Pontificia Universidad Católica de Chile
Instituto de Economía

Examen de Grado
Econometría y Métodos Cuantitativos
Marzo, 2011

Duración	:	180 minutos
Fecha	:	04 de Marzo de 2011
Hora de comienzo	:	15 horas

El examen consta de cuatro preguntas, cada una de ellas de 40 puntos. Todas las partes deben ser respondidas. Usted dispone de 20 minutos para leer el examen y 160 minutos para responderlo, por lo que cada punto corresponde aproximadamente a un minuto.

Parte I (40 puntos)

1. Suponga el siguiente modelo de Regresión lineal: $Y=\alpha+\beta X+\varepsilon$ que cumple todos los supuestos habituales con la excepción que se cree que presenta heteroscedasticidad. Se dispone de las siguientes observaciones para x e y:

X = 1; 2; -3; 0

Y = 4; 3; -6; -1

a) Calcule el estimador de β. (5 puntos)
b) Calcule s^2, el estimador de la varianza del modelo. (5 puntos)
c) Calcule $\hat{V}(\hat{\beta})$ la varianza estimada del estimador de β. (5 puntos).
d) Calcule el estimador de White para $\hat{V}(\hat{\beta})$, ¿Qué utilidad tiene este estimador? En esta pregunta no es necesario que haga los cálculos puede dejarlos indicados, pero sea muy preciso en su respuesta.(7 puntos)

2. Un investigador estudia la incidencia de menores que cometen hurtos. Para ello emplea los datos de las 35 comunas de Santiago y su zona de extensión en el 2008.

El investigador supone que los hurtos pueden ser "explicados" por la población de menores en situación de pobreza. La ecuación considerada es $H = \beta_1+\beta_2 M + \varepsilon$ en que H es la proporción de hurtos por 1000 habitantes y M el porcentaje de menores en situación de pobreza por 1000 habitantes pobres. La variable ε es un residuo que satisface todos los supuestos habituales del modelo MICO.
Suponga además que M se ha dividido en dos sub grupos: MM la proporción de menores en situación de pobreza en viven en un hogar monoparental con madre jefe de hogar y MP la proporción de menores que viven con ambos padres. No se presentan otras situaciones así que M = MM + MP.
Un primer investigador efectúa una regresión por MICO de H en MM y MP. Los resultados se presentan en la columna (1) de la tabla adjunta. Un segundo investigador considera que M es la variable fundamental para determinar H y que la distinción entre MM y MP no es importante. La regresión de H en M se presenta en la columna (2) de la tabla. Finalmente, un tercer investigador está de acuerdo con el segundo en el sentido que la variable M es la más importante pero cree que MM puede agregar algún efecto adicional. Esta nueva regresión se presenta en la columna (3).

RESULTADOS DE LAS REGRESIONES PARA LOS TRES MODELOS

	(1)	(2)	(3)
MM	0,30 (0,10)	----	0,10 (0,05)
MP	0,20 (0,10)	----	----
M	----	0,25 (0,08)	0,20 (0,10)
Constante	5,01 (1,21)	4,87 (1,15)	5,01 (1,21)
SCR	160	180	160

En que H es la variable dependiente, los valores entre paréntesis son los errores estándar y RSS es la suma de cuadrados residuales.
Utilizando sus conocimientos teóricos y la información que entrega la tabla:

a) Analice las propiedades de los estimadores de la segunda regresión (sesgo y precisión) bajo el supuesto que el primer investigador tiene la razón. (6 puntos)
b) Analice las propiedades de los estimadores de la primera regresión suponiendo que el segundo modelo está correcto. (6 puntos)
c) Analice las propiedades de los estimadores de la tercera regresión suponiendo que el primer modelo está correcto. (6 puntos)

Parte II (40 puntos)

Una organización de estudios de la salud desea diagnosticar la situación de la obesidad en la población chilena y quiere estimar la siguiente relación:

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

donde *y* es el logaritmo del número de calorías per-cápita consumidas en un día y *x* es el logaritmo del ingreso per cápita del hogar.

1) (6 puntos) Imagine que el econometrista que contratan no puede encontrar el valor exacto del número de calorías consumidas pero puede obtener un estimador usando la cantidad de comida comprada per cápita (en gramos) multiplicado por el número de calorías en promedio por gramo. Este estimador es tal que: $y_i^* = y_i + \upsilon_i, \upsilon_i \sim N\left(0, \sigma_v^2\right)$

   ¿Cómo influye este problema en la estimación del parámetro β usando MCO? Analice formalmente la consistencia del estimador del parámetro β y explique qué sucede con esta propiedad cuando la varianza en el error de medición crece.
2) (8 puntos) Imagine que el econometrista que contratan no puede encontrar el valor exacto del ingreso de la persona pero si tiene el valor auto-reportado en una encuesta. Este valor es tal que: $x_i^* = x_i + \eta_i, \eta_i \sim N\left(0, \sigma_\eta^2\right)$

