



Dirección de Docencia
Universidad de Concepción

Facultad de Ciencias
Económicas y Administrativas

Departamento
de Economía

Ejercicios Propuestos y Resueltos de Econometría

PRIMERA EDICIÓN

Leonardo Salazar Vergara & Manuel Estay Montecinos

EJERCICIOS PROPUESTOS Y RESUELTOS DE ECONOMETRÍA

**LEONARDO SALAZAR VERGARA
MANUEL ESTAY MONTECINOS**

Registro de Propiedad Intelectual N° 2025-P-7269 año 2025

ISBN: 978-956-9280-60-3

**Prohibida la reproducción total o parcial de esta obra
©UNIVERSIDAD DE CONCEPCION**

Ejercicios Propuestos y Resueltos de Econometría

Ejercicios Propuestos y Resueltos de Econometría

PRIMERA EDICIÓN

Leonardo Salazar & Manuel Estay

DEPARTAMENTO DE ECONOMÍA

Prefacio

La econometría es una disciplina esencial en el análisis de fenómenos económicos, pues permite aplicar métodos cuantitativos para entender las relaciones entre variables y evaluar teorías económicas con base en datos reales. Su importancia se ha incrementado en las últimas décadas, dado el creciente volumen de información disponible y la necesidad de realizar predicciones más precisas para la toma de decisiones en distintos ámbitos, desde la política pública hasta las estrategias empresariales.

El presente libro de Ejercicios Propuestos y Resueltos de Econometría tiene como objetivo servir como una herramienta práctica para aquellas y aquellos estudiantes de pregrado que cursen las asignaturas iniciales de econometría. A través de una serie de problemas resueltos y propuestos, los lectores tendrán la oportunidad de enfrentar diferentes situaciones empíricas y poner en práctica los conceptos adquiridos en clase.

El enfoque de este libro no es únicamente brindar soluciones numéricas, sino también explicar el razonamiento detrás de cada paso, proporcionando así una comprensión profunda de los métodos utilizados. Se busca, además, motivar a las y los estudiantes a pensar respecto a los supuestos involucrados y la interpretación de los resultados, de modo que puedan aplicar estos conocimientos de manera efectiva en sus investigaciones o proyectos profesionales.

El libro aborda temas clave de la econometría dentro de la variedad de los modelos de regresión lineal simple y múltiple estimados mediante Mínimos Cuadrados Ordinarios. Las soluciones detalladas permitirán a las y los lectores verificar sus propios cálculos y mejorar su capacidad de interpretación de los resultados, mientras que los ejercicios propuestos ofrecerán la oportunidad de reforzar lo aprendido a través de la práctica.

Este libro es el resultado de un proyecto docencia *Escribe.DOC* (2023047) de la Universidad de Concepción y está basado en la recopilación de ejercicios de certámenes y ejercicios de clases teóricas y prácticas que se han usado en los cursos de pregrado de econometría. El libro está diseñado para estudiantes de econometría de la carrera Ingeniería Comercial de la Universidad de Concepción.

Confiamos en que el material aquí presentado contribuya al desarrollo de una comprensión sólida de la econometría y a la formación de analistas capaces de afrontar los desafíos económicos del mundo actual.

Cualquier error puede ser comunicado a los autores a través de correo electrónico.

—Leonardo Salazar (lesalaza@udec.cl) & Manuel Estay (mestay@udec.cl)

Índice general

Prefacio	v
1 Ejercicios Propuestos y Resueltos	1
2 Ejercicios Propuestos	33
Agradecimientos	41

Capítulo 1

Ejercicios Propuestos y Resueltos

1. Analice si las siguientes proposiciones son verdaderas o falsas. Sólo justifique las falsas:

1.1 En presencia de heteroscedasticidad es esperable que R^2 del modelo sea alto y que sobrestime al verdadero R^2 .

- Respuesta: Falso: Lo anterior es cierto cuando existe autocorrelación y/o multicolinealidad.

1.2 Se tienen dos modelos de regresión estimados mediante Mínimos Cuadrados Ordinarios. Modelo 1: $\hat{Y}_i = 0,8 + 2,5X_{i2}$ $\bar{R}^2 = 0,7$ y Modelo 2: $\ln \hat{Y}_i = 0,5 + 0,3X_{i2} - 1,3X_{i3}$ $\bar{R}^2 = 0,9$. Entonces, se puede afirmar que el Modelo 2 es mejor que el Modelo 1 ya que $\bar{R}^2$ del Modelo 2 es mayor al $\bar{R}^2$ del Modelo 1.

- Respuesta: Falso. Cuando la variable dependiente no es la misma, no se debe usar el $\bar{R}^2$ para hacer comparaciones ($Y_i \neq \ln Y_i$).

1.3 En presencia de multicolinealidad los estimadores de Mínimos Cuadrados Ordinarios son los mejores estimadores lineales e insesgados (MELI).

- Respuesta: Verdadero.

1.4 En presencia de heteroscedasticidad es muy probable que. (i) el R^2 sea muy alto, (ii) existan muy pocos β_s significativos y que (iii) la hipótesis nula de la prueba de significancia global sea rechazada.

- Respuesta: Falso. Lo indicado es más probable que ocurra en presencia de multicolinealidad.

1.5 Se tiene el siguiente modelo estimado mediante Mínimos Cuadrados Ordinarios: $\hat{Y}_t = 0,9 + 0,3X_{t2} - 2,5X_{t3} + 0,7Y_{t-1}$ $T = 80$ $DW = 0,2$, donde DW es el estadístico de Durbin-Watson. Entonces se puede asegurar que los residuos de la regresión sufren de autocorrelación serial positiva.

- Respuesta: Falso. La información proporcionada no permite evaluar la existencia de autocorrelación serial. El estadístico de Durbin-Watson, DW, no se debe usar cuando el modelo incorpora como variable explicativa rezagos de la variable dependiente. Esto es lo que ocurre en el modelo planteado ya que en la parte derecha aparece el primer rezago de la variable Y_t , es decir Y_{t-1} .

1.6 Una forma de corregir la no normalidad de los residuos es estimar el modelo mediante el método de la varianza de Newey-West.

- Respuesta: Falso. El Método de Newey-West sirve para corregir las varianzas y covarianzas de los parámetros estimados $\hat{\beta}_s$ en presencia de autocorrelación y heteroscedasticidad. Una forma de corregir la no normalidad es mediante el aumento del tamaño de la muestra y/o eliminación de datos atípicos.

1.7 El estadístico t calculado de una prueba de significancia individual no puede ser negativo debido a que la desviación estándar del parámetro siempre debe ser positiva.

- Respuesta: Falso. El estadístico t calculado para el parámetro β_i en una prueba de significancia individual, se define como $t_{calculado} = \hat{\beta}_i / \sqrt{\text{var}(\hat{\beta}_i)}$. Por lo tanto, el t calculado puede ser negativo si $\hat{\beta}_i$ es negativo.

1.8 Si el Error Cuadrático Medio (ECM) de un estimador es igual 0,6 y el sesgo es igual a -0,2, entonces la varianza de dicho estimador debe ser igual a 0,7.

- Respuesta: Falso. La varianza debe ser igual a 0,56.

$$ECM = [\text{Sesgo}]^2 + \text{Var}$$

$$0,6 = [-0,2]^2 + \text{Var}$$

$$\text{Var} = 0,6 - [-0,2]^2$$

$$\text{Var} = 0,56$$

1.9 El supuesto de no existencia de autocorrelación es necesario para demostrar que los estimadores de mínimos cuadrados ordinarios son insesgados.

- Respuesta: Falso. El supuesto de no existencia de autocorrelación es necesario para demostrar la eficiencia de los estimadores de mínimos cuadrados ordinarios.

1.10 Si el modelo de regresión a estimar es $Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \mu_i$, entonces la regresión auxiliar para efectuar la prueba de heteroscedasticidad de White es:

$$\hat{\mu}_i^2 = \gamma_1 + \gamma_2 X_{i2} + \gamma_3 X_{i3} + \gamma_4 X_{i2}^2 + \gamma_5 X_{i3}^2 + \varepsilon_i$$

- Respuesta: Falso, la regresión auxiliar es $\hat{\mu}_i^2 = \gamma_1 + \gamma_2 X_{i2} + \gamma_3 X_{i3} + \gamma_4 X_{i2}^2 + \gamma_5 X_{i3}^2 + \gamma_6 (X_{i2} \cdot X_{i3}) + \varepsilon_i$.

1.11 En el modelo de regresión $Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \mu_i$, donde Y_i es el ingreso del individuo, $X_{i2} = 1$ si el individuo es mujer (= 0 en otro caso), $X_{i3} = 1$ si el individuo es hombre (= 0 en otro caso) y X_{i4} es la escolaridad del individuo, es muy probable que β_2 y β_3 sean altamente significativos debido a que el ingreso depende en gran medida del sexo de la persona.

- Respuesta: Falso, el modelo no puede ser estimado debido a que existe multicolinealidad perfecta (trampa variable dicotómica).

1.12 Si en el modelo de regresión $\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_{t2} + \hat{\beta}_3 X_{t3} + \hat{\beta}_4 X_{t4}$ se quiere probar la hipótesis

$$H_0 \begin{cases} \beta_2 & = 0,7 \\ \beta_4 + \beta_3 & = 1,0 \end{cases}$$

entonces se puede aplicar una prueba t , F y/o χ^2 para evaluar dicha hipótesis.

- Respuesta: Falso, en este caso el número de restricciones es igual a 2, por lo tanto, no se puede aplicar una prueba “t”; solo F y/o χ^2 se pueden aplicar.

1.13 Suponga que el dueño de un negocio está interesado en estimar una función de demanda de su único producto que vende. Una empresa de estudios propone estimar la demanda a través de la siguiente regresión: $\ln(Y_t) = \beta_1 + \beta_2^2 \ln(P_t) + \beta_3 \ln(M_t) + \mu_t$ donde Y_t corresponde a la demanda del producto (medido en unidades), P_t es el precio del producto (medido en pesos) y M_t es el ingreso de los consumidores (medido en pesos). La empresa de estudios propone usar el método de Mínimos Cuadrados Ordinarios para estimar dicha función e indica que por Ley de la Demanda se espera que β_2 sea negativo y que, por lo tanto, β_2^2 sea positivo. ¿Lo propuesto e indicado por la empresa de estudios es correcto?

- Respuesta: Falso. La función a estimar no es lineal en los parámetros y no pueden ser estimada a través de Mínimos Cuadrados Ordinarios. Adicionalmente si $\beta_2^2 > 0$,

entonces $\beta_2 > 0$ ó $\beta_2 < 0$.

