Examen de Introducción a la Econometría

Universidad Carlos III de Madrid

2^a Convocatoria

Curso 2005/2006

Conteste las preguntas siguientes en 2 horas y media

1. [2 puntos] Las desviaciones porcentuales sobre la tendencia de la renta agregada bruta disponible de las familias, Y_d, se descomponen en desviaciones del consumo agregado y del ahorro, $Y_d = C + S$. A partir de una muestra de tamaño $n = 100$ de datos de series temporales, se obtienen los siguientes momentos muestrales: $\hat{E}_n(C) = 1,\ \hat{E}_n(Y_d) = 0,75,\ \widehat{Cov}_n(Y_d, C) = 0,8,\ \widehat{Var}_n(Y_d) = 1,2,\ \widehat{Var}_n(S) = 2/3$.

a. [1 punto] Obtenga por el principio de analogía un estimador del predictor lineal óptimo del consumo agregado dada la renta disponible:

$$C = a_c + b_c Y_d$$

así como de la varianza del error de predicción, $\sigma_c^2 = Var(C - b_c Y_d)$.

b. [0,5 puntos] Considere ahora la predicción del ahorro dada la renta disponible:

$$S = a_s + b_s Y_d$$

Demuestre que $a_c + a_s = 0,\ b_c + b_s = 1$, y que $\sigma_c^2 = \sigma_s^2$.

c. [0,5 puntos] Si le dan un valor de renta disponible, ¿qué variable preferiría predecir, el consumo o el ahorro? Justifique su respuesta.

2. [2 puntos] Se quiere explicar la nota final obtenida en la selectividad, *selec*, en función de las horas semanales medias de estudio, *horas* y de la asignación a un programa de especial de ayuda al estudio, por medio de un modelo de regresión lineal

$$selec = \beta_0 + \beta_1 horas + \beta_2 prog + u \tag{1}$$

donde *prog* es una variable dummy de tal forma que $prog_i = 1$ si el estudiante i es asignado al programa especial de ayuda al estudio.

a. [0,5 puntos] Interprete los coeficientes del modelo (1). ¿Cuál sería la relación entre los coeficientes del modelo (1) y los del modelo lineal basado solamente en *horas*,

$$selec = \delta_0 + \delta_1 horas + v, \tag{2}$$

si se supone que la asignación al programa se ha efectuado aleatoriamente?

b. [0,5 puntos] Si en en cambio, se propusiese la asignación al programa favoreciendo a aquellos estudiantes con menor hábito de estudio, ¿cuál sería la relación esperada entre β_1 y δ_1?

c. [0,5 puntos] Explica como contrastarías la efectividad del programa de ayuda al estudio dada una muestra. Si sólo hay recursos para incluir a un 10% de los alumnos en el estudio, ¿qué política esperas que sea más efectiva para contrastar la efectividad del programa de ayuda al estudio, la asignación propuesta en (a) o la propuesta en (b)?

d. [0,5 puntos] Se quiere modificar el modelo (1) para permitir que la efectividad de tiempo de estudio aumente con la inclusión en el programa. Proponga el modelo adecuado para ello e indique como contrastaría en dicho caso la existencia de un beneficio por la asignación al programa para los estudiantes que dedican 20 horas semanales a estudiar en media.

3. [3 puntos] Se dispone de una muestra de 550 empresas textiles y del sector automovilístico con información sobre el valor de su producción Y (en millones de euros), las unidades de trabajo empleadas N (en miles de horas trabajadas) y las unidades de capital utilizado en la producción K. También se tiene una variable, *textil*, que toma valor 1 si la empresa pertenece al sector textil y cero si pertenece al sector automovilístico. Se supone que la función de producción de las empresas es una Cobb-Douglas (es decir, $Y_i = AN_i^{\beta_1} K_i^{\beta_2} e^{U_i}$); aplicando logaritmos se obtiene una versión lineal de esta función de producción $\ln Y_i = \ln A + \beta_1 \ln N_i + \beta_2 \ln K_i + U_i$. Por tanto, se han estimado los siguientes modelos por Mínimos Cuadrados Ordinarios:

$$\widehat{\ln Y_i} = \underset{(0{,}6601)}{-0,5241} \underset{(0{,}1391)}{+0,7366} \ln N_i \underset{(0{,}0731)}{+0,4453} \ln K_i \tag{3}$$
$$n = 550 \quad R^2 = 0,1225 \quad SSR = 541,4375$$

$$\widehat{\ln Y_i} = \underset{(0{,}9505)}{0,6514} \underset{(1{,}3059)}{-1,9003}\ textil \underset{(0{,}2003)}{+0,5334} \ln N_i \underset{(0{,}0950)}{+0,3321} \ln K_i \underset{(0{,}2747)}{+0,3904}\ textil * \ln N_i \underset{(0{,}1492)}{+0,1214}\ textil * \ln K_i \tag{4}$$
$$n = 550 \quad R^2 = 0,1565 \quad SSR = 520,4230$$

$$\widehat{\ln(Y_i/N_i)} = \underset{(0{,}0835)}{0,0518} \underset{(0{,}1097)}{+0,3846}\ textil \underset{(0{,}0893)}{+0,3530} \ln(K_i/N_i) \underset{(0{,}1387)}{+0,0164}\ textil * \ln(K_i/N_i) \tag{5}$$
$$n = 550 \quad R^2 = 0,0949 \quad SSR = 524,2166$$

$$\widehat{\ln(Y_i/N_i)} = \underset{(0{,}0547)}{0,2884} \underset{(0{,}0683)}{+0,4131} \ln(K_i/N_i) \tag{6}$$
$$n = 550 \quad R^2 = 0,0625 \quad SSR = 542,9477$$

$$\widehat{\ln Y_i} = \underset{(0{,}5205)}{2,7303} \underset{(0{,}9312)}{-4,3011}\ textil \underset{(0{,}0770)}{+0,1854} \ln(K_i/N_i) \underset{(0{,}2525)}{+0,6809}\ textil * \ln N_i \underset{(0{,}1089)}{+0,3889} (textil * \ln K_i + \ln N_i) \tag{7}$$
$$n = 550 \quad R^2 = 0,1460 \quad SSR = 526,9304$$

a. [0,75 puntos] Suponga inicialmente que creemos que los parámetros de la función de producción son iguales en ambos sectores industriales. Contraste con un nivel de significación del 5% la hipótesis de que existen rendimientos constantes a escala.

b. [0,75 puntos] Ahora suponga por contra que creemos que existen rendimientos constantes a escala. Contraste con un nivel de significación del 5% la hipótesis de que las funciones de producción son iguales en ambos sectores.

c. [1 punto] Contraste con un nivel de significación del 5% la hipótesis conjunta de que en ambos sectores existen rendimientos constantes a escala (aunque los parámetros de las funciones de producción son diferentes).

d. [0,5 puntos] Un experto industrial afirma que si una empresa del sector automovilístico aumenta las unidades de capital utilizadas en un 2% y además las unidades de trabajo empleadas en un 1% entonces el cambio (porcentual) en su producción será, en promedio, el mismo que el de una empresa textil que aumenta las unidades de capital utilizadas en un 1% y mantiene el trabajo constante. Contraste con un nivel de significación del 5% si existe evidencia empírica a favor de la aseveración del experto.

4. [**3 puntos**] Una empresa multinacional ha sido acusada de discriminar a sus empleadas pagándoles salarios menores que a los empleados. El presidente solicita una muestra aleatoria simple de 250 empleados y empleadas, todos/as con la misma experiencia y años de educación. En la muestra, se disponen de las siguientes variables:

$$salario_i = \text{salario mensual del empleado/a } i \text{ (en euros)}$$

$$mujer_i = \begin{cases} 1 \text{ si el individuo } i \text{ es mujer} \\ 0 \text{ en caso contrario} \end{cases}$$

$$directiv_i = \begin{cases} 1 \text{ si el individuo } i \text{ tiene un puesto de trabajo directivo (toma de decisiones)} \\ 0 \text{ en caso contrario} \end{cases}$$

