



Ejercicio 5

Estimación del Modelo de Regresión Lineal General

Pilar González y Susan Orbe

Dpto. Economía Aplicada III (Econometría y Estadística)

Contenido

- 1 Ejercicio 5.1. Alquiler de sombrillas (sombrillas.gdt).
- 2 Ejercicio 5.2. Casas rurales (NekaturBizkaia.gdt).
- 3 Ejercicio 5.3. Leche de soja (soja.gdt).

Contenido

- 1 Ejercicio 5.1. Alquiler de sombrillas (sombrillas.gdt).
- 2 Ejercicio 5.2. Casas rurales (NekaturBizkaia.gdt).
- 3 Ejercicio 5.3. Leche de soja (soja.gdt).

Ejercicio 5.1. Alquiler de sombrillas.

Primera Parte. Modelo de regresión lineal simple.

Estima un modelo de regresión lineal simple en el que el número de sombrillas alquiladas solamente dependa de la temperatura de forma lineal.

$$S_t = \alpha + \beta T_t + u_t \quad t = 1, \dots, 22 \quad (1)$$

- Estima el modelo por MCO y guarda tanto los residuos como los valores estimados.
- Escribe la función de regresión muestral.
- Interpreta el coeficiente que acompaña a la variable temperatura e indica si posee el signo esperado.
- Comenta el coeficiente de determinación del modelo.
- Obtén y comenta el gráfico de la serie a analizar junto a la estimada.

Ejercicio 5.1. Alquiler de sombrillas.

- g. ¿Cuántas sombrillas se han alquilado durante la temporada?
- h. ¿Cuál es la media semanal de sombrillas alquiladas?, ¿en cuánto se estima la media semanal de sombrillas alquiladas? ¿La media estimada difiere de la real?, ¿por qué?
- i. ¿En cuánto se estima el número de sombrillas alquiladas en la primera semana de agosto?
- j. ¿Cuál es el error de estimación que se comete al estimar el número de sombrillas alquiladas de la última semana de septiembre?, ¿cómo llamamos a ese error?, ¿a qué se debe ese error?
- k. Si la temperatura media de una semana fuera de 26 grados centígrados, ¿en cuánto se estima el número de sombrillas alquiladas de esa semana?
- l. Si la temperatura media de una semana a otra se elevara en 2 grados centígrados, ¿en cuánto se estima la diferencia en el número de sombrillas alquiladas?

Ejercicio 5.1. Alquiler de sombrillas.

Segunda parte. Modelo de regresión lineal general.

Estima un modelo de regresión lineal general en el que el número de sombrillas alquiladas dependa de la temperatura y del precio de alquiler de forma lineal.

$$S_t = \beta_1 + \beta_2 T_t + \beta_3 P_t + v_t \quad t = 1, \dots, 22 \quad (2)$$

- Estima el modelo por MCO y guarda tanto los residuos como los valores estimados.
- Escribe la función de regresión muestral.
- Compara los modelos (1) y (2). ¿En qué se diferencian?
- Interpreta los coeficientes estimados que acompañan a las variables temperatura y precio. Indica si poseen los signos esperados.
- ¿Coinciden las estimaciones $\hat{\beta}$ y $\hat{\beta}_1$ de los modelos (1) y (2) respectivamente?, ¿por qué?
- Comenta el coeficiente de determinación del modelo (2) y compáralo con el obtenido en el modelo (1).

Ejercicio 5.1. Alquiler de sombrillas.

- g. Obtén el gráfico de la serie a analizar junto a la estimada. Coméntalo y compáralo con el obtenido en el modelo (1).
- h. Considera las variables involucradas en el modelo (2), los valores estimados que se obtienen de este modelo y los residuos. Obtén los estadísticos principales y comenta los resultados obtenidos.
- i. Si la temperatura media de una semana fuera de 39 grados centígrados, ¿en cuánto se estima el número de sombrillas alquiladas de esa semana?, ¿y si el precio medio de alquiler por día fuera de 13 euros?
- j. Si la empresa familiar decidiera cobrar el mismo precio de alquiler a lo largo de la temporada, ¿qué efectos tendría sobre los resultados de estimación?

Ejemplo 5.1. Alquiler de sombrillas.

Tercera parte. Modelo de regresión lineal general.