- 1.14 La teoría económica, en realidad, no es muy útil cuando se debe estimar un modelo de regresión debido a que para calcular el valor de los parámetros se necesita una fórmula matemática y no una teoría económica.

- Respuesta: Falso. La teoría económica es necesaria ya que indica las variables que intervienen en una regresión. Además indica cual variable es la explicada y cuales son las explicativas.

- 1.15 Si en un modelo de regresión lineal simple se verifica que $\hat{\beta}_2 > 0$, entonces $Cov(X, Y) < 0$.

- Respuesta. Falso: En el modelo de regresión lineal simple, $Y_i = \beta_1 + \beta_2 X_i + \mu_i$, estimado mediante mínimos cuadrados ordinarios, se verifica que

$$\hat{\beta}_2 = \frac{Cov(Y, X)}{Var(X)}$$

Luego, si $\hat{\beta}_2 > 0$ se debe necesariamente verificar que $Cov(Y, X) > 0$ ya que, por definición, $Var(X) > 0$.

- 1.16 El criterio de eficiencia permite comparar cualquier tipo de estimador para evaluar el que posee la menor varianza.

- Respuesta: Falso. El criterio de eficiencia sólo permite comparar la varianza de estimadores que son insesgados.

- 1.17 El modelo de regresión $Y_i = \beta_1 + \beta_2 \sqrt{X_i}$ no puede ser estimado por mínimos cuadrados ordinarios ya que la variable Y no es lineal en X .

- Respuesta: Falso. El modelo debe ser lineal en los parámetros β_1, β_2 para ser estimado mediante Mínimos Cuadrados Ordinarios. En el presente caso se satisface dicha condición.

- 1.18 El supuesto de homoscedasticidad implica que la $Cov(\mu_i, \mu_j) = 0$ para todo $i \neq j$.

- Respuesta: Falso. El supuesto de homoscedasticidad implica que la varianza de las perturbación estocástica, μ_i , es constante $Var(\mu_i) = \sigma^2$ para todo i . El enunciado corresponde al supuesto de no existencia de autocorrelación.

- 1.19 Una empresa necesita estimar una función de ventas. Para hacer lo anterior, le proporciona datos de sus ventas anuales de su único artículo (Q), precio promedio anual de

dicho artículo (P_Q), ingresos promedio anual de sus consumidores (M) y del precio promedio anual de su principal sustituto (P_S). Usted propone estimar la siguiente función de demanda $Q_t = \beta_1 + \beta_2 P_{Q,t} + \beta_3 P_{S,t} + \beta_4 M_t + \mu_t$. Si la empresa le proporciona datos anuales entre 1990 y 1992 (incluyendo ambos años), entonces la función puede ser estimada mediante Mínimos Cuadrados Ordinarios.

- Respuesta: Falso. El tamaño de la muestra es $T = N = 3$ y el número de parámetros a estimar es $K = 4$. Por lo tanto se viola la condición de rango ya que $N < K$ y el modelo no puede ser estimado mediante Mínimos Cuadrados Ordinarios.

1.20 Si la función de regresión poblacional está dada por $Y_i = \beta X_i + \mu_i$ (no posee intercepto), entonces el estimador de mínimos cuadrados ordinarios β de será igual a $\hat{\beta} = \frac{\sum_{i=1}^n Y_i^2}{\sum_{i=1}^n X_i^2}$.

- Respuesta: Falso.

$$\min_{\{\hat{\beta}\}} \sum \hat{\mu}_i^2 = \sum (Y_i - \hat{\beta} X_i)^2$$

$$\frac{\partial \sum \hat{\mu}_i^2}{\partial \hat{\beta}} = 0 \implies 2 \sum (Y_i - \hat{\beta} X_i) \cdot -X_i = 0$$

$$\sum (Y_i X_i - \hat{\beta} X_i^2) = 0$$

$$\sum Y_i X_i - \hat{\beta} \sum X_i^2 = 0$$

$$\hat{\beta} = \frac{\sum Y_i X_i}{\sum X_i^2}$$

1.21 La hipótesis nula de test de Jarque-Bera (JB) indica que no existe heteroscedasticidad en los residuos de una regresión lineal.

- Respuesta: Falso. La hipótesis nula de la prueba JB indican que los residuos de la regresión de Mínimos Cuadrados Ordinarios poseen distribución Normal.

1.22 En la regresión lineal, $Y = X\beta + \mu$, estimada a través de Mínimos Cuadrados Ordinarios, el supuesto de homoscedaticidad del error es necesario para demostrar el insesgamiento de $\hat{\beta}$.

- Respuesta: Falso. El supuesto de homoscedaticidad es necesario para demostrar la eficiencia del estimador. Para demostrar el insesgamiento, se requiera $E[\mu] = 0$.

1.23 Si en un modelo de regresión lineal simple se verifica que $R^2 = 0$, entonces se debe cumplir que $Var(X) = 0$.

- Respuesta: Falso. Si $Var(X) = 0$, el R^2 se indetermina debido a que $R^2 = \frac{[Cov(Y,X)]^2}{Var(Y)Var(X)}$. Luego, si $R^2 = 0$, se debe verificar que $Cov(Y, X) = 0$, lo cual implicaría que no existe una relación lineal entre las variables.

1.24 Cuando el tamaño de la muestra es $N = \infty$ entonces $\bar{R}^2 < R^2$.

- Respuesta: Falso. Cuando el tamaño de la muestra es infinito, la pérdida de grados de libertad no es significativa. Por lo tanto $\bar{R}^2 = R^2$.

1.25 Se denomina autocorrelación espacial cuando se encuentra autocorrelación en una regresión que involucre el uso de datos de corte transversal.

- Respuesta: Verdadero.

1.26 Suponga que usted estimó el siguiente modelo mediante Mínimos Cuadrados Ordinarios: $Y_t = \beta_1 + \beta_2 X_t + \mu_t$, donde Y_t es el ingreso por ventas mensual (medido en pesos) y X_t es el precio de venta promedio mensual (medido en pesos por unidad). Sin embargo, debido a un error de datos, usted olvidó agregar un monto de dinero fijo, "F", que se recibió cada mes. Esto es, la nueva variable dependiente del modelo a estimar debería ser $Y_t^* = Y_t + F$. Entonces al usar esta nueva variable, se obtendrá un nuevo estimador $\hat{\beta}_2^*$ tal que $\hat{\beta}_2^* = \hat{\beta}_2$.

- Respuesta: Falso Modelo original: $\hat{\beta}_2 = \frac{Cov(Y,X)}{Var(X)}$
 Modelo modificado: $\hat{\beta}_2^* = \frac{Cov(Y^*,X)}{Var(X)} = \frac{E[(Y^* - E(Y^*))(X - E(X))]}{Var(X)} = \frac{E[(Y + F - E(Y) - F)(X - E(X))]}{Var(X)}$

$$\hat{\beta}_2^* = \frac{E[(Y + F - E(Y) - F)(X - E(X))]}{Var(X)} =$$

$$\hat{\beta}_2^* = \frac{E[(Y - E(Y))(X - E(X))]}{Var(X)}$$

$$\hat{\beta}_2^* = \frac{Cov(Y, X)}{Var(X)}$$

$$\hat{\beta}_2^* = \hat{\beta}_2$$

Si sólo a las observaciones de la variable Y se le agrega una constante, el valor $\hat{\beta}_2$ estimado no experimentara cambio.

- 1.27 En el modelo $Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \mu_i$, estimado mediante mínimos cuadrados ordinarios, la hipótesis:

$$H_0 \begin{cases} \beta_2 + \beta_3 &= 1 \\ \beta_4^2 &= -5 \end{cases}$$

puede ser evaluada mediante una prueba F y/o χ^2 , pero no mediante una prueba t .

- Respuesta: Verdadero, restricciones no lineales, como $\beta_4^2 = -5$, pueden ser probadas; el requerimiento es que el modelo sea lineal en los parámetros.

- 1.28 El factor inflacionario de la varianza, FIV, de la variable X en el modelo de regresión, estimado mediante Mínimos Cuadrados Ordinarios, $Y_i = \beta_1 + \beta_2 X_i + \mu_i$, debe ser igual a $FIV = 5$ debido a que en el modelo de regresión simple, por definición, no existe multicolinealidad.

- Respuesta Si bien en el modelo de regresión lineal simple, por definición, no existe multicolinealidad, el FIV debe ser igual a 1.

$$FIV = \frac{1}{1 - R_X^2}$$

R_X^2 es el R^2 del modelo auxiliar, donde la variable X se regresa sobre una constante y el resto de las variables explicativas en el modelo (que en el presente caso no existen). Esto es, $X_i = \alpha_1 + \varepsilon_i$, lo cual implicará que $R_X^2 = 0$. Por lo tanto, el $FIV = \frac{1}{1 - R_X^2} = \frac{1}{1 - 0} = 1$. Por lo tanto, la varianza “no se infla/amplifica”.

2. Suponga que una empresa que se dedica a la venta de casas está interesada en estimar la siguiente ecuación vía Mínimos Cuadrados Ordinarios:

$$PRE_i = \beta_1 + \beta_2 MC_i + \mu_i$$

donde PRE corresponde al precio de venta que posee la casa y MC corresponde a sus respectivos metros cuadrados. Sin embargo, la empresa sabe que el precio también depende si la casa está ubicada en un barrio de altos ingresos o uno de bajos ingresos.

Se pide: Reescriba el modelo a estimar de forma que éste sea capaz de capturar el diferencial en el precio de la casa por estar ubicada en un sector de altos ingresos (por diferencias en pendientes e intercepto). En su nuevo modelo, interprete el significado de todos los parámetros (excepto β_1).

■ Solución:

Sea $D_i = 1$ si la casa está ubicada en un barrio de altos ingresos y 0 en otro caso (casa ubicada en un barrio de bajos ingresos). Luego, el modelo que captura las diferencias de acuerdo a la ubicación de la casa es:

$$PRE_i = \beta_1 + \beta_2 MC_i + \beta_3 D_i + \beta_4 (D_i \cdot MC_i) + \varepsilon_i$$

La interpretación de los parámetros es como sigue:

- β_2 es el efecto marginal que ejercen los metros cuadrados de la casa en su precio de venta. Se espera que $\beta_2 > 0$.
- β_3 es el diferencial en el precio de venta de una casa por estar ubicada en un sector de altos ingresos respecto a una ubicada en un sector de bajos ingresos. Se espera que $\beta_3 > 0$.
- β_4 es el diferencial en el efecto marginal que ejercen los metros cuadrados en el precio de venta de una casa ubicada en un sector de altos ingresos. Se espera que $\beta_4 > 0$.

