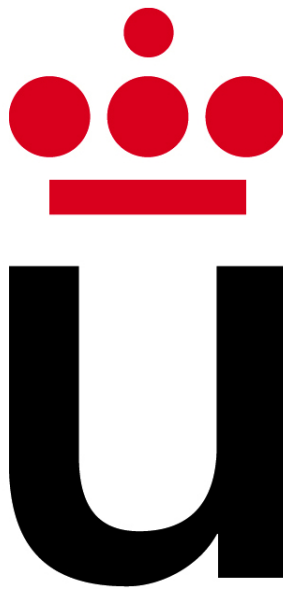


TRABAJO FINAL
Análisis de Datos Avanzados
ANÁLISIS DE SERIES TEMPORALES

José Ignacio Escribano



MÓSTOLES, 20 DE MARZO DE 2016

Índice

Índice de figuras	c
Índice de tablas	d
1. Introducción	1
2. Resolución de las series temporales	1
2.1. Índice de empleo de un determinado país	2
2.2. Venta de cigarros puros de una empresa tabacalera	12
3. Conclusiones	25
4. Código EViews	26
4.1. Índice de empleo	26
4.2. Venta de puros	26

Índice de figuras

1.	Serie temporal empleo	2
2.	Correlograma de la serie temporal empleo	3
3.	Serie transformada tomando una diferencia regular	3
4.	Correlograma de la serie temporal transformada dempleo	4
5.	Estimación del modelo ARIMA(1,1,0)	5
6.	Estimación del modelo ARIMA(1,1,2)	6
7.	Estimación del modelo ARIMA(0,1,2)	7
8.	Correlograma de los residuos del modelo ARIMA(1,1,0)	8
9.	Residuos del modelo ARIMA(1,1,0)	9
10.	Correlograma de los residuos del modelo ARIMA(0,1,2)	10
11.	Residuos del modelo ARIMA(0,1,2)	11
12.	Serie temporal puros	12
13.	Correlograma de la serie temporal puros	13
14.	Serie temporal puros	14
15.	Correlograma de la serie temporal diferenciada dpuros	15
16.	Serie temporal ddpuros12	16
17.	Correlograma de la serie temporal ddpuros12	17
18.	Estimación del modelo ARIMA(0, 1, 1) \times (0, 1, 1) ₁₂	18
19.	Estimación del modelo ARIMA(2, 1, 0) \times (0, 1, 1) ₁₂	19
20.	Residuos del modelo ARIMA(0, 1, 1) \times (0, 1, 1) ₁₂	20
21.	Correlograma de los residuos del modelo ARIMA(0, 1, 1) \times (0, 1, 1) ₁₂	21
22.	Residuos del modelo ARIMA(2, 1, 0) \times (0, 1, 1) ₁₂	22
23.	Correlograma de los residuos del modelo ARIMA(2, 1, 0) \times (0, 1, 1) ₁₂	23
24.	Predicciones (e intervalo de confianza) de ventas de puros para el año 1997	24
25.	Serie temporal de puros junto a las predicciones para el año 1997	24

Índice de tablas

1.	Principales características de la función de autocorrelación y de autocorrelación parcial de los principales modelos estacionarios	1
2.	Análisis de la estimación de los modelos ARIMA(1,1,0) y ARIMA(0,1,2)	7
3.	Predicciones del índice de empleo para el año 1994	10
4.	Análisis de la estimación de los modelos ARIMA(0, 1, 1) \times (0, 1, 1) ₁₂ y ARIMA(2, 1, 0) \times (0, 1, 1) ₁₂	19
5.	Predicciones de ventas de puros para el año 1997	22

1. Introducción

En este caso práctico utilizaremos la metodología Box-Jenkins para analizar dos series temporales. La primera es el índice de empleo de un determinado país, y la segunda es el volumen de ventas mensual de puros de una empresa tabacalera. En ambos casos, se trata de obtener un modelo que se ajuste lo máximo posible a la serie temporal.

La metodología Box-Jenkins recoge los pasos necesarios para obtener el modelo más adecuado de serie temporal:

1. Especificación inicial: consiste en determinar el orden de integración de la serie temporal y naturaleza de diferencias que se requerirán para convertir en estacionaria la serie temporal. En este paso se usa el análisis gráfico de la serie, además de los correlogramas simple y parcial de la serie. Una vez hecho lo anterior, habrá que decidir los órdenes de los polinomios autorregresivo y de medias móviles. De nuevo, se hará uso del correlograma simple y parcial de la serie. La Tabla 1 recoge las principales características de la función de autocorrelación y de autocorrelación parcial de los principales modelos estacionarios.

Tabla 1: Principales características de la función de autocorrelación y de autocorrelación parcial de los principales modelos estacionarios

Modelo	Función de autocorrelación	Función de autocorrelación parcial
AR(p)	Decrecimiento rápido hacia cero, sin llegar a anularse	p primera autocorrelaciones distintas de cero, y el resto cero
MA(q)	q primeras autocorrelaciones significativas, y el resto cero	Decrecimiento rápido hacia cero, sin llegar a anularse
ARMA(p,q)	Decrecimiento rápido hacia cero, sin llegar a anularse	Decrecimiento rápido hacia cero, sin llegar a anularse

2. Estimación: en este paso, se procede a estimar los modelos propuestos, normalmente mediante máxima verosimilitud o mínimos cuadrados no lineales.
3. Chequeo o validación: en este paso, se validan los posibles modelos y se escoge el que parezca más adecuado para describir la serie temporal.
4. Utilización del modelo: el modelo escogido se puede utilizar para predecir futuros valores de la serie, entre otras opciones.

2. Resolución de las series temporales

A continuación, aplicamos la metodología Box-Jenkins para obtener un modelo que se adecue a cada una de las series temporales planteadas.

2.1. Índice de empleo de un determinado país

La primera serie temporal es el índice de empleo de un determinado país. La serie está corregida de estacionalidad y tiene frecuencia trimestral. El período muestral abarca desde el primer trimestre del año 1962 hasta el cuarto trimestre del año 1994.

Comenzamos representando la serie temporal (Figura 1).

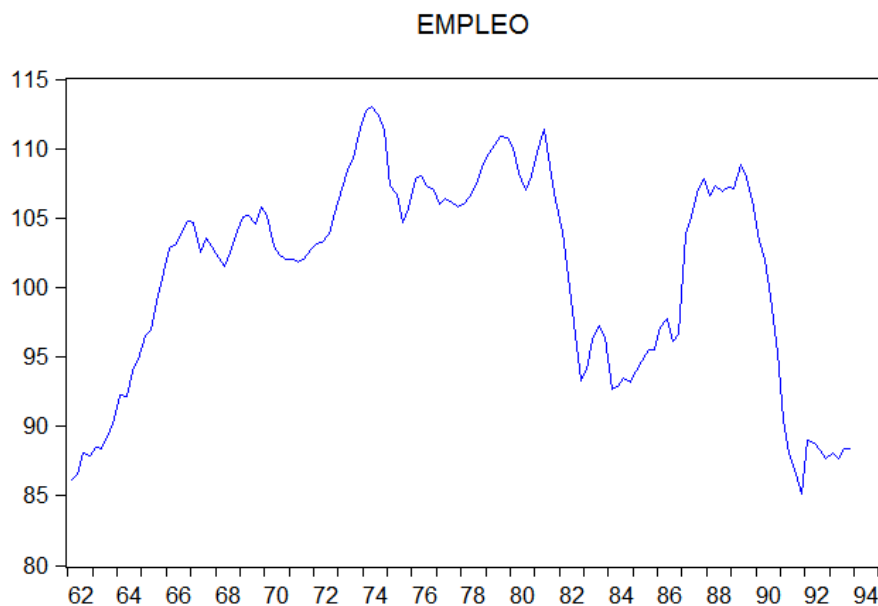


Figura 1: Serie temporal empleo

Se observa que podría haber tendencia en la serie original. Para verificarlo, usamos el correlograma de la serie que se puede ver en la Figura 2.

Como se observa un decrecimiento lento en la parte positiva del eje X, estamos ante una serie que presenta tendencia, por lo que estamos ante una serie no estacionaria. Tenemos que eliminar la tendencia, haciendo uso de las diferencias regulares de la serie original.

Tomamos la primera diferenciación para convertir la serie en estacionaria, que llamamos dempleo. Representamos la nueva serie para comprobar que hemos eliminado la tendencia de la serie original.

La nueva serie parece indicar que estamos ante una serie estacionaria, ya que hemos eliminado la tendencia tomando una diferencia regular, y la serie carecía de estacionalidad de acuerdo al enunciado. Esto se puede comprobar mirando el correlograma de esta nueva serie (Figura 4).

