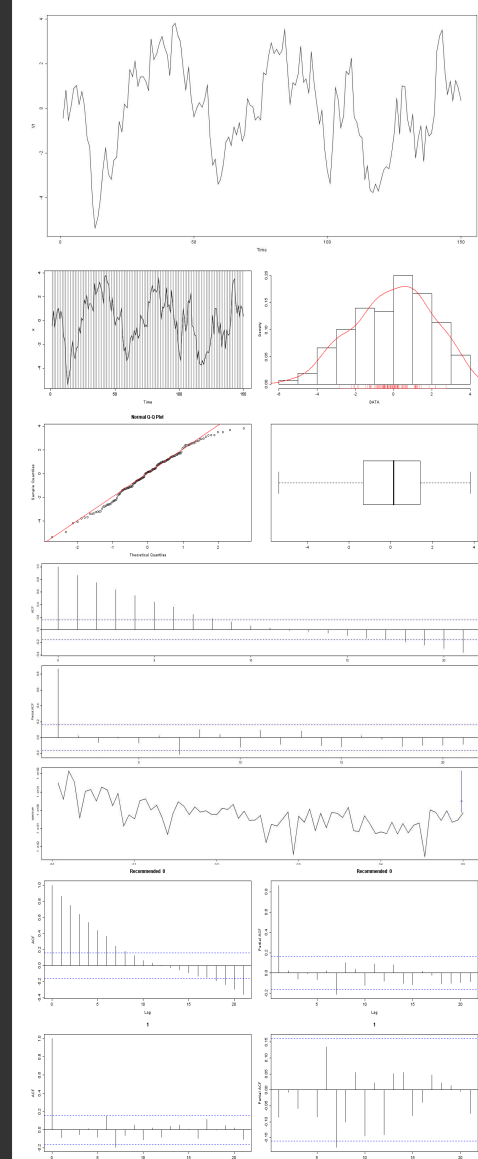


# Análisis de Series Temporales.

Análisis de 10 series de datos aplicando técnicas y métodos estadísticos de análisis de series temporales con R.



Juan Antonio Breña Moral

# Índice.

1. Introducción
2. El motor estadístico R
3. Análisis de las series de estudio
  - a. Serie 1
  - b. Serie 2
  - c. Serie 3
  - d. Serie 4
  - e. Serie 5
  - f. Serie 6
  - g. Serie 7
  - h. Serie 8
  - i. Serie 9
  - j. Serie 10
  - k. Conclusiones
4. Conclusiones
5. Bibliografía
6. Anexos
  - a. Scripts desarrollados para el desarrollo de los análisis

## 1. Introducción.

Dado el conjunto de datos, proporcionados por el profesor, se pretende estudiar el comportamiento de los datos presentes en las diferentes muestras empleando análisis de series temporales. Para ello, nos apoyaremos en el del motor Estadístico R y los Scripts generados para dicho fin.

Los objetivos a conseguir en cada una de las series son:

1. Utilizar herramientas numéricas y graficas para el análisis de las series.
2. Estimar el modelo AR, MA, ARMA o ARIMA para cada una de ellas (tener en cuenta posibles transformaciones)
3. Estimar los parámetros de los modelos, así como intervalos de confianza para los mismos.
4. Realizar estimaciones e inferencias sobre las predicciones de las 30 observaciones de cada serie.

## 2. El motor estadístico R

R es un motor estadístico desarrollado para realizar análisis estadístico por Ross Ihaka y Robert Gentleman. R es considerado como un dialecto del lenguaje S creado por los Laboratorios AT & T Bell. S se comercializa como el programa S-PLUS comercializado por Insightful.

R se distribuye gratuitamente bajo los términos GNU, General Public Licence. Su desarrollo y distribución son llevados a cabo por varios estadísticos conocidos con el sobrenombre de Grupo Nuclear de Desarrollo R.

R está disponible para máquinas Unix y Linux, o como archivos binarios precompilados para Windows, Linux (Debian, Mandrake, RedHat, SuSe), Macintosh y Alpha Unix.

Los archivos necesarios para instalar R, ya sea desde las fuentes o binarios pre-compilados, se distribuyen desde el sitio de internet Comprehensive R Archive Network (CRAN) R posee muchas funciones para análisis Los resultados de análisis estadísticos se muestran en la pantalla, y algunos resultados intermedios (como valores P-, coeficientes de regresión, residuales, ...) se pueden guardar, exportar a un archivo, o ser utilizados en análisis posteriores.

## 3. Análisis de las series de estudio

**Nota:** Las series dadas para trabajar la asignatura, se desconoce a priori el contexto del cual se extrajeron.

El procedimiento para analizar cada serie antes de concluir cuales de ellas se modelaran con modelos AR, MA, ARMA y ARIMA es el siguiente:

1. Análisis Descriptivo de los datos
2. Análisis de serie temporal compuesto de:
  - a. Autocorrelaciones
  - b. Espectrograma
  - c. Periodograma

Una vez se ha analizado los datos a través de estos procedimientos, se concluirá los posibles modelos que aceptan los datos. Una vez se determinan los posibles modelos, se procede a la modelización de los mismos y realizar comparaciones de eficiencia.

### 3.1 Serie 1

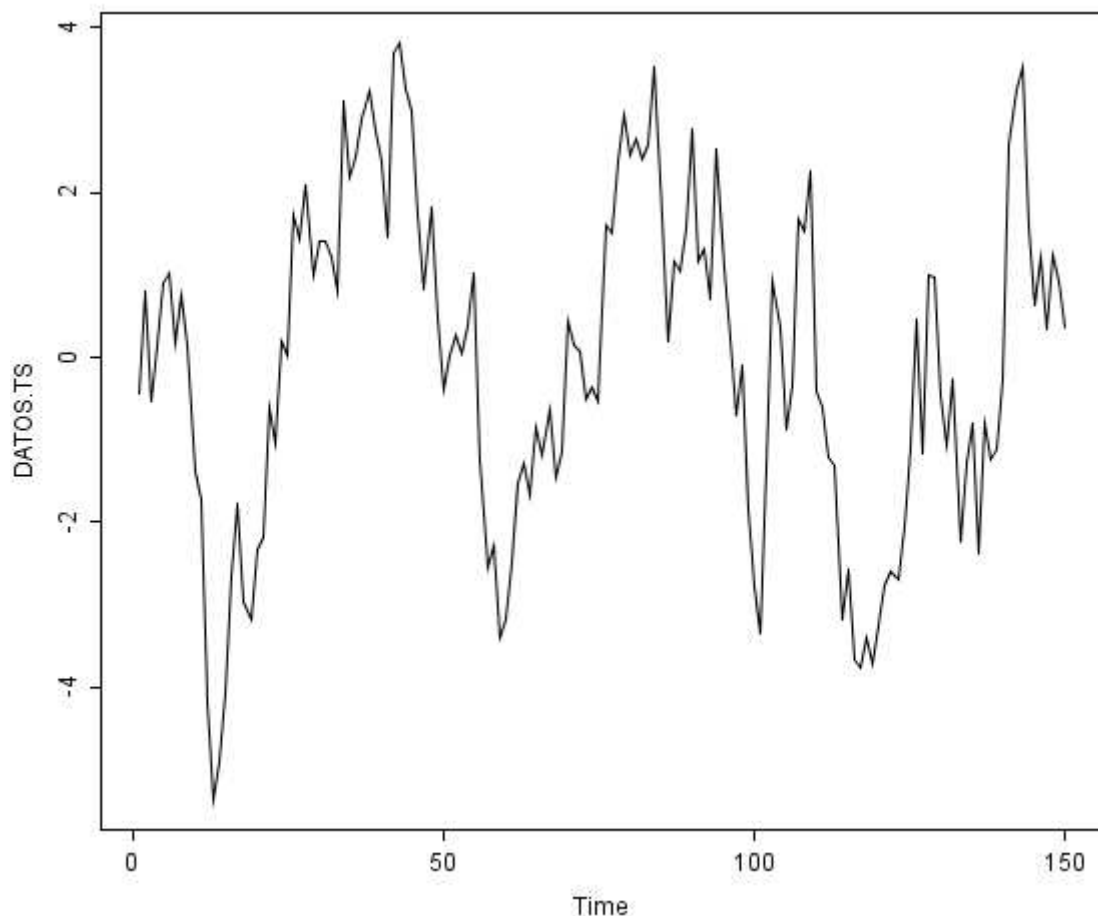
Dada la serie, se procederá a realizar el estudio de análisis de series temporales.

#### Análisis descriptivo de una serie temporal.

Todo análisis de serie temporal univariante, comienza con la presentación de un grafico donde se muestra la evolución de la variable a lo largo del tiempo.

```
#Estableciendo el directorio de trabajo
setwd("C:/DATOS/DOCUMENTACION/2006/DOCTORADO/STAT_METHODS/TIME_SERIES/SERIES/Series/");
#Cargando libreria de analisis de normalidad
library("nortest");
#librerias de time series
library("dyn");
library("ArDec");
library("forecast");
library("fBasics");
library("fCalendar");
library("fSeries");
library("tseries");
#Cargando libreria de EDA, desarrollada por Juan Antonio Breña Moral
source("JAB.EDA.txt");
source("JAB.TS.VIEW.txt");
```

Los resultados son los siguientes:



*Figura 1.1*

Los resultados de la representación grafica de los datos muestran un comportamiento extraño en cuanto a la amplitud de los ciclos. Se detecta una amplitud de ciclo mayor en el ciclo segundo y tercero entendiendo que los datos empiezan mostrando el final de un ciclo. Dichos ciclos (segundo y tercero) son mayores que la amplitud de ciclo cuarto, quinto y sexto. Esta diferencia de amplitudes en los ciclos me hace pensar como si los datos presentados mostrasen 2 fases dentro de la naturaleza de la información suministrada por tanto, a partir del dato 100, los datos empezarían otra fase y por tanto, si el objetivo del análisis de series temporales es el de predecir, yo me decantaría por modelar la ultima fase apreciada anteriormente. Esta hipótesis de partida se descarta al ejecutar el Script de EDA, que a continuación se presenta:

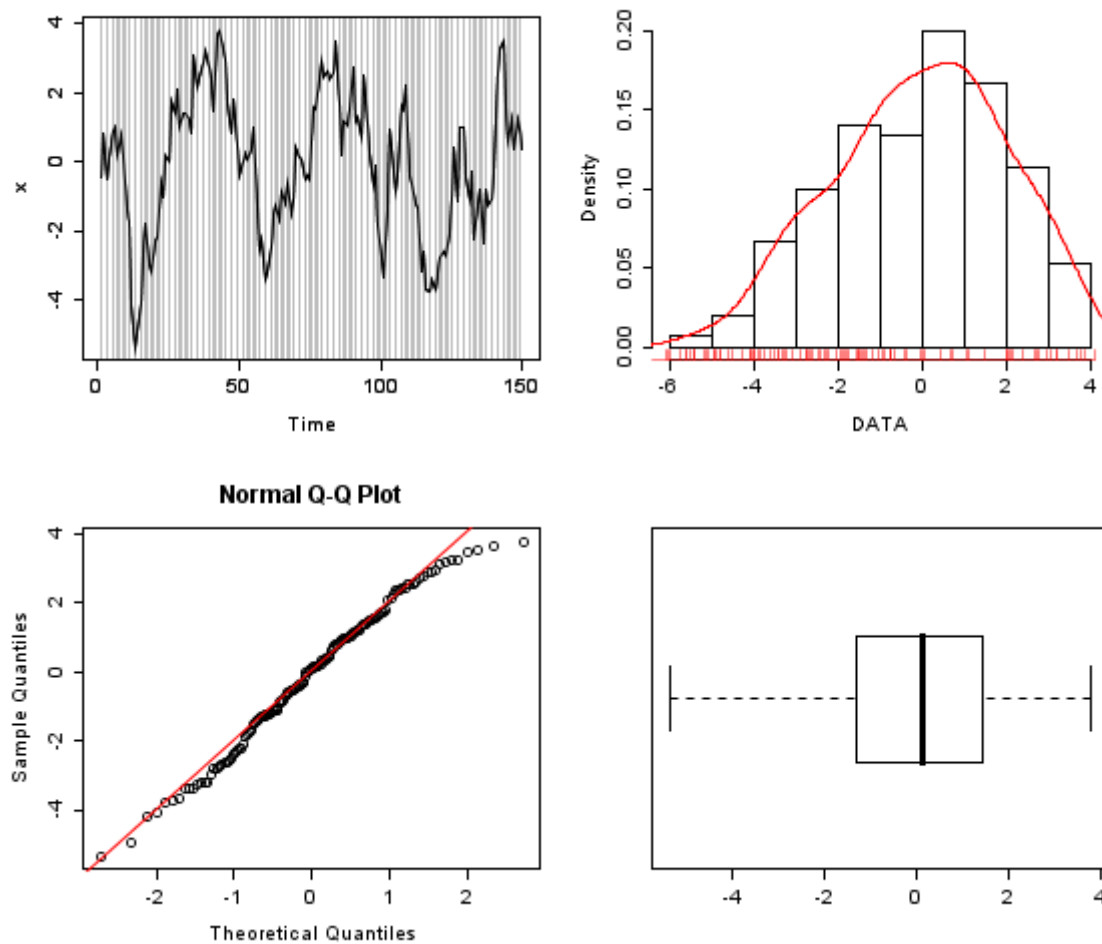


Figura 1.2

[1] "JAB: EDA VIEW 1.1"

The decimal point is at the |

```
-5 | 4
-4 | 9
-4 | 21
-3 | 877
-3 | 44432220
-2 | 88766655
-2 | 433221
-1 | 987755
-1 | 443332222221111
-0 | 988876665555
-0 | 44443331
0 | 0001112222234444
0 | 55678888999
1 | 000000122223444
1 | 55566677778
```

2 | 1234444  
 2 | 556678899  
 3 | 01233  
 3 | 5578

MEAN	MEDIAN	SD	KURTOSIS	SKEWNESS
-0.06355497	0.13620879	2.01945013	-0.60020371	-0.24538213

[1] "Normal Test, with p-value = 0,05"

Shapiro-Wilk	Anderson-Darling	Cramer-von	Pearson
0.0962893	0.1943319	0.2594305	0.1156905

Shapiro-Francia	Jarque-Bera
0.1784554	0.1693652

[1] "OK" "OK" "OK" "OK" "OK" "OK"

Como se observa en los resultados, los datos presentan una cierta simetría, no perfecta. No se observan largas colas. Los datos pasan los test de Normalidad 95% y presentan una cierta estacionalidad en torno a la mediana, tomando este valor debido a una alta variabilidad observada por el valor de la desviación típica

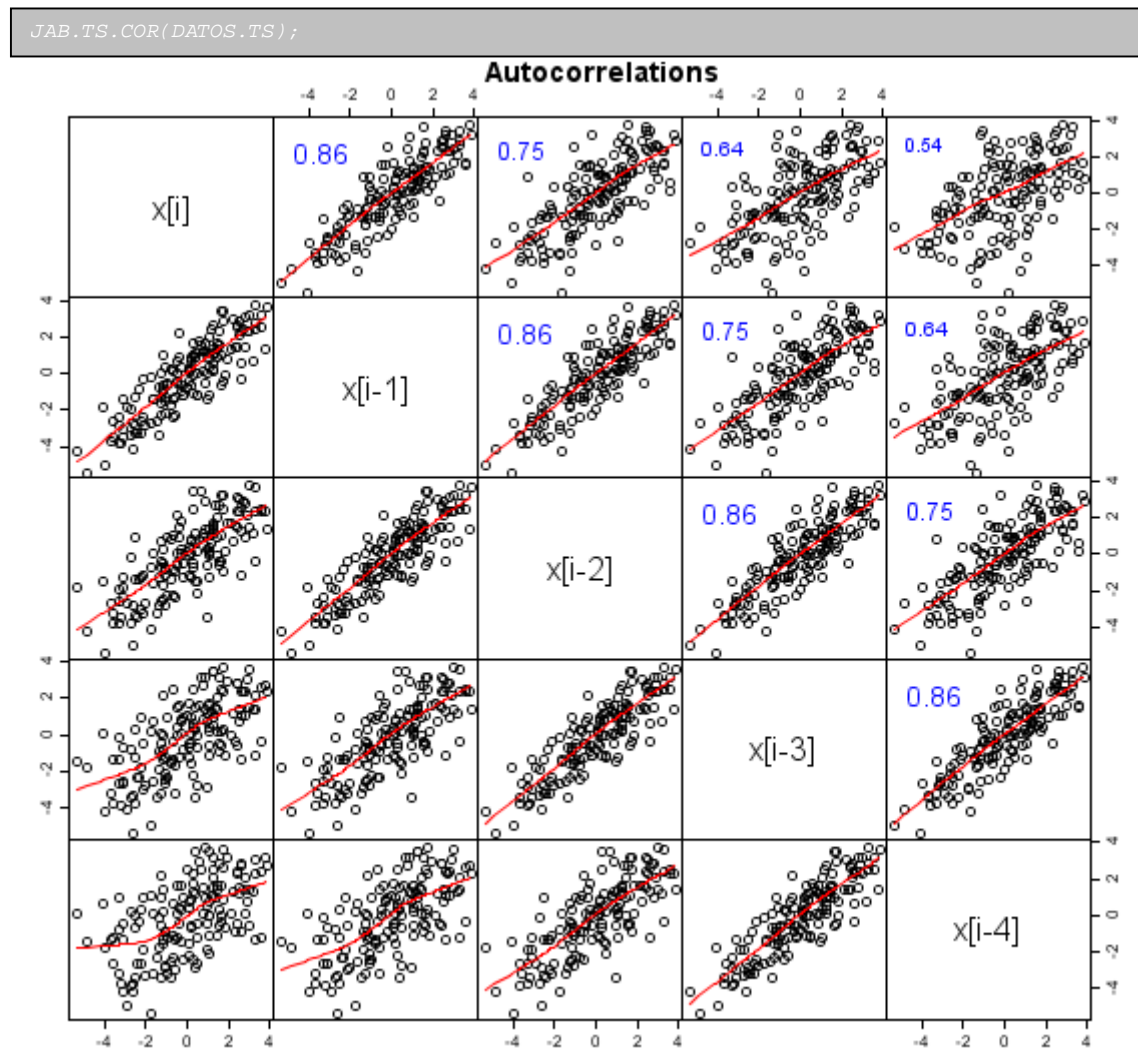


Figura 1.3

Como se observan en las autocorrelaciones, en todos los ordenes, los datos parecen correlados.

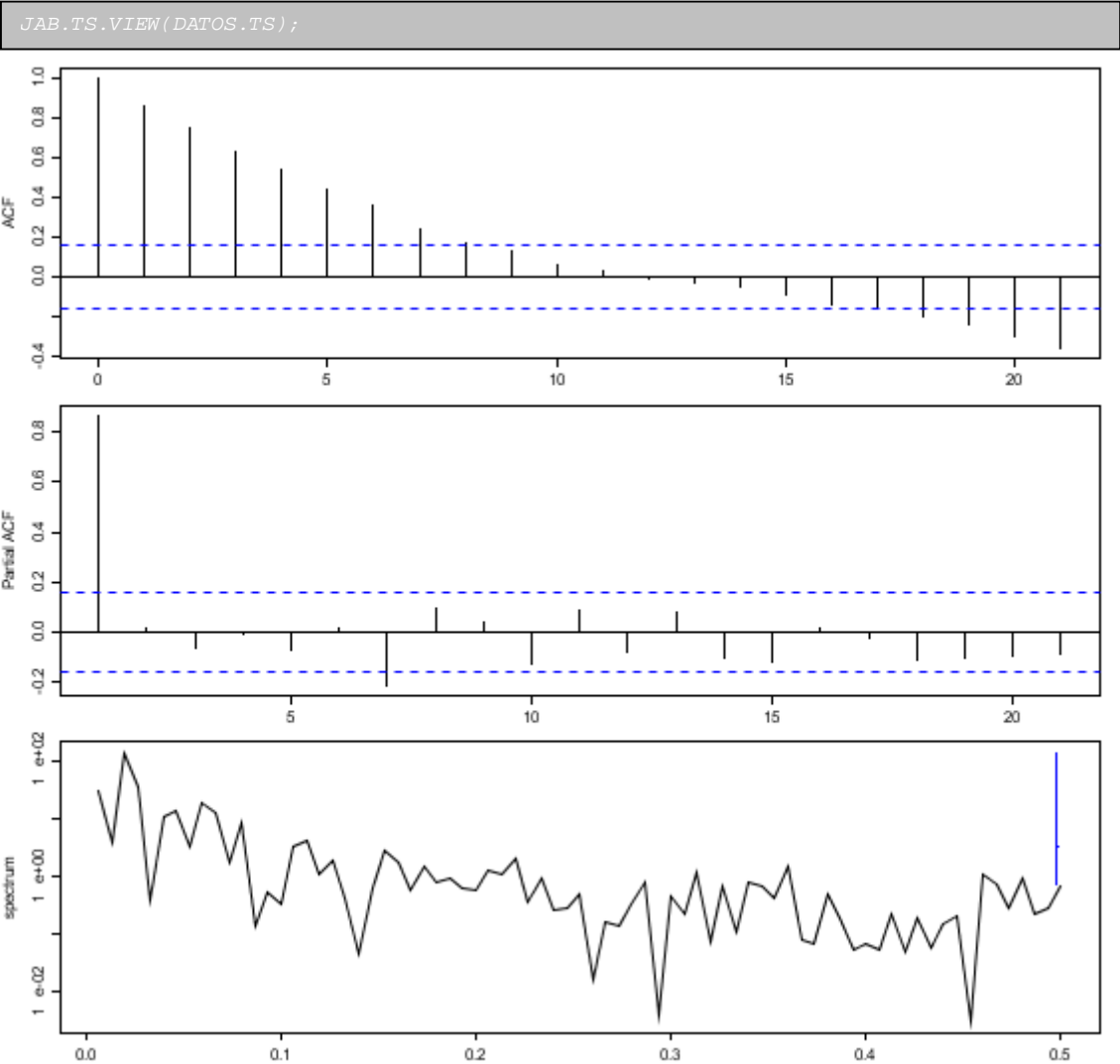


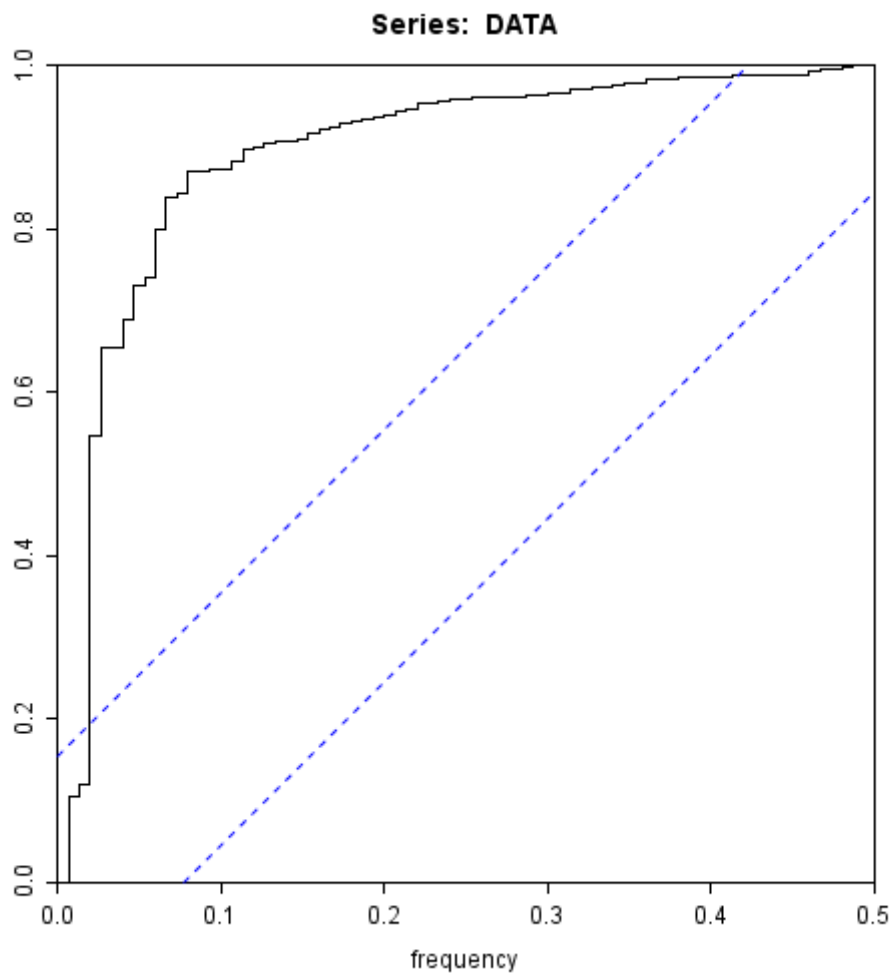
Figura 1.4

Como se observa en los resultados, la funcion de autocorrelacion total muestra un comportamiento en el cual existen al principio una serie de valores no nulos que se van amortiguando a lo largo del tiempo. En el grafico de la autocorrelacion parcial, se muestra un unico valor no nulo, lo cual nos indica como si la serie pudiese ser modelada a traves de un modelo AR(1)

Esto se contrasta con la tabla dada en el libro de Daniel Peña, pagina 158.

TIPO	FAS	FAP
AR(p)	Muchos coeficientes no nulos	Primeros p no nulos, resto 0
MA(q)	Primeros p no nulos, resto 0	Muchos coeficientes no nulos
ARMA(p,q)	Muchos coeficientes no nulos	Muchos coeficientes no nulos

A continuación, se representara el periodograma.



*Figura 1.5*

DUDA, ¿como interpreto esto?

