

Solución Ejercicio 5.

Estimación del Modelo de Regresión Lineal General.

Ejercicio 5.1 Alquiler de sombrillas

Primera parte

a. Modelo a estimar: $S_t = \beta_1 + \beta_2 T_t + u_t \quad t = 1, \dots, 22$

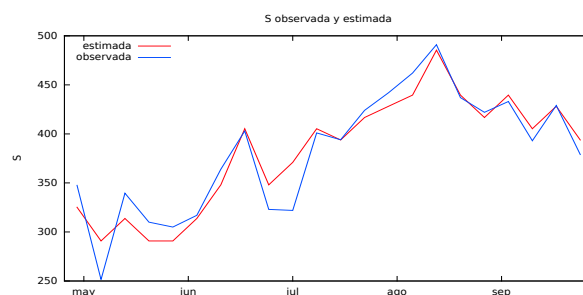
Modelo 1: MCO, usando las observaciones 2012-04-30–2012-09-24 ($T = 22$)

Variable dependiente: S

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	27,6858	26,9302	1,0281	0,3162
T	11,4418	0,860534	13,2962	0,0000
Media de la vble. dep.	381,3409	D.T. de la vble. dep.	60,52005	
Suma de cuad. residuos	7817,177	D.T. de la regresión	19,77015	
R^2	0,898368	R^2 corregido	0,893286	
$F(1, 20)$	176,7876	Valor p (de F)	2,17e-11	
Log-verosimilitud	-95,82005	Criterio de Akaike	195,6401	
Criterio de Schwarz	197,8222	Hannan-Quinn	196,1541	
$\hat{\rho}$	0,126241	Durbin-Watson	1,662734	

b. FRM: $\hat{S}_t = 27,6858 + 11,4418 T_t$

- c. Se estima que la variación en el número de sombrillas alquiladas es de 11,4418 sombrillas cuando la temperatura aumenta en un grado centígrado. Tiene el signo esperado porque a mayor temperatura, más sombrillas alquiladas.
- d. Se explica el 89,8368 % de las variaciones de las sombrillas alquiladas de la muestra a través de las variaciones de la temperatura.
- e. El ajuste parece adecuado, se recoge el comportamiento a largo plazo de la serie de alquiler de sombrillas.



g. En Añadir -- Definir nueva variable escribir $total = sum(S)$

Se han alquilado 8389,5 sombrillas durante la temporada.

- h. Seleccionamos la variable S y con el botón derecho del ratón optamos por estadísticos principales. La media semanal de sombrillas alquiladas es 381,34 y coincide con la media semanal estimada de sombrillas alquiladas porque es una de las propiedades de la FRM.
- i. Guardamos los valores estimados y luego consultamos dichos valores. Se estima que el número de sombrillas alquiladas en la primera semana de agosto es de 439,59 unidades.
- j. Guardamos los residuos del modelo y consultamos el último valor. El error de estimación es de -14,82286 sombrillas, se ha sobreestimado el último valor de la muestra. A este error se le denomina residuo y se debe al error de estimación derivado de estimar los coeficientes del modelo y que la perturbación es impredecible.
- k. Si la temperatura media de una semana fuera de 26 grados centígrados, el número estimado de sombrillas alquiladas de esa semana sería de 325,1726 unidades.
- l. Si la temperatura media de una semana a otra se elevara en 2 grados centígrados, la diferencia estimada en el número de sombrillas alquiladas sería de $2 \times \hat{\beta}_2 = 22,8836$ unidades.