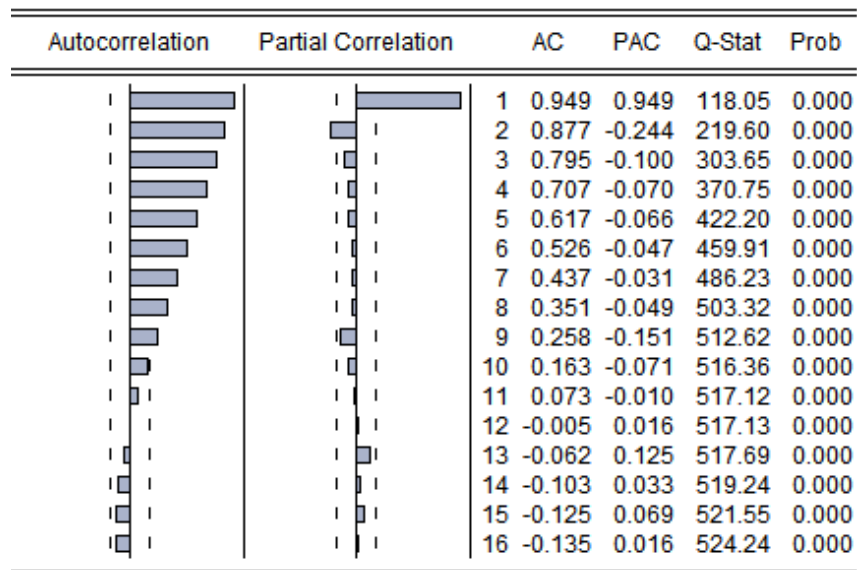


Figura 2: Correlograma de la serie temporal empleo

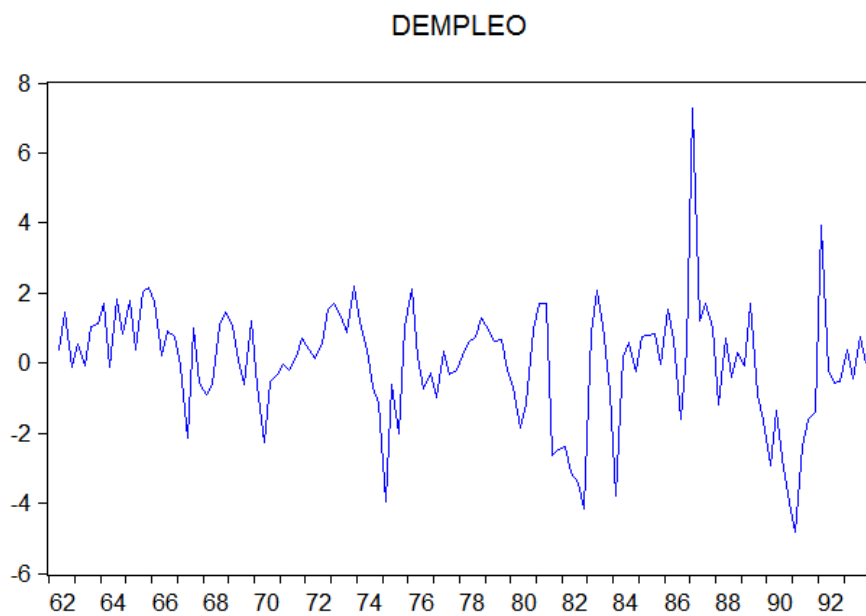


Figura 3: Serie transformada tomando una diferencia regular

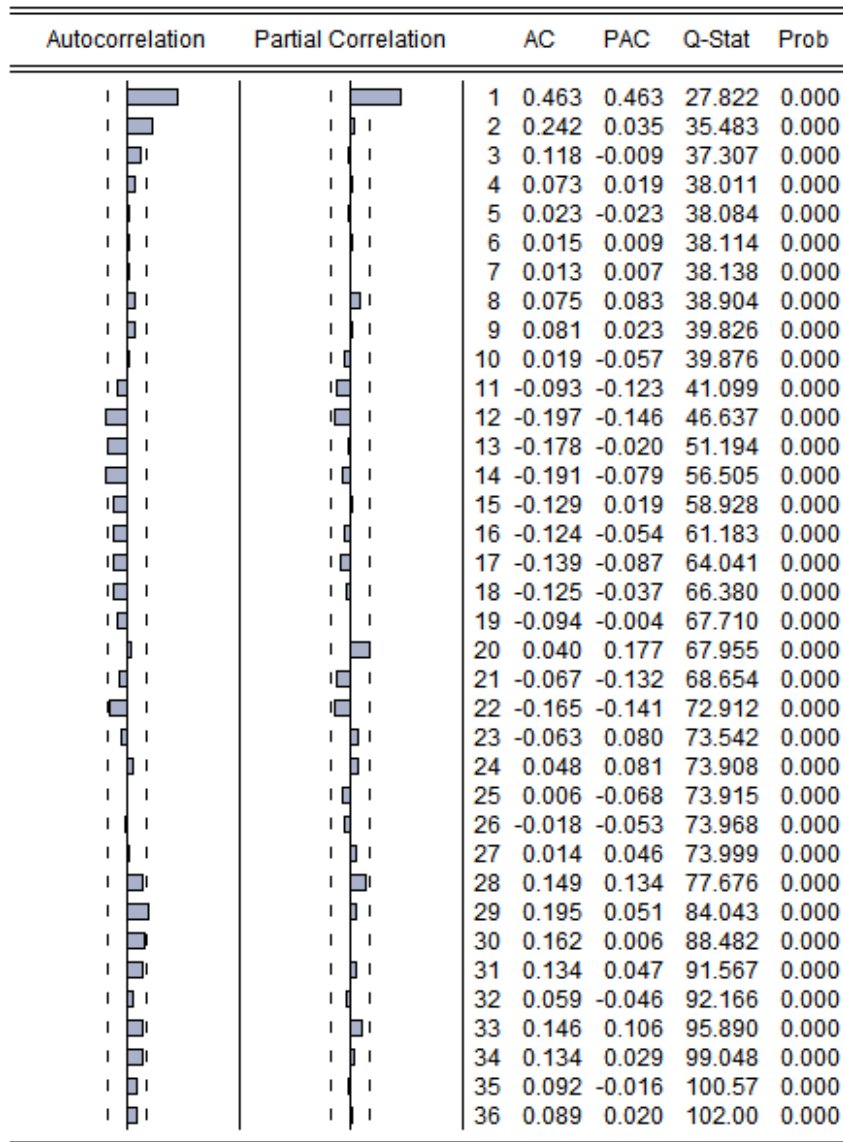


Figura 4: Correlograma de la serie temporal transformada dempleo

Así pues, tenemos que $d = 1$ y $D = 0$, y

$$\text{dempleo} = (1 - B)\text{empleo}$$

Es decir, dempleo es un modelo integrado de orden 1.

Observando el correlograma de la Figura 4 podemos sugerir que la serie temporal puede venir dada por un modelo AR(1), ya que la función de autocorrelación decrece rápidamente hacia cero, sin llegar a anularse, y en la función de autocorrelación parcial,

hay un valor no nulo positivo, y el resto es cero. También podría tratarse de un modelo MA(2), ya que en la función de autocorrelación hay dos valores no nulos positivos y el resto es cero, y en la función de autocorrelación parcial hay un decrecimiento rápido sin llegar a anularse. Así planteamos los siguientes modelos:

1. ARIMA(1,1,0) , que escrito en forma de ecuación es

$$(1 - \theta_1 B)(1 - B)\text{empleo} = a_t$$

2. ARIMA(0,1,2) , que escrito en forma de ecuación es

$$(1 - \theta_1 B)(1 - B)\text{empleo} = (1 - \theta_1 B - \theta_2 B^2)a_t$$

3. ARIMA(1,1,2) , que escrito en forma de ecuación es

$$(1 - B)\text{empleo} = (1 - \theta_1 B - \theta_2 B^2)a_t$$

Estimamos cada uno de los modelos propuestos. En primer lugar estimamos el modelo ARIMA(1,1,0). La salida de EViews de este modelo se muestra en la Figura 5.

Dependent Variable: D(EMPLEO,1)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/14/16 Time: 12:10
Sample: 1962Q2 1993Q4
Included observations: 127
Convergence achieved after 12 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.459214	0.064247	7.147611	0.0000
SIGMASQ	2.139284	0.157305	13.59957	0.0000
R-squared	0.214012	Mean dependent var		0.017041
Adjusted R-squared	0.207724	S.D. dependent var		1.656314
S.E. of regression	1.474284	Akaike info criterion		3.631709
Sum squared resid	271.6891	Schwarz criterion		3.676500
Log likelihood	-228.6135	Hannan-Quinn criter.		3.649907
Durbin-Watson stat	2.024943			
Inverted AR Roots	.46			

Figura 5: Estimación del modelo ARIMA(1,1,0)

Se observa que todos los parámetros del modelo son significativos, por lo que lo consideraremos adecuado. Además, se obtiene un valor de R^2 ajustado de 0.214.

Estimamos nuestro segundo modelo, es decir, el modelo ARIMA(0,1,2) usando EViews. La salida que produce este programa se puede ver en la Figura 6.

Dependent Variable: D(EMPLEO,1)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/14/16 Time: 12:28
Sample: 1962Q2 1993Q4
Included observations: 127
Convergence achieved after 12 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	0.434809	0.075247	5.778406	0.0000
MA(2)	0.194397	0.093151	2.086909	0.0389
SIGMASQ	2.157131	0.187482	11.50579	0.0000
R-squared	0.207454	Mean dependent var		0.017041
Adjusted R-squared	0.194671	S.D. dependent var		1.656314
S.E. of regression	1.486378	Akaike info criterion		3.655626
Sum squared resid	273.9556	Schwarz criterion		3.722812
Log likelihood	-229.1323	Hannan-Quinn criter.		3.682923
Durbin-Watson stat	1.962206			
Inverted MA Roots	-.22-.38i	-.22+.38i		

Figura 6: Estimación del modelo ARIMA(1,1,2)

Observando la salida de EViews, tenemos que todos los parámetros del modelo son significativos, por lo que consideraremos este modelo adecuado para representar la serie de tiempo del índice de empleo. Además, tiene un coeficiente R^2 ajustado de 0.195.