3. De los datos de 46 estados de la Unión Americana para 1992 se obtuvieron los siguientes resultados de regresión:

$$\ln C_i = 4,30 - 1,34 \ln P_i + 0,17 \ln Y_i \quad \bar{R}^2 = 0,27$$

(0,91) (0,32) (0,20)

donde C es el consumo total de cigarros (paquetes por año), P es el precio real por paquete de cigarro, Y es el ingreso disponible real per capita. Los valores en paréntesis corresponden a la desviación estándar del respectivo parámetro.

- 3.1 ¿Cuál es la elasticidad precio de la demanda de cigarros? ¿Es estadísticamente significativa? Si ese es el caso, ¿es estadísticamente distinta de 1 (elasticidad unitaria)? Use un 5 % de significancia. La elasticidad precio estimada de la demanda es igual a $-1,34$.

■ Solución

Para analizar la significancia estadística se puede plantear $H_0 : \beta_2 = 0$ $H_1 : \beta_2 \neq 0$.

$$t_{\text{calculado}} = \frac{-1,34}{0,32} = -4,2$$

$$t_{\text{crítico}} = t_{46-3, 1-0,05/2} = t_{43, 0,975} = 1,96$$

Por lo tanto, se rechaza la hipótesis nula y la elasticidad precio de la demanda es estadísticamente distinta de cero. Esto es, el precio de los cigarros ejerce un efecto significativo en la cantidad demandada de dicho producto.

Para analizar si la elasticidad es unitaria, se puede plantear $H_0 : \beta_2 = -1$ $H_1 : \beta_2 \neq -1$.

$$t_{calculado} = \frac{-1,34 - (-1)}{0,32} = -1,06$$

$$t_{crítico} = t_{46-3,1-0,05/2} = t_{43,0,975} = 1,96$$

Por lo tanto, al 95% de confianza, no se puede rechazar que la elasticidad precio de la demanda de cigarros sea unitaria. Esto es, por ejemplo, si el precio de los cigarros se incrementa en un 1% entonces la cantidad consumida de cigarros disminuirá en un 1%.

3.2 ¿Cuál es la elasticidad ingreso de la demanda de cigarros? ¿Es estadísticamente significativa? Si no es así, ¿cuáles serían las razones?

■ Solución:

La elasticidad ingreso estimada de la demanda es igual a 0,17.

Para analizar la significancia estadística se puede plantear $H_0 : \beta_3 = 0$ $H_1 : \beta_3 \neq 0$.

$$t_{calculado} = \frac{0,17}{0,20} = 0,85$$

$$t_{crítico} = t_{46-3,1-0,05/2} = t_{43,0,975} = 1,96$$

Por lo tanto, no se puede rechazar la hipótesis nula. Esto es, el ingreso no ejerce un efecto significativo en la demanda de cigarros. Desde un punto de vista económico, se puede decir que los cigarros son un bien neutro. Una de las razones principales es la forma en la cual se miden las variables. Mientras que el ingreso está medido en forma per capita, el consumo de cigarros es total. Para relacionar ambas variables, se debería usar la misma unidad de medida.

4. Con el objetivo de medir el efecto que tiene la publicidad en las ventas, una empresa, utilizando datos anuales, estima dicha relación y proporciona la siguiente información:

$$\hat{Y}_t = \underset{(0,15)}{125} + \underset{(0,02)}{0,8} X_t \quad R^2 = 0,89 \quad d = 0,26 \quad t = 1980, 1981, \dots, 1990$$

donde: Y corresponde a la venta de pollos, medida en miles de kilos, X corresponde al gasto en publicidad, medido en millones de pesos y d corresponde al estadístico de Durbin-Watson. El valor entre paréntesis corresponde a la desviación estándar del parámetro estimado.

En base a la regresión, ¿recomendaría usted incrementar el gasto en publicidad con el objetivo de incrementar el nivel de ventas? Use un 5% de significancia. Justifique adecuadamente su respuesta.

■ Solución:

El análisis de significancia individual indica que, con un 95% de confianza, el parámetro β_2 es estadísticamente significativo. Esto es, la publicidad ejerce un efecto significativo y positivo en la venta de pollos. Sin embargo, antes de hacer cualquier recomendación de política se debe evaluar el cumplimiento de los supuestos clásicos y uno de ellos es la ausencia de autocorrelación. En base a la información se puede evaluar la existencia de autocorrelación serial mediante la prueba de Durbin-Watson (d).

H_0 : no existe autocorrelación positiva

H_0^* : no existe autocorrelación negativa

En el presente caso, $N = T = 11$, $K = 2$, $K' = 1$ y $\alpha = 0,05$. Luego, se tienen los siguiente valores críticos:

$$d = 0,26 \sim d_{11,0,05,1} \begin{cases} d_L = 0,927 \\ d_U = 1,324 \end{cases}$$

Así, en base a los datos, y usando un 95% de confianza, se puede rechazar H_0 y los residuos sufren de autocorrelación positiva. Por lo tanto, no se debe recomendar aumentar la publicidad para incrementar las ventas de pollos sin antes corregir la autocorrelación en los residuos de la regresión.

5. Con el objetivo de analizar el salario de los profesores de escuelas públicas, por zona geográfica, (Norte – Centro – Sur) se estimó el siguiente modelo, con datos de 51 profesores (el valor entre paréntesis corresponde a la desviación estándar de cada parámetro):

$$\hat{Y}_i = 26158,62 - 1734,473D_{i2} - 3264,615D_{i3}$$

(1128,523) (1435,953) (1499,615)

donde:

Y_i = Salario (promedio) de los profesores de las escuelas publicas en la zona i (dólares)

$D_2 = 1$ es el profesor residente en la zona norte y 0 en otro caso

$D_3 = 1$ es el profesor residente en la zona sur y 0 en otro caso

5.1 Interprete el coeficiente asociado a la variable D_3 . Haga referencia a su signo y valor.

■ Solución:

$\hat{\beta}_3 = -3264,6$. Este coeficiente corresponde al diferencial de salario de un profesor que trabaja en la zona sur comparado con un profesor que trabaja en la zona centro. Luego, se puede mencionar que un profesor de la zona sur recibe \$3264,6 USD menos que un profesor que trabaja en la zona centro. Esta diferencia significativa al 5% de significancia. $t_{calculado} = \frac{-3264,615}{1499,615} = 2,17 \sim t_{48,0.975} = 1,96$.

5.2 ¿Será recomendable implementar una política pública de forma que el salario de los profesores de la zona norte se iguale al de los profesores de la zona centro? Justifique adecuadamente su respuesta

■ Solución:

No seria recomendable en base a la información entregada, el parámetro que captura la diferencia mencionada es β_2 es cual no es significativo al 5%

$$H_0 : \beta_2 = 0 \quad H_1 : \beta_2 \neq 0 \quad t_{calculado} = \frac{-1.734,473}{1.435,953} = 1,20 \sim t_{48,0.975} = 1,96$$

Por lo tanto, no se puede rechazar la hipótesis nula. Desde un punto de vista estadístico, no existe una diferencia significativa entre el salario de un profesor de la zona norte y uno de la zona centro. La política no es necesaria ni recomendable.

6. Suponga que la siguiente relación estimada representa la relación entre el gasto en investigación y desarrollo, $I\&D$, (medido en US\$) y el nivel de ventas, V , (medido en US\$) de 18 empresas del sector servicios (el valor entre paréntesis corresponde a la desviación estándar del parámetro respectivo):

$$I\&D_i = 192,9931 + 0,0319V_i \quad R^2 = 0,4783$$

(533,9317) (0,0083)

Además, con los residuos de la regresión anterior se proporciona la siguiente estimación:

$$\hat{\mu}_i^2 = -6219665 + 229,3508V_i - 0,000537V_i^2 \quad R^2 = 0,4570$$

¿Estaría usted de acuerdo con una política de estímulo de ventas con el objetivo de aumentar el gasto en investigación y desarrollo? Justifique su respuesta.

■ Solución:

Significancia de β_2 ($H_0 : \beta_2 = 0$, $H_1 : \beta_2 \neq 0$)

$$t_{calculado} = \frac{0,0319}{0,0083} = 3,84 \sim t_{16,0.975} = 2,12$$

por lo tanto se rechaza la hipótesis nula, el parámetro β_2 es estadísticamente significativo al 5%. Esto es, las ventas ejercen un efecto significativo en el gasto en investigación y desarrollo.

Heteroscedasticidad: Con los datos entregados se puede ejecutar la prueba de White.

H_0 : no existe Heteroscedasticidad ($\gamma_2 = \gamma_3 = 0$)

H_1 : existe heteroscedasticidad

$n \cdot R^2 = 18 \cdot 0,4570 = 8,2 \sim \chi^2_{k', 1-\alpha} = \chi^2_{2, 0.95} = 5,99$ por lo tanto, se rechaza la hipótesis nula y los residuos del modelo que explican el costo en investigación y desarrollo sufren de heteroscedasticidad.

Política: no se puede recomendar la aplicación de la política; a pesar que el efecto de las ventas en el gasto en investigación y desarrollo es significativo, la heteroscedasticidad presente en la regresión invalida dicha conclusión. Por lo tanto, antes de recomendar la aplicación de la política, se debe solucionar el problema de la heteroscedasticidad.

7. Suponga que usted debe estimar el siguiente modelo de regresión usando Mínimos Cuadrados Ordinarios: $Y_i = \beta X_i + \mu_i$. Notar que este modelo no posee intercepto. Se pide obtener $\hat{\beta}$ a través del método MCO. Muestre detalladamente el desarrollo.

