Econometría 06216
Examen Parcial #1
Respuestas Sugeridas
Cali, Sábado 03 de Septiembre de 2011

Profesores: Julio César Alonso - Carlos Giovanni González

Estudiante: ______________________
Código: __________________________

Instrucciones:

1. Lea cuidadosamente todas las preguntas e instrucciones.
2. Este examen consta de **6 páginas**; además, deben tener 2 páginas de fórmulas.
3. El examen consta de 4 preguntas que suman un total de 100 puntos. El valor de cada una de las preguntas está 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. No se dará crédito por respuestas consignadas en las hojas de preguntas.
5. El examen está diseñado para dos horas, 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. No se aceptarán reclamos de exámenes no escritos a lapicero (escritos a lápiz).
9. Al finalizar su examen entregue sus hojas de respuesta, así como las hojas de preguntas.
10. ¡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. Consigne su respuesta en la hoja de respuestas suministrada.).

1.1. Si el producto de dos matrices *A*B* es conformable, entonces $A+B$ existirá. Es más, $A+B=B+A$.

1.2. Al comparar dos estimadores lineales consistentes de α (un parámetro que representa una pendiente de un modelo lineal), por ejemplo $\hat{\alpha}$ y $\tilde{\alpha}$, se encuentra que $\hat{\alpha}$ tiene mayor varianza que $\tilde{\alpha}$. Un estudiante del curso de econometría afirma que: "$\hat{\alpha}$ no puede ser el estimador de máxima verosimilitud de α" ¿Es esta afirmación cierta?

1.3. Suponga que se cuenta con el siguiente modelo:

$$\frac{w_i}{y_i} = \left(sen^2(\delta) + \cos^2(\delta)\right)^{\alpha} \cdot y_i \cdot X_i^{3\beta_1} \cdot S_i^{\beta_3} \cdot Z_i^{\beta_4} \cdot Q_i^{\alpha+\beta_3} \varepsilon_i^{\beta_4}$$

y se desea encontrar todos los parámetros representados por betas. Entonces, un investigador afirma que "El método de mínimos cuadrados ordinarios es una buena herramienta para encontrar estimadores insesgados de los betas del modelo". ¿Es ésta afirmación verdadera o falsa?

1.4. Al estimar un modelo para la demanda de "Pispirispis" en la ciudad de Cali, se obtuvo la siguiente Tabla ANOVA. Pero lastimosamente, el perro del estudiante que estaba al frente del estudio, se comió las partes más "sustanciosas" de la Tabla. Las partes que se perdieron de la correspondiente tabla fueron remplazadas con "XXX". Un compañero del estudiante afirma: "Fresco, dado que el modelo tiene intercepto, es obvio que ese modelo no era bueno para explicar la demanda de Pispirispis" .. ¿Es esta afirmación falsa o verdadera?

Fuente de variación	*SS*	*G de L*	*MS*
Regresión	XXX	XXX	250
Error	100	XXX	XXX
TOTAL	2600	110	

1.5. Otro estudiante al momento de prepararse para su parcial encontró una matriz $X^T X$, pero no encontró el modelo al cual pertenecía esta muestra, pero si encontró en el enunciado información que le permitía saber con certeza que el modelo tiene un intercepto.

A. $X^T X = \begin{bmatrix} 10 & 5 & 20 & 40 \\ 5 & 5 & 15 & 35 \\ 20 & 15 & 20 & 25 \\ 40 & 35 & 25 & 30 \end{bmatrix}$

El estudiante le afirma a su amigo, que no tenía mucho tiempo para estudiar, que "la anterior matriz $X^T X$ corresponde a un modelo estimado con 10 observaciones y que contiene con seguridad una variable dummy que afecta el intercepto del modelo." ¿Es esta afirmación verdadera o falsa?

2. Selección Múltiple (15 puntos en total, 5 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. Consigne su respuesta en la hoja de respuestas suministrada.).

2.1.¿Cuál de las siguientes opciones corresponde a las propiedades del estimador MCO?:

a. Las propiedades serian sesgado y con la mínima varianza posible cuando se le compara con los otros posibles estimadores lineales.
b. Las propiedades serian i. Relación lineal entre Y y X; ii. X's no estocásticas y linealmente independientes entre sí, y iii. El error tiene media cero, varianza constante y no esta auto-correlacionado.
c. Las propiedades serian insesgadez y eficiencia.
d. Las propiedades serian consistencia y eficiencia.
e. Ninguna de las anteriores.

2.2.Un economista está interesado en estimar un modelo que tienen todas sus variables expresadas en logaritmos. Donde, (q_i) es la cantidad demandada del bien, (P_i) es el precio del bien y (R_i) es la renta. El modelo estimado finalmente por el economista fue el siguiente:

$$\ln(q_i) = \beta_1 + \beta_2 \ln(P_i) + \beta_3 \ln(R_i) + \varepsilon_i$$

Otro economista, está interesado en estimar el mismo modelo con las variables en logaritmos y además decide incluir dos variables dummy. Donde D_{1i} es igual a 1 si el individuo es mujer y 0 en otro caso; y D_{2i} es igual a 1 si el individuo es trabajador por cuenta propia y 0 en otro caso.

$$ln(q_i) = \beta_1 + \beta_2 D_{1i} + \beta_3 ln(P_i) + \beta_4 \ln(R_i) + \beta_5 D_{2i} + \varepsilon_i$$

Teniendo en cuenta que el interés del primer economista es que todas las variables estén expresadas en logaritmos. ¿Cuál es la razón por la cual el segundo economista no le calcula el logaritmo a todas las variables?

a. El logaritmo sólo se calculará si la teoría económica lo indica así.
b. El logaritmo sólo se calculará si la forma funcional del modelo inherentemente lineal lo indica así.
c. El logaritmo sólo se calculará dependiendo del interés del investigador.
d. No es necesario calcularlo ya que el ln(1)= 0 y ln(0)= 1 y por tanto da lo mismo sacar el logaritmo o no.
e. Ninguna de las anteriores.

