Series Temporales y Predicción

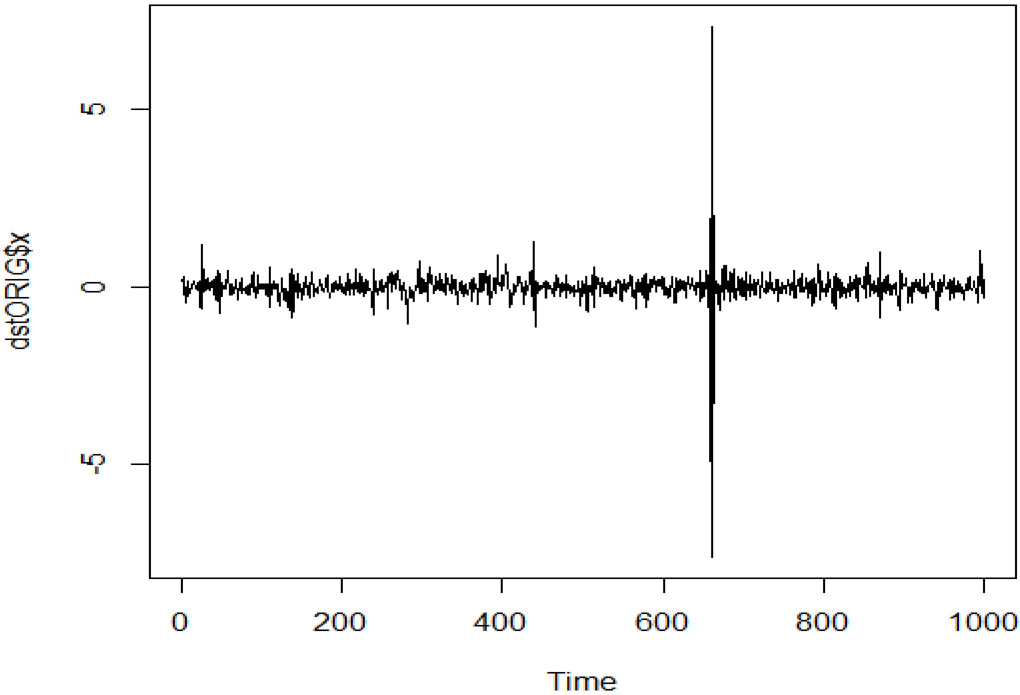
Práctica 12

Modelos ARCH y GARCH 2

Práctica 1

Cargad los datos del fichero “dst.txt”, correspondientes al índice DST que mide perturbaciones en la magnetosfera terrestre provocados por las tempestades solares.

a) Dibujad la serie y comentad los aspectos más relevantes. ¿Es un proceso estacionario? ¿Por qué? Prueba el test de Dickey-Fuller.



Observamos que la serie no tiene tendencia pero la varianza no es constante y se produce un salto importante alrededor de t = 620, el que hace sospechar que la varianza tiene una estructura. Lo analizaremos más a fondo en las próximas cuestiones.

Podría ser estacionario porque la serie está centrada en el 0 pero la variancia fluctúa en intervalos muy desiguales.

El test de Dickey-Fuller nos indica estacionariedad

```
        Augmented Dickey-Fuller Test

data:  dstORIG$x
Dickey-Fuller = -9.628, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(dstORIG$x) : p-value smaller than printed p-value
```

b) ¿Tiene tendencia?

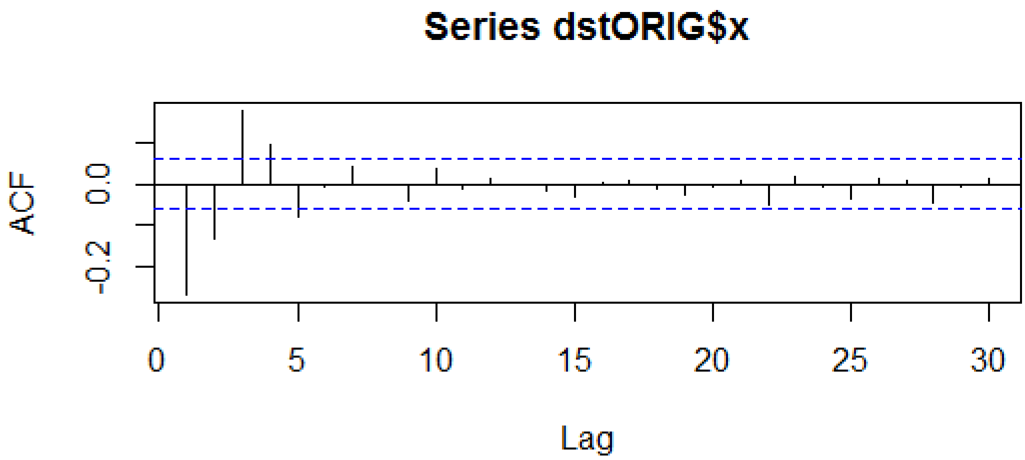
No parece tener tendencia. La serie se mueve alrededor del valor 0.