Estima un modelo de regresión lineal general en el que el número de sombrillas alquiladas dependa de la temperatura, del precio de alquiler y de si ha habido semanas con viento o no.

$$S_t = \gamma_1 + \gamma_2 T_t + \gamma_3 P_t + \gamma_4 VB_t + w_t \quad t = 1, \dots, 22 \quad (3)$$

- Estima el modelo por MCO y escribe la FRM.
- Compara los modelos (2) y (3). ¿En qué se diferencian?
- ¿Cuál es el número de sombrillas alquiladas en las semanas que ha habido viento?, ¿y en las que no?
- Interpreta el coeficiente estimado que acompaña a la variable ficticia VB . ¿Tiene el signo esperado?
- ¿En cuánto se estima la variación en el número de sombrillas alquiladas si el precio aumenta en un euro manteniendo el resto de características constantes?
- ¿En cuánto estimas el número medio de sombrillas alquiladas si el precio medio semanal de alquiler es 7 euros y la temperatura media semanal es de 30 grados?, ¿y si a lo largo de la semana ha habido viento?, ¿y si no la ha habido?

Contenido

- 1 Ejercicio 5.1. Alquiler de sombrillas (sombrillas.gdt).
- 2 Ejercicio 5.2. Casas rurales (NekaturBizkaia.gdt).
- 3 Ejercicio 5.3. Leche de soja (soja.gdt).

Ejercicio 5.2. Casas rurales.

Primera parte. Modelo I.

Relaciona el precio por habitación con el número de habitaciones y el precio del desayuno por persona:

$$PR_i = \alpha_1 + \alpha_2 HB_i + \alpha_3 PD_i + u_i \quad (4)$$

- Estima el modelo (4) por MCO y escribe la FRM.
- Interpreta los coeficientes estimados de HB y PD.
- ¿Cuál es el precio estimado de una noche de habitación si el desayuno está incluido en el precio de la habitación y la casa rural dispone de 10 habitaciones?
- Si el desayuno está incluido en el precio de la habitación, ¿cuál es la variación en el precio estimado entre una casa rural que dispone de 15 habitaciones y otra que dispone de 10?
- Obtén y comenta el gráfico de la serie observada y la estimada.
- ¿Crees que el modelo especificado es adecuado para estos datos?

Ejercicio 5.2. Casas rurales.

Segunda parte. Modelo II.

Se generaliza el modelo anterior incluyendo las variables WIFI y LOC:

$$PR_i = \lambda_1 + \lambda_2 HB_i + \lambda_3 PD_i + \lambda_4 WIFIG_i + \lambda_5 WIFIS_i + \lambda_6 LOCC_i + u_i \quad (5)$$

donde *WIFIG* toma valor uno si la casa rural ofrece conexión gratuita y cero en caso contrario; *WIFIS* toma valor 1 si la casa rural ofrece conexión sujeta a un pago adicional; y, *LOCC* toma valor 1 si la casa rural está en el centro de la ciudad.

- Estima el modelo (5) por MCO y escribe la FRM.
- Interpreta los coeficientes estimados de las variables ficticias *WIFIG* y *WIFIS*.
¿Tienen los signos esperados?
- ¿Cuál es el precio estimado de una noche de habitación si el precio por persona del desayuno es de 3 euros, posee 6 habitaciones y tiene conexión wifi?, ¿y si la conexión se ofrece de forma gratuita?, ¿y si hay un coste fijo de 2 euros?
- ¿Cuál es el precio estimado de una noche para una habitación que tiene las mismas características que la primera observación de la muestra?, ¿difiere del precio real?, ¿por qué?

Ejercicio 5.2. Casas rurales.

Tercera parte. Modelo III.

A continuación se especifica el siguiente modelo:

$$\begin{aligned} PR_i &= \beta_1 + \beta_2 HB_i + \beta_3 PD_i + \beta_4 WIFIG_i + \beta_5 PNR_i + \\ &+ \beta_6 PLR_i + \beta_7 LGR_i + u_i \end{aligned} \quad (6)$$

donde PNR , PLR y LGR toman valor uno si la casa rural está a menos de un kilómetro de algún parque natural, de alguna playa y de algún lago o embalse respectivamente y cero en caso contrario.

- Estima el modelo (6) por MCO y escribe la FRM. ¿En qué se diferencia del modelo anterior? ¿Cuántas variables explicativas tiene?
- Observa cuáles son las variables ficticias incluidas y cuáles han sido excluidas. Interpreta los coeficientes estimados de las variables ficticias PNR , PLR y LGR . ¿Tienen los signos esperados?
- ¿Crees que este modelo se ajusta mejor a los datos?, ¿por qué?

Contenido

- 1 Ejercicio 5.1. Alquiler de sombrillas (sombrillas.gdt).
- 2 Ejercicio 5.2. Casas rurales (NekaturBizkaia.gdt).
- 3 Ejercicio 5.3. Leche de soja (soja.gdt).

Ejercicio 5.3. Leche de soja.

Primera parte. Organización del fichero de datos.