2.3.Dos economistas están participando en una convocatoria de CONCIENCIAS para financiar un proyecto de investigación sobre la estimación de la ecuación de salarios para la ciudad de Cali. Uno de ellos, el economista más reconocido por su trayectoria profesional decide escribir el siguiente modelo:

$$ln(W_i) = \beta_1 + \beta_2 Educ_i + \beta_3 Exp_i + \alpha_1 D_{1i} + \varepsilon_i$$

$$D_{1i} = \begin{cases} 1 = \text{si el individuo i es mujer} \\ 0 = \quad o.w. \end{cases}$$

Por otro lado, el otro economista quien es menos reconocido decide escribir un modelo alternativo para estimar la misma ecuación de salarios.

$$ln(W_i) = \beta_1 + \beta_2 Educ_i + \beta_3 Exp_i + \alpha_1 D_{1i} + \varepsilon_i$$

$$D_{1i} = \begin{cases} -3 = \text{si el individuo i es mujer} \\ 0 = \quad o.w. \end{cases}$$

El jurado de la convocatoria de CONCIENCIAS debe decidir a quién otorgarle los recursos para la investigación. Pero antes debe tomar una decisión sobre los dos modelos que proponen los economistas. Así, si el jurado compara los dos modelos podría decir que:

a. En este caso los dos modelos tienen el mismo significado.
b. En este caso los dos modelos tienen un significado diferente.
c. En este caso el modelo del economista más reconocido es más consistente.
d. En este caso el modelo del economista más reconocido es el más parsimonioso.

e. Ninguna de las anteriores.

3. (30 puntos)

El jefe de estudios económicos de la compañía de telecomunicaciones de la ciudad desea estimar un modelo que describa el gasto en telefonía celular de los individuos. Uno de los investigadores de la oficina de estudios económicos de la compañía le recomienda al jefe de estudios que en caso de disponer de toda la información y según la teoría económica el modelo debería incluir las siguientes variables:

La variable de interés (o variable dependiente) para el análisis debe ser el gasto en telefonía celular de los individuos (GR_i) (medido en dólares por mes); y algunas de las variables independientes que se deberían incluir serian: una variable dummy de género, que tomaría el valor de 1 si el individuo es mujer y 0 en otro caso (D_{1i}), el ingreso del individuo i $(ingr_i)$ (medido en miles de dólares por mes) y la edad en años del individuo i $(edad_i)$.

a. Escriba el modelo estimado por el econometrista que corresponde a la estimación presentadas en la Tabla 1. **(5 Puntos).**
b. Interprete los coeficientes del modelo estimado teniendo en cuenta su significancia. Los resultados se presentan en la Tabla 1. **(7 puntos).**
c. Encuentre los valores que fueron reemplazados por "XXX". No es necesario efectuar el cálculo, pero sí mostrar con qué cantidades se puede encontrar dicho número y explicar en una línea de que se trata. **(6 puntos en total, 2 puntos cada uno).**
d. ¿Qué tan bueno es el modelo estimado? **(3 puntos en total)**
e. El jefe de estudios económicos cree que el gasto en telefonía celular de las personas mayores de 25 años tienen una elasticidad ingreso mayor que las personas menores de esta edad. En especial, cree que la elasticidad ingreso del gasto en telefonía celular es tres veces más grande para los mayores de 25 que para los menores de 25. Por otra parte, uno de los investigadores de la oficina de estudios económicos afirma que este tipo de "afirmaciones" es imposible de contrastar usando la econometría.

   Partiendo del modelo estimado en la Tabla 1., explique brevemente si la afirmación del jefe se puede contrastar usando la econometría. Además, plantee un modelo que permita comprobar si esta hipótesis es cierta. Explique claramente por qué el modelo planteado sirve para este fin. Sea lo más claro posible, y muestre que fórmulas emplearía, que valores debe remplazar en las fórmulas, así como la manera en que tomaría la decisión. **(9 Puntos)**

4. (30 puntos)

Una empresa en su proceso de desarrollo de nuevos productos y servicios desea conocer la disponibilidad de pago de sus clientes por un nuevo servicio de cuidado de plantas a

domicilio (y_i medido en millones de pesos) por medio de un modelo de regresión simple que emplea los años de educación (X_{1i} medido en años), la proporción de años que faltan para su retiro (X_{3i} medido como la razón entre los años que faltan para el retiro y los años totales que trabajará el individuo multiplicado por cien) y el precio mensual por el cuidado de una planta (X_{2i} medido en miles de pesos) como variables explicativas. En especial el modelo será :

$$y_i = \beta_1 + \beta_2 \ln\left(X_{1i}\right) + \beta_3 \frac{1}{X_{2i}} + \beta_4 X_{3i}^2 + \varepsilon_i \qquad (1)$$

donde ε_i representa una termino aleatorio de error.

