División Académica de Actuaría, Estadística y Matemáticas

Departamento de Estadística

Econometría

Al finalizar el curso de Econometría el alumno será capaz de:

- Definir Econometría y sus nexos con la Teoría Económica y la Estadística.
- Identificar datos de corte transversal y datos de series de tiempo.
- Identificar e interpretar los elementos del modelo de regresión lineal múltiple.
- Identificar los supuestos del modelo y sus consecuencias en las propiedades estadísticas de los estimadores de los parámetros del modelo.
- Aplicar las técnicas de inferencia estadística en un modelo estimado y obtener conclusiones.
- Manejar las propiedades básicas del modelo de regresión lineal múltiple con herramientas matriciales.
- Hacer el diagnóstico acerca del cumplimiento de los supuestos del modelo, señalar las consecuencias en caso de que no se cumplan los supuestos y realizar, cuando sea posible, correcciones al modelo de regresión lineal.
- Hacer acopio de elementos para juzgar un modelo de regresión lineal.
- Calcular e interpretar pronósticos puntuales y por intervalo con un modelo de regresión lineal.
- Visualizar algunos conceptos y realizar cálculos selectos de la materia con ayuda del paquete estadístico `R`.

Con el propósito de indicar el nivel requerido para acreditar el curso, en este documento se reúnen las versiones originales de los exámenes finales departamentales de la materia de Econometría aplicados por el Departamento de Estadística del ITAM en el periodo del semestre de Otoño 2019 a Primavera 2021, con sus respectivas respuestas dadas por los coordinadores en el semestre correspondiente. .

Cualquier observación favor de dirigirla a:

José Manuel Lecuanda Ontiveros
`manuel@lecuanda.com`
Coordinador de la materia
Semestre Otoño 2021

Índice general

1. Examen Final (Otoño 2019) 3
2. Examen Final (Otoño 2019) - Solución 14
3. Examen Final (Otoño 2020) 25
4. Examen Final (Otoño 2020) - Solución 33
5. Examen Final (Primavera 2021) 41
6. Examen Final (Primavera 2021) - Solución 51

1. Examen Final (Otoño 2019)

Econometría
7 de diciembre de 2019

Nombre: _____ CU: _____

INSTRUCCIONES:

1. Lea con cuidado cada pregunta antes de responderla.
2. Marque la respuesta correcta para cada pregunta en la hoja de respuestas anexa. Tome en cuenta que para cada pregunta sólo existe una única respuesta correcta.
3. Tome su tiempo para responder cada pregunta. Cuenta con tiempo suficiente para realizar el examen. No hay prisa.
4. Este es un examen a libro cerrado. Cualquier material escrito o electrónico adicional a este examen está prohibido. No es necesario ningún formulario.
5. No desengrape su examen. Puede utilizar las hojas blancas adicionales para sus cálculos pero no se calificará procedimiento alguno.
6. Se requiere de una calculadora científica y de las tablas estadísticas anexas. El uso de cualquier otro tipo de equipo así como de teléfonos móviles está estrictamente prohibido. No los use.
7. No hagas trampa, no copies, no robes ideas y no permitas que otros lo hagan. Cualquier conducta fraudulenta derivará en las sanciones especificadas en el reglamento del ITAM.
8. Cualquier prueba estadística o intervalo de confianza, utilice el 95 % de confianza. Los números entre paréntesis abajo de cualquier estimador se refieren a su error estándar. El término ε siempre representa el término estocástico de la regresión.

1. **(3 PUNTOS).** Un ingeniero investiga el uso de un molino de viento para generar electricidad para lo cual cuenta con n observaciones de corriente directa producida por m molinos. El modelo teórico plantea que la energía producida por los molinos medida en Megawatts (Y_{it}) depende de la velocidad del viento correspondiente medida en Millas por Hora (X_{it}). Como un primer paso en el modelado empírico se proponen dos modelos de regresión lineal simple:

   - Modelo 1: $y_{it} = \beta_0 + \beta_1 x_{it} + \varepsilon_{it} \ \forall \ i = 1, \cdots, n; \ t = 1, \cdots, m$
   - Modelo 2: $y_i = \beta_0 + \beta_1 \frac{1}{x_i} + \varepsilon_i \ \forall \ i = 1, \cdots, n$

   Entonces, se puede asegurar que:

   a) El modelo 1 emplea datos de tipo transversal y el modelo 2 datos de tipo longitudinal (serie de tiempo).

   b) El modelo 1 emplea datos de tipo panel y el modelo 2 datos de tipo longitudinal (serie de tiempo).

   c) El modelo 1 emplea datos de tipo longitudinal (serie de tiempo) y el modelo 2 datos de tipo transversal.

   d) El modelo 1 emplea datos de tipo panel y el modelo 2 datos de tipo transversal.

2. **(5 PUNTOS).** Un investigador propone un modelo lineal que busca explicar el comportamiento de una variable Y mediante el comportamiento de una variable X. En primera instancia, se ajusta el modelo utilizando ambas variables medidas en miles de pesos con lo que se estima a β_0 y a β_1 mediante los estimadores de mínimos cuadrados ordinarios $\hat{\beta}_0$ y $\hat{\beta}_1$. Después, ajusta nuevamente el modelo considerando las unidades de la variable dependiente en pesos pero sin alterar las unidades de la variable independiente, y obtiene estimaciones $\hat{\beta}_0^\star$ y $\hat{\beta}_1^\star$. Entonces se puede asegurar que:

   a) $\hat{\beta}_0^\star = \hat{\beta}_0$ y $\hat{\beta}_1^\star = \hat{\beta}_1$

   b) $\hat{\beta}_0^\star = 1000\hat{\beta}_0$ y $\hat{\beta}_1^\star = 1000\hat{\beta}_1$

   c) $\hat{\beta}_0^\star = \frac{1}{1000}\hat{\beta}_0$ y $\hat{\beta}_1^\star = \frac{1}{1000}\hat{\beta}_1$

   d) Ninguna de las anteriores.

3. **(5 PUNTOS).** En el contexto de un modelo de regresión lineal simple de la forma: $y_i = \theta_0 + \theta_1 x_i + \varepsilon_i$, suponga los siguientes modelos:

   - Modelo 1: $y = \beta_0(1 - e^{\beta_1 x})$
   - Modelo 2: $y = \frac{x}{\beta_0 + \beta_1 x}$

   Si se consideran transformaciones y/o cambios de variable para ajustar un modelo de regresión lineal simple, entonces:

   a) Ambos modelos (1) y (2) son linearizables.

   b) El modelo (1) es linearizable pero (2) no lo es.

   c) El modelo (1) no es linearizable pero (2) si lo es.

   d) Ni el modelo (1) ni el modelo (2) son linearizables.

4. **(3 PUNTOS).** Suponiendo el modelo de regresión lineal $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ donde se satisfacen los supuestos usuales para el término estocástico ε_i, esto es:

$$E\left[\varepsilon_i | x_i\right] = 0;\ Var[\varepsilon_i | x_i] = \sigma_\varepsilon^2;\ Cov[\varepsilon_i, \varepsilon_j | x_i] = 0\ \forall\ i \neq j$$

   Entonces es cierto que:

   a) $\sum_{i=1}^n \hat{y}_i \hat{\varepsilon}_i = 0$

   b) $\sum_{i=1}^n x_i \hat{\varepsilon}_i = 0$

   c) $\sum_{i=1}^n y_i = \sum_{i=1}^n \hat{y}_i$

   d) Todas las anteriores.

5. **(5 PUNTOS).** Sea un modelo de regresión lineal simple de la forma $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ donde el supuesto de que $Var[\varepsilon_i | x_i] = \sigma_\varepsilon^2$ no se satisface pero que los otros dos supuestos de Gauss Markov, $E[\varepsilon_i | x_i] = 0$ y $Cov[\varepsilon_i, \varepsilon_j | x_i] = 0$ sí se satisfacen. Considere las siguientes afirmaciones:

   - (1): $\hat{\beta}_0$ y $\hat{\beta}_1$ son insesgados.
   - (2): $Var(\hat{\beta}_0) = \sigma_\varepsilon^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right)$ y $Var(\hat{\beta}_1) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$.

   Luego, se tiene que:

   a) La afirmación (1) es verdadera pero la afirmación (2) es falsa.

   b) La afirmación (1) es falsa pero la afirmación (2) es verdadera.

   c) Ambas afirmaciones (1) y (2) son falsas.

   d) Ambas afirmaciones (1) y (2) son verdaderas.

6. **(3 PUNTOS).** Suponga que se sabe que el efecto marginal de x_i sobre y_i es negativo y no constante. Entonces, un modelo adecuado sería:

   a) $y_i = \beta_0 + \beta_1 x_i + \beta_2 ln(x_i) + \varepsilon_i$

   b) $y_i = \beta_0 + \beta_1 x_i + \beta_2 \frac{1}{\sqrt{x_i}} + \varepsilon_i$ con $\beta_1 < 0$ y $\beta_2 > 0$

   c) $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ con $\beta_1 < 0$

   d) $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i$ con $\beta_2 < 0$

7. **(8 PUNTOS).** Suponga que se desea probar que el puntaje promedio obtenido en la prueba ENLACE por los niños es diferente al obtenido por las niñas. Para ello se estiman los siguientes modelos, con el puntaje estandarizado obtenido por cada alumno en la prueba $ENLACE$ variable dependiente y el promedio de sus calificaciones en el año escolar ($PROM$), la calificación global de la escuela ESC y el número de horas de estudio $HREST$ como variables independientes:

   Para los 366 alumnos de la escuela, con un $R^2 = 0.3516$ y $STC = 131.8936$:

   $$ENLACE_i = \underset{(0.1837)}{1.4909} + \underset{(0.0002)}{0.0012} PROM_i - \underset{(0.0012)}{0.0099} ESC_i + \underset{(0.0008)}{0.0023} HREST_i$$

   Para las 90 niñas de la escuela, con un $R^2 = 0.4014$ y $STC = 32.7493$:

   $$ENLACE_i = \underset{(0.3616)}{1.1273} + \underset{(0.0003)}{0.0018} PROM_i - \underset{(0.0029)}{0.0090} ESC_i + \underset{(0.0022)}{0.0014} HREST_i$$

   Para los 276 niños de la escuela, con un $R^2 = 0.3169$ y $STC = 86.0014$:

   $$ENLACE_i = \underset{(0.2060)}{1.4808} + \underset{(0.0002)}{0.0011} PROM_i - \underset{(0.0014)}{0.0085} ESC_i + \underset{(0.0008)}{0.0023} HREST_i$$

   Luego de realizar la prueba estadística adecuada, puede concluir que:

   a) El modelo de los niños no es diferente al de las niñas por lo que se necesita incluir la variable dicotómica de género del alumno.

   b) El modelo de los niños es diferente al de las niñas por lo que se necesita incluir la variable dicotómica de género del alumno.

   c) El modelo de los niños no es diferente al de las niñas por lo que no se necesita incluir la variable dicotómica de género del alumno.

   d) El modelo de los niños es diferente al de las niñas, por lo que no se necesita incluir la variable dicotómica de género del alumno.

8. **(8 PUNTOS).** Se sabe que una función de producción del tipo Cobb-Douglas dada por $Y_i = CK_i^{\beta_1} L_i^{\beta_2}$ muestra rendimientos constantes a escala si se cumple que $\beta_1 + \beta_2 = 1$. En caso de que sean crecientes o decrecientes, la suma de estos coeficientes será mayor o menor a uno, respectivamente. Suponga utilizando 33 observaciones se estima el siguiente modelo lineal:

$$\widehat{lnY_i} = \underset{(0.2975)}{-6.3773} + \underset{(0.0705)}{0.5709} lnK_i + \underset{(0.1006)}{2.0786} lnL_i$$

con $Cov(\hat{\beta}_1, \hat{\beta}_2) = -0.0056$, $R^2 = 0.9827$ y $SRC = 0.1394$. Entonces puede asegurar que la función de producción estimada:

a) Presenta rendimientos constantes a escala.

b) Presenta rendimientos crecientes a escala.

c) Presenta rendimientos decrecientes a escala.

d) No se pueden determinar sus rendimientos a escala.

9. **(5 PUNTOS).** Suponga que al estimar el modelo

$$S_t = \beta_0 + \beta_1 A_t + \beta_2 P_t + \beta_3 E_t + \beta_4 A_{t-1} + \beta_5 P_{t-1} + \varepsilon_t$$

se obtienen los siguientes resultados:

	$\hat{\beta}$	$e.s.(\hat{\beta})$	t	valor p	FIV
Intercepto	-14.194	18.715	-0.758	0.4592	
A_t	5.361	4.028	1.331	0.2019	36.94
P_t	8.372	3.586	2.334	0.0329	33.47
E_t	22.521	2.142	10.512	0.0000	1.07
A_{t-1}	3.855	3.578	1.077	0.2973	25.91
P_{t-1}	4.125	3.895	1.059	0.3053	43.52

Entonces, el coeficiente de determinación R^2 que se obtiene al ajustar A_{t-1} con respecto al resto de los regresores A_t, P_t, E_t, P_{t-1} está en el intervalo:

a) $[0, 0.25)$

b) $[0.25, 0.50)$

c) $[0.50, 0.75)$

d) $[0.75, 1.0)$

10. **(8 PUNTOS).** Para analizar la discriminación salarial en una empresa, se estima el siguiente modelo:

$$\widehat{ln(SAL_i)} = 1.5 + 0.08ln(ED_i) + 0.05ln(EX_i) + 0.03H_i + 0.02P_i + 0.03H_iP_i$$

donde SAL_i es el salario del trabajador, ED_i y EX_i su educación y experiencia en años, respectivamente, H_i una variable dicotómica que toma el valor de 1 si el individuo es hombre y 0 si es mujer, y P_i una variable dicotómica que toma el valor de 1 si el individuo posee un posgrado y 0 si no. Todos los coeficientes estimados fueron significativos al 95 % de confianza. Entonces, se puede asegurar que:

a) Una mujer con posgrado gana mas que un hombre sin posgrado.

b) Una mujer con posgrado gana menos que una mujer sin posgrado.

c) Un individuo con posgrado siempre gana más, sin importar su género.

d) Un hombre gana más que una mujer pero sólo si no tiene posgrado.

11. **(3 PUNTOS).** Considere las siguientes afirmaciones en relación al análisis de la multicolinealidad en un modelo de regresión lineal múltiple:

- (1): Si en el modelo lineal con 7 variables regresoras, la variable cualitativa X_2 con 4 categorías, se codifica mediante cuatro variables dicotómicas D_1, D_2, D_3 y D_4, y se incluyen en un modelo regresión lineal con un intercepto β_0, entonces el modelo presentará necesariamente problemas de multicolinealidad.
- (2) Si la correlación de la variable dependiente y con uno de los regresores del modelo, digamos la variable x_1 es mayor a 0.9, entonces necesariamente se tiene un problema de multicolinealidad.

Entonces, es cierto que:

a) La afirmación (1) es falsa mientras que la afirmación (2) es correcta.

b) La afirmación (1) es correcta mientras que la afirmación (1) es falsa.

c) Ambas afirmaciones, la (1) y la (2), son son falsas.

d) Ambas afirmaciones, la (1) y la (2), son son correctas.

12. **(3 PUNTOS).** Usted y un colega ajustan un modelo de regresión lineal simple donde, para validar el supuesto de homoscedasticidad para la varianza del error se procede a realizar un análisis de residuales. Su colega propone analizar un gráfico de dispersión de los residuales $\hat{\varepsilon}_i^2$ contra el regresor x_i y usted propone analizar un gráfico de dispersión de los residuales $\hat{\varepsilon}_i$ contra la variable de respuesta y_i. Luego:

    a) Su colega tiene razón y usted está incorrecto.

    b) Su colega está incorrecto y usted tiene razón.

    c) Ambos, usted y su colega, están en lo correcto.

    d) Ambos, usted y su colega, están incorrectos.

