ECONOMETRIA I
Profesor: Alfonso Novales
Octubre 2008
Cuestiones de Tests de clase

Parte I. Correlación. Interpretación del modelo lineal de regresión

1. ¿Qué relación hay entre los coeficientes de correlación lineal $Corr(aX, cY)$ y $Corr(X, Y)$?

2. Demuestre la relación existente entre Cov(X,Y) y Cov(X,kY)

3. ¿Qué relación hay entre los coeficientes de correlación lineal $Corr(aX + b, cY + d)$ y $Corr(X, Y)$?

4. ¿Es correcto comparar la volatilidad de dos variables comparando su desviación típica? ¿Hay alguna manera mejor de hacerlo?

5. Escriba de tres maneras diferentes la definición de la covarianza *muestral* de dos variables. ¿A qué es igual dicha covarianza cuando una de las variables tiene media muestral igual a cero? ¿Cómo cambiarían sus expresiones si estuviésemos hablando de covarianzas poblacionales, y no de covarianzas muestrales?

6. El coeficiente de endeudamiento medio en una muestra de empresas industriales españolas es de 80%, mientras que el coste de financiación medio es de 150 puntos básicos por encima del MIBOR. El coeficiente de correlación lineal entre ambas variables es de $\rho = 0{,}92$. Estamos interesados en la empresa ECONOMUS, pero no sabemos nada sobre ella. Cuál sería nuestra predicción sobre el coste de financiación al que se enfrenta? Suponga ahora que acabamos de enterarnos de que esta empresa tiene un coeficiente de endeudamiento del 40%. ¿Cambiaríamos nuestra predicción acerca de su coste de financiación?

7. ¿Cómo se calcula la estimación numérica de la pendiente β_1 de un modelo de regresión lineal simple: $y = \beta_0 + \beta_1 x + u$? Un compañero suyo le dice que si estima la regresión de x sobre y obtendrá el mismo valor numérico para dicha pendiente ¿Qué cree?

8. Con respecto a las regresiones $y_i = \beta_0 + \beta_1 x_i + u_i$ y $x_i = \delta_0 + \delta_1 y_i + v_i$, estimados por mínimos cuadrados, indique cuál de las siguientes afirmaciones es cierta:
   - ✓ Las estimaciones de las pendientes son inversas una de la otra
   - ✓ Las estimaciones de las pendientes tienen signo opuesto
   - ✓ El producto de las pendientes estimadas es igual al coeficiente de determinación

9. Utilizando datos de nuestro gasto en publicidad y nuestras cifras de ventas durante los últimos años, hemos estimado el modelo: $Ventas_t = 400 + 1{,}50 Publicidad_t + \hat{u}_i$. Si aumentamos el gasto en publicidad en 10.000 euros ¿en cuanto esperamos que aumenten las ventas?

10. Proporcione al menos 2 formas de evaluar la relevancia económica de X sobre Y, una vez que hayamos estimado la regresión lineal simple correspondiente.

11. Indique 5 hipótesis contenidas en el modelo de regresión lineal

12. ¿Cómo se define el residuo de una regresión lineal simple?

13. Utilizando datos anuales en euros para 1970-2002, hemos estimado la siguiente regresión para explicar el precio de las viviendas por su superficie: $\text{Precio } en\ euros_i = 23000 - 0{,}005 \text{ Superficie } en\ m_i^2$.

Si estimásemos con los precios en dólares (es decir, multiplicando el precio en euros por 1,26) ¿cómo cambiarían las estimaciones numéricas de la constante y la pendiente?

14. ¿Qué información proporciona el coeficiente de correlación entre dos variables? ¿Es más informativo que dicho coeficiente sea positivo, o que sea negativo?

15. ¿Cabe esperar que permanezca inalterado el R2 de un modelo de regresión lineal si transformamos las variables en logaritmos?

Parte II. Estimación

1. Deduzca la expresión para la varianza del estimador de mínimos cuadrados $\hat{\beta}_1$ de la pendiente de la recta de regresión $y_i = \beta_0 + \beta_1 x_i + u_i, i = 1, 2, ..., n$

2. ¿Qué entiendes por relevancia económica de una variable explicativa X sobre la variable dependiente Y? Proporciona dos medidas diferentes de la relevancia de X sobre Y.

3. Hemos estimado 2 modelos alternativos para explicar una misma variable dependiente ¿Por qué es interesante calcular las correlaciones entre los residuos de cada modelo y la variable dependiente?

4. Expresión del coeficiente de determinación ¿Porqué está entre 0 y 1?

5. Demuestre que la suma de los residuos de la estimación de mínimos cuadrados suman cero.

6. Demuestre que los residuos de la estimación de mínimos cuadrados de un modelo de regresión lineal simple tienen correlación cero con la variable explicativa. Interprete este resultado.

7. Si estimamos un modelo con poca precisión, ¿qué problema se plantea de cara a contrastar una hipótesis acerca del verdadero valor numérico de dicho parámetro?