Segunda parte

- a. Modelo a estimar: $S_t = \beta_1 + \beta_2 T_t + \beta_3 P_t + v_t \quad t = 1, \dots, 22$

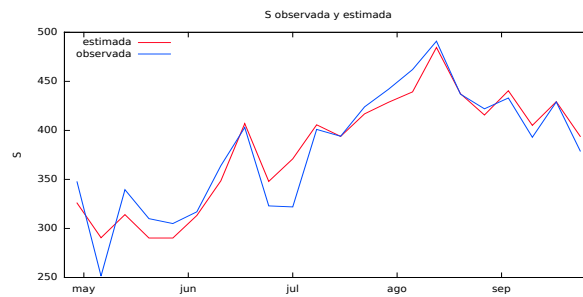
Modelo 2: MCO, usando las observaciones 2012-04-30–2012-09-24 ($T = 22$)
 Variable dependiente: S

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	22,1467	40,7666	0,5433	0,5933
T	11,7892	2,07824	5,6727	0,0000
P	-0,610208	3,30474	-0,1846	0,8555
Media de la vble. dep.	381,3409	D.T. de la vble. dep.	60,52005	
Suma de cuad. residuos	7803,175	D.T. de la regresión	20,26557	
R^2	0,898550	R^2 corregido	0,887871	
$F(2, 19)$	84,14186	Valor p (de F)	3,63e-10	
Log-verosimilitud	-95,80033	Criterio de Akaike	197,6007	
Criterio de Schwarz	200,8738	Hannan-Quinn	198,3717	
$\hat{\rho}$	0,152127	Durbin-Watson	1,616248	

- b. FRM: $\hat{S}_t = 22,1467 + 11,7892 T_t - 0,610208 P_t \quad t = 1, \dots, 22$
- c. Este modelo contiene una variable explicativa más, el precio medio de alquiler de las sombrillas.
- d. Interpretación de coeficientes.
 $\hat{\beta}_2$: Se estima que el número de sombrillas alquiladas se incrementa en 11,7892 unidades cuando la temperatura aumenta en un grado centígrado y el precio se mantiene constante. Tiene el signo esperado porque cuanto más calor hace más sombrillas se alquilan.

$\hat{\beta}_3$: Se estima que el número de sombrillas alquiladas se reduce en 0,610208 unidades cuando el precio aumenta en un euro y la temperatura mantiene constante. Tiene el signo esperado porque cuanto más caro sea el alquiler menos se alquilan. Dado el valor obtenido hace falta incrementar el precio en dos euros para reducir el número de sombrillas alquiladas en una unidad.

- e. No, porque aunque son estimaciones que corresponden a la misma variable dependiente provienen de dos modelos diferentes.
- f. Se explica el 89,8550 % de las variaciones de las sombrillas alquiladas de la muestra a través de las variaciones de la temperatura y el precio. El valor del coeficiente de determinación es mayor que el anterior porque el modelo contiene una variable explicativa más. Esto no implica que el modelo está mejor especificado, habría que analizar la significatividad de esa variable adicional.
- g. No se observan grandes cambios en el gráfico de la variable observada y la estimada en comparación con el modelo anterior. Parece que la variable precio no ha tenido mucha influencia en los resultados que se han obtenido.



h. En base a los estadísticos descriptivos:

Ver -- Estadísticos principales

Trás seleccionar las variables del fichero para las que se quiere calcular estos estadísticos, se obtienen los siguientes resultados:

Estadísticos principales, usando las observaciones 2012-04-30 - 2012-09-24

Variable	Media	Mediana	Mínimo	Máximo
S	381,341	393,500	251,500	491,000
P	8,52273	9,00000	4,50000	15,0000
T	30,9091	32,5000	23,0000	40,0000
yhat2	381,341	399,500	290,248	484,563
uhat2	0,000000	1,93637	-48,9424	25,3681

Variable	Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
S	60,5200	0,158703	-0,249934	-0,700091
P	3,15277	0,369925	0,495965	-0,466571
T	5,01340	0,162198	-0,267677	-1,06131
yhat2	57,3681	0,150438	-0,283110	-1,06232
uhat2	19,2764	1,24342e+015	-0,966128	0,512180

Variable	porc. 5 %	porc. 95 %	Rango IQ	Observaciones ausentes
S	259,525	486,650	107,250	0
P	4,50000	15,0000	4,25000	0
T	23,0000	39,4000	9,25000	0
yhat2	290,248	477,947	105,847	0
uhat2	-47,4590	24,9773	23,5332	0

Tal y como indican las propiedades de la FRM, las medias de la variable endógena y la estimada coinciden y la media de los residuos es cero. La variable alquiler de sombrillas es la que más variabilidad presenta y la variable precio la que menos. El precio medio de alquiler de la temporada es de 8,52273 euros y la temperatura media ha sido de 30,9091 grados centígrados.