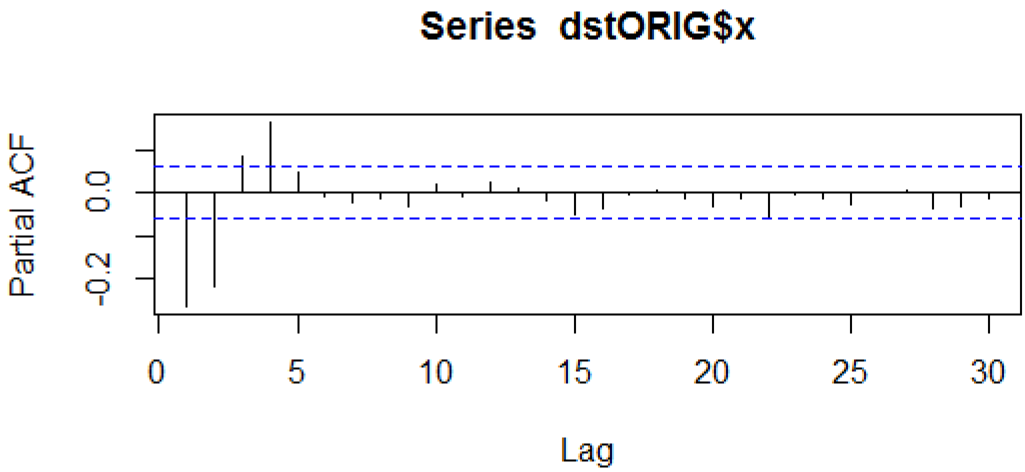
c) ¿Crees que tiene estacionalidad?

Tampoco parece tener una estacionalidad clara

```
> ### Per comprovar-ho, un truc basat en el likelihood ratio test
> fit1 <- ets(dstORIG$x)
> fit2 <- ets(dstORIG$x, model="AZN")
> deviance <- 2*c(logLik(fit1) - logLik(fit2))
> df <- attributes(logLik(fit1))$df - attributes(logLik(fit2))$df
> #P value
> 1-pchisq(deviance,df) ### H0: No és estacional / H1: És estacional
[1] 1
```

d) En base a la eacf, la ACF y la PACF, ¿qué modelo propondrías?





```
> eacf(dstORIG$x)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x x x x o o o o o o  o  o  o
1 x x x x x o o o o o o  o  o  o
2 x x o x o o o o o o o  o  o  o
3 x x x x o o o o o o o  o  o  o
4 x x x o o o x o o o o  o  o  o
5 x x o o x o o o o o o  o  o  o
6 x x o o x o o o o o o  o  o  o
7 x x o x x o o o o o o  o  o  o
```

El ACF y el PACF nos indican un proceso ARMA, aunque no queda claro el orden. En la eacf tampoco se muestra clarament. Por ser tentativos nos podríamos arriesgar con un ARMA(2,2)

e) ¿Cuál es el modelo que propone auto.arima()?

```
> mod1 <- auto.arima(dstORIG$x)
> mod1
Series: dstORIG$x
ARIMA(3,0,2) with zero mean

Coefficients:
         ar1      ar2     ar3      ma1     ma2
      0.3874  -0.3089  0.1582  -0.7159  0.3762
s.e.  0.1105   0.0821  0.0460   0.1106  0.0886

sigma^2 estimated as 0.1764:  log likelihood=-549.19
AIC=1110.38   AICc=1110.47   BIC=1139.83
```