DOS PREGUNTAS. Los cambios en el índice de Volumen de la Industria Manufacturera Man_t se pueden explicar a partir de los cambios en la producción PIB_t de un país. Por ello se estima por mínimos cuadrados ordinarios el siguiente modelo, utilizando información trimestral desde el primer trimestre de 1980 al cuarto trimestre de 1992, inclusive:

$$\widehat{Man}_t = \underset{(74.3642)}{-281.1785} + \underset{(0.0147)}{0.2734} PIB_t$$

Además se sabe que $SRC = 1,121$ y las medias muestrales de las variables son $\overline{Man} = 447$ y $\overline{PIB} = 5,900$. Si el último dato de PIB_t es $5,955$ a precios constantes de 1980 y el gobierno notifica que entre sus criterios de política económica está el que el PIB crezca un 2 %, entonces:

13. **(8 PUNTOS).** El impacto del crecimiento del 2 % en el PIB sobre la manufactura es aproximadamente de:

    a) 447.00

    b) $6,074.00$

    c) $1,379.79$

    d) Ninguna de las anteriores

14. **(8 PUNTOS).** El intervalo de confianza al 95 % para dicho impacto es aproximadamente igual a:

    a) $(445.976, 448.023)$

    b) $(6069.652, 6078.574)$

    c) $(1374.418, 1385.012)$

    d) $(1370.196, 1389.221)$

15. **(3 PUNTOS).** Suponiendo un modelo de regresión lineal $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ para $i = 1, \cdots, n$. con un término estocástico ε_i independiente e idénticamente distribuido como una normal $\mathrm{N}(0, \sigma_\varepsilon^2)$, considere las siguientes dos afirmaciones:

    - (1): Dado $x_i = x_0$, entonces $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_0$ es un predictor insesgado de $E[y_0|x_0]$.
    - (2): La longitud del intervalo de predicción es mínima cuando $x_0 = \bar{x}$.

    Luego,

    a) La afirmación (1) es verdadera pero la afirmación (2) es falsa.

    b) La afirmación (1) es falsa pero la afirmación (2) es verdadera.

    c) Ambas afirmaciones (1) y (2) son falsas.

    d) Ambas afirmaciones (1) y (2) son verdaderas.

Una especificación de Curvas de Engel para gasto en alimentación establecería una relación entre dicho gasto y el gasto total. Un grupo de economistas dispone de datos nacionales y tratan de estimar una Curva de Engel para la alimentación utilizando las siguientes variables:

Variable	Descripción
LAL	Logaritmo natural del gasto anual en alimentación
LGT	Logaritmo natural del gasto anual del hogar
LY	Logaritmo natural del ingreso disponible del hogar (esta variable tiene una correlación positiva y muy alta con LGT)
TAM	Número de miembros del hogar (excluidos los cónyuges)
TAM^2	Número de miembros del hogar (excluidos los cónyuges) al cuadrado
$EDAD$	Edad del marido

El grupo de economistas decide utilizar el siguiente modelo empírico:

$$LAL_i = \beta_0 + \beta_1 LGT_i + \beta_2 TAM_i + \beta_3 TAM_i^2 + \beta_4 EDAD_i + \varepsilon_i$$

16. **(3 PUNTOS).** Suponga que los residuales del modelo anterior cumplen todos los supuestos del modelo clásico de regresión lineal excepto el de homocedasticidad. Considere las siguientes afirmaciones:

- (1): Los estimadores $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ y $\hat{\beta}_4$ no son insesgados.
- (2): Los estimadores $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ y $\hat{\beta}_4$ no son eficientes.
- (3): El R^2 del modelo tendrá sentido hasta que se corrija la heterocedasticidad.

Entonces:

a) Solamente la afirmación (2) es cierta.

b) Solamente las afirmaciones (2) y (3) son ciertas.

c) Las tres afirmaciones (1), (2) y (3) son ciertas.

d) Solamente la afirmación (3) es cierta.

17. **(10 PUNTOS).** Considere el siguiente modelo para una muestra de tamaño $n = 4$:

$$y_i = \beta_1 X_{1_i} + \beta_2 X_{2i} + \varepsilon_i$$

donde $\varepsilon_i \sim i.i.d.(0, \sigma_\varepsilon^2)$ para $i = 1, \cdots, 4$ donde $\bar{Y} = 1$ y además:

$$X'X = \begin{bmatrix} 12 & 4 \\ 4 & 2 \end{bmatrix}, X'Y = \begin{bmatrix} 12 \\ 4 \end{bmatrix}$$

También se obtuvo que: $\hat{\vec{\varepsilon}}' = (1, -1, 0, 0)$, $\hat{\vec{y}}' = (1, 1, 3, -1)$. Analice las siguientes afirmaciones para elegir la opción correcta:

- (1): $\hat{\vec{\beta}}' = (1, 0)$
- (2) $R^2 = 0.80$

Entonces, es cierto que:

a) La afirmación (1) es correcta pero la afirmación (2) es incorrecta.

b) Ambas afirmaciones, (1) y (2), son correctas.

c) Ambas afirmaciones, (1) y (2), son incorrectas.

d) La afirmación (1) es incorrecta pero la afirmación (2) es correcta.

—

TRES PREGUNTAS. Suponga que se quiere estudiar la relación entre desigualdad, medida como la concentración del ingreso y_t, y las siguientes medidas de la libertad económica general:

- Si el derecho de propiedad está protegido: X_1
- Si tienen precios estables: X_2
- Si las barreras al comercio son pocas: X_3
- Si existen otras regulaciones: X_4

Un grupo de economistas trabajan sobre esta relación con datos de Dinamarca para el periodo de 1922 a 2010 por lo que se utilizaron $T = 89$ observaciones anuales. Los economistas estimaron la siguiente regresión lineal

$$y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \beta_3 X_{3,t} + \beta_4 X_{4,t} + \beta_5 y_{t-1} + u_t$$

y obtienen los resultados que se muestran en el Cuadro 1:

Cuadro 1: Resultados de la regresión

	$\hat{\beta}_i$	$e.s.(\hat{\beta}_i)$	$t_{\hat{\beta}_i}$	valor p
Intercepto	-1.5017	0.7982	-1.8810	0.0634*
X_1	0.0590	0.0455	1.2950	0.1988
X_2	0.0526	0.0705	0.7463	0.4576
X_3	0.0547	0.0472	1.1600	0.2492
X_4	-0.0324	0.0765	-0.4243	0.6725
y_{t-1}	1.0293	0.0265	38.8800	0.0000***
$\bar{y}_t$	8.6421		σ_y	3.2071
SRC	13.1278		σ_ε	0.3998
R^2	0.9854		$\bar{R}^2$	0.9846
$F_{(5;83)}$	1,127.938		$p(F)$	0.0000***
Nota:				*p<0.1; **p<0.05; ***p<0.01

Los resultados de las principales pruebas de diagnóstico para los residuales estimados del modelo se presentan en el siguiente cuadro:

Cuadro 2: Pruebas de diagnóstico de los residuales

Prueba	Estadístico	g. de l.	valor p
Jarque-Bera	$JB = 15.494$	2	0.0004
White	$nR^2 = 32.3966$	20	0.0392
Breusch-Pagan	$nR^2 = 28.2221$	5	0.0000
Breusch-Godfrey	$nR^2 = 0.0029$	1	0.9570
Durbin-Watson	$DW = 1.9219$	5	0.1973

18. **(3 PUNTOS).** A partir del análisis de los residuales del modelo, podemos asegurar que:

    a) Rechazamos que los residuales estimados se distribuyen de forma normal

    b) No rechazamos que la distribución de los residuales sea simétrica y tenga curtosis igual a 3

    c) Los parámetros del modelo β_i para $i = 0, \cdots, 5$ serán sesgados

    d) Ninguna de las anteriores

19. **(3 PUNTOS).** A partir del análisis de los residuales del modelo, podemos asegurar que:

    a) Rechazamos la hipótesis de homoscedasticidad dada una probabilidad de error tipo 1 $\alpha = 0.05$

    b) Rechazamos la hipótesis de heteroscedasticidad dada una probabilidad de error tipo 1 $\alpha = 0.05$

    c) Los parámetros del modelo β_i para $i = 0, \cdots, 5$ resultarán sesgados

    d) Tanto b) como c) son ciertas

20. **(3 PUNTOS).** A partir del análisis de los residuales del modelo, podemos asegurar que:

    a) Rechazamos la hipótesis de autocorrelación dada una probabilidad de error tipo 1 $\alpha = 0.05$

    b) Rechazamos la hipótesis de no autocorrelación dada una probabilidad de error tipo 1 $\alpha = 0.05$

    c) Los parámetros del modelo β_i para $i = 0, \cdots, 5$ resultarán sesgados

    d) Tanto b) como c) son ciertas

2. Examen Final (Otoño 2019) - Solución

Econometría
7 de diciembre de 2019

Nombre: _____ CU: _____

INSTRUCCIONES:

1. Lea con cuidado cada pregunta antes de responderla.
2. Marque la respuesta correcta para cada pregunta en la hoja de respuestas anexa. Tome en cuenta que para cada pregunta sólo existe una única respuesta correcta.
3. Tome su tiempo para responder cada pregunta. Cuenta con tiempo suficiente para realizar el examen. No hay prisa.
4. Este es un examen a libro cerrado. Cualquier material escrito o electrónico adicional a este examen está prohibido. No es necesario ningún formulario.
5. No desengrape su examen. Puede utilizar las hojas blancas adicionales para sus cálculos pero no se calificará procedimiento alguno.
6. Se requiere de una calculadora científica y de las tablas estadísticas anexas. El uso de cualquier otro tipo de equipo así como de teléfonos móviles está estrictamente prohibido. No los use.
7. No hagas trampa, no copies, no robes ideas y no permitas que otros lo hagan. Cualquier conducta fraudulenta derivará en las sanciones especificadas en el reglamento del ITAM.
8. Cualquier prueba estadística o intervalo de confianza, utilice el 95 % de confianza. Los números entre paréntesis abajo de cualquier estimador se refieren a su error estándar. El término ε siempre representa el término estocástico de la regresión.

1. **(3 PUNTOS).** Un ingeniero investiga el uso de un molino de viento para generar electricidad para lo cual cuenta con n observaciones de corriente directa producida por m molinos. El modelo teórico plantea que la energía producida por los molinos medida en Megawatts (Y_{it}) depende de la velocidad del viento correspondiente medida en Millas por Hora (X_{it}). Como un primer paso en el modelado empírico se proponen dos modelos de regresión lineal simple:

   - Modelo 1: $y_{it} = \beta_0 + \beta_1 x_{it} + \varepsilon_{it} \ \forall \ i = 1, \cdots, n; \ t = 1, \cdots, m$
   - Modelo 2: $y_i = \beta_0 + \beta_1 \frac{1}{x_i} + \varepsilon_i \ \forall \ i = 1, \cdots, n$

   Entonces, se puede asegurar que:

   a) El modelo 1 emplea datos de tipo transversal y el modelo 2 datos de tipo longitudinal (serie de tiempo).

   b) El modelo 1 emplea datos de tipo panel y el modelo 2 datos de tipo longitudinal (serie de tiempo).

   c) El modelo 1 emplea datos de tipo longitudinal (serie de tiempo) y el modelo 2 datos de tipo transversal.

   d) El modelo 1 emplea datos de tipo panel y el modelo 2 datos de tipo transversal.

2. **(5 PUNTOS).** Un investigador propone un modelo lineal que busca explicar el comportamiento de una variable Y mediante el comportamiento de una variable X. En primera instancia, se ajusta el modelo utilizando ambas variables medidas en miles de pesos con lo que se estima a β_0 y a β_1 mediante los estimadores de mínimos cuadrados ordinarios $\hat{\beta}_0$ y $\hat{\beta}_1$. Después, ajusta nuevamente el modelo considerando las unidades de la variable dependiente en pesos pero sin alterar las unidades de la variable independiente, y obtiene estimaciones $\hat{\beta}_0^\star$ y $\hat{\beta}_1^\star$. Entonces se puede asegurar que:

   a) $\hat{\beta}_0^\star = \hat{\beta}_0$ y $\hat{\beta}_1^\star = \hat{\beta}_1$

   b) $\hat{\beta}_0^\star = 1000\hat{\beta}_0$ y $\hat{\beta}_1^\star = 1000\hat{\beta}_1$

   c) $\hat{\beta}_0^\star = \frac{1}{1000}\hat{\beta}_0$ y $\hat{\beta}_1^\star = \frac{1}{1000}\hat{\beta}_1$

   d) Ninguna de las anteriores.

3. **(5 PUNTOS).** En el contexto de un modelo de regresión lineal simple de la forma: $y_i = \theta_0 + \theta_1 x_i + \varepsilon_i$, suponga los siguientes modelos:

   - Modelo 1: $y = \beta_0(1 - e^{\beta_1 x})$
   - Modelo 2: $y = \frac{x}{\beta_0 + \beta_1 x}$

   Si se consideran transformaciones y/o cambios de variable para ajustar un modelo de regresión lineal simple, entonces:

   a) Ambos modelos (1) y (2) son linearizables.
   b) El modelo (1) es linearizable pero (2) no lo es.
   c) El modelo (1) no es linearizable pero (2) si lo es.
   d) Ni el modelo (1) ni el modelo (2) son linearizables

4. **(3 PUNTOS).** Suponiendo el modelo de regresión lineal $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ donde se satisfacen los supuestos usuales para el término estocástico ε_i, esto es:

$$E\left[\varepsilon_i | x_i\right] = 0;\ Var[\varepsilon_i | x_i] = \sigma_\varepsilon^2;\ Cov[\varepsilon_i, \varepsilon_j | x_i] = 0\ \forall\ i \neq j$$

   Entonces es cierto que:

   a) $\sum_{i=1}^n \hat{y}_i \hat{\varepsilon}_i = 0$
   b) $\sum_{i=1}^n x_i \hat{\varepsilon}_i = 0$
   c) $\sum_{i=1}^n y_i = \sum_{i=1}^n \hat{y}_i$
   d) Todas las anteriores.

5. **(5 PUNTOS).** Sea un modelo de regresión lineal simple de la forma $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ donde el supuesto de que $Var[\varepsilon_i | x_i] = \sigma_\varepsilon^2$ no se satisface pero que los otros dos supuestos de Gauss Markov, $E[\varepsilon_i | x_i] = 0$ y $Cov[\varepsilon_i, \varepsilon_j | x_i] = 0$ sí se satisfacen. Considere las siguientes afirmaciones:

   - (1): $\hat{\beta}_0$ y $\hat{\beta}_1$ son insesgados.
   - (2): $Var(\hat{\beta}_0) = \sigma_\varepsilon^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right)$ y $Var(\hat{\beta}_1) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$.

   Luego, se tiene que:

   a) La afirmación (1) es verdadera pero la afirmación (2) es falsa.
   b) La afirmación (1) es falsa pero la afirmación (2) es verdadera.
   c) Ambas afirmaciones (1) y (2) son falsas.
   d) Ambas afirmaciones (1) y (2) son verdaderas.

