Solución Ejercicio 6.

Inferencia en el Modelo de Regresión Lineal General.

Ejercicio 6.1 Alquiler de sombrillas

Primera parte

a. Modelo a estimar: $S_t = \alpha + \beta T_t + u_t \quad t = 1, \ldots, 22$

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

	Coeficiente	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

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

b. Contraste de significatividad de la variable temperatura.

$$\begin{array}{l} H_0 : \beta = 0 \\ H_a : \beta \neq 0 \end{array} \qquad t = \frac{\hat{\beta} - 0}{\hat{\sigma}_{\hat{\beta}}} \overset{H_0}{\sim} t(T-k)$$

Como $\mid t \mid = 13,2962 > 2,08596 = t_{0,025}(20)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable *temperatura* es individualmente significativa.

c. Contraste bilateral.

$$\begin{array}{l} H_0 : \beta = 20 \\ H_a : \beta \neq 20 \end{array} \qquad t = \frac{\hat{\beta} - 20}{\hat{\sigma}_{\hat{\beta}}} \overset{H_0}{\sim} t(T-k)$$

Como $\mid t \mid = \mid \frac{11,4418-20}{0,860534} \mid = \mid -9,94522 \mid > 2,08596 = t_{0,025}(20)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, si la temperatura media de una semana a la siguiente aumentara en dos grados centígrados, NO sería posible que el número de sombrillas alquiladas aumentase en 20 unidades.

d. Predicción por punto.

$\hat{S}_t = 27,6858 + 11,4418 \times 42 = 508,2414$ sombrillas.

e. Predicción por intervalo.

$IC(\beta_2)_{0,95} = [461,61\,;\,554,87]$

Se estima que el número máximo de sombrillas alquiladas será de 554,87.

Segunda parte

a. Modelo: $S_t = \gamma_1 + \gamma_2\,T_t + \gamma_3\,P_t + \gamma_4 VB_t + v_t \quad t = 1,\ldots,22$

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

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

b. Contraste de significatividad de la variable temperatura.

$$\begin{array}{l} H_0 : \gamma_2 = 0 \\ H_a : \gamma_2 \neq 0 \end{array} \qquad t = \frac{\hat{\gamma}_2 - 0}{\hat{\sigma}_{\hat{\gamma}_2}} \overset{H_0}{\sim} t(T-k)$$

Como $\mid t \mid = 5{,}059 > 2,10092 = t_{0,025}(18)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable *temperatura* es individualmente significativa.

Contraste de significatividad de la variable precio.

$$\begin{array}{l} H_0 : \gamma_3 = 0 \\ H_a : \gamma_3 \neq 0 \end{array} \qquad t = \frac{\hat{\gamma}_3 - 0}{\hat{\sigma}_{\hat{\gamma}_3}} \overset{H_0}{\sim} t(T-k)$$

Como $\mid t \mid = 0,0215 < 2,10092 = t_{0,025}(18)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable *precio* no es individualmente significativa.

Contraste de significatividad de la variable semana con viento.

$$\begin{array}{l} H_0 : \gamma_4 = 0 \\ H_a : \gamma_4 \neq 0 \end{array} \qquad t = \frac{\hat{\gamma}_4 - 0}{\hat{\sigma}_{\hat{\gamma}_4}} \overset{H_0}{\sim} t(T-k)$$

Como $\mid t \mid = 1,0793 < 2,10092 = t_{0,025}(18)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable *semana con viento* no es individualmente significativa.

Contraste de significatividad conjunta de las variables explicativas.

$H_0 : \gamma_2 = \gamma_3 = \gamma_4 = 0$
H_a : alguna identidad no se cumple

$$F = \frac{R^2/(k-1)}{(1-R^2)/T-k} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Como $F = 56,96957 > 3,15991 = \mathcal{F}_{0,05}(3,18)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto las variables explicativas del modelo son conjuntamente significativas.

c. Dado que las variables precio y semana con viento son variables explicativas individualmente no significativas, de existir colinealidad existiría entre estas variables siempre que la significatividad conjunta de ellas fuera significativa.