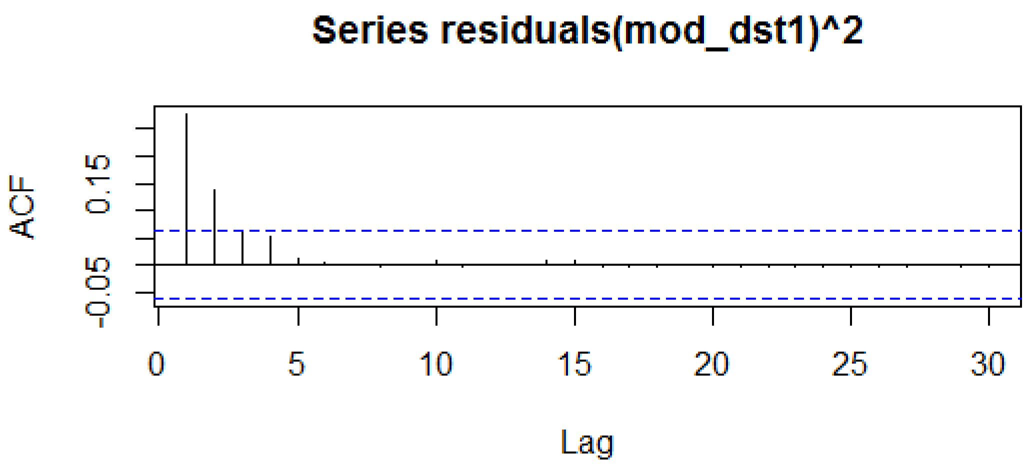
Propone un modelo ARMA(2,3)

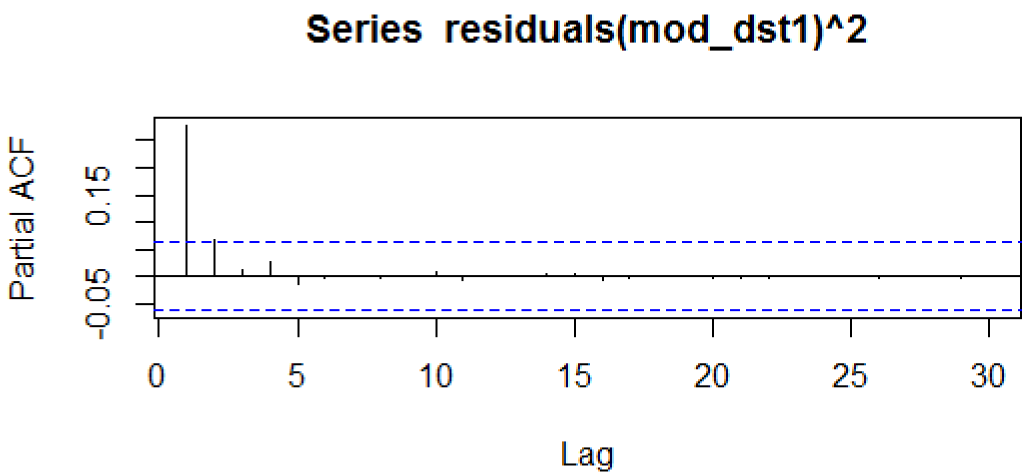
Modelo Y(t) = 0.3874 * Y(t-1) – 0.3089 * Y(t-2) + 0.1582 * Y(t-3) – 0.7159 * E(t-1) + 0.3762 * E(t-2) + E(t)

Donde E(t) se distribuye cómo una normal con esperanza cero y varianza 0.1764

f) Analizad los residuos al cuadrado de la serie. ¿Qué diríais?

```
> ### Els residus quadrats no son independents!
> Box.test(residuals(mod_dst1)^2, type="Ljung-Box")

        Box-Ljung test

data:  residuals(mod_dst1)^2
X-squared = 77.46, df = 1, p-value < 2.2e-16
```






```
> eacf(residuals(mod_dst1)^2)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x x o o o o o o o o  o  o  o
1 x o o o o o o o o o o  o  o  o
2 x o o o o o o o o o o  o  o  o
3 x x x o o o o o o o o  o  o  o
4 x x x x o o o o o o o  o  o  o
5 x o x x o o o o o o o  o  o  o
6 x x x o o o o o o o o  o  o  o
7 x x o o x o o o o o o  o  o  o
> 
```

No es un modelo satisfactorio, porque la ACF y PACF de los residuos al cuadrado muestran una estructura ARMA. La eacf nos parece indicar un modelo ARMA(1,1)

Y el test de Box-Lunjg nos indica que los residuos al cuadrado no son independientes.