Por último, estimamos nuestro último modelo propuesto: el ARIMA(1,1,2). De nuevo, usamos EViews que nos devuelve la salida de la Figura 7.

Observando la salida de este modelo, vemos que todos los parámetros no son significativos, por lo que desechamos este modelo ya que no parece adecuado para describir la serie temporal que estamos tratando.

Sólo tenemos dos modelos que validar: el modelo ARIMA(1,1,0) y el ARIMA(0,1,2). La Tabla 2 muestra una comparativa entre la estimación de los dos modelos.

El modelo ARIMA(1,1,0) tiene un menor error (1.474) que el modelo ARIMA(0,1,2) (1.487). Además tanto los estadísticos de Akaike como de Schwarz son menores en el modelo ARIMA(1,1,0) que en el modelo ARIMA(0,1,2). Lo anterior parece indicar que el modelo más adecuado es el ARIMA(1,1,0).

Dependent Variable: D(EMPLEO,1)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/14/16 Time: 12:36
Sample: 1962Q2 1993Q4
Included observations: 127
Convergence achieved after 12 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.489035	0.327248	1.494388	0.1376
MA(1)	-0.045889	0.344651	-0.133146	0.8943
MA(2)	0.017759	0.192464	0.092271	0.9266
SIGMASQ	2.136406	0.185092	11.54242	0.0000
R-squared	0.215069	Mean dependent var		0.017041
Adjusted R-squared	0.195924	S.D. dependent var		1.656314
S.E. of regression	1.485221	Akaike info criterion		3.661878
Sum squared resid	271.3235	Schwarz criterion		3.751459
Log likelihood	-228.5293	Hannan-Quinn criter.		3.698274
Durbin-Watson stat	1.992379			
Inverted AR Roots	.49			
Inverted MA Roots	.02+.13i	.02-.13i		

Figura 7: Estimación del modelo ARIMA(0,1,2)

Tabla 2: Análisis de la estimación de los modelos ARIMA(1,1,0) y ARIMA(0,1,2)

	ARIMA(1,1,0)	ARIMA(0,1,2)
R^2	0.214	0.207
R^2 ajustado	0.208	0.195
Akaike Info Criterion	3.631	3.656
Schwarz Criterion	3.677	3.723
Error de regresión	1.474	1.487

Para confirmar nuestras sospechas, hacemos un análisis de los residuos de ambos modelos.

Comenzamos con el modelo ARIMA(1,1,0). En la Figura 8 se puede ver el correlograma de los residuos del modelo.

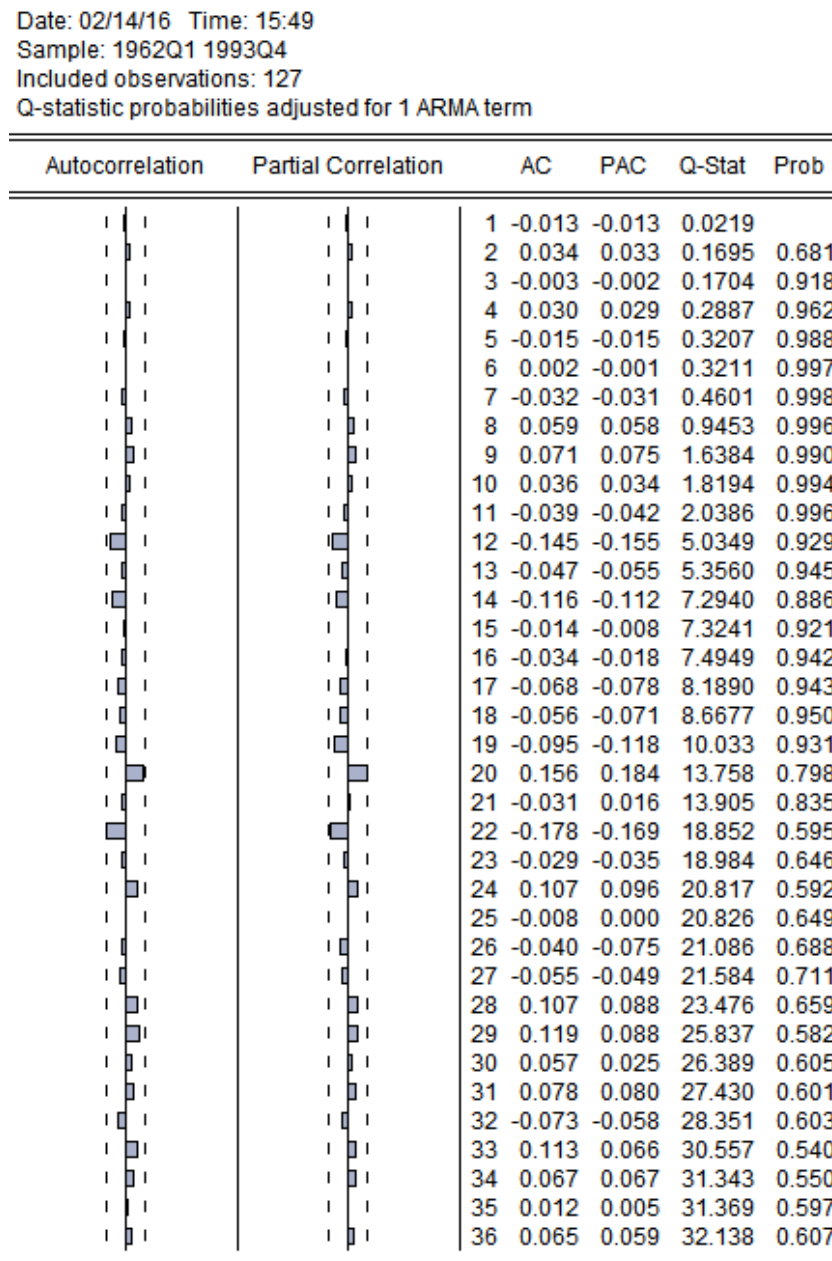


Figura 8: Correlograma de los residuos del modelo ARIMA(1,1,0)

Se puede observar que las autocorrelaciones de los residuos no son significativas y

entran dentro de las bandas de confianza, lo que indica que no son distintas de cero. De la misma forma, el estadístico Q no muestra indicios de autocorrelación de los residuos, por lo que todo parece indicar que estamos ante ruido blanco. Para comprobarlo, representamos el gráfico de residuos, que se puede ver en la Figura 9.

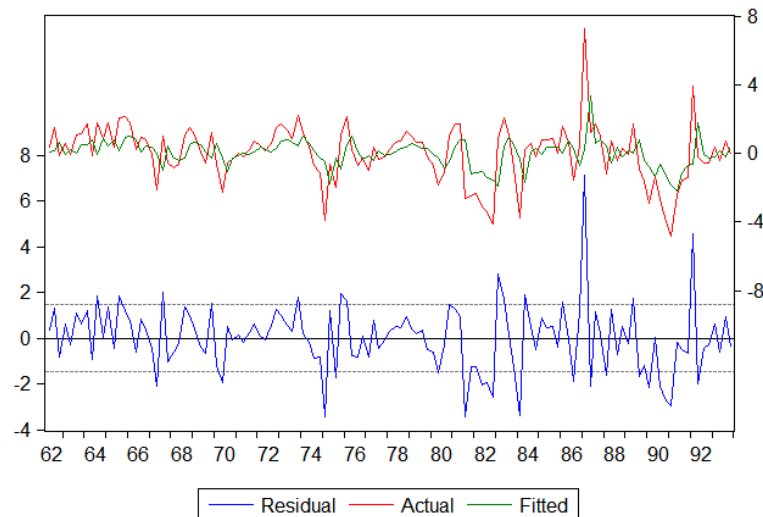


Figura 9: Residuos del modelo ARIMA(1,1,0)

La mayoría de los residuos se encuentran dentro de las bandas de confianza, lo que apoya la teoría de autocorrelación. Todo parece confirmar que los residuos se comportan como ruido blanco, es decir, tienen media 0 y varianza constante.

Notar la presencia de un outlier en el primer trimestre de 1987.

Procedemos de forma similar para comprobar que los residuos del modelo son ruido blanco.

Las Figuras 10 y 11 muestran la presencia de los residuos como ruido blanco.

De todo lo anterior, se deduce que el modelo ARIMA(1,1,0) es superior al modelo ARIMA(0,1,2) ya que tiene mejor coeficiente R^2 , menor error en la estimación, y menores valores en los estadísticos de Akaike y Schwarz.

Así pues, el índice de empleo viene dado por el modelo ARIMA(1,1,0).

Utilizaremos este modelo para predecir los valores de empleo del año siguiente (1994). Los datos para cada uno de los trimestres se pueden ver en la Tabla 3.