8. Hemos estimado un modelo de regresión para explicar el comportamiento de la variable *Y*: $Y = \alpha + \beta X + u$ y otro modelo para explicar el comportamiento de una variable diferente, *Z:* $Z = \delta_0 + \delta_1 X + v_i$. El primer modelo tiene una suma de cuadrados de residuos mucho menor que el segundo ¿Podemos decir que es un modelo mejor?

9. Si el modelo $Y = \alpha + \beta X + u$ tiene un R2 inferior al del modelo*:* $Y = \delta_0 + \delta_1 X + \delta_2 Z + v_i$ ¿Podemos decir que el segundo modelo es mejor?

10. Considere el modelo $y_i = \alpha + \beta x_i + u_i$ y suponga que la media muestral de ambas variables es igual a cero. La estimación de mínimos cuadrados del termino constante en dicho modelo es,
    - ✓ Positiva si la estimación de la pendiente β es mayor que cero
    - ✓ Igual a cero
    - ✓ Negativa si la estimación de la pendiente β es menor que cero
    - ✓ Igual a 1

11. ¿Cuál es el estimador de la varianza del término de error de un modelo de regresión?

12. Obtenga una expresión para el estimador de mínimos cuadrados $\hat{\beta}_1$ de la pendiente de la recta de regresión $y_i = \beta_0 + \beta_1 x_i + u_i, i = 1, 2, ..., n$ en función de los datos de la variable y:

$y_i, i = 1, 2, ..., n$ y, a partir de ella, una expresión en función de los términos de error del modelo $u_i, i = 1, 2, ..., n$.

13. Demuestre que el estimador de mínimos cuadrados $\hat{\beta}_1$ de la pendiente de la recta de regresión $y_i = \beta_0 + \beta_1 x_i + u_i, i = 1, 2, ..., n$ es insesgado.

14. Si para explicar la relación existente entre las variables X e Y contemplamos varios modelos alternativos, ¿qué información nos proporciona la matriz de correlaciones entre los residuos de cada modelo y la variable Y? ¿Y los coeficientes de correlación entre los residuos de dos modelos distintos?

15. Considere el modelo constante: $y_i = \alpha + u_i$.
    - o El estimador de mínimos cuadrados del parámetro α es igual a:
    - o Los residuos de esta regresión son iguales a:
    - o La Suma de Cuadrados de residuos de esta estimación es igual a:
    - o El R2 de esta regresión es:

16. Se desea analizar el salario de los trabajadores de una empresa, para lo cual se especifica el modelo: $S_i = \beta_0 + \beta_1 L_i + \beta_2 H_i + \beta_3 (L_i \times H_i) + u_i \quad i = 1, ..., N$, donde S_i es el salario mensual en miles de euros del trabajador *i*, L_i es una variable ficticia que toma valor 1 si el trabajador es Licenciado y 0 en el resto de los casos, y H_i es una variable ficticia que toma valor 1 si el trabajador es un hombre y cero si es una mujer. Según este modelo:

    A) El salario esperado de una mujer licenciada es igual a β_0.
    B) El salario esperado de un hombre licenciado es igual a $\beta_1 + \beta_3$.
    C) El salario esperado de una mujer licenciada es igual a $\beta_0 + \beta_1 + \beta_3$.
    D) El salario esperado de un hombre licenciado es igual a $\beta_0 + \beta_1 + \beta_2 + \beta_3$.

17. Un monopolista se enfrenta a la siguiente curva de demanda: $Q = \alpha + \beta P + u$ donde Q es la cantidad demandada del producto, P es el precio por unidad y u es una perturbación aleatoria. Dados los precios que fijó en los últimos 15 meses y las cantidades vendidas ha estimado el siguiente modelo:

Tabla B

Dependent Variable: Q
Method: Least Squares
Sample: 1 15
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	20.76912	2.821568	7.360845	0.0000
P	-0.840583	0.237563	-3.538363	0.0036

R-squared	0.490596	Mean dependent var	11.46667
Adjusted R-squared	0.451411	S.D. dependent var	5.356794
S.E. of regression	3.967605	F-statistic	12.52002
Sum squared resid	204.6446	Prob(F-statistic)	0.003637

Sabiendo que $\Pr[t(13) \leq 2.16] = 0.975$ y utilizando todos los decimales de la Tabla B:

A) Un intervalo para β al 95% de confianza es (-1.35372, -0.32745).
B) Se rechaza la hipótesis nula de que la pendiente de la curva de demanda es –1 al 5% de significación.
C) Un intervalo para β al 95% de confianza es (-1.07815, -0.60302).
D) No se rechaza la hipótesis nula de que la pendiente de la curva de demanda es 1 al 5% de significación.

18. ¿Cómo se interpreta la estimación del parámetro β en cada uno de los siguientes modelos?
    - El modelo loglineal: $\ln y_i = \alpha + \beta x_i + u_i,\ i = 1, 2, ..., n$
    - Modelo en logaritmos $\ln y_i = \alpha + \beta \ln x_i + u_i,\ i = 1, 2, ..., n$?
    - Modelo semi-logarítmico $y_i = \alpha + \beta \ln x_i + u_i,\ i = 1, 2, ..., n$?

