

Taller #5
Econometría 06169

Profesor: Julio César Alonso
Monitores: Hernán Betancur
David Valencia

Notas:

- Recuerde que tres preguntas, seleccionadas al azar, serán calificadas.
- Este taller es para ser entregado entre las 8:30 am y 9:30 am del 19 de febrero en mi oficina.

INSTRUCCIONES:

- Este taller debe ser escrito en computador.
- Los cálculos deben ser realizados en EasyReg
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.
- Este taller es un trabajo en pareja. Por tanto el taller debe reflejar **únicamente** el trabajo de la pareja.

Pregunta 1

Usted como investigador de la empresa el maracuya.com se encuentra preocupado porque la empresa no posee una buena idea sobre el comportamiento del consumo de Maracuyá por persona (G_i), por tanto usted inicia un proceso de recolección de información concerniente a 45 individuos en la ciudad de Tangamandapio y al mismo tiempo recolecta el ingreso disponible también para cada individuo (Y_i). Dicha información se encuentra en el archivo T4-01-07a. Los datos se encuentran en miles de pesos por persona. Considerando el siguiente modelo:

$$\ln(G_i) = \alpha + \beta \ln(Y_i) + \varepsilon_i \quad (1)$$

- a. Comente el signo de los coeficientes a priori
- b. Con la información anterior estime el modelo (empleando EasyReg) y muestre sus resultados en una tabla
- c. Interprete los coeficientes estimados.

Pregunta 2

- a. Continuando con el gasto de Maracuyá en Tangamandapio, discuta la significancia de los coeficientes. Discuta el fit del Modelo.
- b. Construya la Tabla ANOVA para este ejercicio.

Pregunta 3

El año pasado la institución encargada de los estudios estadísticos en una isla del caribe recolectó gran cantidad de información a través de un censo. Ud como encargado del departamento de policía, desea determinar que variables explican el comportamiento de robos en la isla. Para dicho estudio usted cuenta con la siguiente información.

$ARRB_i$ Número de arrestos como porcentaje de robos en cada provincia.

$Dens_i$ Densidad de población por cada provincia

$\ln(Rbaño)_i$ Logaritmo natural de la tasa anual de robos por cada 100.000 habitantes

$Despl_i$ Porcentaje de la población que es desplazada en cada provincia.

$Ingr_i$ Ingreso per capita en cada provincia medido en dólares.

El modelo a estimar es el siguiente:

$$\ln(Rbaño)_i = \beta_0 + \beta_1 ARRB_i + \beta_2 Dens_i + \beta_3 Despl_i + \beta_4 \ln(Ingr)_i + \varepsilon_i$$

Los datos se encuentran en el archivo T4-01-07a.xls.

- a) ¿Qué tipo de datos estamos utilizando?
- b) Explique los signos de los coeficientes a priori
- c) Estime el modelo y muestre los resultados de dicha estimación.
- d) Analice la significancia individual y conjunta de los coeficientes

Pregunta 4

Continuando con el ejercicio anterior

- a) Mientras recopilaba información para hacer este estudio, usted encontró que estudios previos mostraban que el coeficiente que acompaña al porcentaje de personas desplazadas tenía un efecto tres veces mayor que el coeficiente asociado al ingreso per capita. Verifique esta información con respecto a su estimación y muestre el procedimiento.

Pregunta 5

A la hora de presentar sus resultados en el congreso de la isla, un importante y reconocido economista, afirma que su modelo es bien intencionado pero no está correctamente formulado. En especial, el economista afirma que la variable arrestos no debería ir en el modelo. Finalizada esta intervención un matemático de una importante universidad refuta lo que dice el estadístico y afirma que la variable que sobra es la concerniente a la densidad de la población porque según él dicha variable no aporta al modelo. Ya a punto de usted defender su modelo, el presidente del congreso que muy joven había sido profesor de econometría refuta la intervención de los dos anteriores y afirmar que la variable que sobra es la concerniente a la densidad de la población porque según él dicha variable no aporta al modelo. Ya a punto de usted defender su modelo, el presidente del congreso que muy joven había sido profesor de econometría refuta la intervención de los dos anteriores y afirmar que la variable que sobra es la concerniente a la densidad de la población porque según él dicha variable no aporta al modelo. ¿Cuál de estas personas está en lo correcto?

Taller #5

Respuestas Sugeridas
Econometría 06169**Profesor: Julio César Alonso****Monitores: Hernán Betancur**
David Valencia**Notas:**

- Recuerde que tres preguntas, seleccionadas al azar, serán calificadas.
- Este taller es para ser entregado entre las 8:30 am y 9:30 am del 19 de febrero en mi oficina.

INSTRUCCIONES:

- Este taller debe ser escrito en computador.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.

Este taller es un trabajo en pareja. Por tanto el taller debe reflejar **únicamente** el trabajo de la pareja.

Pregunta 1

Usted como investigador de la empresa el maracuya.com se encuentra preocupado porque la empresa no posee estimaciones optimas sobre el consumo de Maracuyá por persona (G_i), por tanto ud. inicia un proceso de recolección de información concerniente a 45 individuos en la ciudad de Tangamandapio y al mismo tiempo recolecta el ingreso disponible también para cada individuo (Y_i). Dicha información se encuentra en el archivo T4-01-07a. Los datos se encuentran en miles de pesos por persona.