Después de realizar 100 encuestas, obtiene los siguientes resultados:

$$X^T X = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 20 & -20 & 0 \\ 0 & -20 & 40 & 40 \\ 0 & 0 & 40 & 60 \end{bmatrix} \qquad X^T y = \begin{bmatrix} 60 \\ -15 \\ 10 \\ 5 \end{bmatrix}$$

a) ¿Qué características estadísticas tienen los estimadores MCO si se cumplen los supuestos del Teorema de Gauss-Markov **(4 puntos)**
b) Encuentre (si es posible) los mejores valores estimados para las siguientes cantidades **(8 puntos)**
   - Valor esperado del precio del precio mensual por el cuidado de una planta **(2 puntos)**
   - Valor esperado de la disponibilidad de pago de los clientes por el nuevo servicio de cuidado de plantas a domicilio **(2 puntos)**
   - Valor esperado de los años de educación **(2 puntos)**
   - Valor esperado de la proporción de años que faltan para su retiro **(2 puntos)**

c) Calcule los mejores estimadores lineales insesgados para los parámetros del modelo (1) **(10 puntos)**
d) Explique el significado de los coeficientes estimados **(8 puntos)**
e) Un investigador cree que la elasticidad precio de la disponibilidad a pagar es diferente para aquellos clientes con ingresos superiores a dos millones de pesos mensuales. Es más, está casi seguro que es dos veces más que la elasticidad para aquellos cuyo ingreso es inferior o igual a dos millones de pesos. ¿cómo podría probar si el investigador tiene o no la razón. Explique claramente cómo lo haría, que fórmulas emplearía y que valores puede remplazar en dichas fórmulas y cómo tomaría la decisión. No es necesario que realice los cálculos. **(10 puntos)**

Tabla 1. Resultados de EasyReg.

Dependent variable:
Y = GR (i)

Characteristics:
GR (i)
First observation = 1
Last observation = 100
Number of usable observations: 100
X variables:
X(1) = ingr (i)
X(2) = ingr (i)^2
X(3) = edad(i)
X(4) = 1
Model:
Y = b(1)X(1) +.....+ **XXX** + U,
where U is the error term, satisfying
E[U|X(1),...,X(4)] = 0.

OLS estimation results

Parameters	Estimate	t-value (S.E.) [p-value]	H.C. t-value (H.C. S.E.) [H.C. p-value]
b(1)	0.9452	5.7316	1.638
		XXX	(0.15260)
		[0.00006]	[0.00027]
b(2)	-0.00006	-6	-0.677
		(0.00001)	(0.0612)
		[0.00014]	[0.25136]
b(3)	-0.12	-4.8701	-1.284
		(0.02464)	(0.0115)
		[0.00080]	[0.14711]
b(4)	100.4409	39.828	43.231
		(142.62431)	(131.39687)
		[0.00000]	[0.00000]

Effective sample size (n): 100
Variance of the residuals: 72318.51487238
Standard error of the residuals (SER): 268.92101977
Residual sum of squares (RSS): 6942577.42774893
(Also called SSR = Sum of Squared Residuals)
Total sum of squares (TSS): 16455716.1564733
R-square: 0.6781
Adjusted R-square: 0.6649
Overall F test: F(3,96) = 43.85
p-value = 0.00000

Significance levels:	10%	5%
Critical values:	2.14	2.7
Conclusions:	reject	**XXX**

Econometría 06216
Examen Parcial #1
Respuestas Sugeridas
Cali, Sábado 03 de Septiembre de 2011

Profesores: Julio César Alonso - Carlos Giovanni González

Estudiante: _______________________
Código: ___________________________

Instrucciones:

1. Lea cuidadosamente todas las preguntas e instrucciones.
2. Este examen consta de **6 páginas**; además, deben tener 2 páginas de fórmulas.
3. El examen consta de 4 preguntas que suman un total de 100 puntos. El valor de cada una de las preguntas está 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. No se dará crédito por respuestas consignadas en las hojas de preguntas.
5. El examen está diseñado para dos horas, 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. No se aceptarán reclamos de exámenes no escritos a lapicero (escritos a lápiz).
9. Al finalizar su examen entregue sus hojas de respuesta, así como las hojas de preguntas.
10. ¡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. Consigne su respuesta en la hoja de respuestas suministrada.).

1.1. Si el producto de dos matrices $A*B$ es conformable, entonces $A+B$ existirá. Es más, $A+B=B+A$.

Falso. Un contra ejemplo basta para demostrar que esta afirmación es falsa. Por ejemplo supongamos que A es nxm y B es $m\times m$ y $n\neq m$, AB existe, pero $A+B$ no existe

1.2. Al comparar dos estimadores lineales consistentes de α (un parámetro que representa una pendiente de un modelo lineal), por ejemplo $\hat{\alpha}$ y $\tilde{\alpha}$, se encuentra que $\hat{\alpha}$ tiene mayor varianza que $\tilde{\alpha}$. Un estudiante del curso de econometría afirma que: "$\hat{\alpha}$ no puede ser el estimador de máxima verosimilitud de α" ¿Es esta afirmación cierta?

Falso, al encontrar que $\hat{\alpha}$ es consistente, este podría ser insesgado o sesgado, lo mismo ocurre para $\hat{\alpha}$. Y al este tener $\hat{\alpha}$ una varianza más grande que $\tilde{\alpha}$, podría ser que el primero fuese insesgado y el segundo sesgado, y por tanto $\hat{\alpha}$ aún puede ser el estimador MELI, el cual corresponde también al estimador de Máxima verosimilitud para las pendientes.

1.3. Suponga que se cuenta con el siguiente modelo:

$$\frac{w_i}{y_i}=\left(sen^2\left(\delta\right)+\cos^2\left(\delta\right)\right)^{\alpha}\cdot y_i\cdot X_i^{3\beta_1}\cdot S_i^{\beta_3}\cdot Z_i^{\beta_4}\cdot Q_i^{\alpha+\beta_3}\varepsilon_i^{\beta_4}$$

y se desea encontrar todos los parámetros representados por betas. Entonces, un investigador afirma que "El método de mínimos cuadrados ordinarios es una buena herramienta para encontrar estimadores insesgados de los betas del modelo". ¿Es ésta afirmación verdadera o falsa?

