{ educ}_i + \beta_4^{III} \text{ exper}_i + \beta_5^{III} (\text{exper}_i)^2 + \varepsilon_i^{III}$$

Ecuación [4]:

$$\ln\_ \text{ wage}_i = \beta_1^{IV} + \beta_2^{IV} \text{ female}_i + \beta_3^{IV} \text{ educ}_i + \beta_4^{IV} \text{ exper}_i + \beta_5^{IV} (\text{exper}_i)^2 + \beta_6^{IV} \text{ female}_i \times \text{educ}_i + \varepsilon_i^{IV}$$

Se pide:

- 1) Determine el salario esperado en dólares para hombres y mujeres utilizando la ecuación [1]. Someta a prueba la hipótesis nula de que el coeficiente de la variable *female* es igual a cero indicando cuál es la hipótesis alternativa (en todos los contrastes solicitados en este ejercicio utilice un nivel de significación del 5%). Interprete económicamente el hecho de que dicho coeficiente (teórico) sea o no significativamente distinto de 0. Explique cómo se interpreta la estimación puntual del mismo.
- 2) Explique formalmente qué problema se generaría si a la ecuación [1] se le agregara un regresor *male*, que vale 1 si es varón y 0 en otro caso.
- 3) Indique cómo se interpreta en la ecuación [2] el coeficiente asociado a la variable *female*, indicando la diferencia entre éste y el obtenido en la ecuación [1]. Interprete y ofrezca una justificación de la diferencia en términos de la magnitud de ambos coeficientes.
- 4) Indique cómo se interpreta en la ecuación [3] el coeficiente asociado a la variable *female*, indicando la diferencia entre éste y el obtenido en la ecuación [2].
- 5) En la ecuación [3] se incluyó como regresor adicional la experiencia al cuadrado. Explique qué tipo de comportamiento se está intentando recoger al introducir dicho regresor. Señale si cambiaría (y en caso afirmativo cómo) la magnitud del coeficiente asociado a la variable experiencia al cuadrado si a ésta se la dividiera por 10.
- 6) Notar que en la ecuación [4] las variables *female* y *female x educ* resultan individualmente significativas. Explique cómo interpreta dicho resultado.

## EJERCICIO 2

Utilice los datos incluidos en ECH.gdt. La información proviene de la Encuesta de Ingresos y Gastos de los Hogares realizada por el INE en el año 2006. El objeto es analizar la elasticidad ingreso del gasto en vivienda.

1. Analice los principales estadísticos de las variables gasto de vivienda (*gvivienda*) e ingreso total del hogar (*ing\_tot*).
2. Aplique la transformación logarítmica a las variables anteriores. ¿Por qué razón se generaron valores ausentes? ¿Qué consecuencia tiene este hecho en el análisis a realizar? ¿Por qué razón se realiza la transformación logarítmica?
3. Observe gráficamente la relación entre las variables transformadas. Estime por MCO la siguiente regresión:  
$$\ln gviv_i = \beta_1 + \beta_2 \ln ing\_tot_i + \varepsilon_i, i = 1, 2, \dots [1]$$
4. ¿Cuál es la interpretación del coeficiente  $\beta_2$ , asociado a la variable  $\ln ing\_tot$ ? Realice la prueba de significación de ese coeficiente.
5. Someta a prueba la hipótesis de que  $\beta_2 < 1$ . ¿Qué se pretende verificar con ese contraste?

Hasta ahora sólo se tuvo en cuenta el ingreso del hogar. No obstante, es pertinente preguntarse si la elasticidad ingreso del gasto en vivienda es la misma para diferentes subconjuntos de la población y, en particular, si es diferente en Montevideo respecto del Interior.

6. ¿Qué supuesto del MRLC se estaría violando si resulta que la elasticidad es diferente según se haga referencia a Montevideo o al Interior?
7. Realice la estimación de la siguiente ecuación:  
$$\ln gviv_i = \beta_1 + \delta_1 MVD_i + \beta_2 \ln ing\_tot_i + \delta_2 MVD_i \times \ln ing\_tot_i + \varepsilon_i [2],$$
 donde la variable  $MVD_i$  toma el valor 1 si el hogar es montevideano y 0 en caso contrario. En adelante será el Modelo 2.
8. Interprete los coeficientes de la ecuación del Modelo 2 estimado. Comente.
9. Para someter a prueba la hipótesis de igualdad de las elasticidades entre distintos subconjuntos de la población se propone utilizar el contraste de cambio estructural de Chow. ¿En qué consiste dicha prueba?
10. Estime el modelo considerando sólo las observaciones correspondientes a Montevideo y luego haga la estimación considerando sólo las observaciones correspondientes al Interior.
11. Escriba (si es posible) la relación entre los coeficientes de los modelos [1], [2], [3] y [4].
12. Realice el contraste de cambio estructural de Chow por dos caminos alternativos. En ambos casos establezca: (i) las hipótesis nula y alternativa del contraste, (ii) la forma del estadístico de contraste y su distribución de probabilidad (en este último caso indique el rol que juega el tamaño de la muestra), (iii) el valor y la región crítica del contraste. Calcule si es necesario el valor del estadístico y concluya.