Por lo que un modelo GARCH(1,1) podría ser adecuado.

g) Si la posibilidad de la estructura GARCH en los residuos de estos datos parece razonable, ajustad un modelo más adecuado mediante la instrucción *garch()* del paquete tseries o bien la instrucción *garchFit()* del paquete fGarch. Prueba las diferentes órdenes hasta encontrar el mejor modelo.

```
Title:
 GARCH Modelling

Call:
 garchFit(formula = ~garch(1, 1), data = dstORIG$x)

Mean and Variance Equation:
 data ~ garch(1, 1)
<environment: 0x0bfd0f10>
 [data = dstORIG$x]

Conditional Distribution:
 norm

Coefficient(s):
         mu        omega       alpha1        beta1
-0.01004123   0.02793788   0.67781095   0.00000001

Std. Errors:
 based on Hessian

Error Analysis:
         Estimate  Std. Error  t value Pr(>|t|)
mu     -1.004e-02   6.066e-03   -1.655   0.0978 .
omega   2.794e-02   2.849e-03    9.805   <2e-16 ***
alpha1  6.778e-01   7.197e-02    9.418   <2e-16 ***
beta1   1.000e-08   4.387e-02    0.000   1.0000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
 29.20258    normalized:  0.02920258

Description:
 Fri Jan 08 12:38:02 2021 by user: POR740051
```

El modelo propuesto por la función garch de R es un GARCH(1,1)

Y(t) = -0.01004 + E(t)

donde E(t) sigue una normal con esperanza 0 y varianza sigma^2(t)

donde sigma^2(t) = 0.02794 + 0.6778 * E(t-1) + 0.000000001 * sigma^2(t-1)

h) Con el resultado obtenido en el test de los coeficientes, ¿propondrías un modelo alternativo? En caso afirmativo, comparando el AIC de los modelos, ¿cuál te parece mejor?

El coeficiente beta1 aparece significativamente igual a 0. Podríamos descartarlo y pensar en un modelo GARCH(1,0)

```
Title:
 GARCH Modelling

Call:
 garchFit(formula = ~garch(1, 0), data = dstORIG$x)

Mean and Variance Equation:
 data ~ garch(1, 0)
<environment: 0x0e858490>
 [data = dstORIG$x]

Conditional Distribution:
 norm

Coefficient(s):
       mu      omega     alpha1
-0.010041   0.027938   0.677811

Std. Errors:
 based on Hessian

Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.010041    0.005983   -1.678   0.0933 .
omega   0.027938    0.002019   13.840   <2e-16 ***
alpha1  0.677811    0.067936    9.977   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
 29.20258    normalized:  0.02920258

Description:
 Fri Jan 08 12:45:04 2021 by user: POR740051
```

El modelo propuesto para un GARCH(1,0) es

Y(t) = -0.01004 + E(t)

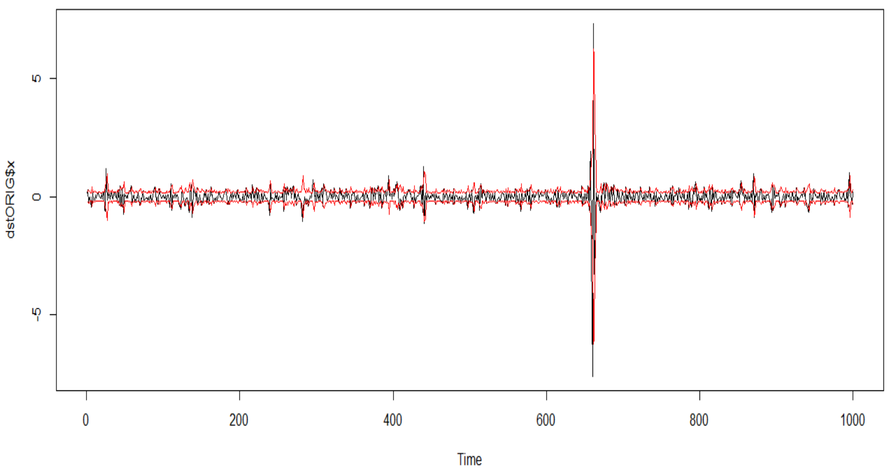
donde E(t) sigue una normal con esperanza 0 y varianza sigma^2(t)

donde sigma^2(t) = 0.02794 + 0.6778 * E(t-1)

Comparando los dos AICs, da mejor el último modelo propuesto GARCH(1,0)

```
> AIC(mod_dst2_a)
[1] -25.50572
> AIC(mod_dst2_b)
[1] -51.81597
```

i) Utilizando la instrucción *predict()* del paquete tseries, dibujad un intervalo de confianza para las estimaciones en todo el periodo juntamente con las observaciones reales. ¿Qué podéis decir sobre la bondad del modelo?

