EXAMEN DE ECONOMETRÍA I

SOLUCIONES SEPTIEMBRE

CURSO 2000/2001

PROBLEMA 1

En un estudio sobre determinación de los salarios de los hombres se ha obtenido, empleando una muestra de 1867 individuos:

$$\begin{aligned} \log W_i &= \underset{(0.034)}{5.83} + \underset{(0.002)}{0.06}\, S_i + \underset{(0.001)}{0.02}\, X_i + \hat{U}_i \qquad (1) \\ R^2 &= 0.3325 \qquad \widehat{\mathbf{U}}'\widehat{\mathbf{U}} = 365.461, \end{aligned}$$

donde W es el salario-hora, S es el número de años de estudio, X son los años de experiencia laboral y los valores entre paréntesis son los errores estándar.

1. **Contraste al 5% si el efecto de los años de educación sobre los salarios es nulo. (1/3)**

   Queremos contrastar:

   $$H_0 : \beta_2 = 0 \text{ versus } H_1 : \beta_2 \neq 0.$$

   Estadístico del contraste:

   $$t = \frac{\widehat{\beta_2}}{\sqrt{\widehat{var}(\widehat{\beta_2})}} = \frac{0.06}{0.002} = 30.$$

   Distribución del estadístico bajo H_0 :

   $$t \underset{aprox.}{\sim} N(0,1).$$

   Criterio del contraste al 5% de significación: Rechazar H_0 si $t > 1.96$ ($\Pr(N(0,1) > 1.96) = 0.025$). Por tanto, rechazamos H_0.

2. **También se ha estimado el siguiente modelo:**

   $$\begin{aligned} \log W_i &= \underset{(0.035)}{5.93} + \underset{(0.0008)}{0.0022}(2S_i + X_i) + \hat{V}_i, \\ \tilde{R}^2 &= 0.2742 \qquad \widehat{\mathbf{V}}'\widehat{\mathbf{V}} = 397.4249, \end{aligned}$$

   donde se ha supuesto que el efecto sobre los salarios de un año adicional de educación es el doble que el de un año adicional de experiencia laboral. Contraste al 5% dicha hipótesis. (1/3)

   Queremos contrastar:

   $$H_0 : \beta_2 = 2\beta_3 \text{ versus } H_1 : \beta_2 \neq 2\beta_3.$$

   Estadístico del contraste con la información disponible:

   $$\begin{aligned} F &= \frac{(\widehat{\mathbf{V}}'\widehat{\mathbf{V}} - \widehat{\mathbf{U}}'\widehat{\mathbf{U}})/q}{\widehat{\mathbf{U}}'\widehat{\mathbf{U}}/(n-k)} = \frac{(397.4249 - 365.461)}{365.461/1864} \approx 163 \\ &= \frac{(R^2 - \tilde{R}^2)/q}{(1-R^2)/(n-k)} = \frac{(0.3325 - 0.2742)}{(1-0.3325)/1864} \approx 163 \end{aligned}$$

   Distribución del estadístico bajo H_0 :

   $$F \underset{aprox.}{\sim} \chi^2_1.$$

   Criterio del contraste al 5% de significación: Rechazar H_0 si $F > 3.85$ ($\Pr(\chi^2_1 > 3.84) = 0.05$). Por tanto, rechazamos la hipótesis nula.

3. **La teoría del capital humano sugiere que el efecto de la experiencia no es lineal y que, por ello, debe incluirse como regresor el cuadrado de la experiencia.**

   El modelo de interés es,

   $$\log W_i = \gamma_1 + \gamma_2 S_i + \gamma_3 X_i + \gamma_4 X_i^2 + E_i.$$

(a) **De ser cierta esta teoría, ¿qué propiedades tendrían las estimaciones del modelo (1)? (1/6)**

Los estimadores MCO del modelo (1) pueden ser inconsistentes cuando $\gamma_4 \neq 0$.

(b) **Contraste al 5% si hay evidencia en contra de esta teoría sabiendo que:**

$$\begin{aligned} \log W_i &= 5.62 + 0.06 S_i + 0.05 X_i - 0.0007 X_i^2 + \hat{E}_i, \\ \bar{R}^2 &= 0.3705 \qquad \widehat{\mathbf{E}}'\widehat{\mathbf{E}} = 344.6945. \end{aligned}$$

(1/6)

Queremos contrastar:

$$H_0 : \gamma_4 = 0 \text{ versus } H_1 : \gamma_4 \neq 0.$$

Estadístico del contraste con la información disponible:

$$\begin{aligned} F &= \frac{(\widehat{\mathbf{U}}'\widehat{\mathbf{U}} - \widehat{\mathbf{E}}'\widehat{\mathbf{E}})/q}{\widehat{\mathbf{E}}'\widehat{\mathbf{E}}/(n-k)} = \frac{(365.461 - 344.6945)}{344.6945/1863} \approx 112 \\ &= \frac{(\bar{R}^2 - R^2)/q}{(1-\bar{R}^2)/(n-k)} = \frac{(0.3705 - 0.3325)}{(1-0.3705)/1863} \approx 112 \end{aligned}$$

Distribución del estadístico bajo H_0 :

$$F \underset{aprox.}{\sim} \chi_1^2.$$

Criterio del contraste al 5% de significación: Rechazar H_0 si $F > 3.85$ ($\Pr(\chi_1^2 > 3.84) = 0.05$). Por tanto, rechazamos H_0.