## Modelado de la serie temporal.

Como se observo en el anterior apartado, la serie de datos 1, parecia ser modelada a traves de un AR(1) debido al estudio de los graficos de autocorrelacion parcial y total:



```
JAB.TS.ACFS(DATOS.TS);
```

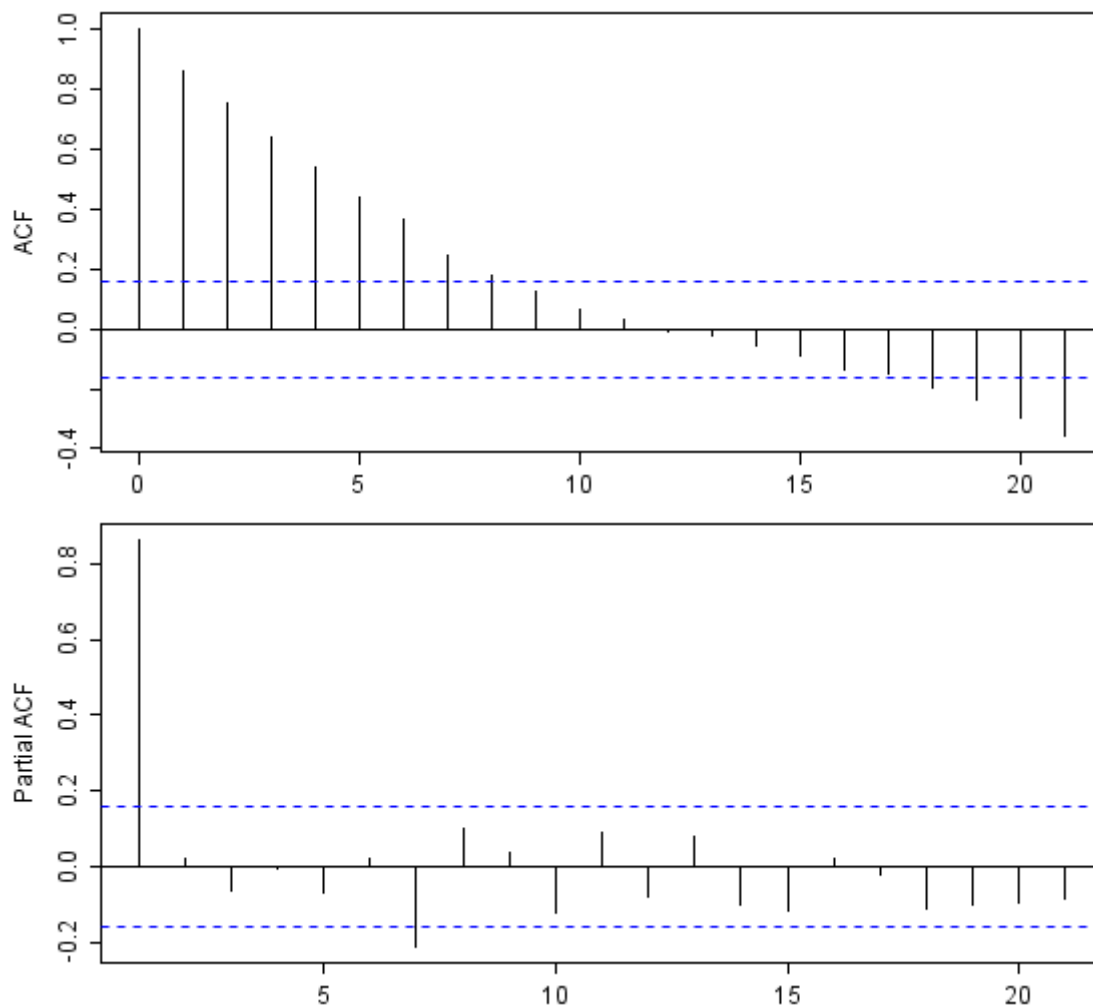


Figura 1.6

El modelo AR se define como:

$$X_t = \alpha_1 X_{t-1} + \dots + \alpha_{t-p} X_{t-p} + Z_t$$

$$\text{d } E(Z_t) = 0, \text{Var}(Z_t) = \sigma_Z^2$$

Para modelar una serie temporal como un modelo AR:

```
DATOS.TS.AR <- ar(DATOS.TS);
DATOS.TS.AR;
```

Call:

ar(x = DATOS.TS)

Coefficients:

1  
0.8638

Order selected 1 sigma^2 estimated as 1.042

Como se observa, el algoritmo selecciona tambien un modelo de orden 1

```
predict(DATOS.TS.AR, n.ahead=30)
```

\$pred

Time Series:

Start = 151

End = 180

Frequency = 1

```
[1] 0.2998444000 0.2503494267 0.2075956659 0.1706649648 0.1387642231  
[6] 0.1112083605 0.0874056048 0.0668447830 0.0490843440 0.0337428757  
[11] 0.0204909145 0.0090438696 -0.0008440884 -0.0093853071 -0.0167632123  
[16] -0.0231362473 -0.0286412753 -0.0333965188 -0.0375040984 -0.0410522259  
[21] -0.0441170986 -0.0467645359 -0.0490513923 -0.0510267791 -0.0527331182  
[26] -0.0542070541 -0.0554802400 -0.0565800181 -0.0575300064 -0.0583506064
```

\$se

Time Series:

Start = 151

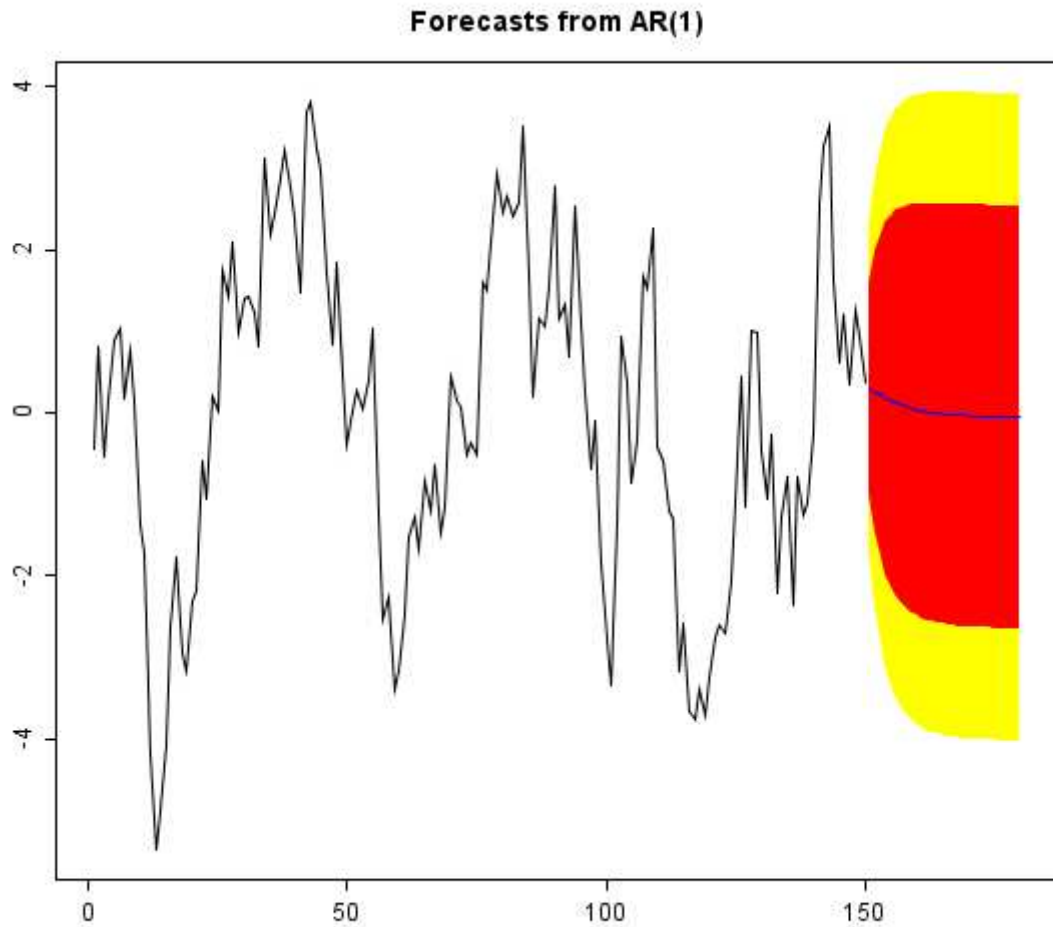
End = 180

Frequency = 1

```
[1] 1.020901 1.349039 1.549245 1.683188 1.776562 1.843154 1.891316 1.926467  
[9] 1.952283 1.971325 1.985414 1.995863 2.003623 2.009394 2.013689 2.016888  
[17] 2.019271 2.021048 2.022373 2.023360 2.024097 2.024647 2.025057 2.025362  
[25] 2.025591 2.025761 2.025888 2.025983 2.026053 2.026053
```

Para observar la prediccion dada por el sistema:

```
plot(forecast(DATOS.TS.AR,30,conf=c(80,95)));
```



*Figura 1.7*

Como se observa, el sistema proporciona, bandas de confianza.

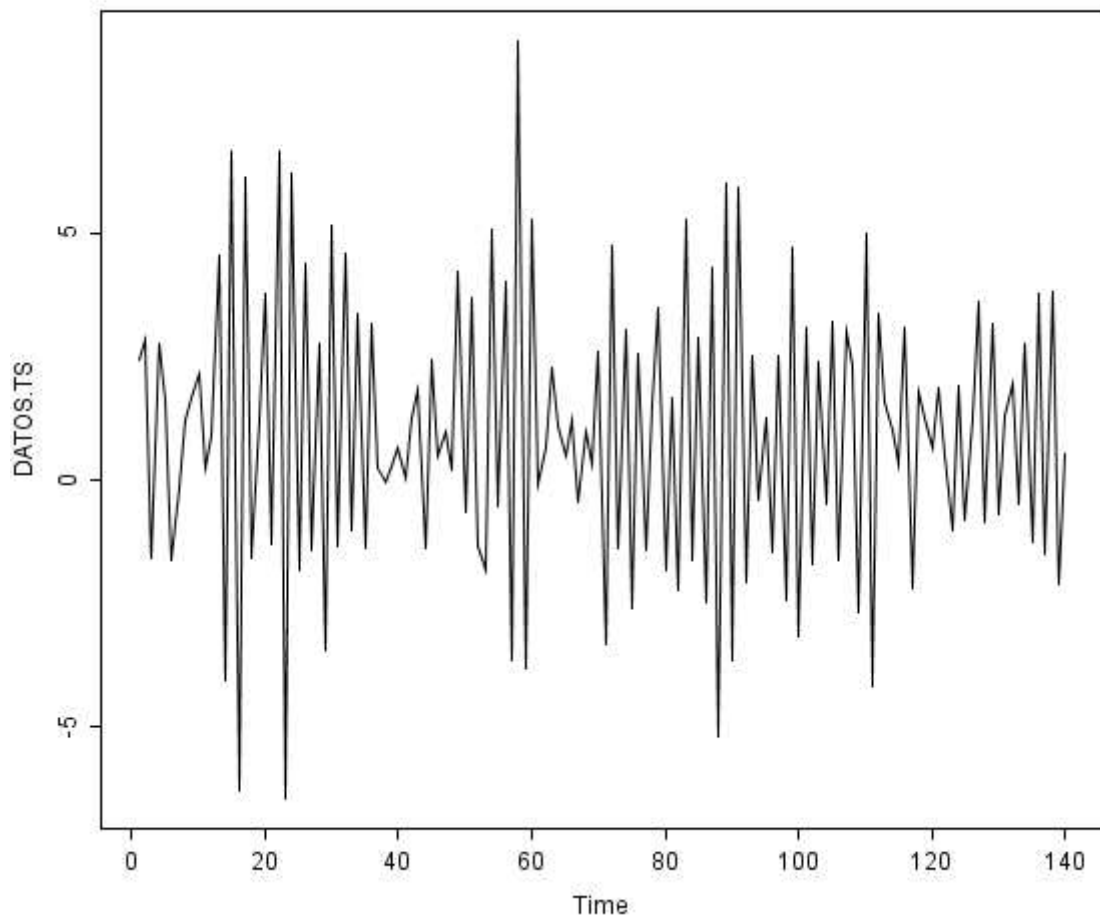
## **3.2 Serie 2**

Dada la serie, se procederá a realizar el estudio de análisis de series temporales.

### **Análisis descriptivo de una serie temporal.**

Todo análisis de serie temporal univariante, comienza con la presentación de un gráfico donde se muestra la evolución de la variable a lo largo del tiempo.

Los resultados son los siguientes:



*Figura 2.1*

Como se observa, los datos son estacionarios en un valor que luego calcularemos al realizar un análisis exploratorio de datos, EDA, pero me preocupa la variabilidad de los mismos y la falta de homogeneidad de los ciclos.

Al realizar el análisis exploratorio de datos

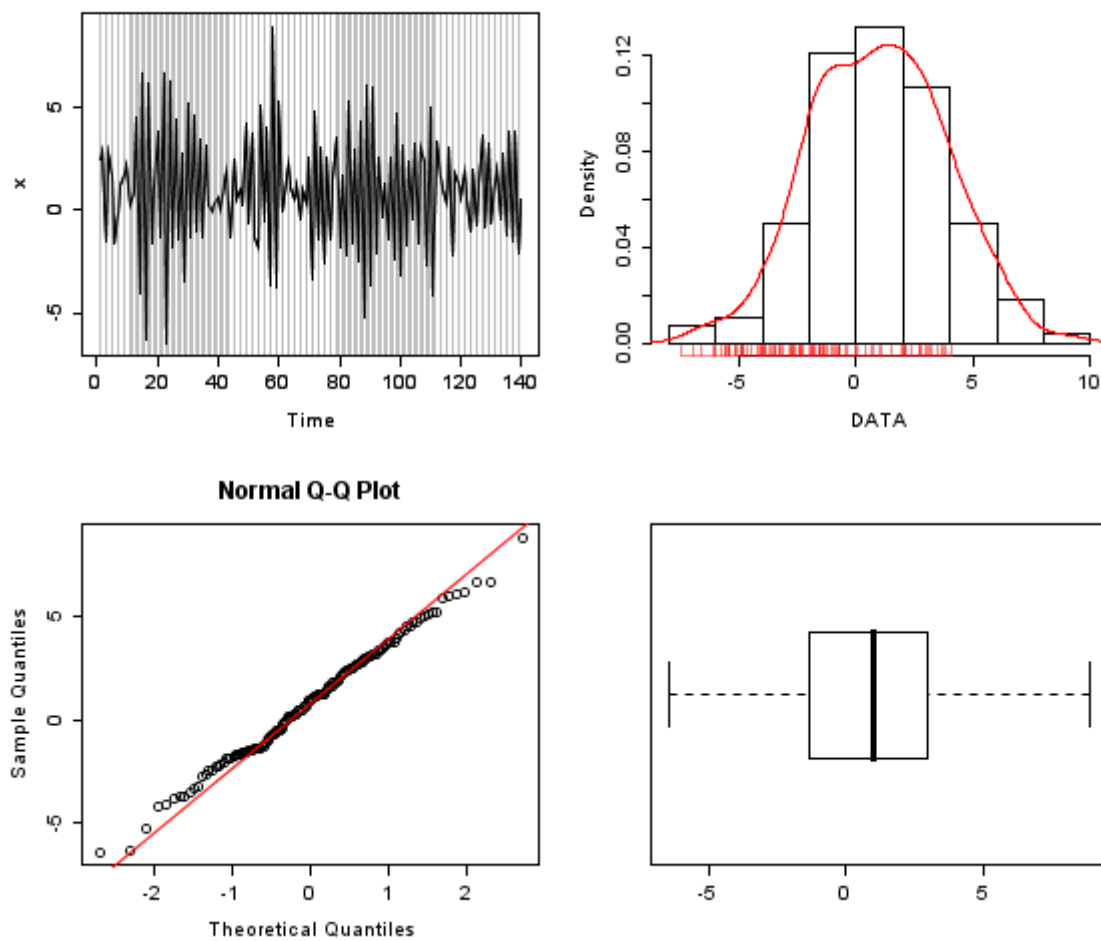


Figura 2.2

[1] "JAB: EDA VIEW 1.1"

The decimal point is at the |

```
-6 | 43
-5 | 2
-4 | 20
-3 | 866531
-2 | 76542211
-1 | 8887666665444444333200
-0 | 887755554210
0 | 122233355556679
1 | 0001222223335667788999
2 | 133444556677889
3 | 001122244567888
4 | 02345678
5 | 011339
6 | 01266
7 |
8 | 9
```

MEAN	MEDIAN	SD	KURTOSIS	SKEWNESS
0.93474299	1.00416215	2.83087240	-0.23739941	0.00467346

[1] "Normal Test, with p-value = 0,05"

Shapiro-Wilk	Anderson-Darling	Cramer-von	Pearson
0.9149104	0.7335455	0.6711437	0.4174399
Shapiro-Francia	Jarque-Bera		
0.8664068	0.8922255		

[1] "OK" "OK" "OK" "OK" "OK" "OK"

Como se observa tanto en los gráficos con los resultados numéricos, los datos tienden a ser simétricos como se observa en el histograma, el grafico de cajas y patillas, gráficos de hojas y ramas y los valores de la mediana y media así como los valores de skewness, pero la variabilidad es alta. Aun asi, se observa colas largas y una alta variabilidad. Los test de normalidad pasan a un 95% Se observa estacionalidad con unos periodos pequeños pero mas homogéneos.

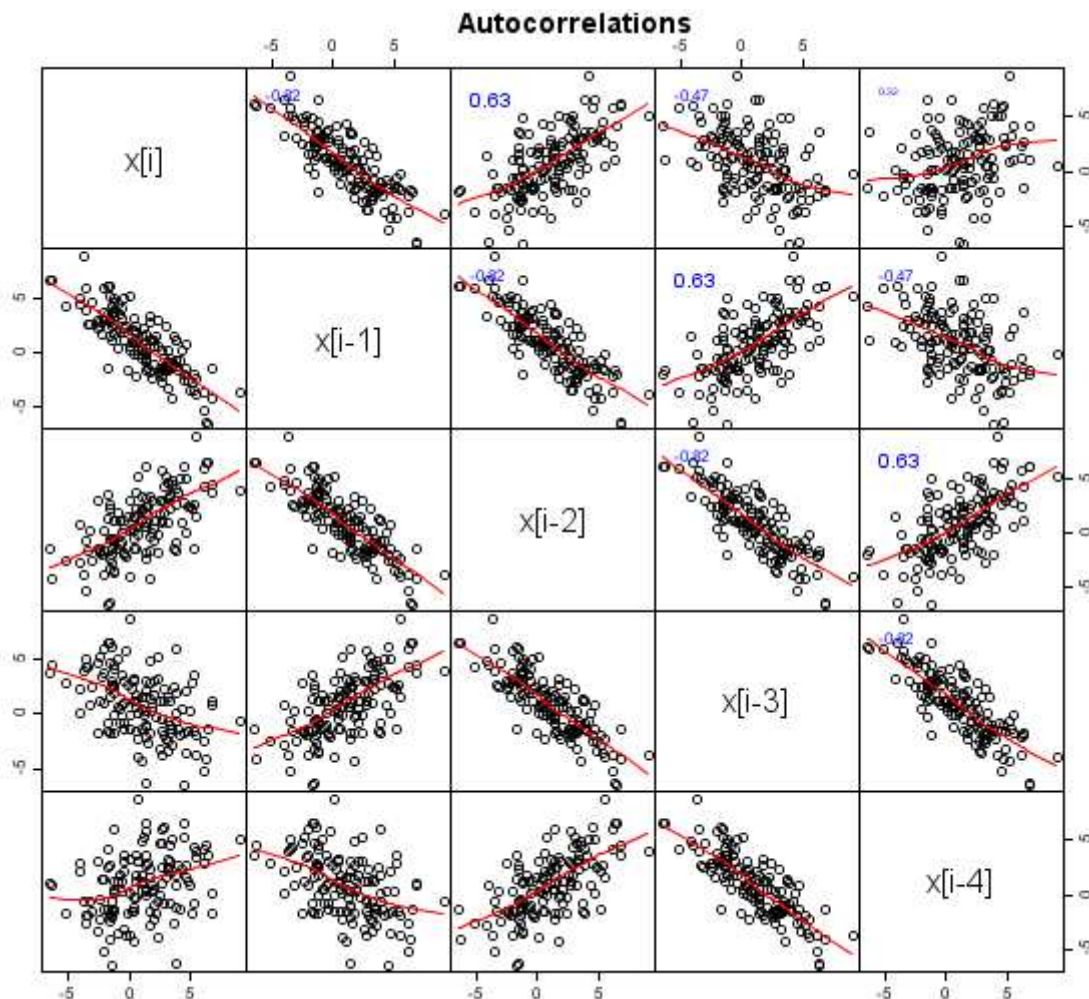
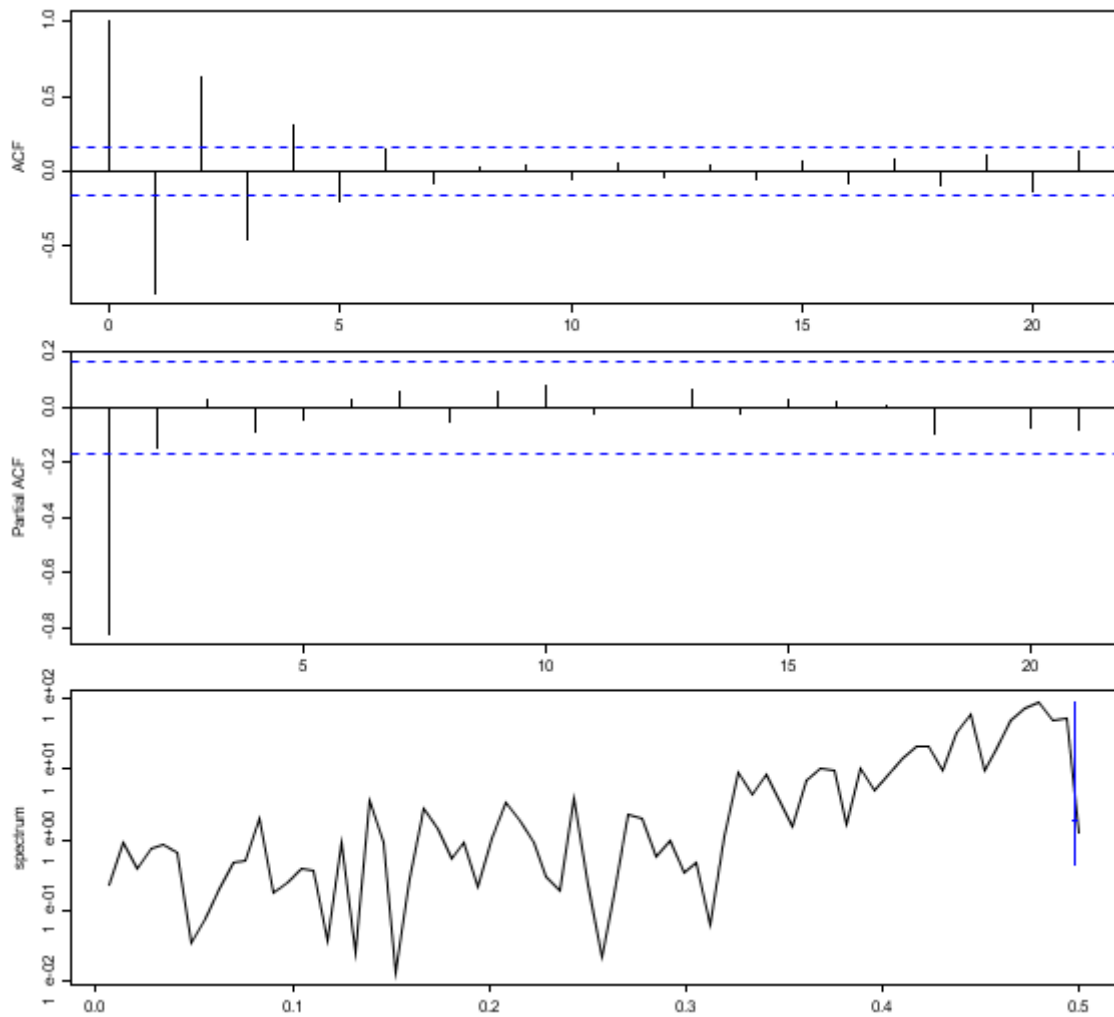


Figura 2.3

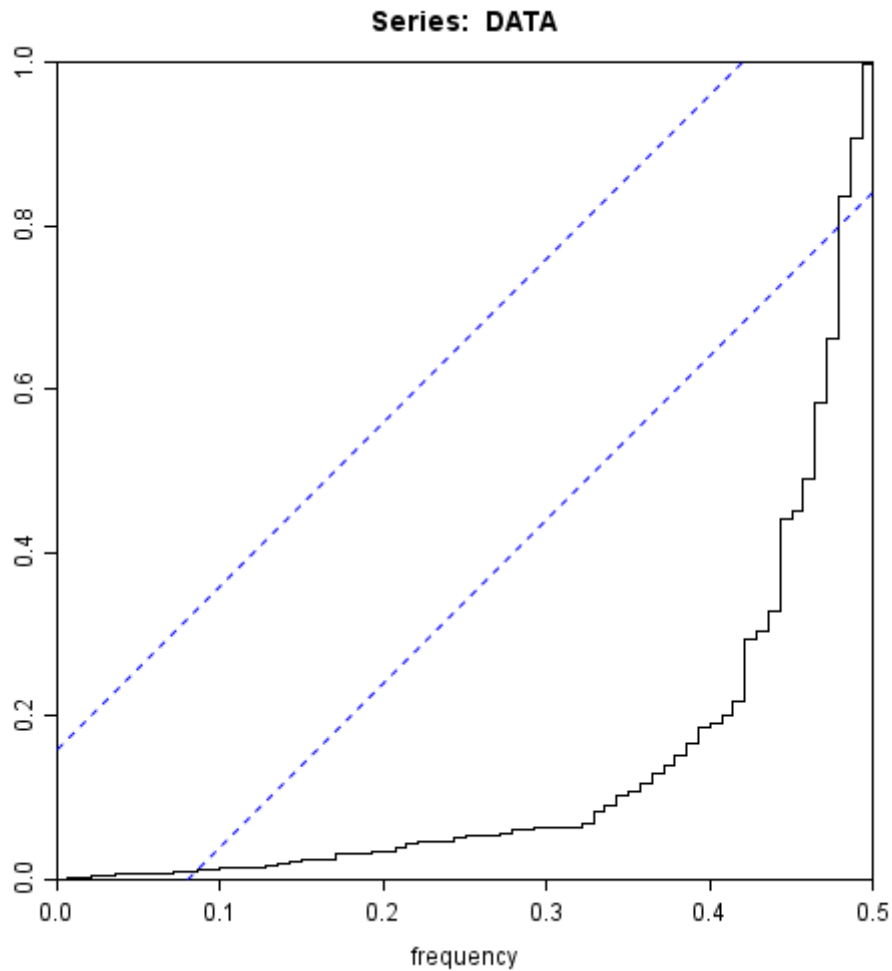
Como se observa en el grafico de autocorrelaciones, se observa correlacion, aunque no perfecta.



*Figura 2.4*

Como se observa, en los graficos de autocorrelacion total y parcial, nos encontramos ante otro modelo AR(1) debido a que el resultado de la funcion de autocorrelacion total oscila amortiguándose, mientras que la funcion de autocorrelacion parcial, solo tiene un valor no nulo.