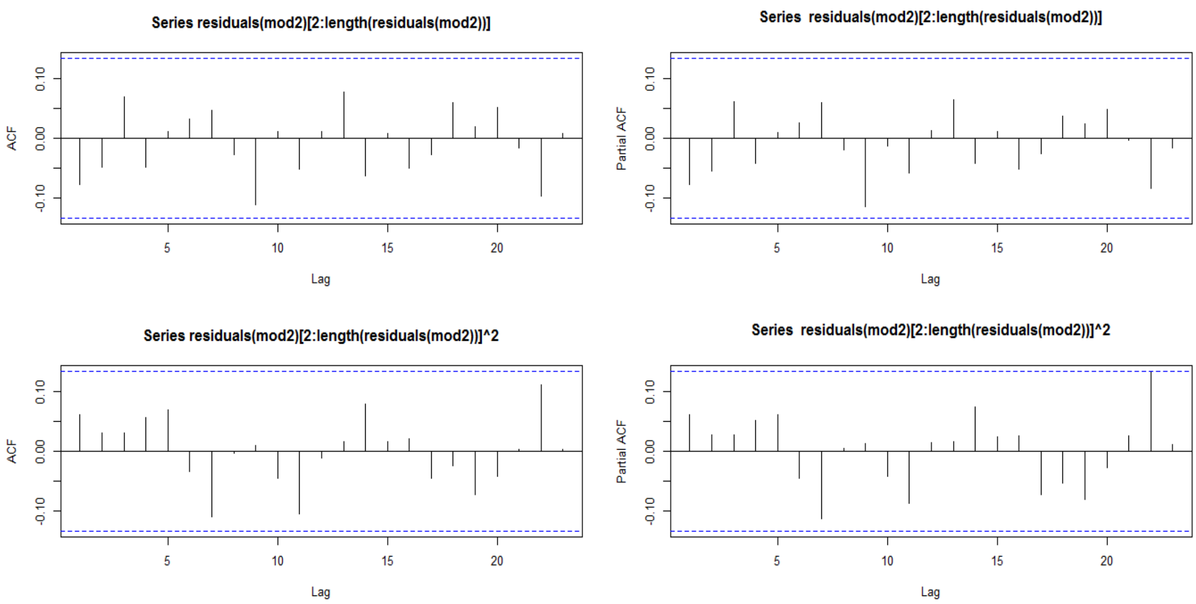


j)

El modelo parece incorporar adecuadamente la estructura en la varianza de los residuos, y en la mayoría de casos les observaciones se encuentran dentro de los límites del intervalo de confianza.

Además ahora los residuos siguen una normal y los residuos al cuadrado no presentan ninguna estructura.

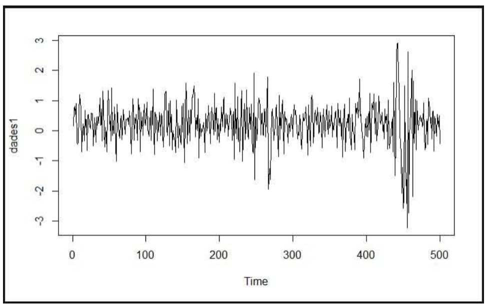
```
> Box.test(residuals(mod_dst2_b), type="Ljung-Box")

        Box-Ljung test

data:  residuals(mod_dst2_b)
X-squared = 0.20248, df = 1, p-value = 0.6527

> Box.test(residuals(mod_dst2_b)^2, type="Ljung-Box")

        Box-Ljung test

data:  residuals(mod_dst2_b)^2
X-squared = 0.58818, df = 1, p-value = 0.4431
```



Práctica 2

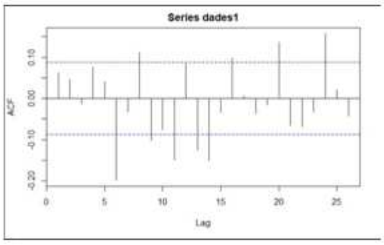
Cargad la serie temporal PRC2_data1.csv que corresponde al incremento de ventas porcentual diario de una cadena de grandes almacenes.

a) Grafica la serie temporal. ¿Es estacionaria? ¿Cómo lo justificarías rápidamente?