19. El modelo loglineal se especifica: $\ln y_i = \alpha + \beta x_i + u_i,\ i = 1, 2, ..., n$ ¿Cómo se interpreta la estimación del parámetro β? ¿Cómo se interpretaría dicho parámetro en el modelo en logaritmos $\ln y_i = \alpha + \beta \ln x_i + u_i,\ i = 1, 2, ..., n$? ¿Son comparables los R2 de ambos modelos?

20. (3 puntos) Demuestre que el estimador de mínimos cuadrados del modelo de regresión simple $y_i = \alpha + \beta x_i + u_i,\ i = 1, 2, ..., n$ es el estimador lineal e insesgado de mínima varianza.

21. ¿Qué es precisión de un estimador? ¿De qué factores y en qué sentido depende la precisión del estimador de mínimos cuadrados de la pendiente de un modelo de regresión lineal simple?

22. El término de error de un modelo de regresión tiene una distribución Normal. Escriba la distribución de probabilidad del estimador de mínimos cuadrados de la pendiente de dicho modelo. Si la varianza del término de error es desconocida, la distribución Normal no le sirve para contrastar hipótesis ¿qué puede hacer?

23. Define el coeficiente de determinación de un modelo de regresión. Escribe una expresión que es equivalente (casi siempre) a su definición. ¿En qué se basa la igualdad entre ambas expresiones?

24. ¿Cómo se calcula el coeficiente de determinación R2 de la regresión lineal simple? ¿Qué relación tiene con el coeficiente de correlación de las dos variables del modelo de regresión?

25. (3 puntos) Pruebe que el estimador de mínimos cuadrados de la pendiente de una regresión lineal simple $y_i = \beta_0 + \beta_1 x_i + u_i$ puede escribirse como una combinación lineal de las observaciones de la variable y_i, $\hat{\beta} = \sum_{i=1}^{n} \alpha_i y_i$ para determinadas ponderaciones α_i.

    Proporcione las expresiones de estas ponderaciones, y demuestre a que es igual: a) su suma, b) la suma de sus cuadrados, y c) la suma $\sum_{i=1}^{n} \alpha_i x_i$.

Utilice esta representación del estimador para calcular su varianza. Especifique las hipótesis que necesita y en qué momento las utiliza.

26. ¿Cuándo pueden compararse los R2 de dos regresiones distintas?

27. Pruebe que, después de estimar por mínimos cuadrados un modelo de regresión lineal simple, se tiene que la Suma Total (ST) es igual a la Suma Explicada (SE), más la Suma de los Cuadrados de los Residuos (SCR).

28. Pruebe que tanto la suma de los residuos generados por el estimador de mínimos cuadrados en un modelo de regresión lineal simple, como el coeficiente de correlación lineal entre los residuos de mínimos cuadrados y la variable explicativa, son igual a cero

29. (2 puntos) Considere el modelo $y_i = \alpha + \beta x_i + u_i$ y suponga que se decide dividir las observaciones de la variable por 100, obteniendo una nueva variable $y_i^* = y_i / 100$. A continuación, estimamos el modelo $y_i^* = \alpha^* + \beta^* x_i + u_i^*$ y proporciones una respuesta *justificada* a la pregunta: ¿Cómo varían:
    - ✓ el estimador de la pendiente del modelo de regresión,
    - ✓ la estimación del término independiente,
    - ✓ los residuos del modelo estimado,
    - ✓ la estimación de la desviación típica de los residuos,
    - ✓ la desviación típica de dicho estimador,
    - ✓ el coeficiente de determinación R2

30. Considere los modelos $y_i = \alpha + \beta x_i + u_i$ y $y_i^* = \alpha^* + \beta^* x_i^* + u_i^*$ donde $y_i^* = b y_i, x_i^* = c + b x_i$, siendo *b* y *c* números conocidos. En estas condiciones,
    - ✓ $\hat{\alpha}^* = \hat{\alpha} + c\overline{x}$ y $\hat{\beta}^* = \hat{\beta}$
    - ✓ $\hat{\alpha}^* = b\hat{\alpha}$ y $\hat{\beta}^* = b\hat{\beta}$
    - ✓ $\hat{\alpha}^* = \hat{\alpha}$ y $\hat{\beta}^* = \hat{\beta}$
    - ✓ $\hat{\alpha}^* = b\hat{\alpha} - c\hat{\beta}$ y $\hat{\beta}^* = \hat{\beta}$
    - ✓ $\hat{\alpha}^* = \hat{\alpha}$ y $\hat{\beta}^* = c + b\hat{\beta}$

31. Si consideramos las variables de un modelo de regresión lineal simple en desviaciones respecto de sus respectivas medias muestrales, cómo varían, respecto de una regresión utilizando las variables originales,
    - ✓ el estimador de mínimos cuadrados de la pendiente,
    - ✓ su varianza,
    - ✓ el estimador del término independiente,
    - ✓ los residuos del modelo,
    - ✓ el coeficiente de determinación?

32. ¿Cómo se interpreta la pendiente de un modelo de regresión simple cuando alguna de las variables está en logaritmos? ¿Y cuando sólo la variable independiente (explicativa) está en logaritmos? ¿Y cuando es la variable dependiente la única que está en logaritmos?