- i. Si la temperatura media de una semana fuera de 39 grados centígrados, número estimado de sombrillas alquiladas de esa semana sería $\hat{S}_t = 481,9255 - 0,610208P_t$.

Y si el precio medio de alquiler por día fuera de 13 euros, entonces sería $\hat{S}_t = 473,9928$.

- j. Si la empresa familiar decidiera cobrar el mismo precio de alquiler a lo largo de la temporada, entonces la tercera columna de la matriz de datos sería constante. Esto implica que la matriz $X'X$ no es de rango completo y que no sería posible estimar todos los coeficientes del modelo de forma individual. Este problema se conoce como multicolinealidad perfecta.

Tercera parte

- a. Modelo a estimar: $S_t = \beta_1 + \beta_2 T_t + \beta_3 P_t + \beta_4 VB_t + v_t \quad t = 1, \dots, 22$

Modelo 3: MCO, usando las observaciones 2012-04-30–2012-09-24 ($T = 22$)

Variable dependiente: S

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	45,0017	45,7826	0,9829	0,3387
T	11,0405	2,18248	5,0587	0,0001
P	-0,0715119	3,32814	-0,0215	0,9831
VB	-10,5178	9,74517	-1,0793	0,2947
Media de la vble. dep.		381,3409	D.T. de la vble. dep.	60,52005
Suma de cuad. residuos		7328,892	D.T. de la regresión	20,17822
R^2		0,904716	R^2 corregido	0,888835
$F(3, 18)$		56,96957	Valor p (de F)	2,18e-09
Log-verosimilitud		-95,11056	Criterio de Akaike	198,2211
Criterio de Schwarz		202,5853	Hannan-Quinn	199,2492
$\hat{\rho}$		0,107235	Durbin-Watson	1,684163

FRM: $\hat{S}_t = 45,0017 + 11,0405 T_t - 0,0715119 P_t - 10,5178 VB_t \quad t = 1, \dots, 22$

- b. Se ha añadido la variable explicativa *semana con viento*. Esta variable es cualitativa y ha sido introducida en el modelo a través de la variable ficticia VB que toma el valor 1 cuando la observación pertenece a una semana con viento y cero en caso contrario.
- c. El número de sombrillas alquiladas en las semanas que ha habido viento es 3132 (obtenido mediante $\text{sum}(S*VB)$). Cuando no ha habido viento, el número de sombrillas alquiladas ha sido 5257,5 (obtenido mediante $\text{sum}(S) - \text{sum}(S*VB)$).

- d. Se estima que la diferencia de sombrillas alquiladas entre una semana en la que NO ha habido viento y otra en la que SÍ ha habido viento es de 10,5178 sombrillas manteniendo el resto de características (temperatura y precio de la sombrilla) constantes.
- e. Se estima que el número de sombrillas alquiladas cuando el precio aumenta en un euro manteniendo el resto de características constantes disminuye en 0,0715119 unidades.
- f. Se estima que el número medio de sombrillas alquiladas cuando el precio medio semanal de alquiler es de 7 euros y la temperatura media semanal es de 30 grados sea $\hat{S}_t = 375,71611 - 10,5178 VB_t$ sombrillas.

Si ha habido viento: $\hat{S}_t = 365,1983167$ sombrillas.

Si no ha habido viento: $\hat{S}_t = 375,71611$ sombrillas.