Falso,si bien el modelo podría reparametrizarse como sigue:

Falso,si bien el modelo podría reparametrizarse como sigue:

$$\frac{w_i}{y_i} = \left(sen^2(\delta) + \cos^2(\delta)\right)^{\alpha} \cdot y_i \cdot X_i^{3\beta_1} \cdot S_i^{\beta_3} \cdot Z_i^{\beta_4} \cdot Q_i^{\alpha+\beta_3} \varepsilon_i^{\beta_4}$$
$$\frac{w_i}{y_i^2} = X_i^{3\beta_1} \cdot S_i^{\beta_3} \cdot Z_i^{\beta_4} \cdot Q_i^{\alpha+\beta_3} \varepsilon_i^{\beta_4}$$
$$\ln\left(\frac{w_i}{y_i^2}\right) = \ln\left(X_i^{3\beta_1} \cdot S_i^{\beta_3} \cdot Z_i^{\beta_4} \cdot Q_i^{\alpha+\beta_3} \varepsilon_i^{\beta_4}\right)$$
$$\ln(w_i) - 2\ln(y_i) = 3\beta_1 \ln(X_i) + \beta_3 \ln(S_i) + \beta_4 \ln(Z_i) + (\alpha + \beta_3)\ln(Q_i) + \beta_4 \ln(\varepsilon_i)$$
$$z_i = \gamma_1 \ln(X_i) + \beta_3 \ln(S_i) + \beta_4 \ln(Z_i) + \gamma_2 \ln(Q_i) + \mu_i$$
donde
$$z_i = \ln(w_i) - 2\ln(y_i)$$
$$\gamma_1 = 3\beta_1$$
$$\gamma_2 = \alpha + \beta_3$$
$$\mu_i = \beta_4 \ln(\varepsilon_i)$$

El modelo reparametrizado permite encontrar todos los betas, pero no tiene intercepto y por tanto no garantiza que el supuesto que el error tenga media se cumpla y por tanto los estimadores MCO no necesariamente serán insesgados.

1.4. Al estimar un modelo para la demanda de "Pispirispis" en la ciudad de Cali, se obtuvo la siguiente Tabla ANOVA. Pero lastimosamente, el perro del estudiante que estaba al frente del estudio, se comió las partes más "sustanciosas" de la Tabla. Las partes que se perdieron de la correspondiente tabla fueron remplazadas con "XXX". Un compañero del estudiante afirma: "Fresco, dado que el modelo tiene intercepto, es obvio que ese modelo no era bueno para explicar la demanda de Pispirispis" .. ¿Es esta afirmación falsa o verdadera?

Fuente de variación	*SS*	*G de L*	*MS*
Regresión	XXX	XXX	250
Error	100	XXX	XXX
TOTAL	2600	110	

Falso, la tabla ANOVA permite comprobar la hipótesis conjunta de significancia global del modelo. De hecho la tabla permite el rechazo de la hipótesis nula, a cualquiera de los niveles de significación "habituales". Por otro lado, el R cuadrado del modelo es relativamente grande (25/26). La tabla correcta debe ser:

Fuente de variación	*SS*	*G de L*	*MS*	*F*
Regresión	2500	10	250	250
Error	100	100	1	
TOTAL	2600	110		

1.5. Otro estudiante al momento de prepararse para su parcial encontró una matriz $X^T X$, pero no encontró el modelo al cual pertenecía esta muestra, pero si encontró en el enunciado información que le permitía saber con certeza que el modelo tiene un intercepto.

A. $X^T X = \begin{bmatrix} 10 & 5 & 20 & 40 \\ 5 & 5 & 15 & 35 \\ 20 & 15 & 20 & 25 \\ 40 & 35 & 25 & 30 \end{bmatrix}$

El estudiante le afirma a su amigo, que no tenía mucho tiempo para estudiar, que "la anterior matriz $X^T X$ corresponde a un modelo estimado con 10 observaciones y que contiene con seguridad una variable dummy que afecta el intercepto del modelo." ¿Es esta afirmación verdadera o falsa?

Verdadero. Noten que si el modelo fuese $y_i = \beta_1 + \beta_2 D_i + \beta_3 X_{1i} + \beta_4 X_{2i} + \varepsilon_i$*, tiene que ser cierto que el segundo elemento de la primera fila debe ser una suma de unos y ceros y por tanto no puede ser mayor que 10. En este caso ocurre esto. Por otro lado, el segundo elemento de la diagonal corresponde a la suma de unos y ceros elevados al cuadrado que sería igual a la suma de esos unos y ceros. Esto ocurre en esta situación. De hecho, sólo es posible que 10 números sumados sean iguales a la suma de los mismos 10 números elevados al cuadrado si estos números corresponden a ceros y unos.*

2. Selección Múltiple (15 puntos en total, 5 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. Consigne su respuesta en la hoja de respuestas suministrada.).

2.1. ¿Cuál de las siguientes opciones corresponde a las propiedades del estimador MCO?:

a. Las propiedades serian sesgado y con la mínima varianza posible cuando se le compara con los otros posibles estimadores lineales.

b. Las propiedades serian i. Relación lineal entre Y y X; ii. X's no estocásticas y linealmente independientes entre sí, y iii. El error tiene media cero, varianza constante y no esta auto-correlacionado.
c. Las propiedades serian insesgadez y eficiencia.
d. Las propiedades serian consistencia y eficiencia.
e. Ninguna de las anteriores.

Respuesta sugerida: e. Noten que si no se cumplen las condiciones del Teorema de Gauss-Markov, entonces la opción c no es válida. Por eso la mejor respuesta es la e.