33. Con respecto a las regresiones $y_i = \beta_0 + \beta_1 x_i + u_i$ y $x_i = \delta_0 + \delta_1 y_i + v_i$, estimados por mínimos cuadrados, indique cuál de las siguientes afirmaciones es cierta:
    - ✓ Las estimaciones de las pendientes son inversas una de la otra
    - ✓ Las estimaciones de las pendientes tienen signo opuesto

- ✓ El valor del R2 es el mismo en ambos modelos
- ✓ El valor de la Suma de cuadrados de los residuos es el mismo en ambos modelos
- ✓ El producto de las pendientes es igual al coeficiente de determinación

34. Si en el modelo $y_i = \alpha + \beta x_i + u_i$, el término de error tiene esperanza cero y varianza constante σ_u^2, y los parámetros son asimismo constantes a lo largo de la muestra, el estimador de mínimos cuadrados de β es,
    - ✓ Un número que coincide con el verdadero valor del parámetro desconocido β
    - ✓ Una constante cuya varianza, de acuerdo con teoremas conocidos, es mínima
    - ✓ Una variable aleatoria cuya distribución de probabilidad esta centrada en torno a cero
    - ✓ Una variable aleatoria cuya varianza es cero

Parte III Contrastación de hipótesis

1. ¿Qué debe suceder para que rechacemos la hipótesis nula en un contraste estadístico de hipótesis?

2. ¿Qué debe observarse para proceder a realizar un contraste de hipótesis unilateral sobre el valor numérico de un coeficiente de un modelo de regresión?

3. ¿Qué información está recogida en el estadístico *t* correspondiente al contraste de hipótesis acerca del valor numérico de un coeficiente en un modelo de regresión? ¿Qué significa un estadístico *t* menor que 2.0 en valor absoluto cuando contrastamos una hipótesis nula al 5% de significación?

4. Escriba la expresión del estadístico para el contraste de la capacidad explicativa global (o la significación global) del modelo. ¿Cuál es su distribución de probabilidad?

5. ¿Qué aspecto tiene la región crítica para el contraste de la hipótesis nula: $H_0 : \beta = 1$ frente a la alternativa: $H_1 : \beta < 1$?

6. Escriba la expresión del estadístico para el contraste de la capacidad explicativa global (o la significación global) del modelo $Y_i = \delta_0 + \delta_1 X_i + \delta_2 Z_i + u_i, i = 1,2,...,n$. ¿Cuál es su distribución de probabilidad?

7. Considere un modelo del tipo $Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + U_i$ en el que se cumplen todas las hipótesis clásicas del MLG. Si $\hat{\beta}_2$ y $\hat{\beta}_3$ son los estimadores MCO de β_2 y de β_3, respectivamente, indique cuál de los estadísticos siguientes sigue una distribución *t* de Student bajo la hipótesis de que $\beta_2 - 3\beta_3 = 1$:

    A) $$\frac{\hat{\beta}_2 - 3\hat{\beta}_3 - 1}{\sqrt{\widehat{\mathrm{Var}}[\hat{\beta}_2]} + 3\sqrt{\widehat{\mathrm{Var}}[\hat{\beta}_3]}}.$$

    B) $$\frac{(\hat{\beta}_2 - 3\hat{\beta}_3 - 1)^2}{\widehat{\mathrm{Var}}[\hat{\beta}_2] + 9\widehat{\mathrm{Var}}[\hat{\beta}_3] - 6\widehat{\mathrm{Cov}}[\hat{\beta}_2, \hat{\beta}_3]}.$$

    C) $$\frac{\hat{\beta}_2 - 3\hat{\beta}_3 - 1}{\sqrt{\widehat{\mathrm{Var}}[\hat{\beta}_2] + 9\widehat{\mathrm{Var}}[\hat{\beta}_3] - 6\widehat{\mathrm{Cov}}[\hat{\beta}_2, \hat{\beta}_3]}}.$$

D) $\dfrac{(\hat{\beta}_2 - 3\hat{\beta}_3 - 1)^2}{\hat{\mathrm{Var}}[\hat{\beta}_2] + 9\,\hat{\mathrm{Var}}[\hat{\beta}_3]}$.

8. Habitualmente, se juzga el contenido informativo de una variable explicativa X a través del valor absoluto del estadístico *t* de Student asociado a su coeficiente. Si es mayor que 2.0 (suponiendo T>50) se dice que la variable X tiene efecto significativo sobre Y, afirmando lo contrario cuando el valor absoluto del estadístico *t* de Student es inferior a 2.0. ¿A qué limitaciones se enfrenta dicha interpretación?

9. ¿Cómo se calcula un rango de efectos de la variable explicativa sobre la variable dependiente de un modelo de regresión a un nivel de confianza determinado, por ejemplo del 95%?

10. En un modelo de regresión simple $y_i = \alpha + \beta x_i + u_i$, hemos estimado $\hat{\beta} = 0,60\,(0,15)$ con T = 100. Obtenga la región crítica para el contraste de la hipótesis $H_0 : \beta = 1,0$ frente a la alternativa $H_1 : \beta < 1,0$ (Test unilateral) al 90% de confianza (10% de significación). Calcule el valor-p del contraste del punto anterior. ¿Rechazaría H_0?