■ Solución:

$$\min_{\{\hat{\beta}\}} \sum \hat{\mu}_i^2 = \sum (Y_i - \hat{\beta} X_i)^2$$

$$\frac{\partial \sum \hat{\mu}_i^2}{\partial \hat{\beta}} = 0 \rightarrow 2 \sum (Y_i - \hat{\beta} X_i) \cdot -X_i = 0$$

$$\begin{aligned}\sum (Y_i X_i - \hat{\beta} X_i^2) &= 0 \\ \sum Y_i X_i - \hat{\beta} \sum X_i^2 &= 0 \\ \hat{\beta} &= \frac{\sum Y_i X_i}{\sum X_i^2}\end{aligned}$$

8. Se sabe que en el modelo de regresión lineal $Y_i = \beta_1 + \beta_2 X_i + \mu_i$, el estimador de MCO de β_2 está dado por $\hat{\beta}_2 = \frac{\sum_{i=1}^N y_i x_i}{\sum_{i=1}^N x_i^2}$. Demuestre que lo anterior también puede ser expresado como:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^N Y_i x_i}{\sum_{i=1}^N X_i^2 - N\bar{X}^2}. \text{ Muestre detalladamente su desarrollo.}$$

■ Solución:

$$\begin{aligned}\hat{\beta}_2 &= \frac{\sum y_i x_i}{\sum x_i^2} = \frac{\sum (Y_i - \bar{Y}) x_i}{\sum (X_i - \bar{X})^2} = \frac{\sum Y_i x_i - \sum \bar{Y} x_i}{\sum (X_i^2 - 2X_i \bar{X} + \bar{X}^2)} \\ \hat{\beta}_2 &= \frac{\sum Y_i x_i - \bar{Y} \sum x_i}{\sum X_i^2 - 2\bar{X} \sum X_i + N\bar{X}^2} = \frac{\sum Y_i X_i - \bar{Y} \sum x_i}{\sum X_i^2 - 2\bar{X} N\bar{X} + N\bar{X}^2} \\ \hat{\beta}_2 &= \frac{\sum Y_i x_i}{\sum X_i^2 - N\bar{X}^2}\end{aligned}$$

9. Con el objetivo de analizar la relación que existe entre el puntaje obtenido por alumnos en la prueba de matemática para el acceso a la universidad ($psumat$), medido en puntos, y el nivel de ingreso de la familia de donde proviene los alumnos ($ingreso$), medidos en miles pesos, se estimaron, mediante Mínimos Cuadrados Ordinarios, las siguientes regresiones usando información de 10 alumnos:

$$\text{Modelo 1: } \ln(psumat_i) = \beta_1 + \beta_2 \ln(ingreso_i) + \mu_i$$

$$\text{Modelo 2: } psumat_i = \alpha_1 + \alpha_2 ingreso_i + \varepsilon_i$$

$$\text{Modelo 3: } psumat_i = \gamma_1 + \gamma_2 \ln(ingreso_i) + v_i$$

El resultado de los parámetros estimados fue: $\hat{\beta}_2 = 0,12$; $\hat{\alpha}_2 = 0,001$ y $\hat{\gamma}_2 = 61,46$. Se pide interpretar cada uno de los estimadores indicados.

■ Solución:

Modelo 1 [log-log]: $\hat{\beta}_2 = 0,12$, si el ingreso familiar aumenta en un 1%, el puntaje de la prueba aumentará en un 0,12%

Modelo 2 [lin-lin]: $\hat{\alpha}_2 = 0,001$, si el ingreso familiar aumenta en un \$1.000, entonces el puntaje de la prueba aumentará en 0,001 puntos

Modelo 3[lin-log]: $\hat{\gamma}_2 = 61,46$, si el ingreso familiar aumenta en un 1%, entonces el puntaje de la prueba aumentará en $61,46/100 = 0,61$ puntos.

10. Asuma que una importante tienda comercial que se dedica a la venta de un único bien (Z) le pide que estime la función de demanda representativa de dicho bien ($Z_i = \beta_1 + \beta_2 P_i + \mu_i$). Para lo anterior, le proporciona los siguientes datos asociados a las unidades demandadas del bien Z y a el precio de venta P (medido en \$ por unidad) obtenidos durante 50 meses de observación:

$$\sum_{i=1}^{50} (Z_i - \bar{Z})(P_i - \bar{P}) = \sum_{i=1}^{50} z_i p_i = -45 \quad \bar{Z} = 30 \quad \bar{P} = 15$$

$$\sum_{i=1}^{50} (Z_i - \bar{Z})^2 = \sum_{i=1}^{50} z_i^2 = 200 \quad \sum_{i=1}^{50} (P_i - \bar{P})^2 = \sum_{i=1}^{50} p_i^2 = 80$$

se pide:

- 10.1 Estime la función de demanda solicitada por la empresa (determine el valor de β_1 y β_2). Interprete cuidadosamente los parámetros estimados desde un punto de vista matemático y económico. Haga referencia a su signo.

■ Solución:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^{50} z_i p_i}{\sum_{i=1}^{50} p_i^2} = \frac{-45}{80} = -0,5625 \quad \hat{\beta}_1 = \bar{Z} - \hat{\beta}_2 \bar{P} = 30 - (-0,5625) 15 = 38,43$$

Función de demanda estimada: $\hat{Z}_i = 38,43 - 0,5625 P_i$.

Interpretación Matemática:

$\hat{\beta}_1 = 38,43$ es el intercepto de la función de demanda.

$\hat{\beta}_2 = -0,56$ es la pendiente de la función de demanda.

Interpretación Económica:

$\hat{\beta}_1 = 38,43$ es el efecto medio de todas aquellas variables que afectan la demanda del bien Z (distintas al precio P) y que se fueron asumidas constantes durante el análisis.

$\hat{\beta}_2 = -0,56$ es el efecto marginal que ejerce el precio del bien Z en la cantidad demandada del mismo. El signo negativo es esperado debido a que corresponde a una función de demanda. El coeficiente indica que, por ejemplo, al aumentar el precio en \$1, entonces la cantidad demandada disminuirá en 0,56 unidades.

10.2 Calcule el R^2 . Interpretelo.

■ Solución:

$$R^2 = \hat{\beta}_2^2 \frac{Var(P)}{Var(Z)} = (-0,5625)^2 \frac{\frac{80}{50-1}}{\frac{200}{50-1}} = 0,1265 = 12,65\%$$

Aproximadamente, un 12,65% de las variaciones que experimenta la cantidad demandada de Z es explicada por las variaciones que experimenta el precio del bien.

10.3 Calcule la varianza de $\hat{\beta}_2$.

■ Solución:

$$Var(\hat{\beta}_2) = \frac{\hat{\sigma}_\mu^2}{\sum_{i=1}^{50} p_i^2} = \frac{\frac{\sum_{i=1}^{50} \hat{\mu}_i^2}{N-K}}{\sum_{i=1}^{50} p_i^2} = \frac{\frac{174,6875}{50-2}}{80} \approx 0,04549$$

$$\begin{aligned} \sum_{i=1}^{50} \hat{\mu}_i^2 &= \sum_{i=1}^{50} (z_i - \hat{\beta}_2 x_i)^2 = \sum_{i=1}^{50} z_i^2 + \hat{\beta}_2^2 \sum_{i=1}^{50} p_i^2 - 2\hat{\beta}_2 \sum_{i=1}^{50} z_i p_i \\ \sum_{i=1}^{50} \hat{\mu}_i^2 &= 200 + (-0,5625)^2 80 - 2 \cdot (-0,5625) \cdot -45 \\ \sum_{i=1}^{50} \hat{\mu}_i^2 &\approx 174,6875 \end{aligned}$$

10.4 Analice la significancia individual de β_2 . Use un 5 % de significancia.

■ Solución:

$$H_0 : \beta_2 = 0 \quad H_1 : \beta_2 \neq 0$$

$$t_{\text{calculado}} = \frac{\hat{\beta}_2}{\sqrt{\text{Var}(\hat{\beta}_2)}} = \frac{-0,5625}{\sqrt{0,04549}} = -2,64 \sim t_{48,0.975} = 1,96$$

Entonces se puede, al 5% de significancia, rechazar la hipótesis nula. Esto es, el precio del bien Z ejerce un efecto significativo en la cantidad demandada del bien Z .

11. Demuestre que si en el modelo de regresión lineal $Y_i = \beta_1 + \beta_2 X_i + \mu_i$ se cumplen los supuestos clásicos, entonces el estimador de Mínimos Cuadrados Ordinarios $\hat{\beta}_1$ será insesgado y lineal en Y .

■ Solución:

Una de las formas para demostrar que $E(\hat{\beta}_1) = \beta_1$ es:

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X} = \frac{\sum Y_i}{n} - \hat{\beta}_2 \bar{X} \quad , Y_i = \beta_1 + \beta_2 X_i + \mu_i$$

$$\hat{\beta}_1 = \frac{\sum (\beta_1 + \beta_2 X_i + \mu_i)}{n} - \hat{\beta}_2 \bar{X}$$

$$\hat{\beta}_1 = \frac{\sum \beta_1}{n} + \beta_2 \frac{\sum X_i}{n} + \frac{\sum \mu_i}{n} - \hat{\beta}_2 \bar{X}$$

$$\hat{\beta}_1 = \beta_1 + \beta_2 \bar{X} + \frac{\sum \mu_i}{n} - \hat{\beta}_2 \bar{X} \quad /E(\cdot)$$

$$E(\hat{\beta}_1) = E(\beta_1) + E(\beta_2 \bar{X}) + E\left(\frac{\sum \mu_i}{n}\right) - \bar{X} E(\hat{\beta}_2)$$

$$E(\hat{\beta}_1) = \beta_1 + \beta_2 \bar{X} - \beta_2 \bar{X}$$

$$E(\hat{\beta}_1) = \beta_1$$

12. Un comerciante desea estimar la función de demanda del único bien que vende. Para lo anterior, el comerciante proporciona información anual, para el período 1990-2018, de la cantidad demandada del bien (Y) y del precio (X) de dicho bien. La cantidad se mide en unidades y el precio en pesos. En base a los datos se estimó la Regresión 1 (R1) y Regresión 2 (R2) que se presentan, respectivamente, en la Figura 1.1 y Figura 1.2. Se pide:

12.1 En base a R1: (i) escriba la función de demanda estimada, (ii) interprete matemáticamente y económicamente los coeficientes estimados $\hat{\beta}_1$ y $\hat{\beta}_2$ y, finalmente, (iii) interprete el R^2 .

■ Solución:

- Función de demanda estimada: $\hat{Y}_i = 1020,58 - 1,007X_i$
- Interpretación parámetros: $\hat{\beta}_1 = 1020,58$ corresponde al efecto medio de todas aquellas variables, distintas al precio, que ejercen un efecto en la cantidad demandada pero que se asumieron constante al estimar el modelo. Matemáticamente es el intercepto de la función. $\hat{\beta}_2 = -1,007$ es el efecto marginal que ejerce el precio en la cantidad demandada, por ejemplo: si el precio aumenta en \$1, la cantidad demandada disminuirá en aproximadamente 1 unidad, matemáticamente corresponde a la pendiente de la función de demanda.
- Interpretación R^2 : $R^2 = 0,5137 \approx 51,37\%$. Dicho valor indica que, en el periodo 1990-2018, un 51,37% de las variaciones que experimento la cantidad demandada pueden ser explicadas por las variaciones del precio del bien.

12.2 En base a R2 interprete el coeficiente estimado $\hat{\beta}_2$ y analice su significancia individual al 5 %. Plantee claramente la hipótesis nula y alterna e interprete el resultado obtenido.