PROBLEMA 2

La siguiente ecuación relaciona los salarios con el nivel de educación:

$$\ln(\widehat{Salarios}) = \underset{(0,185)}{-0,185} + \underset{(0,014)}{0,11}\,(Educación) \quad R^2 = 0,118 \quad n = 428, \tag{2}$$

donde los coeficientes han sido estimados por MCO utilizando una muestra aleatoria de tamaño $n = 428$, los números entre paréntesis son los errores estándar.

En economía laboral el modelo de regresión (**??**) se considera mal especificado, por no tener en cuenta otras variables significativas en la determinación del salario, entre las que destaca la inteligencia (CI). Sería más razonable estimar por MCO,

$$\ln(Salarios) = \beta_1 + \beta_2(Educación) + \beta_3(CI) + U, \tag{3}$$

suponiendo que $E(U) = 0$, y que las variables explicativas son exógenas.

1. **Si $\beta_3 \neq 0$ en el modelo (3), ¿podemos realizar inferencias válidas a partir del modelo estimado (??)? Justifique su respuesta. (1/4)**

   El estimador MCO que afecta a la variable $Educación$ en (**??**) será en general inconsistente. Es consistente cuando $Educación$ y CI están incorrelacionados.

2. **Si la variable "$Educación$" esta correlacionada de forma positiva con la variable "CI" ¿cómo "esperaría" que fuera el parámetro β_2 en el modelo (3) en relación con el valor estimado $(0,11)$ en la ecuación (??)? (1/4)**

   El sesgo asintótico del estimador MCO que afecta a la variable $Educación$ en (**??**) es positivo, por lo que podemos esperar que el verdadero parámetro de la regresión sea inferior de lo que sugiere la estimación del modelo sin incluir la variable explicativa CI.

3. **En una nota a pié de página del artículo mencionado, figura la siguiente ecuación estimada por MCO:**

$$\widehat{Educación} = \underset{(0,28)}{10,24} + \underset{(0,029)}{0,269}(EducaciónMadre) \quad R^2 = 0,173 \quad n = 428 \tag{4}$$

¿Qué condiciones ha de cumplir la variable "$EducaciónMadre$" para ser un buen instrumento de la variable "$Educación$" en el modelo (??)? A la vista de la ecuación estimada (4), ¿podemos inferir que se cumple alguna de estas condiciones? Justifique sus respuestas. (1/4)

La variable instrumental $EducacionMadre$ no ha de estar correlacionada con el término de error, pero ha de estarlo con la variable $Eduacación$. Un contraste de significación, basado en el estadístico t, sugiere que la variable instrumental $EducacionMadre$ está correlacionada con la variable $Educacion$.

4. **En otra nota a pié de página aparece la estimación del modelo (??) por variables instrumentales, utilizando la variable "$EducaciónMadre$" como instrumento de la variable "$Educación$":**

$$\ln(\widehat{Salarios}) = \underset{(0{,}446)}{0,441} + \underset{(0{,}035)}{0,055}(Educación) \quad R^2 = 0,093 \quad n = 428. \tag{5}$$

Compare los resultados obenidos en (5) con los obtenidos en (??), utilizando la inferencia estadística como base de su análisis. (1/4)

El valor estimado del parámetro en (5) se ha reducido a la mitad con respecto al obtenido en (??), y el contraste de significación sugiere que no es significativo. Por tanto, no podemos asegurar que el nivel de educación afecta a los salarios basándonos en la evidencia empírica disponible.

PROBLEMA 3

Se ha planteado el siguiente modelo para explicar las ventas trimestrales y_t de una empresa:

$$Y_t = \alpha_1 + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \alpha_4 D_{4t} + U_t \quad t = 1, ..., T \tag{6}$$

donde $E(U_t) = 0$, $E(U_t^2) = \sigma_U^2\ \forall t$, $E(U_t U_{t-s}) = 0$, $\forall s \neq 0$.

$$D_{jt} = \begin{cases} 1 & \text{si } t \text{ pertenece al trimestre } j\text{-ésimo} \\ 0 & \text{en otro caso} \end{cases} \qquad j = 1, 2, 3, 4.$$

1. **Interprete los coeficientes del modelo (6). (1/5)**

   $\alpha_1 = E(D_{1t} Y_t)$ y $\alpha_1 + \alpha_i = E(D_{it} Y_t)$, $i = 2, ..., 4$. Esto es, α_1 es la media de las ventas en el primer trimestre y α_i, $i = 2, ..., 4$ es la diferencia entre media de las ventas en el correspondiente trimestre y la media de ventas en el primer trimestre.

2. **Si se considera el modelo alternativo:**

   $$Y_t = \beta_0 + \beta_1 D_{1t} + \beta_2 D_{2t} + \beta_3 D_{3t} + \beta_4 D_{4t} + E_t \quad t = 1, ..., T$$

   donde $E(E_t) = 0$, $E(E_t^2) = \sigma_E^2\ \forall t$, $E(E_t E_{t-s}) = 0$, $\forall s \neq 0$. ¿Existe algún problema para estimar dicho modelo? Justifique su respuesta. (1/5)

   Existe colinealidad perfecta, $\sum_{i=1}^4 D_{it} = 1$ para todo t. Por tanto, los coeficientes del modelo no están identificados.

3. **Dado el siguiente modelo:**

   $$Y_t = \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \gamma_4 D_{4t} + V_t \quad t = 1, ..., T; \tag{7}$$

   donde $E(V_t) = 0$, $E(V_t^2) = \sigma_V^2\ \forall t$, $E(V_t V_{t-s}) = 0$, $\forall s \neq 0$. Comente qué relación existe entre los coeficientes del modelo (6) y los del modelo (7). (1/5)

   $\gamma_1 = \alpha_1$ y $\gamma_i = \alpha_1 + \alpha_i$, $i = 2, ..., 4$.

4. **Estime por MCO los parámetros del modelo (7), sabiendo que se dispone de una muestra de 40 datos trimestrales (10 años), y dados los siguientes estadísticos calculados a partir de la muestra:**

   $$\sum_{t=1}^{40} D_{1t} = \sum_{t=1}^{40} D_{2t} = \sum_{t=1}^{40} D_{3t} = \sum_{t=1}^{40} D_{4t} = 10;$$
   $$\sum_{t=1}^{40} Y_t D_{1t} = 20; \quad \sum_{t=1}^{40} Y_t D_{2t} = 30; \quad \sum_{t=1}^{40} Y_t D_{3t} = 70; \quad \sum_{t=1}^{40} Y_t D_{4t} = 30.$$

   (1/5)

   Los valores estimados de γ_i, $i = 1, ..., 4$ son: $\hat{\gamma}_i = \frac{\sum_{t=1}^{40} D_{it} Y_t}{\sum_{t=1}^{40} D_{it}} = \begin{cases} 20/10 = 2 \text{ si } i = 1 \\ 30/10 = 3 \text{ si } i = 2 \\ 70/10 = 7 \text{ si } i = 3 \\ 30/10 = 3 \text{ si } i = 4. \end{cases}$

5. **Sabiendo que para la estimación anterior $\sum_{t=1}^{40} \hat{V}_t^2 = 720$, contraste al 5% que la ventas durante el primer trimestre son iguales en media a las del segundo cuatrimestre. (1/5)**

Notesé que,

$$\hat{\gamma} = \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \\ \hat{\gamma}_3 \\ \hat{\gamma}_4 \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix} + \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 10 \end{bmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^{40} D_{1t}U_t \\ \sum_{t=1}^{40} D_{2t}U_t \\ \sum_{t=1}^{40} D_{3t}U_t \\ \sum_{t=1}^{40} D_{2t}U_t \end{pmatrix}$$

$$\hat{\gamma} \underset{aprox}{\sim} N\left(\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix}, \sigma^2 \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 10 \end{bmatrix}^{-1} \right)$$

Por tanto, para contrastar,

$$H_0 : \gamma_1 = \gamma_2 \text{ versus } H_1 : \gamma_1 \neq \gamma_2,$$

podemos utilizar el estadísitco,

$$t = \frac{|\widehat{\gamma}_1 - \widehat{\gamma}_2|}{sd\left(\widehat{\gamma}_1 - \widehat{\gamma}_2\right)} \underset{aprox.}{\overset{H_0}{\sim}} N(0,1)$$

con

$$\begin{aligned} sd\left(\widehat{\gamma}_1 - \widehat{\gamma}_2\right) &= \sqrt{Var(\widehat{\gamma}_1) + Var(\widehat{\gamma}_2) - 2Cov(\widehat{\gamma}_1, \widehat{\gamma}_2)} \\ &= \sqrt{s^2\left(\frac{1}{\sum d_{i1}} + \frac{1}{\sum d_{i2}}\right)} = \\ &= \sqrt{\frac{720}{40-4}\left(\frac{1}{10} + \frac{1}{10}\right)} = \sqrt{4} = 2 \end{aligned}$$

Por tanto

$$t = \frac{|\widehat{\gamma}_1 - \widehat{\gamma}_2|}{sq\left(\widehat{\gamma}_1 - \widehat{\gamma}_2\right)} = \frac{|2-3|}{2} = \frac{1}{2} = 0.5$$

Bajo H_0: $t \underset{aprox}{\sim} N\left(0,1\right)$. No podemos rechazar H_0 a ningún nivel de significación razonable.