6. **(3 PUNTOS).** Suponga que se sabe que el efecto marginal de x_i sobre y_i es negativo y no constante. Entonces, un modelo adecuado sería:

   a) $y_i = \beta_0 + \beta_1 x_i + \beta_2 ln(x_i) + \varepsilon_i$

   b) $y_i = \beta_0 + \beta_1 x_i + \beta_2 \frac{1}{\sqrt{x_i}} + \varepsilon_i$ con $\beta_1 < 0$ y $\beta_2 > 0$

   c) $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ con $\beta_1 < 0$

   d) $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i$ con $\beta_2 < 0$

7. **(8 PUNTOS).** Suponga que se desea probar que el puntaje promedio obtenido en la prueba ENLACE por los niños es diferente al obtenido por las niñas. Para ello se estiman los siguientes modelos, con el puntaje estandarizado obtenido por cada alumno en la prueba $ENLACE$ variable dependiente y el promedio de sus calificaciones en el año escolar ($PROM$), la calificación global de la escuela ESC y el número de horas de estudio $HREST$ como variables independientes:

   Para los 366 alumnos de la escuela, con un $R^2 = 0.3516$ y $STC = 131.8936$:

   $$ENLACE_i = \underset{(0.1837)}{1.4909} + \underset{(0.0002)}{0.0012} PROM_i - \underset{(0.0012)}{0.0099} ESC_i + \underset{(0.0008)}{0.0023} HREST_i$$

   Para las 90 niñas de la escuela, con un $R^2 = 0.4014$ y $STC = 32.7493$:

   $$ENLACE_i = \underset{(0.3616)}{1.1273} + \underset{(0.0003)}{0.0018} PROM_i - \underset{(0.0029)}{0.0090} ESC_i + \underset{(0.0022)}{0.0014} HREST_i$$

   Para los 276 niños de la escuela, con un $R^2 = 0.3169$ y $STC = 86.0014$:

   $$ENLACE_i = \underset{(0.2060)}{1.4808} + \underset{(0.0002)}{0.0011} PROM_i - \underset{(0.0014)}{0.0085} ESC_i + \underset{(0.0008)}{0.0023} HREST_i$$

   Luego de realizar la prueba estadística adecuada, puede concluir que:

   a) El modelo de los niños no es diferente al de las niñas por lo que se necesita incluir la variable dicotómica de género del alumno.

   b) El modelo de los niños es diferente al de las niñas por lo que se necesita incluir la variable dicotómica de género del alumno.

   c) El modelo de los niños no es diferente al de las niñas por lo que no se necesita incluir la variable dicotómica de género del alumno.

   d) El modelo de los niños es diferente al de las niñas, por lo que no se necesita incluir la variable dicotómica de género del alumno.

8. **(8 PUNTOS).** Se sabe que una función de producción del tipo Cobb-Douglas dada por $Y_i = CK_i^{\beta_1} L_i^{\beta_2}$ muestra rendimientos constantes a escala si se cumple que $\beta_1 + \beta_2 = 1$. En caso de que sean crecientes o decrecientes, la suma de estos coeficientes será mayor o menor a uno, respectivamente. Suponga utilizando 33 observaciones se estima el siguiente modelo lineal:

$$\widehat{lnY_i} = \underset{(0.2975)}{-6.3773} + \underset{(0.0705)}{0.5709} lnK_i + \underset{(0.1006)}{2.0786} lnL_i$$

con $Cov(\hat{\beta}_1, \hat{\beta}_2) = -0.0056$, $R^2 = 0.9827$ y $SRC = 0.1394$. Entonces puede asegurar que la función de producción estimada:

a) Presenta rendimientos constantes a escala.

b) Presenta rendimientos crecientes a escala.

c) Presenta rendimientos decrecientes a escala.

d) No se pueden determinar sus rendimientos a escala.

9. **(5 PUNTOS).** Suponga que al estimar el modelo

$$S_t = \beta_0 + \beta_1 A_t + \beta_2 P_t + \beta_3 E_t + \beta_4 A_{t-1} + \beta_5 P_{t-1} + \varepsilon_t$$

se obtienen los siguientes resultados:

	$\hat{\beta}$	$e.s.(\hat{\beta})$	t	valor p	FIV
Intercepto	-14.194	18.715	-0.758	0.4592	
A_t	5.361	4.028	1.331	0.2019	36.94
P_t	8.372	3.586	2.334	0.0329	33.47
E_t	22.521	2.142	10.512	0.0000	1.07
A_{t-1}	3.855	3.578	1.077	0.2973	25.91
P_{t-1}	4.125	3.895	1.059	0.3053	43.52

Entonces, el coeficiente de determinación R^2 que se obtiene al ajustar A_{t-1} con respecto al resto de los regresores A_t, P_t, E_t, P_{t-1} está en el intervalo:

a) $[0, 0.25)$

b) $[0.25, 0.50)$

c) $[0.50, 0.75)$

d) $[0.75, 1.0)$

10. **(8 PUNTOS).** Para analizar la discriminación salarial en una empresa, se estima el siguiente modelo:

$$\widehat{ln(SAL_i)} = 1.5 + 0.08ln(ED_i) + 0.05ln(EX_i) + 0.03H_i + 0.02P_i + 0.03H_iP_i$$

donde SAL_i es el salario del trabajador, ED_i y EX_i su educación y experiencia en años, respectivamente, H_i una variable dicotómica que toma el valor de 1 si el individuo es hombre y 0 si es mujer, y P_i una variable dicotómica que toma el valor de 1 si el individuo posee un posgrado y 0 si no. Todos los coeficientes estimados fueron significativos al 95 % de confianza. Entonces, se puede asegurar que:

a) Una mujer con posgrado gana mas que un hombre sin posgrado.

b) Una mujer con posgrado gana menos que una mujer sin posgrado.

c) Un individuo con posgrado siempre gana más, sin importar su género.

d) Un hombre gana más que una mujer pero sólo si no tiene posgrado.

11. **(3 PUNTOS).** Considere las siguientes afirmaciones en relación al análisis de la multicolinealidad en un modelo de regresión lineal múltiple:

- (1): Si en el modelo lineal con 7 variables regresoras, la variable cualitativa X_2 con 4 categorías, se codifica mediante cuatro variables dicotómicas D_1, D_2, D_3 y D_4, y se incluyen en un modelo regresión lineal con un intercepto β_0, entonces el modelo presentará necesariamente problemas de multicolinealidad.
- (2) Si la correlación de la variable dependiente y con uno de los regresores del modelo, digamos la variable x_1 es mayor a 0.9, entonces necesariamente se tiene un problema de multicolinealidad.

Entonces, es cierto que:

a) La afirmación (1) es falsa mientras que la afirmación (2) es correcta.

b) La afirmación (1) es correcta mientras que la afirmación (1) es falsa.

c) Ambas afirmaciones, la (1) y la (2), son son falsas.

d) Ambas afirmaciones, la (1) y la (2), son son correctas.

12. **(3 PUNTOS).** Usted y un colega ajustan un modelo de regresión lineal simple donde, para validar el supuesto de homoscedasticidad para la varianza del error se procede a realizar un análisis de residuales. Su colega propone analizar un gráfico de dispersión de los residuales $\hat{\varepsilon}_i^2$ contra el regresor x_i y usted propone analizar un gráfico de dispersión de los residuales $\hat{\varepsilon}_i$ contra la variable de respuesta y_i. Luego:

    a) Su colega tiene razón y usted está incorrecto.

    b) Su colega está incorrecto y usted tiene razón.

    c) Ambos, usted y su colega, están en lo correcto.

    d) Ambos, usted y su colega, están incorrectos.

DOS PREGUNTAS. Los cambios en el índice de Volumen de la Industria Manufacturera Man_t se pueden explicar a partir de los cambios en la producción PIB_t de un país. Por ello se estima por mínimos cuadrados ordinarios el siguiente modelo, utilizando información trimestral desde el primer trimestre de 1980 al cuarto trimestre de 1992, inclusive:

$$\widehat{Man}_t = \underset{(74.3642)}{-281.1785} + \underset{(0.0147)}{0.2734} PIB_t$$

Además se sabe que $SRC = 1,121$ y las medias muestrales de las variables son $\overline{Man} = 447$ y $\overline{PIB} = 5,900$. Si el último dato de PIB_t es $5,955$ a precios constantes de 1980 y el gobierno notifica que entre sus criterios de política económica está el que el PIB crezca un 2 %, entonces:

13. **(8 PUNTOS).** El impacto del crecimiento del 2 % en el PIB sobre la manufactura es aproximadamente de:

    a) 447.00

    b) $6,074.00$

    c) $1,379.79$

    d) Ninguna de las anteriores

14. **(8 PUNTOS).** El intervalo de confianza al 95 % para dicho impacto es aproximadamente igual a:

    a) $(445.976, 448.023)$

    b) $(6069.652, 6078.574)$

    c) $(1374.418, 1385.012)$

    d) $(1370.196, 1389.221)$

15. **(3 PUNTOS).** Suponiendo un modelo de regresión lineal $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ para $i = 1, \cdots, n$. con un término estocástico ε_i independiente e idénticamente distribuido como una normal $\mathrm{N}(0, \sigma_\varepsilon^2)$, considere las siguientes dos afirmaciones:

- (1): Dado $x_i = x_0$, entonces $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_0$ es un predictor insesgado de $E[y_0|x_0]$.
- (2): La longitud del intervalo de predicción es mínima cuando $x_0 = \bar{x}$.

Luego,

a) La afirmación (1) es verdadera pero la afirmación (2) es falsa.

b) La afirmación (1) es falsa pero la afirmación (2) es verdadera.

c) Ambas afirmaciones (1) y (2) son falsas.

d) Ambas afirmaciones (1) y (2) son verdaderas.

Una especificación de Curvas de Engel para gasto en alimentación establecería una relación entre dicho gasto y el gasto total. Un grupo de economistas dispone de datos nacionales y tratan de estimar una Curva de Engel para la alimentación utilizando las siguientes variables:

Variable	Descripción
LAL	Logaritmo natural del gasto anual en alimentación
LGT	Logaritmo natural del gasto anual del hogar
LY	Logaritmo natural del ingreso disponible del hogar (esta variable tiene una correlación positiva y muy alta con LGT)
TAM	Número de miembros del hogar (excluidos los cónyuges)
TAM^2	Número de miembros del hogar (excluidos los cónyuges) al cuadrado
$EDAD$	Edad del marido

El grupo de economistas decide utilizar el siguiente modelo empírico:

$$LAL_i = \beta_0 + \beta_1 LGT_i + \beta_2 TAM_i + \beta_3 TAM_i^2 + \beta_4 EDAD_i + \varepsilon_i$$

16. **(3 PUNTOS).** Suponga que los residuales del modelo anterior cumplen todos los supuestos del modelo clásico de regresión lineal excepto el de homocedasticidad. Considere las siguientes afirmaciones:

- (1): Los estimadores $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ y $\hat{\beta}_4$ no son insesgados.
- (2): Los estimadores $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ y $\hat{\beta}_4$ no son eficientes.
- (3): El R^2 del modelo tendrá sentido hasta que se corrija la heterocedasticidad.

Entonces:

a) Solamente la afirmación (2) es cierta.

b) Solamente las afirmaciones (2) y (3) son ciertas.

c) Las tres afirmaciones (1), (2) y (3) son ciertas.

d) Solamente la afirmación (3) es cierta.

17. **(10 PUNTOS).** Considere el siguiente modelo para una muestra de tamaño $n = 4$:

$$y_i = \beta_1 X_{1_i} + \beta_2 X_{2i} + \varepsilon_i$$

donde $\varepsilon_i \sim i.i.d.(0, \sigma_\varepsilon^2)$ para $i = 1, \cdots, 4$ donde $\bar{Y} = 1$ y además:

$$X'X = \begin{bmatrix} 12 & 4 \\ 4 & 2 \end{bmatrix}, X'Y = \begin{bmatrix} 12 \\ 4 \end{bmatrix}$$

También se obtuvo que: $\hat{\vec{\varepsilon}}' = (1, -1, 0, 0)$, $\hat{\vec{y}}' = (1, 1, 3, -1)$. Analice las siguientes afirmaciones para elegir la opción correcta:

- (1): $\hat{\vec{\beta}}' = (1, 0)$
- (2) $R^2 = 0.80$

Entonces, es cierto que:

a) La afirmación (1) es correcta pero la afirmación (2) es incorrecta.

b) Ambas afirmaciones, (1) y (2), son correctas.

c) Ambas afirmaciones, (1) y (2), son incorrectas.

d) La afirmación (1) es incorrecta pero la afirmación (2) es correcta.

—

TRES PREGUNTAS. Suponga que se quiere estudiar la relación entre desigualdad, medida como la concentración del ingreso y_t, y las siguientes medidas de la libertad económica general:

- Si el derecho de propiedad está protegido: X_1
- Si tienen precios estables: X_2
- Si las barreras al comercio son pocas: X_3
- Si existen otras regulaciones: X_4

Un grupo de economistas trabajan sobre esta relación con datos de Dinamarca para el periodo de 1922 a 2010 por lo que se utilizaron $T = 89$ observaciones anuales. Los economistas estimaron la siguiente regresión lineal

$$y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \beta_3 X_{3,t} + \beta_4 X_{4,t} + \beta_5 y_{t-1} + u_t$$

y obtienen los resultados que se muestran en el Cuadro 1:

Cuadro 3: Resultados de la regresión

	$\hat{\beta}_i$	$e.s.(\hat{\beta}_i)$	$t_{\hat{\beta}_i}$	valor p
Intercepto	-1.5017	0.7982	-1.8810	0.0634*
X_1	0.0590	0.0455	1.2950	0.1988
X_2	0.0526	0.0705	0.7463	0.4576
X_3	0.0547	0.0472	1.1600	0.2492
X_4	-0.0324	0.0765	-0.4243	0.6725
y_{t-1}	1.0293	0.0265	38.8800	0.0000***
$\bar{y}_t$	8.6421		σ_y	3.2071
SRC	13.1278		σ_ε	0.3998
R^2	0.9854		$\bar{R}^2$	0.9846
$F_{(5;83)}$	1,127.938		$p(F)$	0.0000***
Nota:				*p<0.1; **p<0.05; ***p<0.01

Los resultados de las principales pruebas de diagnóstico para los residuales estimados del modelo se presentan en el siguiente cuadro:

Cuadro 4: Pruebas de diagnóstico de los residuales

Prueba	Estadístico	g. de l.	valor p
Jarque-Bera	$JB = 15.494$	2	0.0004
White	$nR^2 = 32.3966$	20	0.0392
Breusch-Pagan	$nR^2 = 28.2221$	5	0.0000
Breusch-Godfrey	$nR^2 = 0.0029$	1	0.9570
Durbin-Watson	$DW = 1.9219$	5	0.1973

18. **(3 PUNTOS).** A partir del análisis de los residuales del modelo, podemos asegurar que:

    a) Rechazamos que los residuales estimados se distribuyen de forma normal

    b) No rechazamos que la distribución de los residuales sea simétrica y tenga curtosis igual a 3

    c) Los parámetros del modelo β_i para $i = 0, \cdots, 5$ serán sesgados

    d) Ninguna de las anteriores

19. **(3 PUNTOS).** A partir del análisis de los residuales del modelo, podemos asegurar que:

    a) Rechazamos la hipótesis de homoscedasticidad dada una probabilidad de error tipo 1 $\alpha = 0.05$

    b) Rechazamos la hipótesis de heteroscedasticidad dada una probabilidad de error tipo 1 $\alpha = 0.05$

    c) Los parámetros del modelo β_i para $i = 0, \cdots, 5$ resultarán sesgados

    d) Tanto b) como c) son ciertas

20. **(3 PUNTOS).** A partir del análisis de los residuales del modelo, podemos asegurar que:

    a) Rechazamos la hipótesis de autocorrelación dada una probabilidad de error tipo 1 $\alpha = 0.05$

    b) Rechazamos la hipótesis de no autocorrelación dada una probabilidad de error tipo 1 $\alpha = 0.05$

    c) Los parámetros del modelo β_i para $i = 0, \cdots, 5$ resultarán sesgados

    d) Tanto b) como c) son ciertas

3. Examen Final (Otoño 2020)

Econometría
10 de diciembre de 2020

Nombre: _____ CU: _____

INSTRUCCIONES:

A continuación, se presentan 14 preguntas de opción múltiple (no se calificará procedimiento). Por favor, cualquier cálculo numérico se debe desarrollar con tres decimales de precisión.