■ Solución:

- Interpretación $\hat{\beta}_2 = -0,109$ corresponde a la elasticidad-precio de la demanda. Por ejemplo, si el precio aumenta en un 1 %, la cantidad demandada disminuirá en 0,109 %
- Significancia ($\alpha = 0,05$)

$$H_0 : \beta_2 = 0$$

$$H_1 : \beta_2 \neq 0$$

$$t_{calculado} = \frac{\hat{\beta}_2}{\sqrt{Var(\hat{\beta}_2)}} = \frac{-0,1069}{0,0205} = -5,34$$

$$t_{crítico} = t_{N-K, 1-\frac{\alpha}{2}} = t_{29-2, 1-\frac{0,05}{2}} = 2,05$$

usando un 95 % de confianza, se puede rechazar H_0 , (lo cual confirma el p-valor de 0,0 % en la regresión estimada). Dado dicho resultado se puede afirmar que el precio ejerce un efecto significativo (distinto a cero) en la cantidad demandada.

- 12.3 Si la cantidad (Y) es ahora medida en miles de unidades y el precio (X) conserva su unidad de medida, ¿A cuánto sería igual el nuevo $\hat{\beta}_2$ si la base de comparación es R1? Muestre todos los pasos de su resultado.

■ Solución:

Nuevo Modelo: $Y_i^* = \beta_1^* + \beta_2^* X_i^*$, donde $Y_i^* = \frac{1}{1000} Y_i$, $X_i^* = X_i$

$$\hat{\beta}_2^* = \frac{Cov(Y^*, X^*)}{Var(X^*)} = \frac{Cov\left(\frac{1}{1000}Y, X\right)}{Var(X)} = \frac{1}{1000} \cdot \frac{Cov(Y, X)}{Var(X)}$$

$$\hat{\beta}_2^* = \frac{1}{1000} \cdot \hat{\beta}_2 = \frac{1}{1000} \cdot (-1,007)$$

$$\hat{\beta}_2^* = -0,001007$$

- 12.4 En base a R2 construya un intervalo de confianza al 10% de significancia para β_2 . Interprete el resultado.

■ Solución:

$$\alpha = 10\% = 0,10 \rightarrow t_{N-K, 1-\alpha/2} = t_{29-2, 1-0.1/2} = t_{27, 0.95} = 1,70$$

Luego, el intervalo de confianza está dado por:

$$\left[\hat{\beta}_2 - t_{N-K, 1-\alpha/2} \sqrt{Var(\hat{\beta}_2)}, \hat{\beta}_2 + t_{N-K, 1-\alpha/2} \sqrt{Var(\hat{\beta}_2)} \right]$$

$$\left[-0,109 - 1,70 \cdot 0,020 \quad , \quad -0,109 + 1,70 \cdot 0,020 \right]$$

$$[-0,143 \quad , \quad -0,075]$$

Por lo tanto, con un 90% de confianza, se puede indicar que la elasticidad precio de la demanda puede oscilar en un rango inelástico dado por los valores -0,143 y -0,075. Además, como el valor cero no pertenece al intervalo de confianza, la elasticidad precio de la demanda es estadísticamente distinta de cero al 10% de significancia.

- 12.5 Suponga que el comerciante le pide asesoría para incrementar el ingreso por ventas. ¿Qué política de precio le recomendaría usted: aumentar o disminuir el precio? Argumente estadísticamente y económicamente su decisión.

■ Solución:

De acuerdo a R2, la demanda es inelástica, $|\hat{\beta}_2| < 1$, y es significativa. En este caso se recomienda aumentar el precio ya que la cantidad demandada disminuirá en una forma menos que proporcional. Esto tendrá como efecto final un aumento en el ingreso por ventas.

Figura 1.1

$$\text{Regresión R1: } Y_i = \beta_1 + \beta_2 X_i + \mu_i$$

Dependent Variable: Y
 Method: Least Squares
 Date: 09/29/18 Time: 22:02
 Sample: 1990 2018
 Included observations: 29

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1020.587	18.86018	54.11333	0.0000
X	-1.007541	0.188623	-5.341564	0.0000
R-squared	0.513797	Mean dependent var		919.8487
Adjusted R-squared	0.495789	S.D. dependent var		1.367047
S.E. of regression	0.970710	Akaike info criterion		2.844895
Sum squared resid	25.44152	Schwarz criterion		2.939191
Log likelihood	-39.25098	Hannan-Quinn criter.		2.874427
F-statistic	28.53231	Durbin-Watson stat		1.917077
Prob(F-statistic)	0.000012			

Figura 1.2

Regresión R2: $\ln(Y_i) = \beta_1 + \beta_2 \ln(X_i) + \varepsilon_i$

Dependent Variable: LNY
 Method: Least Squares
 Date: 10/01/18 Time: 19:15
 Sample: 1990 2018
 Included observations: 29

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.329154	0.094652	77.43271	0.0000
LN X	-0.109653	0.020554	-5.334782	0.0000
R-squared	0.513162	Mean dependent var	6.824208	
Adjusted R-squared	0.495131	S.D. dependent var	0.001487	
S.E. of regression	0.001056	Akaike info criterion	-10.80140	
Sum squared resid	3.01E-05	Schwarz criterion	-10.70711	
Log likelihood	158.6203	Hannan-Quinn criter.	-10.77187	
F-statistic	28.45990	Durbin-Watson stat	1.915214	
Prob(F-statistic)	0.000012			

13. Demuestre que si en el modelo de regresión lineal $Y_i = \beta_1 + \beta_2 X_i + \mu_i$ se cumplen los supuestos clásicos, entonces:

$$\text{Var}(\hat{\beta}_1) = \sigma_\mu^2 \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n x_i^2} \right]$$

■ Solución:

$$\begin{aligned} E(\hat{\beta}_1) &= E(\bar{Y} - \hat{\beta}_2 \bar{X}) \\ \hat{\beta}_1 &= \frac{\sum_{i=1}^N Y_i}{N} - \hat{\beta}_2 \bar{X} \\ \hat{\beta}_1 &= \frac{1}{N} \left[\sum_{i=1}^N (\beta_1 + \beta_2 X_i + \mu_i) \right] - \hat{\beta}_2 \bar{X} \\ \hat{\beta}_1 &= \frac{1}{N} N \beta_1 + \beta_2 \frac{\sum_{i=1}^N X_i}{N} + \frac{1}{N} \sum_{i=1}^N \mu_i - \hat{\beta}_2 \bar{X} \\ \hat{\beta}_1 &= \beta_1 + \beta_2 \bar{X} + \frac{\sum_{i=1}^N \mu_i}{N} - \hat{\beta}_2 \bar{X} \rightarrow \hat{\beta}_1 - \beta_1 = -\bar{X} (\hat{\beta}_2 - \beta_2) + \frac{\sum_{i=1}^N \mu_i}{N} \end{aligned}$$

Luego, la varianza de del estimador $\hat{\beta}_1$ se puede determinar como:

$$\begin{aligned}
Var(\hat{\beta}_1) &= E\left[\hat{\beta}_1 - E(\hat{\beta}_1)\right]^2 \\
Var(\hat{\beta}_1) &= E\left[\hat{\beta}_1 - \beta_1\right]^2 \\
Var(\hat{\beta}_1) &= E\left[-\bar{X}(\hat{\beta}_2 - \beta_2) + \frac{\sum_{i=1}^N \mu_i}{N}\right]^2 \\
Var(\hat{\beta}_1) &= E\left[\bar{X}^2(\hat{\beta}_2 - \beta_2)^2 + \left(\frac{\sum_{i=1}^N \mu_i}{N}\right)^2 - 2\bar{X}(\hat{\beta}_2 - \beta_2)\frac{\sum_{i=1}^N \mu_i}{N}\right] \\
Var(\hat{\beta}_1) &= \underbrace{\bar{X}^2 E(\hat{\beta}_2 - \beta_2)^2}_{Var(\hat{\beta}_2)} + \frac{1}{N^2} E\left(\sum_{i=1}^N \mu_i\right)^2 - 2\bar{X}(\hat{\beta}_2 - \beta_2)\frac{\sum_{i=1}^N E(\mu_i)}{N} \\
Var(\hat{\beta}_1) &= \bar{X}^2 \sigma_\mu^2 \frac{1}{\sum_{i=1}^N x_i^2} + \frac{1}{N^2} N \sigma_\mu^2 \\
Var(\hat{\beta}_1) &= \sigma_\mu^2 \left[\frac{\bar{X}^2}{\sum_{i=1}^N x_i^2} + \frac{1}{N} \right]
\end{aligned}$$

14. Con el objetivo de analizar la relación que existe entre el puntaje de la prueba de aptitud (ACT) para el acceso a las universidades en Estados Unidos y el promedio de notas en la universidad (GPA), una universidad propone estimar la regresión $GPA_i = \beta_1 + \beta_2 ACT_i + \mu_i$ mediante el método de Mínimos Cuadrados Ordinarios. Para hacer lo anterior, la universidad usa información de 8 alumnos y obtiene los siguientes resultados:

$$\overline{ACT} = 25,875 \quad \overline{GPA} = 2,673 \quad \sum_{i=1}^8 (ACT_i - \overline{ACT})^2 = 56,875$$

$$\sum_{i=1}^8 (GAP_i - \overline{GPA})^2 = 6,673$$

$$\sum_{i=1}^8 (ACT_i - \overline{ACT})(GAP_i - \overline{GPA}) = -4,3507 \quad Var(GAP) = 0,612$$

$$Var(ACT) = 8,125$$

$$\sum_{i=1}^8 (ACT_i - \widehat{ACT}_i)^2 = 1,015 \quad \sum_{i=1}^8 (GAP_i - \widehat{GAP}_i)^2 = 4,0156$$

En base a los resultados se pide:

- 14.1 Determinar el valor de $\hat{\beta}_1$ y $\hat{\beta}_2$. Interpretar el valor obtenido de $\hat{\beta}_2$.

■ Solución

$$\hat{\beta}_2 = \frac{\sum_{i=1}^8 (ACT_i - \overline{ACT})(GPA_i - \overline{GPA})}{\sum_{i=1}^8 (ACT_i - \overline{ACT})^2} = \frac{-4,3507}{56,875} = -0,0764$$

$$\hat{\beta}_1 = \overline{GPA} - \hat{\beta}_2 \overline{ACT} = 2,673 - (-0,0764) 25,875$$

$$\hat{\beta}_1 = 4,6498$$

$\hat{\beta}_2$: corresponde al efecto marginal que ejerce el puntaje de la prueba de ingreso en el promedio de notas de la universidad. Como el signo es negativo, existe una relación negativa entre el puntaje y el promedio de notas en la universidad.

14.2 ¿En cuánto cambiará el GPA si ACT aumenta en 5 puntos? Indique claramente si el cambio corresponde a un aumento o una disminución.