   ¿Cómo influye este problema en la estimación del parámetro β usando MCO? Analice formalmente la consistencia del estimador del parámetro β y explique qué sucede con esta propiedad cuando la varianza en el error de medición crece.
3) (10 puntos) Imagine que el econometrista que contratan no puede encontrar el valor exacto del ingreso de la persona pero si tiene el valor auto-reportado pero en este caso se sabe que la gente de mayores ingresos en general no reporta todo lo que ganan en las encuestas por temor a tener que pagar más impuestos y la gente pobre exagera cuánto ganan para no parecer indigente frente al entrevistador . ¿Cómo influye este problema en la estimación del parámetro β usando MCO? Analice la consistencia del estimador del parámetro β y compare con su respuesta anterior.
4) (10 puntos) Suponga que *y* se encuentra bien medido pero tiene problemas con la medición de *x*. Ahora además de x_i^* tiene otra medida del ingreso dado por:

   $\hat{x}_i = x_i + \mu_i, \mu_i \sim N(0, \sigma_\mu^2), Cov(\mu_i, \eta_i) = 0.$

   Derive el estimador de 2SLS usando la segunda medida como variable instrumental y determine si este estimador converge a β o no. Explique cuáles de los supuestos son necesarios para garantizar este resultado.
5) (6 puntos) Usted es contratado por la organización y como ha aprendido mucho durante sus cursos de pre-grado encuentra una fuente de datos donde puede obtener el valor exacto de las dos variables. A la organización le gustaría usar sus resultados e interpretar al estimador de β como la elasticidad ingreso de la demanda por calorías. Mencione 2 problemas que pueden hacer que esta interpretación no sea adecuada.

Parter III (40 puntos)

En una regresión con los datos de un país se obtienen los siguientes resultados:

$$\Delta Y_t = \underset{(0,25)}{10} + \underset{(0,05)}{0,2Y_{t-1}} + \underset{(0,15)}{0,5X_t} + \underset{(0,2)}{0,6\,X_{t-1}}$$

$R^2 = 0,95$

$N = 50$

Las desviaciones estándar estimadas aparecen entre paréntesis debajo de cada coeficiente.

a) (7 puntos) Observe atentamente los valores de los estimadores. ¿Qué comentario le merecen estos resultados en cuanto a la existencia de una relación de largo plazo entre X e Y?

b) (6 puntos) Si Ud. estimara, con los mismos datos, la regresión:

$Y_t = a + b\,X_t + \varepsilon_t$

¿Qué rango de valores esperaría para el estadígrafo de Durbin Watson?, ¿por qué?

c) (9 puntos) En otra regresión, para otro país, se obtiene:

$$\Delta Y_t = \underset{(0,3)}{10} - \underset{(0,05)}{0,2Y_{t-1}} + \underset{(0,15)}{0,4X_t} + \underset{(0,2)}{0,6\,X_{t-1}}$$

Dados los coeficientes estimados, ¿existiría una relación estimada de largo plazo?, ¿por qué?, ¿cuál sería esa relación?

d) (6 puntos) Reescriba el modelo anterior en el formato de corrección de errores

e) (8 puntos) Explique cómo haría un test para corroborar la existencia de una relación de largo plazo entre Y_t y X_t.

f) (4 puntos) ¿Qué diferencia hay entre "existencia de una relación de largo plazo" y "existencia de cointegración"?

Parte IV (40 puntos).

Considere un individuo que busca empleo en el mercado laboral y el tiempo que demora en recibir una oferta sigue una distribución distribución exponencial con una tasa igual a 1/θ. Es decir,

$$f(y;\theta)=\frac{1}{\theta}\exp(-y/\theta)$$

donde θ >0, e y>0 indexa el tiempo medido en días.

a) ¿Cuál es el tiempo esperado en el cual el individuo recibiría una oferta de trabajo? Indique el procedimiento para obtener dicha expresión. (5 puntos)

b) Suponga que la probabilidad de seguir buscando (sin recibir una oferta) a los 10 días de empezar la búsqueda de empleo es igual a 0.3 ¿Cuál es la probabilidad de seguir buscando (sin recibir una oferta) en 20 días, condicional en que no se ha recibido una oferta en los primeros 10 días? (5 puntos)

c) Ahora suponga que la tasa de ocurrencia de las ofertas no es común para los individuos sino que depende de características observables de ellos. Así, suponga que la distribución condicional del tiempo en que tarda en llegar una oferta es

$$f(y_i \mid x_i;\beta)=\frac{1}{x_i'\beta}\exp(-y_i/x_i'\beta)$$

Donde x_i es un vector de k x 1 de características y β es un vector de parámetros de k x1. Asuma que $x_i'\beta>0$.

Escriba el log de la función de verosimilitud para una muestra de n observaciones independientes. (5 puntos)

d) Compute las condiciones de primer orden (CPO) para la estimación del parámetro β. Muestre que dicha condición de primer orden puede ser escrita de la siguiente forma (10 puntos):

$$\frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{(x_i'\hat{\beta})^2}u_i(\hat{\beta})=0$$

donde $u_i(\hat{\beta})$ es un seudo-residuo que Ud. debe encontrar. Muestre además que $E[u_i(\beta)\mid x_i]=0$. (5 puntos)

e) Dado que la CPO de la pregunta anterior no permite una solución analítica cerrada para $\hat{\beta}$, explique un algorítmo numérico y sus supuestos que permitan encontrar una solución numérica para $\hat{\beta}$. Discuta los potenciales problemas que pueden amenzar la convergencia de dicho algoritmo. (10 puntos)