1. **(4 puntos)**. Considere el modelo de regresión lineal simple:

$$Y_i = \beta_0 + \beta_1 X_i + u_i \text{ con } u_i \sim \text{i.i.d. } N(0, \sigma^2 X_i) \text{ para } i = 1, \cdots, n$$

   Sea $\hat{\beta}_1$ el estimador de mínimos cuadrados ordinarios para β_1. Entonces es cierto que:

   a) $\hat{\beta}_1$ es insesgado y eficiente
   b) $\hat{\beta}_1$ es sólo insesgado
   c) $\hat{\beta}_1$ es sólo eficiente
   d) $\hat{\beta}_1$ no es ni insesgado ni eficiente

2. **(6 puntos)**. Suponga que se estiman los siguientes modelos de regresión lineal simple mediante el método de mínimos cuadrados ordinarios.

   - **Modelo 1**: $ln(Y_i) = \beta_0 + \beta_1 ln(X_i) + \varepsilon_i$, $\varepsilon_i \sim$ i.i.d. $N(0, \sigma_\varepsilon^2)$, $i = 1, \cdots, n$
   - **Modelo 2**: $ln\left(\frac{Y_i}{X_i}\right) = \alpha_0 + \alpha_1 ln(X_i) + \nu_i$, $\nu_i \sim$ i.i.d. $N(0, \sigma_\nu^2)$, $i = 1, \cdots, n$

   Entonces, los estimadores de los parámetros de estas regresiones mediante el método de mínimos cuadrados ordinarios entre sí cumplen que:

   a) $\hat{\beta}_0 \neq \hat{\alpha}_0$, $\hat{\beta}_1 \neq \hat{\alpha}_1$ y $\hat{\varepsilon}_i \neq \hat{\nu}_i$
   b) $\hat{\beta}_0 = \hat{\alpha}_0$, $\hat{\beta}_1 - 1 = \hat{\alpha}_1$ y $\hat{\varepsilon}_i = \hat{\nu}_i$
   c) $\hat{\beta}_0 \neq \hat{\alpha}_0$, $\hat{\beta}_1 = \hat{\alpha}_1$ y $\hat{\varepsilon}_i = \hat{\nu}_i$
   d) $\hat{\beta}_0 = \hat{\alpha}_0$, $\hat{\beta}_1 \neq \hat{\alpha}_1$ y $\hat{\varepsilon}_i \neq \hat{\nu}_i$

3. **(6 puntos)**. Considera un modelo de regresión lineal múltiple:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t,\ \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } t = 1, 2, \cdots, 30$$

con datos de series de tiempo de la variable endógena Y_t= Ventas de automóviles y las variables exógenas X_{1t} = Precio y X_{2t} = Año de fabricación. Después de estimar los parámetros del modelo por el método de mínimos cuadrados ordinarios, se obtuvo el estadístico de Durbin-Watson $d = 1.27$. Mientras que la tabla respectiva de valores críticos para un tamaño de muestra $n = 30$, brindó los valores $d_L = 1.284$ y $d_U = 1.567$ al nivel de significancia del 5 %, y $d_L = 1.070$ y $d_U = 1.339$ al nivel de significancia del 1 %. Como resultado se concluye que:

   a) No existe autocorrelación positiva, al nivel de significancia del 1 %
   b) Existe autocorrelación negativa, al nivel de significancia del 1 %
   c) No se puede concluir si existe autocorrelación negativa, al nivel de significancia del 5 %
   d) Existe autocorrelación positiva, al nivel de significancia del 5 %

4. **(4 puntos)**. Considera el siguiente modelo de regresión lineal simple

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \text{ con } \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, 10$$

en el que Y_i= Salario mensual por empleado y X_i= Productividad promedio en la industria de manufactura de bienes no-duraderos, con cifras en miles de pesos. Los datos corresponden a industrias manufactureras mexicanas de bienes no-duraderos en 2010, clasificados por el número de empleados en la industria o establecimiento. Los resultados de la estimación del modelo por el método de mínimos cuadrados ordinarios fueron:

$$\widehat{Y}_i = \underset{(936.479)}{1992.345} + \underset{(0.099)}{0.232} X_i \text{ con } R^2 = 0.437$$

Los valores entre paréntesis son las desviaciones estándar de los parámetros estimados. Para verificar un supuesto sobre los errores del modelo, se estimó la siguiente regresión auxiliar:

$$\widehat{\varepsilon_i}^2 = \underset{(38.319)}{35.817} - \underset{(4.216)}{2.8099} ln(X_i) \text{ con } R^2 = 0.059$$

Los datos entre paréntesis son las desviaciones estándar de los parámetros estimados. Indica cuál supuesto se intenta verificar con la regresión auxiliar y cuál es la conclusión que se alcanza.

   a) Supuesto de no-correlación en los errores y conclusión de que hay autocorrelación
   b) Supuesto de homoscedasticidad de los errores y como conclusión se valida el supuesto
   c) Supuesto de linealidad en la relación funcional y conclusión de que existe curvatura
   d) Supuesto de no multicolinealidad y conclusión de que el supuesto es válido

5. **(12 puntos)**. Su compañera Alicia propone la siguiente función de demanda de pollos:

$$ln(y_i) = \beta_0 + \beta_1 ln(x_{1i}) + \beta_2 ln(x_{2i}) + \beta_3 ln(x_{3i}) + \beta_4 ln(x_{4i}) + \varepsilon_i$$

con $\varepsilon_i \sim$ i.i.d. $N(0, \sigma^2)$ y donde y_i = consumo de pollo per-cápita, x_{1i}= ingreso real disponible per-cápita en dólares, x_{2i}= precio real del pollo por kg en dólares, x_{3i} = precio real del cerdo por kg en dólares, y x_{4i}= precio real de la carne de res por kg. en dólares. Usando información anual sobre la demanda de pollo durante el periodo de 1990-2012, Alicia estimó el siguiente modelo mediante el método de mínimos cuadrados ordinarios:

$$\widehat{ln(y_i)} = \underset{(0.1557)}{2.189} + \underset{(0.0833)}{0.343}\, ln(x_{1i}) - \underset{(0.1109)}{0.505}\, ln(x_{2i}) + \underset{(0.0977)}{0.149}\, ln(x_{3i}) + \underset{(0.0997)}{0.911}\, ln(x_{4i})$$

con $R^2 = 0.982$. Sin embargo, su compañero Pedro, experto en el mercado de venta de pollo al menudeo afirma que la verdadera función de demanda es la siguiente:

$$\widehat{ln(y_i)} = \underset{(0.1162)}{2.033} + \underset{(0.0247)}{0.452}\, ln(x_{1i}) - \underset{(0.0635)}{0.377}\, ln(x_{2i})$$

con $R^2 = 0.980$. En ambos modelos, los valores entre paréntesis se refieren a la desviación estándar de los parámetros estimados. Evalúe las siguientes afirmaciones para identificar la opción correcta. Evalúe las siguientes afirmaciones para identificar la opción correcta.

I) La Suma de Cuadrados Totales es la misma en los dos modelos estimados

II) A partir del coeficiente de determinación se pude concluir que los modelos son comparables y que el modelo estimado por Alicia ajusta mejor que el estimado por Pedro.

III) La elasticidad precio de la demanda estimada por Alicia es mayor en valor absoluto que la estimada por Pedro.

IV) El estadístico de prueba para probar la hipótesis de que la especificación del modelo de Pedro es el apropiado toma un valor aproximado de 1.00 con 2,18 grados de libertad y dado en Error de Tipo I $\alpha = 5\,\%$ se concluye que Pedro puede seguir gozando de su reputación de experto en el mercado de venta de pollo al menudeo.

Luego,

a) Las cuatro afirmaciones son verdaderas

b) Las cuatro afirmaciones son falsas

c) I) y IV) son verdaderas, pero II) y III) son falsas

d) I), II) y III) son verdaderas, pero IV) es falsa

6. **(6 puntos)**. Considera el siguiente modelo de regresión lineal simple

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t,\ \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } t = 1, 2, \cdots, 13$$

con Y_t= Salario por hora en dólares, X_t= Años de escolaridad, para t = 13 distintos años de estudio, que cubren valores $X_1 = 6$ a $X_{13} = 18$. Se obtuvieron así los siguientes resultados $\bar{X} = 12$ y $\sum_{i=1}^{13}(X_i - \bar{X})^2 = 182$ y la regresión estimada:

$$\widehat{Y_i} = \underset{(0.931)}{-0.014} + \underset{(0.070)}{0.724} X_i \text{ con } R^2 = 0.906$$

con la desviación estándar de los parámetros estimados en paréntesis y una varianza de los residuales $\hat{\sigma}^2 = 0.893$. Encuentra el pronóstico puntual y su intervalo de predicción del 95 %, correspondiente al salario por hora de un individuo con $X_0 = 20$ años de estudio.

a) El pronóstico puntual es 14.466 y por intervalo es $(12.805, 15.827)$

b) El pronóstico puntual es 12.466 y por intervalo es $(9.977, 16.959)$

c) El pronóstico puntual es 14.466 y por intervalo es $(11.984, 16.948)$

d) El pronóstico puntual es 12.466 y por intervalo es $(10.805, 13.827)$

7. **(12 puntos)**. Considere el siguiente modelo de regresión lineal múltiple:

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \text{ con } \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2)$$

con las cuatro observaciones para y_i, x_{1i} y x_{2i} que se presentan en la siguiente tabla:

y_i	x_{1i}	x_{2i}
2	1	0
0	1	0
3	3	1
-1	-1	-1

Analice las siguientes afirmaciones para elegir la opción correcta:

I) $\widehat{Var(\hat{\beta}_1 + \hat{\beta}_2)} = \frac{7}{4}$

II) $\widehat{Var(\hat{\beta}_1 + \hat{\beta}_2)} = \frac{3}{2}$

III) $\widehat{Cov(\hat{\beta}_1, \hat{\beta}_2)} = -\frac{1}{2}$

Entonces:

a) I) es verdadera, pero II) y III) son falsas

b) I) y III) son verdaderas, pero II) es falsa

c) II) y III) son verdaderas, pero I) es falsa

d) I) y III) son falsa, pero II) es verdadera

8. **(9 puntos)**. Se desea estimar el gasto público en educación $EDUC$ en función del Producto Interno Bruto PIB. Para ello, se cuenta con datos (en millones de Euros) correspondientes a 20 países. Los resultados son los siguientes:

$$\widehat{EDUC}_i = -21436 + 0.049PIB_i$$

mientras que el coeficiente de asimetría y curtosis de los residuales son iguales a -0.411 y 2.666, respectivamente. Analice las siguientes afirmaciones para elegir la opción correcta:

I) El estadístico apropiado para probar la hipótesis de normalidad toma un valor aproximadamente igual a 0.656 y dado un error de Tipo I, $\alpha = 5\,\%$, se concluye que no se cumple el supuesto de normalidad y por lo tanto no se puede realizar inferencia por desconocer su distribución exacta.

II) El estadístico apropiado para probar la hipótesis de normalidad toma un valor aproximadamente igual a 0.656 y dado un error de Tipo I, $\alpha = 90\,\%$, se concluye que se cumple el supuesto de normalidad y por lo tanto si se puede realizar inferencia.

III) El estadístico apropiado para probar la hipótesis de normalidad toma un valor aproximadamente igual a 3.50 y dado un error de Tipo I, $\alpha = 5\,\%$, se concluye que se cumple el supuesto de normalidad y por lo tanto no se puede realizar inferencia.

IV) El estadístico apropiado para probar la hipótesis de normalidad toma un valor aproximadamente igual a 3.50 y dado un error de Tipo I, $\alpha = 90\,\%$, se concluye que no se cumple el supuesto de normalidad y por lo tanto no se puede realizar inferencia por desconocer su distribución exacta.

Entonces,

a) I) y II) son verdaderas, pero III) y IV) son falsas

b) II) es verdaderas, pero I), III) y IV) son falsas

c) IV) es verdaderas, pero I), II) y III) son falsas

d) I), II), III) y IV) son falsas

9. **(10 puntos)**. Se ha estimado la función de producción Y_i con insumos de capital K_i y trabajo L_i, para una muestra aleatoria de 53 firmas:

$$\widehat{ln(Q_i)} = 1.37 + \underset{(0.257)}{0.632} ln(K_i) + \underset{(0.219)}{0.452} ln(L_i)$$

con $R^2 = 0.980$ y $\widehat{Cov(\hat{\beta}_1, \hat{\beta}_2)} = 0.055$. Los números en paréntesis son las desviaciones estándar de los parámetros estimados. Analice las siguientes afirmaciones para elegir la opción correcta:

I) El estadístico para probar la hipótesis de que las elasticidades del producto con respecto al capital y al trabajo son iguales versus que sean diferentes es aproximadamente igual a 2.857 y dado un error de Tipo I, $\alpha = 5\,\%$, se concluye que las elasticidades son diferentes

II) El intervalo de confianza del 95 % para la elasticidad del producto con respecto al capital es aproximadamente igual a $(0.116, 1.148)$

III) El estadístico para probar la hipótesis de la existencia de rendimientos constantes a escala versus rendimientos crecientes es aproximadamente igual a 0.389 y dado un error de Tipo I, $\alpha = 5\,\%$, se concluye que la función de producción de rendimientos a escala constantes

Entonces:

a) I) y II) son verdaderas, pero III) es falsa

b) I) y III) son verdaderas, pero II) es falsa

c) Las tres son verdaderas

d) Ninguna es verdadera

10. **(6 puntos)**. Suponga el modelo de regresión lineal simple:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \text{ con } \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, n$$

Entonces es cierto que:

a) $\sum_{i=1}^{n} \hat{y}_i \hat{\varepsilon}_i = 0$

b) $\sum_{i=1}^{n} x_i \hat{\varepsilon}_i = 0$

c) $\sum_{i=1}^{n} \hat{y}_i = \sum_{i=1}^{n} y_i$

d) Todas las anteriores

11. **(6 puntos)**. Considérese el siguiente modelo de regresión lineal, en donde los parámetros fueron estimados mediante el método de mínimos cuadrados ordinarios:

$$\widehat{ln(C_i)} = \underset{(0.427)}{-0.467} + \underset{(0.749)}{0.804} ln(YD_i) + \underset{(0.175)}{0.201} ln(W_i) - \underset{(0.0007)}{0.002}\, ln(R_i)$$

en donde C_i es el consumo real, YD_i es el ingreso personal disponible, W_i es la riqueza real y R_i es la tasa de interés real en una economía hipotética. Los números entre paréntesis se refieren a las desviaciones estándar de los parámetros estimados. Además, se cuenta con la siguiente información:

Regresor	$ln(YD_i)$	$ln(W_i)$	$ln(R_i)$
FIV	4	20	2.717

Entonces, el coeficiente de determinación que describe la medida de bondad de ajuste del logaritmo natural de la riqueza en términos del logaritmo natural del ingreso personal disponible y del logaritmo natural de la tasa de interés es aproximadamente igual a:

a) 0.57

b) 0.75

c) 0.95

d) 0.63

12. **(4 puntos)**. Considere el siguiente modelo de regresión lineal simple:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \text{ con } \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, n$$