■ Solución:

Del resultado se tiene que $\frac{\partial GPA}{\partial ACT} = -0,0764$. Por lo tanto, si el puntaje aumenta en 5 puntos ($\partial ACT = 5$), el cambio que experimentará el promedio de notas en la universidad será:

$$\partial GPA = -0,0764 \cdot 5 = -0,382$$

luego, el promedio de notas en la universidad disminuirá en 0,38 puntos si el puntaje en la prueba de ingreso aumenta en 5 puntos.

14.3 Determine las varianzas de $\hat{\beta}_1$ y $\hat{\beta}_2$.

■ Solución:

$$\bullet \hat{\sigma}_\mu^2 = \frac{\sum_{i=1}^8 \hat{\mu}_i^2}{N-K} = \frac{\sum_{i=1}^8 (GPA_i - \widehat{GPA}_i)^2}{N-K} = \frac{4,0156}{8-2} = 0,6692.$$

$$\bullet Var(\hat{\beta}_2) = \frac{\hat{\sigma}_\mu^2}{\sum_{i=1}^8 (ACT_i - \overline{ACT})^2} = \frac{0,6692}{56,875} = 0,0117.$$

$$\bullet Var(\hat{\beta}_1) = \hat{\sigma}_\mu^2 \left[\frac{1}{N} + \overline{ACT}^2 \frac{1}{\sum_{i=1}^8 (ACT_i - \overline{ACT})^2} \right] = 0,6692 \left[\frac{1}{8} + (25,875)^2 \frac{1}{56,875} \right] = 7,9612.$$

15. Con el objetivo de explicar el promedio de notas en la universidad, se llevó a cabo un estudio donde se recopiló información de 141 estudiantes respecto a las siguientes variables:

- COLGPA: promedio de notas en la universidad, medido en puntos.
- PC: variable que indica si el alumno tenía computador en la universidad (su valor es 1 si tenía computador y 0 si no lo tenía)
- HSGPA: promedio de notas de la enseñanza media, medido en puntos.
- ACT: puntaje obtenido en la prueba de ingreso a la universidad, medido en puntos.
- SKIPPED: número de clases por semana a las cuales el alumno decidió faltar, medido en número de clases.

En base a los datos se estimaron tres modelos:

$$\text{Modelo 1: } COLGPA_i = \beta_1 + \beta_2 HSGPA_i + \beta ACT_i + \mu_i$$

$$\text{Modelo 2: } COLGPA_i = \beta_1 + \beta_2 HSGPA_i + \beta ACT_i + \beta_4 PC_i + \nu_i$$

$$\text{Modelo 3: } COLGPA_i = \beta_1 + \beta_2 HSGPA_i + \beta ACT_i + \beta_4 PC_i + \beta_5 SKIPPED_i + \kappa_i$$

donde μ_i , ν_i y κ_i son los errores de los modelos.

El resultado de la estimación del Modelo 1, Modelo 2 y Modelo 3, se presentan, respectivamente, en la Figura 1.3, Figura 1.4 y Figura 1.5.

Figura 1.3

Modelo 1

Dependent Variable: COLGPA

Method: Least Squares

Sample: 1 141

Included observations: 141

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.286328	0.340822	3.774191	0.0002
HSGPA	0.453456	0.095813	4.732721	0.0000
ACT	0.009426	0.010777	0.874627	0.3833
R-squared	0.176422	Mean dependent var	3.056738	
Adjusted R-squared	0.164486	S.D. dependent var	0.372310	
S.E. of regression	0.340316	Akaike info criterion	0.703161	
Sum squared resid	15.98244	Schwarz criterion	0.765901	
Log likelihood	-46.57287	Hannan-Quinn criter.	0.728656	
F-statistic	14.78073	Durbin-Watson stat	1.885351	
Prob(F-statistic)	0.000002			

Figura 1.4

Modelo 2

Dependent Variable: COLGPA

Method: Least Squares

Sample: 1 141

Included observations: 141

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.263520	0.333126	3.792925	0.0002
HSGPA	0.447242	0.093647	4.775801	0.0000
ACT	0.008659	0.010534	0.821987	0.4125
PC	0.157309	0.057287	2.745961	0.0068
R-squared	0.219386	Mean dependent var	3.056738	
Adjusted R-squared	0.202292	S.D. dependent var	0.372310	
S.E. of regression	0.332527	Akaike info criterion	0.663768	
Sum squared resid	15.14868	Schwarz criterion	0.747421	
Log likelihood	-42.79566	Hannan-Quinn criter.	0.697762	
F-statistic	12.83426	Durbin-Watson stat	1.869922	
Prob(F-statistic)	0.000000			

Figura 1.5

Modelo 3

Dependent Variable: COLGPA

Method: Least Squares

Sample: 1 141

Included observations: 141

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.356509	0.327502	4.141986	0.0001
HSGPA	0.412952	0.092431	4.467678	0.0000
ACT	0.013344	0.010444	1.277706	0.2035
PC	0.124439	0.057312	2.171276	0.0316
SKIPPED	-0.071034	0.026249	-2.706104	0.0077
R-squared	0.259271	Mean dependent var	3.056738	
Adjusted R-squared	0.237484	S.D. dependent var	0.372310	
S.E. of regression	0.325109	Akaike info criterion	0.625507	
Sum squared resid	14.37467	Schwarz criterion	0.730073	
Log likelihood	-39.09822	Hannan-Quinn criter.	0.667999	
F-statistic	11.90070	Durbin-Watson stat	1.875252	
Prob(F-statistic)	0.000000			

15.1 Usando el Modelo 3, interprete cada uno de los coeficientes estimados haciendo referencia a su signo y magnitud.

■ Solución:

* $\hat{\beta}_2 = 0,41$. Si el promedio de notas de la enseñanza media aumenta en 1 punto, entonces el promedio de notas en la universidad aumentará en 0,41 puntos. El efecto es positivo y es el esperado.

* $\hat{\beta}_3 = 0,01$. Si el puntaje de la prueba de acceso aumenta en 1 punto, entonces se espera que el promedio de notas en la universidad aumente en 0,01 puntos. El efecto es positivo y es el esperado.

* $\hat{\beta}_4 = 0,12$. Si el alumno tiene computador, entonces se espera que el promedio de notas en la universidad aumente en 0,12 puntos. El efecto es positivo y esperado.

* $\hat{\beta}_5 = -0,07$. Si el alumno decide faltar a una clase en la semana, se espera que su promedio de notas disminuya en 0,07 puntos. El efecto es negativo y esperado.

15.2 Al comparar el Modelo 1 con el Modelo 2, ¿Es conveniente incorporar la variable PC? Justifique estadísticamente su respuesta.

■ Solución:

Sí es conveniente. Del resultado del Modelo 2 se puede ver que el p-value de PC es igual a 0,68% por lo cual la variable PC es estadísticamente significativa al 1%. Esto es, poseer un computador ayuda a explicar el promedio de notas en la universidad.

15.3 Analice la significancia global del Modelo 3. Plantee claramente la hipótesis nula a evaluar. Interprete el resultado obtenido.

■ Solución:

$$H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

$$H_1 : \text{a lo menos un } \beta_i \neq 0$$

Del resultado del Modelo 3 se tiene que $F_{calculado} = 11,90$ y posee un p-value de un 0%. Por lo tanto se rechaza la hipótesis nula y el modelo es globalmente significativo.

15.4 Suponga que la universidad le solicita a usted evaluar la conveniencia de entregar un computador a todos los alumnos e implementar, simultáneamente, una política de asistencia obligatoria (que busca reducir el número de clases a la cuales el alumno decide faltar) para mejorar el promedio de notas. Evalúe estadísticamente la política, plantee la hipótesis a evaluar e interprete el resultado obtenido. ¿Apoyaría usted la política simultánea que la universidad quiere implementar?

■ Respuesta:

Para evaluar si la política posee un efecto significativo se debe evaluar la significancia conjunta de β_4 y β_5 (el número de restricciones lineales es $q = 2$). Esto es:

$$H_0 : \begin{cases} \beta_4 = 0 \\ \beta_5 = 0 \end{cases}$$

$$F_{calculado} = \frac{\frac{SCR_R - SCR}{q}}{\frac{SCR}{N-K}} = \frac{\frac{15,98244 - 14,37467}{2}}{\frac{14,37467}{141-5}} = 7,60$$

$$F_{[2,\infty]0,95} = 3,0$$

Luego, como $F_{calculado} = 7,60 > F_{crítico} = 3,0$, se puede rechazar la hipótesis nula planteada al 95% de confianza. Lo anterior implica que la política propuesta para la universidad será efectiva. Esto es, a través de simultáneamente implementar la asistencia obligatoria y entregar un computador a cada alumno, se observará un aumento en el promedio de notas de la universidad.

16. Demuestre que el $F_{calculado}$ de la prueba de significancia global puede ser escrito en términos del R^2 del modelo. Esto es, se pide demostrar que:

$$F_{calculado} = \frac{\frac{SCE}{K-1}}{\frac{SCR}{N-K}} = \frac{\frac{R^2}{K-1}}{\frac{1-R^2}{N-K}}$$

Para estimar la relación ingreso-educación se encuesta a 680 personas en la región metropolitana y se estima la siguiente ecuación:

$$\hat{Y}_i = 18361 + 2125X_i$$

$$R^2 = 0.9$$

$$SCR = 3870$$

donde Y es el nivel de ingreso del individuo (\$) y X son los años de educación. Además, con el objetivo de evaluar un posible cambio estructural, la regresión fue estimada por sectores económicos. Los datos asociados a las regresiones de cada sector económico son:

Sector Económico	N	SCR
Construcción	180	675
Servicios	300	970
Agrícola	200	810

En base a los datos entregados, y usando la prueba F de Chow, pruebe si existen diferencias estructurales entre los sectores económicos antes descritos. Plantee claramente la hipótesis nula a evaluar. Use un 5% de significancia.

■ Solución:

Paso 1: $SCR_R = 3870$

$$\text{Paso 2: } SCR = SCR^{cons} + SCR^{serv} + SCR^{agri} = 675 + 970 + 810 = 2455$$

$$\text{Paso 3: } H_0 : \beta^{cons} = \beta^{serv} = \beta^{agri} \text{ (estabilidad estructural; } Z = 3 \text{ grupos)}$$

$$F_{calculado} = \frac{\frac{SCR_R - SCR}{K(Z-1)}}{\frac{SCR}{N-ZK}} = \frac{\frac{3870 - 2455}{2(3-1)}}{\frac{2455}{680-3 \cdot 2}} = 97,11$$

$$F_{calculado} \sim F_{(2(3-1), 680-3 \cdot 2)0.95} = 2,37$$

Por lo tanto, al 95% de confianza se puede rechazar la hipótesis nula de estabilidad estructural. Luego, la relación ingreso y años de educación es distinta para los grupos analizados. Se recomienda estimar la regresión base para cada sector analizado ya que existirán diferencia en interceptos y/o pendientes.