2.2. Un economista está interesado en estimar un modelo que tienen todas sus variables expresadas en logaritmos. Donde, (q_i) es la cantidad demandada del bien, (P_i) es el precio del bien y (R_i) es la renta. El modelo estimado finalmente por el economista fue el siguiente:

$$\ln(q_i) = \beta_1 + \beta_2 \ln(P_i) + \beta_3 \ln(R_i) + \varepsilon_i$$

Otro economista, está interesado en estimar el mismo modelo con las variables en logaritmos y además decide incluir dos variables dummy. Donde D_{1i} es igual a 1 si el individuo es mujer y 0 en otro caso; y D_{2i} es igual a 1 si el individuo es trabajador por cuenta propia y 0 en otro caso.

$$ln(q_i) = \beta_1 + \beta_2 D_{1i} + \beta_3 ln(P_i) + \beta_4 \ln(R_i) + \beta_5 D_{2i} + \varepsilon_i$$

Teniendo en cuenta que el interés del primer economista es que todas las variables estén expresadas en logaritmos. ¿Cuál es la razón por la cual el segundo economista no le calcula el logaritmo a todas las variables?

a. El logaritmo sólo se calculará si la teoría económica lo indica así.
b. El logaritmo sólo se calculará si la forma funcional del modelo inherentemente lineal lo indica así.
c. El logaritmo sólo se calculará dependiendo del interés del investigador.
d. No es necesario calcularlo ya que el ln(1)= 0 y ln(0)= 1 y por tanto da lo mismo sacar el logaritmo o no.
e. Ninguna de las anteriores.

Respuesta sugerida e.

2.3. Dos economistas están participando en una convocatoria de CONCIENCIAS para financiar un proyecto de investigación sobre la estimación de la ecuación de salarios

para la ciudad de Cali. Uno de ellos, el economista más reconocido por su trayectoria profesional decide escribir el siguiente modelo:

$$ln(W_i) = \beta_1 + \beta_2 Educ_i + \beta_3 Exp_i + \alpha_1 D_{1i} + \varepsilon_i$$

$$D_{1i} = \begin{cases} 1 = \text{si el individuo i es mujer} \\ 0 = \quad o.w. \end{cases}$$

Por otro lado, el otro economista quien es menos reconocido decide escribir un modelo alternativo para estimar la misma ecuación de salarios.

$$ln(W_i) = \beta_1 + \beta_2 Educ_i + \beta_3 Exp_i + \alpha_1 D_{1i} + \varepsilon_i$$

$$D_{1i} = \begin{cases} -3 = \text{si el individuo i es mujer} \\ 0 = \quad o.w. \end{cases}$$

El jurado de la convocatoria de CONCIENCIAS debe decidir a quién otorgarle los recursos para la investigación. Pero antes debe tomar una decisión sobre los dos modelos que proponen los economistas. Así, si el jurado compara los dos modelos podría decir que:

a. En este caso los dos modelos tienen el mismo significado.
b. En este caso los dos modelos tienen un significado diferente.
c. En este caso el modelo del economista más reconocido es más consistente.
d. En este caso el modelo del economista más reconocido es el más parsimonioso.
e. Ninguna de las anteriores.

Respuesta sugerida a.

3. (30 puntos)

El jefe de estudios económicos de la compañía de telecomunicaciones de la ciudad desea estimar un modelo que describa el gasto en telefonía celular de los individuos. Uno de los investigadores de la oficina de estudios económicos de la compañía le recomienda al jefe de estudios que en caso de disponer de toda la información y según la teoría económica el modelo debería incluir las siguientes variables:

La variable de interés (o variable dependiente) para el análisis debe ser el gasto en telefonía celular de los individuos (GR_i) (medido en dólares por mes); y algunas de las variables independientes que se deberían incluir serian: una variable dummy de género, que tomaría el valor de 1 si el individuo es mujer y 0 en otro caso (D_{1i}), el ingreso del individuo i $\left(ingr_i\right)$ (medido en miles de dólares por mes) y la edad en años del individuo i $\left(edad_i\right)$.

a. Escriba el modelo estimado por el econometrista que corresponde a la estimación presentadas en la Tabla 1. **(5 Puntos).**

El modelo estimado en la Tabla 1 fue:

$$GR_i = \beta_1 + \beta_2\left(ingr_i\right) + \beta_3\left(ingr_i^2\right) + \beta_4 edad_i + \varepsilon_i$$

i=1,2,...100.

b. Interprete los coeficientes del modelo estimado teniendo en cuenta su significancia. Los resultados se presentan en la Tabla 1. **(7 puntos).**

$\hat{\beta}_1$=100.4409 Es significativo con un nivel de confianza del 99%. 100 dólares corresponde a la parte del gasto mensual en telefonía celular que no depende ni de la edad ni del ingreso. (1 punto)

$\hat{\beta}_4$=-0.12. Es significativo al 99%. Ante un aumento en un año en la edad del individuo se espera que en promedio el gasto mensual en telefonía celular disminuya en 12 centavos de dólar. (2 punto)

Ahora, noten que:

$$\frac{\partial GR_i}{\partial ingr_i} = \beta_2 + 2\beta_3\left(ingr_i\right)$$

Dado que ambos coeficientes son significativos (el primero con un nivel de confianza del 99% y el segundo con un nivel de confianza del 99%), tenemos las siguientes interpretaciones:

$\hat{\beta}_2$= 0.9452. La parte constante del efecto marginal de un aumento de mil dólares en el ingreso es de 0.94 dólares. (2 puntos)

$\hat{\beta}_3$ = -0.00006. Ante un aumento de mil dólares en el ingreso se espera que en promedio el efecto marginal de esos mil dólares sobre el gasto en telefonía celular disminuya en 2*(0.00006)($ingr_i$) dólares. (2 puntos)

c. Encuentre los valores que fueron reemplazados por "XXX". No es necesario efectuar el cálculo, pero sí mostrar con qué cantidades se puede encontrar dicho número y explicar en una línea de que se trata. **(6 puntos en total, 2 puntos cada uno).**

(1)XXX= b(4)X(4)
Es el intercepto del modelo.

(2)XXX= reject

Es la decisión que reporta el EasyReg de la prueba F:

(3)XXX= 0.1649
Es el error estándar asociado al coeficiente $\hat{\boldsymbol{\beta}}_{\mathbf{2}}$.

$$ee_{\hat{\beta}_2} = \frac{\hat{\beta}_2}{t} = \frac{0.9452}{5.7316} = 0.1649$$

d. ¿Qué tan bueno es el modelo estimado? **(3 puntos en total)**

Para probar que tan bueno es el modelo se plantean las siguientes hipótesis a contrastar $H_0: \beta_2 = \beta_3 = \beta_4 = 0$ vs $H_a: No\ H_0$. Esto se puede comprobar mediante una prueba F, donde el valor-p=0.0000, dado que éste es menor a un nivel de significancia de 1%, entonces se rechaza la hipótesis, por lo que el modelo es significativo en su conjunto. (1 punto).

Ahora bien, R^2el indica que el 67,81% de la variabilidad del gasto mensual en telefonía celular es explicada por el modelo. (1 punto) Por tanto se puede concluir que el modelo presenta un buen ajuste. . (1 punto)

e. El jefe de estudios económicos cree que el gasto en telefonía celular de las personas mayores de 25 años tienen una elasticidad ingreso mayor que las personas menores de esta edad. En especial, cree que la elasticidad ingreso del gasto en telefonía celular es tres veces más grande para los mayores de 25 que para los menores de 25. Por otra parte, uno de los investigadores de la oficina de estudios económicos afirma que este tipo de "afirmaciones" es imposible de contrastar usando la econometría.

   Partiendo del modelo estimado en la Tabla 1., explique brevemente si la afirmación del jefe se puede contrastar usando la econometría. Además, plantee un modelo que permita comprobar si esta hipótesis es cierta. Explique claramente por qué el modelo planteado sirve para este fin. Sea lo más claro posible, y muestre que fórmulas emplearía, que valores debe remplazar en las fórmulas, así como la manera en que tomaría la decisión. **(9 Puntos)**

Incluyendo variables dummy en el modelo de regresión lineal múltiple que interactúen con las pendientes (cambio estructural) es posible probar hipótesis de este estilo usando restricciones del tipo RB=c.

En este caso es necesario crear la siguiente variable dummy:

$$D_i = \begin{cases} 1 & si\, i > 25\, años \\ 0 & o.w. \end{cases}$$

(1 punto por plantear la variable dummy)

Entonces, el modelo será:

$$GR_i = \beta_1 + \beta_2 (ingr_i) + \beta_3 (ingr_i^2) + \beta_4 edad_i + \alpha_1 D_i (ingr_i) + \alpha_2 D_i (ingr_i^2) + \varepsilon_i$$

(2 puntos por plantear el modelo)

En este caso el modelo permitirá comprobar la hipótesis pues tenemos que la elasticidad ingreso (evaluada en la media) del gasto en recreación (η) es :

$$\eta = \frac{\partial GR}{\partial ingr} \cdot \frac{\overline{GR}}{\overline{ingr}}$$

En este caso tendremos que:

$$E[GR_i] = \begin{cases} \beta_1 + (\beta_2 + \alpha_1)(ingr_i) + (\beta_3 + \alpha_2)(ingr_i^2) + \beta_4 edad_i & si\ i > 25\ años \\ \beta_1 + \beta_2 (ingr_i) + \beta_3 (ingr_i^2) + \beta_4 edad_i & o.w. \end{cases}$$

Por tanto nuestra hipótesis nula será:

$$Ho: Elasticidad(mayores\ de\ 25\ años) = 3 \cdot Elasticidad(menores\ de\ 25\ años)$$

$$Ho: \left.\frac{\partial GR}{\partial ingr}\right|_{mayores\ de\ 25\ años} \cdot \left.\frac{\overline{GR}}{\overline{ingr}}\right|_{mayores\ de\ 25\ años} = 3 \cdot \left[\left.\frac{\partial GR}{\partial ingr}\right|_{menores\ de\ 25\ años} \cdot \left.\frac{\overline{GR}}{\overline{ingr}}\right|_{menores\ de\ 25\ años} \right]$$

$$Ho: \left((\beta_2 + \alpha_1) + (\beta_3 + \alpha_2) \cdot \overline{ingr}_{>25años}\right) \cdot \frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} = 3 \cdot \left[\left(\beta_2 + \beta_3 \cdot \overline{ingr}_{<25años}\right) \cdot \frac{\overline{GR}_{<25años}}{\overline{ingr}_{<25años}} \right]$$

$$Ho: \beta_2 \left[\frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} - 3 \frac{\overline{GR}_{<25años}}{\overline{ingr}_{<25años}} \right] + \beta_3 \left[\overline{GR}_{>25años} - 3\overline{GR}_{<25años} \right] + \alpha_1 \frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} + \alpha_2 = 0$$

(3 puntos por plantear la hipótesis)

Esta hipótesis se puede escribir de la forma $R\beta = C$ donde:

$$R = \begin{bmatrix} 0 & \left[\frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} - 3 \frac{\overline{GR}_{<25años}}{\overline{ingr}_{<25años}} \right] & \left[\overline{GR}_{>25años} - 3\overline{GR}_{<25años} \right] & 0 & \frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} & \overline{GR}_{>25años} \end{bmatrix} \text{ y}$$

$$C = 0$$

Y por tanto la hipótesis alterna será: $H_A:\ no\ H_0$

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

En este caso el denominador será: 100-5=95. El numerador es:

$$\left(-\left[\hat{\beta}_2 \left[\frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} - 3 \frac{\overline{GR}_{<25años}}{\overline{ingr}_{<25años}} \right] + \hat{\beta}_3 \left[\overline{GR}_{>25años} - 3\overline{GR}_{<25años} \right] + \hat{\alpha}_1 \frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} + \hat{\alpha}_2 \right]\right) \bullet$$