A partir de esa muestra, el presidente obtiene la siguiente tabla con medias de los salarios para hombres y mujeres (y errores estándar de las respectivas medias entre paréntesis):

Tabla 1

Hombres	Mujeres
1915, 57	2132, 93
(49,53)	(79,91)

Algunos trabajadores, que estudiaron y comprendieron los conceptos básicos de la Econometría, no están satisfechos con esta evidencia. Solicitan la misma muestra y, a partir de ella, estiman varios modelos econométricos por Mínimos Cuadrados Ordinarios; como el presidente no los entiende muy bien, confeccionan la siguiente la tabla con medias de salarios (y errores estándar de las respectivas medias entre paréntesis):

Tabla 2

	Hombres	Mujeres
En trabajos directivos	2994, 43	2497, 97
	(9,22)	(9,71)
En trabajos no directivos	1501, 87	992, 16
	(6,34)	(18,98)

a. [**0,75 puntos**] Formule un modelo econométrico (incluyendo las variables disponibles en la encuesta y/o variables construidas a partir de ellas) tal que uno de los parámetros pueda interpretarse como la diferencia en el salario promedio entre hombres y mujeres. Utilizando los datos de la Tabla 1, ofrezca valores para la estimación de todos los parámetros en este modelo.

b. [**0,75 puntos**] Formule otro modelo econométrico (incluyendo las variables disponibles en la encuesta y/o variables construidas a partir de ellas) tal que uno de los parámetros pueda interpretarse como la diferencia en el salario promedio entre hombres en puestos directivos y mujeres en puestos directivos. Utilizando los datos de la Tabla 2, ofrezca valores para la estimación de todos los parámetros en este nuevo modelo.

c. [**0,5 puntos**] A partir del modelo formulado en el apartado (a) y de los datos de la Tabla 1, contraste con un nivel de significación del 5% la hipótesis de que los hombres y las mujeres tienen el mismo salario.
[Pista: puesto que se utiliza una muestra aleatoria simple, la covarianza (condicional en educación y experiencia) entre el salario medio de un hombre y el salario medio de una mujer es, *ceteris paribus*, cero. Es decir, si $\overline{y}^h$ e $\overline{y}^m$ son el salario medio, *ceteris paribus,* de un hombre y de una mujer respectivamente, entonces $Cov\left(\overline{y}^h, \overline{y}^m\right) = 0$ (donde $Cov(\cdot)$ debe entenderse como covarianza condicional en educación y experiencia)]

d. [**0,5 puntos**] A partir del modelo formulado en el apartado (b) y de los datos de la Tabla 2, contraste con un nivel de significación del 5% la hipótesis de que los hombres en puestos directivos y las mujeres en puestos directivos tienen el mismo salario.
[Pista: igual que en el apartado anterior, si $\overline{y}^{hd}$, $\overline{y}^{md}$, $\overline{y}^{hn}$ e $\overline{y}^{mn}$ son los salarios medio de hombre directivos, mujeres directivas, hombres no directivos y mujeres no directivas, respectivamente, entonces $Cov\left(\overline{y}^{md}, \overline{y}^{mn}\right) = Cov\left(\overline{y}^{md}, \overline{y}^{hd}\right) = Cov\left(\overline{y}^{md}, \overline{y}^{hn}\right) = Cov\left(\overline{y}^{mn}, \overline{y}^{hd}\right) = Cov\left(\overline{y}^{mn}, \overline{y}^{hn}\right) = Cov\left(\overline{y}^{hd}, \overline{y}^{hn}\right) = 0$]

e. [0,5 puntos] Contraste con un nivel de significación del 5% la hipótesis de que los diferenciales salariales entre hombres y mujeres no cambian según el tipo de trabajo; es decir, que la diferencia en el salario promedio entre hombres directivos y mujeres directivas es la misma que la diferencia en el salario promedio entre hombres no directivos y mujeres no directivas.

VALORES CRÍTICOS:

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

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

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 $\chi^2_{(q)}/q$.