

Examen de Introducción a la Econometría
Universidad Carlos III de Madrid

1ª Convocatoria

Curso 2005/2006

Conteste las preguntas siguientes en 2 horas y media

1. [2 puntos] Sean tres variables aleatorias discretas U , X e Y tales que $Y = X + U$. El siguiente cuadro describe la función de masa de probabilidad conjunta de X y U :

	$U = -7$	$U = 3$
$X = 5$	0,2	0,3
$X = 15$	0,1	0,4

- a. [0,75 puntos] Obtenga los siguientes momentos poblacionales: $E(X)$, $E(U)$, $Cov(X, U)$, $Var(X)$, $Var(U)$, $E(Y)$, $Var(Y)$ y $Cov(X, Y)$.
- b. [0,75 puntos] Calcule el predictor lineal óptimo de la variable Y dado X . Es decir, calcule el valor de los parámetros a y b del $PLO(Y|X) = a + bX$ y ofrezca el valor predicho para Y en función del valor de X según esta fórmula. Demuestre que el predictor óptimo es lineal, es decir, $E(Y|X) = PLO(Y|X)$.
- c. [0,5 puntos] Calcule $E(X|Y)$.
2. [3 puntos] Considere el modelo lineal

$$\log(price) = \delta_0 + \delta_1 age + u$$

donde $price$ es el precio de las casas en un determinado municipio y age es la edad de la casa (en años). Se dispone de una muestra de 321 viviendas y se estima la ecuación por MCO.

- a. [1 punto] Si el intervalo de confianza al 95% para el parámetro δ_1 se sitúa en $(-0,0067524, -0,0040476)$, ¿cuál sería la estimación resultante del parámetro δ_1 ? ¿Cuál es la interpretación de la estimación de este coeficiente? ¿Es la variable age significativa al 5%? ¿Y al 1%?
- b. [0,5 puntos] Si la desviación típica muestral de la de la variable age es igual a 32,51507343, ¿cuál es la covarianza muestral entre las variables $\log(price)$ y age ?
- c. [1 punto] El R^2 de la anterior estimación es de 0,161051. La teoría económica indica que la variable $rooms$ (número de habitaciones de la vivienda) también debería incluirse como variable explicativa en la ecuación. Si se estima una nueva ecuación donde se incluye esta variable explicativa (además de age) y su R^2 es igual a 0,385150, ¿se puede rechazar la hipótesis nula de que la variable $rooms$ es no significativa?
- d. [0,5 puntos] Si realizamos el ajuste de la variable $rooms$ sobre la variable age obtenemos el siguiente resultado:

Dependent Variable: ROOMS
 Method: Least Squares
 Sample: 1 321
 Included observations: 321

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AGE	-0.001417			
C	6.611182			
F-statistic		0.838128		
Prob(F-statistic)		0.360624		

¿Se puede rechazar la hipótesis de que la estimación de la pendiente de la variable age es la misma en el modelo de regresión "corto" que en el "largo"?

3. [3 puntos] Las empresas pueden financiarse mediante la emisión de acciones (fondos propios) o mediante la emisión de bonos (deuda). Una teoría económica afirma que la ratio de deuda sobre fondos propios, $100 * Deuda / (Fondos Propios)$, dependerá de los tipos impositivos que afronte la empresa. En concreto,

- las empresas preferirán utilizar el endeudamiento cuanto mayores sean los impuestos sobre los beneficios de las empresas, *ceteris paribus*
- cuantos mayores sean los impuestos sobre las ganancias de capital (*ceteris paribus*), mayor será el endeudamiento de las empresas
- cuantos mayores sean los impuestos de la renta (*ceteris paribus*), menor será el endeudamiento de las empresas

Usted dispone de datos para 51 estados americanos con información sobre los tipos impositivos de los impuestos sobre el beneficio empresarial, $tprof$, sobre las ganancias de capital, $tcap$, y sobre la renta, $tinc$; además, se conoce la ratio (media) de deuda sobre fondos propios de las empresas ubicadas en cada estado, $debrat$. Todas las variables están medidas en puntos porcentuales. Con la muestra, disponible se han estimado por Mínimos Cuadrados Ordinarios los siguientes modelos:

$$\widehat{debrat}_i = 69,3662 + 1,6223 tprof_i - 1,1214 tinc_i + 1,0777 tcap_i \quad (1)$$

(9,8700) (1,0607) (0,7028) (1,3427)

$n = 51 \quad R^2 = 0,350553 \quad SSR = 7328,5176$

$$(\widehat{debrat}_i - tcap_i) = 81,7726 + 0,9886 (tprof_i - tcap_i) - 0,3085 (tinc_i - tcap_i) \quad (2)$$

(4,8050) (0,9749) (0,4201)

$n = 51 \quad R^2 = 0,063316 \quad SSR = 7649,2574$

$$\widehat{debrat}_i = 69,4269 + 1,5778 tprof_i - 1,1302 (tinc_i - tcap_i) \quad (3)$$

(9,6765) (0,3989) (0,6680)

$n = 51 \quad R^2 = 0,350524 \quad SSR = 7328,8376$

$$\widehat{debrat}_i = 69,3662 + 1,6223 tprof_i - 1,1214 (tinc_i - tcap_i) - 0,0437 tcap_i \quad (4)$$

(9,8700) (1,0607) (0,7028) (0,9651)