   Contraste de significatividad conjunta de las variables precio y semana con viento.

   $H_0 : \gamma_3 = \gamma_4 = 0$
   H_a : alguna identidad no se cumple

$$F = \frac{SCR_R - SCR_{NR}}{SCR_R}\frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

   Como $F = 0,599622 < 3,55456 = \mathcal{F}_{0,05}(2,18)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto las variables explicativas precio y semana de viento no son conjuntamente significativas.

   Las variables precio y semana con viento no son ni individualmente ni conjuntamente significativas. Por tanto no existen problemas de multicolinealidad alta.

b. Predicción por intervalo

   Semana con viento: $IC(Y_p)_{0,95} = [440,14\,;\, 530,57]$

   Se estima que el número máximo de sombrillas alquiladas es 530 sombrillas.

   Semana sin viento: $IC(Y_p)_{0,95} = [422,59\,;\, 549,58]$

   Se estima que el número máximo de sombrillas alquiladas es 549 sombrillas.

e. El modelo (2). Las variables precio y semana con viento resultan no significativas tanto individualmente como conjuntamente por lo que incorporar las restricciones verdaderas al modelo $\beta_3 = \beta_4 = 0$ reduce la varianza del estimador.

Ejercicio 6.2 Casas rurales

Primera parte

Modelo: $PR_i = \alpha_1 + \alpha_2\, HB_i + \alpha_3\, PD_i + u_i$

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

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

a. Contraste de significatividad conjunta de las variables explicativas.

$H_0 : \alpha_2 = \alpha_3 = 0$
H_a : alguna identidad no se cumple

$$F = \frac{R^2/(k-1)}{(1-R^2)/T-k} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Como $F = 3,189724 > 3,12391 = \mathcal{F}_{0,05}(2,72)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto las variables explicativas del modelo son conjuntamente significativas.

b. Contraste de significatividad de la variable número de habitaciones.

$H_0 : \alpha_2 = 0$
$H_a : \alpha_2 \neq 0$

$$t = \frac{\hat{\alpha}_2 - 0}{\hat{\sigma}_{\hat{\alpha}_2}} \overset{H_0}{\sim} t(T-k)$$

Como $\mid t \mid = 1,888 < 1,99346 = t_{0,025}(72)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable número de habitaciones no es individualmente significativa.

c. Intervalo de confianza de un coeficiente.

$IC(\alpha_3)_{0,95} = [-0,692164\,;\, 3,68333]$

Si el precio del desayuno por persona aumenta en un euro y el número de habitaciones que oferta la casa rural se mantiene, la variación estimada del precio medio de la habitación oscila entre -0,692164 y 3,68333 euros.

d. Predicción por punto.

$\widehat{PR}_i = 38,4321 + 2,26766 \times 10 + 1,49558 \times 3 = 65,59544$ euros.

e. Predicción por intervalo.

$IC(PR_p)_{0,95} = [33,345\,;\, 97,846]$

Si una casa rural tiene 10 habitaciones y ofrece desayunos por tres euros, se estima que el precio de la habitación oscila entre 33,345 y 97,846 euros.

Segunda parte

a. Modelo: $PR_i = \lambda_1 + \lambda_2\, HB_i + \lambda_3\, PD_i + \lambda_4\, WIFIG_i + \lambda_5\, WIFIS_i + \lambda_6\, LOCC_i + u_i$

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. Contraste de significatividad conjunta de las variables wifi y localización.

$H_0 : \lambda_4 = \lambda_5 = \lambda_6 = 0$
H_a : alguna identidad no se cumple

$$F = \frac{SCR_R - SCR_{NR}}{SCR_R}\frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Como $F = 1,328 < 2,73749 = \mathcal{F}_{0,05}(3,69)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto las variables explicativas wifi y localización no son conjuntamente significativas.

c. Contraste unilateral.

$H_0 : \lambda_6 \leq 0$
$H_a : \lambda_6 > 0$

$$t = \frac{\hat{\alpha}_6 - 0}{\hat{\sigma}_{\hat{\alpha}_6}} \overset{H_0}{\sim} t(T-k)$$

Como $t = 0,3887 < 1,66724 = t_{0,05}(69)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, el precio de las habitaciones de las casas rurales que se sitúan en los centros urbanos no es más caro.

d. Contraste de significatividad individual de la variable wifi.

$H_0 : \lambda_4 = \lambda_5 = 0$
H_a : alguna identidad no se cumple

$$F = \frac{SCR_R - SCR_{NR}}{SCR_R}\frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Como $F = 1,98942 < 3,12964 = \mathcal{F}_{0,05}(2,69)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto la variable explicativa “tener o no conexión a wifi” no es individualmente significativa.

e. Contraste de igualdad de coeficientes.

$H_0 : \lambda_4 = \lambda_5$
H_a : alguna identidad no se cumple
$$F = \frac{SCR_R - SCR_{NR}}{SCR_R}\frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Como $F = 1,04678 < 3,97981 = \mathcal{F}_{0,05}(1,69)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto existe evidencia a favor de que lo importante es poder acceder a internet y no si es gratuito o no.

Tercera parte

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

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

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

b. Contraste de significatividad de que la casa rural tenga wifi gratuito.

$H_0 : \beta_4 = 0$
$H_a : \beta_4 \neq 0$
$$t = \frac{\hat\beta_4 - 0}{\hat\sigma_{\hat\beta_4}} \overset{H_0}{\sim} t(T-k)$$

Como $\mid t \mid = 2,717 > 1,99547 = t_{0,025}(68)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable wifi es individualmente significativa.

c. Contraste de significatividad conjunta de las variables cercanía a un parque natural, cercanía a una playa y cercanía a un lago o embalse.

$H_0 : \beta_5 = \beta_6 = \beta_7 = 0$
H_a : alguna identidad no se cumple
$$F = \frac{SCR_R - SCR_{NR}}{SCR_R}\frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Como $F = 5,76258 > 2,7395 = \mathcal{F}_{0,05}(3,68)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto las variables cercanía a un parque natural, cercanía a una playa y cercanía a un lago o embalse son conjuntamente significativas.

d. Contraste de significatividad individual de las variables cercanía a un parque natural, cercanía a una playa y cercanía a un lago o embalse.

$$H_0 : \beta_i = 0 \qquad\qquad t = \frac{\hat{\beta}_i - 0}{\hat{\sigma}_{\hat{\beta}_i}} \overset{H_0}{\sim} t(T-k)$$
$$H_a : \beta_i \neq 0$$

$i = 5$: Como $\mid t \mid = 0,8476 < 1,99547 = t_{0,025}(68)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable cercanía a un parque natural no es individualmente significativa.

$i = 6$: Como $\mid t \mid = 3,613 > 1,99547 = t_{0,025}(68)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable cercanía a una playa es individualmente significativa.

$i = 7$: Como $\mid t \mid = 1,539 < 1,99547 = t_{0,025}(68)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable cercanía a un lago o embalse no es individualmente significativa.

e. Resultados de estimación del modelo aumentado.

Modelo 4: MCO, usando las observaciones 1–75
Variable dependiente: PR

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	33,6327	6,80200	4,9445	0,0000
HB	1,31946	1,12334	1,1746	0,2443
PD	1,66487	1,08630	1,5326	0,1301
WIFIG	9,14238	3,35367	2,7261	0,0082
PNR	3,35482	3,96390	0,8463	0,4004
PLR	16,3609	4,52546	3,6153	0,0006
LGR	11,6257	7,91546	1,4687	0,1466
LOCC	1,89981	4,55032	0,4175	0,6776

Media de la vble. dep.	56,13893	D.T. de la vble. dep.	14,98446
Suma de cuad. residuos	11520,47	D.T. de la regresión	13,11287
R^2	0,306644	R^2 corregido	0,234204
$F(7,67)$	4,233071	Valor p (de F)	0,000641
Log-verosimilitud	−295,2101	Criterio de Akaike	606,4202
Criterio de Schwarz	624,9601	Hannan–Quinn	613,8230

Contraste de significatividad de la variable localización.

$$H_0 : \beta_8 = 0 \qquad\qquad t = \frac{\hat{\beta}_8 - 0}{\hat{\sigma}_{\hat{\beta}_8}} \overset{H_0}{\sim} t(T-k)$$
$$H_a : \beta_8 \neq 0$$

Como $\mid t \mid = 0,4175 < 1,99601 = t_{0,025}(67)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable localización no es significativa, no hay diferencias entre el precio de las habitaciones de casas rurales situadas en el centro urbano y las alejadas.

f. La variable localización ha resultado ser no significativa por lo que no tiene ningún efecto sobre las propiedades del estimador MCO empleado en la estimación del modelo anterior. Mantienen las propiedades de linealidad en u, insesgadez y varianza mínima de entre los estimadores lineales e insesgados.

g. Contraste de significatividad conjunta de las variables número de habitaciones, precio de desayuno, cercanía a un parque natural, cercanía a un lago o embalse y localización.

$H_0 : \beta_2 = \beta_3 = \beta_5 = \beta_7 = \beta_8 = 0$
H_a : alguna identidad no se cumple

$$F = \frac{SCR_R - SCR_{NR}}{SCR_R} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Como $F = 1,52656 < 2,35166 = \mathcal{F}_{0,05}(5,67)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto las variables número de habitaciones, precio de desayuno, cercanía a un parque natural, cercanía a un lago o embalse y localización no son conjuntamente significativas.

El modelo propuesto es: $PR_i = \beta_1 + \beta_4\, WIFIG_i + \beta_6\, PLR_i + u_i$

Los resultados de estimación del modelo restringido:

Modelo 5: MCO, usando las observaciones 1–75
Variable dependiente: PR

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	47,9589	2,43513	19,6946	0,0000
WIFIG	11,2175	3,15271	3,5580	0,0007
PLR	16,0018	4,45227	3,5941	0,0006

Media de la vble. dep.	56,13893	D.T. de la vble. dep.	14,98446
Suma de cuad. residuos	12832,91	D.T. de la regresión	13,35046
R^2	0,227656	R^2 corregido	0,206202
$F(2,72)$	10,61133	Valor p (de F)	0,000091
Log-verosimilitud	−299,2559	Criterio de Akaike	604,5118
Criterio de Schwarz	611,4642	Hannan–Quinn	607,2878

El estimador empleado en este modelo es lineal en u, insesgado porque las restricciones impuestas son verídicas y tiene menor varianza que el empleado en el modelo inicial.

Ejercicio 6.3 Leche de soja

Primera parte

a. Modelo: $V_t = \beta_1 + \beta_2\, P_t + \beta_3\, G_t + \beta_4\, G_t^2 + u_t$

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

	Coeficiente	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

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

b. Contraste de significatividad conjunta de las variables explicativas.

$H_0 : \beta_2 = \beta_3 = \beta_4 = 0$
$H_a :$ alguna identidad no se cumple

$$F = \frac{R^2/(k-1)}{(1-R^2)/T-k} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Como $F = 148,1277 > 2,63854 = \mathcal{F}_{0,05}(3,266)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto las variables explicativas del modelo son conjuntamente significativas.

Contraste de significatividad de la variable precio.

$H_0 : \beta_2 = 0$
$H_a : \beta_2 \neq 0$

$$t = \frac{\hat{\beta}_2 - 0}{\hat{\sigma}_{\hat{\beta}_2}} \overset{H_0}{\sim} t(T-k)$$

Como $\mid t \mid = 17,20 > 1,96892 = t_{0,025}(266)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable precio es individualmente significativa.

Contraste de significatividad de la variable gasto en publicidad.

$H_0 : \beta_3 = \beta_4 = 0$
$H_a :$ alguna identidad no se cumple

$$F = \frac{SCR_R - SCR_{NR}}{SCR_R} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Como $F = 92,5594 > 3,02973 = \mathcal{F}_{0,05}(2,266)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto la variable gasto en publicidad es significativa.

c. Contraste bilateral.

$$H_0 : \beta_4 = 0 \qquad t = \frac{\hat{\beta}_4 - 0}{\hat{\sigma}_{\hat{\beta}_4}} \overset{H_0}{\sim} t(T-k)$$
$$H_a : \beta_4 \neq 0$$

Como $\mid t \mid = 7,262 < 1,96892 = t_{0,025}(266)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la relación entre las variables ventas y gasto en publicidad no es lineal sino cuadrática.

d. Contraste general.

$$H_0 : \beta_2 = -0,015 \qquad F = \frac{SCR_R - SCR_{NR}}{SCR_R}\frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$
$$H_a : \beta_2 \neq -0,015$$

Como $F = 289,593 > 3,87666 = \mathcal{F}_{0,05}(1,266)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, un aumento de 50 céntimos en el precio de la leche de soja no puede acarrear una disminución en las ventas de 750 envases.

e. Predicción por intervalo.

$IC(V_p)_{0,95} = [13,5156\,;\, 71,7971]$

Si el precio del envase se fijara en 75 céntimos y se realizara un gasto en publicidad de 20000 euros, se podría llegar entre 13,5156 y 71,7971 miles de envases. Por tanto vender 25 mil envases es viable.

Segunda parte

a. $V_t = \beta_1 + \beta_2 P_t + \beta_3 G_t + \beta_4 G_t^2 + \beta_5 time + \beta_6 D_{1t} + \beta_7\, D_{2t} + \beta_8\, D_{3t} + \ldots + \beta_{16} D_{11t} + u_t$

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

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−384,992	23,5086	−16,3766	0,0000
P	2,65667	0,126825	20,9475	0,0000
GP	2,16316	0,274338	7,8850	0,0000
sq_GP	−0,00628330	0,00106251	−5,9136	0,0000
time	0,496684	0,0149453	33,2334	0,0000
dm1	−0,353509	1,36885	−0,2583	0,7964
dm2	−0,891238	1,36890	−0,6511	0,5156
dm3	−0,872756	1,36878	−0,6376	0,5243
dm4	−1,21630	1,36889	−0,8885	0,3751
dm5	−1,63315	1,36942	−1,1926	0,2341
dm6	−2,12044	1,36981	−1,5480	0,1229
dm7	−1,53919	1,38388	−1,1122	0,2671
dm8	−1,90023	1,38350	−1,3735	0,1708
dm9	−1,51783	1,38333	−1,0972	0,2736
dm10	−1,08713	1,38346	−0,7858	0,4327
dm11	−0,827367	1,38321	−0,5982	0,5503

Media de la vble. dep.	120.8475	D.T. de la vble. dep.	16.87912
Suma de cuad. residuos	5345.186	D.T. de la regresión	4.587378
R^2	0.930255	R^2 corregido	0.926137
$F(15, 254)$	225.8569	Valor p (de F)	2.6e–137
Log-verosimilitud	−786.1599	Criterio de Akaike	1604.320
Criterio de Schwarz	1661.895	Hannan–Quinn	1627.439
$\hat{\rho}$	0.815407	Durbin–Watson	0.367882

b. Ventas estimadas para enero:

$$\hat{V}_{enero} = -385,345509 + 2,65667P_t + 2,16316G_t - 0,00628330G_t^2 + 0,496684time$$

Ventas estimadas para agosto:

$$\hat{V}_{enero} = -386,89223 + 2,65667P_t + 2,16316G_t - 0,00628330G_t^2 + 0,496684time - 1,90023$$

c. Contraste de significatividad de la variable tendencia.

$$\begin{array}{l} H_0 : \beta_5 = 0 \\ H_a : \beta_5 \neq 0 \end{array} \qquad t = \frac{\hat{\beta}_5 - 0}{\hat{\sigma}_{\hat{\beta}_5}} \overset{H_0}{\sim} t(T-k)$$

Como $\mid t \mid = 33,23 > 1,96935 = t_{0,025}(254)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la variable tendencia es individualmente significativa.

d. Contraste de significatividad de la variable estacionalidad.

$$\begin{array}{l} H_0 : \beta_6 = \ldots \beta_{16} = 0 \\ H_a : \text{alguna identidad no se cumple} \end{array} \qquad F = \frac{SCR_R - SCR_{NR}}{SCR_R} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Como $F = 0,410257 > 1,82647 = \mathcal{F}_{0,05}(11, 254)$, no se rechaza H_0 a un nivel de significatividad del 5 %. Por tanto la variable estacionalidad no es significativa.

e. Resultados de estimación del modelo aumentado.

Modelo 5: MCO, usando las observaciones 1990:01–2012:06 (T = 270)
Variable dependiente: V

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−250,520	30,6844	−8,1644	0,0000
P	−1,06176	0,605467	−1,7536	0,0807
GP	2,66103	0,267845	9,9350	0,0000
sq_GP	−0,00772453	0,00101699	−7,5954	0,0000
time	0,513666	0,0141954	36,1855	0,0000
dm1	−0,310565	1,27623	−0,2433	0,8079
dm2	−0,917931	1,27627	−0,7192	0,4727
dm3	−0,884234	1,27615	−0,6929	0,4890
dm4	−1,20676	1,27626	−0,9455	0,3453
dm5	−1,58079	1,27677	−1,2381	0,2168
dm6	−2,03810	1,27718	−1,5958	0,1118
dm7	−1,66481	1,29038	−1,2902	0,1982
dm8	−1,98003	1,28994	−1,5350	0,1260
dm9	−1,52685	1,28972	−1,1839	0,2376
dm10	−1,12275	1,28985	−0,8704	0,3849
dm11	−0,841019	1,28960	−0,6522	0,5149
sq_P	0,0191809	0,00306307	6,2620	0,0000

Media de la vble. dep.	120.8475	D.T. de la vble. dep.	16.87912
Suma de cuad. residuos	4627.906	D.T. de la regresión	4.276929
R^2	0.939614	R^2 corregido	0.935796
$F(16, 253)$	246.0466	Valor p (de F)	5.1e–144
Log-verosimilitud	−766.7075	Criterio de Akaike	1567.415
Criterio de Schwarz	1628.588	Hannan–Quinn	1591.980
$\hat{\rho}$	0.791304	Durbin–Watson	0.428458

Contraste bilateral.

$$\begin{array}{l} H_0 : \beta_{17} = 0 \\ H_a : \beta_{17} \neq 0 \end{array} \qquad t = \frac{\hat{\beta}_{17} - 0}{\hat{\sigma}_{\hat{\beta}_{17}}} \overset{H_0}{\sim} t(T-k)$$

Como $\mid t \mid= 6,2620 > 1,96938 = t_{0,025}(253)$, se rechaza H_0 a un nivel de significatividad del 5 %. Por lo tanto, la relación entre las ventas y el precio no es lineal, sino cuadrática.

f. Tras contrastar nuevamente la significatividad de la variable estacional en el modelo aumentado y obtener que no es significativa estimaría el siguiente modelo:

$$V_t = \beta_1 + \beta_2 P_t + \beta_3 G_t + \beta_4 G_t^2 + \beta_5 time + \beta_6 P_t^2 + u_t \quad t = 1990:1, \ldots, 2012:6$$

Los resultados de estimación son:

Modelo 6: MCO, usando las observaciones 1990:01–2012:06 ($T = 270$)
Variable dependiente: V

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−251.137	30.3113	−8.2853	0.0000
P	−1.06533	0.598435	−1.7802	0.0762
GP	2.66920	0.264313	10.0986	0.0000
sq_GP	−0.00776953	0.00100340	−7.7432	0.0000
time	0.512564	0.0140071	36.5931	0.0000
sq_P	0.0191490	0.00302805	6.3239	0.0000

Media de la vble. dep.	120.8475	D.T. de la vble. dep.	16.87912
Suma de cuad. residuos	4724.480	D.T. de la regresión	4.230338
R^2	0.938354	R^2 corregido	0.937187
$F(5, 264)$	803.7083	Valor p (de F)	1.9e–157
Log-verosimilitud	−769.4957	Criterio de Akaike	1550.991
Criterio de Schwarz	1572.582	Hannan–Quinn	1559.661
$\hat{\rho}$	0.789293	Durbin–Watson	0.431438