Ejercicio 5.2 Casas rurales

Primera parte

- a. El modelo a estimar es $PR_i = \alpha_1 + \alpha_2 HB_i + \alpha_3 PD_i + u_i$

Modelo 1: MCO, usando las observaciones 1–75
Variable dependiente: PR

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	38,4321	7,22899	5,3164	0,0000
HB	2,26766	1,20082	1,8884	0,0630
PD	1,49558	1,09746	1,3628	0,1772
Media de la vble. dep.	56,13893	D.T. de la vble. dep.	14,98446	
Suma de cuad. residuos	15263,15	D.T. de la regresión	14,55982	
R^2	0,081392	R^2 corregido	0,055875	
$F(2, 72)$	3,189724	Valor p (de F)	0,047064	
Log-verosimilitud	-305,7595	Criterio de Akaike	617,5189	
Criterio de Schwarz	624,4714	Hannan–Quinn	620,2950	

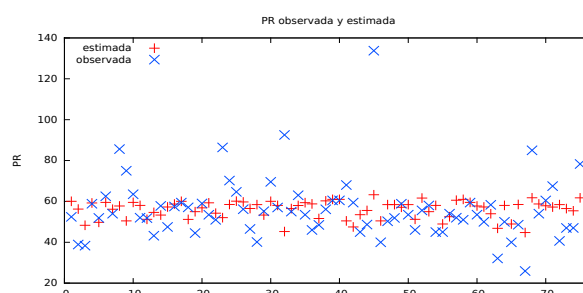
FRM: $\widehat{PR}_i = 38,4321 + 2,26766 HB_i + 1,49558 PD_i$

- b. Interpretación de los coeficientes.

$\hat{\alpha}_2$: La variación estimada en el precio de la habitación cuando la casa rural dispone de una habitación adicional es de 2,26766 euros manteniendo el precio del desayuno constante.

$\hat{\alpha}_3$: La variación estimada en el precio de la habitación cuando el precio del desayuno aumenta en un euro es de 1,49558 euros manteniendo el número de habitaciones constante.

- c. El precio estimado de una noche de habitación cuando el desayuno está incluido en el precio y la casa rural dispone de 10 habitaciones es de 61,1087 euros.
- d. Si el desayuno está incluido en el precio de la habitación, la variación en el precio estimado entre una casa rural que dispone de 15 habitaciones y otra que dispone de 10 es de $\hat{\alpha}_2 \times 5 = 11,3383$ euros.
- e. Se observa que el ajuste es un poco escaso dado que solamente se ajusta el nivel de la serie de precios.



f. No, dado el valor del coeficiente de determinación (0,081392) el ajuste es muy pobre.

Segunda parte

a. El modelo a estimar es:

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

Modelo 2: MCO, usando las observaciones 1–75

Variable dependiente: PR

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	40,5761	7,39661	5,4858	0,0000
HB	1,94192	1,21303	1,6009	0,1140
PD	0,559911	1,21918	0,4593	0,6475
WIFIG	6,98544	3,65362	1,9119	0,0600
WIFIS	-5,75696	12,0827	-0,4765	0,6352
LOCC	2,11170	5,43209	0,3887	0,6987
Media de la vble. dep.		56,13893	D.T. de la vble. dep.	14,98446
Suma de cuad. residuos		14429,98	D.T. de la regresión	14,46133
R^2		0,131536	R^2 corregido	0,068604
$F(5, 69)$		2,090130	Valor p (de F)	0,077001
Log-verosimilitud		-303,6545	Criterio de Akaike	619,3089
Criterio de Schwarz		633,2138	Hannan–Quinn	624,8610

FRM:

$$\widehat{PR}_i = 40,5761 + 1,94192 HB_i + 0,559911 PD_i + 6,98544 WIFIG_i - 5,75696 WIFIS_i + 2,11170 LOCC_i$$

b. Interpretación de coeficientes.