- a. Comente el signo de los coeficientes a priori
Dado que tenemos el gasto en Maracuyá en función de la renta disponible, se puede asumir que el signo asociado al β será positivo dado que el Maracuyá puede clasificarse como un bien normal. El intercepto sería como el gasto en Maracuyá a un nivel de renta cero, lo cual carece de sentido y por tanto de interpretación económica.
- b. Con la información anterior estime el siguiente modelo y muestre sus resultados
$$\ln(G_i) = \alpha + \beta \ln(Y_i) + \varepsilon_i \quad (1)$$

La Estimacion hecha vía EasyReg arroja los siguientes resultados

Tabla 1, Resultados para el modelo 1

VARIABLE DEPENDIENTE: Ln(Gi)		
Estadísticos t entre paréntesis		
MCO		
Constante	-2.56 (-3.32)	***
Ln(Yi)	0.687 (6.40)	***
R ²	0.5190	
R ² Ajustado	0.5063	
# de Obs.	40	

(*) nivel de significancia: 10%

(**) nivel de significancia: 5%

(***) nivel de significancia: 1%

MCO: Mínimos Cuadrados Ordinarios

c. Interprete los coeficientes asociados al modelo

Para interpretar β_1 en el modelo $\ln(G_i) = \alpha + \beta \ln(Y_i) + \varepsilon_i$, se debe derivar la función como sigue a continuación:

$$\beta_1 = \frac{\partial \ln(G_i)}{\partial \ln(Y_i)} \approx \frac{\Delta \ln(G_i)}{\Delta \ln(Y_i)} = \frac{\frac{\Delta G_i}{G_i}}{\frac{\Delta Y_i}{Y_i}} = \frac{\Delta G_i}{\Delta Y_i} \cdot \frac{Y_i}{G_i} = E_{G,Y}$$

Podemos entonces interpretar β_1 como la elasticidad renta de la demanda. Por lo tanto cuando la renta crece un 1%, la demanda lo hará un 0.6865%. El coeficiente α correspondiente al intercepto carece de interpretación económica.

Pregunta 2

a. Continuando con el gasto de Maracuyá en Tangamandapio, discuta la significancia de los coeficientes. Discuta el fit del Modelo.

Al analizar los datos se puede observar que el primer coeficiente, el intercepto posee un nivel de significancia de 99% lo que permite asumir que este es relevante a la hora de explicar el gasto en Maracuyá. El coeficiente asociado a la renta personal igualmente tiene una significancia del 99% lo que permite entonces asumir que los coeficientes estimados son diferentes de cero. El fit del modelo, como medida para determinar que tanto el gasto en Maracuyá es explicado por la renta es mostrado a través del R2, el cual en este caso es de 0.5190 lo que nos permite afirmar que el modelo explica de una manera regular la relación entre el gasto en Maracuyá y la renta según este modelo.

b. Construya la Tabla ANOVA para este ejercicio.

La tabla ANOVA está formada por las siguientes ecuaciones:

Tabla 2. Tabla ANOVA.

Fuente de la Variación	SS	Grados de libertad	MS
Regresión	2.683847	1	2.683847
Error	0.265758	40 - 2	0.00699
Total	5.579597	39	

Pregunta 3

El año pasado la institución encargada de los estudios estadísticos en una isla del caribe recolectó gran cantidad de información a través de un censo. Ud como encargado del departamento de policía, desea determinar que variables explican el comportamiento de robos en la isla. Para dicho estudio usted cuenta con la siguiente información.

$ARRB_i$ Numero de arrestos como porcentaje de robos en cada provincia.

$Dens_i$ Densidad de población por cada provincia

$Ln(Rbaño)_i$ Logaritmo natural de la tasa anual de robos por cada 100.000 habitantes

$Despl_i$ Porcentaje de la población que es desplazada en cada provincia.

$Ingr_i$ Ingreso per capita en cada provincia medido en dólares.

El modelo a estimar es el siguiente:

$$Ln(Rbaño)_i = \beta_0 ARRB_i + \beta_1 Dens_i + \beta_2 Despl_i + \beta_3 Ln(Ingr)_i + \epsilon_i$$

Los datos se encuentran en el archivo T4-01-07a.xls.

a) Explique los signos de los coeficientes a priori

$ARRB_i$ Se esperaría que a mayor numero de arrestos como porcentaje de robos el numero de robos en las provincias disminuyera. Por tanto el signo seria negativo.

$Dens_i$ Se esperaría que a mayor densidad de población, mayor cantidad de robos. El signo seria positivo.

$Despl_i$ Se esperaría que a mayor porcentaje de desplazados, mayor cantidad de robos. El signo seria positivo.

$Ln(Ingr)_i$ Se esperaría que a mayor ingreso per capita, menos cantidad de robos. El signo seria negativo.

b) Estime el modelo y muestre los resultados de dicha estimación.