Date: 02/14/16 Time: 15:54
Sample: 1962Q1 1993Q4
Included observations: 127
Q-statistic probabilities adjusted for 2 ARMA terms

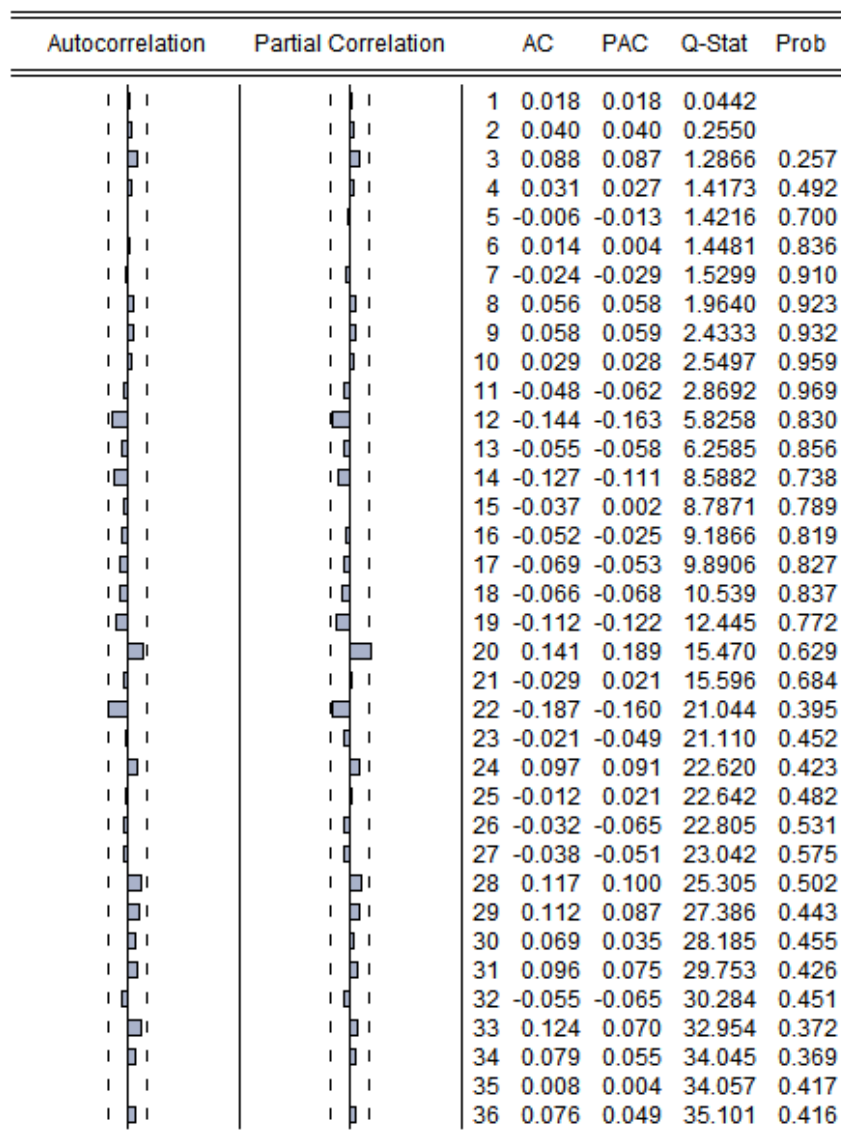


Figura 10: Correlograma de los residuos del modelo ARIMA(0,1,2)

Tabla 3: Predicciones del índice de empleo para el año 1994

Trimestre	Predicción
1/1994	88.34407
2/1994	88.33593
3/1994	88.33219
4/1994	88.33048

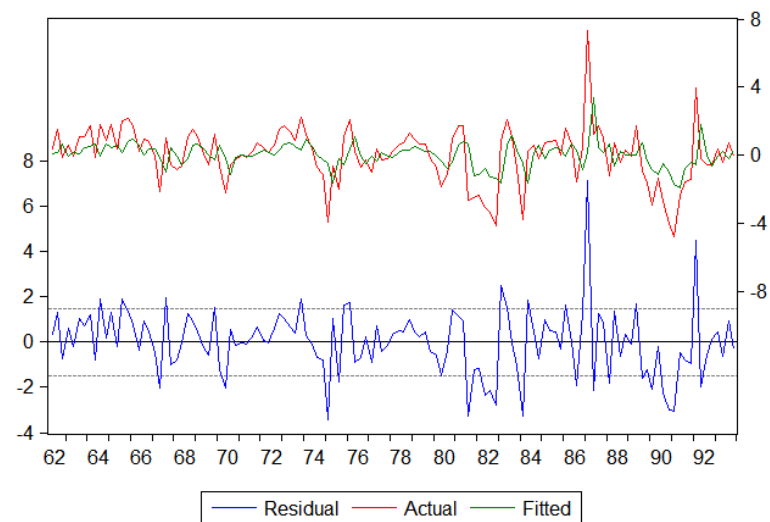


Figura 11: Residuos del modelo ARIMA(0,1,2)

Según estas predicciones, el índice de empleo del 1994 se mantuvo constante, en torno al 88 %.

2.2. Venta de cigarros puros de una empresa tabacalera

La segunda serie temporal es el volumen de ventas mensual de puros de una empresa tabacalera. El período de la serie abarca desde enero de 1989 hasta diciembre de 1996.

Comenzamos representando la serie temporal (Figura 1).

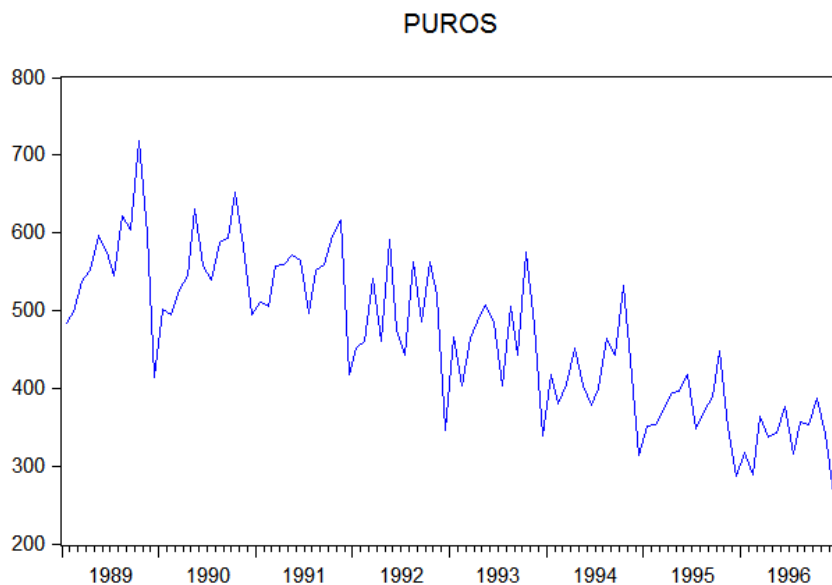


Figura 12: Serie temporal puros

Se puede ver una fuerte tendencia decreciente a lo largo de la serie temporal. Además, se observa una estacionalidad de los datos: la venta aumenta entre los meses de enero y septiembre, y desciende en los meses de octubre a diciembre.

Nos aseguramos de la presencia de tendencia y estacionalidad mirando el correlograma de la serie (Figura 13).

Se observa el decaimiento tanto en la parte regular como en la parte estacional, es decir, en los retardos múltiplos de 12, por lo que necesitamos eliminar la tendencia y la estacionalidad para que nuestra serie sea estacionaria.

Comenzamos eliminando la tendencia. Para ello, tomamos una diferencia regular de la serie, que llamaremos $dpuros$. La Figura 14 muestra nuestra nueva serie con una diferencia regular.

Date: 02/14/16 Time: 12:54
Sample: 1989M01 1996M12
Included observations: 96

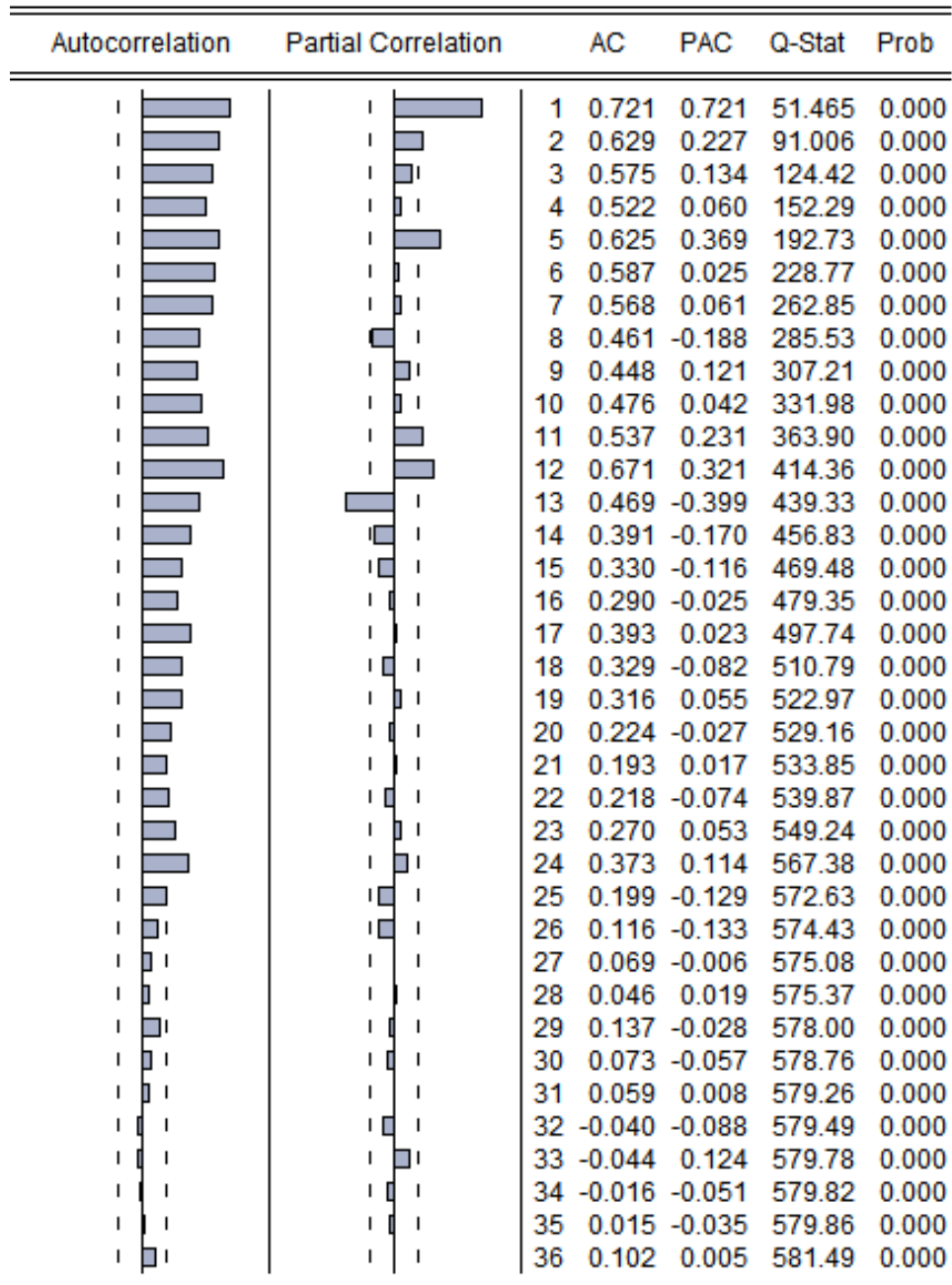


Figura 13: Correlograma de la serie temporal puros

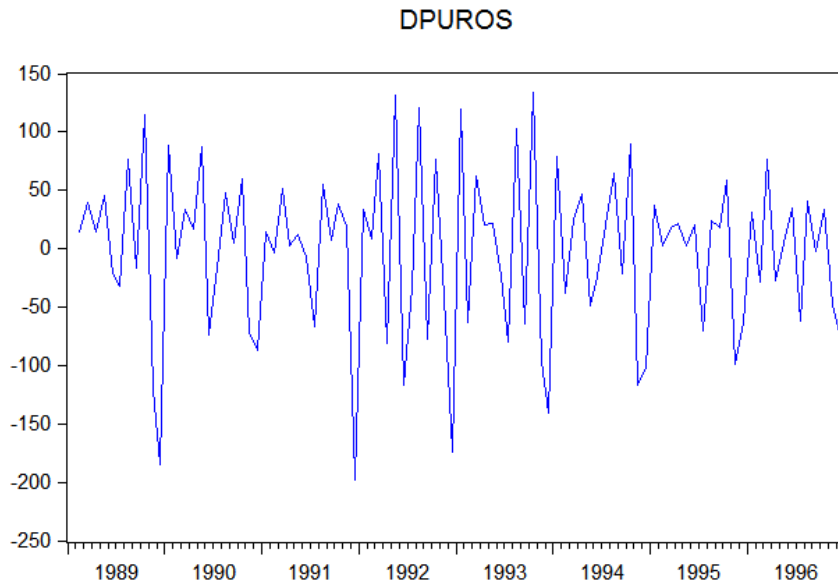


Figura 14: Serie temporal puros

Observamos que se ha eliminado la tendencia, pero no así la estacionalidad. Para asegurarnos observamos de nuevo el correlograma de esta nueva serie temporal (Figura 15).

En el correlograma se puede observar que se ha corregido la tendencia, pero no la estacionalidad en los retardos múltiplos de 12. Así que tomamos una diferencia estacional de período estacional 12. A esta nueva variable la llamamos *ddpuros12*, que se muestra en la Figura 16.

Se puede ver que esta nueva serie tiene tanto ausencia de tendencia como de estacionalidad. Para asegurarnos, vemos el correlograma de esta serie (Figura 17).

El correlograma no deja duda de que esta nueva serie es estacionaria.

Por tanto, para convertir la serie en estacionaria hemos tenido que aplicar la siguiente transformación:

$$\text{ddpuros12} = (1 - B)(1 - B^{12}) \cdot \text{puros}$$

Es decir, tenemos que $d = D = 1$.

Debemos establecer el modelo generador de la serie.

El correlograma de la Figura 17 sugiere que el modelo puede estar generado por un

Date: 02/14/16 Time: 12:57
Sample: 1989M01 1996M12
Included observations: 95

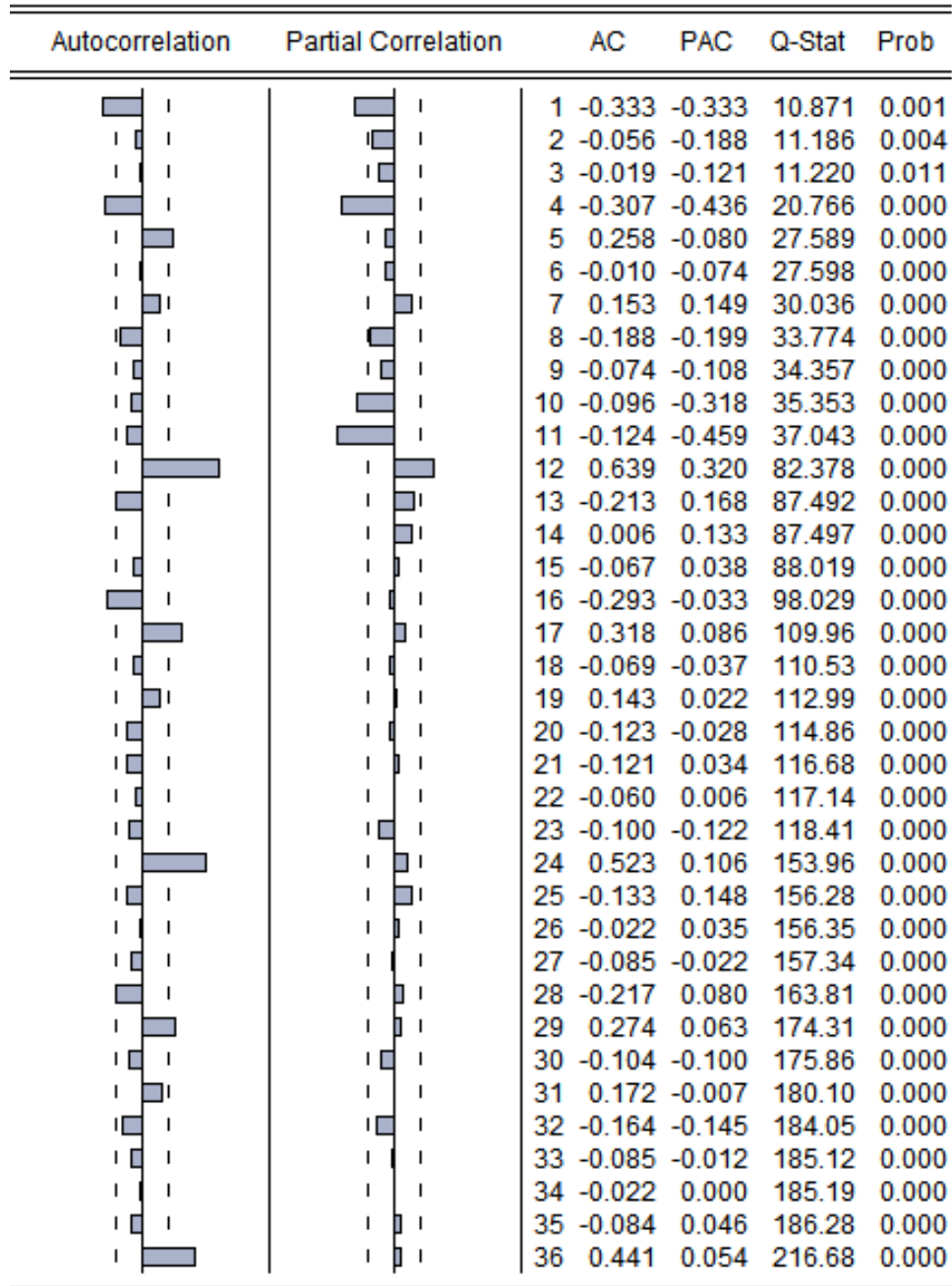


Figura 15: Correlograma de la serie temporal diferenciada dpuros

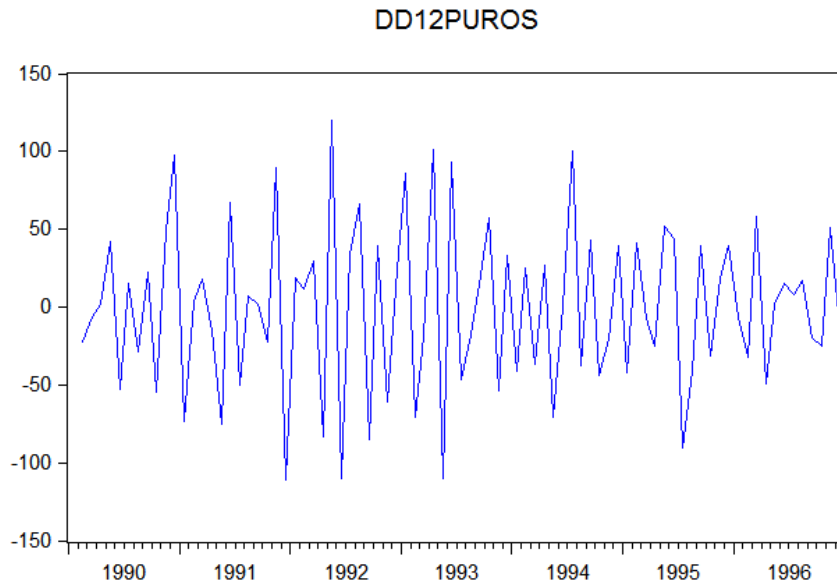


Figura 16: Serie temporal ddpuros12

MA(1) estacional ya que la función de autocorrelación presenta un valor no nulo en parte negativa del eje X en los múltiplos de 12 retardos, y la función de autocorrelación parcial muestra un decrecimiento lento en la parte negativa del eje X.

En la parte regular, el modelo puede estar dado por un AR(2), ya que la función de autocorrelación muestra un decrecimiento lento a lo largo de todos los retardos, y la función de autocorrelación parcial, dos valores no nulos negativos en la parte del eje X. También puede ser generado por un MA(1) ya que la función de autocorrelación presenta un valor no nulo negativo en el eje X, y la función de autocorrelación parcial presenta un decrecimiento lento en todos los retardos.

Así, los modelos candidatos a generar la serie son:

1. ARIMA(0, 1, 1) \times (0, 1, 1)₁₂, que escrito en forma de ecuación es

$$(1 - B)(1 - B^{12})\text{puros} = (1 - \theta_1 B)(1 - \Theta_1 B^{12})a_t$$

2. ARIMA(2, 1, 0) \times (0, 1, 1)₁₂, que escrito en forma de ecuación es

$$(1 - \phi_1 B - \phi_2 B^2)(1 - B)(1 - B^{12})\text{puros} = (1 - \theta_1 B^{12})a_t$$

Necesitamos estimar ambos modelos y validarlos.

Date: 02/14/16 Time: 13:02
Sample: 1989M01 1996M12
Included observations: 83

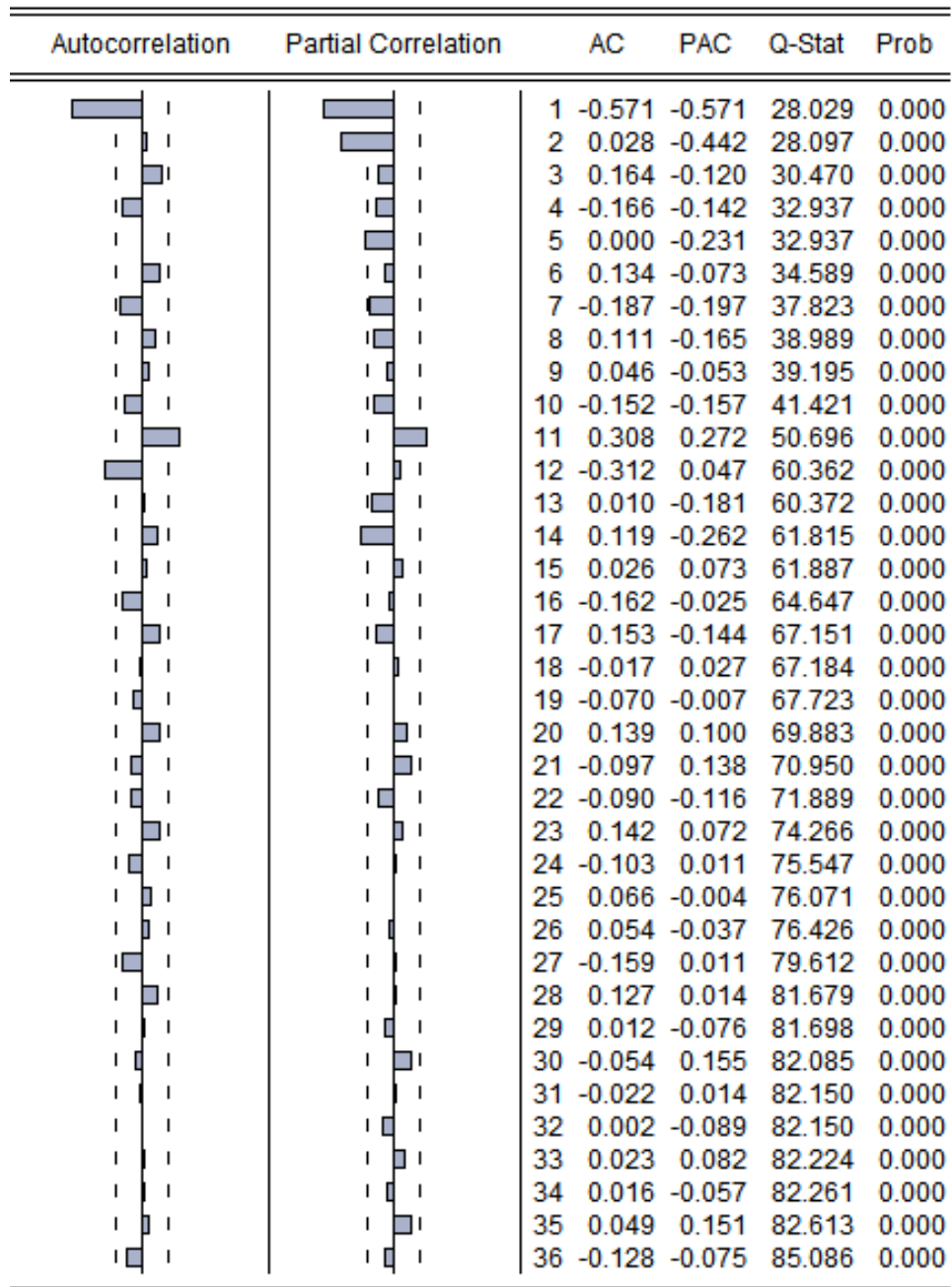


Figura 17: Correlograma de la serie temporal ddpuros12

Comenzamos con el modelo $ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$. En la Figura 18 se muestra el resultado de la estimación de este modelo en EViews.

Dependent Variable: D(PUROS,1,12)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/19/16 Time: 21:32
Sample: 1990M02 1996M12
Included observations: 83
Convergence achieved after 21 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	-0.842796	0.065574	-12.85262	0.0000
SMA(12)	-0.616497	0.156795	-3.931867	0.0002
SIGMASQ	1045.706	178.3908	5.861882	0.0000
R-squared	0.620837	Mean dependent var	-0.481928	
Adjusted R-squared	0.611358	S.D. dependent var	52.83531	
S.E. of regression	32.93812	Akaike info criterion	9.948582	
Sum squared resid	86793.61	Schwarz criterion	10.03601	
Log likelihood	-409.8661	Hannan-Quinn criter.	9.983705	
Durbin-Watson stat	2.276776			
Inverted MA Roots	.96	.84	.83+.48i	.83-.48i
	.48+.83i	.48-.83i	.00+.96i	-.00-.96i
	-.48+.83i	-.48-.83i	-.83-.48i	-.83+.48i
	-.96			

Figura 18: Estimación del modelo $ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$

Se puede observar que todos los términos son significativos, y que tiene un coeficiente R^2 ajustado de 0.620.

Si ahora estimamos el modelo 2, es decir, el modelo $ARIMA(2, 1, 0) \times (0, 1, 1)_{12}$, se tiene el resultado de la Figura 19.

Se puede observar que todos los coeficientes de este modelo son significativos, y que tiene un coeficiente R^2 ajustado de 0.623.

A continuación, validaremos el modelo comparando tanto los resultados de la estimación como los residuos de ambos modelos. En la Tabla 4 se puede ver una comparativa entre ambos modelos.

Method: ARMA Maximum Likelihood (OPG - BHHH)

Date: 02/19/16 Time: 21:41

Sample: 1990M02 1996M12

Included observations: 83

Convergence achieved after 22 iterations

Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.862410	0.107217	-8.043558	0.0000
AR(2)	-0.493444	0.111440	-4.427903	0.0000
MA(12)	-0.703127	0.148611	-4.731318	0.0000
SIGMASQ	1000.483	177.9938	5.620885	0.0000
R-squared	0.637235	Mean dependent var		-0.481928
Adjusted R-squared	0.623459	S.D. dependent var		52.83531
S.E. of regression	32.42129	Akaike info criterion		9.952185
Sum squared resid	83040.08	Schwarz criterion		10.06876
Log likelihood	-409.0157	Hannan-Quinn criter.		9.999017
Durbin-Watson stat	2.025038			
Inverted AR Roots	-.43+.55i	-.43-.55i		
Inverted MA Roots	.97	.84+.49i	.84-.49i	.49+.84i
	.49-.84i	.00+.97i	-.00-.97i	-.49-.84i
	-.49+.84i	-.84-.49i	-.84+.49i	-.97

Figura 19: Estimación del modelo $ARIMA(2, 1, 0) \times (0, 1, 1)_{12}$

Tabla 4: Análisis de la estimación de los modelos $ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$ y $ARIMA(2, 1, 0) \times (0, 1, 1)_{12}$

	$ARIMA(2, 1, 0) \times (0, 1, 1)_{12}$	$ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$
R^2	0.620	0.637
R^2 ajustado	0.611	0.623
Akaike Info Criterion	9.949	9.952
Schwarz Criterion	10.036	10.068
Error de regresión	32.938	32.421

Se puede observar que el valor de los coeficientes R^2 y R^2 ajustado son mayores en el modelo $\text{ARIMA}(2, 1, 0) \times (0, 1, 1)_{12}$ que en el modelo $\text{ARIMA}(2, 1, 0) \times (0, 1, 1)_{12}$. Además, también es menor tanto el Akaike Info Criterion como el Schwarz Criterion en el modelo $\text{ARIMA}(2, 1, 0) \times (0, 1, 1)_{12}$. Sin embargo, el modelo $\text{ARIMA}(2, 1, 0) \times (0, 1, 1)_{12}$ tiene un menor error de regresión que el modelo $\text{ARIMA}(0, 1, 1) \times (0, 1, 1)_{12}$.

A continuación, pasamos a analizar los residuos de los modelos. Comenzamos con el modelo $\text{ARIMA}(0, 1, 1) \times (0, 1, 1)_{12}$. Las Figuras 20 y 21 muestra los residuos y el correlograma de estos. Se puede observar que la mayoría caen dentro de las bandas de confianza, por lo que es de suponer que los residuos se comportan como ruido blanco. El correlograma parece también contrastarlo.

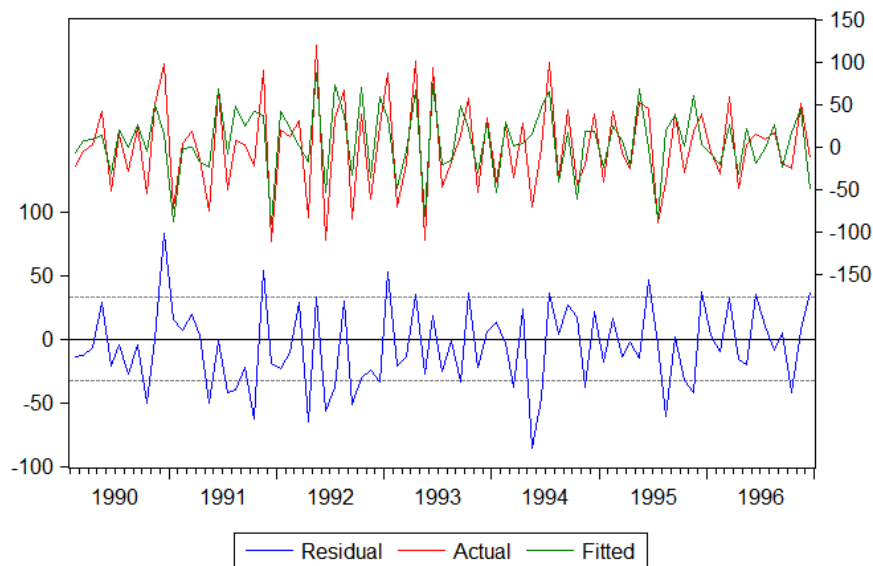


Figura 20: Residuos del modelo $\text{ARIMA}(0, 1, 1) \times (0, 1, 1)_{12}$

De forma similar, comprobamos que los residuos del modelo $\text{ARIMA}(2, 1, 0) \times (0, 1, 1)_{12}$ (Figuras 22 y 23).

A tenor de los resultados, estamos ante dos modelos muy parecidos, pero nos decantaremos por el modelo $\text{ARIMA}(0, 1, 1) \times (0, 1, 1)_{12}$, ya que es el que menor valor tiene en los Akaike Info Criterion y Schwarz Criterion, aunque tiene un mayor error de regresión (cercano a 0.5).

Por tanto, la venta de puros viene dada por el modelo $\text{ARIMA}(2, 1, 0) \times (0, 1, 1)_{12}$.

Usaremos este modelo para predecir las ventas de puros en el año 1997. Estos datos se pueden ver en la Tabla 5.

Date: 02/19/16 Time: 21:38
Sample: 1989M01 1996M12
Included observations: 83
Q-statistic probabilities adjusted for 2 ARMA terms

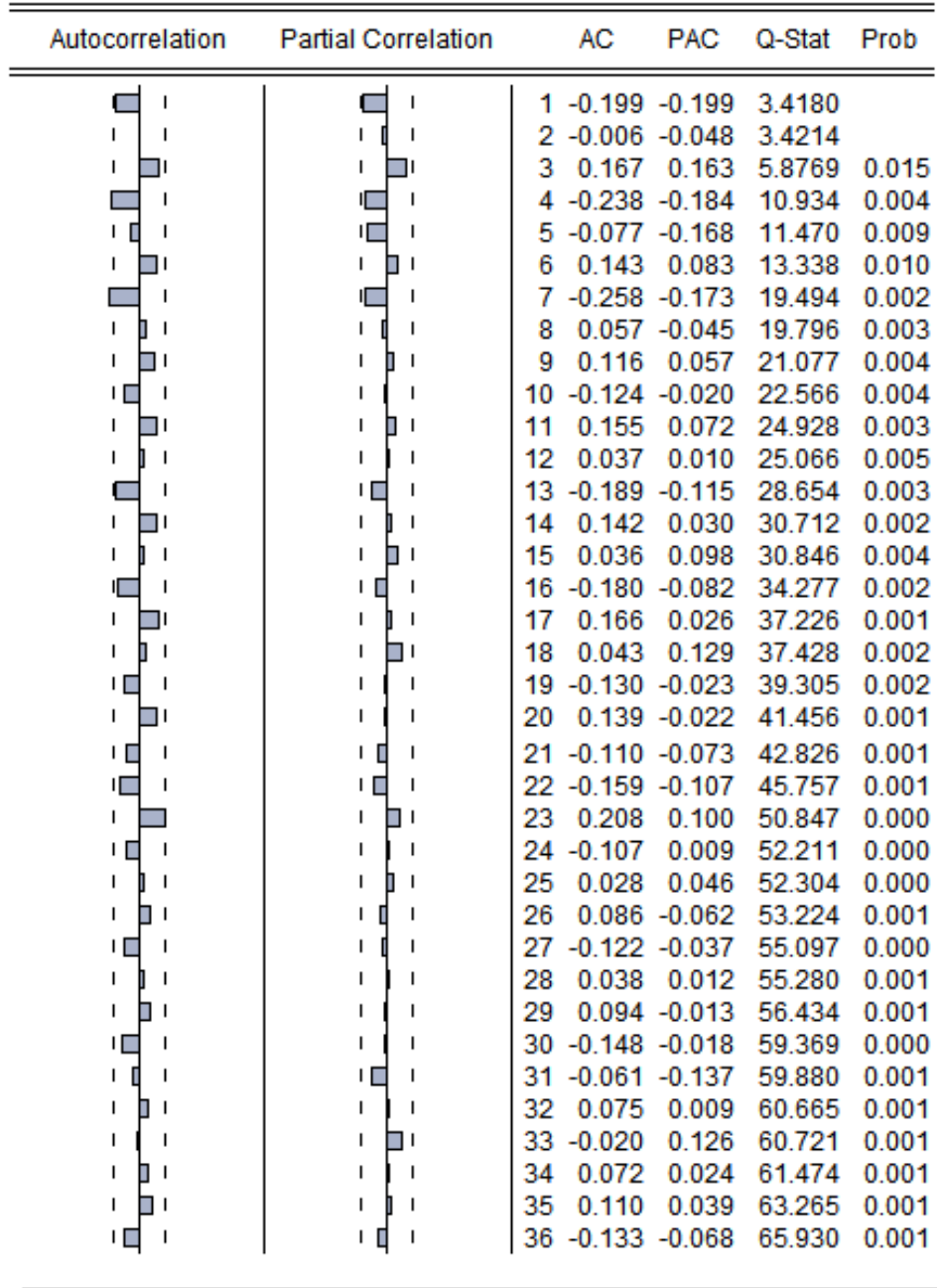


Figura 21: Correlograma de los residuos del modelo $ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$

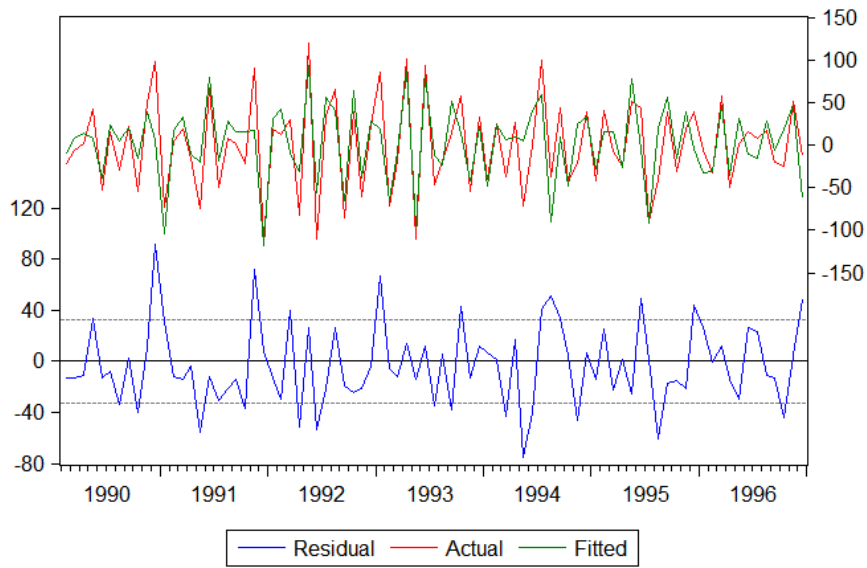


Figura 22: Residuos del modelo $ARIMA(2, 1, 0) \times (0, 1, 1)_{12}$

Tabla 5: Predicciones de ventas de puros para el año 1997

Mes	Predicción
Enero	278.693
Febrero	252.391
Marzo	328.391
Abril	301.391
Mayo	306.391
Junio	341.391
Julio	279.391
Agosto	320.391
Septiembre	318.391
Octubre	352.391
Noviembre	304.391
Diciembre	228.391

De forma más gráfica, se puede ver en la Figura 24. Si comparamos la serie temporal inicial junto con las predicciones, observamos la clara tendencia descendente que se mantiene en el año 1997 (Figura 25).