Una empresa vendedora de leche de soja quiere conocer el comportamiento de sus ventas (V , en miles de envases de un litro) en función de su precio (P , en céntimos de euros por litro) y de los gastos en publicidad (G , en cientos de euros). Para ello dispone de datos mensuales desde enero de 1990 a junio de 2012 sobre las ventas realizadas (V , en miles de euros), el precio (P , en euros) y los gastos en publicidad (G , en decenas de euros). Los datos correspondientes se encuentran en el fichero soja.gdt.

- Abre el fichero de datos y organiza la estructura de los datos teniendo en cuenta que son datos de series temporales con una frecuencia mensual.
- En el fichero de datos Y corresponde a las ventas, $X1$ al precio y $X2$ al gasto realizado en publicidad. Edita los atributos de las variables, incorporando la nomenclatura dada en el enunciado del problema y la información sobre las unidades de medida.
- Guarda el fichero de datos en el archivo ventas-soja.gdt.

Ejercicio 5.3. Leche de soja.

Segunda parte. $V = f(P)$.

Modelo de regresión lineal simple que relaciona las ventas con el precio de forma lineal:

$$V_t = \gamma_1 + \gamma_2 P_t + u_t \quad t = 1990 : 1, \dots, 2012 : 6. \quad (7)$$

- Analiza la información de los datos. ¿Cuál es el rango de ventas a lo largo de estos años? ¿Cuáles son las ventas medias del periodo?
- Obtén el coeficiente de correlación muestral de las variables ventas y precio.
- Estima el modelo por MCO y escribe la función de regresión muestral.
- Interpreta el coeficiente estimado de la variable precio. ¿Tiene el signo esperado?
- ¿En cuánto estimas que variarán las ventas si el precio de la leche de soja aumenta en 30 céntimos?
- Interpreta el coeficiente de determinación.
- Estima la varianza de las perturbaciones del modelo, σ^2 .
- ¿Cuál es la varianza del estimador $\hat{\gamma}_2^{MCO}$?
- Obtén y comenta el gráfico de la serie observada junto a la estimada.

Ejercicio 5.3. Leche de soja.

Tercera parte. $V = f(P, G)$.

Modelo de regresión lineal general en el que las ventas se determinan en función del precio y el gasto en publicidad:

$$V_t = \beta_1 + \beta_2 P_t + \beta_3 G_t + \beta_4 G_t^2 + u_t \quad t = 1990 : 1, \dots, 2012 : 6. \quad (8)$$

- ¿Cuántas variables explicativas tiene el modelo? ¿En que se diferencia del modelo anterior?
- ¿Cumple el modelo la hipótesis básica de linealidad?
- Obtén la matriz de correlaciones de las variables ventas, precio y gasto. ¿Tienen los signos esperados?
- Estima el modelo por MCO y escribe la función de regresión muestral.

Ejercicio 5.3. Leche de soja.

- e. ¿En cuánto estimas que variarán las ventas si el gasto en publicidad aumenta en 100 euros mientras que el precio de la leche de soja se mantiene constante?, ¿y si el gasto de publicidad en ese momento es de 1500 euros?, ¿y si el gasto de publicidad en ese momento es de 15000 euros?
- f. ¿En cuánto estimas que variarán las ventas si el precio por litro de leche aumenta en medio euro manteniendo el gasto en publicidad constante?, ¿y si en ese momento costaba 125 céntimos?, ¿y si en ese momento costaba 80 céntimos?
- g. ¿Cuáles son las ventas estimadas según este modelo para el diciembre de 1990? ¿En cuánto se diferencia del valor real?
- h. Si en julio del 2012 el gasto en publicidad realizado fuera de 14600 euros y el precio por litro de leche de soja se vendiera a 1,23 euros, ¿en cuánto se estima el número de envases vendidas?
- i. Estima la matriz de covarianzas del estimador MCO.
- j. Obtén y comenta el gráfico de la serie observada junto a la estimada.

Ejercicio 5.3. Leche de soja.

Cuarta parte. Tendencia.

Se propone generalizar el modelo anterior incluyendo una variable que recoja la tendencia presente en las ventas:

$$V_t = \alpha_1 + \alpha_2 P_t + \alpha_3 G_t + \alpha_4 G_t^2 + \alpha_5 t + u_t \quad t = 1990 : 1, \dots, 2012 : 6. \quad (9)$$

- Añade a la base de datos una tendencia y estima por MCO el modelo (9).
- Obtén el gráfico de la serie a analizar junto a la estimada. Comenta lo que observas.
- Interpreta el coeficiente estimado que acompaña a la variable tendencia.
- En base a todos los resultados obtenidos, ¿crees que la variable tendencia ha aportado información adicional al modelo?, ¿por qué?