17. Suponga que con el objetivo de medir la reacción que experimenta el consumo del pollo, un supermercado, con datos anuales que abarcan el período 1960 – 1982, estima la siguiente ecuación:

$$\ln(CP_t) = \underset{\hat{\beta}_1}{2,51} - \underset{\hat{\beta}_2}{0,78} \ln(PP_t) + \underset{\hat{\beta}_3}{0,47} \ln(PC_t) + \underset{\hat{\beta}_4}{0,43} \ln(PV_t) + \hat{\mu}_t$$

donde:

- CP_t = Consumo de kilos de pollo en el año t
- PP_t = Precio promedio del kilos de pollo en el año t, medido en pesos
- PC_t = Precio promedio del kilos de cerdo en el año t, medido en pesos
- PV_t = Precio promedio del kilos de vacuno en el año t, medido en pesos

Al estimar la regresión, vía MCO, se obtuvieron los siguientes datos:

$$Var(\hat{\beta}) = \begin{bmatrix} 0,003 & -0,017 & 0,007 & 0,0008 \\ & 0,013 & -0,005 & -0,002 \\ & & 0,006 & -0,003 \\ & & & 0,005 \end{bmatrix} \quad R^2 = 0,965683$$

- 17.1 Analice la significancia individual del parámetro β_2 (utilice un 90 % de confianza), plante claramente la hipótesis a evaluar e interprete su resultado. Además, analice si es recomendable una política de aumento o disminución de precio con el objetivo de aumentar el ingreso total del supermercado por concepto de venta de pollo.

- Solución:

$$H_0 : \beta_2 = 0$$

$$H_1 : \beta_2 \neq 0$$

En el ejercicio, $\alpha = 0,1$, $N = T = 23$ y $K = 4$. Luego:

$$t_{calculado} = \frac{\hat{\beta}_2 - 0}{\sqrt{Var(\hat{\beta}_2)}} = \frac{-0,78}{\sqrt{0,013}} = -6,84$$

$$t_{critico} = t_{1-\alpha/2, N-K} = t_{0.95, 19} = 1,73$$

Luego, dado que $|t_{calculado}| > t_{critico}$, se puede rechazar la hipótesis nula al 90% de confianza. Esto es, el precio del pollo tiene un efecto significativo en su cantidad demandada. En este caso, como la demanda es inelástica, una políticas para aumentar el ingreso total es aumentar el precio del kilo de pollo.

17.2 Analice la siguiente hipótesis e interprete su resultado (utilice un 95 % de confianza):

$$H_0 : \beta_4 = 1$$

$$H_1 : \beta_4 \neq 1$$

■ Solución:

En este caso, se tiene $\alpha = 0,05$, $N = T = 23$ y $K = 4$. Luego:

$$t_{calculado} = \frac{\hat{\beta}_4 - 1}{\sqrt{Var(\hat{\beta}_4)}} = \frac{0,43 - 1}{\sqrt{0,005}} = -8,06$$

$$t_{critico} = t_{1-\alpha/2, N-K} = t_{0.975, 19} = 2,09$$

Luego, dado que $|t_{calculado}| > t_{critico}$, se puede rechazar la hipótesis nula al 95% de confianza. Esto es, la elasticidad precio cruzada de la demanda de pollo con el precio del vacuno es distinta a 1. Luego, si el precio del vacuno aumenta, por ejemplo, en un 1%, la demanda de pollo aumentará, pero en menos que un 1%.

17.3 Analice la significancia Global del modelo (utilice un 5 % de significancia), plantee claramente la hipótesis a evaluar e interprete el resultado.

■ Respuesta

En este caso, se tiene $\alpha = 0,05$, $N = T = 23$ y $K = 4$. Luego:

$$H_0 : \beta_2 = \beta_3 = \beta_4 = 0$$

$$H_1 : \text{a lo menos un } \beta_i \neq 0$$

$$F_{\text{calculado}} = \frac{R^2}{1 - R^2} \frac{N - K}{K - 1} = \frac{0,965683}{1 - 0,965683} \frac{23 - 4}{4 - 1}$$

$$F_{\text{calculado}} = 178,22$$

$$F_{\text{crítico}} = F_{(K-1, N-K)1-\alpha} = F_{(3,19)0,95} = 3,10$$

Luego, como $F_{\text{calculado}} > F_{\text{crítico}}$, se puede rechazar la hipótesis nula al 95% de confianza. Por lo tanto, el modelo de demanda de pollo es globalmente significativo.

18. Suponga que con el objetivo analizar los factores que determinan el tiempo que un individuo se demora en llegar en auto desde su casa a su trabajo, se estima el modelo $time_i = \beta_1 + \beta_2 depart_i + \beta_3 reds_i + \beta_4 trains_i + \mu_i$, donde $time$ es el tiempo (minutos) que demora un individuo en llegar a su trabajo, $depart$ es la hora de salida desde la casa (minutos a partir de las 6 am), $reds$ es el número de semáforos en rojo de camino al trabajo y $trains$ es el número de cruce de trenes donde el individuo debió detenerse. El resultado de la regresión, estimado a través de Mínimo Cuadrados Ordinarios, se presenta en la Figura 1.6:

Figura 1.6

Regresión Tiempo en Auto

Dependent Variable: TIME

Method: Least Squares

Included observations: 249

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	20.87014	1.675760	(a)	0.0000
DEPART	0.368125	0.035102	10.48733	0.0000
REDS	(b)	0.185038	8.225030	0.0000
TRAINS	3.023696	0.634012	4.769148	0.0000
R-squared	0.534561	Mean dependent var	42.95582	
Adjusted R-squared	(c)	S.D. dependent var	9.176257	
S.E. of regression	(d)	Akaike info criterion	6.534446	
Sum squared resid	9719.536	Schwarz criterion	6.590951	
Log likelihood	-809.5385	Hannan-Quinn criter.	6.557190	
F-statistic	(e)	Durbin-Watson stat	1.652439	
Prob(F-statistic)	0.000000			

Se pide: Obtener los valores faltantes indicados por los rectángulos (a)-(e). Muestre todos sus cálculos.

■ Solución:

$$(a) \ t = \frac{\hat{\beta}_1}{\sqrt{Var(\hat{\beta}_1)}} = \frac{20,87014}{1,675760} = 12,45.$$

$$(b) \ \hat{\beta}_3 = t \cdot \sqrt{Var(\hat{\beta}_3)} = 8,225030 \cdot 0,185038 = 1,52.$$

$$(c) \ \bar{R}^2 = 1 - (1 - R^2) \left(\frac{n-1}{n-k} \right) = 1 - (1 - 0,5345661) \left(\frac{249-1}{249-4} \right) = 0,5288.$$

$$(d) \ \sqrt{\hat{\sigma}_\mu^2} = \sqrt{\frac{SCR}{N-K}} = \sqrt{\frac{9719,536}{249-4}} = 6,298.$$

Capítulo 2

Ejercicios Propuestos

1. Con el objetivo de medir el efecto que tiene el número de cigarros (*cigarro*) que fumó al día una madre embarazada sobre el peso del recién nacido (*peso*), se obtuvieron los siguientes datos.

Peso recién nacido (kilos) pesos	Cigarros (número de cigarros por día)
3,0	20
3,5	10
4,0	0
3,9	2

En base a los datos, estime el siguiente modelo de regresión usando MCO: $\text{peso}_i = \beta_1 + \beta_2 \text{cigarro}_i + \mu_i$ e interprete, posteriormente, $\hat{\beta}_1$ y $\hat{\beta}_2$.

2. Basado en una muestra de 10 observaciones, se obtuvieron los siguientes resultados para estimar el modelo de regresión

$$Y_i = \beta_1 + \beta_2 X_i + \mu_i :$$
$$\sum_{i=1}^{10} Y_i = 1.110 \quad \sum_{i=1}^{10} X_i = 1.700 \quad \sum_{i=1}^{10} X_i Y_i = 205.500$$
$$\sum_{i=1}^{10} X_i^2 = 322.000 \quad \sum_{i=1}^{10} Y_i^2 = 132.100$$

donde el coeficiente de correlación es $r = 0,9758$. Al verificar por segunda vez los cálculos, se encontró que se habían registrado erróneamente dos pares de observaciones

Y	X	en lugar de	Y	X
90	120		80	110
140	220		150	210

¿Cuál será el efecto de este error en r ? Obtenga el coeficiente de correlación correcto.

3. Considérese los siguientes resultados de regresión $Y_t = \beta_1 + \beta_2 X_t + \mu_t$ para la economía de los Estados Unidos durante el período 1968-1987, donde Y es el gasto de Estados Unidos en bienes importados y X es el ingreso personal disponible, ambos medidos en miles de millones de dólares de 1982:

	$\hat{\beta}_1$	$\hat{\beta}_2 X_t$	R^2
$\hat{Y}_t =$	-261,09	0,2453 X_t	0,9388
$DS =$	31,327	()	
$t_{calculado} =$	()	16,616	

- 3.1 Complete la información faltante. Esto es, la información que debe estar en entre paréntesis.
- 3.2 Interprete el parámetro estimado $\hat{\beta}_2$, interprete el R^2 y analice la significancia individual de β_2 usando un 5 % de significancia.
4. Demuestre que en el modelo de regresión simple se verifica que:

$$\underbrace{\sum_{i=1}^n \hat{\mu}_i^2}_{SCR} = SCT (1 - R^2)$$

5. Una economía posee información histórica de su producto interno bruto real , empleo y capital real para los años 1955-1974. En base a la información se estimó la siguiente función de producción Cobb-Douglas: $\log(Y_t) = \beta_1 + \beta_2 \log(X_{2t}) + \beta_3 \log(X_{3t}) + \mu_t$, donde $\log(\cdot)$ corresponde al logaritmo natural de la variable. Los resultados de la regresión se presentan en la Figura 2.1 y, además, se presentan los resultados de la prueba Breusch-Godfrey en la Figura 2.2.