$n = 51 \quad R^2 = 0,350553 \quad SSR = 7328,5176$

- a. [1 punto] Fíjese en los resultados del Modelo (1). Si no existieran impuestos, ¿tendrían las empresas en promedio más Deuda que Fondos Propios? ¿Cuál sería su respuesta si un estado fija un tipo impositivo de 15 puntos porcentuales para las ganancias de capital, de 20 puntos porcentuales en el impuesto sobre la renta y de 30 puntos porcentuales en el impuesto sobre los beneficios empresariales?
- b. [0,75 puntos] Contraste con un nivel de significación del 1% la hipótesis de que conjuntamente las cuestiones fiscales son relevantes en las decisiones de endeudamiento de las empresas. ¿Sería la conclusión diferente si hubiera hecho tres contrastes de significatividad individual con el mismo nivel de significación cada uno? ¿Por qué?
- c. [0,75 puntos] Contraste, con un nivel de significación del 1%, la hipótesis de que un aumento simultáneo de todos los impuestos en un punto porcentual provocará un aumento de un punto en la ratio de deuda sobre fondos propios. Ofrezca el valor tanto del estadístico F como el del estadístico t que se pueden utilizar para realizar este contraste.

- d. [0,5 puntos] Para los resultados del Modelo (1), se conoce que la matriz de varianzas y covarianzas de los estimadores de MCO son

$$\text{Var} \begin{pmatrix} \widehat{\beta}_{tprof} \\ \widehat{\beta}_{tinc} \\ \widehat{\beta}_{tcap} \end{pmatrix} = \begin{pmatrix} 1,1251014 & -0,0399419 & -0,9069212 \\ -0,0399419 & 0,4939748 & -0,6826735 \\ -0,9069212 & -0,6826735 & 1,802722 \end{pmatrix}$$

Un gobierno anuncia una reforma fiscal consistente en reducir el impuesto de la renta en 5 puntos porcentuales y aumentar el impuesto sobre ganancias de capital en 10 puntos (manteniendo constante el impuesto sobre los beneficios). Ofrezca un intervalo de confianza al 90% para el cambio esperado, en promedio, en la ratio de deuda sobre fondos propios. ¿Considera probable, con un nivel de confianza del 90%, que esta política provoque una caída de 10 puntos en la ratio de deuda sobre fondos públicos?

4. [2 puntos] La Oficina de Estadísticas Laborales de Estados Unidos realiza todos los años una encuesta para estudiar la situación del mercado de trabajo norteamericano. Entre las variables que se pueden construir a partir de la encuesta se encuentran el logaritmo neperiano del salario nominal a la hora (*lwage*), los años efectivos de educación universitaria (en exceso de 16.5 años) (*educa*), y una variable artificial (dummy) que toma valor 1 si el encuestado es mujer (*female*) y 0 en caso contrario. Se disponen de datos para el año 2000 para trabajadores con experiencia laboral menor de un año y estudios universitarios (al menos 16,5 años de educación). Entre estos trabajadores, la variable *educa* solo toma tres valores: 0, 1, 5 y 3, 5. Además, se genera una nueva variable interactuando *educa* y *female*, $edufem = educa * female$.

Con los datos disponibles se realizan las siguientes regresiones MCO (errores estándar en paréntesis):

Regresión 1:

$$\widehat{lwage} = 2,676852 + 0,1310272educa$$

(,041913) (,0206192)

$n = 501; \quad RSS = 97,342215946$

Regresión 2:

$$\widehat{lwage} = 2,73752 + 0,1874562educa - 0,1217951female - 0,1090569edufem$$

(,0287494) (,0590482) (,080746) (,0828375)

$n = 501; \quad RSS = 94,649281968$

- a. [0,75 puntos] Interprete el valor del coeficiente de *educa* en la regresión 1. Calcule el diferencial salarial medio exacto, en tantos por ciento, entre los trabajadores con estudios superiores universitarios de postgrado (es decir, con 20 años de educación) y los trabajadores con una titulación de tres años (es decir, con 16.5 años de educación).
- b. [0,75 puntos] Interprete el valor del coeficiente de *edufem* en la regresión 2. ¿Son las diferencias por sexo en los rendimientos a la educación significativas? Justifique su respuesta.
- c. [0,5 puntos] Indique detalladamente cómo contrastaría que los rendimientos a la educación superior de las mujeres no son lineales respecto a los años en educación superior.

VALORES CRÍTICOS:

$N(0, 1)$
$\Pr(N(0, 1) > 2, 576) = 0, 005$
$\Pr(N(0, 1) > 2, 326) = 0, 01$
$\Pr(N(0, 1) > 1, 960) = 0, 025$
$\Pr(N(0, 1) > 1, 645) = 0, 05$
$\Pr(N(0, 1) > 1, 282) = 0, 10$

$\chi^2_{(1)}$	$\chi^2_{(2)}$	$\chi^2_{(3)}$
$\Pr(\chi^2_{(1)} > 6, 63) = 0, 01$	$\Pr(\chi^2_{(2)} > 9, 21) = 0, 01$	$\Pr(\chi^2_{(3)} > 11, 34) = 0, 01$
$\Pr(\chi^2_{(1)} > 3, 84) = 0, 05$	$\Pr(\chi^2_{(2)} > 5, 99) = 0, 05$	$\Pr(\chi^2_{(3)} > 7, 81) = 0, 05$
$\Pr(\chi^2_{(1)} > 2, 71) = 0, 10$	$\Pr(\chi^2_{(2)} > 4, 61) = 0, 10$	$\Pr(\chi^2_{(3)} > 6, 25) = 0, 10$

Recordamos que una t de Student con n grados de libertad se comporta como un $N(0, 1)$ para n razonablemente grande ($n > 30$). Por otro lado, una F de Fisher con q grados de libertad en el numerador y n grados de libertad en el denominador se comporta como una $\chi^2_{(q)}/q$.