

Taller #4
Regresión Simple
Econometría 06169

Profesores: Julio César Alonso
Diego Yépez

Notas:

- Recuerde que sólo dos preguntas, seleccionadas al azar, serán calificadas.
- Este taller es para ser entregado en los primeros 10 minutos de la clase del próximo 31 de agosto.

INSTRUCCIONES:

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

1. Dado el modelo

$$\ln(y) = \beta_1 + \beta_2 \ln(X_i) + \varepsilon_i \quad (1.1)$$

demuestre que el coeficiente asociado con la variable independiente corresponde a una elasticidad.

Un investigador desea determinar la relación entre el salario (w_i , medido como el ingreso por hora promedio en dólares) y el nivel educativo (edu_i , medido en años de educación). Para esto se cuenta con una muestra de 526 individuos. El investigador cree que el modelo adecuado para describir esta relación es:

$$\ln(w_i) = \beta_1 + \beta_2 edu_i + \varepsilon_i \quad (1.2)$$

Empleando la información disponible en el archivo T3-02-04.xls, y usando Excel responda las siguientes preguntas.

2. Responda:

- a. ¿qué tipo de datos estamos empleando?
- b. Calcule los coeficientes del modelo por medio del método MCO. (muestre las fórmulas y cálculos realizados).
- c. Interprete el significado de los coeficientes estimados

3. Continuando con la pregunta anterior, responda:

- a. ¿Es el coeficiente asociado con la educación significativo?
- b. Grafique cada una de la observaciones y la respectiva línea de regresión estimada (emplee Excel)
- c. ¿Qué tan bueno es el modelo estimado? Explique.

Taller #4
Regresión Simple
Respuestas Sugeridas
Econometría 06169

Profesores: Julio César Alonso
Diego Yépez

Notas:

- o Recuerde que sólo dos preguntas, seleccionadas al azar, serán calificadas.
- o Este taller es para ser entregado en los primeros 10 minutos de la clase del próximo 31 de agosto.

INSTRUCCIONES:

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

1. Dado el modelo

$$\ln(y) = \beta_1 + \beta_2 \ln(X_i) + \varepsilon_i \quad (0.1)$$

demuestre que el coeficiente asociado con la variable independiente corresponde a una elasticidad.

Para demostrar que el coeficiente asociado con la variable independiente corresponde a una elasticidad será necesario varias manipulaciones algebraicas. Noten que (0.1) se puede escribir como:

$$y_i = \gamma X_i^{\beta_2} \mu$$

Donde, $\beta_1 = \ln(\gamma)$ y $\varepsilon_i = \ln(\mu_i)$. Ahora, podemos derivar con respecto a X_i . Es decir,

$$\frac{\partial y_i}{\partial X_i} = \frac{\partial}{\partial X_i} [\gamma X_i^{\beta_2} \mu] = \beta_2 \gamma X_i^{\beta_2 - 1} \mu$$

$$\frac{\partial y_i}{\partial X_i} = \beta_2 X_i^{-1} [\gamma X_i^{\beta_2} \mu] = \beta_2 \frac{Y_i}{X_i}$$

Ahora, manipulando un poco esta expresión, tenemos:

$$\beta_2 = \frac{\partial y_i}{\partial X_i} \frac{X_i}{Y_i} = \frac{\partial y_i}{\partial X_i} \frac{X_i}{Y_i} \cdot \frac{100}{100} = \frac{\Delta\% y_i}{\Delta\% X_i}$$

Q.E.D.

Un investigador desea determinar la relación entre el salario (w_i , medido como el ingreso por hora promedio en dólares) y el nivel educativo (edu_i , medido en años de

educación). Para esto se cuenta con una muestra de 526 individuos. El investigador cree que el modelo adecuado para describir esta relación es:

$$\ln(w_i) = \beta_1 + \beta_2 edu_i + \varepsilon_i \quad (0.2)$$

Empleando la información disponible en el archivo T3-02-04.xls, y usando Excel responda las siguientes preguntas.

2. Responda:
- a. ¿qué tipo de datos estamos empleando?

Estos son datos de corte transversal

- b. Calcule los coeficientes del modelo por medio del método MCO. (muestre las fórmulas y cálculos realizados).

Definamos $y_i = \ln(w_i)$ y $X_i = edu_i$, entonces tenemos que:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = 0.082806$$

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x} = 0.58377$$

- c. Interprete el significado de los coeficientes estimados

Noten que el modelo (0.2) se puede reescribir como:

$$w_i = e^{\beta_1 + \beta_2 edu_i + \varepsilon_i}$$

Entonces, derivando obtenemos:

$$\frac{\partial w_i}{\partial edu_i} = \frac{\partial}{\partial edu_i} [e^{\beta_1 + \beta_2 edu_i + \varepsilon_i}] = \beta_2 e^{\beta_1 + \beta_2 edu_i + \varepsilon_i} = \beta_2 w_i$$

Manipulando algebraicamente esta expresión, obtenemos:

$$\beta_2 = \frac{\partial w_i}{\partial edu_i} \frac{1}{w_i} = \frac{\partial w_i / w_i}{\partial edu_i}$$

$$\beta_2 \cdot 100 = \frac{\partial w_i / w_i}{\partial edu_i} \cdot 100 = \frac{\Delta\% w_i}{\Delta\% edu_i}$$

Es decir, que $0.082806 \cdot 100 = 8.28$ corresponde al cambio porcentual del salario por un cambio de una unidad en la educación. En otras palabras, cuando la educación aumenta en un año, el salario aumentará en un 8.28%.

Por otro lado, el intercepto en este caso carece de interpretación económica pues cuando la educación es cero entonces tendremos que:

$$\ln(w_i) = \beta_1$$

$$w_i = e^{\beta_1}$$

Por tanto no tiene mucho sentido en este caso el intercepto.

3. Continuando con la pregunta anterior, responda:

a. ¿Es el coeficiente asociado con la educación significativo?

Noten que el estadístico t para probar la hipótesis nula que $\beta_2 = 0$ versus la alterna que $\beta_2 \neq 0$ corresponde a:

$$t_c = \frac{\hat{\beta}_2}{s_{\hat{\beta}_2}}$$

donde

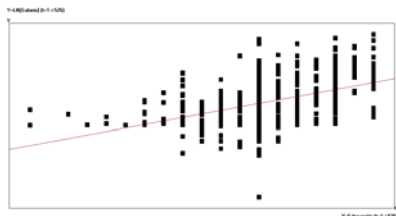
$$s_{\hat{\beta}_2} = \sqrt{\frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2} S^2}$$

En este caso ese t calculado corresponde a 10.935 que es muy superior al estadístico t con 524 grados de libertad y un nivel de significancia del 99%. Por tanto se puede rechazar la hipótesis nula de que el coeficiente es cero. Por tanto la educación si afecta los ingresos.

b. Grafique cada una de la observaciones y la respectiva línea de regresión estimada (emplee Excel)

A continuación se presenta el gráfico de cada una de las observaciones y la respectiva línea de regresión estimada.

Gráfica 1. Observaciones y la línea de regresión estimada



c. ¿Qué tan bueno es el modelo estimado? Explique.

En este caso, el coeficiente de determinación está dado por:

$$r^2 = \frac{SSR}{SST} = 0.184253$$

Es decir el 18.4253 % de la variabilidad del **logaritmo del salario** es explicado por el modelo. Por tanto el modelo no es muy bueno. Aunque hay que anotar que típicamente en datos de corte transversal el coeficiente de determinación normalmente es muy bajo.