El periodograma puestra el siguiente aspecto:



*Figura 2.5*

**NO SE INTERPRETARLO**

## Modelado de la serie temporal.

Como se observo en el anterior punto, los datos muestran una natural predisposición a un modelado a traves de un AR(1) debido a esto:



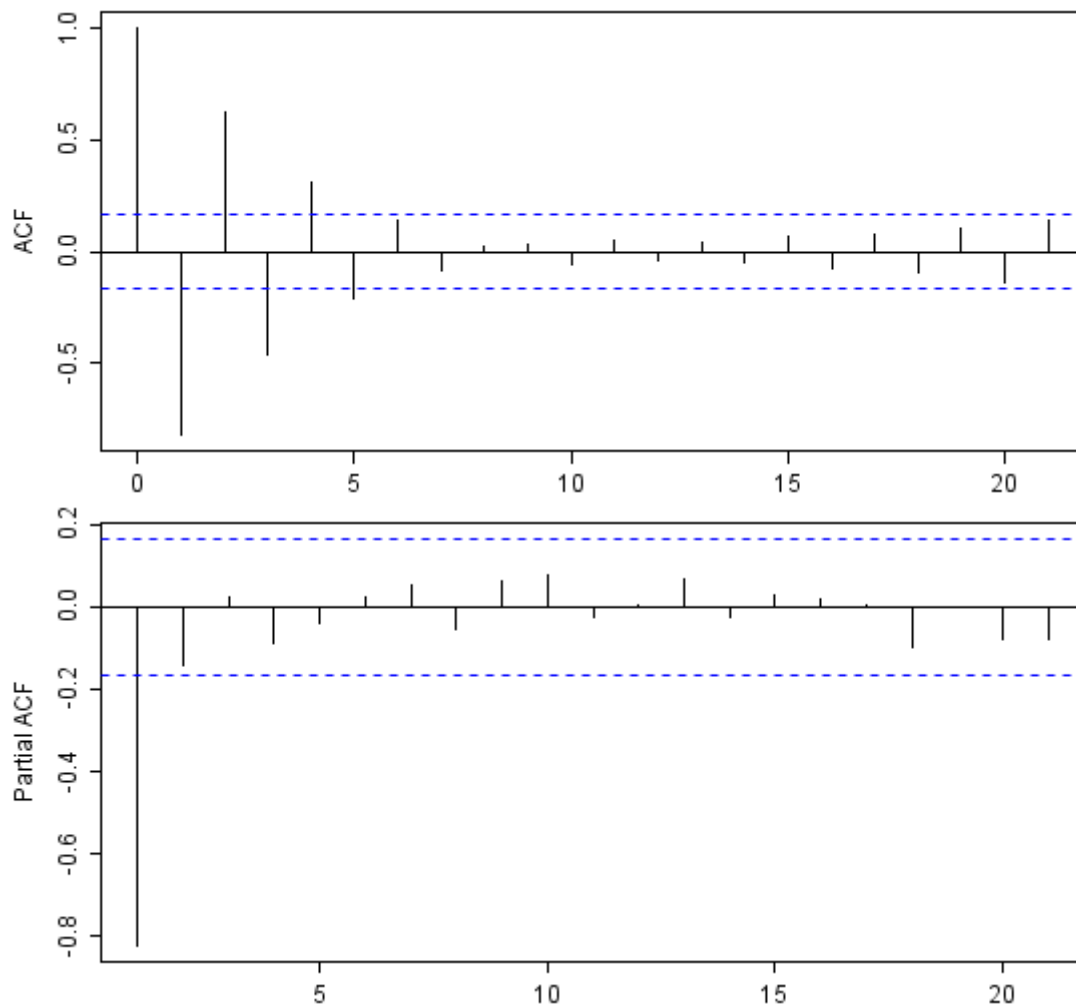


Figura 2.6

Como se observa, esto se parece bastante a la representación teorica de un AR(1):

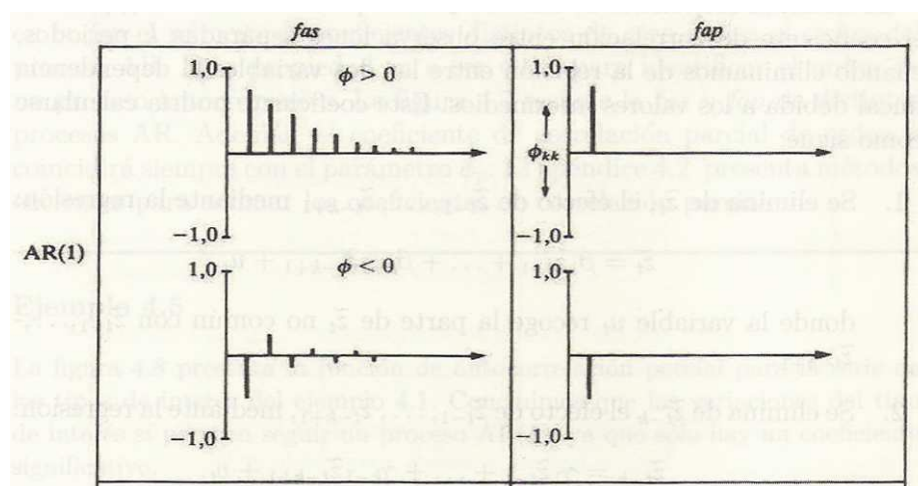


Figura 2.7

El modelo AR se define como:

$$X_t = \alpha_1 X_{t-1} + \dots + \alpha_{t-p} X_{t-p} + Z_t$$

$$d E(Z_t) = 0, \text{Var}(Z_t) = \sigma_z^2$$

Los resultados de la modelizacion, son los siguientes:

Call:

ar(x = DATOS.TS, aic = TRUE, order.max = 1)

Coefficients:

1  
-0.8218

Order selected 1 sigma^2 estimated as 2.621

Los datos de prediccion son:

\$pred

Time Series:

Start = 141

End = 170

Frequency = 1

[1] 1.2824962 0.6489699 1.1695827 0.7417589 1.0933315 0.8044197 1.0418387  
[8] 0.8467350 1.0070653 0.8753107 0.9835827 0.8946080 0.9677247 0.9076396  
[15] 0.9570157 0.9164399 0.9497839 0.9223828 0.9449002 0.9263961 0.9416022  
[22] 0.9291063 0.9393751 0.9309365 0.9378710 0.9321724 0.9368554 0.9330071  
[29] 0.9361695 0.9335707

\$se

Time Series:

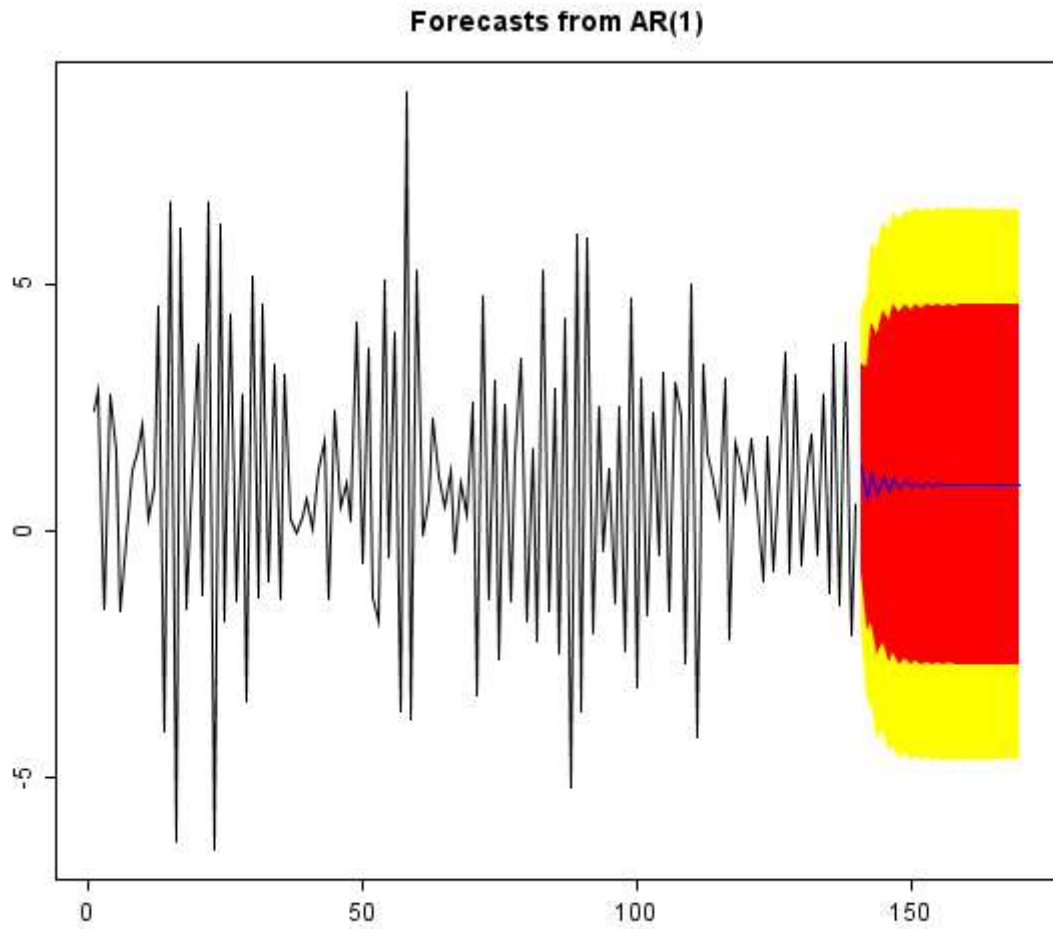
Start = 141

End = 170

Frequency = 1

[1] 1.618921 2.095427 2.363482 2.528476 2.634058 2.703026 2.748622 2.778990  
[9] 2.799312 2.812952 2.822125 2.828304 2.832468 2.835277 2.837173 2.838452  
[17] 2.839316 2.839899 2.840292 2.840558 2.840737 2.840859 2.840940 2.840996  
[25] 2.841033 2.841058 2.841075 2.841087 2.841095 2.841095

Por ultimo, el grafico de prediccion se muestra a continuaci3n:



*Figura 2.8*

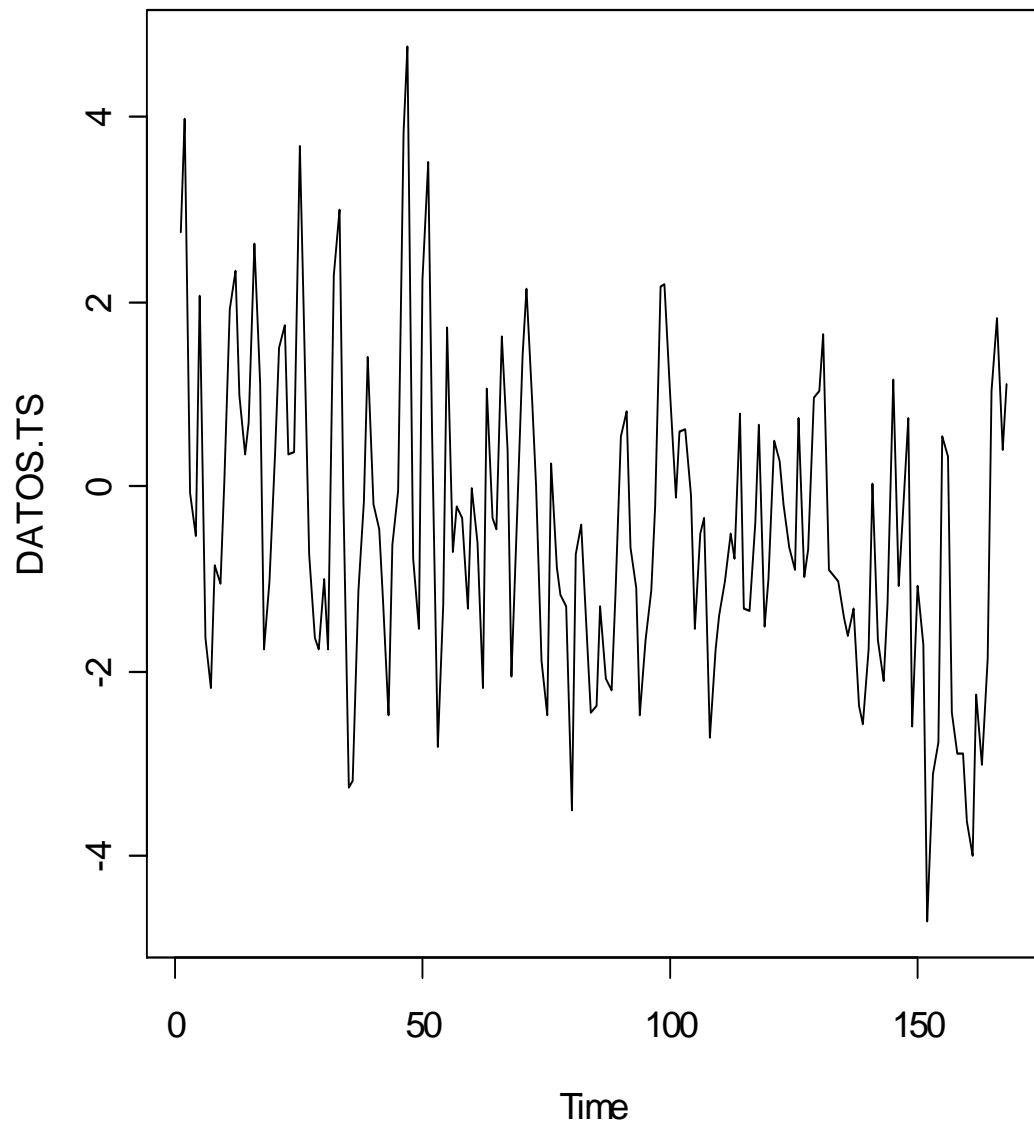
### 3.3 Serie 3

Dada la serie, se procederá a realizar el estudio de análisis de series temporales.

#### Análisis descriptivo de una serie temporal.

Todo análisis de serie temporal univariante, comienza con la presentación de un grafico donde se muestra la evolución de la variable a lo largo del tiempo.

Los resultados son los siguientes:



*Figura 3.1*

Como se observa en la grafica sacada a traves del Motor Estadístico R, los datos parecen estacionarios, como después podremos observar en el los resultados grafico-numericos del EDA.

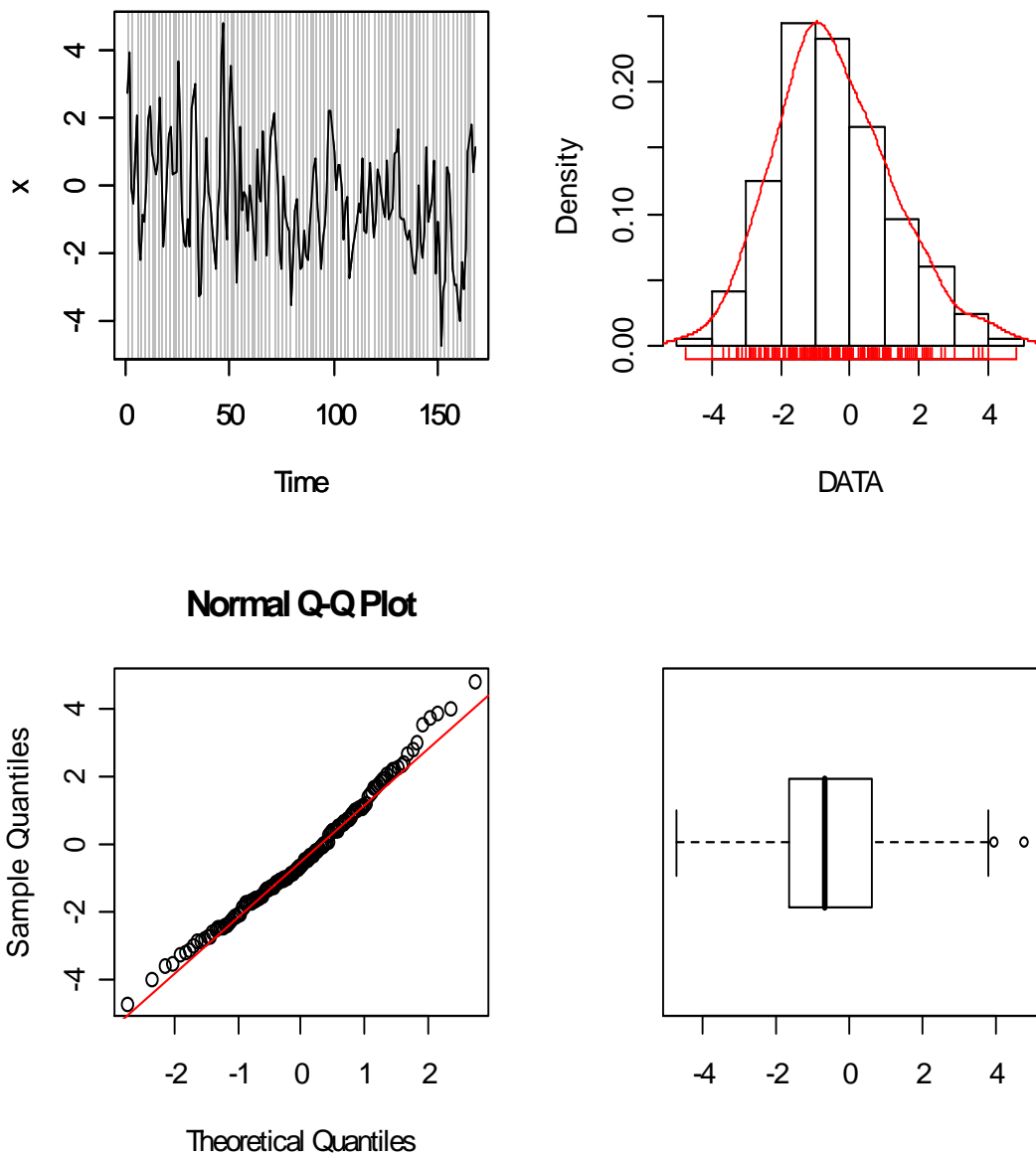


Figura 3.2

[1] "JAB: EDA VIEW 1.1"

The decimal point is at the |

```
-4 | 7
-4 | 0
-3 | 65
-3 | 3210
-2 | 99887665555
-2 | 4443222110
-1 | 98888887776666655
-1 | 4433333332221111110000000
-0 | 9999887777776655555
-0 | 444333322222211100
0 | 001333444444
0 | 5566667777889
```

```

1 | 00000111144
1 | 56778899
2 | 1122233
2 | 67
3 | 0
3 | 578
4 | 0
4 | 8

```

```

      MEAN   MEDIAN      SD  KURTOSIS  SKEWNESS
-0.4399769 -0.6422915  1.6875614  0.1419879  0.4134895
[1] "Normal Test, with p-value = 0,05"
      Shapiro-Wilk Anderson-Darling      Cramer-von      Pearson
      0.14289080      0.10333599      0.08937338      0.70170317
Shapiro-Francia      Jarque-Bera
      0.12054569      0.07808588
[1] "OK" "OK" "OK" "OK" "OK" "OK"

```

Como se observa claramente en el grafico de cajas y patillas, existe una serie de outliers que pueden distorsionar el análisis EDA. Si se observan los resultados de los test de normalidad, todos ellos fallan y a excepción del test de Pearson, rechazan la hipótesis nula a un 85% aceptan que los datos se aproximan a una normal. Se observa una cierta simetría en los datos, hecho que se demuestra empíricamente a través de la observación del Histograma.

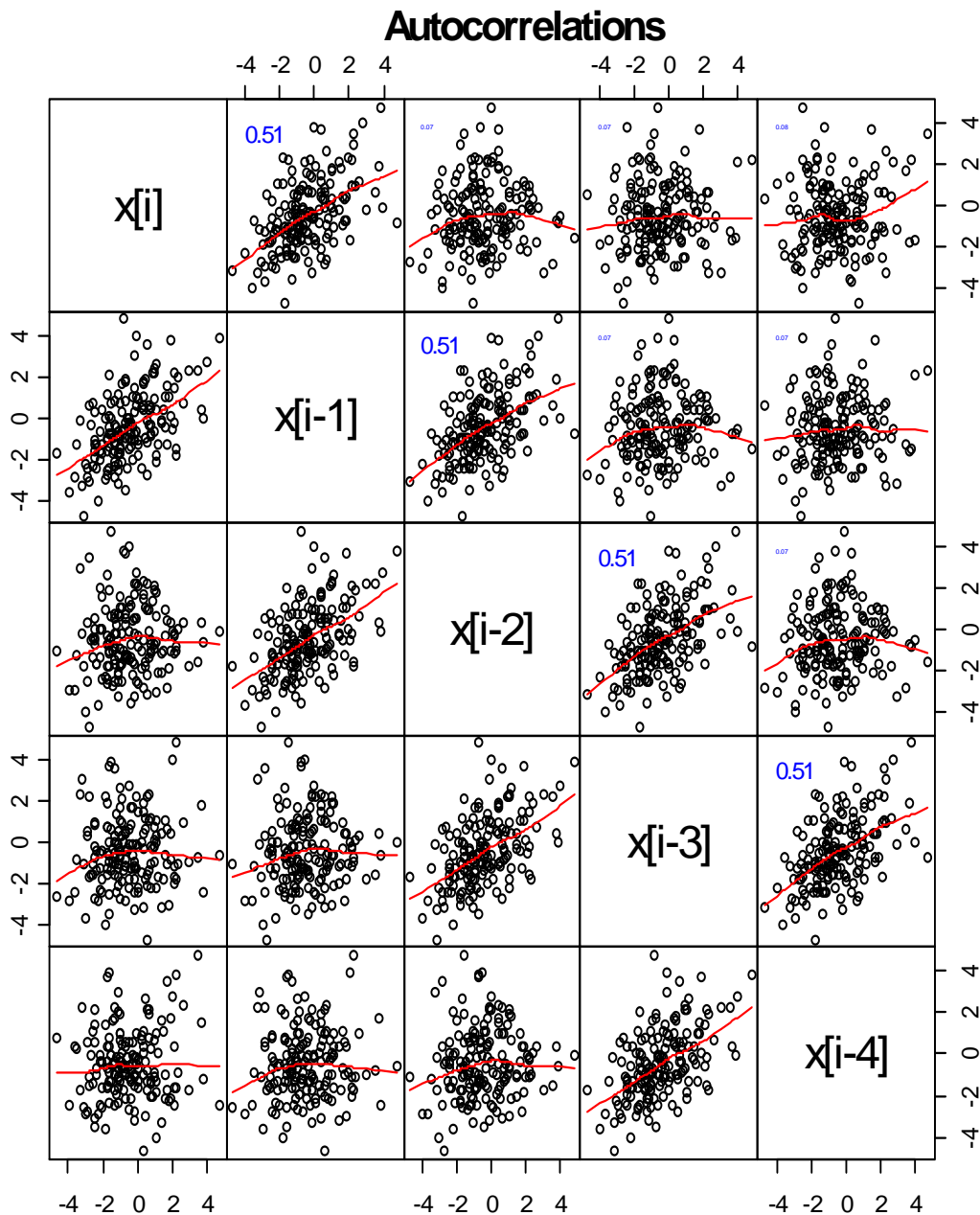
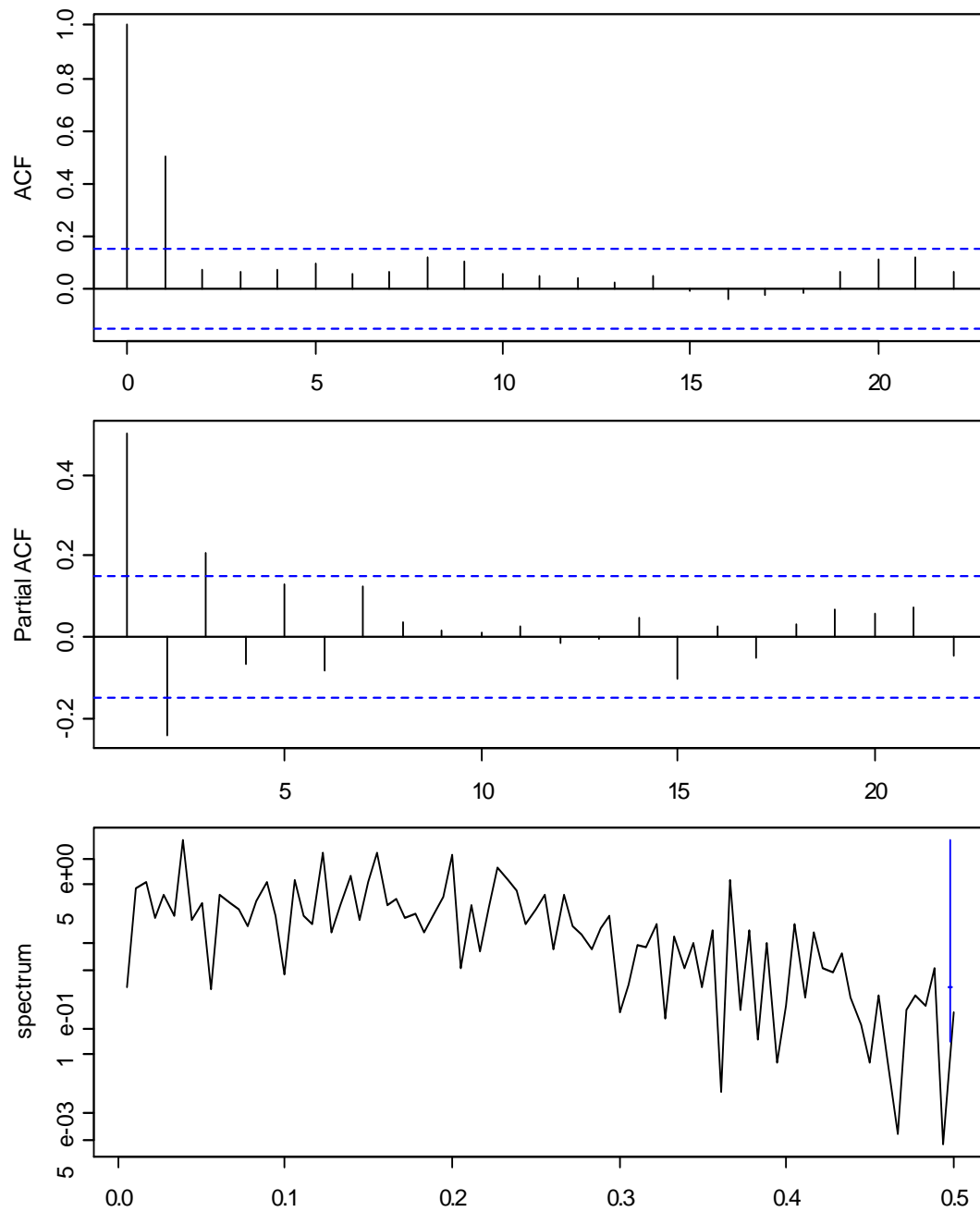


Figura 3.3

Si observamos las autocorrelaciones, podemos observar como no estan excesivamente correlados los datos propuestos. Esto se ratifica en los coeficientes dados ya que las autocorrelaciones son como mucho de un 51%

A continuación procederemos a analizar las funciones de autocorrelacion total, parcial y espectrograma.



