Econometría 06216
Examen Parcial #1
Grupo 3
Respuestas Sugeridas
Cali, Lunes 5 de Marzo de 2007

Profesor: Julio César Alonso

Estudiante: ________________________
Código: ______________________

Instrucciones:

1. Lea cuidadosamente todas las preguntas e instrucciones.
2. Este examen consta de **5** páginas; además, deben tener 1 página de fórmulas.
3. El examen consta de 4 preguntas que suman un total de 100 puntos. El valor de cada una de las preguntas esta expresado al lado de cada pregunta.
4. Escriba su respuesta en las hojas suministradas, marque cada una de las hojas con su nombre. NO responda en las hojas de preguntas.
5. El examen esta diseñado para una hora, pero ustedes tienen 3 horas para trabajar en él.
6. Recuerde que no se tolerará ningún tipo de deshonestidad académica. En especial usted no puede emplear ningún tipo de ayuda diferente a la que se le entrega con este examen.
7. El uso de calculadoras está prohibido
8. Al finalizar su examen entregue sus hojas de respuesta, así como las horas de preguntas.
9. Asigne su tiempo de forma eficiente!

Suerte.

1 Falso o Verdadero (25 puntos en total, 5 puntos cada subparte)

Diga si las siguientes afirmaciones son falsas o verdaderas y explique en dos o tres líneas su respuesta. (No se dará ningún crédito a respuestas sin justificación.)

a) Si bien el siguiente modelo $w_i = sen(\alpha) \cdot X_i^{3\beta_1} \cdot Y_i^{\beta_3\beta_4} \cdot \theta \cdot Z_i^{\beta_4^{\gamma}} \cdot \varepsilon_i$, no es lineal desde el punto de vista matemático, si se puede emplear los estimadores MCO para encontrar todos los coeficientes del modelo.

Falso, Noten que en este caso el modelo es linealizable:

$$w_i = sen(\alpha) \cdot X_i^{3\beta_1} \cdot Y_i^{\beta_3\beta_4} \cdot \theta Z_i^{\beta_4^{\gamma}} \cdot \varepsilon_i$$

$$\ln(w_i) = \ln\left(sen(\alpha) \cdot X_i^{3\beta_1} \cdot Y_i^{\beta_3\beta_4} \cdot \theta Z_i^{\beta_4^{\gamma}} \cdot \varepsilon_i\right)$$

$$\ln(w_i) = \ln\left(sen(\alpha) + \theta\right) + 3\beta_1 \ln(X_i) + \beta_3\beta_4 \ln(Y_{ii}) + \beta_4^{\gamma} \ln(Z_i) + \ln(\varepsilon_i)$$

Si bien el modelo si se puede estimar por MCO, no podríamos estimar TODOS los coeficientes. Por ejemplo, γ y θ no pueden ser encontrados.

b) Después de estimar el siguiente modelo $y_i = \beta_0 + \beta_1 X_i + \alpha W_i + \alpha \ln(Z_i) + \varepsilon_i$, publiqué en mi página Web la siguiente Tabla Anova. Pero lastimosamente la Tabla no quedó bien cargada y se omitió un número que fue remplazado por XXX. Un estudiante me envió un correo con la siguiente afirmación: "Creo que uno de los números que falta es128.". ¿Es esta afirmación falsa o verdadera?

Fuente de variación	***SS***	***G de L***	***MS***
Regresión	XXX	XXX	50
Error	300	XXX	2.4
TOTAL	450	XXX	

Verdadero, pues los números que faltan son: 150, 3, 128 y 125.

c) Si bien el estimador de máxima verosimilitud para los β s es igual al del método de MCO, emplear el método de máxima verosimilitud o el de mínimos cuadrados ordinarios para estudiar los β s de un modelo lineal implica los mismo resultado cuantitativos y cualitativos

Falso, aunque los estimadores para los β s son iguales para los dos métodos, existe una diferencia en el estimador de la varianza del error. En especial, para muestras pequeñas tendremos que la varianza del error por el método de máxima verosimilitud es sesgado y por tanto lo será la matriz de varianzas y covarianzas de los betas estimados. Así los resultados, paras las pruebas de hipótesis serán diferentes en ambos casos.

d) Sean X, Y y c dos variables aleatorias y una constante, respectivamente. Además suponga que $E[X] = c^2$, $E[Y] = c$, entonces X y cY son estadísticamente independientes.