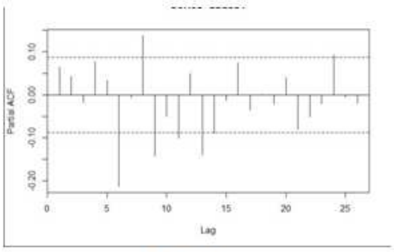


Si observamos la gráfica, podemos ver que alrededor del 450 la varianza tiene un salto lo suficientemente grande que nos indica que la varianza no es constante y por lo tanto posiblement se trate de un proceso no estacionario.

b) Comenta el output de las funciones acf, pacf y eacf. ¿Qué crees que implica sobre la estructura de la serie?



(a) Acf



(b) PACF

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	o	o	o	o	o	x	o	x	x	o	x	o	x	x
1	x	o	o	o	o	x	o	o	x	o	x	o	x	x
2	x	x	o	o	o	x	x	o	o	o	x	o	x	o
3	x	x	x	o	o	o	o	o	x	x	o	x	x	o
4	x	x	o	x	o	o	o	o	x	o	x	o	x	x
5	x	x	o	x	x	o	o	o	o	o	x	o	x	x
6	o	x	x	x	x	x	o	o	o	o	o	o	x	o
7	o	x	x	x	x	x	o	o	o	o	o	o	x	x

(c) EACF

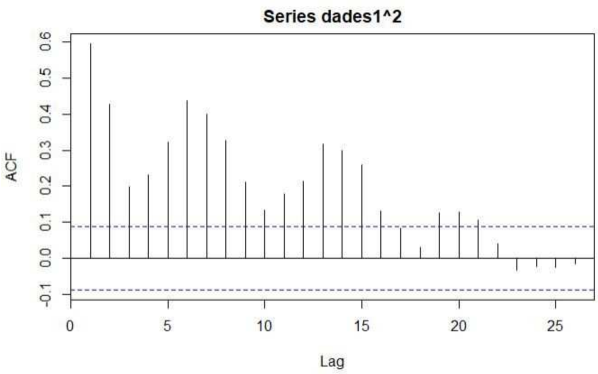
```
Series: dades1
ARIMA(0,0,0) with non-zero mean

Coefficients:
        mean
      0.2250
s.e.  0.0316

sigma^2 estimated as 0.4999:  log likelihood=-535.65
AIC=1075.29   AICc=1075.32   BIC=1083.72
```

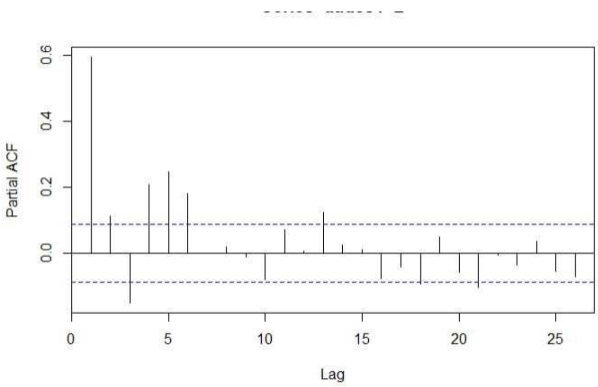
(d) Auto.arima

Si miramos l'acf, pacf, eacf podría hacernos pensar en un ruido blanco. De hecho la auto.arima nos recomienda un ARMA(0,0).

c) Ajusta el modelo que creas que es más conveniente para esta serie y comentad los términos/coeficientes.

Ahora bien si miramos las funciones PACF i ACF de los residuos al cuadrado observamos una estructura. Por lo que podemos pensar en un modelo GARCH.





El mejor modelo parece ser un GARCH(1,1)

	ARCH(1)	GARCH(1,1)	GARCH(2,1)	GARCH(1,2)
AIC	1018.605	1004.724	1057.895	1004.747

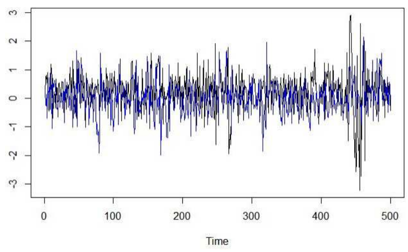
d) Da la expresión formal del modelo.