*Figura 3.4*

Como se puede apreciar tras evaluar detenidamente los graficos, los datos se podrían modelizar como un MA(2) debido a la forma que se observa en la funcion de autocorrelacion parcial y total:



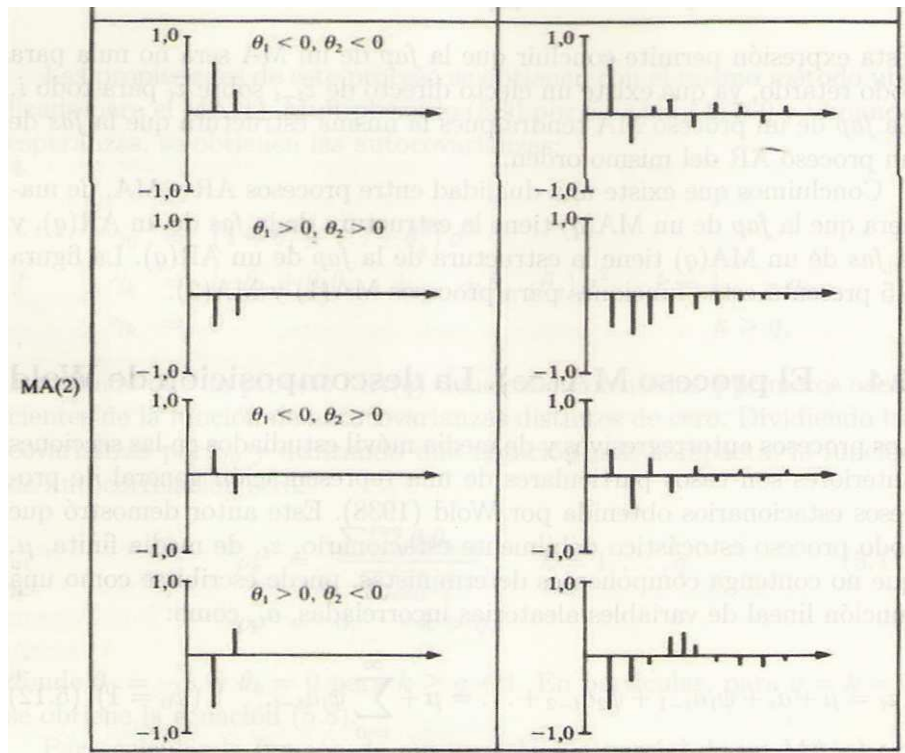
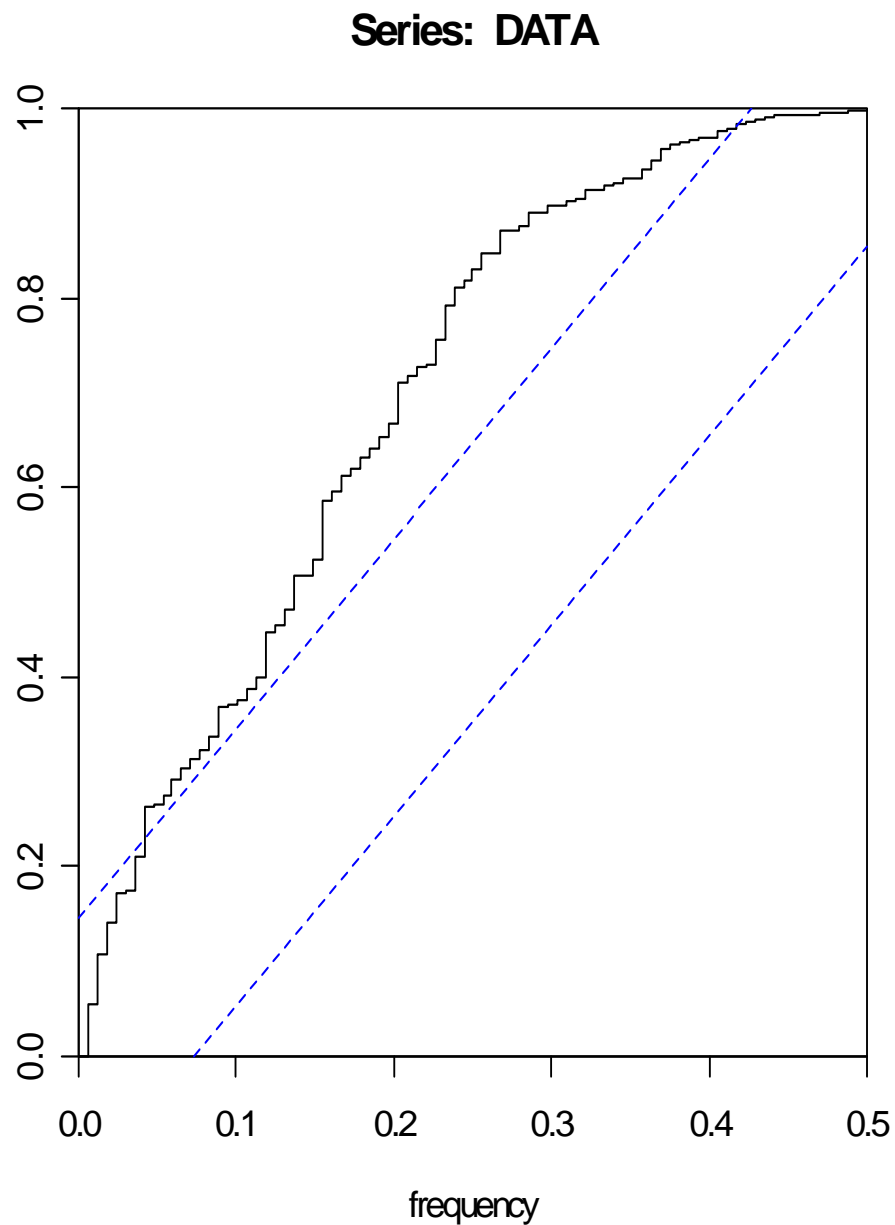


Figura 3.4

El periodograma de los datos es:

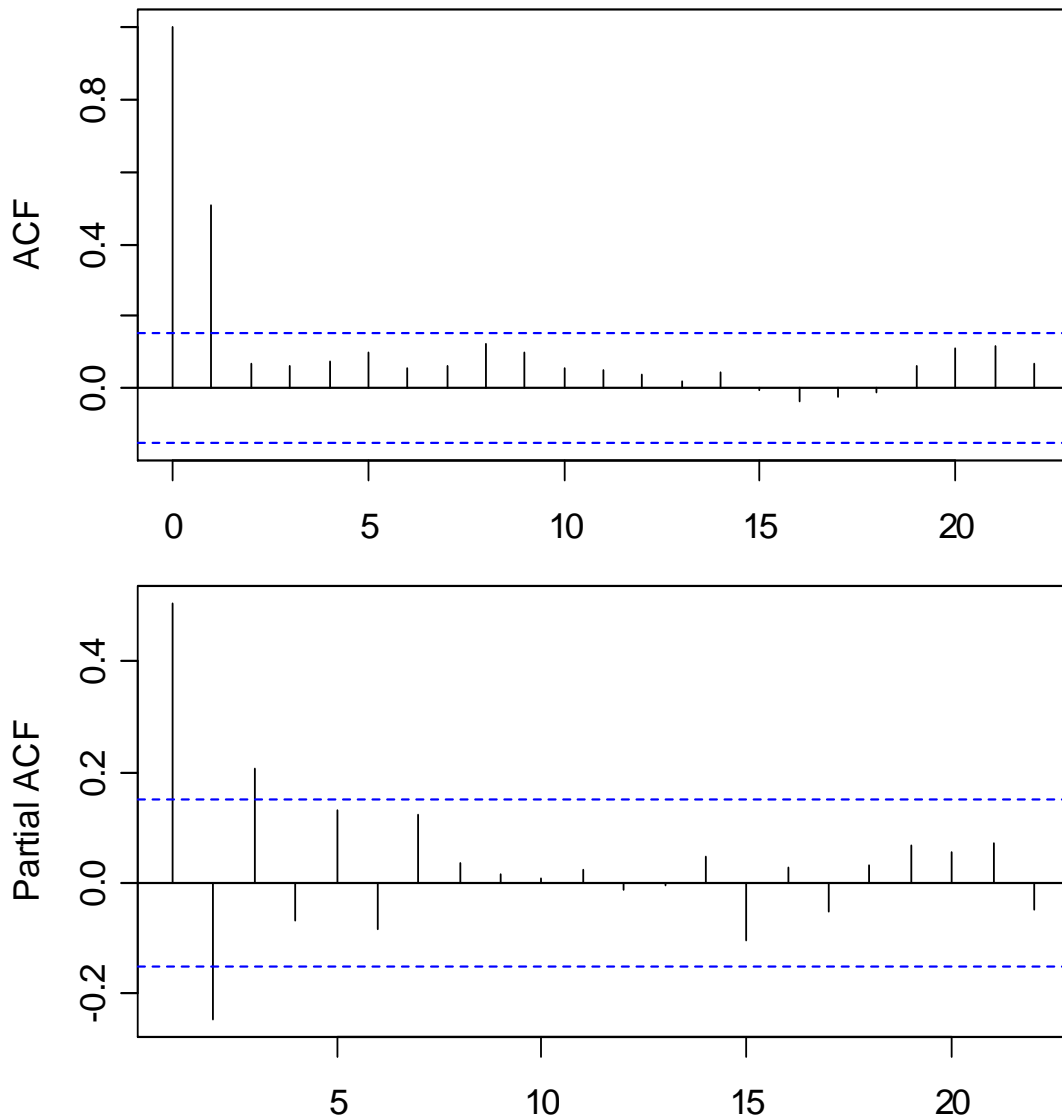


*Figura 3.5*

No se explicar el significado.

### Modelado de la serie temporal.

Como se observo en el anterior punto, los datos muestran una natural predisposición a un modelado a traves de un MA(2) debido a esto:



Por tanto, modelaremos los datos a través de dicho modelo teórico

$${}^tX = a + {}^tW + \sum_{k=1}^q \beta_k {}^{t-k}W$$

```
DATOS.TS.MA <- arma(DATOS.TS, order = c(0,2));
DATOS.TS.MA;
DATOS.TS.MA$coef
DATOS.TS.MA.PREDICTION_30 <- predict(DATOS.TS.MA, n.ahead=30);
DATOS.TS.MA.PREDICTION_30;
plot(forecast(DATOS.TS.MA,30,conf=c(80,95)));
```

NOTA: He tenido que modelar un proceso MA como un ARMA con parámetro p nulo debido a que no he encontrado la función Ma en R en la actualidad.

Los resultados del modelado son los siguientes:

```
arma(x = DATOS.TS, order = c(0, 2))
```

Coefficient(s):

ma1	ma2	intercept
0.69904	0.03027	-0.45734

Los datos de prediccion:

\$pred

Time Series:

Start = 169

End = 198

Frequency = 1

```
[1] 1.830747 2.542203 3.253658 3.965114 4.676569 5.388025 6.099481
[8] 6.810936 7.522392 8.233847 8.945303 9.656759 10.368214 11.079670
[15] 11.791126 12.502581 13.214037 13.925492 14.636948 15.348404 16.059859
[22] 16.771315 17.482770 18.194226 18.905682 19.617137 20.328593 21.040048
[29] 21.751504 22.462960
```

\$se

Time Series:

Start = 169

End = 198

Frequency = 1

```
[1] 2.429948 5.433530 9.092034 13.309376 18.020980 23.180231
[7] 28.751537 34.706605 41.022251 47.679031 54.660338 61.951772
[13] 69.540692 77.415884 85.567313 93.985932 102.663523 111.592587
[19] 120.766236 130.178122 139.822365 149.693501 159.786431 170.096389
[25] 180.618901 191.349759 202.284996 213.420863 224.753809 236.280467
```

**NOTA: No hay grafico de prediccion debido que me falla la funcion por algun motivo.**

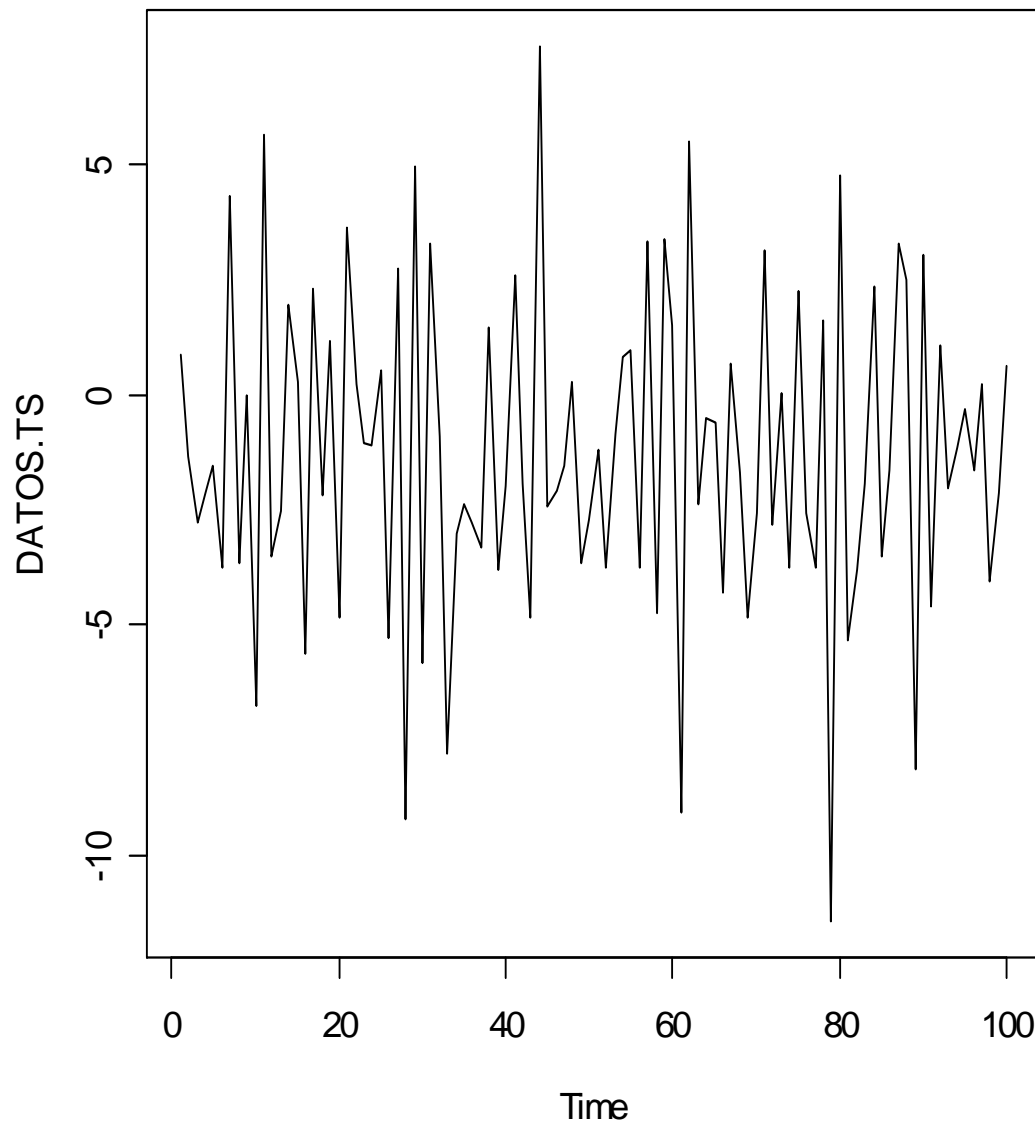
## 3.4 Serie 4

Dada la serie, se procederá a realizar el estudio de análisis de series temporales.

### Análisis descriptivo de una serie temporal.

Todo análisis de serie temporal univariante, comienza con la presentación de un grafico donde se muestra la evolución de la variable a lo largo del tiempo.

Los resultados son los siguientes:



*Figura 4.1*

Como se observa en los datos procedentes de la cuarta muestra, los datos parecen estacionarios aun valor que luego calcularemos. Por otro lado, se observa una gran variabilidad.

Los resultados de un análisis exploratorio de datos son los siguientes:

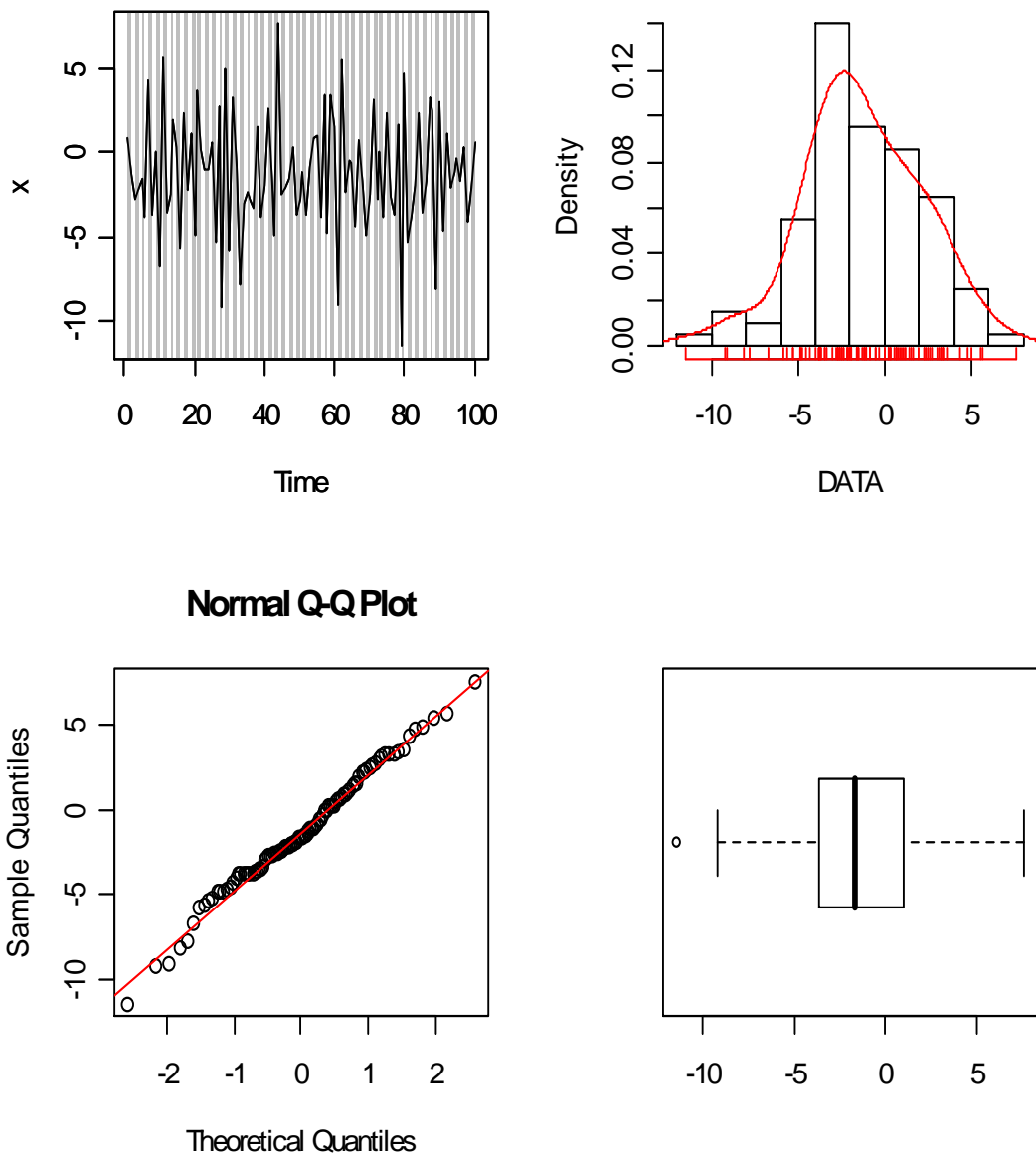


Figura 4.2

[1] "JAB: EDA VIEW 1.1"

The decimal point is at the |

```
-10 | 4
-8 | 211
-6 | 88
-4 | 86338887631
-2 | 88888877755308777655443211100
-0 | 997665532111986530
0 | 02333667890125569
2 | 3345670133346
4 | 38057
6 | 6
```

MEAN	MEDIAN	SD	KURTOSIS	SKEWNESS
------	--------	----	----------	----------

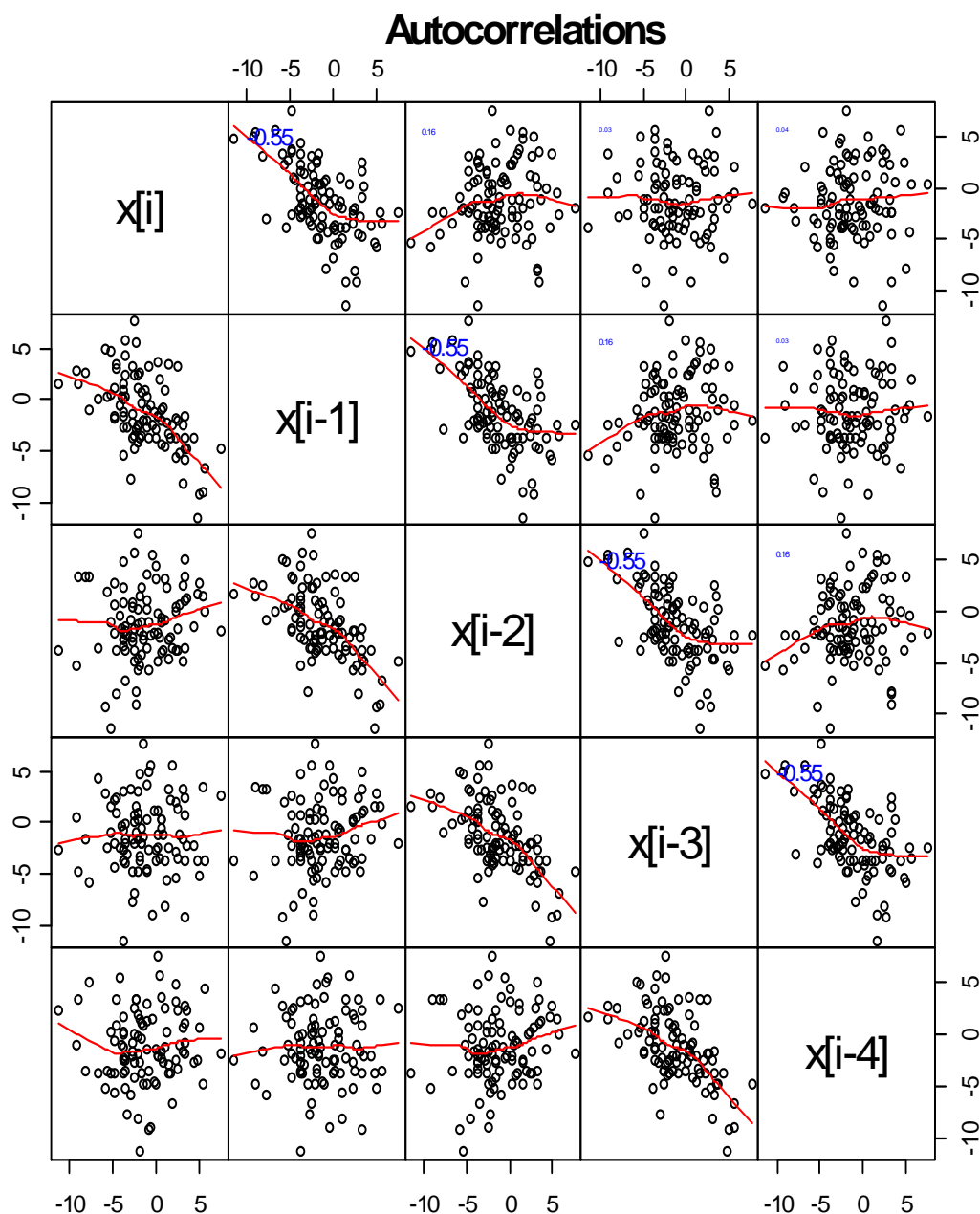
-1.27413859 -1.63717338 3.46255901 0.13591643 -0.08991347

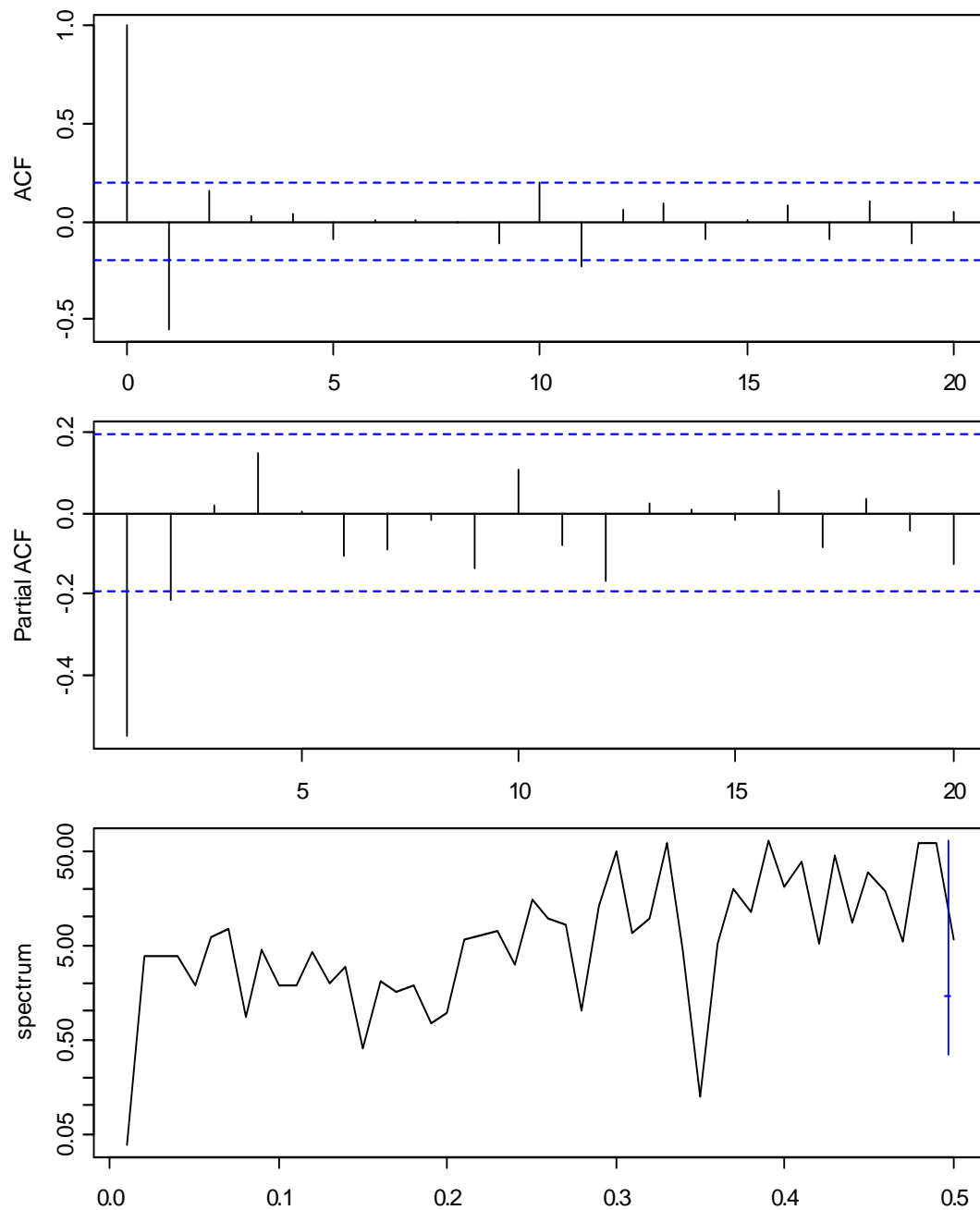
[1] "Normal Test, with p-value = 0,05"

Shapiro-Wilk	Anderson-Darling	Cramer-von	Pearson
0.6768497	0.4039575	0.3665614	0.1449581
Shapiro-Francia	Jarque-Bera		
0.4918085	0.8586250		

[1] "OK" "OK" "OK" "OK" "OK" "OK"

Como se observa en los resultados del análisis, se observa una muy alta variabilidad y hasta un dato extremo. Sería interesante estudiar la calidad del muestreo o si es un dato real. Por otro lado, los resultados de la media y mediana son contrarios, aunque debido a la alta variabilidad, me inclinaría a usar como indicador medio, la mediana, debido a la alta variabilidad.