Falso, pues para que exista independencia de requiere que $E[XcY] = cE[XY] = cE[X]E[Y]$ y esto no necesariamente es cierto en este caso.

e) Sea A una matriz de cualquier dimensión (n x m), entonces siempre que $Ran(A) = n$ entonces la inversa de A existirá.

Falso, pues cuando $n \neq m$, tendremos que la matriz no es cuadrada y por tanto no existirá inversa.

2 Selección Múltiple (15 puntos en total, 3 puntos cada subparte)

Determine cuál de las siguientes respuestas es la correcta. Escoja la mejor opción y explique en dos o tres líneas su respuesta. (No se dará ningún crédito a respuestas sin justificación.)

2.1. Considere el siguiente modelo de regresión estimado:

$$\hat{Y} = \underset{(0.11)}{0.32} + \underset{(0.2)}{1.54} \cdot X_1,$$

donde los números entre paréntesis son los errores estándar. Suponga que la muestra es de tamaño 220 y considere la siguiente prueba de hipótesis:

$H_0 : \beta_2 = 2.0$ Vs $H_A : \beta_2 \neq 2.0$

Entonces, el estadístico t (redondeando a tres dígitos) será:

a) 2.077
b) 5.677
c) -2.3
d) Ninguno de los anteriores

Respuesta: c)

2.2. Se sabe que dos estimadores para la pendiente en un modelo de regresión simple (β_2) corresponden a $\hat{\beta}_2^{Método1} = 2/3 \cdot \hat{\beta}_2^{MCO}$ y $\hat{\beta}_2^{Método2} = 1/3 \cdot \hat{\beta}_2^{MCO}$, donde $\hat{\beta}_2^{MCO}$ es el correspondiente estimador de MCO y los otros estimadores corresponde a dos métodos diferentes., entonces se puede decir que:

a. El estimador $\hat{\beta}_2^{Método3} = \hat{\beta}_2^{Método1} + \hat{\beta}_2^{Método2}$ es un estimador consistente de β_2.
b. $\hat{\beta}_2^{Método1}$ es un estimador con un sesgo mayor que $\hat{\beta}_2^{Método2}$.
c. Tanto $\hat{\beta}_2^{Método1}$ y $\hat{\beta}_2^{Método2}$ son estimadores eficientes de β_2.
d. Ninguna de las anteriores.

Respuesta: a)

2.3. Considere el siguiente modelo estimado: $\hat{H}_i = 1.32 + 2.02 \cdot W_i - 2.38 \cdot C_i$, i = 1,...,n, donde H_i es el número de horas por semana laboradas por el trabajador i (medido en horas), W_i representa el salario (medidos en miles de pesos por hora) del trabajador i y C_i es el número de niños menores de 10 años que tiene el trabajador i. Entonces, de acuerdo a esta ecuación estimada de oferta de trabajo, si el trabajador i recibe $21,520 por hora y no tiene hijos, entonces el número de horas ofrecidas será (redondeando):

a. 45.3 horas

b. 43.5 horas
c. 44.8 horas
d. 42.1 horas

Respuesta: c)

2.4. Recuerde que la variable W_i en el modelo estimado presentado en Pregunta anterior está medida en miles de pesos por hora. Suponga que ahora la variable W_i fuera medida en pesos por hora. Entonces, el modelo estimado cambiará de la siguiente manera:
a. el coeficiente estimado asociado con W_i será 20.202.
b. el coeficiente estimado asociado con C_i será 23.802.
c. el coeficiente estimado asociado con W_i será 0.202.
d. el coeficiente estimado asociado con W_i será 0.00202.