El modelo ajustado es

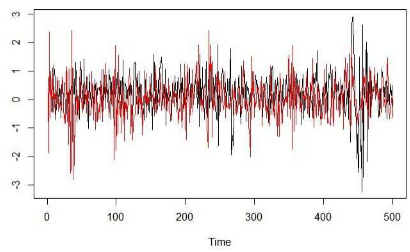
$$r_t = a_t - 0.2250435 \qquad a_t = \sigma_y \epsilon_t$$

$$\sigma_t^2 = 0.086564 + 0.298411 a_{t-1}^2 + 0.486549 \sigma_{t-1}^2$$

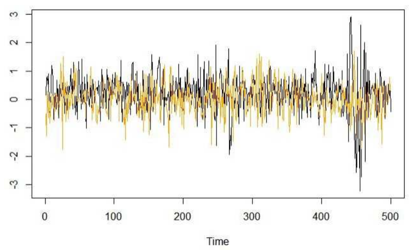
e) Dad las predicciones de varios modelos y comprobad si los resultados son similares. Mostrad los resultados de las predicciones gráficamente.



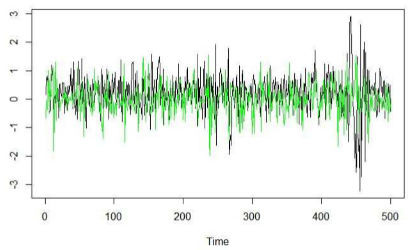
(a) ARCH(1)



(b) GARCH(1,1)



(c) GARCH(1,2)



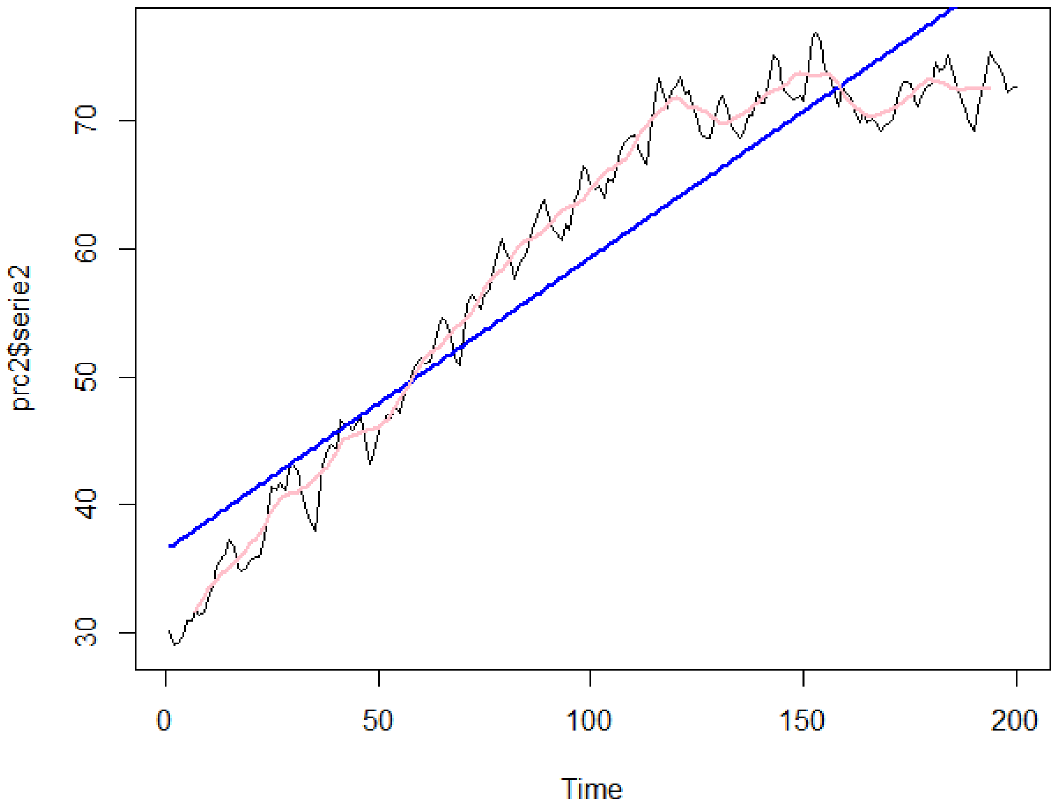
(d) GARCH(2,1)

Mirando los gráficos parece que el modelo GARCH(1,1) pueda predecir mejor las varianzas de los datos.

