

Métodos Estadísticos para Economía y Gestión (IN540-2)
Primavera 2008 - Semestre II

Universidad de Chile
Departamento de Ingeniería Industrial
Profesor: Mattia Makovec (mmakovec@dii.uchile.cl)

El archivo *salarios_tarea1.xls* contiene observaciones sobre los salarios mensuales en US\$ (*Wage*) de 935 hombres empleados en Estados Unidos. Además, el archivo incluye las siguientes variables:

Educ	años de educación
Exper	años de experiencia laboral
Tenure	años de experiencia con el actual empleador
Age	edad
Married	dummy =1 si casado
Black	dummy =1 si de etnia afro-americana
South	dummy =1 si vive en un estado del sur de EE.UU.
Urban	dummy =1 si vive en una area metropolitana
Meduc	años de educación de la madre
Feduc	años de educación del padre
sibs	numero de hermanos/hermanas
Brthord	orden de nacimiento
IQ	cociente intelectual

Preguntas

1. Estime con el método de mínimos cuadrados ordinarios la siguiente ecuación de salario:

$$\log(Wage)_i = \beta_0 + \beta_1 Educ_i + \beta_2 Exper_i + \beta_3 Exper_i^2 + \beta_4 Tenure_i + \beta_5 Age_i + \varepsilon_i \quad (1)$$

$i = 1, \dots, 935$, donde se supone que se cumplen todos los supuestos básicos del modelo lineal general incluyendo la hipótesis de normalidad de los errores.

- (a) Comente los resultados obtenidos y la significatividad individual de los coeficientes estimados. ¿Son los coeficientes estimados significativos al 1%? ¿Y al 5%? ¿Y al 10%? ¿Cual es el valor estimado de la varianza de los errores ε_i en la regresión (1)?
- (b) Defina y realice un test de significatividad global de la regresión, especificando claramente la hipótesis nula, la hipótesis alternativa y la regla de decisión por un nivel de significación $\alpha = 0.05$.
- (c) Realizando la regresión auxiliar apropiada, demuestre que el estimador de mínimos cuadrados ordinarios de β_1 en la ecuación (1), $\hat{\beta}_1$, es numericamente equivalente al estimador de mínimos cuadrados ordinarios de δ_1 en la regresión lineal de $\log(Wage)_i$ sobre \hat{r}_{i1} :

$$\log(Wage)_i = \delta_0 + \delta_1 \hat{r}_{i1} + \eta_i$$

siendo \hat{r}_{i1} los residuos estimados obtenidos de la regresión lineal de $Educ_i$ sobre las demás variables explicativas incluidas en la ecuación (1).

2. Estime con el método de mínimos cuadrados ordinarios la siguiente especificación alternativa de la ecuación (1):

$$\log(Wage)_i = \alpha_0 + \alpha_1 Educ_i + \alpha_2 Exper_i + \alpha_3 Tenure_i + \alpha_4 Age_i + \varepsilon_i \quad (2)$$

- (a) Comente los resultados obtenidos con respecto al apartado anterior. Defina y realice el test estadístico apropiado para testear la hipótesis que el efecto de $Exper$ sobre el logaritmo del salario sea lo mismo que lo de $Tenure$, indicando claramente la hipótesis nula, la hipótesis alternativa y la regla de decisión por un nivel de significación $\alpha = 0.05$.
- (b) ¿Cual especificación se debería preferir entre la ecuación (1) y la ecuación (2)? Justifique la respuesta utilizando **todos** los tests apropiados en este contexto analizados en clase (utilice como nivel de significación $\alpha = 0.05$).
3. Ampliar el modelo especificado en la ecuación (2) para incluir los efectos sobre el salario de: 1) estado civil; 2) raza; 3) estado de residencia; 4) residencia en una area metropolitana.
- (a) Realice la regresión apropiada para el estimar el modelo ampliado y comente los resultados.
- (b) Demuestre cual es la mejor especificación entre la del apartado anterior y la de la ecuación (2) utilizando **todos** los tests apropiados en este contexto analizados en clase (utilice como nivel de significación $\alpha = 0.05$).
- (c) Utilizando la especificación del apartado anterior, calcule (cada vez manteniendo fijas las demás características) la diferencia percentual aproximada en el salario estimado entre: 1) hombres de etnia afro-americana que viven en areas metropolitanas y hombres de etnia afro-americana que viven en otras areas; 2) hombres de etnia *no* afro-americana que viven en areas metropolitanas y hombres de etnia *no* afro-americana que viven en otras areas; 3) hombres de etnia afro-americana que viven en areas metropolitanas y hombres de etnia *no* afro-americana que viven en areas metropolitanas; 4) hombres de etnia afro-americana que viven en areas *no* metropolitanas y hombres de etnia *no* afro-americana que viven en areas *no* metropolitanas. Comente los resultados obtenidos.

4. Explique por qué es apropiado añadir a los regresores de la especificación del apartado anterior la variable IQ . Cuales son las potenciales consecuencias de omitir IQ sobre el estimador del coeficiente de la variable $Educ$?

5. Estime la siguiente especificación de la ecuación de salario:

$$\begin{aligned} \log(Wage)_i = & \delta_0 + \delta_1 Educ_i + \delta_2 Exper_i + \delta_3 Tenure_i + \delta_4 Age_i + \delta_5 Black + \\ & + \delta_6 Urban + \delta_7 Meduc + \delta_8 Feduc + \delta_9 sibs + \delta_{10} Brthord + \delta_{11} IQ + \varepsilon_i \end{aligned} \quad (3)$$

Comente los resultados obtenidos. Cual es ahora la variación percentual *aproximada* de un año adicional de educación sobre el salario? Explique por qué es diferente del resultado obtenido en el primer apartado. ¿Son las nuevas variables *Meduc Feduc sibs* y *Brthord* individualmente significativas? Son conjuntamente significativas? Defina claramente la hipótesis nula, la hipótesis alternativa y la regla de decisión por un nivel de significación $\alpha = 0.05$.

6. Indique al menos dos tipos de variables no incluidas en la especificación (3) y no disponibles en el archivo de datos que podrían ser añadidas a los regresores utilizados en esta ecuación para mejorar la estimación de $\log(Wage)$.