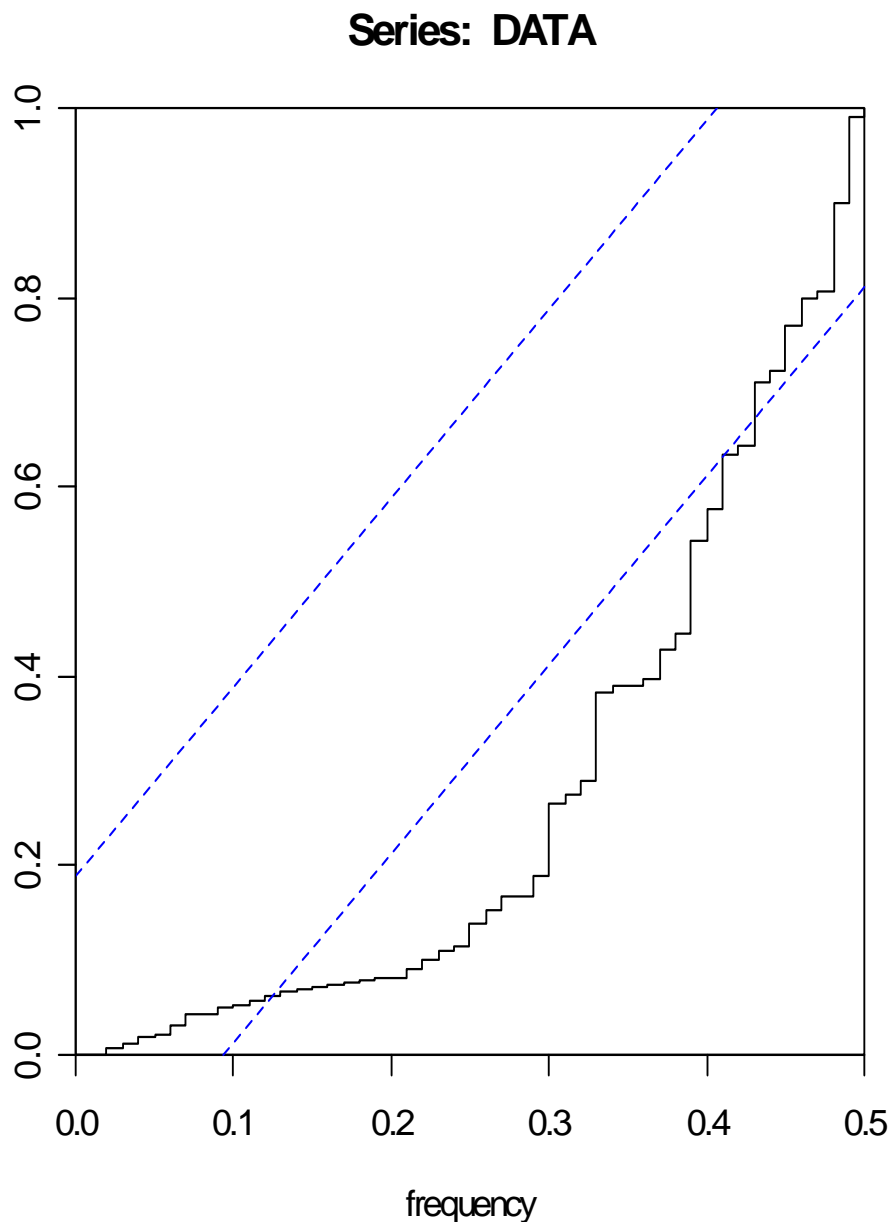


### VER PERIODOGRAMA

Observando los resultados, se puede apreciar que tanto la funcion de autocorrelacion parcial y total tiene coeficientes no nulos, esto indica que seria conveniente modelarla como un proceso ARMA. En este caso podria ser un ARMA(2,2) o un ARMA(2,1)

TIPO	FAS	FAP
AR(p)	Muchos coeficientes no nulos	Primeros p no nulos, resto 0
MA(q)	Primeros p no nulos, resto 0	Muchos coeficientes no nulos
ARMA(p,q)	Muchos coeficientes no nulos	Muchos coeficientes no nulos





## Modelado de la serie temporal.

Como se observo en el anterior punto, los datos muestran una natural predisposición a un modelado a traves de un ARMA(2,2) o un ARMA(2,1) debido a esto:

Un modelo ARMA(p,q) se modela de la siguiente forma:

$${}^tX = a + {}^tW + \sum_{k=1}^q \beta_k {}^{t-k}W + \sum_{k=1}^p b_k {}^{t-k}X$$

### Caso 1: ARMA(2,2):

Los resultados son:

```
DATOS.TS.ARMA <- arma(DATOS.TS, order = c(2,2));  
summary(DATOS.TS.ARMA);
```

Call:

```
arma(x = DATOS.TS, order = c(2, 2))
```

Model:

ARMA(2,2)

Residuals:

Min	1Q	Median	3Q	Max
-8.823	-1.913	0.232	2.049	6.314

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t )
ar1	-0.02275	0.22390	-0.102	0.919062
ar2	-0.23254	0.16271	-1.429	0.152955
ma1	-0.64907	0.19297	-3.364	0.000769 ***
ma2	0.53315	0.15989	3.335	0.000854 ***
intercept	-1.65982	0.46453	-3.573	0.000353 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Fit:

sigma<sup>2</sup> estimated as 7.692, Conditional Sum-of-Squares = 746.18, **AIC = 497.81**

## Caso 2: ARMA(2,1):

Los resultados son:

Call:

```
arma(x = DATOS.TS, order = c(2, 1))
```

Model:

ARMA(2,1)

Residuals:

Min	1Q	Median	3Q	Max
-8.6672	-2.0658	0.3841	1.8978	6.8712

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t )
ar1	-0.7320	0.2886	-2.537	0.0112 *
ar2	-0.2472	0.1761	-1.404	0.1603
ma1	0.0574	0.2879	0.199	0.8420
intercept	-2.5763	0.6615	-3.894	9.84e-05 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Fit:

sigma<sup>2</sup> estimated as 8.022, Conditional Sum-of-Squares = 778.11, **AIC = 500**

Como se observa en los resultados el modelo ARMA(2,2) es el que mejor se ajusta, debido a que tiene un AIC menor. El Criterio de Akaike, se define:

$$AIC(M) = T \log(\tilde{\sigma}_a^2) + 2M$$

El criterio de información de Akaike pondera entre la función logaritmo de máxima verosimilitud, usando la varianza residual, y el número de parámetros en el modelo. El modelo a escoger es donde AIC es mínimo.

NOTA: No tengo la predicción, puesto que no la encuentro, me da que la voy a tener que programar. El gráfico no es problema cuando tenga los datos.

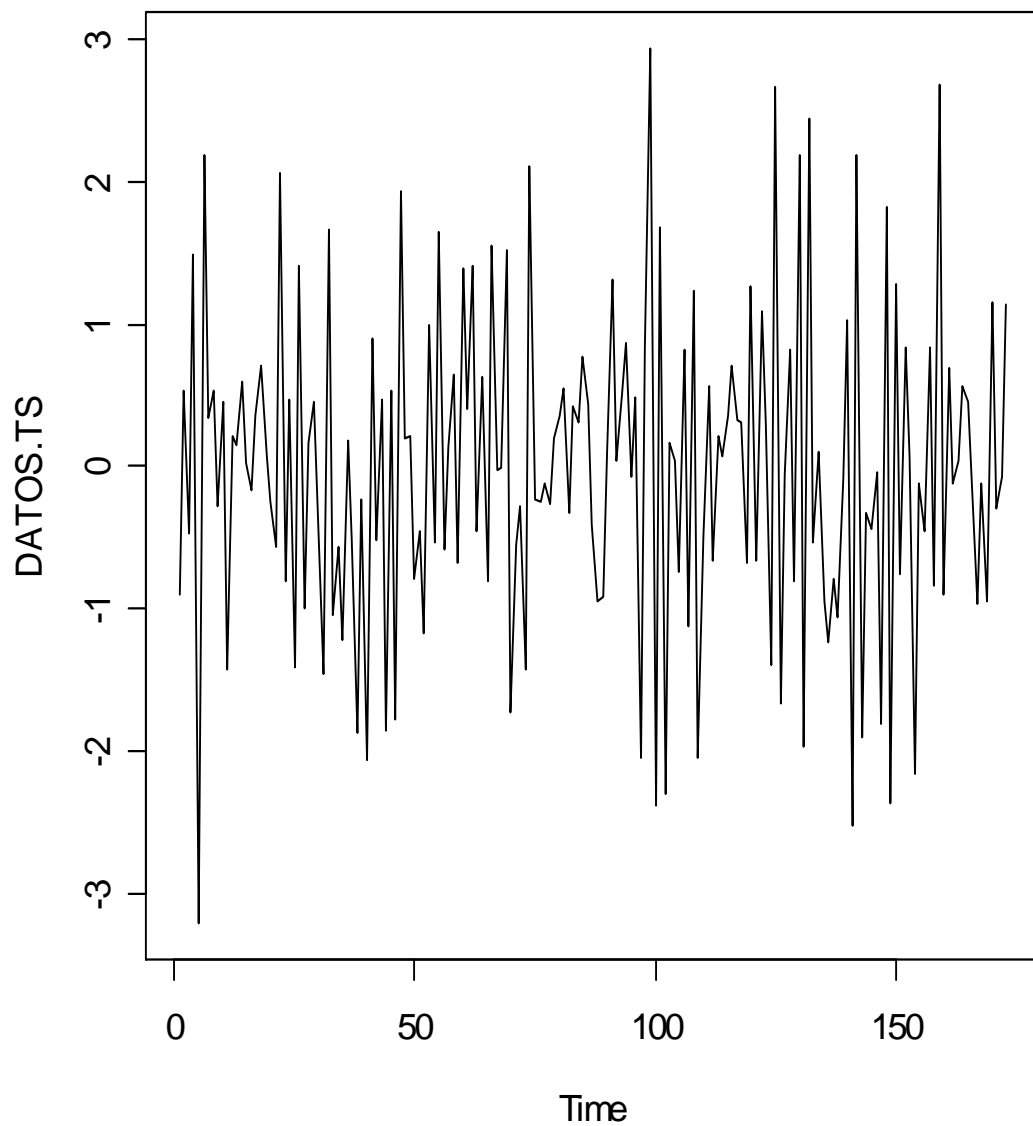
### 3.5 Serie 5

Dada la serie, se procederá a realizar el estudio de análisis de series temporales.

#### Análisis descriptivo de una serie temporal.

Todo análisis de serie temporal univariante, comienza con la presentación de un gráfico donde se muestra la evolución de la variable a lo largo del tiempo.

Los resultados son los siguientes:



*Figura 5.1*

Al igual que la serie 4, esta serie presenta a simple vista una alta variabilidad en los datos así como un comportamiento a estudiar a partir del dato 80 aproximadamente. Los resultados del Análisis Exploratorio de datos es el siguiente:

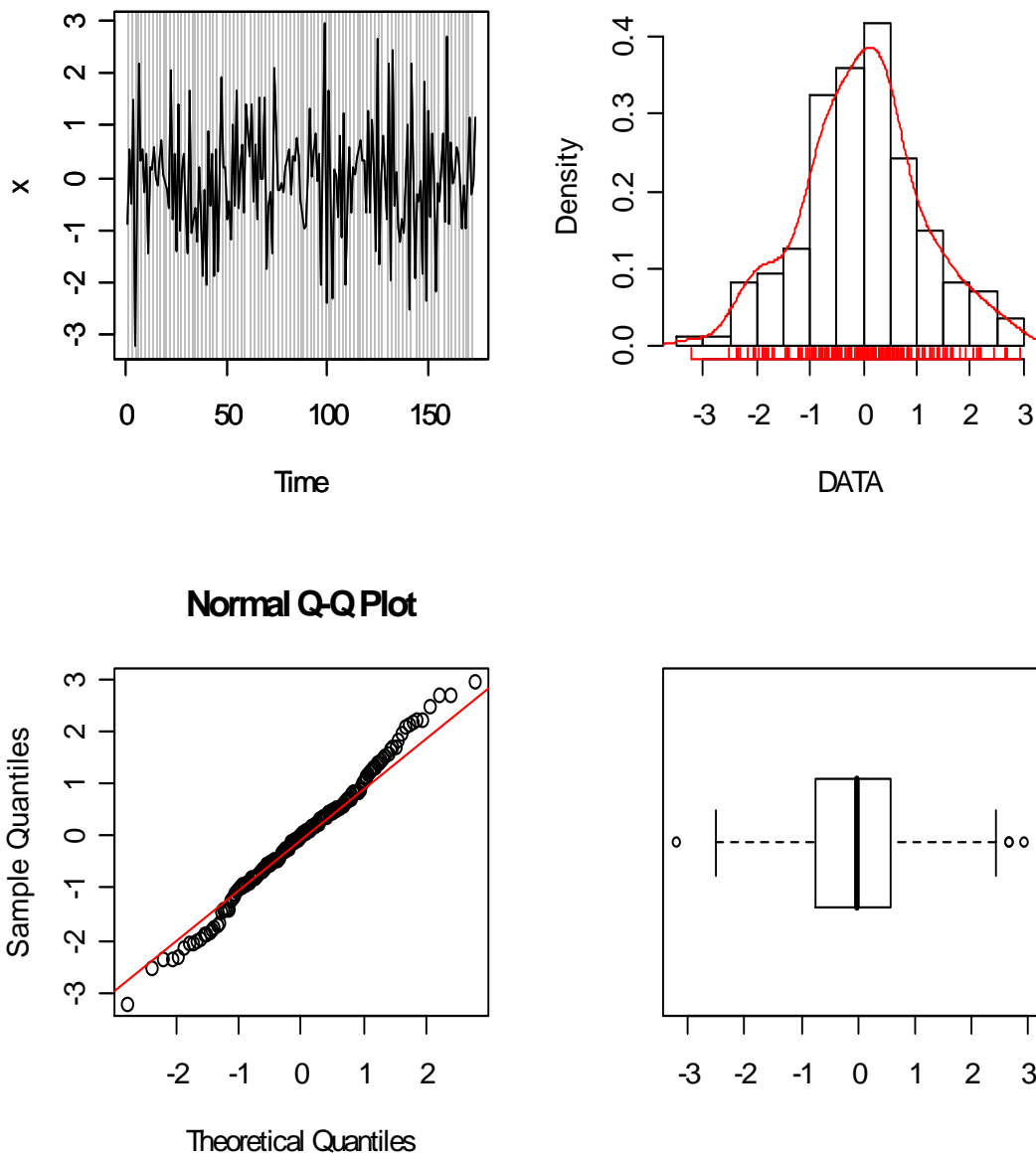


Figura 5.2

[1] "JAB: EDA VIEW 1.1"

The decimal point is at the |

```
-3 | 2
-2 | 5
-2 | 44321000
-1 | 99988775
-1 | 444422211000
-0 | 99999988888887777776666555555555
-0 | 4433333322222111111110000
0 | 00001111112222222233333444444
0 | 555555555556666677788888899
1 | 001112333444
1 | 55667789
2 | 112224
```

```

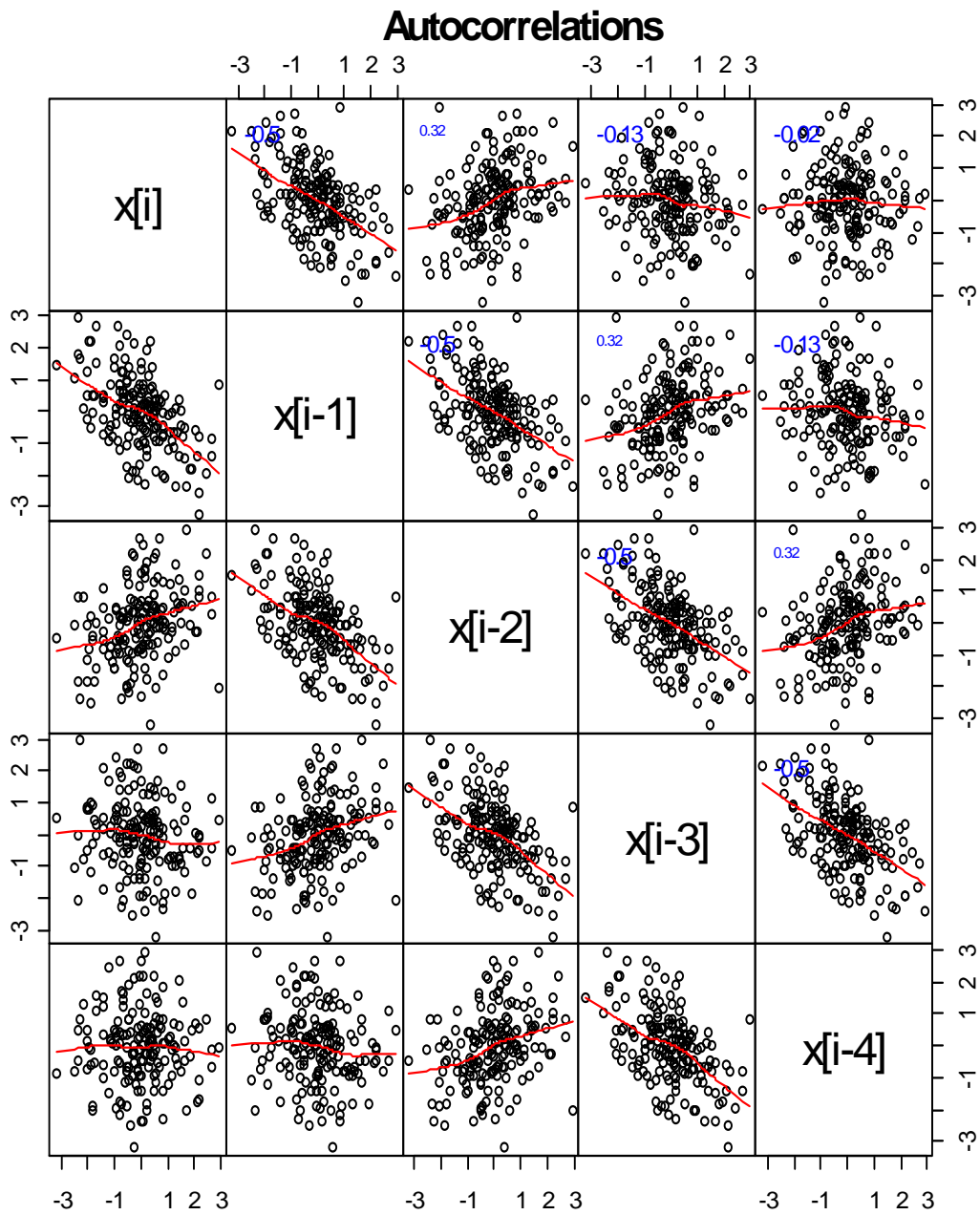
      MEAN      MEDIAN      SD      KURTOSIS      SKEWNESS
-0.029547523 -0.002593627 1.130702168 0.054315641 0.025951189
[1] "Normal Test, with p-value = 0,05"
      Shapiro-Wilk Anderson-Darling      Cramer-von      Pearson
      0.6312825      0.3566558      0.3487105      0.6456378
Shapiro-Francia      Jarque-Bera
      0.6020619      0.9617214
[1] "OK" "OK" "OK" "OK" "OK" "OK"

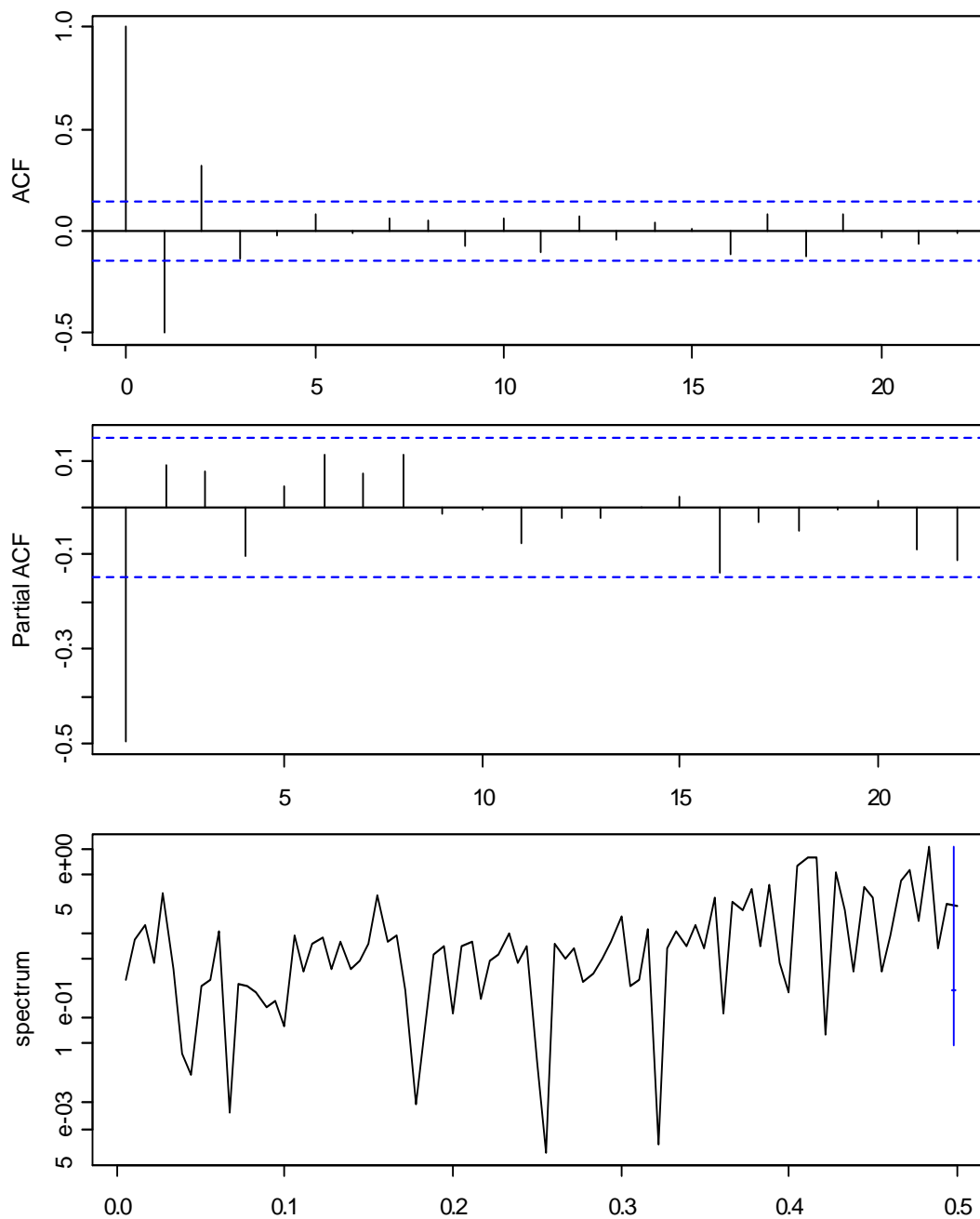
```

Como se observa en los gráficos y resultados, los datos son muy prometedores debido a que se observa simetría y la variabilidad no era tan alta como aparentemente se pensaba. Tras evaluar estos datos, considero que estos datos pueden ser seguir siendo analizados. De todas formas, se han detectado algunos valores anómalos en las colas. Mi pregunta es: ¿Es necesario borrarlos?

De todas formas, la descomposición clásica mediante medias móviles no ha sido posible realizarla debido a que no soy capaz de determinar la estacionalidad. Este procedimiento lo he realizado en R y Statgraphics.

Los test de normalidad rechazan la hipótesis nula a un 45%, esto me preocupa.





Como se observa en el grafico, se observa varios coeficientes no nulos al principio en el grafico de la funcion de autocorrelacion parcial, mientras que en el caso de la parcial se observa uno solo, esto indica, las siguientes posibilidades:

1. AR(2)
2. MA(1)
3. ARMA(2,1)

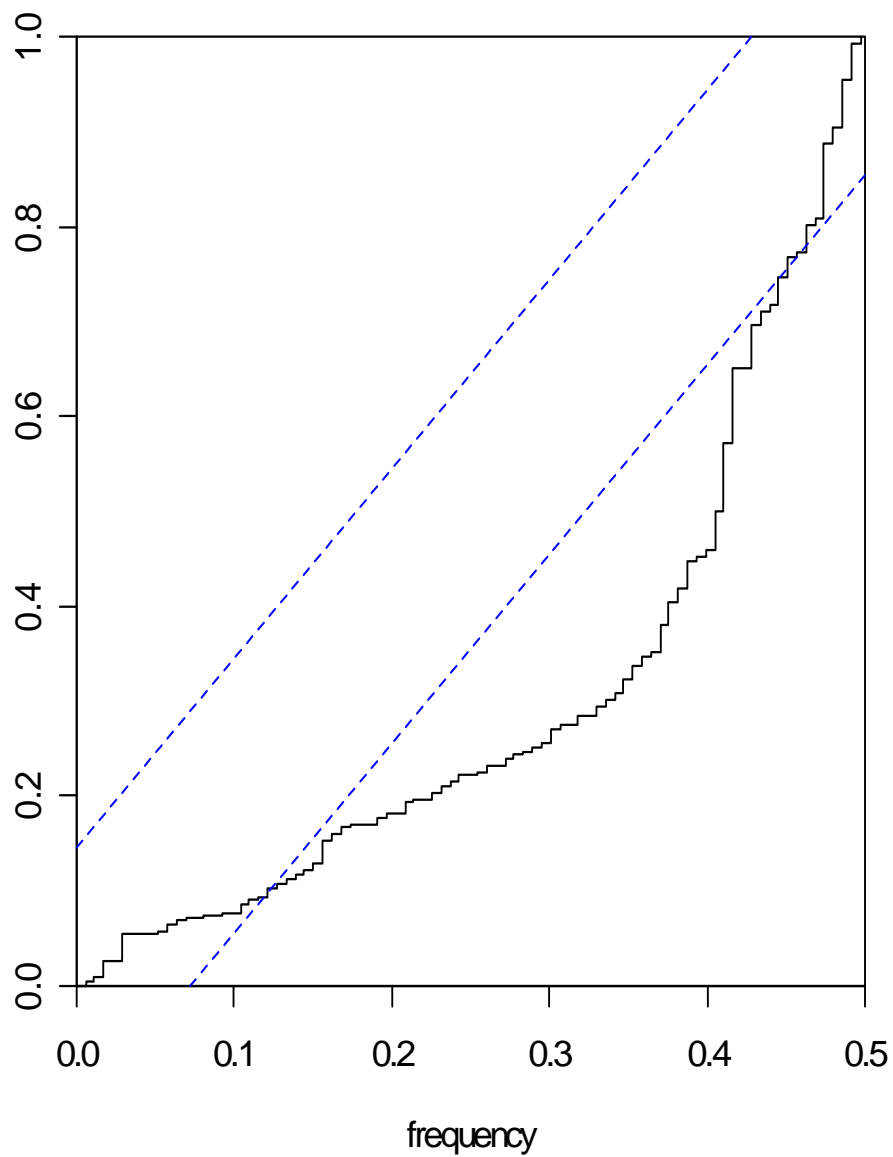
Mas adelante observaremos, cual es el mejor modelo.

TIPO	FAS	FAP
AR(p)	Muchos coeficientes no nulos	Primeros p no nulos, resto 0
MA(q)	Primeros p no nulos, resto 0	Muchos coeficientes no nulos



ARMA(p,q)	Muchos coeficientes no nulos	Muchos coeficientes no nulos
-----------	------------------------------	------------------------------

### Series: DATA



Ni idea

### Modelado de la serie temporal.

Como se observo en el anterior punto, los datos muestran una natural predisposición a las siguientes opciones:

1. AR(2)
2. MA(1)
3. ARMA(2,1)

A continuación, analizaremos cada una de las opciones y nos decidiremos por aquella opcion con un AIC menor.

### Caso 1: AR(2)

```
DATOS.TS.AR <- ar(DATOS.TS, aic = TRUE);# , order.max=1
DATOS.TS.AR;
DATOS.TS.AR$aic
```

Call:

```
ar(x = DATOS.TS, aic = TRUE, order.max = 1)
```

Coefficients:

```
1
-0.498
```

Order selected 1 sigma^2 estimated as 0.967

```
0 1
AIC: 47.31628 0.00000
```

Como se observa, en los resultados, el mejor modelo AR para los datos es de orden 1, como bien se habia analizado empíricamente. La funcion ar, si no le das ningun limite de orden, te calcula hasta orden 23. el mejor modelo, es el de orden 1.