### EJERCICIO 3

Se toma como ejemplo (con datos simulados) la especificación de una ecuación salarial para profesionales universitarios. Las variables utilizadas son las siguientes:

VARIABLE	DEFINICIÓN
Salario	Salario nominal mensual
Exper	Años de experiencia laboral
Mujer	1 si es mujer, 0 en otro caso
Varon	1 si es varón, 0 en otro caso
Medicina	1 si tiene título de medicina, 0 en otro caso
Economía	1 si tiene título de economía, 0 en otro caso
Ingeniería	1 si tiene título de ingeniería, 0 en otro caso
Humanidades	1 si tiene título de humanidades, 0 en otro caso
Posgrado	1 si tiene título de posgrado, 0 en otro caso

A continuación siguen los resultados de la estimación para distintas especificaciones del modelo:

Dependent Variable: LOG(SALARIO)				
Included Observations: 600				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12.72541	0.016746	759.907441	0,0000
EXPER	0.017224	0.002950	5.83864407	0,0000
EXPER^2	-0.000691	0.000118	-5.8559322	0,0000
MUJER	-0.089584	0.011561	-7.74881066	0,0000
R-squared	0.141438	Mean dependent var		12.75184
Adjusted R-squared	0.137117	S.D. dependent var		0.152242
S.E. of regression	0.14142	Akaike info criterion		-1.067527
Sum squared resid	11.91969	Schwarz criterion		-1.038215
Log likelihood	324.2582	F-statistic		32.72804
Durbin-Watson stat	1.440394	Prob(F-statistic)		0,0000

Dependent Variable: LOG(SALARIO)				
Included Observations: 600				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12.63583	0.016382	771.3438	0.00000
EXPER	0.017224	0.00295	5.838669	0.00000
EXPER^2	-0.000691	0.000118	-5.88111	0.00000
VARON	0.089584	0.011561	7.748953	0.00000
R-squared	0.141438	Mean dependent var		12.75184
Adjusted R-squared	0.137117	S.D. dependent var		0.152242
S.E. of regression	0.14142	Akaike info criterion		-1.067527
Sum squared resid	11.91969	Schwarz criterion		-1.038215
Log likelihood	324.2582	F-statistic		32.72804
Durbin-Watson stat	1.440394	Prob(F-statistic)		0,0000

Dependent Variable: LOG(SALARIO)				
Included Observations: 600				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
EXPER	0.017224	0.00295	5.838669	0.00000
EXPER^2	-0.000691	0.000118	-5.88111	0.00000
VARON	12.72541	0.016746	759.8943	0.00000
MUJER	12.63583	0.016382	771.3438	0.00000
R-squared	0.141438	Mean dependent var		12.75184
Adjusted R-squared	0.137117	S.D. dependent var		0.152242
S.E. of regression	0.14142	Akaike info criterion		-1.067527
Sum squared resid	11.91969	Schwarz criterion		-1.038215
Log likelihood	324.2582	F-statistic		32.72804
Durbin-Watson stat	1.440394	Prob(F-statistic)		0.00000

Se pide:

- 1) ¿Cómo se interpreta el coeficiente asociado a la variable MUJER en la primera salida?
- 2) Observe el coeficiente estimado asociado a la variable VARON en la segunda especificación y examine que el salario medio esperado para Hombres es de 336182 y para Mujeres de 307376, o sea que las Mujeres perciben en promedio el 91.43% del salario que percibiría un hombre con su misma experiencia. Compare este resultado con el de la tercera estimación.
- 3) ¿Por qué no se incluye constante en la tercera estimación del modelo?

Se quiere ahora analizar el efecto que tiene sobre el salario el egreso de distintas carreras. Los resultados de las estimaciones de dos especificaciones alternativas son los siguientes:

Dependent Variable: LOG(SALARIO)				
Included Observations: 600				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12.55238	0.018774	668.5909	0.00000
EXPER	0.018685	0.002754	6.784199	0.00000
EXPER^2	-0.000747	0.00011	-6.799556	0.00000
MEDICINA	0.152526	0.019177	7.953631	0.00000
ECONOMIA	0.099588	0.016024	6.214852	0.00000
INGENIERO	0.193054	0.015929	12.1197	0.00000
R-squared	0.253759	Mean dependent var		12.75184
Adjusted R-squared	0.247477	S.D. dependent var		0.152242
S.E. of regression	0.132067	Akaike info criterion		-1.201071
Sum squared resid	10.36031	Schwarz criterion		-1.157101
Log likelihood	366.3212	F-statistic		40.39786
Durbin-Watson stat	1.889841	Prob(F-statistic)		0.00000

Dependent Variable: LOG(SALARIO)				
Included Observations: 600				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
EXPER	0.018685	0.002754	6.784199	0.00000
EXPER^2	-0.000747	0.00011	-6.799556	0.00000
MEDICINA	12.70491	0.019656	646.3554	0.00000
ECONOMIA	12.65197	0.016349	773.89	0.00000
INGENIERO	12.74544	0.016168	788.3309	0.00000
HUMANIDADES	12.55238	0.018774	668.5909	0.00000
R-squared	0.253759	Mean dependent var		12.75184
Adjusted R-squared	0.247477	S.D. dependent var		0.152242
S.E. of regression	0.132067	Akaike info criterion		-1.201071
Sum squared resid	10.36031	Schwarz criterion		-1.157101
Log likelihood	366.3212	F-statistic		40.39786
Durbin-Watson stat	1.889841	Prob(F-statistic)		0.00000

- 1) ¿Cómo se interpretan los coeficientes asociados a las variables ficticias de los distintos títulos en ambas salidas? ¿Qué diferencial de salario esperado hay entre un economista y un ingeniero que tengan idénticos años de experiencia?
- 2) Exprese las hipótesis nula y alternativa apropiadas para las siguientes pruebas (en caso de contar con los datos necesarios realice las pruebas):
  - no existen diferencias salariales para los graduados en medicina y humanidades;
  - no existen diferencias salariales para los graduados en medicina y economía;
  - no existen diferencias salariales para los graduados en medicina y economía respecto a los graduados en humanidades;
  - los distintos títulos no generan diferencias salariales.
- 3) Especifique un modelo que permita evaluar el efecto conjunto que tienen el género y la experiencia sobre el salario.

### EJERCICIO 3

Con datos trimestrales sobre Ventas y Utilidades de una empresa manufacturera se estimó por MCO el modelo siguiente:

$$\text{Utilidades}_t = \beta_0 + \beta_1 \text{Ventas}_t + u_t$$

La muestra fue de 24 observaciones, desde el primer trimestre de 1995 al cuarto trimestre del 2000. Las variables Ventas y Utilidades fueron deflactadas con el índice de precios al consumidor. Los resultados de la regresión lineal realizada son los siguientes:

Variable dependiente: Utilidades

Included observations: 24

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.543495	1.825754	3.583995	0.0017
Ventas	0.041222	0.011878	3.470538	0.0022
R-squared	0.353789	Mean dependent var		12.82500
Adjusted R-squared	0.324416	S.D. dependent var		1.428666
S.E. of regression	1.174276	Akaike info criterion		3.238837
Sum squared resid	30.33635	Schwarz criterion		3.337008
Log likelihood	-36.86604	F-statistic		12.04463
Durbin-Watson stat	1.117061	Prob(F-statistic)		0.002172

Al observar los datos de Ventas y Utilidades se concluyó que ambas variables durante el año alcanzan guarismos más altas en el segundo trimestre que en los restantes, quizás por un efecto estacional. Para investigarlo se realizó la regresión siguiente:

Included observations: 24

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.630881	1.695470	3.910939	0.0009
D2	1.314967	0.632272	2.079747	0.0513
D3	-0.218014	0.626084	-0.348219	0.7315
D4	0.173991	0.647936	0.268532	0.7912
Ventas	0.038563	0.011374	3.390592	0.0031
R-squared	0.531649	Mean dependent var		12.82500
Adjusted R-squared	0.433049	S.D. dependent var		1.428666
S.E. of regression	1.075731	Akaike info criterion		3.166929
Sum squared resid	21.98674	Schwarz criterion		3.412357
Log likelihood	-33.00315	F-statistic		5.391966
Durbin-Watson stat	0.387680	Prob(F-statistic)		0.004490

D2, D3 y D4 son variables Dummy incluidas para reflejar el efecto sobre el término independiente del factor estacional mencionado.  $D_i = 1$  en el  $i$ -ésimo trimestre; 0 en el resto; con  $i = 2, 3, 4$ .

Se pide:

- 1) Exprese el modelo estimado en la segunda regresión y cómo queda conformada la matriz X en este caso. ¿Podría incluirse una variable D1 que valga uno en el primer trimestre y cero en el resto?
- 2) Compruebe que por cada peso que aumentan las ventas, en los dos modelos, las utilidades esperadas estimadas se incrementan en aproximadamente 0.04.
- 3) Contraste la hipótesis de que los coeficientes de D2, D3 y D4 son conjuntamente no significativamente distintos de cero. Especifique la hipótesis nula, la matriz de restricciones, R, y el estadístico utilizado.
- 4) ¿Cómo escribiría el modelo a ser estimado en la segunda regresión si desea investigar coeficientes diferenciales de la variable ventas según los trimestres? ¿Cómo queda conformada la matriz X en este caso?

#### EJERCICIO 4

Utilizando datos correspondientes a la economía americana durante el período enero 1975 – diciembre 2000 (datos trimestrales), se ha estimado la siguiente función de demanda de rosas:

$$\log Y_t = \beta_1 + \beta_2 \log X_{2t} + \beta_3 \log X_{3t} + \beta_4 \log X_{4t} + \beta_5 D_1 + \beta_6 D_2 + \beta_7 D_3 + u_t$$

Donde las variables son:

- Y: cantidad demandada de rosas (medida en docenas)
- $X_2$ : precio promedio al por mayor de los claveles (U\$/docena)
- $X_3$ : precio promedio al por mayor de las rosas (U\$/docena)
- $X_4$ : ingreso disponible familiar semanal (promedio, en U\$)
- $D_1, D_2$  y  $D_3$  son variables ficticias que han sido definidas de forma que toman el valor uno si la observación corresponde al primer, segundo o tercer trimestre respectivamente.

Basándose en los resultados que se proporcionan en la tabla que sigue, responda a las siguientes preguntas:

- 1) Observando los coeficientes de las variables: ¿en qué trimestre se demandan más rosas?
- 2) ¿Por qué no se incluyó una  $D_4$  en la especificación del modelo econométrico?
- 3) Analice la bondad de ajuste y contraste la significación global del modelo.

Variable Dependiente: LY				
Regresores: 1, LX <sub>2</sub> , LX <sub>3</sub> , LX <sub>4</sub> , D <sub>1</sub> , D <sub>2</sub> , D <sub>3</sub>				
Muestra: 1975.1 – 2000.4 No. Observaciones: 104				
Regresores	Coefficiente	Desv. Típica	Estadístico t	Prob > t
1	5.389710	0.339587	15.87	0.0000
LX <sub>2</sub>	0.602212	0.026776	22.49	0.0000
LX <sub>3</sub>	-0.450580	0.037258	-12.09	0.0000
LX <sub>4</sub>	0.349478	0.033149	10.54	0.0000
D <sub>1</sub>	1.526778	0.077966	19.58	0.0000
D <sub>2</sub>	0.732776	0.079113	9.26	0.0000
D <sub>3</sub>	-1.561168	0.079586	-19.62	0.0000
Media Variable Dependiente	0.9753	Desv. Típica Var. Depend.	1.4348	
Error Típico Regresión	0.2787	Suma Cuadrados Residuos	7.5336	
R cuadrado	0.9645	R cuadrado corregido	0.9623	
Logaritmo Verosimilitud	-11.0685	Criterio AIC	0.3475	
Estadístico F (6, 97)	438.8403	Prob > F	0.0000	
Estadístico Durbin-Watson	1.6549	Est. Autocorrelación	0.1725	

## EJERCICIO 5

Se desea analizar los determinantes del monto de alquiler que se cobra por las viviendas en Uruguay. Para ello se dispone de la Encuesta Continua de Hogares para el año 2012, donde se cuenta con las siguientes variables:

<b>numhab:</b>	Número de habitaciones en la vivienda
<b>numdorm:</b>	Número de habitaciones utilizadas para dormir
<b>numbaños:</b>	Número de baños en la vivienda
<b>casa:</b>	Variable binaria (1=casa)
<b>apto:</b>	Variable binaria (1=apartamento)
<b>numprob:</b>	Número de problemas en la vivienda (goteras, humedades, grietas, etc.)
<b>numprob2:</b>	Variable <i>numprob</i> al cuadrado
<b>mdeo:</b>	Variable binaria (1=Montevideo)
<b>intsur:</b>	Variable binaria (1=Departamento del Interior y al Sur del Río Negro)
<b>asentam:</b>	Variable binaria (1=Vivienda ubicada en Asentamiento Irregular)
<b>asentam_mdeo:</b>	Variable de interacción entre <i>asentam</i> y <i>mdeo</i>

Se estima un modelo cuya salida es la siguiente:

Source	SS	df	MS	Number of obs = 120462		
Model	2.2819e+12	10	2.2819e+11	F( 10,120451) = 7325.55		
Residual	3.7520e+12	120451	31149580	Prob > F = 0.0000		
Total	6.0339e+12	120461	50089864.3	R-squared = 0.3782		
				Adj R-squared = 0.3781		
				Root MSE = 5581.2		

  

alquiler	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
numhab	964.9304	20.10127	48.00	0.000	925.5323	1004.329
numdorm	-512.4182	27.77583	-18.45	0.000	-566.8583	-457.978
numbaños	5984.709	37.06243	161.48	0.000	5912.067	6057.351
casa	-634.3849	44.18185	-14.36	0.000	-720.9806	-547.7892
numprob	-1728.703	178.3719	-9.69	0.000	-2078.309	-1379.097
numprob2	94.02991	7.419977	12.67	0.000	79.48688	108.5729
mdeo	3182.734	51.07786	62.31	0.000	3082.623	3282.846
intsur	397.0093	45.97766	8.63	0.000	306.8938	487.1248
asentam	-324.0618	125.4632	-2.58	0.010	-569.9676	-78.1559
asent_mdeo	-3373.212	151.8753	-22.21	0.000	-3670.885	-3075.539
_cons	2861.641	1060.012	2.70	0.007	784.0337	4939.248

Se pide:

- 1) Interprete los coeficientes asociados a las variables *numhab*, *numdorm* y *numbaños*. ¿Son razonables los signos de los coeficientes? Justifique.
- 2) Interprete el coeficiente asociado a la variable *casa*. Suponga que se estima una regresión con las mismas variables pero agregando la variable *apto* y quitando la constante. ¿Cuánto sería el coeficiente asociado a *casa*? ¿Y el coeficiente asociado a *apto*?
- 3) Interprete conjuntamente los coeficientes asociados al número de problemas de la vivienda. ¿Cuál sería el efecto esperado sobre el alquiler de pasar de 0 a 1 problema? ¿Y de pasar de 4 a 5 problemas en la vivienda?
- 4) Interprete los coeficientes asociados a las variables *mdeo* y *intsur*. Señale cuál es la variable



omitida y explique detalladamente el porqué de tal omisión.

- 5) ¿En cuánto se reduce en promedio el monto de un alquiler si la vivienda se encuentra ubicada en un asentamiento en Montevideo? ¿Y si se trata de un asentamiento en el interior? Detalle los cálculos que son necesarios para obtener el intervalo de confianza asociado a este último efecto.

## EJERCICIO 6

Con datos de una muestra de hogares, se estudió para un país europeo el gasto de consumo familiar según la raza del jefe de hogar. El modelo especificado fue el siguiente:

$$C_i = \beta_0 + \beta_1 Y_i + \beta_2 N_i + \beta_3 O_i + u_i$$

Donde:

$C_i$  = Consumo de la familia  $i$ -ésima

$Y_i$  = Ingreso disponible de la familia  $i$

$N_i$  = Variable dummy que vale 1 para las familias con jefe de hogar de raza negra y 0 para los demás casos

$O_i$  = Variable dummy que vale 1 para las familias con jefe de hogar de raza asiática y 0 en los demás casos

$u_i$  = Término de perturbación (independientes, con distribución Normal, media 0, y varianza  $\sigma^2$ )

Se pide:

- 1) ¿Cómo contrastaría la hipótesis de que la raza no tiene efecto sobre el consumo?
- 2) Construya un modelo en el que pueda estudiar la discriminación racial en el mercado laboral en lo que respecta al ingreso (por ejemplo, cada grupo racial tiene una constante y un  $\beta_1$  diferente). ¿Cómo contrastaría la hipótesis implícita en 1., bajo el supuesto de que el ingreso familiar no depende del origen racial del jefe de hogar ?