$\hat{\lambda}_4$: La diferencia estimada entre los precios de habitación de una casa rural que ofrece wifi gratis frente a una que no dispone de wifi, es de 6,98544 euros manteniendo el resto de características iguales. Se espera que su signo sea positivo y que se eleve el precio ya que se está ofreciendo un servicio añadido.

$\hat{\lambda}_5$: La diferencia estimada entre los precios de habitación de una casa rural que ofrece wifi sujeta a un pago adicional frente a una que no dispone de wifi, es de -5,75696 euros manteniendo el resto de características iguales. No se espera que su signo sea positivo ya que se está ofreciendo la posibilidad de optar por un servicio.

c. El precio estimado de una noche de habitación cuando el precio por persona del desayuno es de 3 euros, posee 6 habitaciones y tiene conexión wifi es

$$\widehat{PR}_i = 53,907353 + 6,98544 WIFIG_i - 5,75696 WIFIS_i + 2,11170 LOCC_i \text{ euros.}$$

Y cuando la conexión se ofrece de forma gratuita es

$$\widehat{PR}_i = 60,892793 + 2,11170 LOCC_i \text{ euros.}$$

Y si hay un coste fijo de 2 euros $\widehat{PR}_i = 48,150393 + 2,11170 LOCC_i \text{ euros.}$

- d. El precio estimado de la primera observación de la muestra es 62,20859 euros mientras que la real es de 52,430 euros y por lo tanto el valor estimado difiere de la real. Esta diferencia, denominada residuo, se debe a los errores de estimación y al hecho de que exista la perturbación en el modelo.

Tercera parte

- a. El modelo a estimar es

$$PR_i = \beta_1 + \beta_2 HB_i + \beta_3 PD_i + \beta_4 WIFIG_i + \beta_5 PNR_i + \beta_6 PLR_i + \beta_7 LGR_i + u_i$$

Modelo 3: MCO, usando las observaciones 1–75
Variable dependiente: PR

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	33,9408	6,72069	5,0502	0,0000
HB	1,36156	1,11199	1,2244	0,2250
PD	1,62738	1,07599	1,5124	0,1351
WIFIG	9,02246	3,32099	2,7168	0,0084
PNR	3,33934	3,93959	0,8476	0,3996
PLR	16,1587	4,47207	3,6133	0,0006
LGR	12,0185	7,81149	1,5386	0,1285
Media de la vble. dep.	56,13893	D.T. de la vble. dep.	14,98446	
Suma de cuad. residuos	11550,44	D.T. de la regresión	13,03301	
R^2	0,304840	R^2 corregido	0,243503	
$F(6, 68)$	4,969876	Valor p (de F)	0,000285	
Log-verosimilitud	-295,3075	Criterio de Akaike	604,6151	
Criterio de Schwarz	620,8375	Hannan–Quinn	611,0925	

FRM:

$$\widehat{PR}_i = 33,9408 + 1,36156 HB_i + 1,62738 PD_i + 9,02246 WIFIG_i + 3,33934 PNR_i + 16,1587 PLR_i + 12,0185 LGR_i$$

Este modelo se diferencia del anterior en dos aspectos. Primero, tiene tres variables explicativas añadidas. Para introducir las nuevas variables “cercanía a un parque natural”, “cercanía a un lago o embalse” y “cercanía a una playa”, solo se han introducido, respectivamente, las variables ficticias PNR , LGR , PLR , por lo que solo se discrimina entre dos grupos de agroturismos: los que están a menos de un kilómetro del servicio considerado, y el resto, que se considera que están medianamente lejos.

Segundo, la variable explicativa cualitativa común wifi difiere en el número de variables ficticias incorporadas, en este caso solamente se incluye el hecho de ofrecer wifi gratis versus los que o no la tienen o es de pago.

El modelo incorpora seis variables explicativas.

- b. Interpretación de coeficientes.

$\hat{\beta}_5$: La diferencia estimada entre los precios de habitación de una casa rural que se sitúa a menos de un kilómetro de un parque natural frente a una que no, es de

3,33934 euros manteniendo el resto de características iguales. Se espera que su signo sea positivo y que esta cercanía eleve el precio.

$\hat{\beta}_6$: La diferencia estimada entre los precios de habitación de una casa rural que se sitúa a menos de un kilómetro de una playa frente a una que no, es de 16,1587 euros manteniendo el resto de características iguales. Se espera que su signo sea positivo y que esta cercanía eleve el precio.

$\hat{\beta}_7$: La diferencia estimada entre los precios de habitación de una casa rural que se sitúa a menos de un kilómetro de un lago o embalse frente a una que no, es de 12,0185 euros manteniendo el resto de características iguales. Se espera que su signo sea positivo y que esta cercanía eleve el precio.

- c. Sí, parece que el ajuste es algo mejor. En este modelo el coeficiente de determinación es 0,304840 con lo que se triplica el valor anterior. No obstante habría que analizar la significatividad de las variables.

Ejercicio 5.3 Leche de soja

Primera parte

Consultar fichero ventas-soja.gdt.

Segunda parte

En esta parte trabajamos con el siguiente modelo:

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

- a. Estadísticos principales de la variable ventas.

Estadísticos principales, usando las observaciones 1990:01 - 2012:06 para la variable V (270 observaciones válidas)

Media	Mediana	Mínimo	Máximo
120,848	119,963	76,2000	171,506
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
16,8791	0,139673	0,228360	0,729406
porc. 5 %	porc. 95 %	Rango IQ	Observaciones ausentes
89,3864	152,127	19,1710	0

Rango: 76,2 - 171,51.

Ventas medias del periodo: 120,85 miles de envases de un litro.

- b. El coeficiente de correlación muestral de las variables ventas y precio es: $\text{corr}(V, P) = -0,60412271$.
- c. Resultados de estimación:

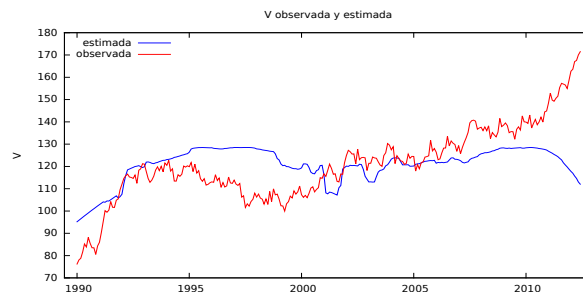
Modelo 1: MCO, usando las observaciones 1990:01–2012:06 ($T = 270$)

Variable dependiente: V

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	204,889	6,82123	30,0370	0,0000
P	-0,846206	0,0681840	-12,4106	0,0000
Media de la vble. dep.	120,8475	D.T. de la vble. dep.	16,87912	
Suma de cuad. residuos	48668,71	D.T. de la regresión	13,47589	
R^2	0,364964	R^2 corregido	0,362595	
$F(1, 268)$	154,0235	Valor p (de F)	3,01e-28	
Log-verosimilitud	-1084,353	Criterio de Akaike	2172,707	
Criterio de Schwarz	2179,903	Hannan–Quinn	2175,597	
$\hat{\rho}$	0,994798	Durbin–Watson	0,042141	

FRM: $\hat{V}_t = 204,889 - 0,846206 P_t$

- d. $\hat{\gamma}_2$. Se estima que las ventas disminuyen en 846,206 envases ante un aumento del precio un céntimo de euro. Tiene el signo negativo esperado de forma que cuanto más caro sea menos se vende.
- e. Se estima una disminución en las ventas de $30 \times \hat{\gamma}_2 = 25,38618$ miles de envases de un litro si el precio de la leche de soja aumenta en 30 céntimos.
- f. Se explica el 36,4964 % de la variabilidad de las ventas de soja con la variabilidad del precio de forma lineal.
- g. $\hat{\sigma}^2 = \frac{SCR}{T-k} = 181,597$.
- h. $\widehat{Var}(\hat{\gamma}_2^{MCO}) = 0,0681840^2 = 0,00464905$.
- i. Se observa que el ajuste es muy malo, no se recoge si siquiera el comportamiento promedio de la serie de alquiler.