Respuesta: d) (Pero existía una omisión de un cero, así que será válido cualquier respuesta consecuente con esto)

2.5. Para el modelo ~~estimado~~ $Y_i = \alpha_0 + \alpha_1 X_i + \varepsilon_i$ es posible afirmar:

a) ε_i es medible directamente.
b) ε_i representa la distancia entre el punto observado (data point) y la línea real de regresión.
c) ε_i representa el error estimado.
d) Ninguna de las anteriores.

Respuesta: B)

3 (30 puntos)

Como miembro del grupo de análisis económico del Banco central de una pequeña economía, se le encarga continuar con el trabajo del econometrista principal que se acaba de retirar. El último trabajo del econometrista principal era estimar la siguiente función de demanda de dinero:

$$M1_t = \beta_1 + \beta_2 \ln\left(TB3mo_t\right) + \beta_3 \ln\left(GDP_t\right) + \varepsilon_t \quad (1)$$

Donde $M1_t$ y GDP_t representan la cantidad de dinero (medida por M1) y el PIB, respectivamente. Ambas cantidades están medidas en millones de moneda local. Finalmente, $TB3mo_t$ representa la tasa de interés de los bonos del tesoro a 3 meses.

El Econometrista principal, antes de retirarse estimó el modelo empleando datos trimestrales y dejó los resultados que se reportan al final del examen. Lastimosamente, no se cuenta con la base de datos originalmente empleada por el econometrista, y lo único con que se cuenta es con unos resultados parciales. Específicamente, unos de los resultados no son legibles y se remplazaron con “XXX”.

A partir de esta información, responda las siguientes preguntas.

a) Encuentre los valores perdidos (no es necesario que realice todas las operaciones, pero si es importante mostrar que números emplearía y como para obtener los valores perdidos).**(8 Puntos, 2 puntos cada valor perdido)**

9.02= 39.00536/ 4.322863

$$R^2 = 0.755 = \frac{(1326511691\text{-}324358591.9)}{1326511691}$$

$$F_C = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} = 254.8957628$$

Reject, pues F calculado es mayor que el individual.

b) Comente la significancia individual y conjunta de los coeficientes **(4 Puntos, 2 puntos cada coeficiente)**

Los coeficientes son conjunamente significativos (1 punto)
Unicamente el intercepto es significativo.

(se esperaba que ustedes mostrarán porque llegan a su conclusión)

c) Teniendo en cuanta su respuesta en la pregunta anterior, y los resultados reportados, interprete los coeficientes estimados del modelo.**(6 Puntos, 2 puntos cada coeficiente)**

Noten que según estos resultados tenemos que:

$$M1_t = \beta_1 + \beta_2 \ln(TB3mo_t) + \beta_3 \ln(GDP_t) + \varepsilon_t$$

$\hat{\beta}_1$= no tiene interpretación económica.

$\hat{\beta}_2$ no es significativo, por tanto esto implica que la tasa de interés no tiene efecto sobre la demanda de dinero.

$\hat{\beta}_3$ = no es significativo, por tanto esto implica que los cambios en el PIB no tienen efecto sobre la demanda de dinero.

d) Dado que la muestra contiene información trimestral desde 1964:2, el econometrista quería comprobar si era cierto que la elasticidad ingreso de la demanda por dinero es mayor para el proceso de democratización del sistema financiero que inició de el primero de enero de 1990 que antes de dicho proceso. Plantee un modelo que permita determinar dicho cambio **(muestre claramente porque su modelo si puede servir a comprobar dicha hipótesis. 8 Puntos)**, y muestre claramente cuál sería la hipótesis nula y alterna que probaría, así como los valores que tendría que emplear para llegar a la respuesta. **(4 Puntos por esta parte de la respuesta (12Puntos en total)**

Dado que $\frac{\beta_3}{100} = \frac{\partial M2_t}{\Delta \% PIB}$ y lo que deseamos es conocer la elasticidad, entonces dicha elasticidad será igual a:

$$\eta = \frac{\partial M2_t}{\Delta \% PIB} \frac{100}{\overline{M2}_t} = \frac{\beta_3}{100} \frac{100}{\overline{M2}} = \frac{\beta_3}{\overline{M2}}$$

Esto quiere decir que debemos escribir un modelo que permita el comprobar si existe un cambio en la pendiente con respecto a ln(PIB) o no. Así, el modelo será:

$$M2_t = \gamma_1 + \gamma_2 TB3mo_t + \gamma_3 \ln(GDP_t) + \alpha D_t \ln(GDP_t) + \varepsilon_t$$

Donde:

$$D_t = \begin{cases} 1 & si\ t \geq 1990:1 \\ 0 & o.w. \end{cases}$$

Es fácil demostrar que:

$$M2_t = \begin{cases} \gamma_1 + \gamma_2 TB3mo_t + (\gamma_3 + \alpha)\ln(GDP_t) & si\ t \geq 1990:1 \\ \gamma_1 + \gamma_2 TB3mo_t + \alpha \ln(GDP_t) & o.w. \end{cases}$$

Así, la tendremos que:

$$\eta = \begin{cases} \frac{(\gamma_3 + \alpha)}{\overline{M2}} & si\ t \geq 1990:1 \\ \frac{\gamma_3}{\overline{M2}} & si\ t < 1990:1 \end{cases}$$

Entonces para comprobar la hipótesis bajo estudio tenemos que:

$$H_0 : \alpha \leq 0$$
$$H_A : \alpha > 0$$

Y para esto únicamente necesitamos estimar el modelo con la variable dummy y emplear el t-calculado y el de la tabla (o su correspondiente p-value) para tomar la decisión.

(El siguiente análisis no se esperaba) Ahora bien si se quería ser más preciso se tiene que

$$\eta = \begin{cases} \frac{(\gamma_3 + \alpha)}{\overline{M2}_{\forall t \geq 1990:1}} & si\ t \geq 1990:1 \\ \frac{\gamma_3}{\overline{M2}_{\forall t < 1990:1}} & si\ t < 1990:1 \end{cases}$$

Entonces para comprobar la hipótesis bajo estudio tendríamos que la hipótesis alterna sería:

$$\frac{(\gamma_3+\alpha)}{\overline{M2}_{\forall t\geq 1990:1}}-\frac{\gamma_3}{\overline{M2}_{\forall t<1990:1}}>0$$

$$\gamma_3\left[\frac{1}{\overline{M2}_{\forall t\geq 1990:1}}-\frac{1}{\overline{M2}_{\forall t<1990:1}}\right]+\frac{\alpha}{\overline{M2}_{\forall t\geq 1990:1}}>0$$

Dado que esta es un hipótesis que involucra dos coeficientes, y que no contamos con más herramientas, se debería emplear una prueba F, es decir:

$$H_0: R\beta = C$$

$$H_0:\begin{bmatrix}0 & 0 & 0 & \left[\frac{1}{\overline{M2}_{\forall t\geq 1990:1}}-\frac{1}{\overline{M2}_{\forall t<1990:1}}\right] & \frac{1}{\overline{M2}_{\forall t\geq 1990:1}}\end{bmatrix}\begin{bmatrix}\gamma_1\\ \gamma_2\\ \gamma_3\\ \alpha\end{bmatrix}=0$$

Versus la hipótesis alterna de que H_0 no es correcta. Noten que en caso de rechazar H_0 se tiene que las elasticidades serán diferentes, y entonces únicamente se podrá comparar cual de las dos elasticidades es mayor numéricamente.

Y para esto necesitamos estimar el modelo con la variable dummy, las medias para M2 en los dos periodos, y los demás elementos de la siguiente fórmula:

$$F_c=\frac{\left(c-R\hat{\beta}\right)^T\left(R\left(X^TX\right)^{-1}R^T\right)^{-1}\left(c-R\hat{\beta}\right)/r}{SSE/n-k}$$

4 (30 puntos)

Una empresa embotelladora de té helado a nivel nacional desea estimar la demanda de su producto estrella en un departamento que ha venido experimentando una de las épocas más secas de su historia. De hecho en este departamento, solo ha llovido un día en el último año. Usted como consultor, cree que la cantidad vendida (y_t) del producto estrella (en 100,000 unidades) sigue la siguiente relación.

$$y_t=\beta_1+\beta_2\ln\left(X_{2t}\right)+\beta_3X_{3t}+\varepsilon_t \qquad t = 1, 2, ...$$

donde X_{2t} representa el número de días de lluvia en el periodo t y X_{3t} denota la tasa de desempleo (medido como (Desocupados/PEA)*100). (Los datos son anuales)

a) ¿Qué se debe cumplir, para obtener estimadores MELI (BLUE) para los parámetros β por el método de máxima verosimilitud? **(5 puntos)**

Se debe cumplir:

- Relación lineal entre la variable dependiente y los regresores.
- Los regresores deben ser no estocásticos y linealmente independientes entre si
- Los errores deben:
  - Seguir una distribución normal

- Tener media cero
- Varianza constante
- Y no estar autocorrelacionados

b) Después de realizar las transformaciones y operaciones aritméticas del caso, se obtienen las siguientes matrices que corresponden al equivalente de la matriz $X^T X$ y $X^T y$:

$$X^T X = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 30 & 5 \\ 0 & 5 & 1 \end{bmatrix} \quad X^T y = \begin{bmatrix} 10 \\ 5 \\ 2 \end{bmatrix}$$

Explique claramente a que corresponde cada uno de los elementos de estas dos matrices. (Por ejemplo, explique a partir de que sumatoria sale el 1 que corresponde al último elemento de la matriz $X^T X$, y así sucesivamente con cada elemento de las dos matrices) **(5 puntos – medio punto cada uno)**

En este caso tenemos: $n = 10$, $\sum_{t=1}^{n} \ln(X_{2t}) = \sum_{t=1}^{n} X_{3t} = 0$, $\sum_{t=1}^{n} \ln(X_{2t}) X_{3t} = 5$, $\sum_{t=1}^{n} (\ln(X_{2t}))^2 = 30$ $\sum_{t=1}^{n} (X_{3t})^2 = 1$

c) Encuentre los estimadores de ***máxima verosimilitud*** de los betas del modelo **(8 Puntos).**

En este caso tenemos que:

$$\beta_{hat} = \left(X^T X\right)^{-1} \cdot X^T y$$

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T y = \begin{bmatrix} 1/10 & 0 & 0 \\ 0 & 1/5 & -1 \\ 0 & -1 & 6 \end{bmatrix} \begin{bmatrix} 10 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix}$$

d) Interprete el significado de cada uno de los coeficientes estimados.**(6 Puntos – 2 puntos cada uno)**

$\hat{\beta}_2 = -1$

Un aumento del uno por ciento en el número de días festivos lluviosos disminuirá la demanda de unidades del producto estrella en 1000.0 unidades (1/100 * 100,000 unidades)

$\hat{\beta}_3 = 7$ Un aumento de un punto porcentual en la tasa de desempleo generará un aumento 7000 unidades del producto estrella.

$\hat{\beta}_1 = 1$. En este caso, dadas las características del problema este intercepto si tiene sentido. La parte de la demanda del producto estrella que no es explicado por la tasa de desempleo cuando hay solo un día lluvioso al año es de 100,000 unidades.

e) La empresa también desea determinar si los días de lluvia o la tasa de desempleo pesan o no lo mismo al momento de explicar la demanda de este producto. Explique como podría comprobar esta idea y muestre la mayor cantidad de información, así como las fórmulas que emplearía para tomar la decisión.**(6 Puntos)**

Como se desea comparar el peso que tiene cada variable para explicar la demanda del producto, será necesario emplear los coeficientes estandarizados. Es decir debemos determinar si:

$$\beta_2^E = \beta_3^E$$

$$\beta_2 \frac{s_{X_2}}{s_y} = \beta_3 \frac{s_{X_3}}{s_y}$$

$$\beta_2 \frac{s_{X_2}}{s_y} - \beta_3 \frac{s_{X_3}}{s_y} = 0$$

$$\begin{bmatrix} 0 & \frac{s_{X_2}}{s_y} & -\frac{s_{X_3}}{s_y} \end{bmatrix} \beta = 0$$

$$\frac{1}{s_y}\begin{bmatrix} 0 & s_{X_2} & -s_{X_3} \end{bmatrix} \beta = 0$$

$$\begin{bmatrix} 0 & s_{X_2} & -s_{X_3} \end{bmatrix} \beta = 0$$

$$\begin{bmatrix} 0 & \sqrt{20} & -\sqrt{12} \end{bmatrix} \beta = 0$$

$$\begin{bmatrix} 0 & \sqrt{4\cdot 5} & -\sqrt{3\cdot 4} \end{bmatrix} \beta = 0$$

$$\begin{bmatrix} 0 & \sqrt{5} & -\sqrt{4} \end{bmatrix} \beta = 0$$

Así, la hipótesis es de la forma $R\beta = C$. Donde $R = \begin{bmatrix} 0 & \sqrt{5} & -\sqrt{4} \end{bmatrix}$. (Nonten que dadas las caracterísicas de la matriz $X^T X$ tenemos que las medias de ambas variables es cero y por tanto sus varianzas corresponden a la suma cuadrada.

Así se debe emplear la formula $F_c = \dfrac{\left(c - R\hat{\beta}\right)^T \left(R\left(X^T X\right)^{-1} R^T\right)^{-1} \left(c - R\hat{\beta}\right) \big/ r}{S^2}$, con $c = 0$ y $r = 1$ para comprobar la hipótesis nula de que $R\beta = C$, versus la alterna no Ho. Se rechazará la hipótesis nula si el F calculado es mayor que el f de la tabla con 1 grado de libertad en el numerador y 7 grados de libertad en el denominador.

Resultados de EasyReg para la pregunta 3.

```
Characteristics:
M1
  First available observation = 1(=1964.1)
  Last available observation  = 169(=2006.1)
  First chosen observation = 2(=1964.2)
  Last chosen observation  = 169(=2006.1)
  Number of usable chosen observations: 168
Subsample characteristics:
  Sample mean:   1.9144605E+003

X variables:
X(1) = lnGDP
X(2) =ln(TB3mo)
X(3) = 1

Model:
Y = b(1)X(1) + b(2)X(2) + b(3)X(3)  + U,
where U is the error term, satisfying
E[U|X(1),X(2),X(3)] = 0.

OLS estimation results
Parameters     Estimate       t-value    H.C. t-value
                              (S.E.)     (H.C. S.E.)
                              [p-value]  [H.C. p-value]
b(1)           39.00536       XXX          12.42
                           (4.322863)     (3.139923)
                           [0.36690]      [0.21415]
b(2)        -3694.18782       -0.296       -0.355
                         (12480.62673)  (10395.61254)
                           [0.76723]      [0.72232]
b(3)         1714.31992        5.372        5.903
                           (319.10549)    (290.41237)
                           [0.00000]      [0.00000]

Effective sample size (n):                      168
Variance of the residuals:                1965809.647603
Standard error of the residuals (SER):          1402.073339
Residual sum of squares (RSS):           324358591.85452
(Also called SSR = Sum of Squared Residuals)
Total sum of squares (TSS):               1326511691
R-square:                                 XXX

Overall F test: F(2,165) = XXX
p-value = 0.57936
Significance levels:      5%
Critical values:        3.05
```

Conclusions:	XXX