Se tiene la siguiente tabla de ANOVA para el modelo estimado mediante el método de mínimos cuadrados ordinarios $\widehat{y}_i = 2627.82 - 37.15x_i$:

Fuente	Suma de Cuadrados	grados de libertad	F	valor p
$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	1,527,334.95	1	165.21	0.0000
$\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$	166,402.65	18		
$\sum_{i=1}^{n}(y_i - \bar{y})^2$	1,693,737.60	19		

Con estos datos se obtienen las siguientes conclusiones, seleccione la que es incorrecta:

a) El valor de $\hat{\sigma}^2 = 9,244.59$

b) Se rechaza la hipótesis $H_0 : \beta_1 = 0$ para un error de tipo I, $\alpha = 5\,\%$

c) La suma de cuadrados media total es: $1,536,579.64$

d) El valor de R^2 es 0.9017

13. **(5 puntos)**. En un modelo de regresión lineal múltiple en donde el modelo 1 tiene más regresores que el modelo 2, entonces siempre se cumple que:

    a) R^2 del modelo 1 es mayor que el R^2 del modelo 2
    b) $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ es mayor en el modelo 1 que en el modelo 2
    c) $\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$ es menor en el modelo 1 que en el modelo 2
    d) Las tres opciones son correctas

14. **(10 puntos)**. Considere el siguiente modelo de regresión lineal simple:

$$y_i = \beta_0 + \varepsilon_i \text{ con } E[\varepsilon_i] = 0 \text{ y } Var[\varepsilon_i] = \sigma^2 z_i$$

    Analice las siguientes afirmaciones para elegir la opción correcta:

    I) El estimador de mínimos cuadrados generalizados para β_0 es:

$$\widehat{\beta}_0^{MCG} = \frac{\sum_{i=1}^{n}(y_i/z_i^2)}{\sum_{i=1}^{n}(1/z_i^2)}$$

    II) $Var[\widehat{\beta}_0^{MCG}] = \frac{\sigma^2}{\sum_{i=1}^{n}(1/z_i^2)}$

    Luego:

    a) I) es falsa y II) es verdadera
    b) II) es falsa y I) es verdadera
    c) I) y II) son verdaderas
    d) I) y II) son falsas

4. Examen Final (Otoño 2020) - Solución

Econometría
10 de diciembre de 2020

Nombre: _____ CU: _____

INSTRUCCIONES:

A continuación, se presentan 14 preguntas de opción múltiple (no se calificará procedimiento). Por favor, cualquier cálculo numérico se debe desarrollar con tres decimales de precisión.

1. **(4 puntos)**. Considere el modelo de regresión lineal simple:

$$Y_i = \beta_0 + \beta_1 X_i + u_i \text{ con } u_i \sim \text{i.i.d. } N(0, \sigma^2 X_i) \text{ para } i = 1, \cdots, n$$

   Sea $\hat{\beta}_1$ el estimador de mínimos cuadrados ordinarios para β_1. Entonces es cierto que:

   a) $\hat{\beta}_1$ es insesgado y eficiente
   b) $\hat{\beta}_1$ es sólo insesgado
   c) $\hat{\beta}_1$ es sólo eficiente
   d) $\hat{\beta}_1$ no es ni insesgado ni eficiente

2. **(6 puntos)**. Suponga que se estiman los siguientes modelos de regresión lineal simple mediante el método de mínimos cuadrados ordinarios.

   - **Modelo 1**: $ln(Y_i) = \beta_0 + \beta_1 ln(X_i) + \varepsilon_i$, $\varepsilon_i \sim$ i.i.d. $N(0, \sigma_\varepsilon^2)$, $i = 1, \cdots, n$
   - **Modelo 2**: $ln\left(\frac{Y_i}{X_i}\right) = \alpha_0 + \alpha_1 ln(X_i) + \nu_i$, $\nu_i \sim$ i.i.d. $N(0, \sigma_\nu^2)$, $i = 1, \cdots, n$

   Entonces, los estimadores de los parámetros de estas regresiones mediante el método de mínimos cuadrados ordinarios entre sí cumplen que:

   a) $\hat{\beta}_0 \neq \hat{\alpha}_0$, $\hat{\beta}_1 \neq \hat{\alpha}_1$ y $\hat{\varepsilon}_i \neq \hat{\nu}_i$
   b) $\hat{\beta}_0 = \hat{\alpha}_0$, $\hat{\beta}_1 - 1 = \hat{\alpha}_1$ y $\hat{\varepsilon}_i = \hat{\nu}_i$
   c) $\hat{\beta}_0 \neq \hat{\alpha}_0$, $\hat{\beta}_1 = \hat{\alpha}_1$ y $\hat{\varepsilon}_i = \hat{\nu}_i$
   d) $\hat{\beta}_0 = \hat{\alpha}_0$, $\hat{\beta}_1 \neq \hat{\alpha}_1$ y $\hat{\varepsilon}_i \neq \hat{\nu}_i$

3. **(6 puntos)**. Considera un modelo de regresión lineal múltiple:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t,\ \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } t = 1, 2, \cdots, 30$$

con datos de series de tiempo de la variable endógena Y_t= Ventas de automóviles y las variables exógenas X_{1t} = Precio y X_{2t} = Año de fabricación. Después de estimar los parámetros del modelo por el método de mínimos cuadrados ordinarios, se obtuvo el estadístico de Durbin-Watson $d = 1.27$. Mientras que la tabla respectiva de valores críticos para un tamaño de muestra $n = 30$, brindó los valores $d_L = 1.284$ y $d_U = 1.567$ al nivel de significancia del 5 %, y $d_L = 1.070$ y $d_U = 1.339$ al nivel de significancia del 1 %. Como resultado se concluye que:

   a) No existe autocorrelación positiva, al nivel de significancia del 1 %
   b) Existe autocorrelación negativa, al nivel de significancia del 1 %
   c) No se puede concluir si existe autocorrelación negativa, al nivel de significancia del 5 %
   d) Existe autocorrelación positiva, al nivel de significancia del 5 %

4. **(4 puntos)**. Considera el siguiente modelo de regresión lineal simple

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \text{ con } \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, 10$$

en el que Y_i= Salario mensual por empleado y X_i= Productividad promedio en la industria de manufactura de bienes no-duraderos, con cifras en miles de pesos. Los datos corresponden a industrias manufactureras mexicanas de bienes no-duraderos en 2010, clasificados por el número de empleados en la industria o establecimiento. Los resultados de la estimación del modelo por el método de mínimos cuadrados ordinarios fueron:

$$\widehat{Y}_i = \underset{(936.479)}{1992.345} + \underset{(0.099)}{0.232} X_i \text{ con } R^2 = 0.437$$

Los valores entre paréntesis son las desviaciones estándar de los parámetros estimados. Para verificar un supuesto sobre los errores del modelo, se estimó la siguiente regresión auxiliar:

$$\widehat{\varepsilon_i}^2 = \underset{(38.319)}{35.817} - \underset{(4.216)}{2.8099} ln(X_i) \text{ con } R^2 = 0.059$$

Los datos entre paréntesis son las desviaciones estándar de los parámetros estimados. Indica cuál supuesto se intenta verificar con la regresión auxiliar y cuál es la conclusión que se alcanza.

   a) Supuesto de no-correlación en los errores y conclusión de que hay autocorrelación
   b) Supuesto de homoscedasticidad de los errores y como conclusión se valida el supuesto
   c) Supuesto de linealidad en la relación funcional y conclusión de que existe curvatura
   d) Supuesto de no multicolinealidad y conclusión de que el supuesto es válido

5. **(12 puntos)**. Su compañera Alicia propone la siguiente función de demanda de pollos:

$$ln(y_i) = \beta_0 + \beta_1 ln(x_{1i}) + \beta_2 ln(x_{2i}) + \beta_3 ln(x_{3i}) + \beta_4 ln(x_{4i}) + \varepsilon_i$$

con $\varepsilon_i \sim$ i.i.d. $N(0, \sigma^2)$ y donde y_i = consumo de pollo per-cápita, x_{1i}= ingreso real disponible per-cápita en dólares, x_{2i}= precio real del pollo por kg en dólares, x_{3i} = precio real del cerdo por kg en dólares, y x_{4i}= precio real de la carne de res por kg. en dólares. Usando información anual sobre la demanda de pollo durante el periodo de 1990-2012, Alicia estimó el siguiente modelo mediante el método de mínimos cuadrados ordinarios:

$$\widehat{ln(y_i)} = \underset{(0.1557)}{2.189} + \underset{(0.0833)}{0.343}\, ln(x_{1i}) - \underset{(0.1109)}{0.505}\, ln(x_{2i}) + \underset{(0.0977)}{0.149}\, ln(x_{3i}) + \underset{(0.0997)}{0.911}\, ln(x_{4i})$$

con $R^2 = 0.982$. Sin embargo, su compañero Pedro, experto en el mercado de venta de pollo al menudeo afirma que la verdadera función de demanda es la siguiente:

$$\widehat{ln(y_i)} = \underset{(0.1162)}{2.033} + \underset{(0.0247)}{0.452}\, ln(x_{1i}) - \underset{(0.0635)}{0.377}\, ln(x_{2i})$$

con $R^2 = 0.980$. En ambos modelos, los valores entre paréntesis se refieren a la desviación estándar de los parámetros estimados. Evalúe las siguientes afirmaciones para identificar la opción correcta. Evalúe las siguientes afirmaciones para identificar la opción correcta.

I) La Suma de Cuadrados Totales es la misma en los dos modelos estimados

II) A partir del coeficiente de determinación se pude concluir que los modelos son comparables y que el modelo estimado por Alicia ajusta mejor que el estimado por Pedro.

III) La elasticidad precio de la demanda estimada por Alicia es mayor en valor absoluto que la estimada por Pedro.

IV) El estadístico de prueba para probar la hipótesis de que la especificación del modelo de Pedro es el apropiado toma un valor aproximado de 1.00 con 2,18 grados de libertad y dado en Error de Tipo I $\alpha = 5\,\%$ se concluye que Pedro puede seguir gozando de su reputación de experto en el mercado de venta de pollo al menudeo.

Luego,

a) Las cuatro afirmaciones son verdaderas

b) Las cuatro afirmaciones son falsas

c) I) y IV) son verdaderas, pero II) y III) son falsas

d) I), II) y III) son verdaderas, pero IV) es falsa

6. **(6 puntos)**. Considera el siguiente modelo de regresión lineal simple

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t,\ \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } t = 1, 2, \cdots, 13$$

con Y_t= Salario por hora en dólares, X_t= Años de escolaridad, para t = 13 distintos años de estudio, que cubren valores $X_1 = 6$ a $X_{13} = 18$. Se obtuvieron así los siguientes resultados $\bar{X} = 12$ y $\sum_{i=1}^{13}(X_i - \bar{X})^2 = 182$ y la regresión estimada:

$$\widehat{Y_i} = \underset{(0.931)}{-0.014} + \underset{(0.070)}{0.724} X_i \text{ con } R^2 = 0.906$$

con la desviación estándar de los parámetros estimados en paréntesis y una varianza de los residuales $\hat{\sigma}^2 = 0.893$. Encuentra el pronóstico puntual y su intervalo de predicción del 95 %, correspondiente al salario por hora de un individuo con $X_0 = 20$ años de estudio.

a) El pronóstico puntual es 14.466 y por intervalo es $(12.805, 15.827)$

b) El pronóstico puntual es 12.466 y por intervalo es $(9.977, 16.959)$

c) El pronóstico puntual es 14.466 y por intervalo es $(11.984, 16.948)$

d) El pronóstico puntual es 12.466 y por intervalo es $(10.805, 13.827)$

7. **(12 puntos)**. Considere el siguiente modelo de regresión lineal múltiple:

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \text{ con } \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2)$$

con las cuatro observaciones para y_i, x_{1i} y x_{2i} que se presentan en la siguiente tabla:

y_i	x_{1i}	x_{2i}
2	1	0
0	1	0
3	3	1
-1	-1	-1

Analice las siguientes afirmaciones para elegir la opción correcta:

I) $\widehat{Var(\hat{\beta}_1 + \hat{\beta}_2)} = \frac{7}{4}$

II) $\widehat{Var(\hat{\beta}_1 + \hat{\beta}_2)} = \frac{3}{2}$

III) $\widehat{Cov(\hat{\beta}_1, \hat{\beta}_2)} = -\frac{1}{2}$

Entonces:

a) I) es verdadera, pero II) y III) son falsas

b) I) y III) son verdaderas, pero II) es falsa

c) II) y III) son verdaderas, pero I) es falsa

d) I) y III) son falsa, pero II) es verdadera

8. **(9 puntos)**. Se desea estimar el gasto público en educación $EDUC$ en función del Producto Interno Bruto PIB. Para ello, se cuenta con datos (en millones de Euros) correspondientes a 20 países. Los resultados son los siguientes:

$$\widehat{EDUC}_i = -21436 + 0.049PIB_i$$

mientras que el coeficiente de asimetría y curtosis de los residuales son iguales a -0.411 y 2.666, respectivamente. Analice las siguientes afirmaciones para elegir la opción correcta:

I) El estadístico apropiado para probar la hipótesis de normalidad toma un valor aproximadamente igual a 0.656 y dado un error de Tipo I, $\alpha = 5\,\%$, se concluye que no se cumple el supuesto de normalidad y por lo tanto no se puede realizar inferencia por desconocer su distribución exacta.

II) El estadístico apropiado para probar la hipótesis de normalidad toma un valor aproximadamente igual a 0.656 y dado un error de Tipo I, $\alpha = 90\,\%$, se concluye que se cumple el supuesto de normalidad y por lo tanto si se puede realizar inferencia.

III) El estadístico apropiado para probar la hipótesis de normalidad toma un valor aproximadamente igual a 3.50 y dado un error de Tipo I, $\alpha = 5\,\%$, se concluye que se cumple el supuesto de normalidad y por lo tanto no se puede realizar inferencia.

IV) El estadístico apropiado para probar la hipótesis de normalidad toma un valor aproximadamente igual a 3.50 y dado un error de Tipo I, $\alpha = 90\,\%$, se concluye que no se cumple el supuesto de normalidad y por lo tanto no se puede realizar inferencia por desconocer su distribución exacta.

Entonces,

a) I) y II) son verdaderas, pero III) y IV) son falsas

b) II) es verdaderas, pero I), III) y IV) son falsas

c) IV) es verdaderas, pero I), II) y III) son falsas

d) I), II), III) y IV) son falsas

9. **(10 puntos)**. Se ha estimado la función de producción Y_i con insumos de capital K_i y trabajo L_i, para una muestra aleatoria de 53 firmas:

$$\widehat{ln(Q_i)} = 1.37 + \underset{(0.257)}{0.632} ln(K_i) + \underset{(0.219)}{0.452} ln(L_i)$$

con $R^2 = 0.980$ y $\widehat{Cov(\hat{\beta}_1, \hat{\beta}_2)} = 0.055$. Los números en paréntesis son las desviaciones estándar de los parámetros estimados. Analice las siguientes afirmaciones para elegir la opción correcta:

I) El estadístico para probar la hipótesis de que las elasticidades del producto con respecto al capital y al trabajo son iguales versus que sean diferentes es aproximadamente igual a 2.857 y dado un error de Tipo I, $\alpha = 5\,\%$, se concluye que las elasticidades son diferentes

II) El intervalo de confianza del 95 % para la elasticidad del producto con respecto al capital es aproximadamente igual a $(0.116, 1.148)$

III) El estadístico para probar la hipótesis de la existencia de rendimientos constantes a escala versus rendimientos crecientes es aproximadamente igual a 0.389 y dado un error de Tipo I, $\alpha = 5\,\%$, se concluye que la función de producción de rendimientos a escala constantes

Entonces:

a) I) y II) son verdaderas, pero III) es falsa

b) I) y III) son verdaderas, pero II) es falsa

c) Las tres son verdaderas

d) Ninguna es verdadera

10. **(6 puntos)**. Suponga el modelo de regresión lineal simple:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \text{ con } \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, n$$

Entonces es cierto que:

a) $\sum_{i=1}^{n} \hat{y}_i \hat{\varepsilon}_i = 0$

b) $\sum_{i=1}^{n} x_i \hat{\varepsilon}_i = 0$

c) $\sum_{i=1}^{n} \hat{y}_i = \sum_{i=1}^{n} y_i$

d) Todas las anteriores

11. **(6 puntos)**. Considérese el siguiente modelo de regresión lineal, en donde los parámetros fueron estimados mediante el método de mínimos cuadrados ordinarios:

$$\widehat{ln(C_i)} = \underset{(0.427)}{-0.467} + \underset{(0.749)}{0.804} ln(YD_i) + \underset{(0.175)}{0.201} ln(W_i) - \underset{(0.0007)}{0.002}\, ln(R_i)$$

en donde C_i es el consumo real, YD_i es el ingreso personal disponible, W_i es la riqueza real y R_i es la tasa de interés real en una economía hipotética. Los números entre paréntesis se refieren a las desviaciones estándar de los parámetros estimados. Además, se cuenta con la siguiente información:

Regresor	$ln(YD_i)$	$ln(W_i)$	$ln(R_i)$
FIV	4	20	2.717

Entonces, el coeficiente de determinación que describe la medida de bondad de ajuste del logaritmo natural de la riqueza en términos del logaritmo natural del ingreso personal disponible y del logaritmo natural de la tasa de interés es aproximadamente igual a:

a) 0.57

b) 0.75

c) 0.95

d) 0.63

12. **(4 puntos)**. Considere el siguiente modelo de regresión lineal simple:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \text{ con } \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, n$$

Se tiene la siguiente tabla de ANOVA para el modelo estimado mediante el método de mínimos cuadrados ordinarios $\widehat{y}_i = 2627.82 - 37.15x_i$:

Fuente	Suma de Cuadrados	grados de libertad	F	valor p
$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	1,527,334.95	1	165.21	0.0000
$\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$	166,402.65	18		
$\sum_{i=1}^{n}(y_i - \bar{y})^2$	1,693,737.60	19		

Con estos datos se obtienen las siguientes conclusiones, seleccione la que es incorrecta:

a) El valor de $\hat{\sigma}^2 = 9,244.59$

b) Se rechaza la hipótesis $H_0 : \beta_1 = 0$ para un error de tipo I, $\alpha = 5\,\%$

c) La suma de cuadrados media total es: $1,536,579.64$

d) El valor de R^2 es 0.9017

13. **(5 puntos)**. En un modelo de regresión lineal múltiple en donde el modelo 1 tiene más regresores que el modelo 2, entonces siempre se cumple que:

   a) R^2 del modelo 1 es mayor que el R^2 del modelo 2
   b) $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ es mayor en el modelo 1 que en el modelo 2
   c) $\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$ es menor en el modelo 1 que en el modelo 2
   d) Las tres opciones son correctas

14. **(10 puntos)**. Considere el siguiente modelo de regresión lineal simple:

$$y_i = \beta_0 + \varepsilon_i \text{ con } E[\varepsilon_i] = 0 \text{ y } Var[\varepsilon_i] = \sigma^2 z_i$$

   Analice las siguientes afirmaciones para elegir la opción correcta:

   I) El estimador de mínimos cuadrados generalizados para β_0 es:

$$\widehat{\beta}_0^{MCG} = \frac{\sum_{i=1}^{n}(y_i/z_i^2)}{\sum_{i=1}^{n}(1/z_i^2)}$$

   II) $Var[\widehat{\beta}_0^{MCG}] = \frac{\sigma^2}{\sum_{i=1}^{n}(1/z_i^2)}$

   Luego:

   a) I) es falsa y II) es verdadera
   b) II) es falsa y I) es verdadera
   c) I) y II) son verdaderas
   d) I) y II) son falsas

5. Examen Final (Primavera 2021)

Econometría
20 de mayo de 2021

Nombre: _____ CU: _____

INSTRUCCIONES:

A continuación, se presentan 12 preguntas de opción múltiple (no se calificará procedimiento) y 2 problemas a desarrollar. En los 2 problemas a desarrollar indica claramente el inciso que estás resolviendo. Si tienes dos respuestas para una pregunta tacha la que consideres incorrecta, de lo contrario la ambigüedad anulará ambas. Un número o un resultado sin la debida justificación no dan crédito.

PREGUNTA 1

Sean los siguientes modelos de regresión lineal múltiple, donde ln representa el logaritmo natural:

(1) $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ con $\varepsilon_i \sim$ i.i.d.$(0, \sigma^2)$ para $i = 1, 2, \cdots, n$

(2) $y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \nu_i$ con $\nu_i \sim$ i.i.d.$(0, \sigma^2)$ para $i = 1, 2, \cdots, n$

(3) $y_i = \beta_0 + \beta_1 ln|x_{i1}| + \beta_2 ln|x_{i2}| + \eta_i$ con $\eta_i \sim$ i.i.d.$(0, \sigma^2)$ para $i = 1, 2, \cdots, n$

Considere las siguientes afirmaciones para elegir la opción correcta:

(I) Bajo los supuestos del modelo de regresión lineal, la suma de los residuos derivados de la estimación mediante el método de mínimos cuadrados ordinarios de los modelos (1), (2) y (3) son cero.

(II) El R^2 de la estimación de mínimos cuadrados ordinarios del modelo (1) siempre es mayor que el R^2 resultante del modelo (3).

(III) Si la variable $x_{i1} = 4$ para toda la muestra en los tres modelos, se produce multicolinealidad exacta en los modelos (1) y (3).

Entonces,

a) Las tres afirmaciones son falsas

b) (I) y (II) son falsas y (III) es verdadera

c) (I) y (II) son verdaderas y (III) es falsa

d) (I) y (III) es verdadera y (II) es falsa

PREGUNTA 2

En el modelo de regresión lineal múltiple:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i,\ i = 1, 2, \cdots, n$$

indica en cuál de los siguientes casos el término estocástico del modelo presenta heteroscedasticidad:

a) $Var[\varepsilon_i] = \sigma_i^2$ para toda $i = 1, 2, \cdots, n$

b) $\varepsilon_i = 10 + \nu_i$ donde $Var[\nu_i] = 5$ para toda $i = 1, 2, \cdots, n$

c) $Var[\varepsilon_i] = 2\sigma^2$ para toda $i = 1, 2, \cdots, n$

d) $\varepsilon_i = \rho\varepsilon_{i-1} + \eta_i$ donde $|\rho| < 1$ y $\nu_i \sim$ i.i.d.$(0, \sigma_\eta^2)$ para toda $i = 1, 2, \cdots, n$

PREGUNTA 3

Considere el siguiente modelo de regresión lineal múltiple :

$$C_i = \beta_0 + \beta_1 R_i + \beta_2 I_i + \varepsilon_i \text{ con } \varepsilon_i \sim \text{i.i.d.} N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, n$$

donde C_i es el consumo de la economía española, R_i la renta e I_i los impuestos pagados por el sector privado (todas las variables están medidas en millones de euros). Utilizando datos trimestrales desde el primer trimestre de 1974 hasta el cuarto trimestre de 1998, se obtiene la siguiente estimación (las desviaciones estándar de los parámetros estimados se reportan entre paréntesis):

$$\widehat{C_i} = \underset{(40)}{145} + \underset{(0.26)}{0.84}\, R_i - \underset{(0.22)}{0.53}\, I_i$$

con un $R^2 = 0.76$ y $\widehat{Cov}\left[\hat{\beta}_1, \hat{\beta}_1\right] = -0.05$. Se desea probar la hipótesis nula de que la propensión marginal a consumir (β_1) es unitaria frente a la alternativa de que es inferior a la unidad. Se concluye que:

a) No existe información suficiente para probar la hipótesis

b) La prueba de hipótesis se puede realizar y se rechaza la hipótesis nula en favor de la alternativa al 5 % de significancia

c) La prueba de hipótesis se puede realizar y no se rechaza la hipótesis nula en contra de la alternativa al 5 % de significancia

PREGUNTA 4

Suponga el modelo $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ cumple que $Var[\varepsilon_i] = \frac{\sigma^2}{z_i^2}$ con $z_i \neq 0$ para toda $i = 1, 2, \cdots, n$. Entonces, en el modelo que podríamos asegurar que el término estocástico η_i será homoscedástico es:

a) $\frac{y_i}{z_i} = \frac{\beta_0}{z_i} + \beta_1 \frac{x_i}{z_i} + \eta_i$

b) $y_i z_i = \beta_0 z_i + \beta_1 x_i z_i + \eta_i$

c) $\frac{y_i}{z_i^2} = \frac{\beta_0}{z_i^2} + \beta_1 \frac{x_i}{z_i^2} + \eta_i$

d) $y_i z_i^2 = \beta_0 z_i^2 + \beta_1 x_i z_i^2 + \eta_i$

PREGUNTA 5

Considere el siguiente modelo de regresión lineal:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i,\ \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, n$$

Al realizar su estimación mediante el método de mínimos cuadrados ordinarios se obtienen los siguientes resultados:

Parámetro	Estimador	d.s.	t-est.	valor-p
β_0		0.010291	-1.013	0.322
β_1	0.034656	0.006425	5.394	0.0000204
β_2	0.760464	0.121588	6.254	0.0000027
SRC	0.002503		g. de l.	22
R^2	0.971		$\bar{R}^2$	0.9706
F-est.	397.6		valor-p	0.00000

El estadístico Durbin-Watson es de 1.8524. Además, bajo la sospecha de que los errores del modelo presentan heteroscedasticidad, se estima una regresión auxiliar en donde se relacionan los residuos resultantes del modelo al cuadrado en función de una constante, todos los regresores del modelo, sus cuadrados y sus productos cruzados. El R^2 de dicha ecuación auxiliar es de $R^2_{aux} = 0.393$. Evalúe las siguientes afirmaciones para identificar la opción correcta.

(I) Hay evidencia estadística al 95 % de confianza para rechazar la hipótesis nula de que $H_0 : \beta_1 = \beta_2 = 0$

(II) Con $\alpha = 5\,\%$, hay evidencia estadística para afirmar que el término estocástico del modelo presenta autocorrelación de orden uno.

(III) El estimador puntual del parámetro β_0 es aproximadamente igual a -0.0104

Luego,

a) (I), (II) y (III) son verdaderas

b) (I) y (III) son verdaderas y (II) es falsa

c) (I) es verdadera, pero (II) y(III) son falsas

d) (I), (II) y (III) son falsas

PREGUNTA 6

Después de estimar un modelo mediante el método de mínimos cuadrados ordinarios, se obtiene la siguiente información:

- SCR_1 basado en las primeras 30 observaciones = 55 con 25 grados de libertad
- SCR_2 basado en las últimas 30 observaciones = 140 con 25 grados de libertad

Sea σ_i^2 la varianza de los residuales del grupo i. Evalúe las siguientes afirmaciones para identificar la opción correcta:

(I) Para para contrastar la hipótesis $H_0 : \sigma_1^2 = \sigma_2^2$ vs. $H_1 : \sigma_2^2 > \sigma_1^2$ usamos un estadístico de prueba cuya distribución muestral es $\chi^2_{(25)}$ y concluimos que no rechazamos H_0 con $\alpha = 5\,\%$

(II) Para para contrastar la hipótesis $H_0 : \sigma_1^2 = \sigma_2^2$ vs. $H_1 : \sigma_2^2 > \sigma_1^2$ usamos un estadístico de prueba cuya distribución muestral es $F_{(25,25)}$ y concluimos que no rechazamos H_0 con $\alpha = 5\,\%$

Luego:

a) (I) y (II) son verdaderas

b) (I) es verdadera y (II) es falsa

c) (I) es falsa y (II) es verdadera

d) (I) y (II) son falsas

PREGUNTA 7

El modelo de regresión lineal múltiple estimado por mínimos cuadrados ordinarios con $n = 53$ observaciones está dado por:

$$\hat{y}_i = 2 + 3x_{1i} - 2x_{2i}$$

La matriz de varianzas y covarianzas de los estimadores está dada por:

$$\widehat{Cov}(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2) = \begin{pmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

Entonces, al probar la hipótesis nula $H_0 : \beta_1 + 2\beta_2 = 5$ contra la alternativa $H_1 : \beta_1 + 2\beta_2 \neq 5$:

a) Se rechaza la hipótesis nula tanto al $\alpha = 10\,\%$ como al $\alpha = 5\,\%$

b) Se rechaza la hipótesis nula al $\alpha = 10\,\%$ pero no al $\alpha = 5\,\%$

c) No se rechaza la hipótesis nula al $\alpha = 5\,\%$ pero sí al $\alpha = 10\,\%$

d) No se rechaza la hipótesis nula ni al $\alpha = 5\,\%$ ni al $\alpha = 10\,\%$

PREGUNTA 8

Considere el siguiente modelo de regresión lineal simple:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,\ \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, 5$$

que al estimarlo mediante el método de mínimos cuadrados ordinarios se obtiene:

$$\widehat{y}_i = \underset{(0.5477)}{1.0} + \underset{(0.1118)}{1.75}\ x_i$$

Las desviaciones estándar de los parámetros estimados se reportan entre paréntesis. Además, sabemos que $\hat{\sigma}^2 = 0.5$, $\bar{x} = 4$ y $\sum_{i=1}^{5}(x_i - \bar{x})^2 = 40$. Analice las siguientes afirmaciones para elegir la opción correcta:

(I) Dado $x_0 = 10$, el pronóstico para y_0 es 18.5

(II) El intervalo de confianza del 95 % para y_0 es aproximadamente igual a $(15.24, 21.76)$

Entonces,

a) (I) es verdadera y (II) es falsa

b) (I) y (II) son falsas

c) (I) y (II) son verdaderas

d) (I) es falsa y (II) es verdadera

PREGUNTA 9

De acuerdo con ciertas investigaciones, se sabe que los precios de las acciones dependen de la política de dividendos seguida por sus administradores. Un posible modelo para explicar el valor de mercado de las acciones es: $P = \alpha \cdot VC^{\beta_1} \cdot DPA^{\beta_2} \cdot e^{\varepsilon_i}$ donde P es el precio de cotización en bolsa (en pesos mexicanos) de las acciones de una empresa, VC es el valor contable de las mismas, y DPA es el dividendo obtenido por acción. Para estimar el modelo se ha seleccionado una muestra de 22 empresas que cotizan en la Bolsa Mexicana de Valores. El modelo que se estimará por el método de mínimos cuadrados ordinarios es: $ln|P_i| = \beta_0 + \beta_1 ln|VC_i| + \beta_2 ln|DPA_i| + \varepsilon_i$, con $\varepsilon_i \sim$ i.i.d. $N(0, \sigma^2)$. Los resultados de la estimación son los siguientes:

$$\widehat{ln|P_i|} = \underset{(0.4152)}{1.0913} + \underset{(0.0935)}{0.7781} ln|VC_i| + \underset{(0.0820)}{0.1814} ln|DPA_i|$$