Tercera parte

El modelo que vamos a analizar es:

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

- a. El modelo incluye dos variables explicativas: precio y gasto en publicidad. Este modelo se diferencia del anterior en que incluye una variable adicional: el gasto en publicidad.
- b. Sí, el modelo es lineal en los coeficientes. La relación entre el precio y la variable gasto es cuadrática pero esto no afecta a las hipótesis básicas del MRLG.
- c. Matriz de correlaciones.

Coeficientes de correlación, usando las observaciones 1990:01 - 2012:06
 valor crítico al 5 % (a dos colas) = 0,1194 para n = 270

	V	P	GP	
1,0000		-0,6041	0,1614	V
		1,0000	-0,7444	P
			1,0000	GP

Sí tienen los signos esperados. Las ventas se relacionan de forma proporcional a los gastos en publicidad (a más publicidad más ventas) e inversamente proporcional a los precio (cuanto más caro menos ventas).

d. Modelo que vamos a estimar:

$$V_t = \beta_1 + \beta_2 P_t + \beta_3 G_t + \beta_4 G_t^2 + u_t$$

Modelo 2: MCO, usando las observaciones 1990:01–2012:06 ($T = 270$)

Variable dependiente: V

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	57,5860	43,7095	1,3175	0,1888
P	-1,38907	0,0807449	-17,2032	0,0000
GP	3,77879	0,610062	6,1941	0,0000
sq_GP	-0,0166627	0,00229460	-7,2617	0,0000
Media de la vble. dep.	120,8475	D.T. de la vble. dep.	16,87912	
Suma de cuad. residuos	28697,27	D.T. de la regresión	10,38674	
R^2	0,625554	R^2 corregido	0,621331	
$F(3, 266)$	148,1277	Valor p (de F)	1,89e-56	
Log-verosimilitud	-1013,042	Criterio de Akaike	2034,083	
Criterio de Schwarz	2048,477	Hannan-Quinn	2039,863	
$\hat{\rho}$	0,965596	Durbin-Watson	0,084094	

$$\text{FRM: } \hat{V}_t = 57,5860 - 1,38907P_t + 3,77879G_t - 0,0166627G_t^2$$

e. La variación estimada en las ventas si el gasto en publicidad aumenta en 100 euros y se mantiene el precio de la leche de soja constante es de $3,77879 - 2 \times 0,0166627 \times G_t$ miles de envases. Este efecto no es constante a lo largo de la muestra porque depende del nivel de gasto que se realiza en ese momento de tiempo.

Si el gasto de publicidad en ese momento es de 1500 euros, la variación estimada es de $3,77879 - 2 \times 0,0166627 \times 15 = 3,278909$ miles de envases. Se logra aumentar las ventas en 3278,909 envases.

Si el gasto de publicidad en ese momento es de 15000 euros, la variación estimada es de $3,77879 - 2 \times 0,0166627 \times 150 = -1,22002$ miles de envases. No se logra aumentar las ventas, de hecho disminuyen en 1220,002 envases.

f. La variación estimada en las ventas cuando el precio por litro de leche aumenta en un céntimo de euro y se mantiene el gasto en publicidad constante es de $\hat{\beta}_2$ miles de envases. Esta variación es constante a lo largo de la muestra.

Por tanto si el precio aumenta en medio euro, la disminución variación estimada es de $50\hat{\beta}_2 = 69,4535$ miles de envases. Es decir, se estima una disminución de 69453,5 envases.