Figura 2.1

Regresión Cobb-Douglas

Dependent Variable: LOG(Y)

Method: Least Squares

Sample: 1955 1974

Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.652419	0.606198	-2.725873	0.0144
LOG(X2)	0.339732	0.185692	1.829548	0.0849
LOG(X3)	0.845997	0.093352	9.062488	0.0000
R-squared	0.995080	Mean dependent var	12.22605	
Adjusted R-squared	0.994501	S.D. dependent var	0.381497	
S.E. of regression	0.028289	Akaike info criterion	-4.155221	
Sum squared resid	0.013604	Schwarz criterion	-4.005861	
Log likelihood	44.55221	Hannan-Quinn criter.	-4.126064	
F-statistic	1719.231	Durbin-Watson stat	0.425667	
Prob(F-statistic)	0.000000			

Figura 2.2

Prueba de Breusch-Godfrey

Breusch-Godfrey Serial Correlation LM Test:

Null hypothesis: No serial correlation at up to 2 lags

F-statistic	13.96025	Prob. F(2,15)	0.0004
Obs*R-squared	13.01033	Prob. Chi-Square(2)	0.0015

- 5.1 Interprete el valor de los parámetros estimados $\hat{\beta}_2$ y $\hat{\beta}_3$. Además, usando 95 % de confianza, evalúe la significancia individual de β_2 y la significancia global del modelo. En cada caso, plantee claramente la hipótesis nula y alterna
- 5.2 Con el objetivo de evaluar la presencia de autocorrelación en el modelo de la Figura 1, se aplicó la prueba de Breush-Godfrey y se usó $p = 2$. El resultado de la prueba se muestra en la Figura 2. Al respecto se pide: (i) escriba la regresión auxiliar usada en la prueba de Breush-Godfrey. (ii) Determine si existe autocorrelación usando el resultado en la Figura 2. Además, mencione dos formas para solucionar la autocorrelación.
- 5.3 Con el objetivo de probar si existen rendimientos constantes a escalas y que el parámetro $\beta_2 = 0$, se estimó un modelo restringido en base a la siguiente hipótesis nula:

$$H_0 \begin{cases} \beta_2 + \beta_3 &= 1 \\ \beta_2 &= 0,3 \end{cases}$$

Sin embargo, al momento de estimar el modelo, el programa calculó los siguientes tres modelos:

Modelo 1	Modelo 2
$Y_t^* = \alpha_1 + \alpha_2 X_t^* + v_t$	$Y_t^* = \gamma_1 + \gamma_2 X_t^* + \vartheta_t$
SCR=1,4944	SCR=1,8666
donde	donde
$Y_t^* = \log(Y_t) - 0,7 \log(X_{2t})$	$Y_t^* = \log(Y_t) - 0,3 \log(X_{2t})$
$X_t^* = 0,3 [\log(X_{2t}) - \log(X_{3t})]$	$X_t^* = \log(X_{3t}) - \log(X_{2t})$

Modelo 3
$Y_t^* = \delta_1 + v_t$
SCR=2,1889
donde
$Y_t^* = \log(Y_t) - 0,3 \log(X_{2t}) - 0,7 \log(X_{3t})$

Determine y derive algebraicamente el modelo correcto para evaluar H_0 . Una vez derivado el modelo, use un 5% de significancia para evaluar la hipótesis.

6. Asuma que un país, que debe importar todo el petróleo que consume, está interesado en llevar a cabo una política con el objetivo de disminuir el consumo de petróleo destinado a calefacción. Para hacer lo anterior, el gobierno dispone de datos de consumo anual de petróleo en miles de bidones (Y), tasa porcentual de impuesto aplicada a la importación de petróleo X_2 y aislamiento exigido a las viviendas en pulgadas X_3 . Con datos para 7 años, se procede a estimar la siguiente ecuación:

$$\ln Y_t = \beta_1 + \beta_2 \ln X_{2t} + \beta_3 \ln X_{3t} + \mu_t$$

Los resultados de la regresión, usando Mínimos Cuadrados Ordinarios, son los siguientes:

$$\hat{\beta} = \begin{pmatrix} 6,474 \\ -0,068 \\ -0,586 \end{pmatrix} \quad \hat{\sigma}_\mu^2 (\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 0,4644 & -0,1561 & -0,0556 \\ & 0,0613 & 0,0073 \\ & & 0,00221 \end{pmatrix} \quad R^2 = 0,7978$$

En base a dicha información, se pide:

- 6.1 Escribir la función estimada e interpretar, haciendo referencia a su signo, el valor de los parámetros β_2 y β_3 .

- 6.2 Evalúe la significancia individual del parámetro β_3 . Plantee claramente la hipótesis, utilice un 95 % de confianza e interprete su resultado.
- 6.3 Analice la significancia global del modelo. Plantee claramente la hipótesis a evaluar, utilice un 99 % de confianza e interprete su resultado.
- 6.4 Si el gobierno decide, en base a los resultados, implementar una política de reducción de consumo de petróleo vía aumento de la tasa de impuesto a la importación de petróleo, ¿estará usted de acuerdo en la implementación de dicha política? Recuerde que el aumento de la tasa impositiva genera mayores ingresos a las Arcas Fiscales.
- 6.5 Dada la forma del modelo, es muy probable que exista heteroscedasticidad. Comente y argumente su respuesta.
7. Con el objetivo de estimar el efecto que tiene la educación (X), medida en años de escolaridad, en los ingresos promedios (Y), medidos en pesos, se procedió a tomar una muestra. Los resultados, obtenidos a través de Mínimos Cuadrados Ordinarios, son los siguientes (el valor entre paréntesis corresponde a la desviación estándar del parámetro respectivo):

$$\hat{Y}_i = 5520 + 1520X_i$$

$$R^2 = 0,8944 \quad \hat{\sigma}_\mu^2 = 0,8040$$

En base a la regresión, se pide:

- 7.1 Interpretar los valores estimados de β_1 y β_2 . Hacer referencia a sus signos y magnitudes.
- 7.2 Interpretar el R^2 del modelo.
8. Demuestre que si en la función de regresión poblacional $Y_i = \beta_1 + \beta_2 X_i + \mu_i$ el término de error es homoscedástico y posee un valor esperado igual a cero, $\mu_i \sim (0, \sigma_\mu^2)$, entonces el supuesto de homoscedasticidad de la perturbación estocástica, $Var(\mu_i|X_i) = \sigma_\mu^2$, es equivalente al supuesto que $Var(Y_i|X_i) = \sigma_\mu^2$.
9. Para estudiar el consumo de pollo per cápita en Estados Unidos se estima la siguiente regresión usando información anual entre 1960-1982:

$$\ln(Y_t) = \beta_1 + \beta_2 \ln(X_{2t}) + \beta_3 \ln(X_{3t}) + \beta_4 \ln(X_{4t}) + \beta_5 \ln(X_{5t}) + \mu_t$$

donde Y es el consumo de pollo per cápita, X_2 es el ingreso real disponible per cápita, X_3 es el precio real al por mayor del pollo por libra (\$), X_4 es precio real al por mayor de la carne de cerdo por libra (\$) y X_5 es el precio real de la carne de vacuno por libra (\$). El resultado de la regresión se puede encontrar en la Figura 2.3. En base a la regresión, se pide (use un 5 % de significancia en todo el análisis):

- 9.1 Analizar la significancia individual de β_4 (plantee claramente la hipótesis a evaluar) e interprete el coeficiente estimado $\hat{\beta}_4$.
- 9.2 Analice la significancia global del modelo (plantee claramente la hipótesis a evaluar) e interprete el valor del R^2 .
- 9.3 Suponga que se quiere evaluar la hipótesis que la carne de pollo, carne de cerdo y carne de vacuno son bienes no relacionados. Plantee claramente, y evalúe, la hipótesis nula. Use los resultados proporcionados en Figura 2.4 y/o Figura 2.5.

Figura 2.3

Regresión Pollo per capita

Dependent Variable: LOG(Y)

Method: Least Squares

Sample: 1960 1982

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.189792	0.155715	14.06283	0.0000
LOG(X2)	0.342555	0.083266	4.113970	0.0007
LOG(X3)	-0.504592	0.110894	-4.550212	0.0002
LOG(X4)	0.148545	0.099673	1.490334	0.1535
LOG(X5)	0.091105	0.100716	0.904568	0.3776
R-squared	0.982313	Mean dependent var	3.663887	
Adjusted R-squared	0.978383	S.D. dependent var	0.187659	
S.E. of regression	0.027591	Akaike info criterion	-4.152987	
Sum squared resid	0.013703	Schwarz criterion	-3.906140	
Log likelihood	52.75935	Hannan-Quinn criter.	-4.090906	
F-statistic	249.9282	Durbin-Watson stat	1.826069	
Prob(F-statistic)	0.000000			

Figura 2.4

Regresión Pollo Restringida 1

Dependent Variable: LOG(Y)

Method: Least Squares

Sample: 1960 1982

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.243883	0.051338	43.70846	0.0000
LOG(X2)-LOG(X3)	0.488186	0.016657	29.30812	0.0000
LOG(X4)-LOG(X5)	0.042926	0.055166	0.778119	0.4456
R-squared	0.977269	Mean dependent var		3.663887
Adjusted R-squared	0.974996	S.D. dependent var		0.187659
S.E. of regression	0.029674	Akaike info criterion		-4.075997
Sum squared resid	0.017611	Schwarz criterion		-3.927889
Log likelihood	49.87396	Hannan-Quinn criter.		-4.038748
F-statistic	429.9333	Durbin-Watson stat		2.122500
Prob(F-statistic)	0.000000			

Figura 2.5

Regresión Pollo Restringida 2

Dependent Variable: LOG(Y)

Method: Least Squares

Sample: 1960 1982

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.032820	0.116183	17.49673	0.0000
LOG(X2)	0.451528	0.024695	18.28435	0.0000
LOG(X3)	-0.372212	0.063466	-5.864740	0.0000
R-squared	0.980074	Mean dependent var		3.663887
Adjusted R-squared	0.978082	S.D. dependent var		0.187659
S.E. of regression	0.027783	Akaike info criterion		-4.207711
Sum squared resid	0.015437	Schwarz criterion		-4.059603
Log likelihood	51.38868	Hannan-Quinn criter.		-4.170462
F-statistic	491.8681	Durbin-Watson stat		1.875601
Prob(F-statistic)	0.000000			

10. Suponga que una empresa está interesada en analizar si la evaluación del servicio al cliente influye en su ingreso por ventas. Para hacer lo anterior cuenta con información del ingreso por ventas promedio por departamento (\$) (Y), el precio promedio de ventas por departamento (\$), (X) y la calidad de atención del servicio al cliente. Respecto de esta última variable, los conceptos usados son Calidad Muy Buena, Calidad Buena y Calidad Mala. Se pide: plantee la función de regresión para estimar el ingreso por ventas promedio por departamento y que permita analizar el efecto de la evaluación del servicio al cliente. Una vez planteada señale cómo, a partir de dicha regresión, se puede medir el efecto de la atención al cliente en el ingreso por ventas.

Agradecimientos

Los autores agradecen el financiamiento al Proyecto *Escribe.DOC 2023047* de la Universidad de Concepción.