### Caso 1: MA(1)

En este caso, realizaremos un modelo MA(1) a traves de la equivalencia con un modelo ARMA(0,1)

Call:

```
arma(x = DATOS.TS, order = c(0, 1))
```

Model:

ARMA(0,1)

Residuals:

Min	1Q	Median	3Q	Max
-2.6439	-0.6372	0.0811	0.7163	3.0406

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t )
ma1	-0.37966	0.05738	-6.617	3.67e-11 ***
intercept	-0.02788	0.04826	-0.578	0.563

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Fit:

sigma^2 estimated as 1.047, Conditional Sum-of-Squares = 178.97, **AIC = 502.83**

```
> DATOS.TS.MA$coef
ma1 intercept
-0.37965561 -0.02788308
>
```

### Caso 3: ARMA(2,1)

Call:

```
arma(x = DATOS.TS, order = c(2, 1))
```

Model:

ARMA(2,1)

Residuals:

Min	1Q	Median	3Q	Max
-2.4894	-0.6177	0.1116	0.6955	3.4631

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t )
ar1	0.279112	0.189538	1.473	0.141
ar2	0.445902	0.096695	4.611	4.00e-06 ***
ma1	-0.751678	0.191151	-3.932	8.41e-05 ***
intercept	-0.008156	0.019907	-0.410	0.682

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Fit:

sigma^2 estimated as 0.9539, Conditional Sum-of-Squares = 162.16, **AIC = 490.78**

Como se puede observar el resultado indica que el mejor modelo, es el AR(1) debido a que tiene un AIC menor.

Los datos de prediccion son:

\$pred

Time Series:

Start = 174

End = 203

Frequency = 1

```
[1] -0.60902872 0.25905203 -0.17327902 0.04203520 -0.06519793 -0.01179252
[7] -0.03839006 -0.02514366 -0.03174078 -0.02845521 -0.03009153 -0.02927659
[13] -0.02968245 -0.02948032 -0.02958099 -0.02953085 -0.02955582 -0.02954339
[19] -0.02954958 -0.02954650 -0.02954803 -0.02954727 -0.02954765 -0.02954746
[25] -0.02954755 -0.02954751 -0.02954753 -0.02954752 -0.02954752 -0.02954752
```

\$se

Time Series:

Start = 174

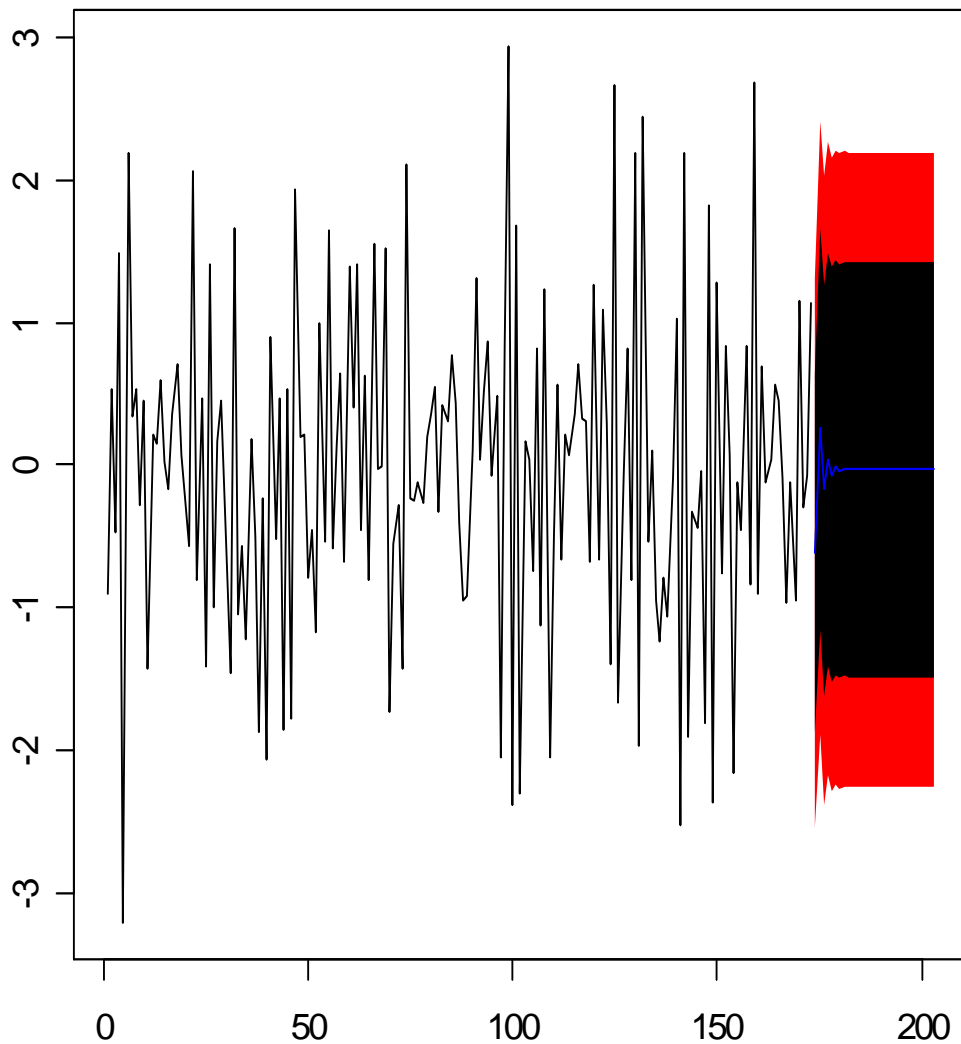
End = 203

Frequency = 1

```
[1] 0.9833617 1.0985672 1.1253181 1.1318554 1.1334711 1.1338715 1.1339707
[8] 1.1339954 1.1340015 1.1340030 1.1340034 1.1340035 1.1340035 1.1340035
[15] 1.1340035 1.1340035 1.1340035 1.1340035 1.1340035 1.1340035 1.1340035
[22] 1.1340035 1.1340035 1.1340035 1.1340035 1.1340035 1.1340035 1.1340035
```

[29] 1.1340035 1.1340035

### Forecasts from AR(1)



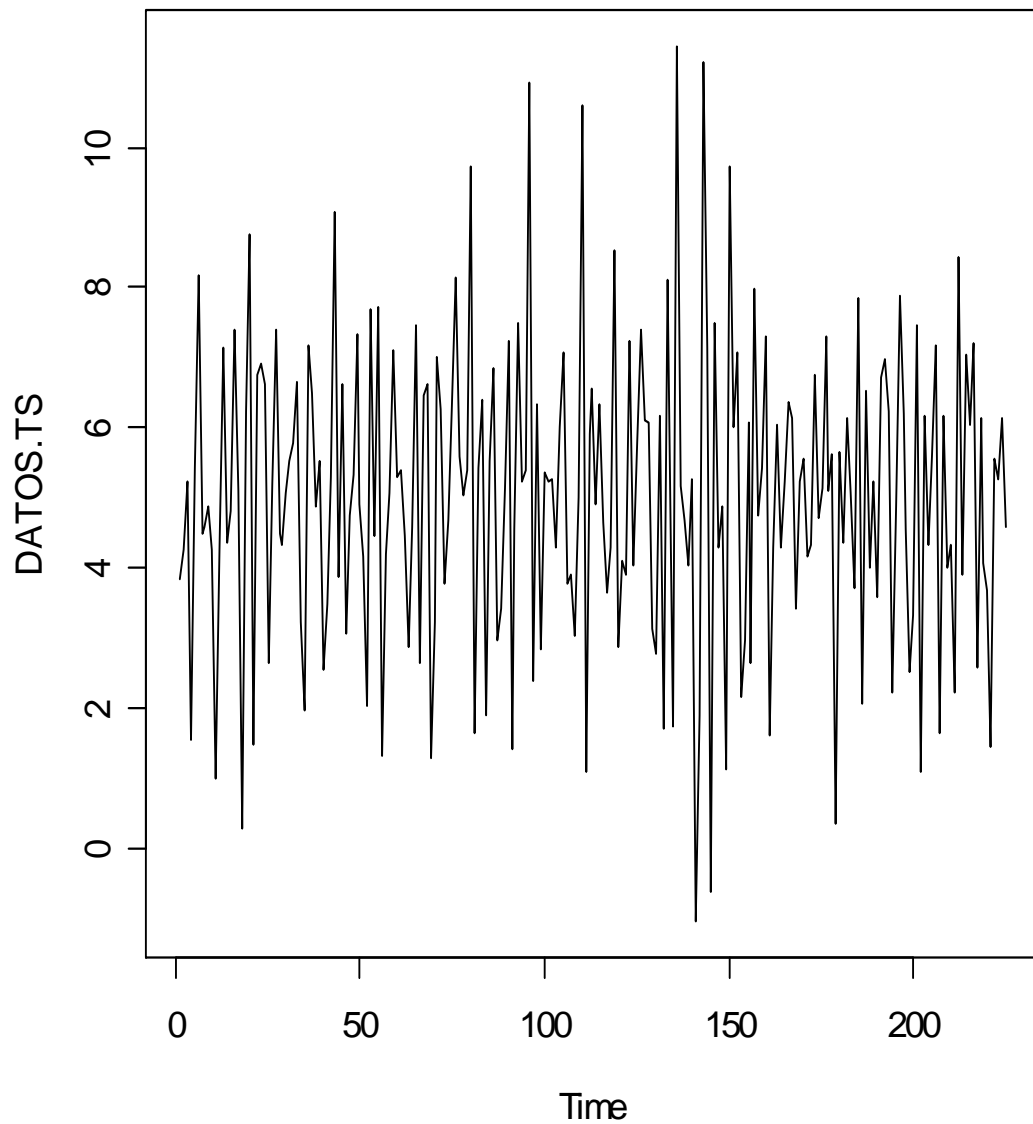
## 3.6 Serie 6

Dada la serie, se procederá a realizar el estudio de análisis de series temporales.

### Análisis descriptivo de una serie temporal.

Todo análisis de serie temporal univariante, comienza con la presentación de un grafico donde se muestra la evolución de la variable a lo largo del tiempo.

Los resultados son los siguientes:



*Figura 6.1*

Como se aprecia en los resultados, los datos parecen estacionarios. Es posible que los varores cercanos al valor 150, generen perturbaciones en el modelo. En futuras actualizaciones del documento, generare un Script que determine Datos Anomalos en las series.

El análisis exploratorio de datos revela los siguientes resultados:

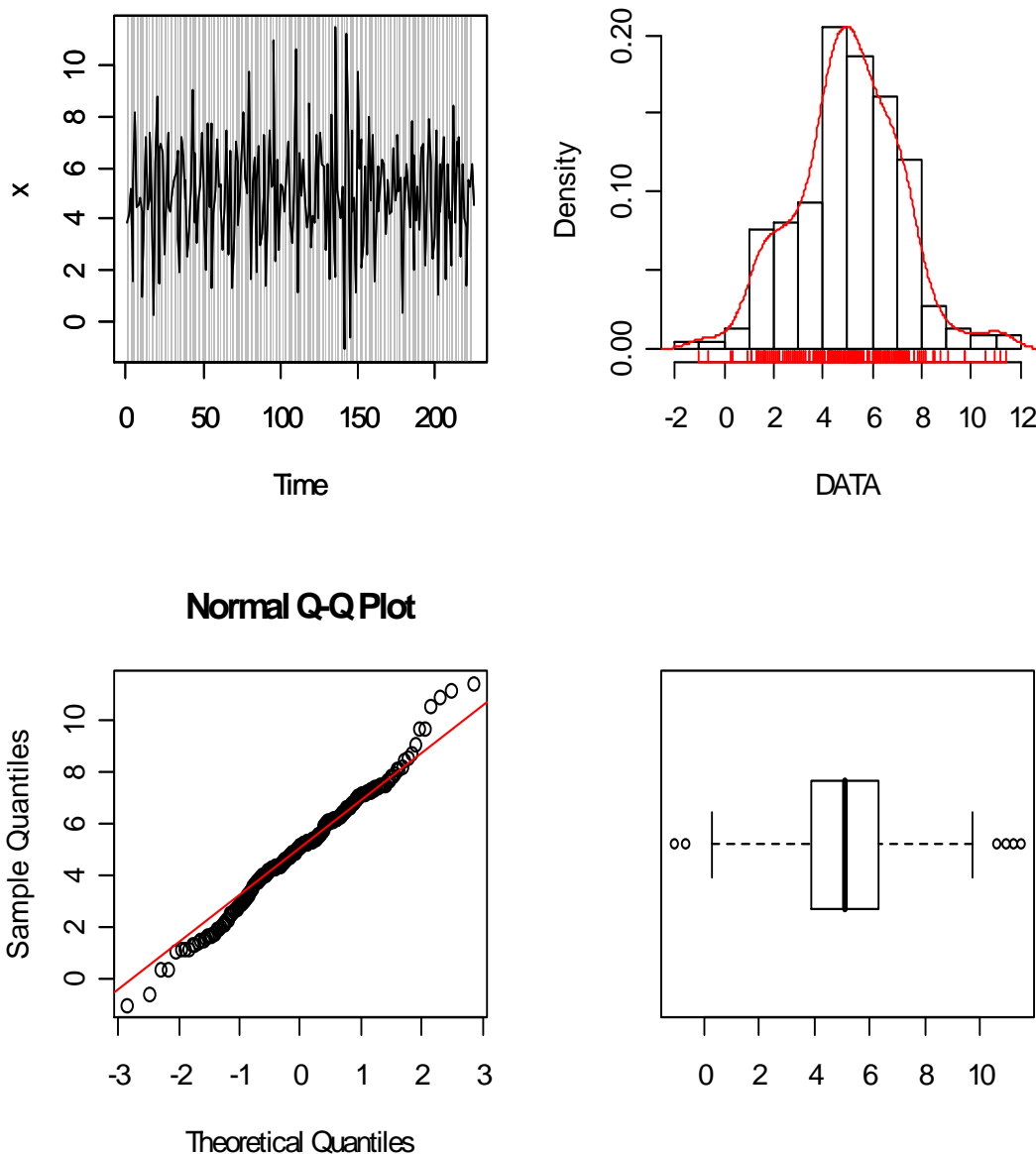


Figura 6.2

[1] "JAB: EDA VIEW 1.1"

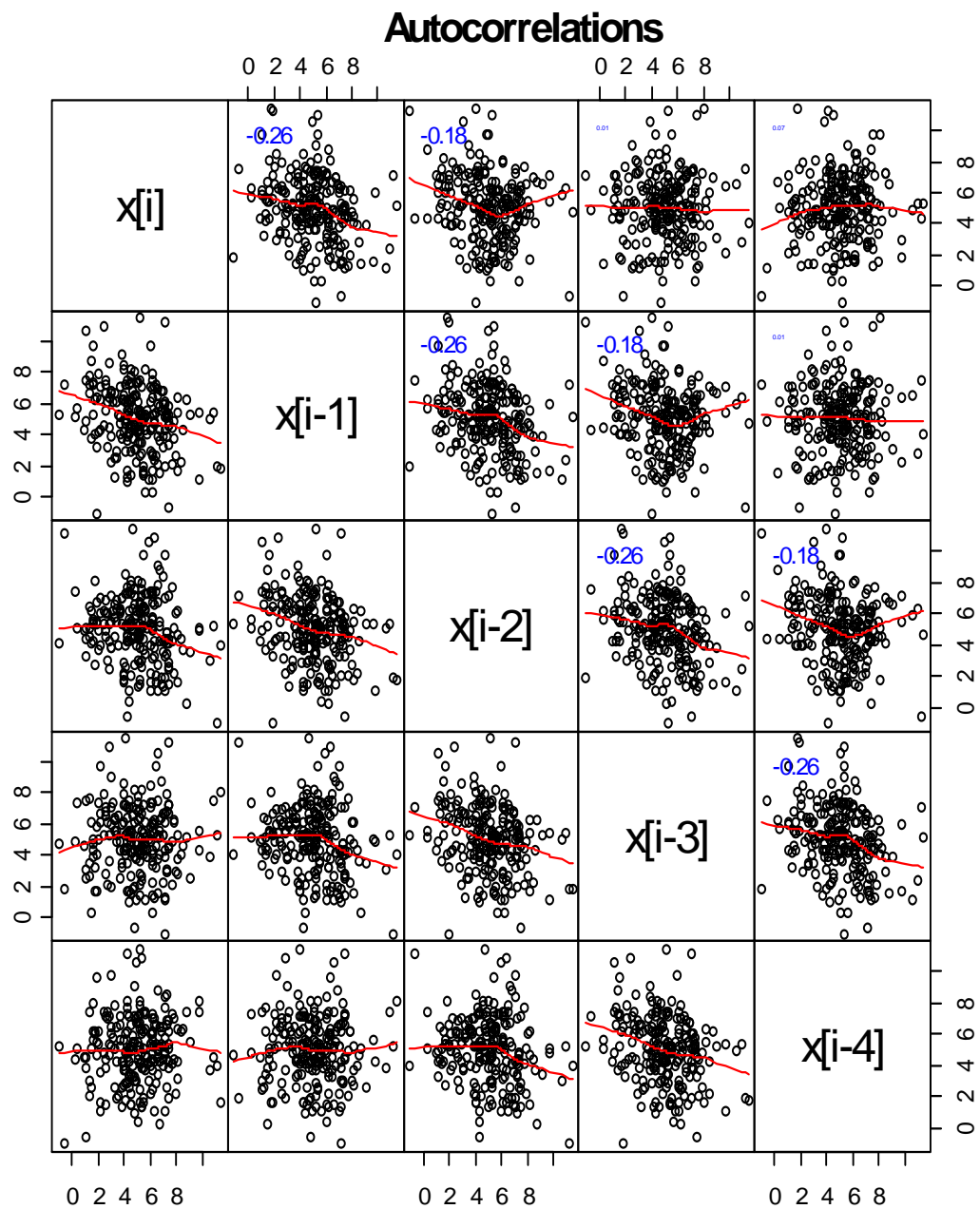
The decimal point is at the |

```
-1 | 0
-0 | 6
0 | 33
1 | 01113344566667899
2 | 00122245566678899
3 | 0001122344567778889999
4 | 0000011222233333333334445556666677777789999999
5 | 0011111112222222233333344444444555666667889
6 | 000011111112222223333445556666778889
7 | 000111122222333344455557789
8 | 0112458
9 | 177
```

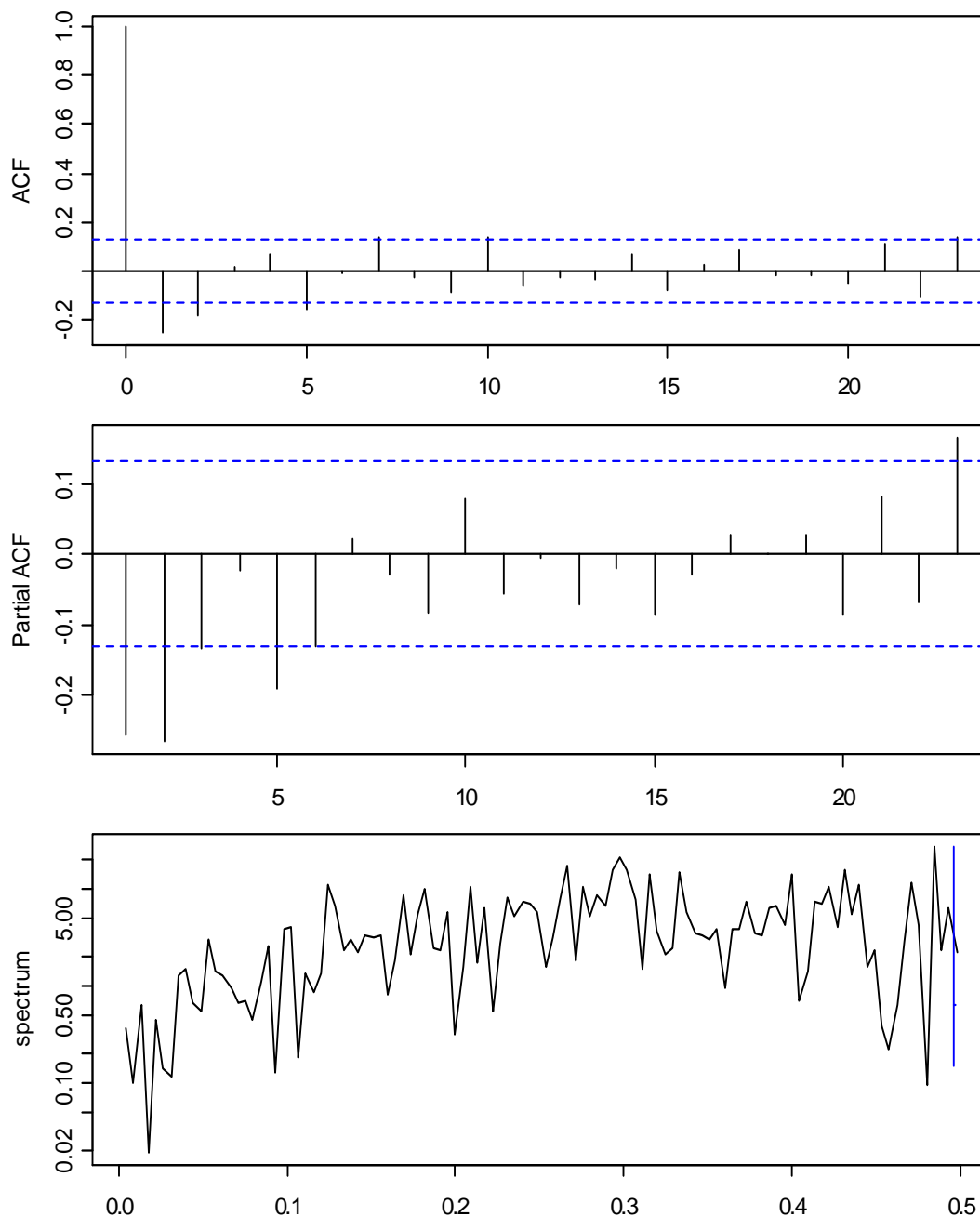
10 | 69  
11 | 24

```
      MEAN   MEDIAN     SD  KURTOSIS  SKEWNESS
5.01488910 5.10232296 2.12712872 0.36919390 0.03304518
[1] ""
[1] "Normal Test, with p-value = 0,05"
      Shapiro-Wilk Anderson-Darling   Cramer-von      Pearson
      0.13767355    0.10080401    0.14316774    0.04017816
Shapiro-Francia   Jarque-Bera
      0.10551267    0.46380599
[1] "OK" "OK" "OK" "X"  "OK" "OK"
```

Como se observa, los datos son simétricos pero se ven afectados por datos anomalos.  
¿Seria necesario eliminar los datos anomalos, para mejorar el modelo? Por otro lado,  
aunque solo pasa un criterio de normalidad, los restantes lo pasan a un 15%. Como se  
observa, los datos son casi simétricos.

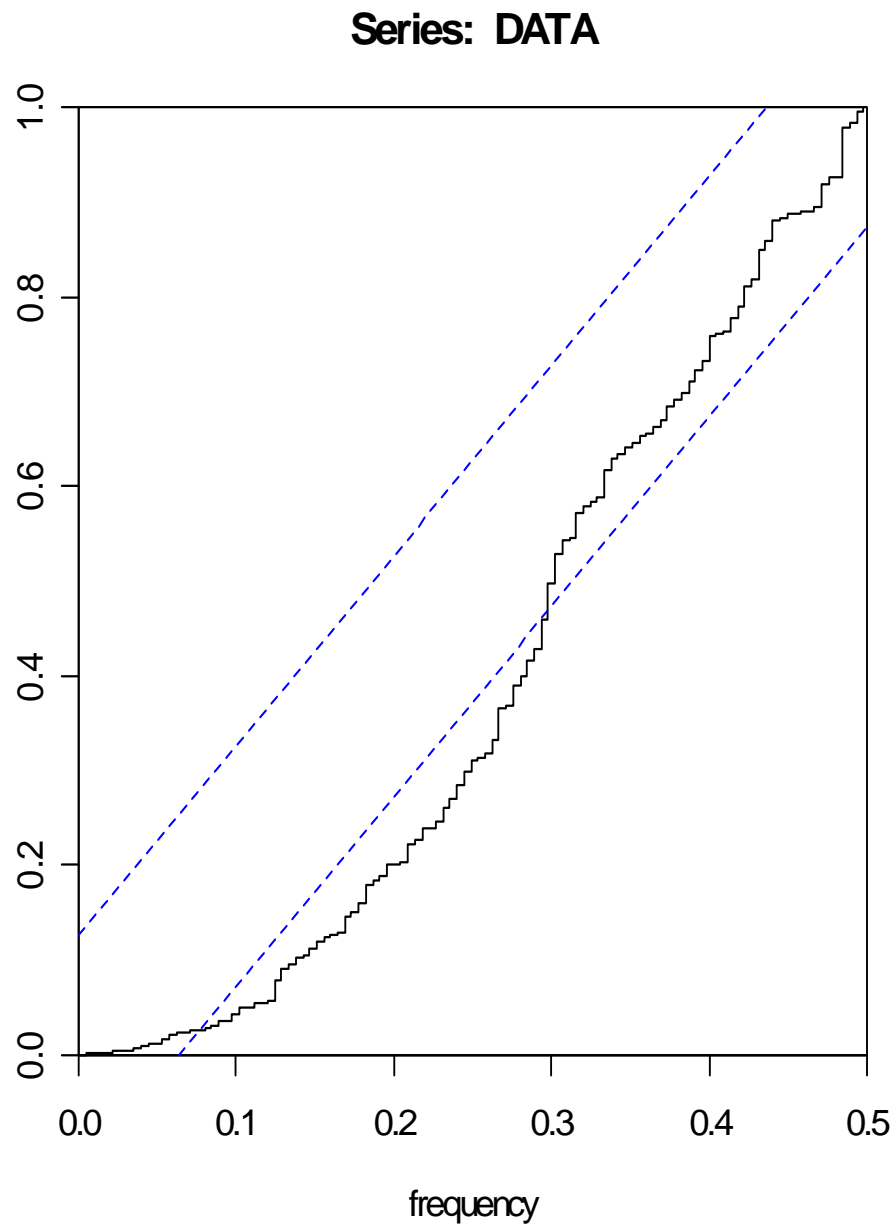






Atendiendo a los resultados, se considera modelar los datos a través de un proceso MA, debido a que en el gráfico de la función de autocorrelación, se observan los primeros coeficientes no nulos y el resto nulos, por otra parte si se observa el gráfico de la función de autocorrelación parcial, se observan, muchos coeficientes no nulos, por tanto, el proceso puede ser MA(4), MA(3), MA(2) o MA(1),

TIPO	FAS	FAP
AR(p)	Muchos coeficientes no nulos	Primeros p no nulos, resto 0
MA(q)	Primeros p no nulos, resto 0	Muchos coeficientes no nulos
ARMA(p,q)	Muchos coeficientes no nulos	Muchos coeficientes no nulos

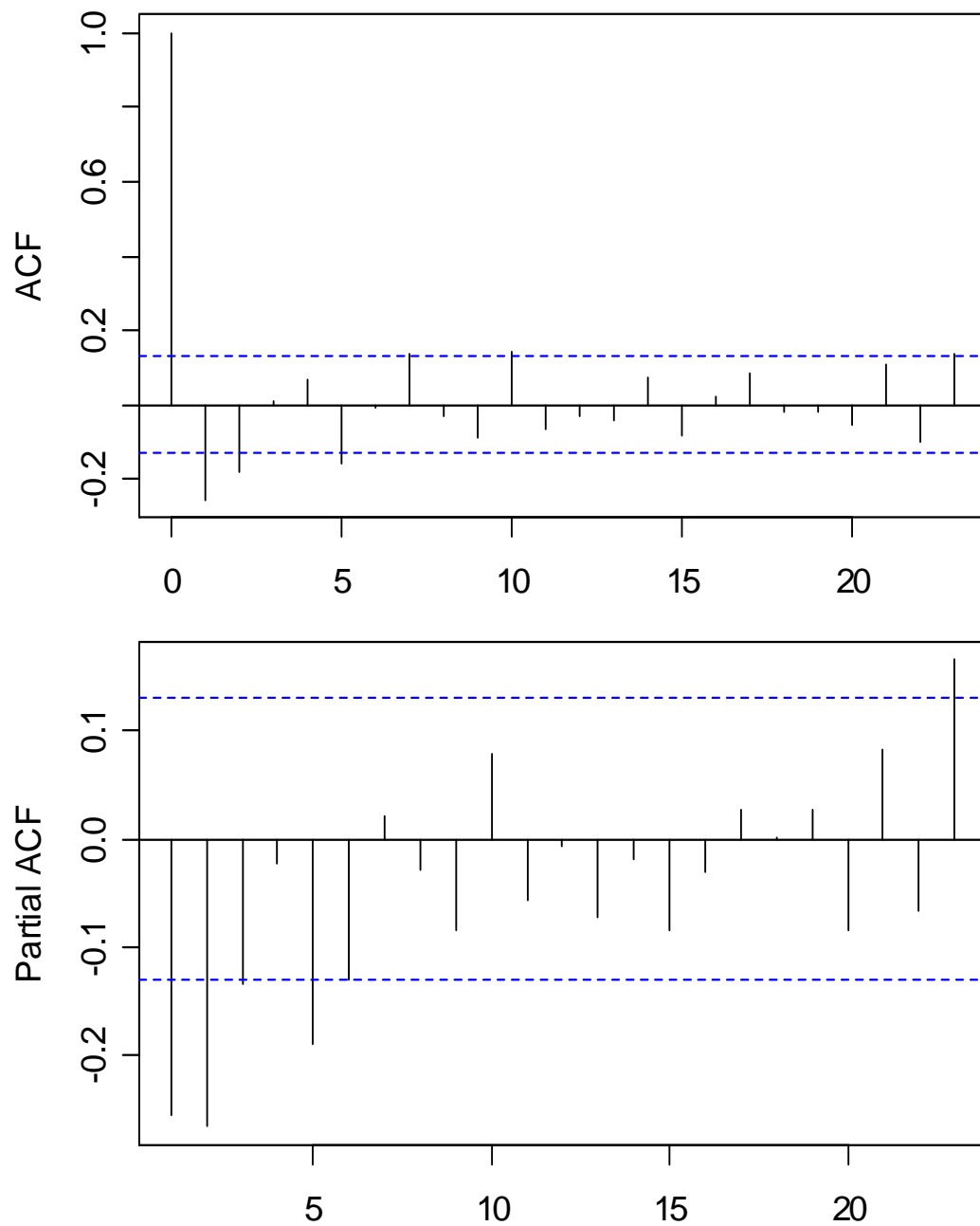


## Modelado de la serie temporal.

Como se observo en el anterior punto, los datos muestran una natural predisposición a las siguientes opciones:

1. MA(1)
2. MA(2)
3. MA(3)
4. MA(4)

A continuación, analizaremos cada una de las opciones y nos decidiremos por aquella opcion con un AIC menor.



### Caso 1: MA(1)

Call:

```
arma(x = DATOS.TS, order = c(0, 1))
```

Model:

ARMA(0,1)

Residuals:

Min	1Q	Median	3Q	Max
-6.21932	-1.18739	-0.05062	1.31536	6.34541

Coefficient(s):

Estimate	Std. Error	t value	Pr(> t )
----------	------------	---------	----------

```

ma1      -0.50808   0.08017  -6.338 2.33e-10 ***
intercept 5.02232   0.06535  76.858 < 2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Fit:  
sigma^2 estimated as 3.957, Conditional Sum-of-Squares = 882.46, **AIC = 952.02**

```

      ma1 intercept
-0.5080798 5.0223208

```

### Caso 2: MA(2)

Call:  
arma(x = DATOS.TS, order = c(0, 2))

Model:  
ARMA(0,2)

Residuals:

Min	1Q	Median	3Q	Max
-5.87292	-1.21803	0.03529	1.32590	6.24806

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t )
ma1	-0.41918	0.06343	-6.609	3.87e-11 ***
ma2	-0.24857	0.06546	-3.797	0.000146 ***
intercept	5.02885	0.04325	116.266	< 2e-16 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Fit:  
sigma^2 estimated as 3.748, Conditional Sum-of-Squares = 832.04, **AIC = 941.79**

```

      ma1      ma2 intercept
-0.4191815 -0.2485719 5.0288457

```

### Caso 3: MA(3)

Call:  
arma(x = DATOS.TS, order = c(0, 3))

Model:  
ARMA(0,3)

Residuals:

Min	1Q	Median	3Q	Max
-5.80777	-1.16484	0.03085	1.32710	6.21342

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t )
--	----------	------------	---------	----------

```

ma1      -0.40660    0.06692   -6.075 1.24e-09 ***
ma2      -0.23715    0.06779   -3.498 0.000468 ***
ma3      -0.04372    0.06312   -0.693 0.488545
intercept 5.03024    0.04085   123.133 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Fit:  
sigma^2 estimated as 3.76, Conditional Sum-of-Squares = 830.97, AIC = **944.51**

```

      ma1      ma2      ma3 intercept
-0.40659750 -0.23714559 -0.04371872  5.03023951

```

#### Caso 4: MA(4)

Call:  
arma(x = DATOS.TS, order = c(0, 4))

Model:  
ARMA(0,4)

Residuals:

Min	1Q	Median	3Q	Max
-5.7713	-1.2255	0.1157	1.3406	6.1009

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t )	
ma1	-0.41857	0.06956	-6.017	1.78e-09	***
ma2	-0.23281	0.07394	-3.149	0.00164	**
ma3	-0.04924	0.07480	-0.658	0.51034	
ma4	-0.02849	0.07648	-0.373	0.70946	
intercept	5.04069	0.03541	142.338	< 2e-16	***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Fit:  
sigma^2 estimated as 3.723, Conditional Sum-of-Squares = 819.12, AIC = **944.3**

```

> DATOS.TS.MA$coef
      ma1      ma2      ma3      ma4 intercept
-0.41856965 -0.23280736 -0.04924055 -0.02849379  5.04068659

```

Conclusiones, como se observa, el mejor modelo es el MA(2) debido a un AIC mas bajo con parámetros:

```

      ma1      ma2 intercept
-0.4191815 -0.2485719  5.0288457

```

### 3.7 Serie 7

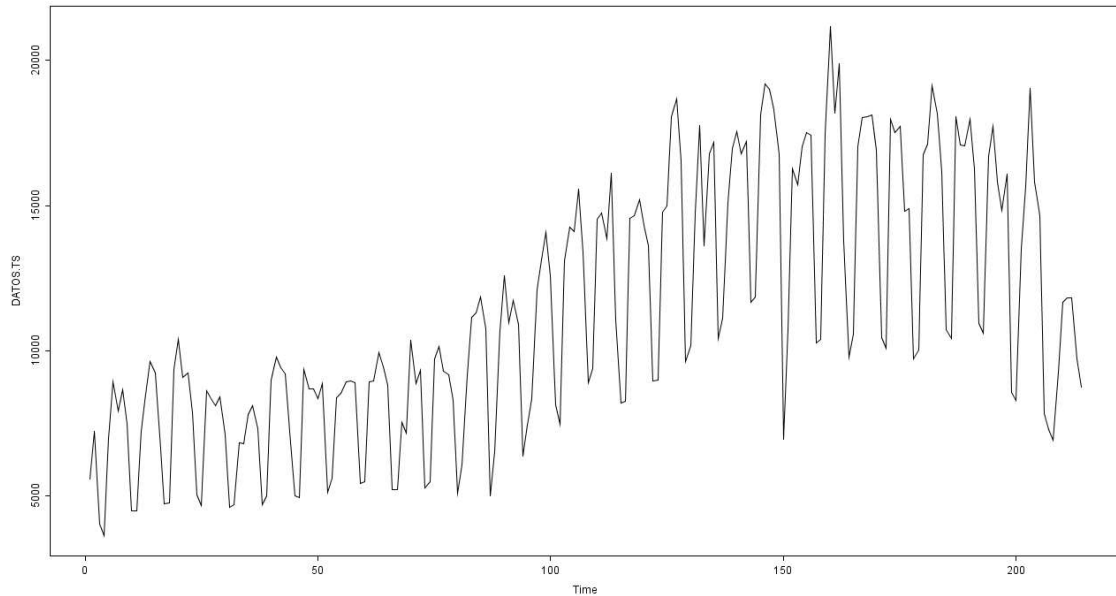
Dada la serie, se procederá a realizar el estudio de análisis de series temporales.

## MEMORIA LARGA

### Análisis descriptivo de una serie temporal.

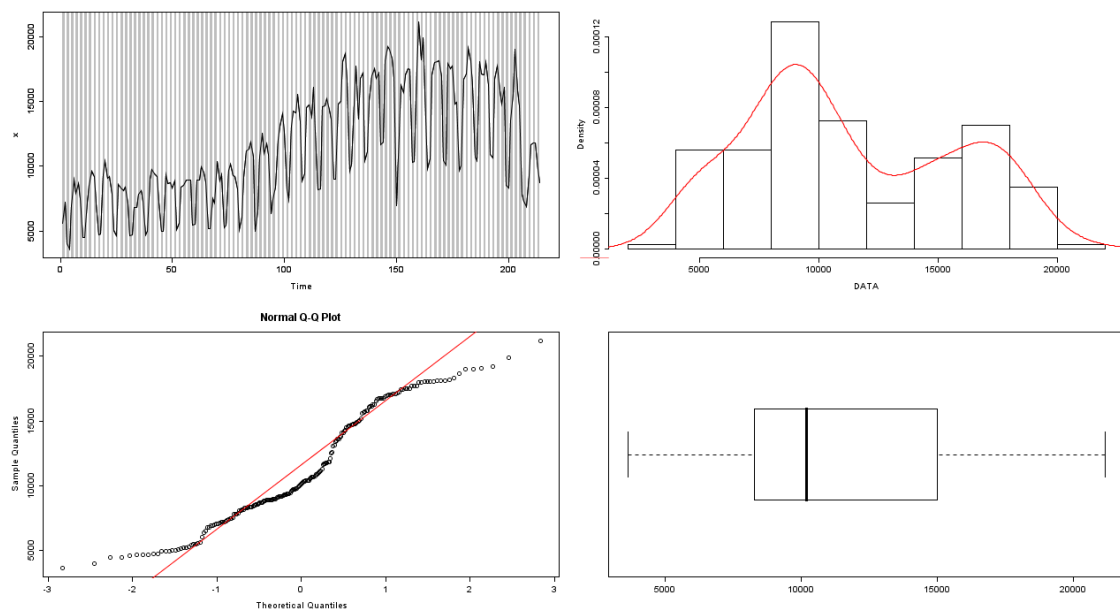
Todo análisis de serie temporal univariante, comienza con la presentación de un gráfico donde se muestra la evolución de la variable a lo largo del tiempo.

Los resultados son los siguientes:



*Figura 7.1*

Como se observa, los datos muestran a una serie no estacionaria de memoria larga. Los resultados del análisis de exploración de datos muestra los siguientes resultados:



*Figura 7.2*

[1] "JAB: EDA VIEW 1.1"

The decimal point is 3 digit(s) to the right of the |

```
3 | 6
4 | 0556777779
5 | 00001122345566
6 | 145889
7 | 001122223345558899
8 | 1122333334445566777789999999
9 | 000001222223334444466777889
10 | 0122344444566777799
11 | 011237778889
12 | 166
13 | 11456689
14 | 11335677788889
15 | 01266788
16 | 112336777889
17 | 0000112244555778
18 | 000111112237
19 | 00129
20 |
21 | 2
```

MEAN	MEDIAN	SD	KURTOSIS	SKEWNESS
11255.257009	10176.500000	4325.433148	-1.078456	0.299687

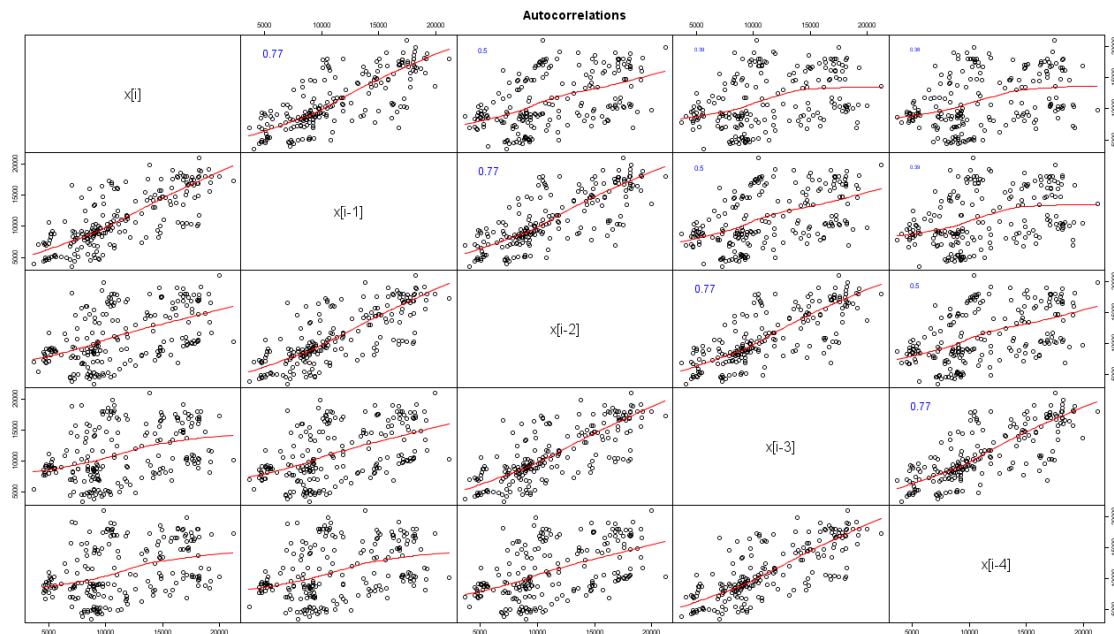
[1] ""

[1] "Normal Test, with p-value = 0,05"

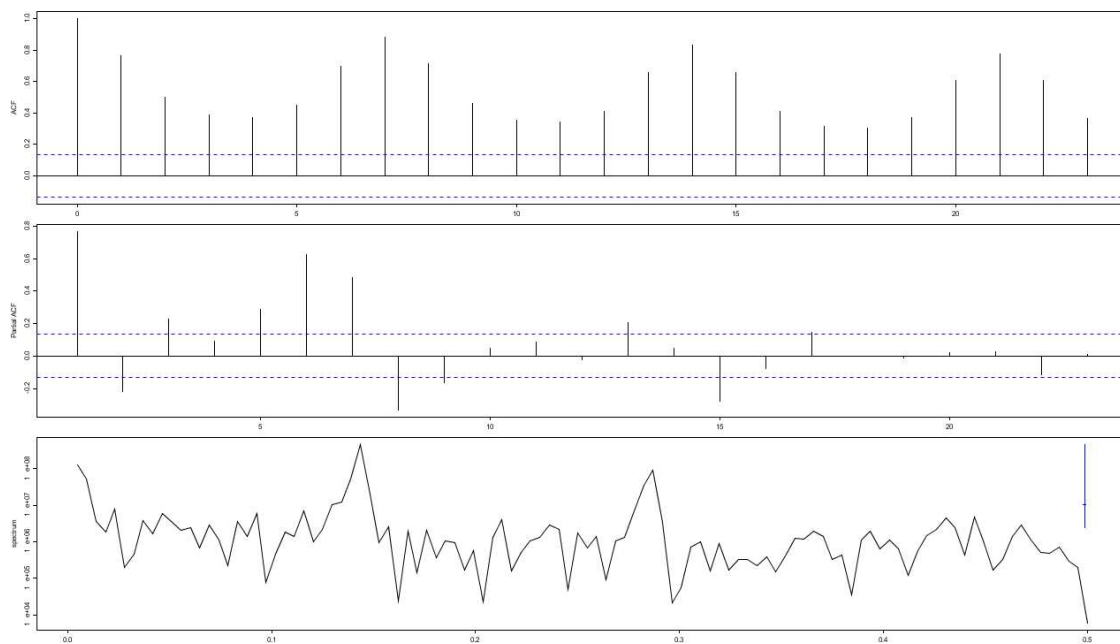
Shapiro-Wilk	Anderson-Darling	Cramer-von	Pearson
2.912403e-07	9.918415e-11	2.271439e-08	5.878509e-14
Shapiro-Francia	Jarque-Bera		
2.818513e-06	1.310683e-03		

[1] "X" "X" "X" "X" "X" "X"

Como se observa, los datos no son simétricos, pero que creo que a través de transformaciones, es posible hacer estacionaria dicha serie y poder ser modelada.



Los datos de las autocorrelaciones no son convincentes.



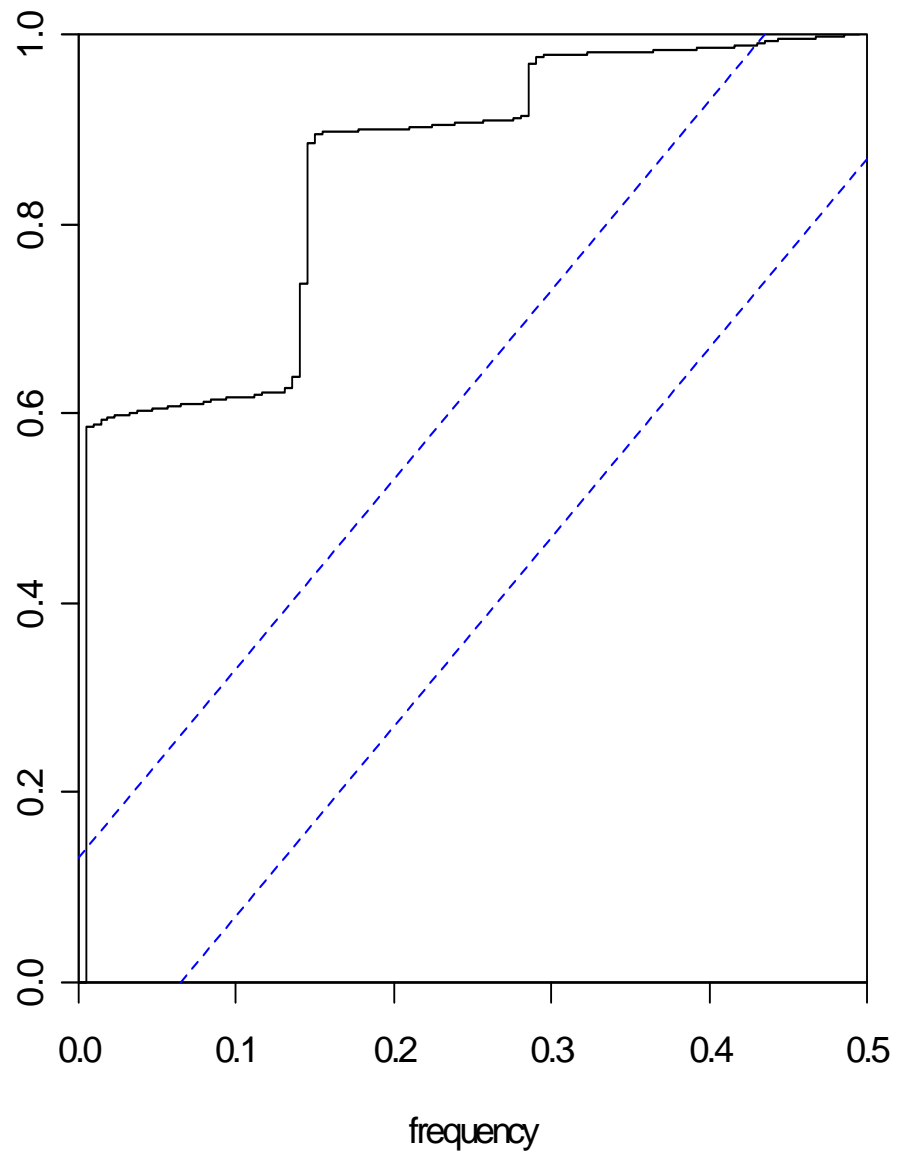
TIPO	FAS	FAP
AR(p)	Muchos coeficientes no nulos	Primeros p no nulos, resto 0
MA(q)	Primeros p no nulos, resto 0	Muchos coeficientes no nulos
ARMA(p,q)	Muchos coeficientes no nulos	Muchos coeficientes no nulos

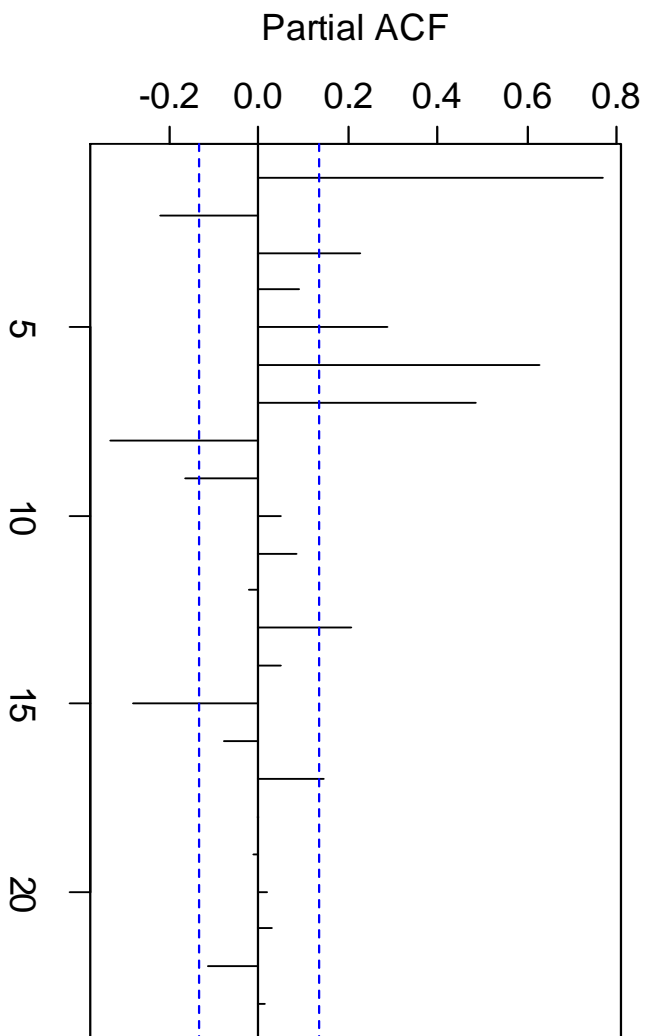
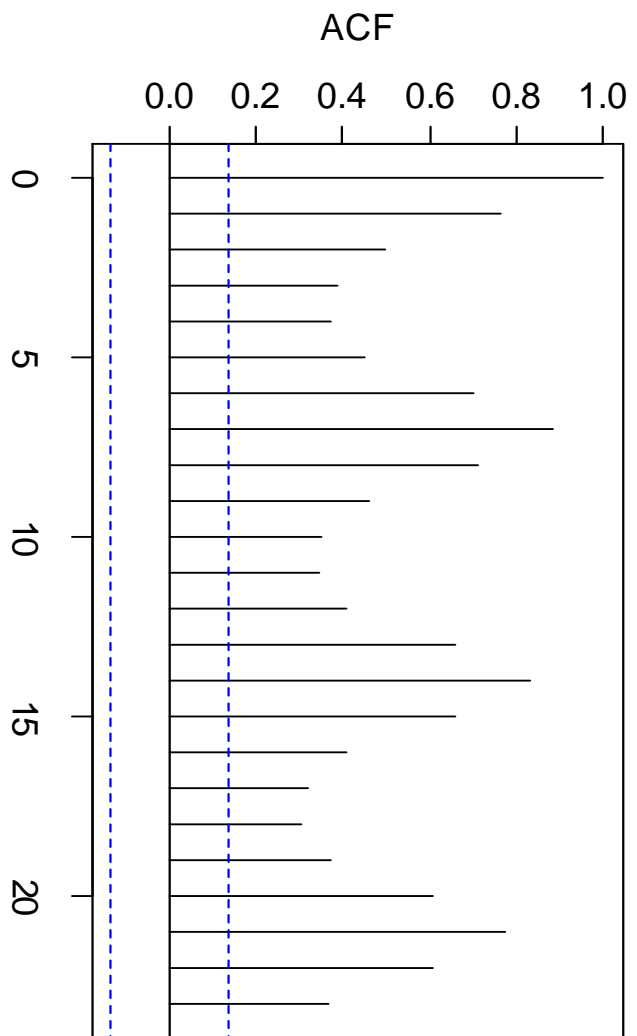
### Conclusión:

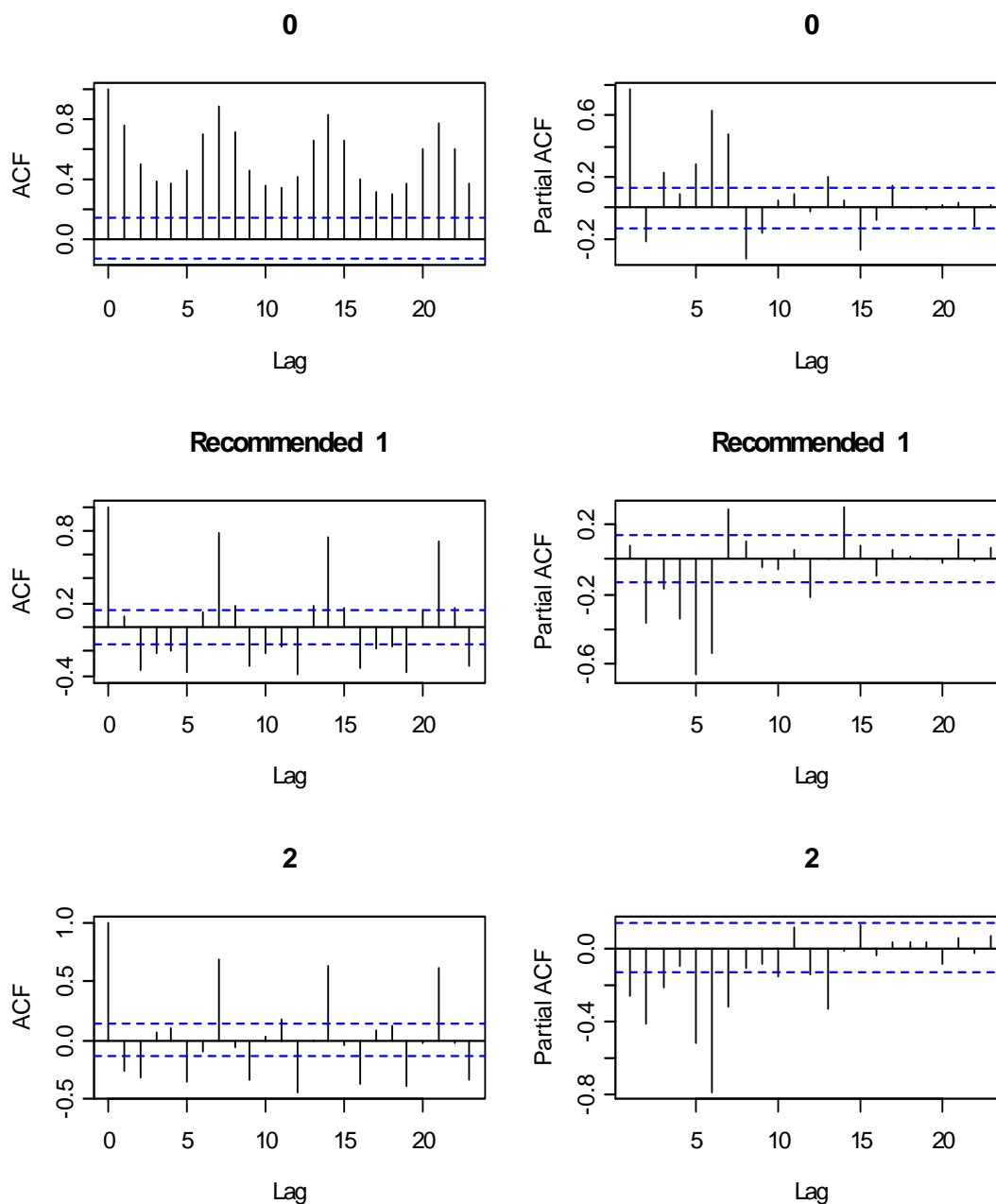
Considero que puede pasar a una siguiente etapa donde intentar generar modelos ARMA o ARIMA



**Series: DATA**







Observando los resultados, nos recomienda, realizar una diferencia a los datos.