Date: 02/19/16 Time: 21:44
Sample: 1989M01 1996M12
Included observations: 83
Q-statistic probabilities adjusted for 3 ARMA terms









































































Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.041	-0.041	0.1427	
		2 -0.073	-0.075	0.6060	
		3 -0.196	-0.204	4.0042	
		4 -0.118	-0.152	5.2551	0.022
		5 -0.103	-0.166	6.2220	0.045
		6 0.077	-0.016	6.7673	0.080
		7 -0.163	-0.267	9.2409	0.055
		8 0.017	-0.119	9.2670	0.099
		9 0.067	-0.037	9.6956	0.138
		10 0.023	-0.115	9.7451	0.203
		11 0.115	0.033	11.034	0.200
		12 0.019	-0.041	11.069	0.271
		13 -0.131	-0.136	12.793	0.235
		14 0.070	0.051	13.289	0.275
		15 0.015	0.006	13.312	0.347
		16 -0.107	-0.122	14.527	0.338
		17 0.102	0.088	15.640	0.336
		18 0.083	0.138	16.382	0.357
		19 -0.061	-0.020	16.793	0.399
		20 0.025	0.035	16.865	0.464
		21 -0.117	-0.040	18.412	0.429
		22 -0.127	-0.109	20.264	0.379
		23 0.149	0.104	22.859	0.296
		24 0.025	0.025	22.933	0.348
		25 0.047	0.041	23.207	0.390
		26 -0.001	-0.020	23.208	0.449
		27 -0.062	-0.005	23.698	0.479
		28 0.049	0.073	24.003	0.519
		29 0.078	0.003	24.790	0.531
		30 -0.128	-0.041	26.979	0.465
		31 -0.115	-0.111	28.780	0.424
		32 0.042	0.009	29.023	0.464
		33 0.052	0.075	29.399	0.497
		34 0.060	-0.053	29.918	0.522
		35 0.099	0.061	31.353	0.499
		36 -0.094	-0.013	32.677	0.483

Figura 23: Correlograma de los residuos del modelo $ARIMA(2, 1, 0) \times (0, 1, 1)_{12}$

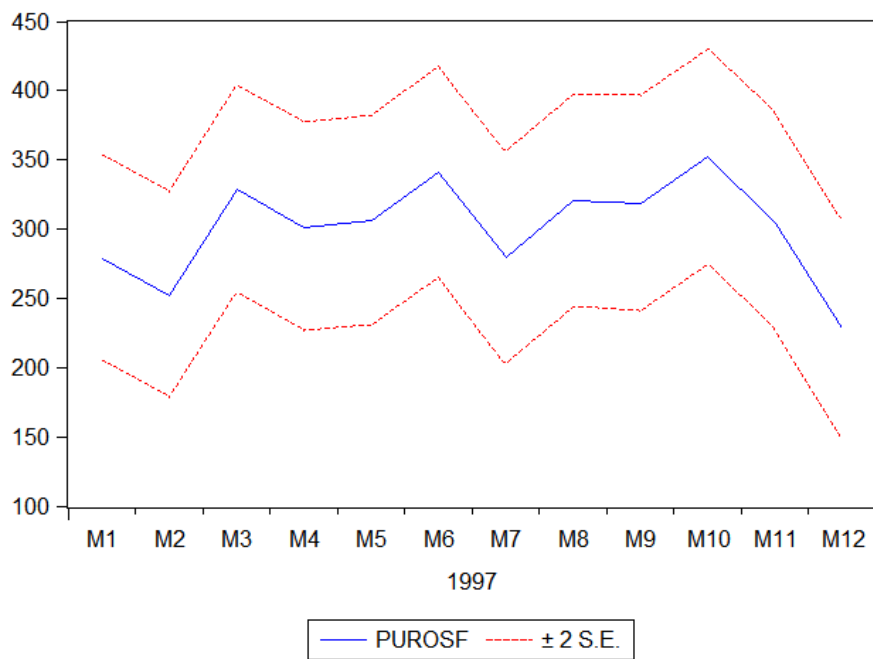


Figura 24: Predicciones (e intervalo de confianza) de ventas de puros para el año 1997

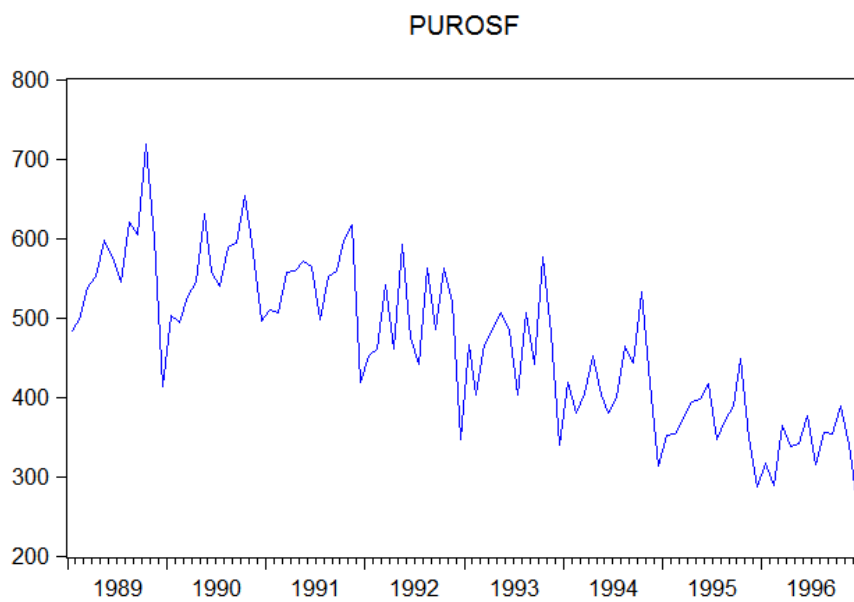


Figura 25: Serie temporal de puros junto a las predicciones para el año 1997

3. Conclusiones

En este trabajo hemos visto cómo aplicar la metodología Box-Jenkins para modelizar distintas series temporales, que trataban distintos temas socioeconómicos: la primera serie era sobre el índice de paro de un país, y la segunda, la venta de puros. En ambos casos, hemos utilizado el software EViews, que gracias a su potencia y sus modelos de series temporales ya implementados, ha permitido ahorrar mucho tiempo estimando, analizando y prediciendo nuestros modelos propuestos.

4. Código EViews

A continuación se muestra el código EViews utilizado para resolver cada una de las series temporales planteadas.

4.1. Índice de empleo

```
empleo.sheet
{%graph}.line
empleo.correl(16)
series d
series dempleo = empleo - empleo(-1)
dempleo.sheet
{%graph}.line
dempleo.correl
{%equation}.ls(optmethod=opg) d(empleo,1) ar(1)
{%equation}.resids(g)
{%equation}.correl
{%equation}.ls(optmethod=opg) d(empleo,1) ma(1) ma(2)
{%equation}.resids(g)
{%equation}.results
{%equation}.correl
{%equation}.resids(g)
dempleo.hist
{%graph}.line
{%equation}.ls(optmethod=opg) d(empleo,1) ar(1)
smpl 1994q1 1994q4
{%equation}.forecast(e, g) empleof
smpl 1962q1 1993q4
{%equation}.forecast
{%equation}.results
empleof.sheet
{%graph}.line
empleo.correl(16)
empleo.sheet
empleof.sheet
```

4.2. Venta de puros

```
puros.sheet
{%graph}.line
puros.correl
series dpuros = puros - puros(-1)
```

```

dpuros.sheet
{%graph}.line
dpuros.correl
series dd12puros = dpuros(puros,1,12)
series dd12puros = d(puros,1,12)
dd12puros.sheet
{%graph}.line
dd12puros.correl
{%equation}.ls(optmethod=opg) dd12puros ma(1) sma(12)
{%equation}.ls(optmethod=opg) d(puros,1,12) ma(1) sma(12)
{%equation}.correl
{%equation}.resids(g)
{%equation}.correl
{%equation}.ls(optmethod=opg) d(puros,1,12) ar(1) ar(2) ma(12)
{%equation}.resids(g)
{%equation}.correl
{%equation}.ls(optmethod=opg) d(puros,1,12) ma(1) sma(1)
dd12puros.sheet
puros.sheet
pagestruct(end=1997M12)
smpl 1997m01 1997m12
{%equation}.forecast(e, g) puros f
smpl 1989m01 1996m12
puros f.sheet
{%graph}.line

```