Hoja de Ejercicios 10
Especificación de Modelos

Estadística-II. INTRODUCCIÓN a la ECONOMETRÍA. UC3M

1. (Ejercicio 6.3, Wooldridge (2006)). Usando los datos de RDCHEM.RAW, se ha obtenido la siguiente ecuación por MCO:

$$\begin{aligned} \widehat{rdintens} &= 2{,}613 + 0{,}00030\; sales - 0{,}0000000070\; sales^2 \\ & \quad (0{,}429)\ (0{,}00014) \qquad\qquad (0{,}0000000037) \\ n &= 32,\ R^2 = 0{,}1484 \end{aligned}$$

   (i) ¿A partir de qué punto el efecto marginal de *sales* sobre *rdintens* comienza a ser negativo?

   (ii) ¿Conviene mantener el término cuadrático en el modelo? ¿Porqué?

   (iii) Definamos *salesbil* como las ventas medidas en billones de dólares: $salesbil = sales/1000$. Volver a escribir la ecuación estimada con *salesbil* y $salesbil^2$ como variables independientes. Asegurarse de presentar los errores estándar y el R-cuadrado. [Pista: Nótese que $salesbil^2 = sales^2/(1000)^2$]

   (iv) A la hora de presentar los resultados, ¿qué ecuación es preferible?

2. (Ejercicio 6.4, Wooldridge (2006)). El siguiente modelo hace depender el rendimiento de la educación total que tienen ambos padres, denominado *pareduc*:

$$\log(wage) = \beta_0 + \beta_1\; educ + \beta_2\; educ \cdot pareduc + \beta_3\; exper + \beta_4\; tenure\ + \mu$$

   (i) Demostrar que, en forma decimal, el rendimiento de un año más de educación en este modelo es:

$$\Delta \log(wage)/\Delta educ = \beta_1 + \beta_2\; pareduc$$

   ¿Qué signo espera para β_2? ¿Porqué?

   (ii) Usando los datos de WAGE2.RAW, la ecuación estimada es:

$$\begin{aligned} \log(\widehat{wage}) &= 5{,}65 + 0{,}047\; educ + 0{,}00078\; educ \cdot pareduc + 0{,}019\; exper + 0{,}010\; tenure \\ & \quad (0{,}13)\ (0{,}010) \qquad (0{,}00021) \qquad\qquad (0{,}004) \qquad (0{,}003) \\ n &= 722,\ R^2 = 0{,}169 \end{aligned}$$

   (Solamente 722 observaciones contienen la información completa sobre la educación de los padres). Interpretar los coeficientes del término de interacción. Puede ser de ayuda elegir dos valores específicos para *pareduc* (por ejemplo, $pareduc = 32$ si ambos padres tienen formación universitaria, o $pareduc = 24$ si ambos padres tienen una educación secundaria) y comparar la estimación del rendimiento de *educ*.

(iii) Cuando añadimos *pareduc* a la ecuación como variable separada, obtenemos:

$$\begin{aligned}
\log(\widehat{wage}) &= \underset{(0,38)}{4,94} + \underset{(0,027)}{0,097}\, educ + \underset{(0,017)}{0,033}\, pareduc - \underset{(0,0012)}{0,0016}\, educ \cdot pareduc + \underset{(0,004)}{0,020}\, exper \\
&\quad + \underset{(0,003)}{0,010}\, tenure \\
n &= 722,\ R^2 = 0,174
\end{aligned}$$

¿El rendimiento de la educación depende ahora positivamente de la educación de los padres? Contrastar la hipótesis nula de que el rendimiento de la educación no depende de la educación de los padres.

3. (Ejercicio 6.10, Wooldridge (2006)). Consideremos un modelo en el que el rendimiento de la educación depende de la experiencia laboral (y viceversa):

$$\log(wage) = \beta_0 + \beta_1\, educ + \beta_2\, exper + \beta_3\, educ \cdot exper \ + \mu$$

(i) Demostrar que el efecto de un año más de educación (en forma decimal), manteniendo *exper* constante, es: $\beta_1 + \beta_3\, exper$.

(ii) Especificar la hipótesis nula de que el rendimiento de la educación no depende del nivel de *exper* . ¿Cuál es la hipótesis alternativa adecuada?

(iii) Usar los datos de WAGE2.RAW para contrastar la hipótesis nula de (ii) contra la alternativa propuesta.

(iv) Sea θ_1 el rendimiento de la educación (en forma decimal), cuando $exper = 10$: $\theta_1 = \beta_1 + 10\beta_3$. Obtener $\widehat{\theta}_1$ y un intervalo de confianza al 95 % para θ_1. (Pista: Escribir $\theta_1 = \beta_1 - 10\beta_3$, introducir esto en la ecuación y reordenar. Esto proporciona la regresión de la que se puede obtener el intervalo de confianza para θ_1).