Práctica 3

Trabaja con la serie temporal PRC2_data2.csv que corresponde al número de abanicas por semana (en miles de €).

a) Grafica la serie temporal, añadiéndole la tendencia (con una media móvil y con la recta de regresión).



b) Comenta los resultados ¿Tiene estacionalidad? ¿Cómo lo comprobarías? Si es que sí, coméntala.

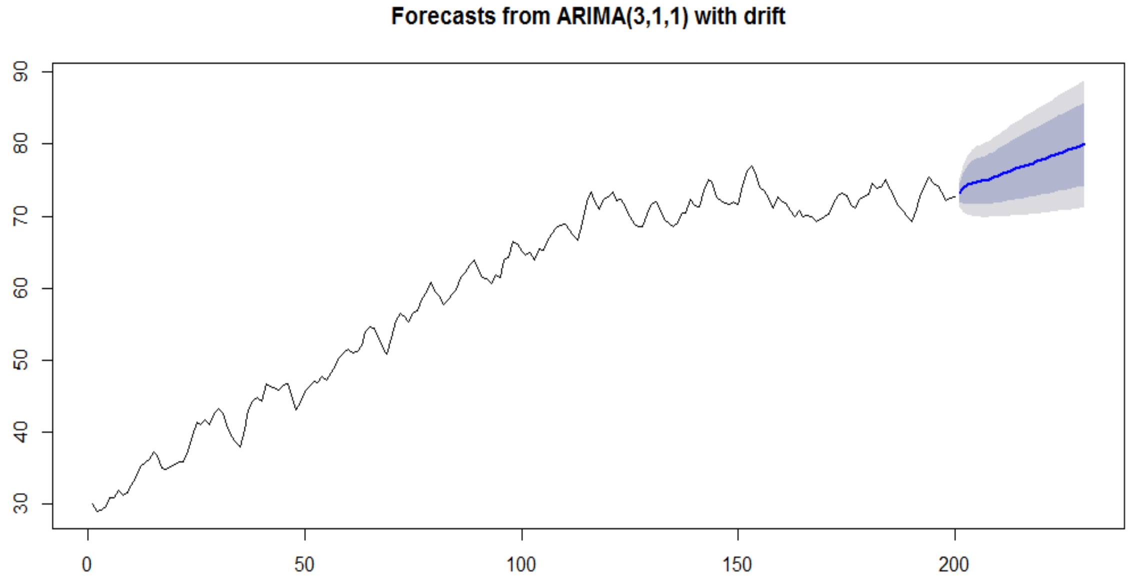
```
> ### Per comprovar-ho, un truc basat en el likelihood ratio test
> fit1 <- ets(prc2$serie2)
> fit2 <- ets(prc2$serie2, model="AZN")
> deviance <- 2*c(logLik(fit1) - logLik(fit2))
> df <- attributes(logLik(fit1))$df - attributes(logLik(fit2))$df
> #P value
> 1-pchisq(deviance,df) ### H0: No és estacional / H1: És estacional
[1] 1
> ### La component estacional no és significativa.
```

c) Ajusta el modelo que creas que es más conveniente por esta serie y comenta los términos/coeficientes, siguiendo las técnicas que hemos visto en clase.

```
Series: prc2$serie2
ARIMA(3,1,1) with drift

Coefficients:
         ar1      ar2      ar3      ma1   drift
      0.8743  -0.2058  -0.1928  -0.5982  0.2178
s.e.  0.1812   0.1164   0.0872   0.1803  0.0538

sigma^2 estimated as 0.9873:  log likelihood=-278.8
AIC=569.6   AICc=570.03   BIC=589.36
> coeftest(mod_autoarima)

z test of coefficients:

       Estimate Std. Error z value  Pr(>|z|)    
ar1    0.874259   0.181242  4.8237 1.409e-06 ***
ar2   -0.205822   0.116376 -1.7686 0.0769600 .  
ar3   -0.192752   0.087206 -2.2103 0.0270839 *  
ma1   -0.598247   0.180264 -3.3187 0.0009043 ***
drift  0.217760   0.053766  4.0502 5.118e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

La función auto.arima propone diferenciar la serie y una vez diferenciada propone un ARMA(3,1).

d) Muestra las predicciones a 30 semanas vista con el modelo que has ajustado, con un intervalo de confianza del 90 y 95%. ¿Crees que se ajusta correctamente?



No parece ajustar bien la tendencia más suavizada de las últimas observaciones.

Si realizamos el modelo sin las últimas observaciones el modelo obtenido nos da las siguientes predicciones, donde la serie original está coloreada en rojo: