Econometría 06216
Examen Parcial #1
Cali, Sábado 27 de Febrero de 2010

Profesores: Julio César Alonso C.
Carlos Giovanni González
Ana Isabel Gallego L.

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 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.
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 respuestas escritas en 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.)

a. Para el modelo $y_{nx1} = \beta X_{nxc} + \varepsilon_{nx1}$ se construyó la siguiente tabla ANOVA. Pero lastimosamente, la Tabla no quedó bien impresa y se omitieron unos números que fueron remplazados por XXX. Después de leer la tabla, un estudiante hace la siguiente afirmación: "Es súper claro que el modelo tiene 4 pendientes y todas son significativas con un 99% de confianza dado que 40 es un número muy grande para la distribución F". ¿Esta afirmación es falsa o verdadera?

Fuente de variación	SS	G de L	MS
Regresión	xxx	xxx	100
Error	250	xxx	2.5
Total	xxx	104	

b. Si el producto de dos matrices $AB = C$ existe, entonces $BA = D$ sin que necesariamente $D = C$.

c. Al comparar dos estimadores lineales insesgados de α , sean $\hat{\alpha}$ y $\tilde{\alpha}$, se encuentra que $\hat{\alpha}$ tiene menor varianza que $\tilde{\alpha}$, por lo tanto, $\hat{\alpha}$ es el estimador MELI de α . ¿Es esta afirmación cierta?

d. Diga si la siguiente afirmación es verdadera o falsa:
"El siguiente modelo puede ser estimado por MCO": $y_i = \tan(X_{1,i}) + \beta \ln(X_{2,i}) + \gamma e^{y_i} + \delta_i + \varepsilon_i$

e. Un amigo mío toma la decisión de casarse porque "algo le dice" que, en Pitillito S.A., a un hombre casado le ofrecen, de entrada, un mejor sueldo que a cualquier otro empleado. Dado que tengo un amigo en Recursos Humanos de Pitillito que me prestó los datos de los empleados, puedo asegurar que el siguiente modelo me permite concluir sobre semejante hipótesis antes de que mi amigo tome tan seria decisión:

$$S_i = \alpha + \beta ex_i + \gamma G_i + \delta EC_i + \varepsilon_i$$

Donde

S_i es el salario ofrecido al empleado i.

ex_i son los años de experiencia previos a la entrada a la compañía

G_i es una dummy que toma el valor de 1 si el individuo es hombre y 0 si es mujer.

EC_i es una dummy que toma el valor de 1 si el individuo es casado y 0 si no lo es.

ε_i es una variable estocástica que cumple con los supuestos del teorema de Gauss-Markov

El profesor de econometría, al ver el modelo, afirma que: "el modelo planteado sí permite determinar si a un hombre casado le ofrecen, de entrada, un mejor sueldo que a cualquier otro empleado." ¿Es esta última 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 Un economista destacado escribe en un documento de trabajo: "Sabemos que los supuestos del modelo de regresión lineal múltiple son: 1). La relación entre Y y X es lineal. 2). Las X´s son no estocásticas y linealmente independientes entre sí. 3).i. $E[\beta] = 0_{nx1}$ ii. $\text{var}[\beta_i] = \sigma^2$ iii. $E[\beta_i, \beta_j] = 0 \quad i \neq j$

De 3)ii y 3) iii, tenemos:

$$E\left[\beta^T, \beta\right] = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix},,$$

¿Cuál de las siguientes afirmaciones es correcta?

a. Los supuestos 3)i y 3)ii son correctos
b. Los supuestos 3)i y 3)iii son correctos
c. Todos los supuestos son correctos
d. Ninguna de las anteriores

2.2 Considere el siguiente modelo de regresión estimado:

$$\hat{Y}_i = 0.32 + 1.76 \cdot X_{1i} + 2.43 \cdot X_{2i} + 1.75 \cdot X_{3i},$$
$$(0.10) \quad (0.16) \quad (0.22) \quad (0.25)$$

donde el número en paréntesis corresponde al error estándar. Ahora considere la siguiente prueba de hipótesis (sobre la pendiente asociada a X_{3i}):

$$H_0: \beta_4 \leq 0.50$$
$$H_A: \beta_4 > 0.50$$

Entonces:

a. el t estadístico calculado para comprobar esta hipótesis es 1.25.
b. el t estadístico calculado para comprobar esta hipótesis es 5.
c. el t estadístico calculado para comprobar esta hipótesis es 3.
d. el t estadístico calculado para comprobar esta hipótesis es 2.50.

2.3 Considere el siguiente modelo:

$$y_i = \beta_0 + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \varepsilon_i,$$

donde y_i representa el salario del i-ésimo graduado de la Universidad Icesi (en millones de pesos). Además:

$$X_{1i} = \begin{cases} 0 & \text{si el i-ésimo graduado tiene un título de posgrado} \\ 1 & \text{ow} \end{cases}.$$

y X_{2i} es el número de años transcurrido desde la graduación de la de la Universidad Icesi.

Suponiendo que un título profesional de posgrado aumenta el salario:

a) esperamos que β_1 sea positivo y mayor que β_2.
b) esperamos que β_1 sea positivo.
c) esperamos que β_1 sea negativo.
d) esperamos que β_1 sea positivo y menor que β_2.

3 **(30 Puntos)**

Una empresa multinacional con muchas plantas de producción en Europa y América, desea realizar un análisis sobre la función de producción de la empresa. Por lo que la junta directiva decide contratar a un economista como asesor principal. El economista, quien es un experto en el tema, decide recomendar dos funciones de producción y que la junta directiva tome la decisión final.

Uno de los modelos recomendado por el economista es:

Modelo 1:
$$y_j = \beta_1 L_j^{\beta_2} K_j^{\beta_3} e^{\beta_4 L_j + \beta_5 K_j} \varepsilon_j$$

El modelo 1 es una función de producción que proviene de una generalización de la conocida función de producción Cobb-Douglas. Donde y_j es el producto en miles de pesos, L_j es el insumo trabajo en miles de personas, y K_j es el insumo capital en miles de pesos.

El otro modelo que recomienda el economista es el siguiente:

Modelo 2:
$$\ln(y_j) = \alpha_1 + \alpha_2 \ln(K_j) + \alpha_3 K_j^2 + \alpha_4 TL_j + \mu_j$$

Donde y_j es el producto en miles de pesos, TL_j es la tasa de trabajadores dedicados exclusivamente al proceso de producción, y K_j es el insumo capital en miles de pesos.

Los dos modelos anteriores (modelo 1 y 2) son “versiones preliminares” y no son las versiones finalmente estimadas. Los resultados de los modelos que se estimaron se encuentran al final del examen.

a. Partiendo del modelo 2, la junta directiva de la empresa está interesada en conocer el efecto (*ceteris paribus*) que tiene en la producción, el que la planta se ubique en el continente americano o en el continente europeo, que son las dos zonas donde la empresa tiene plantas de producción. Escriba un modelo que permita lograr este objetivo y muestre por qué este modelo sí permite comprobar el efecto de la ubicación de la planta. **(6 puntos en total)**.
b. 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 **(6 puntos en total, 2 puntos cada uno).**
c. Interprete, a priori, el efecto que tiene el trabajo sobre la producción en el modelo 1 y discuta los signos o signo esperado. **(5 puntos).**
d. A partir de los coeficientes estimados, interprete el efecto que tiene el capital sobre la producción en el modelo 2, teniendo en cuenta su significancia. **(5 puntos).**
e. ¿Qué tan buenos son los modelos estimados? Justifique cuál sería el mejor modelo (el 1 o el 2) para realizar las conclusiones del estudio, **(8 puntos, 4 puntos por decir qué tan buenos son los modelos y 4 por justificar la elección del modelo).**

4 (30 puntos)
El jefe de la división de planeación del departamento de mercadeo de una cadena de almacenes, espera que las ventas diarias en millones de pesos (y_t) tengan el comportamiento descrito por el siguiente modelo estadístico:

$$y_t = \alpha_1 X_{1,t} + \alpha_2 X_{2,t} + \alpha_3 X_{3,t} + \varepsilon_t \qquad (1.1)$$

donde $X_{1,t}$ =1, $X_{2,t}$ y $X_{3,t}$ representan el inverso del número de reclamos en el día t y el logaritmo del porcentaje de descuento ofrecido en el día t respectivamente. Además, R_t y Des_t representan el número de reclamos y el porcentaje de descuento (medido en puntos porcentuales) ofrecido en el día t, respectivamente. Finalmente, ε_t representa una variable estocástica que cumple los supuestos del teorema de Gauss-Markov.

Los economistas de la división de planeación ya han recogido un buen número de observaciones diarias y se han efectuado los siguientes cálculos:

$$X^T X = \begin{bmatrix} 150 & 0 & 0 \\ 0 & 20 & 40 \\ 0 & 40 & 40 \end{bmatrix} \quad X^T y = \begin{bmatrix} 300 \\ 20 \\ 25 \end{bmatrix}$$

a. ¿Qué supuestos debe cumplir el término de error para que el estimador de Máxima Verosimilitud de las pendientes y el intercepto de un modelo sean MELI? **(3 puntos)**

b. A partir de esta información, explique a qué corresponden los elementos (1,3) y (2,2) de la matriz $X^T X$ Exprese su respuesta ***<u>en términos de las variables</u>*** R_t y Des_t **(4 puntos – dos puntos cada uno)**

c. Encuentre los estimadores de los betas del modelo por el método de MCO. **(6 Puntos).**

d. Interprete el significado de cada uno de los coeficientes estimados ***<u>en términos de las variables</u>*** R_t y Des_t , cuando sea posible. **(7 Puntos)**

e. En la última reunión del departamento de planeación, el director de mercadeo argumentó que: *"Acabo de recibir el reporte diario y veo que tenemos 200 reclamos y el descuento es del 10%. Las ventas diarias no satisfacen a la gerencia. ¡Tenemos que hacer algo! Yo estoy seguro que si logramos que en un día existan 5 reclamos y hacemos un descuento de únicamente el 1%, podríamos doblar las ventas que tenemos hoy"* Pruebe la hipótesis de dicho director. Sea lo más claro posible, y muestre qué fórmulas emplearía, qué valores debe reemplazar en las fórmulas (en caso que esas cantidades estén disponibles rápidamente), así como la manera en que tomaría la decisión. NO ES NECESARIO REALIZAR LOS CÁLCULOS RESPECTIVOS, pero sí debe indicar claramente cómo le daría o no la razón al director de mercadeo. **(10 Puntos)**

Tabla *. Resultados de un modelo estimado en EasyReg (Pregunta # 3).

```
Dependent variable:
Y = ln(Y)
Characteristics:
  First observation = 1
  Last observation  = 300
  Number of usable observations: 300
X variables:
X(1) = ln(K)
X(2) = K^2
X(3) = TL
X(4) = D1
X(5) = 1

Model:
Y = b(1)X(1) +.....+ b(5)X(5)  + U,
where U is the error term, satisfying
E[U|X(1),...,X(5)] = 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)       0.0465403      13.2417      12.024
                          (0.00354)    (0.0027)
                          [0.00000]    [0.00004]
b(2)       0.025414       7.3578       8.029
                          (0.003454)   (0.00287)
                          [0.00004]    [0.00006]
b(3)       XXX            3.208        4.676
                          (0.01224)    (0.00839)
                          [0.00034]    [0.00000]
b(4)       0.055455       0.6054       1.310
                          (0.09152)    (0.07656)
                          [0.19158]    [0.11567]
b(5)       4.549524       5.5713       8.310
                          (0.81592)    (0.49656)
                          [0.00000]    [0.00000]
Notes:
1: S.E. = Standard error
Effective sample size (n):                   XXX
Variance of the residuals:                   0.2546
Standard error of the residuals (SER):       0.18675684
Residual sum of squares (RSS):               XXX
Total sum of squares (TSS):                  4.42150588
R-square:                                    0.8864
Adjusted R-square:                           0.8738
Overall F test: F(4, 295) = 37.92
p-value = 0.00000
Significance levels:      10%       5%
Critical values:          2.76      3.81
Conclusions:              reject    reject
```

Tabla **. Resultados de un modelo estimado en EasyReg (Pregunta # 3).

```
Dependent variable:
Y = ln(Y)
Characteristics:
  First observation = 1
  Last observation  = 300
  Number of usable observations: 300
X variables:
X(1) = ln(L)
X(2) = ln(K)
X(3) = L
X(4) = K
X(5) = 1

Model:
Y = b(1)X(1) +.....+ b(5)X(5)  + U,
where U is the error term, satisfying
E[U|X(1),...,X(5)] = 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)       0.02803        13.0981       12.024
                          (0.00214)     (0.00027)
                          [0.00031]     [0.00004]
b(2)       0.054778       3.216         4.029
                          (0.01702)     (0.00787)
                          [0.00004]     [0.00006]
b(3)       0.0115879      3.7108        4.676
                          (0.00314)     (0.00839)
                          [0.00034]     [0.00020]
b(4)       0.025424       0.8457        1.0310
                          (0.03002)     (0.59656)
                          [0.15248]     [0.12567]
b(5)       -4.9576124     -0.413        -0.310
                          (12.0915)     (12.59656)
                          [0.25540]     [0.24600]
Notes:
1: S.E. = Standard error
Effective sample size (n):                300
Variance of the residuals:                2,1541321
Standard error of the residuals (SER):    0.1867568
Residual sum of squares (RSS):            2,0712154
Total sum of squares (TSS):               4.42150588
R-square:                                 0.8525
Adjusted R-square:                        0.8338
Overall F test: F(4, 295) = 37.92
p-value = 0.00000
Significance levels:     10%      5%
Critical values:         2.76     3.81
Conclusions:             reject   reject
```

Econometría 06216
Examen Parcial #1
Respuestas Sugeridas
Cali, Sábado 27 de Febrero de 2010

Profesores: Julio César Alonso C.
Carlos Giovanni González
Ana Isabel Gallego L.

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 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.
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 respuestas escritas en 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.)

NOTA: Si la razón no es correcta, se asignará solo un punto.

A. Para el modelo $y_{nx1} = \beta X_{nxc} + \varepsilon_{nx1}$ se construyó la siguiente tabla ANOVA. Pero lastimosamente, la Tabla no quedó bien impresa y se omitieron unos números que fueron remplazados por XXX. Después de leer la tabla, un estudiante hace la siguiente afirmación: "Es súper claro que el modelo tiene 4 pendientes y todas son significativas con un 99% de confianza dado que 40 es un número muy grande para la distribución F". ¿Esta afirmación es falsa o verdadera?

Fuente de variación	**SS**	**G de L**	**MS**
Regresión	xxx	xxx	100
Error	250	xxx	2.5
Total	xxx	104	

Falso. La afirmación tiene varios problemas, primero, el modelo fue mal escrito porque en forma matricial debería haber sido: $y_{nx1} = X_{nxc}\beta_{cx1} + \varepsilon_{nx1}$, se sabe que es en forma matricial porque no hay subíndices i o t. Segundo, dado que tenemos una matriz X, no es posible determinar si una de las columnas de la matriz es de unos (1), entonces no podemos saber a partir de la información que tenemos, si el modelo tiene o no tiene intercepto. Por tanto no es clarísimo que el modelo tiene 4 pendientes.
NOTA: La primera razón (Xbeta no conformable, da los 5 puntos), la segunda razón da 4 puntos.

B. Si el producto de dos matrices $AB = C$ existe, entonces $BA = D$ sin que necesariamente $D = C$.

Falso. La afirmación es cierta únicamente si estamos tratando con matrices cuadradas. Si se trata de matrices con dimensiones diferentes, por ejemplo, A es nxn y B es nxm y $n \neq m$, $AB = C$ existe, pero BA no es conformable. Era suficiente dar un contraejemplo para mostrar que la afirmación es falsa.
Aquí valdremos 3 puntos al que diga que es verdadero porque no se cumple la propiedad conmutativa.

C. Al comparar dos estimadores lineales insesgados de α , sean $\hat{\alpha}$ y $\tilde{\alpha}$, se encuentra que $\hat{\alpha}$ tiene menor varianza que $\tilde{\alpha}$, por lo tanto, $\hat{\alpha}$ es el estimador MELI de α ¿Es esta afirmación cierta?
Falso, porque puede haber un $\breve{\alpha}$ insesgado con menor varianza que $\hat{\alpha}$. El estimador MELI sería el de menor varianza de todos los estimadores lineales insesgados de α .

D. Diga si la siguiente afirmación es verdadera o falsa:
"El siguiente modelo puede ser estimado por MCO":

$$y_i = \tan(X_{1,i}) + \beta \ln(X_{2,i}) + \gamma e^{y_i} + \delta_i + \varepsilon_i$$

Verdadero, el modelo podría reparametrizarse como sigue:

$$y_i - \tan(X_{1,i}) = \beta \ln(X_{2,i}) + \gamma e^{y_i} + \delta_i + \varepsilon_i$$

$$z_i = \beta \ln(X_{2,i}) + \gamma X_{3,i} + \mu_i$$

donde

$$z_i = y_i - \tan(X_{1,i})$$

$$X_{3,i} = e^{y_i}$$

$$\mu_i = \delta_i + \varepsilon_i$$

E. Un amigo mío toma la decisión de casarse porque "algo le dice" que, en Pitillito S.A., a un hombre casado le ofrecen, de entrada, un mejor sueldo que a cualquier otro empleado. Dado que tengo un amigo en Recursos Humanos de Pitillito que me prestó los datos de los empleados, puedo asegurar que el siguiente modelo me permite concluir sobre semejante hipótesis antes de que mi amigo tome tan seria decisión:

$$S_i = \alpha + \beta ex_i + \gamma G_i + \delta EC_i + \varepsilon_i$$

Donde

S_i es el salario ofrecido al empleado i.

ex_i son los años de experiencia previos a la entrada a la compañía

G_i es una dummy que toma el valor de 1 si el individuo es hombre y 0 si es mujer.

EC_i es una dummy que toma el valor de 1 si el individuo es casado y 0 si no lo es.

ε_i es una variable estocástica que cumple con los supuestos del teorema de Gauss-Markov

El profesor de econometría, al ver el modelo, afirma que: "el modelo planteado sí permite determinar si a un hombre casado le ofrecen, de entrada, un mejor sueldo que a cualquier otro empleado." ¿Es esta última afirmación verdadera o falsa?
Falso, este modelo permite probar si hay algún diferencial per se en el salario entre un hombre y una mujer; así mismo, permite probar si hay algún diferencial per se, en el salario entre una persona casada y una que no lo está. Pero no permite determinar el diferencial per se por ser un hombre casado. Esto sólo se lograría si se incluyera una interacción entre las dummies, el modelo que permitiría probar la hipótesis es el siguiente:

$$S_i = \alpha + \beta ex_i + \gamma G_i + \delta EC_i + \varphi G_i EC_i + \varepsilon_i$$

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 Un economista destacado escribe en un documento de trabajo: "Sabemos que los supuestos del modelo de regresión lineal múltiple son: 1). La relación entre Y y X es lineal. 2). Las X´s son no estocásticas y linealmente independientes entre sí. 3).i. $E[\beta] = 0_{nx1}$ ii. $\operatorname{var}[\beta_i] = \sigma^2$ iii. $E[\beta_i, \beta_j] = 0 \quad i \neq j$

De 3)ii y 3) iii, tenemos:

$$E\left[\beta^T, \beta\right] = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix},,$$

¿Cuál de las siguientes afirmaciones es correcta?

a. Los supuestos 3)i y 3)ii son correctos
b. Los supuestos 3)i y 3)iii son correctos
c. Todos los supuestos son correctos
d. Ninguna de las anteriores

Respuesta: d. Noten que el grupo de supuestos expresados en 3 son sobre los coeficientes, no sobre los errores. Por tanto está mal.

2.2 Considere el siguiente modelo de regresión estimado:

$$\hat{Y}_i = 0.32 + 1.76 \cdot X_{1i} + 2.43 \cdot X_{2i} + 1.75 \cdot X_{3i},$$
$$(0.10) \quad (0.16) \quad (0.22) \quad (0.25)$$

donde el número en paréntesis corresponde al error estándar. Ahora considere la siguiente prueba de hipótesis (sobre la pendiente asociada a X_{3i}):

$$H_0: \beta_4 \leq 0.50$$
$$H_A: \beta_4 > 0.50$$

Entonces:

a. el t estadístico calculado para comprobar ésta hipótesis es 1.25.
b. el t estadístico calculado para comprobar ésta hipótesis es 5.
c. el t estadístico calculado para comprobar ésta hipótesis es 3.
d. el t estadístico calculado para comprobar ésta hipótesis es 2.50.

Respuesta: b.

2.3 Considere el siguiente modelo:

$$Y_i = \beta_0 + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \varepsilon_i,$$

donde Y_i representa el salario del i-ésimo graduado de la Universidad Icesi (en millones de pesos). Además:

$$X_{1i} = \begin{cases} 0 & si\ el\ \text{i}-esimo\ graduado\ tiene\ un\ titulo\ de\ posgrado \\ 1 & ow \end{cases}.$$

y X_{2i} es el número de años transcurrido desde la graduación de la de la Universidad Icesi.

Suponiendo que un título profesional de posgrado aumenta el salario:

a) esperamos que β_1 sea positivo y mayor que β_2.
b) esperamos que β_1 sea positivo.
c) esperamos que β_1 sea negativo.
d) esperamos que β_1 sea positivo y menor que β_2.

Respuesta: c.

3 **(30 Puntos)**

Una empresa multinacional con muchas plantas de producción en Europa y América desea realizar un análisis sobre la función de producción de la empresa. Por lo que la junta directiva decide contratar a un economista como asesor principal. El economista, quien es un experto en el tema, decide recomendar dos funciones de producción y que la junta directiva y el Gerente tomen la decisión final.

Uno de los modelos recomendado por el economista es:

Modelo 1: $$y_j = \beta_1 L_j^{\beta_2} K_j^{\beta_3} e^{\beta_4 L_j + \beta_5 K_j} \varepsilon_j$$

El modelo 1 es una función de producción que proviene de una generalización de la conocida función de producción Cobb-Douglas. Donde y_j es el producto en miles de pesos, L_j es el insumo trabajo en miles de personas, y K_j es el insumo capital en miles de pesos.

El otro modelo que recomienda el economista es el siguiente:

Modelo 2: $$\ln(y_j) = \alpha_1 + \alpha_2 \ln(K_j) + \alpha_3 K_j^2 + \alpha_4 TL_j + \mu_j$$

Donde y_j es el producto en miles de pesos, TL_j es la tasa de trabajadores dedicados exclusivamente al proceso de producción, y K_j es el insumo capital en miles de pesos.

Los dos modelos anteriores (modelo 1 y 2) son "versiones preliminares" y no son las versiones finalmente estimadas. Los resultados de los modelos que se estimaron se encuentran al final del examen.

a. Partiendo del modelo 2, la junta directiva de la empresa está interesada en conocer el efecto (ceteris paribus) que tiene en la producción, el que la planta se ubique en el continente

americano o en el continente europeo que son las dos zonas donde la empresa tiene plantas de producción. Escriba un modelo que permita lograr este objetivo y muestre por qué este modelo sí permite comprobar el efecto de la ubicación de la planta. **(6 puntos en total)**.

Partiendo del modelo 2: $\ln(y_i) = \alpha_1 + \alpha_2 \ln(K_i) + \alpha_3 \ln(K_i^2) + \alpha_4 TL_i + \mu_i$ y teniendo en cuenta el interés de la junta directiva es necesario incluir una variable dummy que mida el efecto de la ubicación de la planta:

$$D_1 = \begin{cases} 1 & si\ la\ planta\ está\ en\ el\ continente\ americano \\ 0 & O.W. \end{cases}$$

(2 puntos)

El modelo que se debe estimar es:

$$\ln(y_i) = \alpha_1 + \alpha_2 \ln(K_i) + \alpha_3 \ln(K_i^2) + \alpha_4 TL_i + \alpha_5 D_{1,i} + \mu_i$$

(2 puntos)

Donde $i = 1,2, \ldots 300$.

Para mostrar que este modelo sí permite comprobar el efecto de la ubicación de la planta debemos calcular el valor esperado y mostrar que α_5 debe ser diferente de cero si existe una diferencia producto de la ubicación (ceteris paribus):

$$E\left[\ln(y_i)\right] = \begin{cases} (\alpha_1 + \alpha_5) + \alpha_2 \ln(K_i) + \alpha_3 \ln(K_i^2) + \alpha_4 TL_i & si\ la\ planta\ está\ en\ el\ continente\ americano \\ \alpha_1 + \alpha_2 \ln(K_i) + \alpha_3 \ln(K_i^2) + \alpha_4 TL_i & o.w \end{cases}$$

(2 puntos)

b. 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, **(6 puntos en total, 2 puntos cada uno).**

En la tabla (*) hay tres valores que han sido reemplazados por "XXX". Estos valores corresponden a:

i. El coeficiente estimado α_4 que acompaña a la variable TL_i. Para realizar el cálculo en EasyReg tenemos: (Es el que aparece como b(3) en el output de Easyreg)

$$S_{\hat{\alpha}_4} = \frac{\hat{\alpha}_4}{tc_{\hat{\alpha}_4}}$$

$$0.01224 = \frac{\hat{\alpha}_4}{3.208}$$

$$\hat{\alpha}_4 = 0.03926$$

ii. El tamaño de la muestra usada en la estimación es:
n=300.

iii. Suma de los cuadrados de los errores

$$SSE = SST(1-R^2)$$

$$SSE = 4.42150588(1-0.8864) = 0.5022$$

c. Interprete, a priori, el efecto que tiene el trabajo sobre la producción en el modelo 1 y discuta los signos o signo esperado. **(5 puntos).**

El modelo 1 es el siguiente: $y_j = \beta_1 L_j^{\beta_2} K_j^{\beta_3} e^{\beta_4 L_j + \beta_5 K_j} \varepsilon_j$. Así, el modelo linealizado será:

$$\ln(y_j) = \ln(\beta_1) + \beta_2 \ln(L_j) + \beta_3 \ln(K_j) + \beta_4 L_j + \beta_5 K_j + \varepsilon_j$$

$$\ln(y_j) = \gamma + \beta_2 \ln(L_j) + \beta_3 \ln(K_j) + \beta_4 L_j + \beta_5 K_j + \varepsilon_j$$

Donde, $\ln(\beta_1) = \gamma$

Noten que en este caso tendremos:

$$\frac{dy_j}{dL_j} = \frac{d\left[\beta_1 K_j^{\beta_3} e^{\beta_5 K_j} L_j^{\beta_2} e^{\beta_4 L_j} \varepsilon_j\right]}{dL_j}$$

$$\frac{dy_j}{dL_j} = \beta_1 K_j^{\beta_3} e^{\beta_5 K_j} \varepsilon_j \left[\beta_2 L_j^{\beta_2 - 1} \cdot e^{\beta_4 L_j} + \beta_4 L_j^{\beta_2} e^{\beta_4 L_j}\right]$$

$$\frac{dy_j}{dL_j} = \beta_2 \frac{\beta_1 K_j^{\beta_3} L_j^{\beta_2} e^{\beta_5 K_j} e^{\beta_4 L_j} \varepsilon_j}{L_j} + \beta_4 \beta_1 K_j^{\beta_3} L_j^{\beta_2} e^{\beta_4 L_j} e^{\beta_5 K_j} \varepsilon_j$$

$$\frac{dy_j}{dL_j} = \beta_2 \frac{y_j}{L_j} + \beta_4 y_j$$

$$\frac{dy_j}{dL_j} = y_j \left(\frac{\beta_2}{L_j} + \beta_4\right)$$

$$\frac{dy_j / y_j}{dL_j} = \frac{\beta_2}{L_j} + \beta_4$$

$$\frac{dy_j / y_j}{dL_j / L_j} = \beta_2 + \beta_4 L_j$$

$$\frac{\Delta\% y_j}{\Delta\% L_j} = \beta_2 + \beta_4 L_j$$

Por tanto la elasticidad del producto con respecto al trabajo tiene una parte constante y una que varía con el nivel de trabajo. Noten que esta elasticidad puede tomar valores positivos o negativos dependiendo de si estamos en un segmento en el que representan rendimientos marginales positivos o negativos. Así, β_2 es la parte constante de la elasticidad del producto con respecto al trabajo (1.5 puntos). Y β_4 representa el cambio en la elasticidad del producto con respecto al trabajo cuando los trabajadores aumentan en 1000. (2 puntos).

En cuanto al signo, noten que el signo de $\frac{dy_j}{dL_j} = y_j\left(\frac{\beta_2}{L_j} + \beta_4\right)$ dependerá de si la función de producción presenta rendimientos marginales crecientes o decrecientes. En especial, dado que y_j y L_j son positivos el signo de los rendimientos marginales dependerán β_2 y β_4. En general los signos esperados dependerán de si se espera tener rendimientos marginales crecientes o decrecientes. Es decir,

$$\beta_2 > 0 \quad \beta_4 > 0 \qquad \Rightarrow \qquad \frac{dy_j}{dL_j} > 0$$

$$\beta_2 < 0 \quad \beta_4 < 0 \qquad \Rightarrow \qquad \frac{dy_j}{dL_j} < 0$$

$$\beta_2 > 0 \quad \beta_4 < 0 \quad \wedge -\frac{\beta_2}{\beta_4} > L_j \Rightarrow \frac{dy_j}{dL_j} > 0$$

$$\beta_2 > 0 \quad \beta_4 < 0 \quad \wedge -\frac{\beta_2}{\beta_4} < L_j \Rightarrow \frac{dy_j}{dL_j} < 0$$

$$\beta_2 < 0 \quad \beta_4 > 0 \quad \wedge -\frac{\beta_2}{\beta_4} < L_j \Rightarrow \frac{dy_j}{dL_j} > 0$$

$$\beta_2 < 0 \quad \beta_4 > 0 \quad \wedge -\frac{\beta_2}{\beta_4} > L_j \Rightarrow \frac{dy_j}{dL_j} > 0$$

Nota: no se esperaba un análisis tan detallado, con solo determinar que el signo de los coeficientes depende del signo de los rendimientos marginales del trabajo (1.5 puntos)

d. A partir de los coeficientes estimados, interprete el efecto que tiene el capital sobre la producción en el modelo 2, teniendo en cuenta su significancia. **(5 puntos).**

Partiendo del modelo $\ln(y_j) = \alpha_1 + \alpha_2 \ln(K_j) + \alpha_3 K_j^2 + \alpha_4 TL_j + \mu_j$ podemos encontrar que:

$$\ln(y_j) = \alpha_1 + \alpha_2 \ln(K_j) + \alpha_3 K_j^2 + \alpha_4 TL_j + \mu_j$$

$$y_j = e^{\alpha_1 + \alpha_2 \ln(K_j) + \alpha_3 K_j^2 + \alpha_4 TL_j + \mu_j}$$

$$\frac{dy_j}{dK_j} = \left(\frac{\alpha_2}{K_j} + 2 \cdot \alpha_3 K_j\right) e^{\alpha_1 + \alpha_2 \ln(K_j) + \alpha_3 K_j^2 + \alpha_4 TL_j + \mu_j}$$

$$\frac{dy_j}{dK_j} = \left(\frac{\alpha_2}{K_j} + 2 \cdot \alpha_3 K_j\right) y_j$$

$$\frac{dy_j / y_j}{dK_j} = \frac{\alpha_2}{K_j} + 2 \cdot \alpha_3 K_j$$

$$\frac{\Delta \% y_j}{\Delta \% K_j} = \alpha_2 + 2 \cdot \alpha_3 K_j^2$$

Noten que $\hat{\alpha}_2 = 0.0465403$, es significativo al 99% de confianza. Es la parte constante de la elasticidad del producto con respecto al insumo capital ***(2 puntos)***. Y, $\alpha_3 = 0.025414$, es significativo al 99% de confianza. Por lo tanto, un aumento de mil pesos en el capital provoca un aumento de $4 * 0.025414 K_j$ en la elasticidad del producto con respecto al insumo capital ***(2 puntos)***.

e. ¿Qué tan buenos son los modelos estimados? Justifique cuál sería el mejor modelo (el 1 o el 2) para realizar las conclusiones del estudio, **(8 puntos, 4 puntos por decir que tan buenos son los modelos y 4 por justificar la elección del modelo).**

Noten que la pregunta tiene dos partes. Primero se presunta: Qué tan buenos son los modelos estimados:

Modelo 1 (Tabla **) ***(2 puntos)***

Para entender qué tan bueno es el modelo, es necesario en primer lugar revisar la significancia conjunta del modelo. Para esto se plantean las siguientes hipótesis a contrastar $H_0: \gamma_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$ 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. Ahora bien, el $R^2 = 0.8864$, indica que el 88,64% de la variabilidad de la variable dependiente es explicada por el modelo. Por tanto se puede concluir que el modelo presenta un buen ajuste.

Modelo 2 (Tabla *) ***(2 puntos)***

Igualmente lo podemos hacer para el modelo 2. Se plantean las siguientes hipótesis a contrastar $H_0: \gamma_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$ 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. Ahora bien, el $R^2 = 0.8525$, indica que el 85,25% de la variabilidad de la variable dependiente es explicada por el modelo. Por tanto se puede concluir que el modelo presenta un buen ajuste.

La segunda parte de la pregunta es elegir el mejor modelo de entre los dos que se estimaron: Y así para elegir el mejor modelo comparamos el R^2-Ajustado (El R^2 también puede ser empleado en este caso).

[$R^2 - Ajus. = 0.8338$ del modelo 1 (Tabla **)] < [$R^2 - Ajus. = 0.8738$ del Modelo 2 (Tabla *)]

El modelo 2 (Tabla *) presenta un mayor valor de $R^2 - Ajustado$ que el estimado para el modelo 1. Y por tanto el mejor modelo es el 2
(4 puntos)

4 (30 puntos)
El jefe de la división de planeación del departamento de mercadeo de una cadena de almacenes espera que las ventas diarias en millones de pesos (y_t) tengan un comportamiento descrito por el siguiente modelo estadístico: $y_t = \alpha_1 X_{1,t} + \alpha_2 X_{2,t} + \alpha_3 X_{3,t} + \varepsilon_t$ (1.1)
donde $X_{1,t}$ =1,. Además, $X_{2,t}$ y $X_{3,t}$ representa el inverso del número de reclamos en el día t y el logaritmo del porcentaje de descuento ofrecido en el día t. Además, R_t y Des_t representan el

número de reclamos y el porcentaje de descuento (medido en puntos porcentuales) ofrecido en el día t, respectivamente. Finalmente, ε_t representa una variable estocástica que cumple los supuestos del teorema de Gauss-Markov.

Los economistas de la división de planeación ya han recogido un buen número de observaciones diarias y se han efectuado los siguientes cálculos:

$$X^T X = \begin{bmatrix} 150 & 0 & 0 \\ 0 & 20 & 40 \\ 0 & 40 & 40 \end{bmatrix} \quad X^T y = \begin{bmatrix} 300 \\ 20 \\ 25 \end{bmatrix}$$

a) ¿Qué supuestos debe cumplir el término de error para que el estimador de Máxima Verosimilitud de las pendientes y el intercepto de un modelo sean MELI? **(3 puntos)**

Para que los estimadores de máxima verosimilitud para las pendientes y el intercepto sean insesgados se necesita que el término de error tenga:

- media cero (0.5 puntos)
- varianza constante (0.5 puntos)
- no autocorrelación (0.5 puntos)
- sigue una distribución normal. (1.5 punto)

b) A partir de esta información, explique a qué corresponden los elementos (1,3) y (2,2) de la matriz $X^T X$. Exprese su respuesta ***en términos de las variables*** R_t y Des_t **(4 puntos – dos puntos cada uno)**

En este caso se tiene que:

- El elemento (1,3)= $\sum_{i=1}^{150} \left(\ln \left(Des_t \right) \right)$
- El elemento (2,2)= $\sum_{i=1}^{150} \frac{1}{R_t^2} = 20$

Nota si las sumatorias se expresan en Términos de X2 y X3 no se dará crédito, pues la pregunta es lo suficientemente clara.

c) Encuentre los estimadores de los betas del modelo por el método de MCO. **(6 Puntos).**

En este caso tenemos que:

$$\hat{\beta} = \left(X^T X \right)^{-1} X^T y = \begin{bmatrix} 1/150 & 0 & 0 \\ 0 & -1/20 & 1/20 \\ 0 & 1/20 & -1/40 \end{bmatrix} \begin{bmatrix} 300 \\ 20 \\ 25 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/4 \\ 3/8 \end{bmatrix}$$

d) Interprete el significado de cada uno de los coeficientes estimados ***en términos de las variables*** R_t y Des_t, cuando sea posible. **(7 Puntos)**

Noten que el correspondiente modelo es: $y_t = \alpha_1 + \alpha_2 \frac{1}{R_t} + \alpha_3 \ln \left(Des_t \right) + \varepsilon_t$

$\hat{\alpha}_1 = 2$. Carece de interpretación económica (1 puntos)

$\hat{\alpha}_2 = 1/4$ Un aumento de un uno por ciento en el número de reclamos provoca una disminución en las ventas diarias de $\frac{2500}{R_t}$ pesos. (3 puntos)

$$\frac{\partial y_t}{\partial R_t} = -\alpha_2 \frac{1}{R_t^2}$$

$$\frac{\partial y_t}{\partial R_t / R_t} = -\alpha_2 \frac{1}{R_t}$$

$$\frac{\partial y_t}{\Delta \% R_t} = -\frac{\alpha_2}{100} \frac{1}{R_t}$$

$\hat{\alpha}_3 = 3/8$. Un aumento del uno por ciento en el descuento ofrecido en el día t implica un aumento en las ventas diarias de 3750 pesos. (3 puntos)

$$\frac{\partial y_t}{\partial Des_t} = \alpha_3 \frac{1}{Des_t}$$

$$\frac{\partial y_t}{\partial Des_t / Des_t} = \alpha_3$$

$$\frac{\partial y_t}{\Delta \% Des_t} = \frac{\alpha_3}{100}$$

e) En la última reunión del departamento de planeación, el director de mercadeo argumentó que: *"Acabo de recibir el reporte diario y veo que tenemos 200 reclamos y el descuento es del 10%. Las ventas diarias no satisfacen a la gerencia. ¡Tenemos que hacer algo! Yo estoy seguro que si logramos que en un día existan 5 reclamos y hacemos un descunto de únicamente el 1%, podríamos doblar las ventas que tenemos hoy"* Pruebe la hipótesis de dicho director. Sea lo más claro posible, y muestre qué fórmulas emplearía, qué valores debe reemplazar en las fórmulas (en caso que esas cantidades estén disponibles rápidamente), así como la manera en que tomaría la decisión. NO ES NECESARIO REALIZAR LOS CÁLCULOS RESPECTIVOS, pero sí debe indicar claramente cómo le daría o no la razón al director de mercadeo. **(10 Puntos)**

Esta afirmación es equivalente a probar la siguiente hipótesis nula:

$$\alpha_1 + \alpha_2 \frac{1}{5} + \alpha_3 \ln(1) = 2\left(\alpha_1 + \alpha_2 \frac{1}{200} + \alpha_3 \ln(10) \right)$$

Esto es equivalente a

$$-\alpha_1+\left(\frac{1}{5}-\frac{1}{100}\right)\alpha_2+2\bullet\ln\left(10\right)\alpha_3=0$$

$$-\alpha_1+\left(\frac{19}{100}\right)\alpha_2+2\bullet\ln\left(10\right)\alpha_3=0$$

Versus la hipótesis alterna no H_o. (4 puntos por plantear la Ho y la Ha)

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

$R=\begin{bmatrix}-1 & \left(\frac{19}{100}\right) & 2\bullet\ln\left(10\right)\end{bmatrix}$ y $C=0$

Esto implica el siguiente F calculado:

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

$$=\frac{\left(0-\begin{bmatrix}-1 & \left(\frac{19}{100}\right) & 2\bullet\ln(10)\end{bmatrix}\begin{bmatrix}2\\1/4\\3/8\end{bmatrix}\right)^T\left(\begin{bmatrix}-1 & \left(\frac{19}{100}\right) & 2\bullet\ln(10)\end{bmatrix}\begin{bmatrix}1/150 & 0 & 0\\0 & -1/20 & 1/20\\0 & 1/20 & -1/40\end{bmatrix}\begin{bmatrix}-1 & \left(\frac{19}{100}\right) & 2\bullet\ln(10)\end{bmatrix}^T\right)^{-1}\left(c-R\hat{\beta}\right)}{SSE/\left(150-3\right)}$$

(4 puntos)

Este estadístico se debe comparar con el F de la tabla con 1 grado de libertad en el numerador y 147 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. (2 puntos)

Tabla *. Resultados de un modelo estimado en EasyReg (Pregunta # 3).

Dependent variable:
Y = ln(Y)
Characteristics:
First observation = 1
Last observation = 300
Number of usable observations: 300
X variables:
X(1) = ln(K)
X(2) = K^2
X(3) = TL
X(4) = D1
X(5) = 1
Model:
Y = b(1)X(1) +.....+ b(5)X(5) + U,
where U is the error term, satisfying
E[U|X(1),...,X(5)] = 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)	0.0465403	13.2417	12.024
		(0.00354)	(0.0027)
		[0.00000]	[0.00004]
b(2)	0.025414	7.3578	8.029
		(0.003454)	(0.00287)
		[0.00004]	[0.00006]
b(3)	**XXX**	3.208	4.676
		(0.01224)	(0.00839)
		[0.00034]	[0.00000]
b(4)	0.055455	0.6054	1.310
		(0.09152)	(0.07656)
		[0.19158]	[0.11567]
b(5)	4.549524	5.5713	8.310
		(0.81592)	(0.49656)
		[0.00000]	[0.00000]

Notes:
1: S.E. = Standard error

Effective sample size (n):	**XXX**
Variance of the residuals:	0.2546
Standard error of the residuals (SER):	0.18675684
Residual sum of squares (RSS):	**XXX**
Total sum of squares (TSS):	4.42150588
R-square:	0.8864
Adjusted R-square:	0.8738

Overall F test: F(4, 295) = 37.92
p-value = 0.00000

Significance levels:	10%	5%
Critical values:	2.76	3.81
Conclusions:	reject	reject

Tabla **. Resultados de un modelo estimado en EasyReg (Pregunta # 3).

Dependent variable:
Y = ln(Y)
Characteristics:
First observation = 1
Last observation = 300
Number of usable observations: 300
X variables:
X(1) = ln(L)
X(2) = ln(K)
X(3) = L
X(4) = K
X(5) = 1

Model:
Y = b(1)X(1) +.....+ b(5)X(5) + U,
where U is the error term, satisfying
E[U|X(1),...,X(5)] = 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)	0.02803	13.0981	12.024
		(0.00214)	(0.00027)
		[0.00031]	[0.00004]
b(2)	0.054778	3.216	4.029
		(0.01702)	(0.00787)
		[0.00004]	[0.00006]
b(3)	0.0115879	3.7108	4.676
		(0.00314)	(0.00839)
		[0.00034]	[0.00020]
b(4)	0.025424	0.8457	1.0310
		(0.03002)	(0.59656)
		[0.15248]	[0.12567]
b(5)	-4.9576124	-0.413	-0.310
		(12.0915)	(12.59656)
		[0.25540]	[0.24600]

Notes:
1: S.E. = Standard error

Effective sample size (n):	300
Variance of the residuals:	2,1541321
Standard error of the residuals (SER):	0.1867568
Residual sum of squares (RSS):	2,0712154
Total sum of squares (TSS):	4.42150588
R-square:	0.8525
Adjusted R-square:	0.8338

Overall F test: F(4, 295) = 37.92
p-value = 0.00000

Significance levels:	10%	5%
Critical values:	2.76	3.81
Conclusions:	reject	reject