Examen de Econometría I
Universidad Carlos III de Madrid
2aa Convocatoria (1 de Septiembre de 2004)
Curso 2003/2004
Conteste las preguntas siguientes en 2:30 horas
Valores críticos de diferentes distribuciones pueden encontrarse al final del examen.

PREGUNTA 1. Considere la siguiente función de producción Cobb-Douglas:

$$Y = A \cdot K^{\alpha} \cdot L^{\beta} \cdot \exp(U),$$

donde A, α y β son parámetros desconocidos, e Y, K, L y U son variables aleatorias. Dada una población de empresas y tomada una empresa al azar, Y representa la cantidad producida, K la cantidad utilizada de capital, L la cantidad utilizada de trabajo y U es un término de error tal que $E(U|K = k, L = l) = 0$ para todo posible valor de k y l. Supongamos que tenemos una muestra aleatoria (Y_i, K_i, L_i), $i = 1, ..., n$ del vector de variables (Y, K, L), donde n es un valor muy grande, por ej. $n = 5.000$ empresas.

(a) (0,5 puntos) Explique cómo estimaría los parámetros del modelo por mínimos cuadrados ordinarios.

(b) (1,5 puntos) Explique cómo contrastaría al 5% de significación la existencia de rendimientos constantes a escala, utilizando el estadístico de la t. Detalle la hipótesis a contrastar, proponga el estadístico del contraste y la regla de decisión.

SOLUCIÓN:

(a) Consideramos la ecuación en logaritmos

$$\ln Y = \ln A + \alpha \ln K + \beta \ln L + U,$$

y estimamos el modelo por mínimos cuadrados ordinarios.

(b) La hipótesis a contrastar es

$$H_0 : \alpha + \beta = 1$$

en la dirección de la alternativa,

$$H_A : \alpha + \beta \neq 1.$$

Utilizamos el estadístico

$$t = \frac{\hat{\alpha} + \hat{\beta} - 1}{\sqrt{\widehat{Var}(\hat{\alpha}) + \widehat{Var}\left(\hat{\beta}\right) + 2 \cdot \widehat{Cov}\left(\hat{\alpha}, \hat{\beta}\right)}},$$

donde $\hat{\alpha}$ y $\hat{\beta}$ son los estimadores de mínimos cuadrados ordinarios de los parámetros α y β, $\widehat{Var}(\hat{\alpha})$, $\widehat{Var}\left(\hat{\beta}\right)$ y $\widehat{Cov}\left(\hat{\alpha}, \hat{\beta}\right)$ son estimadores de la varianza de $\hat{\alpha}$, de la varianza de $\hat{\beta}$ y la covarianza entre $\hat{\alpha}$ y $\hat{\beta}$, respectivamente. La regla de decisión consiste en rechazar la hipótesis nula cuando el valor del estadístico en valor absoluto es mayor que $Z_{0,05} = 1,96$, donde $\Pr(N(0,1) > Z_{0,05}) = 0,975$.

PREGUNTA 2. Una base de datos proporciona la siguiente información: precio medio en miles de euros de los apartamentos en cada barrio de Madrid (*precio*), número medio de habitaciones por apartamentos en cada barrio (*dorm*), y número de atracos (robos en la calle) en los últimos dos años de cada barrio (*robo*). Se considera el modelo

$$\ln(precio) \;=\; \beta_0 + \beta_1 \ln(dorm) + \beta_2 robo + U,$$

donde $E(U|dorm, robo) = 0$, y $\beta_2 < 0$.

(a) (1 punto) Interprete los coeficientes β_1 y β_2.

(b) (1 punto) Considere el modelo en que se excluye la variable *robo*,

$$\ln(precio) \;=\; \delta_0 + \delta_1 \ln(dorm) + \varepsilon,$$

donde $E(\varepsilon) = 0$. Proporcione una fórmula que relacione δ_1 con β_1.

(c) (1 punto) Suponiendo que tenemos una muestra aleatoria muy grande, 2.000 observaciones, considere el modelo estimado utilizando mínimos cuadrados ordinarios,

$$\ln(\widehat{precio})_i \;=\; \hat{\delta}_0 + \hat{\delta}_1 \ln(dorm)_i, \;\; i = 1, ..., 2.000.$$

¿Es $\hat{\delta}_1$ un estimador consistente de δ_1? Explique en qué circunstancia $\hat{\delta}_1$ tiene un sesgo (asintótico) negativo al estimar β_1.

SOLUCIÓN:

(a) β_1 es la elasticidad esperada del *precio* respecto a *dorm* : se espera que un aumento de un uno por cien en el número de habitaciones en un piso implique un aumento, ceteris paribus, del precio de un β_1 por ciento. Por cada robo adicional en un barrio se espera que, ceteris paribus, el precio de los pisos disminuya en $100 \cdot \beta_2$ %.

(b)

$$\delta_1 = \beta_1 + \beta_2 \frac{Cov[\ln(dorm), robo]}{Var[\ln(dorm)]}$$

(c) $\hat{\delta}_1$ es un estimador consistente de δ_1, pero no de β_1. Teniendo en cuenta que $\beta_2 < 0$, el sesgo (asintótico) de $\hat{\delta}_1$ como estimador de β_1 es positivo cuando $Cov[\ln(dorm), robo] < 0$.

PREGUNTA 3: Considere el modelo

$$E(Salario\,|Sexo,\; Educación) = 9,5 - 4,1\; Sexo + 0,8\; Educación,$$

donde el *Salario* (anual) está expresado en miles de euros, la *Educación* está medida en años y el *Sexo* toma el valor 0 para hombres y 1 para mujeres.

(a) (0,5 puntos) ¿Cuál es el salario esperado de un varón con 10 años de *Educación*?

(b) (0,5 puntos) ¿Cuál es la diferencia esperada en el salario de un hombre y de una mujer con igual nivel de educación?

(c) (0,5 puntos) Suponga que

$$PLO\,(Educación\,|Sexo) = 8,0 + 3,0\; Sexo.$$

¿Cuál sería la pendiente del $PLO(Salario\,|Sexo)$?

(d) (0,5 puntos) ¿Qué medida es preferible para medir discriminación por sexo: el coeficiente del *Sexo* en $E(Salario\,|Sexo,\; Educación)$ o la pendiente del $PLO(Salario\,|Sexo)$?. Justifique la respuesta.

SOLUCIÓN:

(a) $E(Salario\,|Sexo,\; Educación) = 9,5 - 4,1 \times 0 + 0,8 \times 10 = 17,5$ miles de euros = 17.500 euros.

(b) $E(Salario\,|Sexo = 0,\; Educación) - E(Salario\,|Sexo = 1,\; Educación) = 4,1$ miles de euros = 4.100 euros.

(c) $PLO(Salario|Sexo) = \gamma_0 + \gamma_1 Sexo$, donde

$$\begin{aligned}\gamma_0 &= E(Salario) - \gamma_1 E(Sexo)\\ \gamma_1 &= \beta_1 + \beta_2\delta_1\end{aligned}$$

donde β_1, β_2 son los coeficientes de $Sexo$ y $Educación$, respectivamente, en $E\,(Salario\,|Sexo,\ Educación)$ y δ_1 es la pendiente del $PLO\,(Educación\,|Sexo)$, es decir,

$$\delta_1 = \frac{C(Sexo, Educación)}{V(Sexo)}.$$

Por tanto,

$$\gamma_1 = -4,1 + 0,8 \times 3,0 = -1.7$$

(d) Es preferible el coeficiente del $Sexo$ en $E\,(Salario\,|Sexo,\ Educación)$, porque tiene en cuenta que la distribución de los niveles de educación difiere por sexo. Si utilizáramos $PLO(Salario\,|Sexo)$ estaríamos atribuyendo como discriminación el hecho de que la educación está negativamente correlacionada con el sexo.

PREGUNTA 4 Se quiere estudiar la relación que existe entre asistencia a clase y calificaciones finales en una universidad. Considere un modelo que relaciona las calificaciones obtenidas en una universidad (ACT) con el porcentaje de clases atendidas ($ATNDRTE$).

El modelo considerado es muy simple

$$ACT = \beta_0 + \beta_1 ATNDRTE + \varepsilon,$$

donde ε es un término de error con media cero.

(a) (1 punto) Enumere las propiedades que ha de cumplir un buen instrumento de la variable $ATNDRTE$. ¿Puede considerarse como instrumento la distancia desde el domicilio del alumno al campus ($DIST$)? Utilice las salidas en lo posible.

(b) (1 punto) Utilizando las salidas y aplicando los contrastes que considere pertinentes, ¿puede concluirse que el estimador de mínimos cuadrados de β_1 es consistente?

(c) (1 punto) A la vista de sus respuestas, comente sobre el signo de β_1 a partir de las estimaciones por mínimos cuadrados y por variables instrumentales ¿cual de ellos parece el más razonable? ¿podemos concluir que β_1 es no significativo?

SALIDA 1:
Dependent Variable: ACT
Method: Least Squares
Sample: 1 680
Included observations: 680

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	25.12694	0.648325	38.75670	0.0000
ATNDRTE	-0.032024	0.007768	-4.122784	0.0000

SALIDA 2:
Dependent Variable: ATNDRTE
Method: Least Squares
Sample: 1 680
Included observations: 680

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	47.12702	2.872615	16.40562	0.0000
DIST	-13.36898	1.086703	-12.30233	0.0000

SALIDA 3:
RES son los residuos de la salida 2.

Dependent Variable: ACT
Method: Least Squares
Sample: 1 680
Included observations: 680

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.653993	1.316797	6.572001	0.0000
ATNDRTE	0.169580	0.016052	10.56438	0.0000
RES	-0.246607	0.017753	-13.89061	0.0000

SALIDA 4:
Dependent Variable: ACT
Method: Two-Stage Least Squares
Sample: 1 680
Included observations: 680
Instrument list: DIST

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.653993	2.106047	4.109117	0.0000
ATNDRTE	0.169580	0.025673	6.605333	0.0000

SOLUCIÓN:

(a) Un buen instrumento debe no estar correlacionado con el término de error, pero ha de estar correlacionado con $ATNDRTE$. La variable $DIST$ puede ser un buen instrumento porque no debe estar correlacionada con las variables omitidas que entran en el término de error (motivación, capacidad intelectual, renta familiar, etc), pero, según la salida 2 está correlacionada con la variable explicativa del modelo.

(b) A partir de la salida 3 podemos realizar un contraste de Hausman sobre la hipótesis nula : "β_1 es consistente", que se rechaza en vista de la salida 3, donde el coeficiente de los residuos de la forma reducida de $ATNDRTE$ es estadísticamente distinto de cero.

(c) Debemos mirar a la salida 4, que es la que ofrece una estimación consistente, y concluimos que el signo es positivo y significativo, como esperábamos. La estimación por MCO en la salida 1 sugiere un signo negativo, que atenta contra la lógica. Fuera de respuesta: es un buen ejemplo de un resultado espúreo por considerar proyecciones lineales en vez estimar un modelo econométrico.

Valores Críticos

$Z_{0.975} = 1,96$	$Z_{0.95} = 1,645$	$F_{(2,33);0.95} = 3,28$	$F_{(2,33);0.975} = 4,13$	$F_{(2,33);0.99} = 5,31$
$F_{(1,28);0.95} = 4,20$	$F_{(1,30);0.99} = 7,56$	$F_{(1,30);0.95} = 4,17$	$F_{(1,28);0.99} = 7,64$	$\chi^2_{1;0.99} = 6,63$
$\chi^2_{3;0.975} = 9,35$	$\chi^2_{3;0.95} = 7,82$	$\chi^2_{1;0.95} = 3,84$	$\chi^2_{2;0.95} = 5,99$	$\chi^2_{2;0.99} = 9,21$
$t_{100;0.975} = 1,984$	$t_{94;0.975} = 1,985$	$t_{100;0.95} = 1,66$	$t_{94;0.95} = 1,661$	$t_{94;0.995} = 2,629$

$\Pr(Z > Z_\alpha) = \alpha$; $\Pr\left(F_{(q,v)} > F_{(q,v);\alpha}\right) = \alpha$; $\Pr\left(\chi^2_q > \chi^2_{q;\alpha}\right) = \alpha$; $\Pr(t_q > t_{q;\alpha}) = \alpha$.

Z denota una variable aleatoria (v.a) normal con media cero y varianza uno; $F_{(q,v)}$ denota una v.a. F con q grados de libertad en el numerador y v en el denominador; χ^2_q denota una v.a. Ji-cuadrado con q grados de libertad; t_q denota una t de Student con q grados de libertad.