11. Considere el modelo $Y_i = \beta_1 + \beta_2 X_i + U_i \;\; (i = 1, ..., 30)$ en el que se cumplen todas las hipótesis clásicas del MLG. Si $\bar{t}$ representa el valor calculado del estadístico *t* habitual para el contraste de $H_0 : \beta_2 = 1$ frente a $H_1 : \beta_2 > 1$, indique cuál de las afirmaciones siguientes es CIERTA:
    - A) El nivel de significación marginal (*p-value*) asociado con el contraste anterior es igual a $\Pr[t(28) \geq \bar{t}\,]$.
    - B) El valor calculado del estadístico *t* es $\bar{t} = \hat{\beta}_2 \,/\, EE_2$, donde EE_2 representa el error estándar (o desviación típica estimada) del estimador MCO de β_2.
    - C) El nivel de significación marginal (*p-value*) asociado con el contraste anterior es igual a $1 - \Pr[t(28) \geq \bar{t}\,]$.
    - D) El valor calculado del estadístico *t* es $\bar{t} = (\hat{\beta}_2 - 1) \times EE_2$, donde EE_2 representa el error estándar (o desviación típica estimada) del estimador MCO de β_2.

12. (2 puntos) Demuestre a qué se reduce la expresión general del estadístico *F* para la contrastación de hipótesis lineales sobre los parámetros del modelo de regresión, cuando se contrasta la significación global del modelo

13. (1,5 puntos) En el modelo: $y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + u_i, i = 1,2,...,n$ quiere contrastar la restricción: H0: $\beta_1 + 2\beta_2 = 3$ frente a la alternativa: H1: $\beta_1 + 2\beta_2 \neq 3$ siguiendo el procedimiento de sustituir la restricción en el modelo. Explique cómo lo haría (0,5), cuál sería el estadístico utilizado (0,5) y cuál su distribución de probabilidad (0,5).

14. (1,5 puntos) Explique cómo contrastaría la hipótesis anterior mediante el análisis de la *holgura* de dicha restricción. ¿Qué información necesita y cómo la conseguiría en EVIEWS? ¿Cuál sería el estadístico y cuál su distribución de probabilidad? Lleve a cabo el contraste de hipótesis bajo el supuesto de que estimando con 80 datos, ha obtenido $\hat{\beta}_1 = -0,80$; $\hat{\beta}_2 = 3,10$ y la matriz de covarianzas de los estimadores:

30720.4	-140.0	-1033.8
-140.0	0.83	1.04
-1033.8	1.04	1.47

15. (2 puntos) Sea el modelo estimado:

Dependent Variable: TB10
Sample: 1 400

Variable	Coefficient	Std. Error	t-Statistic
C	2.96	0.14	20.24
FF	0.26	0.11	2.36

y denotamos por β el coeficiente de la variable FF. Queremos contrastar la hipótesis H0: $\beta = 0$ frente a la alternativa: H1: $\beta > 0$.
¿Qué forma tiene la región crítica? Haga un gráfico.
Construya dicha Región Crítica para un nivel de significación: $\alpha = 1\%$. ¿Rechazaría H0? Y al 5% y 10%? Calcule el valor-p del contraste.

16. En el modelo: $y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + u_i, i = 1, 2, ..., n$ quiere contrastar la restricción: H0: $\beta_1 + 2\beta_2 = 3$ frente a la alternativa: H1: $\beta_1 + 2\beta_2 \neq 3$ siguiendo el procedimiento de sustituir la restricción en el modelo. Explique cómo lo haría, cuál sería el estadístico utilizado y cuál su distribución de probabilidad.

17. Las preguntas 3 a 6 siguientes se refieren al enunciado siguiente: Utilizando una sección cruzada con datos sobre gasto *per cápita* anual en escuelas públicas (GASTO) y renta *per cápita* anual (RENTA) en 50 estados de los Estados Unidos (con ambas variables medidas en dólares), EViews ha proporcionado la siguiente información (donde LOG en la Tabla C representa el logaritmo neperiano):

Tabla C

Dependent Variable: LOG(GASTO)
Method: Least Squares
Date: 12/16/02 Time: 11:32
Sample: 1 50
Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.645352	0.148536	31.27417	0.0000
RENTA	0.000164	1.93E-05	8.499495	0.0000

R-squared	0.600803	Mean dependent var	5.896203
Adjusted R-squared	0.592487	S.D. dependent var	0.222840
S.E. of regression	0.142254	Akaike info criterion	-1.023230
Sum squared resid	0.971335	Schwarz criterion	-0.946749
Log likelihood	27.58075	F-statistic	72.24142
Durbin-Watson stat	2.205026	Prob(F-statistic)	0.000000

Pregunta 3. Utilizando la información que figura en la Tabla C, ¿son correctas o falsas las afirmaciones?:

A) Un aumento de 1000 dólares en la RENTA implica un aumento esperado en el GASTO aproximadamente igual al 1.64%.