$$\left\{ \begin{bmatrix} 0 & \left[\frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} - 3 \frac{\overline{GR}_{<25años}}{\overline{ingr}_{<25años}} \right] & \left[\overline{GR}_{>25años} - 3\overline{GR}_{<25años} \right] & 0 & \frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} & \overline{GR}_{>25años} \end{bmatrix} \left(X^T X\right)^{-1} \begin{bmatrix} 0 \\ \frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} - 3 \frac{\overline{GR}_{<25años}}{\overline{ingr}_{<25años}} \\ \overline{GR}_{>25años} - 3\overline{GR}_{<25años} \\ 0 \\ \frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} \\ \overline{GR}_{>25años} \end{bmatrix} \right\} \bullet$$

$$\left(-\left[\hat{\beta}_2 \left[\frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} - 3 \frac{\overline{GR}_{<25años}}{\overline{ingr}_{<25años}} \right] + \hat{\beta}_3 \left[\overline{GR}_{>25años} - 3\overline{GR}_{<25años} \right] + \hat{\alpha}_1 \frac{\overline{GR}_{>25años}}{\overline{ingr}_{>25años}} + \hat{\alpha}_2 \right]\right)$$

(2 puntos)

Este estadístico se debe comparar con el F de la tabla con 1 grado de libertad en el numerado y 95 grados de libertad en el denominador y con un nivel de significancia del 1%, 5% o 10% por ciento. Se rechazará si el $F_{calculado}$ es mayor que el de la tabla. (1 punto)

4. (30 puntos)

Una empresa en su proceso de desarrollo de nuevos productos y servicios desea conocer la disponibilidad de pago de sus clientes por un nuevo servicio de cuidado de plantas a domicilio (y_i medido en millones de pesos) por medio de un modelo de regresión simple que emplea los años de educación (X_{1i} medido en años), la proporción de años que faltan para su retiro (X_{3i} medido como la razón entre los años que faltan para el retiro y los años totales que trabajará el individuo multiplicado por cien) y el precio mensual por el cuidado de una planta (X_{2i} medido en miles de pesos) como variables explicativas. En especial el modelo será :

$$y_i = \beta_1 + \beta_2 \ln\left(X_{1i}\right) + \beta_3 \frac{1}{X_{2i}} + \beta_4 X_{3i}^2 + \varepsilon_i \qquad (1)$$

donde ε_i representa una termino aleatorio de error.

Después de realizar 100 encuestas, obtiene los siguientes resultados:

$$X^T X = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 20 & -20 & 0 \\ 0 & -20 & 40 & 40 \\ 0 & 0 & 40 & 60 \end{bmatrix} \qquad X^T y = \begin{bmatrix} 60 \\ -15 \\ 10 \\ 5 \end{bmatrix}$$

a) ¿Qué características estadísticas tienen los estimadores MCO si se cumplen los supuestos del Teorema de Gauss-Markov **(4 puntos)**

Los estimadores MCO serán:
Insesgados (2 puntos)
Eficientes (minima varianza) (2 puntos)

b) Encuentre (si es posible) los mejores valores estimados para las siguientes cantidades **(8 puntos)**
- Valor esperado del precio del precio mensual por el cuidado de una planta **(2 puntos)**
- Valor esperado de la disponibilidad de pago de los clientes por el nuevo servicio de cuidado de plantas a domicilio **(2 puntos)**
- Valor esperado de los años de educación **(2 puntos)**
- Valor esperado de la proporción de años que faltan para su retiro **(2 puntos)**

Respuesta:
- Valor esperado del precio del precio mensual por el cuidado de una planta (2 puntos): No se tiene la información necesaria para calcular la media muestral de esta variable
- Valor esperado de la disponibilidad de pago de los clientes por el nuevo servicio de cuidado de plantas a domicilio (2 puntos): $\bar{y} = \frac{\sum y_i}{n} = \frac{60}{100} = 0.6$
- Valor esperado de los años de educación (2 puntos): No se tiene la información necesaria para calcular la media muestral de esta variable
- Valor esperado de la proporción de años que faltan para su retiro (2 puntos): No se tiene la información necesaria para calcular la media muestral de esta variable

c) Calcule los mejores estimadores lineales insesgados para los parámetros del modelo (1) **(10 puntos)**

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

$$\begin{bmatrix} 1/100 & 0 & 0 & 0 \\ 0 & -1/10 & -3/20 & 1/10 \\ 0 & -3/20 & -3/20 & 1/10 \\ 0 & 1/10 & 1/10 & -1/20 \end{bmatrix} \begin{bmatrix} 60 \\ -15 \\ 10 \\ 5 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 1/2 \\ 5/4 \\ -3/4 \end{bmatrix}$$

d) Explique el significado de los coeficientes estimados **(8 puntos)**

- $\hat{\beta}_1 = 3/5$ No tiene interpretación económica.

- $\hat{\beta}_2 = 1/2$ .Un aumento del uno porciento en los años de educación provocará un aumento de 5 mil pesos (1000000 * (1/2)/100) en la disponibilidad de pago de los clientes por un nuevo servicio de cuidado de plantas a domicilio.
- $\hat{\beta}_3 = 5/4$. Un aumento del uno por ciento en el precio mensual por el cuidado de una planta del provocará una disminución de $\frac{5}{400 \bullet X_{2i}}$ millones de pesos en la disponibilidad de pago de los clientes por un nuevo servicio de cuidado de plantas a domicilio.
- $\hat{\beta}_4 = -3/4$. Un aumento de un punto porcentual en la proporción de años que faltan para su retiro de mil pesos provocará una disminución de $1.5 \bullet X_{3i}$ millones de pesos en la disponibilidad de pago de los clientes por un nuevo servicio de cuidado de plantas a domicilio.

e) Un investigador cree que la elasticidad precio de la disponibilidad a pagar es diferente para aquellos clientes con ingresos superiores a dos millones de pesos mensuales. Es más, está casi seguro que es dos veces más que la elasticidad para aquellos cuyo ingreso es inferior o igual a dos millones de pesos. ¿cómo podría probar si el investigador tiene o no la razón. Explique claramente cómo lo haría, que fórmulas emplearía y que valores puede remplazar en dichas fórmulas y cómo tomaría la decisión. No es necesario que realice los cálculos. **(10 puntos)**