Tabla 3. Modelos Estimados.

VARIABLE DEPENDIENTE: Ln(Rbaño)						
Estadísticos t entre paréntesis						
	Ecuación 1		Ecuación 2		Ecuación 3	
	MCO		Estadístico		Matematico	
					Presidente C.	
Constante	12.68		12.24		10.3884	-
	(11.50)	***	(10.580)	***	(9.697)	***
ARRBi	-0.0023		-		-0.0027	0.00432
	(-8.53)	***	-		(-8.32)	*** (12.30)
Densi	0.00003		0.00003		-	-0.00003
	(6.35)	***	(6.073)	***	-	(-3.49) ***
Despli	0.030		0.0330		0.0555	0.14302
	(5.17)	***	(5.497)	***	(13.46)	*** 16.614 ***
Ln(Ingr)i	-1.171		-1.1526		-0.9335	-
	(-9.94)	***	(-9.312)	***	(-3.21)	***
R ²	0.3340		0.2644		0.2954	-1.0819
R ² Ajustado	0.3301		0.2613		0.2924	-1.0878
F	87.370	***	83.63	***	97.56	***
# de Obs.	702		702		702	702

(*) nivel de significancia: 10%
 (**) nivel de significancia: 5%
 (***) nivel de significancia: 1%
 MCO: Mínimos Cuadrados Ordinarios

c) Analice la significancia individual y conjunta de los coeficientes

Como se puede observar todos los coeficientes presentes en el modelo son significativos individualmente a un nivel del 99%. Al analizar conjuntamente la significancia de los coeficientes a través del F global que arroja EasyReg se puede encontrar que con un 87.37 la hipótesis nula de que todos los coeficientes son iguales a cero es rechazada a un nivel del 5% y 10%. De igual manera, aplicando una prueba de Wald para significancia conjunta se encuentra que la hipótesis de que son iguales a cero se rechaza a un nivel del 5% y 10% al encontrar un valor de Wald igual a 5360.

Pregunta 4

Continuando con el ejercicio anterior

a) Que tipo de datos estamos utilizando?

Los datos son corte transversal.

b) Mientras recopilaba información para hacer este estudio, usted encontró que estudios previos mostraban que el coeficiente que acompaña al porcentaje de

personas desplazadas tenía un efecto tres veces mayor que el coeficiente asociado al ingreso per capita. Verifique esta información con respecto a su estimación y muestre el procedimiento.

Se trata de comprobar la hipótesis nula de: $H_0 : 3\beta_3 = \beta_4$, lo cual se puede expresar como: $H_0 : 3\beta_3 - \beta_4 = 0$ versus la hipótesis alterna de $H_A : \text{No } H_0$. Para comprobar estas hipótesis, podemos reescribirlas de la forma $R\beta = C$, en donde:

$$R = [0 \quad 0 \quad 0 \quad 3 \quad -1] \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} \quad C = [0]$$

Utilizando estos datos, será muy fácil utilizar una prueba de Wald en Easyreg para comprobar las hipótesis aquí planteadas. El estadístico Wald (113.19), contrastado con los valores críticos 2.71 y 3.84 para un nivel de significancia del 5% y 10% respectivamente, rechazan la hipótesis nula, por lo que las apreciaciones del otro trabajo sobre los coeficientes que acompañan al porcentaje de personas desplazada y el Ln de la renta no son verídicas.

Pregunta 5

A la hora de presentar sus resultados en el congreso de la isla, un importante y reconocido estadístico, afirma que su modelo esta correctamente formulado pero que la variable arrestos no debería ir allí. Finalizada esta intervención un matemático de una importante universidad refuta lo que dice el estadístico y afirma que la variable que sobra es la concerniente a la densidad de la población porque según él dicha variable no aporta al modelo. Ya a punto de usted defender su modelo, el presidente del congreso que muy joven había sido profesor de econometría refuta la intervención de los dos anteriores y afirmar que la variable que sobra es la concerniente a la concerniente al ingreso además de que el intercepto es irrelevante. Cual de estas personas esta en lo correcto¹?

Para poder refutar esta información usted debe estimar cada modelo propuesto por esta personas y compararlo por el que usted estimó. Los resultados se pueden ver en el tabla 3.

Con respecto al primer modelo (el planteado por el estadista) se puede refutar fácilmente porque aunque todos los coeficientes son significativos individual y conjuntamente, el R² es menor que el que usted había calculado al igual que el R² ajustado. Con respecto al modelo propuesto por el Matemático se debe utilizar el mismo argumento y para el del presidente se debe refutar a través de la idea de que aunque los coeficientes son significativos, la eliminación del intercepto haría asumir que los robos solo depende de las variables que allí se explican lo cual no es lógico, además de

¹ Utilice F's Globales, t calculados, R² para todos los análisis.

que el nivel de ingreso per individuo es una variable relevante a la hora de explicar actos ilegales tales como el robo, por tanto el modelo especificado no es óptimo. Además ustedes deben explicar que el R^2 para este modelo no es interpretable porque no hay intercepto y al mismo tiempo el F global no es interpretable para este modelo por no existir intercepto.