B) Una reducción de 100 dólares en la RENTA implica un aumento esperado en el GASTO aproximadamente igual al 1.64%.

C) Un aumento de un 1% en la RENTA implica un aumento esperado en el GASTO aproximadamente igual al 0.000164%.

D) Un aumento de 10 dólares en la RENTA implica un aumento esperado en el GASTO aproximadamente igual al 0.164%.

Pregunta 4. Utilizando la información que figura en la Tabla C, **¿es correcto o falso?**:

E) La varianza residual estimada por MCO es 0.020236.

F) La variación observada en la RENTA explica aproximadamente el 60% de la variación observada en el GASTO.

G) La variación observada en la RENTA explica aproximadamente el 14.22% de la variación observada en el LOG(GASTO).

H) La desviación típica muestral de la serie GASTO es 0.222840.

Pregunta 5. Utilizando la información que figura en la Tabla C, si la variable RENTA hubiera estado medida en miles de dólares en vez de en dólares:

A) La estimación de la pendiente del modelo habría sido 0.00164.

B) El valor calculado del estadístico t del contraste de significación de la pendiente no habría cambiado.

C) La estimación de la pendiente del modelo habría sido 0.0164.

D) La estimación del término constante del modelo habría sido 4645.352.

Pregunta 6. Si el valor del estadístico de Jarque-Bera calculado con los residuos de la estimación del modelo de la Tabla C es igual a 3.97, el nivel de significación marginal (*p-value*) asociado con el contraste de normalidad de las perturbaciones del modelo debe calcularse como:

A) $\Pr[\chi^2(3) > 3.97]$.

B) $\Pr[t(48) > 3.97]$.

C) $\Pr[\chi^2(2) > 3.97]$.

D) $\Pr[\chi^2(2) \leq 3.97]$.

16. En el modelo de regresión $y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t$ se quiere contrastar la hipótesis nula: $H_0 : \beta_2 = 1$ frente a la hipótesis alternativa $H_0 : \beta_2 \neq 1$. ¿Cómo llevaría a cabo el contraste de dicha hipótesis? ¿Cuáles serían las sumas de cuadrados de los modelos restringidos y sin restringir? ¿Cuál sería la expresión del estadístico apropiado y cuál sería su distribución de probabilidad?

Parte IV. Modelo de regresión múltiple

1. Escriba la expresión para la varianza del estimador del coeficiente de una variable X en un modelo de regresión que explica el comportamiento de Y utilizando X y Z como variables explicativas. ¿Es esta varianza mayor o menor que la que obtendría si explicase Y utilizando únicamente X como variable explicativa?

2. Para explicar una variable Y dispone de datos sobre 5 variables X1, X2, X3, X4, X5. En principio, creemos que todas ellas pueden contiene información útil sobre Y. Calcula las correlaciones entre estas variables y observa que son todas superiores a 0,30. Describa el procedimiento por el que iría elaborando un modelo de regresión para explicar Y.

3. Dadas variables Y, X, Z, defina el coeficiente de correlación entre Y y X parcial en la información contenida en Z. ¿Cómo lo calcularía?

4. Expresión del coeficiente de determinación ajustado. ¿Por qué se ajusta el coeficiente de determinación?

5. Utilizando datos anuales desde 1940 a 2006, ha estimado un modelo que explica el consumo de gasolina para automóviles en función del número de coches matriculados y de la renta per capita. Le han pedido que opine sobre si la introducción en gasolineras en el año 2001 de carburantes elaborados a partir de productos agrícolas ha reducido el consumo de gasolina. Defina qué debe entenderse por "*una reducción en el consumo de gasolina a partir de dicha fecha*". ¿Cómo podría contrastar esta posibilidad utilizando su modelo estimado?

6. Para caracterizar los determinantes del consumo de gasolina tiene datos sobre el consumo total, en todo el país, en miles de litros, la población, el número de coches matriculados, y el PIB. Razone por qué cabría esperar que todas estas variables muestren unos coeficientes de correlación lineal muy elevados. En este caso, ¿a qué problema nos exponemos estimando el modelo con las variables mencionadas? ¿Cómo debería transformar las variables para estimar un modelo razonable?

7. ¿Cómo contrastaría en el modelo $Y_i = \delta_0 + \delta_1 X_i + \delta_2 Z_i + v_i$, $i = 1,2,...,n$ la hipótesis nula $H_0 : \delta_1 = \delta_2$ frente a la alternativa: $H_0 : \delta_1 \neq \delta_2$?
   a. mediante el estadístico t
   b. mediante sustitución de dicha restricción en el modelo

8. Cómo contrastaría mediante sustitución en el modelo $y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + u_i$ la restricción: $H_0 : \beta_2 + \beta_3 = 0$?

9. ¿Cuál es la expresión general del estadístico para el contraste de múltiples hipótesis sobre los coeficientes de un modelo de regresión? Explique todos los elementos que aparecen en dicha expresión.