Si en ese momento la leche de soja costaba 125 céntimos o 80 céntimos, la disminución estimada sigue siendo la misma: 69,4535 miles de envases.

g. Las ventas estimadas para diciembre de 1990 son de 103,4355 miles de envases. Se diferencia del valor real (84,3639) en -19,0716 miles de envases. Como el valor del residuo es negativo, el modelo ha sobreestimado las ventas de ese momento.

h. Predicción por punto.

$$\hat{V}_{2012:7} = 57,5860 - 1,38907P_{2012:7} + 3,77879G_{2012:7} - 0,0166627G_{2012:7}^2$$

$$\hat{V}_{2012:7} = 57,5860 - 1,38907 \times 123 + 3,77879 \times 146 - 0,0166627 \times 146^2 = 83,2516$$

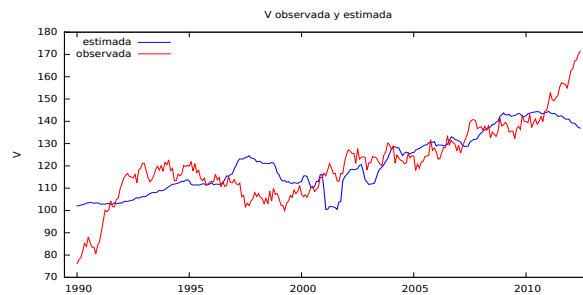
miles de envases.

i. Matriz de covarianzas del estimador MCO.

Matriz de covarianzas de los coeficientes

const	GP	sq_GP	
2977.6	-46.126	0.17546	const
	0.71824	-0.0027443	GP
		1.0527e-005	sq_GP

j. El ajuste ha mejorado respecto al modelo anterior pero aún no se recoge la trayectoria a largo plazo o la tendencia de la serie.



Cuarta parte

a. El modelo a estimar es

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

Modelo 3: MCO, usando las observaciones 1990:01–2012:06 ($T = 270$)

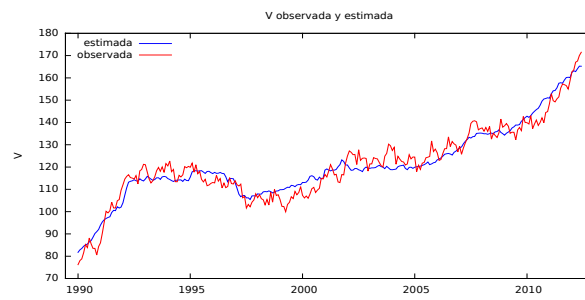
Variable dependiente: V

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	-385.413	23.1683	-16.6353	0.0000
P	2.64647	0.124963	21.1780	0.0000
GP	2.17358	0.270360	8.0396	0.0000
sq_GP	-0.00633613	0.00104691	-6.0522	0.0000
time	0.495534	0.0147224	33.6585	0.0000
Media de la vble. dep.	120.8475	D.T. de la vble. dep.	16.87912	
Suma de cuad. residuos	5440.154	D.T. de la regresión	4.530881	
R^2	0.929016	R^2 corregido	0.927945	
$F(4, 265)$	867.0608	Valor p (de F)	7.5e-151	
Log-verosimilitud	-788.5374	Criterio de Akaike	1587.075	
Criterio de Schwarz	1605.067	Hannan-Quinn	1594.300	
$\hat{\rho}$	0.813395	Durbin-Watson	0.372147	

FRM:

$$\hat{V}_t = -385,413 + 2,64647P_t + 2,17358G_t - 0,00633613G_t^2 + 0,495534 time$$

b. El ajuste ha mejorado mucho respecto al modelo anterior. El comportamiento a largo plazo de la serie de ventas se recoge adecuadamente. Sin embargo aún queda por explicar las pequeñas oscilaciones de la serie.



- c. La tasa de variación anual estimada es de 495,534 envases manteniendo el precio y el gasto en publicidad constante.
- d. Sí, el gráfico de la serie ajustada sugiere que esta variable aporta información relevante para determinar las ventas de leche de soja. Habría que realizar un contraste para verificarlo.