Las desviaciones estándar de los parámetros estimados se reportan entre paréntesis. Entonces, el intervalo de confianza aproximado para la elasticidad del precio de cotización de las acciones con respecto a su valor contable, al 95 % de confianza es:

a) (0.0206,0.3422)

b) (0.0465,0.3164)

c) (0.6243,0.9319)

d) (0.5824,0.9737)

PREGUNTA 10

De acuerdo con ciertas investigaciones, se sabe que los precios de las acciones dependen de la política de dividendos seguida por sus administradores. Un posible modelo para explicar el valor de mercado de las acciones es:

$$P = \alpha \cdot VC^{\beta_1} \cdot DPA^{\beta_2} \cdot e^{\varepsilon_i}$$

donde P es el precio de cotización en bolsa (en pesos mexicanos) de las acciones de una empresa, VC es el valor contable de las mismas, y DPA es el dividendo obtenido por acción. Para estimar el modelo se ha seleccionado una muestra de 22 empresas que cotizan en la Bolsa Mexicana de Valores. El modelo que se estimará por el método de mínimos cuadrados ordinarios es:

$$ln|P_i| = \beta_0 + \beta_1 ln|VC_i| + \beta_2 ln|DPA_i| + \varepsilon_i,\ \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2)$$

Los resultados de la estimación son los siguientes:

$$\widehat{ln|P_i|} = \underset{(0.4152)}{1.0913} + \underset{(0.0935)}{0.7781} ln|VC_i| + \underset{(0.0820)}{0.1814} ln|DPA_i|$$

Las desviaciones estándar de los parámetros estimados se reportan entre paréntesis. De las estimaciones se deduce que:

a) Ante un incremento del 1 % en el valor contable, el precio de cotización medio aumenta en un 0.778 % siendo el parámetro $\hat{\beta}_1$ significativo a una $\alpha = 5\,\%$

b) Ante un incremento de un peso mexicano en el valor contable, el precio de cotización medio aumenta en 0.778 pesos siendo el parámetro $\hat{\beta}_1$ significativo a una $\alpha = 5\,\%$

c) Ante un incremento del 1 % en el valor contable, el precio de cotización medio aumenta en un 0.778 % aunque el parámetro $\hat{\beta}_1$ no es significativo a una $\alpha = 5\,\%$ por lo que no podemos obtener ninguna conclusión válida.

d) Ante un incremento de un peso mexicano en el valor contable, el precio de cotización medio aumenta en 0.778 aunque el parámetro $\hat{\beta}_1$ no es significativo a una $\alpha = 5\,\%$ por lo que no podemos obtener ninguna conclusión válida.

PREGUNTA 11

Lo que se conoce como línea característica del análisis de inversión moderna, es sencillamente la línea de regresión simple obtenida del siguiente modelo:

$$r_t = \beta_0 + \beta_1 rm_t + \varepsilon_t,\ \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2)$$

en donde:

- r_t es la tasa de retorno del título-valor en el tiempo t
- rm_t es a tasa de retorno del portafolio del mercado en el tiempo t
- ε_t es el término estocástico

En este modelo β_1 es conocido como el Coeficiente Beta del título-valor, una medida del riesgo del mercado de un título-valor. Con base en 240 tasas de retorno mensuales, Fogler y Ganapathy obtuvieron la siguiente línea característica para las acciones de IBM en relación con el índice de portafolio del mercado desarrollado en la Universidad de Chicago (la desviación estándar de los parámetros estimados se da entre paréntesis):

$$\widehat{r}_t = \underset{(0.3001)}{0.7264} + \underset{(0.0728)}{1.0598} rm_t$$

con $R^2 = 0.4710$ y $\hat{\sigma}^2 = 0.050$. Se dice que un título-valor cuyo Coeficiente Beta es mayor que uno es un título-valor volátil o agresivo. Entonces:

(I) Con $\alpha = 5\,\%$, las acciones de IBM son valores volátiles en el período bajo estudio

(II) El riesgo del mercado es un factor que explica el $5\,\%$ de la variabilidad en la rentabilidad de IBM

Luego,

a) (I) y (II) son verdaderas

b) (I) es falsa y (II) es verdadera

c) (I) es verdadera y (II) es falsa

d) (I) y (II) son falsas

PREGUNTA 12

En las siguientes afirmaciones, β_j es el coeficiente que acompaña a la variable X_j y R_j^2 es el coeficiente de determinación de la regresión auxiliar de la variable X_j en función del resto de las k variables independientes. FIV_j es el factor inflador de la varianza. Evalúe las siguientes afirmaciones para identificar la opción correcta:

(I) La varianza del estimador de mínimos cuadrados ordinarios de cada uno de los parámetros del modelo de regresión lineal está dada por

$$Var[\hat{\beta}_j] = \frac{\sigma^2}{\sum(X_{ji} - \bar{X}_j)^2(1 - R_j^2)} = \frac{\sigma^2}{\sum(X_{ji} - \bar{X}_j)^2} \cdot FIV_j$$

(II) Si $R^2 \approx 1$ tendríamos un problema de colinealidad en el modelo

(III) Si $R^2 = 1$ tendríamos un problema de colinealidad perfecta en el modelo y el rango$(X) \neq k + 1$, en donde X es la matriz de $k + 1$ columnas.

Luego:

a) (I) es falsa, pero (II) y (III) son verdaderas

b) (I), (II) y (III) son verdaderas

c) (I) y (III) son falsas, pero (II) es verdadera

d) (I), (II) y (III) son falsas

PREGUNTA ABIERTA 1

Se obtienen datos anuales durante el período 1979 a 2017 para cada una de las siguientes variables:

- Q: un índice del PIB de EUA en dólares constantes
- L: un índice de los días laborales
- K: un índice para el insumo de capital

Se estimó la función de producción para el periodo completo y los resultados son los siguientes:

$$ln|Q_t| = -3.8766 + 1.4106ln|L_t| + 0.4162ln|K_t|$$
$$\text{con } R^2 = 0.9937 \text{ y } \hat{\sigma}^2 = 0.03755$$

También se estimó la función de producción para los siguientes dos periodos

(I) 1979-1998

$$ln|Q_t| = -4.0576 + 1.6167ln|L_t| + 0.2197ln|K_t|$$
$$\text{con } R^2 = 0.9759 \text{ y } \hat{\sigma}^2 = 0.04573$$

(II) 1999-2017

$$ln|Q_t| = -1.9564 + 0.8336ln|L_t| + 0.6631ln|K_t|$$
$$\text{con } R^2 = 0.9904 \text{ y } \hat{\sigma}^2 = 0.02185$$

A un nivel de confianza del 95 % pruebe la igualdad de los parámetros de la función de producción en ambos periodos.

PREGUNTA ABIERTA 2

Considere el siguiente modelo

$$E_i = \beta_0 + \beta_1 D_i + \varepsilon_i,\ \varepsilon_i \sim \text{i.i.d. } (0, \sigma^2),\ i = 1, 2, \cdots, n$$

La variable E_i representa el número de años de educación de un individuo y D_i es una variable dicotómica que vale 1 si una persona vive en una zona urbana y 0 si vive en una zona rural. Si utilizamos 300 observaciones de las cuales 150 son de personas que viven en zona urbana y 150 que viven en zona rural, la suma de los años de educación de las personas de zona urbana es 2250 y 1500 para los que viven en zona rural, esto es:

$$\sum_{i=1}^{150} E_i = 2250 \text{ para los que viven en zona urbana}$$
$$\sum_{i=1}^{150} E_i = 1500 \text{ para los que viven en zona rural}$$

a) Escriba el modelo en su forma matricial, esto es: $\vec{Y} = \boldsymbol{X}\vec{\beta} + \vec{\varepsilon}$ (exprese a $\vec{Y}$ con los niveles E_i de cada individuo)

b) Proporcione $\boldsymbol{X}'\boldsymbol{X}$ y $\boldsymbol{X}'\vec{Y}$

c) Proporcione los estimadores de mínimos cuadrados ordinarios de β_0 y β_1

d) Si $\hat{\sigma}^2 = 300$, calcule $\widehat{Var}[\hat{\beta}_0]$ y $\widehat{Var}[\hat{\beta}_1]$

6. Examen Final (Primavera 2021) - Solución

Econometría
20 de mayo de 2021

Nombre: _____ CU: _____

INSTRUCCIONES:

A continuación, se presentan 12 preguntas de opción múltiple (no se calificará procedimiento) y 2 problemas a desarrollar. En los 2 problemas a desarrollar indica claramente el inciso que estás resolviendo. Si tienes dos respuestas para una pregunta tacha la que consideres incorrecta, de lo contrario la ambigüedad anulará ambas. Un número o un resultado sin la debida justificación no dan crédito.

PREGUNTA 1

Sean los siguientes modelos de regresión lineal múltiple, donde ln representa el logaritmo natural:

(1) $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ con $\varepsilon_i \sim$ i.i.d.$(0, \sigma^2)$ para $i = 1, 2, \cdots, n$

(2) $y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \nu_i$ con $\nu_i \sim$ i.i.d.$(0, \sigma^2)$ para $i = 1, 2, \cdots, n$

(3) $y_i = \beta_0 + \beta_1 ln|x_{i1}| + \beta_2 ln|x_{i2}| + \eta_i$ con $\eta_i \sim$ i.i.d.$(0, \sigma^2)$ para $i = 1, 2, \cdots, n$

Considere las siguientes afirmaciones para elegir la opción correcta:

(I) Bajo los supuestos del modelo de regresión lineal, la suma de los residuos derivados de la estimación mediante el método de mínimos cuadrados ordinarios de los modelos (1), (2) y (3) son cero.

(II) El R^2 de la estimación de mínimos cuadrados ordinarios del modelo (1) siempre es mayor que el R^2 resultante del modelo (3).

(III) Si la variable $x_{i1} = 4$ para toda la muestra en los tres modelos, se produce multicolinealidad exacta en los modelos (1) y (3).

Entonces,

a) Las tres afirmaciones son falsas

b) (I) y (II) son falsas y (III) es verdadera

c) (I) y (II) son verdaderas y (III) es falsa

d) (I) y (III) es verdadera y (II) es falsa

PREGUNTA 2

En el modelo de regresión lineal múltiple:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i,\ i = 1, 2, \cdots, n$$

indica en cuál de los siguientes casos el término estocástico del modelo presenta heteroscedasticidad:

a) $Var[\varepsilon_i] = \sigma_i^2$ para toda $i = 1, 2, \cdots, n$

b) $\varepsilon_i = 10 + \nu_i$ donde $Var[\nu_i] = 5$ para toda $i = 1, 2, \cdots, n$

c) $Var[\varepsilon_i] = 2\sigma^2$ para toda $i = 1, 2, \cdots, n$

d) $\varepsilon_i = \rho\varepsilon_{i-1} + \eta_i$ donde $|\rho| < 1$ y $\nu_i \sim$ i.i.d.$(0, \sigma_\eta^2)$ para toda $i = 1, 2, \cdots, n$

PREGUNTA 3

Considere el siguiente modelo de regresión lineal múltiple :

$$C_i = \beta_0 + \beta_1 R_i + \beta_2 I_i + \varepsilon_i \text{ con } \varepsilon_i \sim \text{i.i.d.} N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, n$$

donde C_i es el consumo de la economía española, R_i la renta e I_i los impuestos pagados por el sector privado (todas las variables están medidas en millones de euros). Utilizando datos trimestrales desde el primer trimestre de 1974 hasta el cuarto trimestre de 1998, se obtiene la siguiente estimación (las desviaciones estándar de los parámetros estimados se reportan entre paréntesis):

$$\widehat{C_i} = \underset{(40)}{145} + \underset{(0.26)}{0.84}\, R_i - \underset{(0.22)}{0.53}\, I_i$$

con un $R^2 = 0.76$ y $\widehat{Cov}\left[\hat{\beta}_1, \hat{\beta}_1\right] = -0.05$. Se desea probar la hipótesis nula de que la propensión marginal a consumir (β_1) es unitaria frente a la alternativa de que es inferior a la unidad. Se concluye que:

a) No existe información suficiente para probar la hipótesis

b) La prueba de hipótesis se puede realizar y se rechaza la hipótesis nula en favor de la alternativa al 5 % de significancia

c) La prueba de hipótesis se puede realizar y no se rechaza la hipótesis nula en contra de la alternativa al 5 % de significancia

PREGUNTA 4

Suponga el modelo $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ cumple que $Var[\varepsilon_i] = \frac{\sigma^2}{z_i^2}$ con $z_i \neq 0$ para toda $i = 1, 2, \cdots, n$. Entonces, en el modelo que podríamos asegurar que el término estocástico η_i será homoscedástico es:

a) $\frac{y_i}{z_i} = \frac{\beta_0}{z_i} + \beta_1 \frac{x_i}{z_i} + \eta_i$

b) $y_i z_i = \beta_0 z_i + \beta_1 x_i z_i + \eta_i$

c) $\frac{y_i}{z_i^2} = \frac{\beta_0}{z_i^2} + \beta_1 \frac{x_i}{z_i^2} + \eta_i$

d) $y_i z_i^2 = \beta_0 z_i^2 + \beta_1 x_i z_i^2 + \eta_i$

PREGUNTA 5

Considere el siguiente modelo de regresión lineal:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i,\ \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, n$$

Al realizar su estimación mediante el método de mínimos cuadrados ordinarios se obtienen los siguientes resultados:

Parámetro	Estimador	d.s.	t-est.	valor-p
β_0		0.010291	-1.013	0.322
β_1	0.034656	0.006425	5.394	0.0000204
β_2	0.760464	0.121588	6.254	0.0000027
SRC	0.002503		g. de l.	22
R^2	0.971		$\bar{R}^2$	0.9706
F-est.	397.6		valor-p	0.00000

El estadístico Durbin-Watson es de 1.8524. Además, bajo la sospecha de que los errores del modelo presentan heteroscedasticidad, se estima una regresión auxiliar en donde se relacionan los residuos resultantes del modelo al cuadrado en función de una constante, todos los regresores del modelo, sus cuadrados y sus productos cruzados. El R^2 de dicha ecuación auxiliar es de $R^2_{aux} = 0.393$. Evalúe las siguientes afirmaciones para identificar la opción correcta.

(I) Hay evidencia estadística al 95 % de confianza para rechazar la hipótesis nula de que $H_0 : \beta_1 = \beta_2 = 0$

(II) Con $\alpha = 5\,\%$, hay evidencia estadística para afirmar que el término estocástico del modelo presenta autocorrelación de orden uno.

(III) El estimador puntual del parámetro β_0 es aproximadamente igual a -0.0104

Luego,

a) (I), (II) y (III) son verdaderas

b) (I) y (III) son verdaderas y (II) es falsa

c) (I) es verdadera, pero (II) y(III) son falsas

d) (I), (II) y (III) son falsas

PREGUNTA 6

Después de estimar un modelo mediante el método de mínimos cuadrados ordinarios, se obtiene la siguiente información:

- SCR_1 basado en las primeras 30 observaciones = 55 con 25 grados de libertad
- SCR_2 basado en las últimas 30 observaciones = 140 con 25 grados de libertad

Sea σ_i^2 la varianza de los residuales del grupo i. Evalúe las siguientes afirmaciones para identificar la opción correcta:

(I) Para para contrastar la hipótesis $H_0 : \sigma_1^2 = \sigma_2^2$ vs. $H_1 : \sigma_2^2 > \sigma_1^2$ usamos un estadístico de prueba cuya distribución muestral es $\chi^2_{(25)}$ y concluimos que no rechazamos H_0 con $\alpha = 5\,\%$

(II) Para para contrastar la hipótesis $H_0 : \sigma_1^2 = \sigma_2^2$ vs. $H_1 : \sigma_2^2 > \sigma_1^2$ usamos un estadístico de prueba cuya distribución muestral es $F_{(25,25)}$ y concluimos que no rechazamos H_0 con $\alpha = 5\,\%$

Luego:

a) (I) y (II) son verdaderas

b) (I) es verdadera y (II) es falsa

c) (I) es falsa y (II) es verdadera

d) (I) y (II) son falsas

PREGUNTA 7

El modelo de regresión lineal múltiple estimado por mínimos cuadrados ordinarios con $n = 53$ observaciones está dado por:

$$\hat{y}_i = 2 + 3x_{1i} - 2x_{2i}$$

La matriz de varianzas y covarianzas de los estimadores está dada por:

$$\widehat{Cov}(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2) = \begin{pmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

Entonces, al probar la hipótesis nula $H_0 : \beta_1 + 2\beta_2 = 5$ contra la alternativa $H_1 : \beta_1 + 2\beta_2 \neq 5$:

a) Se rechaza la hipótesis nula tanto al $\alpha = 10\,\%$ como al $\alpha = 5\,\%$

b) Se rechaza la hipótesis nula al $\alpha = 10\,\%$ pero no al $\alpha = 5\,\%$

c) No se rechaza la hipótesis nula al $\alpha = 5\,\%$ pero sí al $\alpha = 10\,\%$

d) No se rechaza la hipótesis nula ni al $\alpha = 5\,\%$ ni al $\alpha = 10\,\%$

PREGUNTA 8

Considere el siguiente modelo de regresión lineal simple:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,\ \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \text{ para } i = 1, 2, \cdots, 5$$

que al estimarlo mediante el método de mínimos cuadrados ordinarios se obtiene:

$$\widehat{y}_i = \underset{(0.5477)}{1.0} + \underset{(0.1118)}{1.75}\ x_i$$

Las desviaciones estándar de los parámetros estimados se reportan entre paréntesis. Además, sabemos que $\hat{\sigma}^2 = 0.5$, $\bar{x} = 4$ y $\sum_{i=1}^{5}(x_i - \bar{x})^2 = 40$. Analice las siguientes afirmaciones para elegir la opción correcta:

(I) Dado $x_0 = 10$, el pronóstico para y_0 es 18.5

(II) El intervalo de confianza del 95 % para y_0 es aproximadamente igual a $(15.24, 21.76)$

Entonces,

a) (I) es verdadera y (II) es falsa

b) (I) y (II) son falsas

c) (I) y (II) son verdaderas

d) (I) es falsa y (II) es verdadera

PREGUNTA 9

De acuerdo con ciertas investigaciones, se sabe que los precios de las acciones dependen de la política de dividendos seguida por sus administradores. Un posible modelo para explicar el valor de mercado de las acciones es: $P = \alpha \cdot VC^{\beta_1} \cdot DPA^{\beta_2} \cdot e^{\varepsilon_i}$ donde P es el precio de cotización en bolsa (en pesos mexicanos) de las acciones de una empresa, VC es el valor contable de las mismas, y DPA es el dividendo obtenido por acción. Para estimar el modelo se ha seleccionado una muestra de 22 empresas que cotizan en la Bolsa Mexicana de Valores. El modelo que se estimará por el método de mínimos cuadrados ordinarios es: $ln|P_i| = \beta_0 + \beta_1 ln|VC_i| + \beta_2 ln|DPA_i| + \varepsilon_i$, con $\varepsilon_i \sim$ i.i.d. $N(0, \sigma^2)$. Los resultados de la estimación son los siguientes:

$$\widehat{ln|P_i|} = \underset{(0.4152)}{1.0913} + \underset{(0.0935)}{0.7781} ln|VC_i| + \underset{(0.0820)}{0.1814} ln|DPA_i|$$

Las desviaciones estándar de los parámetros estimados se reportan entre paréntesis. Entonces, el intervalo de confianza aproximado para la elasticidad del precio de cotización de las acciones con respecto a su valor contable, al 95 % de confianza es:

a) (0.0206,0.3422)

b) (0.0465,0.3164)

c) (0.6243,0.9319)

d) (0.5824,0.9737)

PREGUNTA 10

De acuerdo con ciertas investigaciones, se sabe que los precios de las acciones dependen de la política de dividendos seguida por sus administradores. Un posible modelo para explicar el valor de mercado de las acciones es:

$$P = \alpha \cdot VC^{\beta_1} \cdot DPA^{\beta_2} \cdot e^{\varepsilon_i}$$

donde P es el precio de cotización en bolsa (en pesos mexicanos) de las acciones de una empresa, VC es el valor contable de las mismas, y DPA es el dividendo obtenido por acción. Para estimar el modelo se ha seleccionado una muestra de 22 empresas que cotizan en la Bolsa Mexicana de Valores. El modelo que se estimará por el método de mínimos cuadrados ordinarios es:

$$ln|P_i| = \beta_0 + \beta_1 ln|VC_i| + \beta_2 ln|DPA_i| + \varepsilon_i,\ \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2)$$

Los resultados de la estimación son los siguientes:

$$\widehat{ln|P_i|} = \underset{(0.4152)}{1.0913} + \underset{(0.0935)}{0.7781} ln|VC_i| + \underset{(0.0820)}{0.1814} ln|DPA_i|$$

Las desviaciones estándar de los parámetros estimados se reportan entre paréntesis. De las estimaciones se deduce que:

a) Ante un incremento del 1 % en el valor contable, el precio de cotización medio aumenta en un 0.778 % siendo el parámetro $\hat{\beta}_1$ significativo a una $\alpha = 5\,\%$

b) Ante un incremento de un peso mexicano en el valor contable, el precio de cotización medio aumenta en 0.778 pesos siendo el parámetro $\hat{\beta}_1$ significativo a una $\alpha = 5\,\%$

c) Ante un incremento del 1 % en el valor contable, el precio de cotización medio aumenta en un 0.778 % aunque el parámetro $\hat{\beta}_1$ no es significativo a una $\alpha = 5\,\%$ por lo que no podemos obtener ninguna conclusión válida.

d) Ante un incremento de un peso mexicano en el valor contable, el precio de cotización medio aumenta en 0.778 aunque el parámetro $\hat{\beta}_1$ no es significativo a una $\alpha = 5\,\%$ por lo que no podemos obtener ninguna conclusión válida.

PREGUNTA 11

Lo que se conoce como línea característica del análisis de inversión moderna, es sencillamente la línea de regresión simple obtenida del siguiente modelo:

$$r_t = \beta_0 + \beta_1 rm_t + \varepsilon_t,\ \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2)$$

en donde:

- r_t es la tasa de retorno del título-valor en el tiempo t
- rm_t es a tasa de retorno del portafolio del mercado en el tiempo t
- ε_t es el término estocástico

En este modelo β_1 es conocido como el Coeficiente Beta del título-valor, una medida del riesgo del mercado de un título-valor. Con base en 240 tasas de retorno mensuales, Fogler y Ganapathy obtuvieron la siguiente línea característica para las acciones de IBM en relación con el índice de portafolio del mercado desarrollado en la Universidad de Chicago (la desviación estándar de los parámetros estimados se da entre paréntesis):

$$\widehat{r}_t = \underset{(0.3001)}{0.7264} + \underset{(0.0728)}{1.0598} rm_t$$

con $R^2 = 0.4710$ y $\hat{\sigma}^2 = 0.050$. Se dice que un título-valor cuyo Coeficiente Beta es mayor que uno es un título-valor volátil o agresivo. Entonces:

(I) Con $\alpha = 5\,\%$, las acciones de IBM son valores volátiles en el período bajo estudio

(II) El riesgo del mercado es un factor que explica el $5\,\%$ de la variabilidad en la rentabilidad de IBM

Luego,

a) (I) y (II) son verdaderas

b) (I) es falsa y (II) es verdadera

c) (I) es verdadera y (II) es falsa

d) (I) y (II) son falsas

PREGUNTA 12

En las siguientes afirmaciones, β_j es el coeficiente que acompaña a la variable X_j y R_j^2 es el coeficiente de determinación de la regresión auxiliar de la variable X_j en función del resto de las k variables independientes. FIV_j es el factor inflador de la varianza. Evalúe las siguientes afirmaciones para identificar la opción correcta:

(I) La varianza del estimador de mínimos cuadrados ordinarios de cada uno de los parámetros del modelo de regresión lineal está dada por

$$Var[\hat{\beta}_j] = \frac{\sigma^2}{\sum(X_{ji} - \bar{X}_j)^2(1 - R_j^2)} = \frac{\sigma^2}{\sum(X_{ji} - \bar{X}_j)^2} \cdot FIV_j$$

(II) Si $R^2 \approx 1$ tendríamos un problema de colinealidad en el modelo

(III) Si $R^2 = 1$ tendríamos un problema de colinealidad perfecta en el modelo y el rango$(X) \neq k + 1$, en donde X es la matriz de $k + 1$ columnas.

Luego:

a) (I) es falsa, pero (II) y (III) son verdaderas

b) (I), (II) y (III) son verdaderas

c) (I) y (III) son falsas, pero (II) es verdadera

d) (I), (II) y (III) son falsas

PREGUNTA ABIERTA 1

Se obtienen datos anuales durante el período 1979 a 2017 para cada una de las siguientes variables:

- Q: un índice del PIB de EUA en dólares constantes
- L: un índice de los días laborales
- K: un índice para el insumo de capital

Se estimó la función de producción para el periodo completo y los resultados son los siguientes:

$$ln|Q_t| = -3.8766 + 1.4106ln|L_t| + 0.4162ln|K_t|$$
$$\text{con } R^2 = 0.9937 \text{ y } \hat{\sigma}^2 = 0.03755$$

También se estimó la función de producción para los siguientes dos periodos

(I) 1979-1998

$$ln|Q_t| = -4.0576 + 1.6167ln|L_t| + 0.2197ln|K_t|$$
$$\text{con } R^2 = 0.9759 \text{ y } \hat{\sigma}^2 = 0.04573$$

(II) 1999-2017

$$ln|Q_t| = -1.9564 + 0.8336ln|L_t| + 0.6631ln|K_t|$$
$$\text{con } R^2 = 0.9904 \text{ y } \hat{\sigma}^2 = 0.02185$$

A un nivel de confianza del 95 % pruebe la igualdad de los parámetros de la función de producción en ambos periodos.

Solución. Deseamos realizar una prueba de hipótesis sobre la estabilidad de parámetros. Para ello, calculamos la SRC para cada modelo:

$$\begin{aligned} SRC_R =& \hat{\sigma}^2(n-k-1) = 0.03755(39-3) = 1.3518 \\ SRC_1 =& \hat{\sigma}_1^2(n-k-1) = 0.04573(20-3) = 0.7774 \\ SRC_2 =& \hat{\sigma}_2^2(n-k-1) = 0.02188(19-3) = 0.3496 \\ SRC_{NR} =& SRC_1 + SRC_2 = 0.7774 + 0.3496 = 1.1237 \end{aligned}$$

Entonces el estadístico de prueba F será:

$$F = \frac{(SRC_R - SRC_{NR})/r}{SRC_{NR}/(n-k-1)} = \frac{(1.3518 - 1.1237)/3}{1.2137/33} = \frac{0.07603}{0.03677} = 2.0677$$

Dado que $F_{3,33,0.95} = 2.92$, entonces no se rechaza la hipótesis nula, por lo que los parámetros de la función de producción son iguales en ambos periodos. ■

PREGUNTA ABIERTA 2

Considere el siguiente modelo

$$E_i = \beta_0 + \beta_1 D_i + \varepsilon_i,\ \varepsilon_i \sim \text{i.i.d. } (0, \sigma^2),\ i = 1, 2, \cdots, n$$

La variable E_i representa el número de años de educación de un individuo y D_i es una variable dicotómica que vale 1 si una persona vive en una zona urbana y 0 si vive en una zona rural. Si utilizamos 300 observaciones de las cuales 150 son de personas que viven en zona urbana y 150 que viven en zona rural, la suma de los años de educación de las personas de zona urbana es 2250 y 1500 para los que viven en zona rural, esto es:

$$\sum_{i=1}^{150} E_i = 2250 \text{ para los que viven en zona urbana}$$
$$\sum_{i=1}^{150} E_i = 1500 \text{ para los que viven en zona rural}$$

a) Escriba el modelo en su forma matricial, esto es: $\vec{Y} = \boldsymbol{X}\vec{\beta} + \vec{\varepsilon}$ (exprese a $\vec{Y}$ con los niveles E_i de cada individuo)

b) Proporcione $\boldsymbol{X}'\boldsymbol{X}$ y $\boldsymbol{X}'\vec{Y}$

c) Proporcione los estimadores de mínimos cuadrados ordinarios de β_0 y β_1

d) Si $\hat{\sigma}^2 = 300$, calcule $\widehat{Var}[\hat{\beta}_0]$ y $\widehat{Var}[\hat{\beta}_1]$

Solución. Para escribir el modelo en su forma matricial definimos:

$$\vec{Y} = \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_{n_1} \\ E_{n_1+1} \\ E_{n_1+2} \\ \vdots \\ E_n \end{bmatrix},\ \boldsymbol{X} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{bmatrix},\ \vec{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix},\ \vec{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{n_1} \\ \varepsilon_{n_1+1} \\ \varepsilon_{n_1+2} \\ \vdots \\ \varepsilon_n \end{bmatrix},$$

Por lo que el modelo $\vec{Y} = \boldsymbol{X}\vec{\beta} + \vec{\varepsilon}$ será:

$$\begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_{n_1} \\ E_{n_1+1} \\ E_{n_1+2} \\ \vdots \\ E_n \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{n_1} \\ \varepsilon_{n_1+1} \\ \varepsilon_{n_1+2} \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Con estas matrices definidas, calculamos:

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} n & n_1 \\ n_1 & n_1 \end{bmatrix} = \begin{bmatrix} 300 & 150 \\ 150 & 150 \end{bmatrix}$$

$$\boldsymbol{X}'\vec{Y} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_{n_1} \\ E_{n_1+1} \\ E_{n_1+2} \\ \vdots \\ E_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} E_i \\ \sum_{i=1}^{n_1} E_i \end{bmatrix} = \begin{bmatrix} 3750 \\ 2250 \end{bmatrix}$$

Entonces, los estimadores de mínimos cuadrados ordinarios de β_0 y β_1 se calculan a partir de:

$$\hat{\vec{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\vec{Y} = \begin{bmatrix} 300 & 150 \\ 150 & 150 \end{bmatrix}^{-1} \begin{bmatrix} 3750 \\ 2250 \end{bmatrix} = \begin{bmatrix} \frac{1}{150} & -\frac{1}{150} \\ -\frac{1}{150} & \frac{2}{150} \end{bmatrix} \begin{bmatrix} 3750 \\ 2250 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \end{bmatrix}$$

esto quiere decir que $\widehat{\beta_0} = 10$ y $\widehat{\beta_1} = 5$.

Finalmente, calculamos la matriz de varianzas y covarianzas de los estimadores:

$$\widehat{Var[\hat{\vec{\beta}}]} = \hat{\sigma}^2(\boldsymbol{X}'\boldsymbol{X})^{-1} = 300 \begin{bmatrix} 300 & 150 \\ 150 & 150 \end{bmatrix}^{-1} = 300 \begin{bmatrix} \frac{1}{150} & -\frac{1}{150} \\ -\frac{1}{150} & \frac{2}{150} \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}$$

Por lo que: $\widehat{Var[\hat{\beta}_0]} = 2$, $\widehat{Var[\hat{\beta}_1]} = 4$ y $\widehat{Cov[\hat{\beta}_0, \hat{\beta}_1]} = -2$. ■