10. Hemos especificado un modelo para explicar el comportamiento de una variable y_i. Si incluimos dicha variable asimismo como variable explicativa (a la derecha de la ecuación): $y_i = \alpha + \beta x_i + \delta z_i + \gamma y_i + u_i$, $i = 1,2,...,n$ ¿a qué serán igual los coeficientes estimados para cada una de las variables explicativas? ¿a qué será igual el término constante estimado? ¿a qué serán iguales los residuos? ¿y la Suma de cuadrados de residuos? ¿Y el R2?

11. ¿Por qué se utiliza un coeficiente de determinación corregido de grados de libertad? ¿Cómo se calcula dicho coeficiente?

12. ¿Cuál es el estadístico general para la contrastación de una o más restricciones entre coeficientes de un modelo de regresión múltiple? ¿Cuál es su expresión particular en el caso de que se contraste la significación global del modelo de regresión?

13. (2 puntos) Para analizar la posible existencia de discriminación salarial, se ha estimado el modelo: $salario_i = \hat{\beta}_0 + \hat{\beta}_1 \left(male_i\right) + \hat{\beta}_2 educación_i + \hat{\beta}_3 \left(male_i * educación_i\right) + \hat{u}_i$ donde la variable *male* se ha construido de modo que tome el valor 1 en los trabajadores varones, siendo igual a 0 para las mujeres trabajadoras.¨
    - ¿Qué signo esperaría que tomasen los 4 coeficientes del modelo?

- ¿cómo utilizaría los valores numéricos de los cuatro coeficientes para estimar el salario de un trabajador varón con nivel educativo igual a 2? ¿y el de una mujer de igual nivel educativo? Proporcione las expresiones que utilizaría para dicho cálculo.
- ¿Cómo contrastaría la existencia de discriminación salarial?
-

14. [**Sobre este enunciado se hacen las Preguntas 1 a 5 que aparecen posteriormente**] Utilizando una muestra de 75 hogares o unidades de gasto de la Encuesta de Presupuestos Familiares de 1973-1974 elaborada por el INE se dispone de información sobre el:
- gasto total en consumo de las familias en pesetas (GTINE),
- ingresos totales del hogar según el INE en pesetas (ITINE),
- número de perceptores de ingresos regulares en el hogar (NPER),
- número de personas que forman el hogar o unidad de gasto (NPHO)
- nivel educativo del sustentador principal representado por cuatro variables ficticias:
  - EDU1, toma el valor 1 si el cabeza de familia no tiene estudios y cero en caso contrario;
  - EDU2 toma el valor 1 si el máximo grado educativo del cabeza de familia es el bachillerato y cero en caso contrario; y
  - EDU3 que toma el valor 1 si el máximo grado educativo del cabeza de familia es una diplomatura, y cero en caso contrario
  - EDU4 toma el valor 1 si el máximo grado educativo del cabeza de familia es una licenciatura universitaria, y creo en caso contrario.

  Note que, para cada una de las familias en la muestra, la suma de las cuatro variables EDU1 + EDU2 + EDU3 + EDU4 es igual a 1.

Utilizando EVIEWS se han estimado dos modelos que aparecen en las Tablas 1 y 2:

Tabla 1

Dependent Variable: GTINE
Method: Least Squares
Date: 12/11/02 Time: 11:59
Sample: 1 75
Included observations: 75

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	195834.8	68927.52	2.841170	0.0059
ITINE	0.490579	0.128916	3.805406	0.0003
NPER	3767.697	14870.57	0.253366	0.8007
NPHO	25035.66	10252.14	2.441993	0.0172
EDU1	-171378.3	79325.11	-2.160455	0.0343
EDU2	-185909.8	68679.83	-2.706905	0.0086
EDU3	-164628.1	57214.43	-2.877387	0.0054

R-squared	0.577101	Mean dependent var	261276.6
Adjusted R-squared	0.539786	S.D. dependent var	170340.5
S.E. of regression	115557.4	Akaike info criterion	26.24161
Sum squared resid	9.08E+11	Schwarz criterion	26.45791
Log likelihood	-977.0603	F-statistic	15.46581
Durbin-Watson stat	1.938365	Prob(F-statistic)	0.000000

Tabla 2

Dependent Variable: GTINE
Method: Least Squares
Date: 12/11/02 Time: 11:59
Sample: 1 75
Included observations: 75

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	193536.2	67862.93	2.851869	0.0057
ITINE	0.502111	0.119794	4.191454	0.0001
NPHO	25924.82	9567.242	2.709749	0.0085
EDU1	-168300.6	77856.17	-2.161687	0.0341
EDU2	-181481.8	65967.12	-2.751095	0.0076
EDU3	-161768.9	55708.89	-2.903826	0.0049

R-squared	0.576702	Mean dependent var	261276.6
Adjusted R-squared	0.546028	S.D. dependent var	170340.5
S.E. of regression	114771.1	Akaike info criterion	26.21589
Sum squared resid	9.09E+11	Schwarz criterion	26.40128
Log likelihood	-977.0957	F-statistic	18.80112
Durbin-Watson stat	1.945676	Prob(F-statistic)	0.000000

Pregunta 1: Según la estimación que aparece en la Tabla 1 cuál de las siguientes proposiciones es la CORRECTA:

A) No se puede rechazar al 1 % la hipótesis de que el parámetro que acompaña a la variable NPHO sea estadísticamente igual a cero.
B) Se puede rechazar al 1% la hipótesis de que el parámetro que acompaña a la variable NPHO sea estadísticamente igual a cero.
C) No se puede rechazar al 5% la hipótesis de que el parámetro que acompaña a la variable NPHO sea estadísticamente igual a cero, pero si al 1 %.
D) No se puede rechazar ni al 1% ni al 5% que el parámetro que acompaña a la variable NPHO sea estadísticamente distinto de cero

Pregunta 2: Según la información de la Tabla 2 indique cuál de las siguientes afirmaciones es la CORRECTA:

A) No hay información suficiente para contrastar si todos los parámetros excepto la constante son iguales a cero
B) Hay información suficiente para contrastar si todos los parámetros excepto la constante son iguales a cero y no se puede rechazar esta hipótesis
C) Hay información suficiente para contrastar si todos los parámetros excepto la constante son iguales a cero y se puede rechazar esta hipótesis
D) No hay información suficiente para contrastar si todos los parámetros excepto la constante son iguales a cero porque necesitaría la Suma residual del Modelo Restringido

Pregunta 3: Dada la información que aparece en la Tablas 1 y 2 indique cuál es la afirmación **CORRECTA:**

A) Es mejor el modelo de la Tabla 1 porque tiene un R^2 superior.
B) No tengo información suficiente para identificar cual de los dos modelos tiene mayor capacidad explicativa.
C) Tiene mayor capacidad explicativa el de la Tabla 1 porque en este caso la Suma Total es más elevada que en el caso de la Tabla 2
D) Es mejor el modelo de la Tabla 2 porque el R^2 ajustado es más elevado.

Pregunta 4: Atendiendo a los resultados de las Tabla 2, escriba las ecuaciones que le permitirían estimar el gasto en consumo de 4 familias. Todas ellas están compuestas por 5 personas y tienen dos perceptores de renta, siendo los ingresos mensuales de 400.000 pesetas. El sustentador principal de la primera familia no tiene estudios; el de la segunda familia tiene estudios primarios, el de la tercera familia tiene bachillerato elemental, y el de la cuarta familia tiene estudios superiores.

Pregunta 5: Utilizando la Suma de Cuadrados de los Residuos de los modelos estimados en la Tabla 1 y 2, la hipótesis de que el parámetro que acompaña a la variable NPER es estadísticamente igual a cero:

A) No puede rechazarse al 5 % pero si al 10 %.
B) Puede rechazarse al 5 % y al 10 %.
C) No puede realizarse el contraste utilizando las Sumas de Cuadrados de los Residuos.
D) No se rechaza ni al 5% ni al 10 %

15. En el modelo de regresión $y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t$ se quiere contrastar la hipótesis nula: $H_0 : \beta_2 = \beta_3$ frente a la alternativa $H_0 : \beta_2 \neq \beta_3$ ¿Cómo llevaría a cabo el contraste de dicha hipótesis? ¿Cuáles serían las sumas de cuadrados de los modelos restringidos y sin restringir? ¿Cuál sería la expresión del estadístico apropiado y cuál sería su distribución de probabilidad?

16. Las ventas de un producto a lo largo de una red comercial dependen de la renta per cápita y del precio que la empresa fija para el producto. La empresa fija un precio más alto en las zonas de mayor renta per cápita y lo va reduciendo gradualmente según considera zonas de menor renta. Dispone de datos de renta per cápita y ventas para los 300 establecimientos de la red comercial.
    - ¿Qué coeficiente de correlación esperaría observar entre los datos de precio y renta per cápita?
    - ¿Cómo discutiría qué variable (renta o precio) es más importante para explicar las ventas del producto?
    - Suponga que estima un modelo que explica las ventas mediante el precio del producto y la renta per cápita ¿cómo contrastaría la hipótesis de que las elasticidades precio y renta son iguales y de signo contrario?
    - ¿Tendrá alguna dificultad en la interpretación individual de los coeficientes estimados en la regresión anterior?
    - ¿Cómo podría estimar el efecto del precio del producto sobre las ventas del producto, descontando la relación que la renta per cápita guarda con ambas variables?
    - Especifique un modelo para estimar la elasticidad-precio de las ventas, permitiendo distinta elasticidad en zonas rurales que en zonas urbanas.
    - Describa cómo obtendría la elasticidad-precio en cada uno de los tipos de zona, una vez estimado el modelo.
    - ¿Cómo contrastaría la hipótesis de que la elasticidad-precio es la misma en zonas rurales que en zonas urbanas?
    - El director comercial considera que, incluso a igualdad de renta y precio, las ventas no serían las mismas en una zona rural que en una zona urbana. ¿Cómo recogería este posible efecto en su modelo?

17. Escriba la expresión general del estadístico F para el contraste de hipótesis lineales entre coeficientes del modelo de regresión múltiple. Cuando se quiere contrastar la significación estadística de un solo coeficiente ¿qué relación existe entre el estadístico F y el estadistico t de Student de dicho coeficiente?