Para probar está hipótesis del investigador tenemos que reconocer que la elasticidad precio de la disponibilidad a pagar es:

$$\frac{\Delta\% y_i}{\Delta\% X_{2i}} = \frac{-\beta_3}{y_i \bullet X_{2i}}$$

Por otro lado, para probar la hipótesis tenemos que definir una variable dummy como por ejemplo:

$$D_i = \begin{cases} 1 & si\ i\ tiene\ ingreso\ mayor\ a\ 2\ millones \\ 0 & o.w. \end{cases}$$

Así el modelo será:

$$y_i = \beta_1 + \beta_2 \ln\left(X_{1i}\right) + \beta_3 \frac{1}{X_{2i}} + \beta_4 X_{3i}^2 + \alpha D_i \frac{1}{X_{2i}} + \varepsilon_i$$

(2 puntos)
Es importante mostrar que este modelo implica:

$$E\left[y_i\right]=\begin{cases}\beta_1+\beta_2\ln\left(X_{1i}\right)+\left(\beta_3+\alpha\right)\dfrac{1}{X_{2i}}+\beta_4X_{3i}^2 & si\ D_i=1\\ \beta_1+\beta_2\ln\left(X_{1i}\right)+\beta_3\dfrac{1}{X_{2i}}+\beta_4X_{3i}^2 & o.w.\end{cases}$$

(1 punto)
Esto equivale a que

$$\frac{\Delta\% y_i}{\Delta\% X_{2i}}=\begin{cases}\dfrac{-\left(\beta_3+\alpha\right)}{y_i^{>2}\bullet X_i^{>2}} & si\ i\ tiene\ ingreso\ mayor\ a\ 2\ millones\\ \dfrac{-\beta_3}{y_i^{\le 2}\bullet X_i^{\le 2}} & o.w.\end{cases}$$

Donde, $y_i^{>2}$ y $X_i^{>2}$ representan los valores de ambas variables para los individuos cuyos ingresos están por encima de 2 millones de pesos.

Por tanto, la hipótesis nula y alterna para la hipótesis del investigador son:

$$H_0:2\frac{\beta_3}{y_i^{\le 2}\bullet X_i^{\le 2}}=\frac{\left(\beta_3+\alpha\right)}{y_i^{>2}\bullet X_i^{>2}}$$

$$H_0:2\frac{\beta_3}{y_i^{\le 2}\bullet X_i^{\le 2}}-\frac{\left(\beta_3+\alpha\right)}{y_i^{>2}\bullet X_i^{>2}}=0$$

$$H_0:\left(\frac{2}{y_i^{\le 2}\bullet X_i^{\le 2}}-\frac{1}{y_i^{>2}\bullet X_i^{>2}}\right)\beta_3-\frac{\alpha}{y_i^{>2}\bullet X_i^{>2}}=0$$

$$H_A:No\ H_0$$

(2 puntos)

Esta hipótesis nula se puede escribir de la forma

$$R\beta=c$$

donde

$$R=\begin{bmatrix}0 & 0 & \left(\dfrac{2}{y_i^{\le 2}\bullet X_i^{\le 2}}-\dfrac{1}{y_i^{>2}\bullet X_i^{>2}}\right) & 0 & \dfrac{1}{y_i^{>2}\bullet X_i^{>2}}\end{bmatrix}$$

$$c=\left[0\right]$$

(1 punto)

y para probarla tenemos que emplear un F calculado de la siguiente forma:

$$F_c=\frac{\left(c-R\hat{\beta}\right)^T\left(R\left(X^TX\right)^{-1}R^T\right)^{-1}\left(c-R\hat{\beta}\right)\big/r}{SSE/n-k}=\frac{\left(c-R\hat{\beta}\right)^T\left(R\left(X^TX\right)^{-1}R^T\right)^{-1}\left(c-R\hat{\beta}\right)}{SSE/100-5}$$

(1 punto)

Entonces se rechaza la hipótesis nula si el F_c es mayor que el F de la tabla con un grado de libertad en el numerador y 95 en el denominador y un nivel de confianza de alfa. (2 puntos)

Tabla 1. Resultados de EasyReg.

```
Dependent variable:
Y = GR (i)

Characteristics:
GR (i)
  First observation = 1
  Last observation  = 100
  Number of usable observations: 100
  X variables:
X(1) = ingr (i)
X(2) = ingr (i)^2
X(3) = edad(i)
X(4) = 1
Model:
Y = b(1)X(1) +.....+ XXX + U,
where U is the error term, satisfying
E[U|X(1),...,X(4)] = 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)          0.9452       5.7316      1.638
                           XXX         (0.15260)
                           [0.00006]   [0.00027]
b(2)          -0.00006     -6          -0.677
                           (0.00001)   (0.0612)
                           [0.00014]   [0.25136]
b(3)          -0.12        -4.8701     -1.284
                           (0.02464)   (0.0115)
                           [0.00080]   [0.14711]
b(4)          100.4409     39.828      43.231
                           (142.62431) (131.39687)
                           [0.00000]   [0.00000]

Effective sample size (n):               100
Variance of the residuals:               72318.51487238
Standard error of the residuals (SER):   268.92101977
Residual sum of squares (RSS):           6942577.42774893
(Also called SSR = Sum of Squared Residuals)
Total sum of squares (TSS):              16455716.1564733
R-square:                                0.6781
Adjusted R-square:                       0.6649
Overall F test: F(3,96) = 43.85
p-value = 0.00000
Significance levels:   10%     5%
Critical values:       2.14    2.7
Conclusions:           reject  XXX
```