## Caso 2: Arima(5,1,0)

```
DATOS.TS.BEST_ARIMA <- best.arima(DATOS.TS ,d = ndiffs(DATOS.TS));
DATOS.TS.BEST_ARIMA;
```

Series: DATOS.TS  
ARIMA(5,1,0) model

Coefficients:

	ar1	ar2	ar3	ar4	ar5
	-0.2334	-0.5558	-0.4521	-0.3410	-0.6582
s.e.	0.0509	0.0477	0.0523	0.0475	0.0502

sigma^2 estimated as 3528396: log likelihood = -1909.7, **aic = 3831.4**

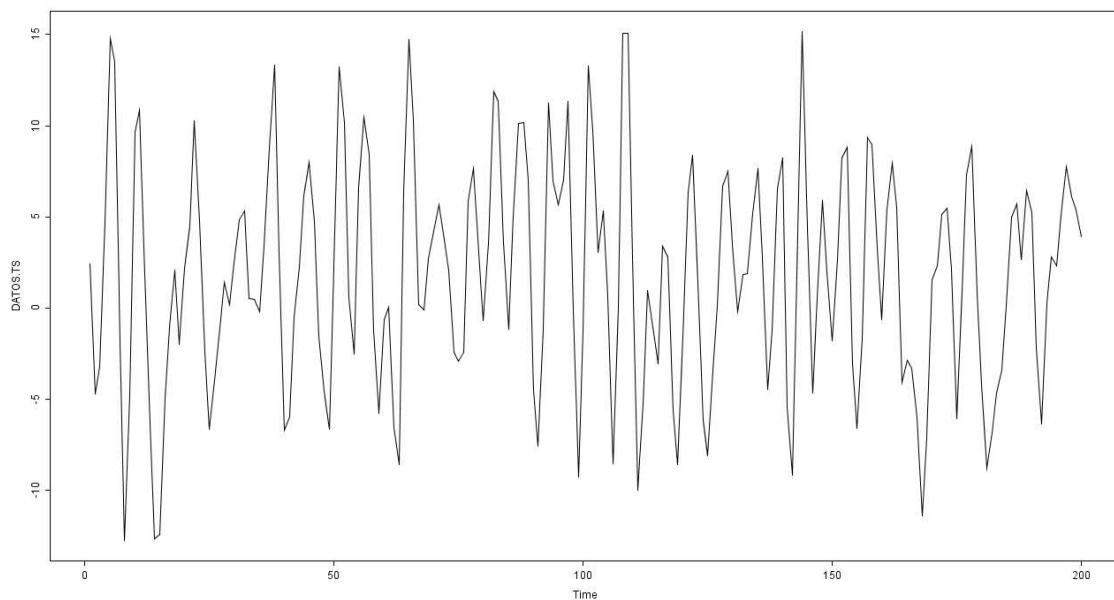
## 3.8 Serie 8

Dada la serie, se procederá a realizar el estudio de análisis de series temporales.

### Análisis descriptivo de una serie temporal.

Todo análisis de serie temporal univariante, comienza con la presentación de un grafico donde se muestra la evolución de la variable a lo largo del tiempo.

Los resultados son los siguientes:



*Figura 8.1*

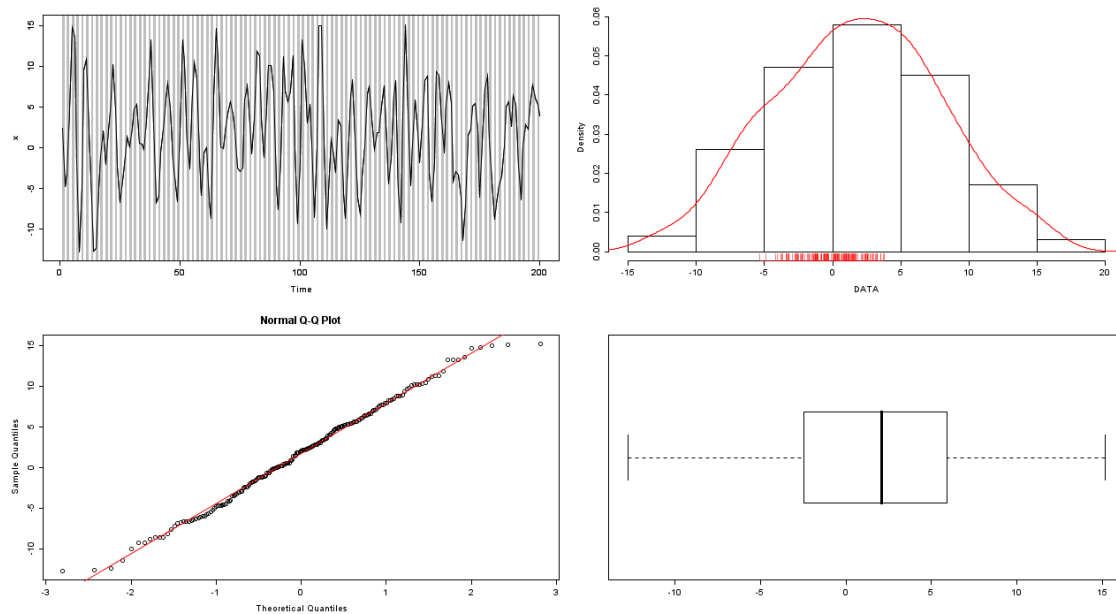


Figura 8.2

[1] "JAB: EDA VIEW 1.1"

The decimal point is at the |

```
-12 | 764
-10 | 40
-8 | 3286651
-6 | 61877665432100
-4 | 864208776655210
-2 | 44321099544420
-0 | 8876532222110776522110
0 | 112222555566990445999
2 | 011222333456778880113445569
4 | 01245778901112233445577799
6 | 1134556700035777
8 | 0023458899368
10 | 12233492339
12 | 2335
14 | 78012
```

MEAN	MEDIAN	SD	KURTOSIS	SKEWNESS
1.75469548	2.08595329	6.13217319	-0.49271238	-0.02867296

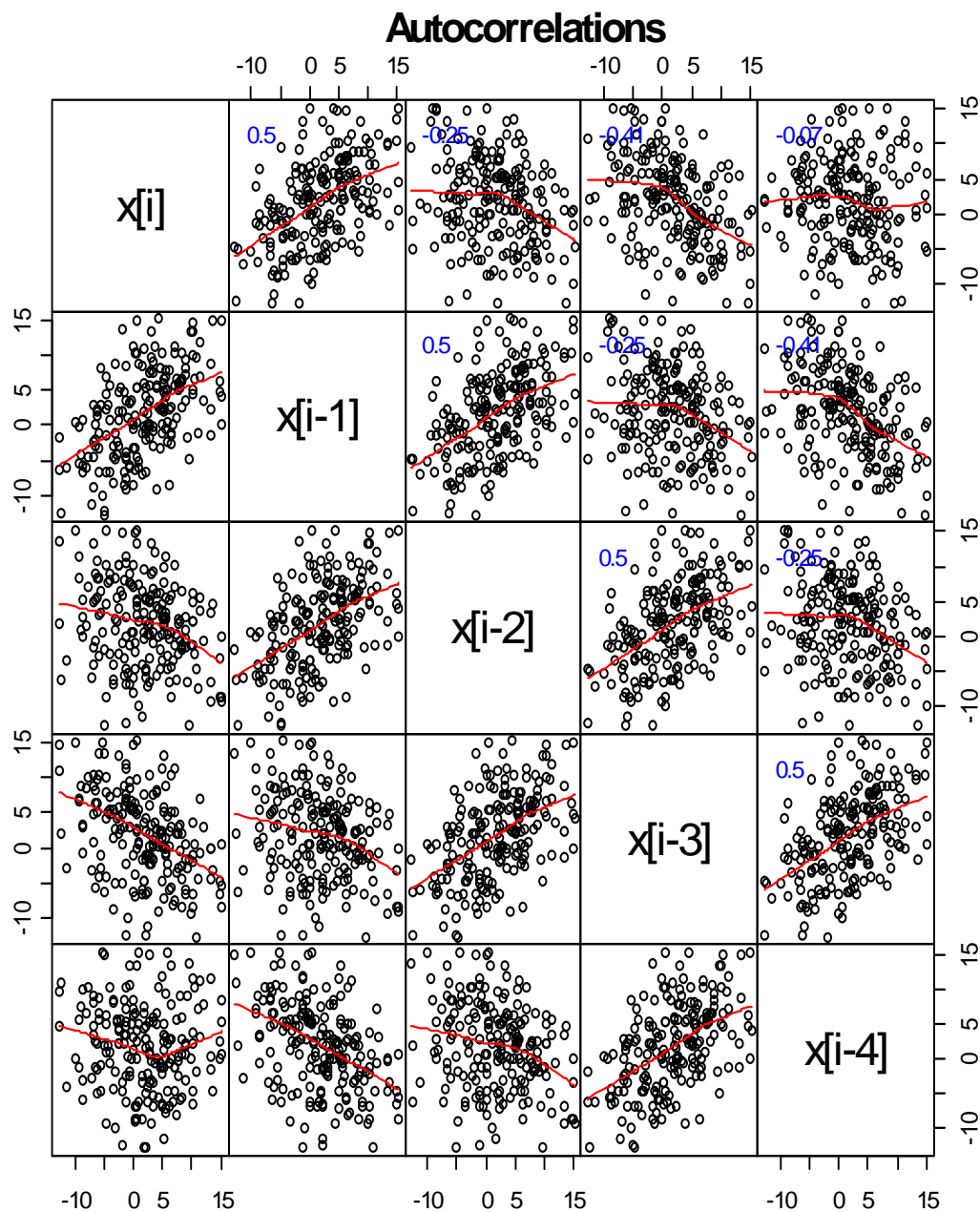
[1] ""

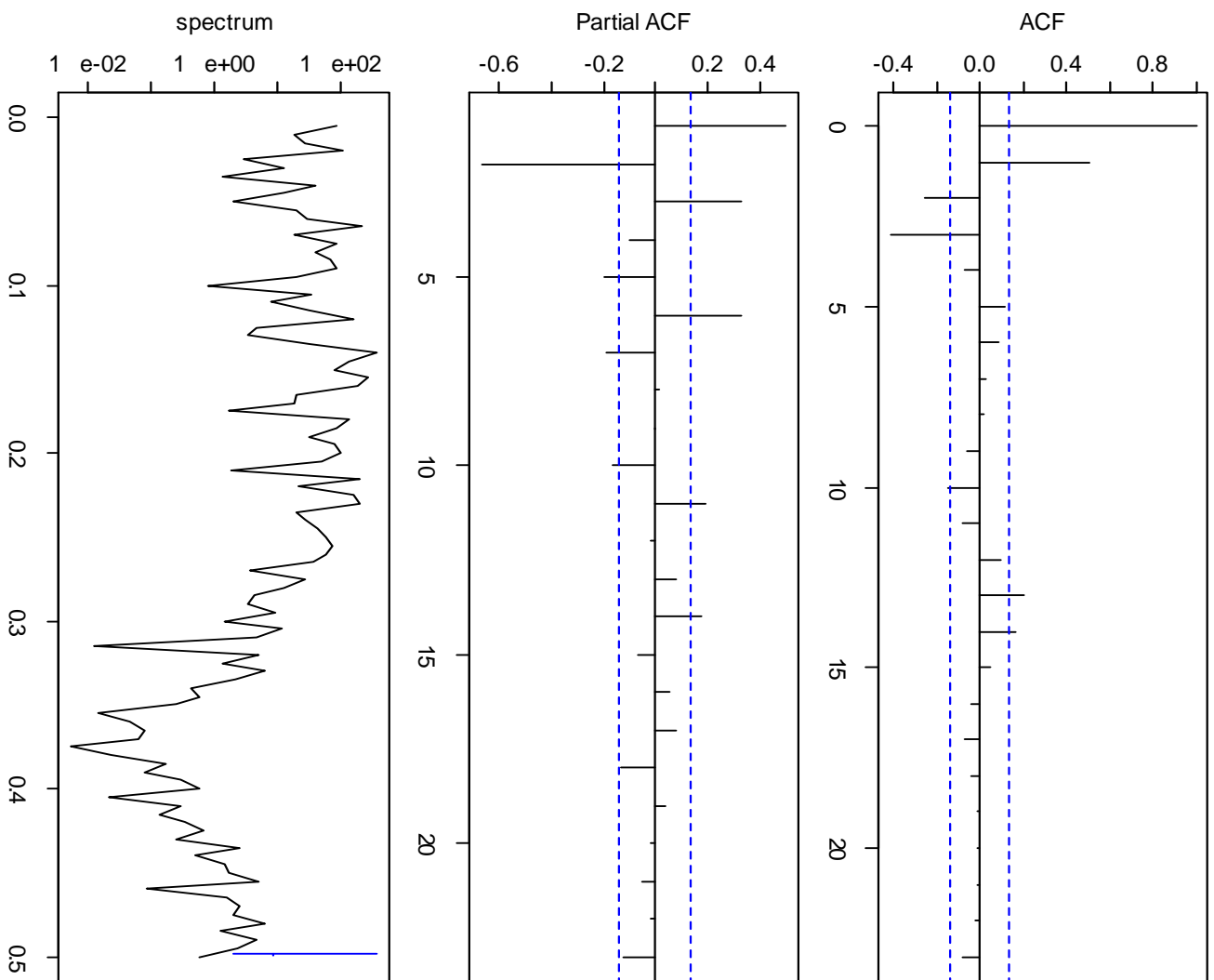
[1] "Normal Test, with p-value = 0,05"

Shapiro-Wilk	Anderson-Darling	Cramer-von	Pearson
0.3934106	0.8107897	0.8322955	0.8355085
Shapiro-Francia	Jarque-Bera		
0.6789782	0.3967825		

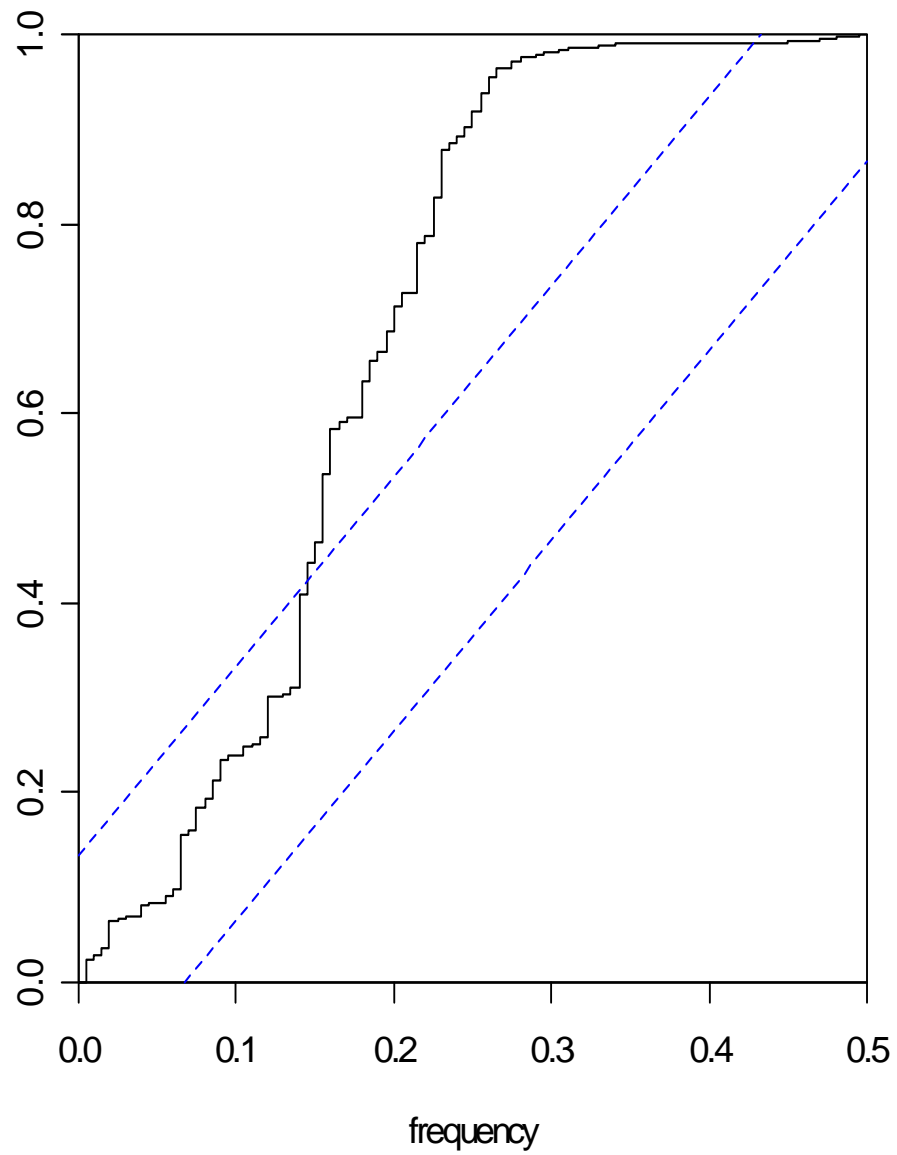
[1] "OK" "OK" "OK" "OK" "OK" "OK"

Se observa una cierta simetría y colas largas. Lo que mas me preocupa de la serie es la alta variabilidad que puede hacer que aunque se pueda modelar la serie, los rangos de predicción sean altos.

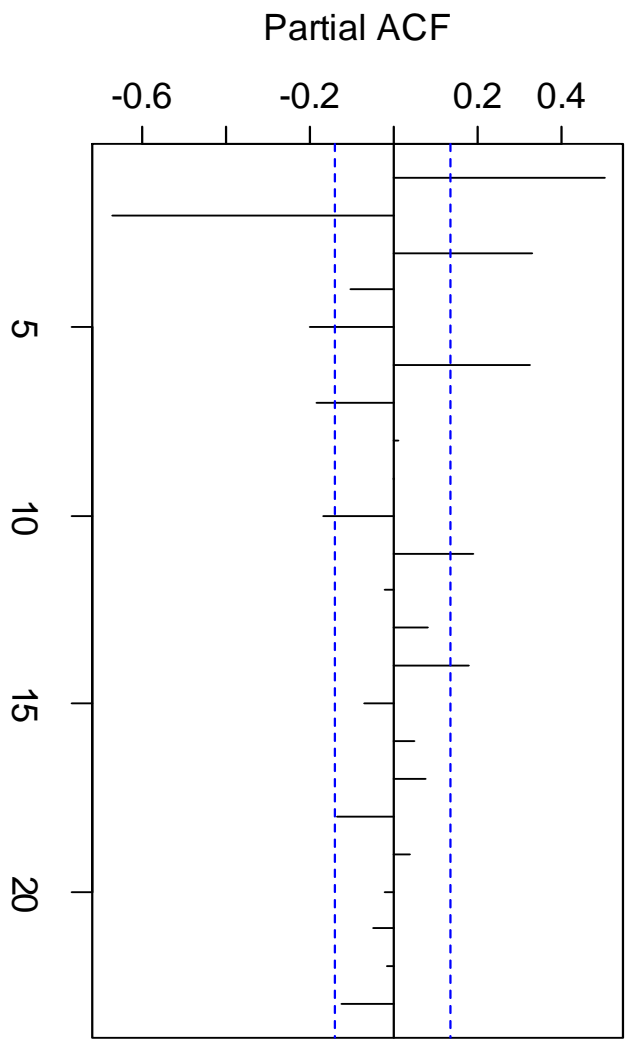
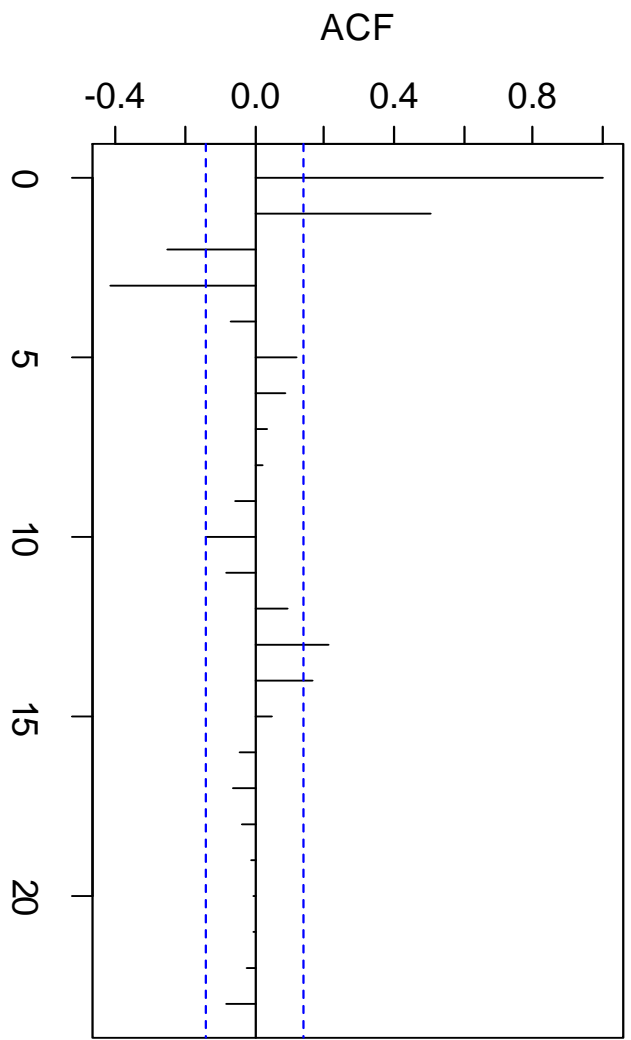


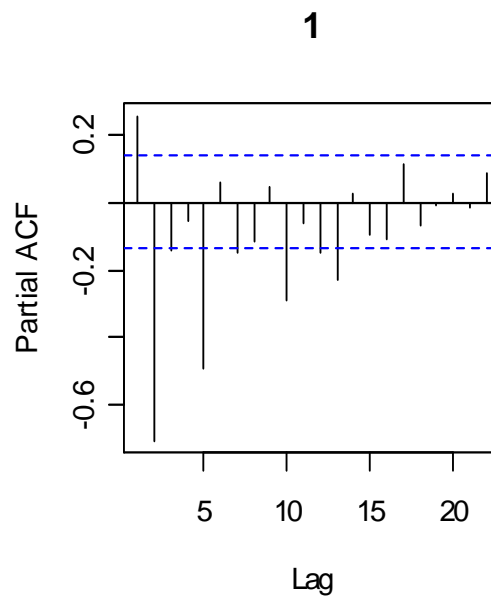
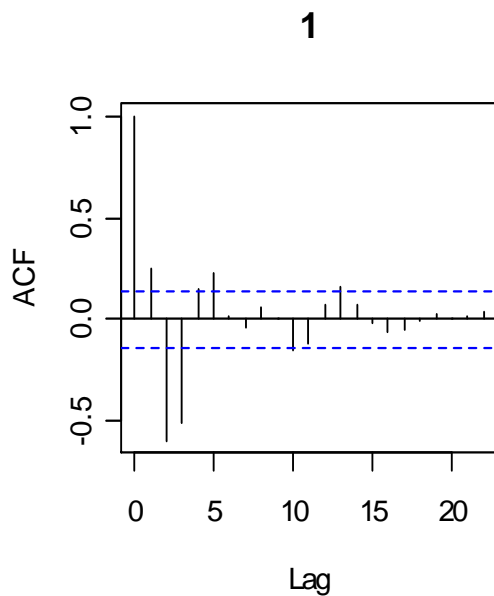
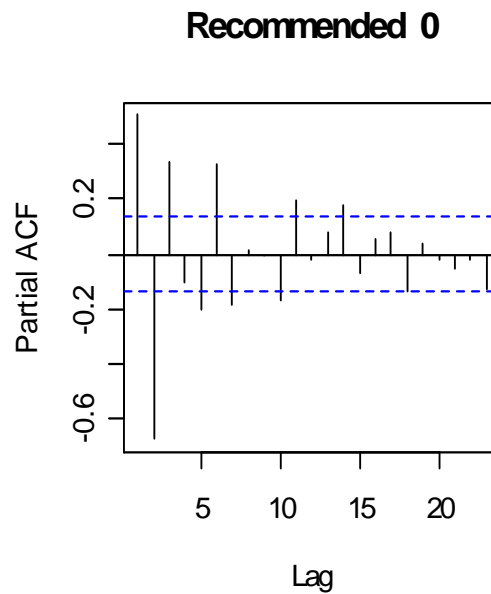
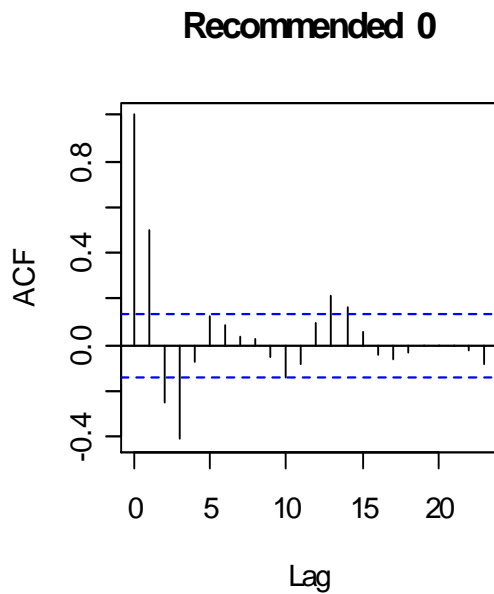


**Series: DATA**









Series: DATOS.TS  
ARIMA(3,0,2) model

Coefficients:

	ar1	ar2	ar3	ma1	ma2	intercept
	-0.0436	-0.3542	-0.2147	1.3813	0.9425	1.7639
s.e.	0.0760	0.0679	0.0720	0.0384	0.0346	0.4476

sigma<sup>2</sup> estimated as 9.44: log likelihood = -511.27, **aic = 1036.55**

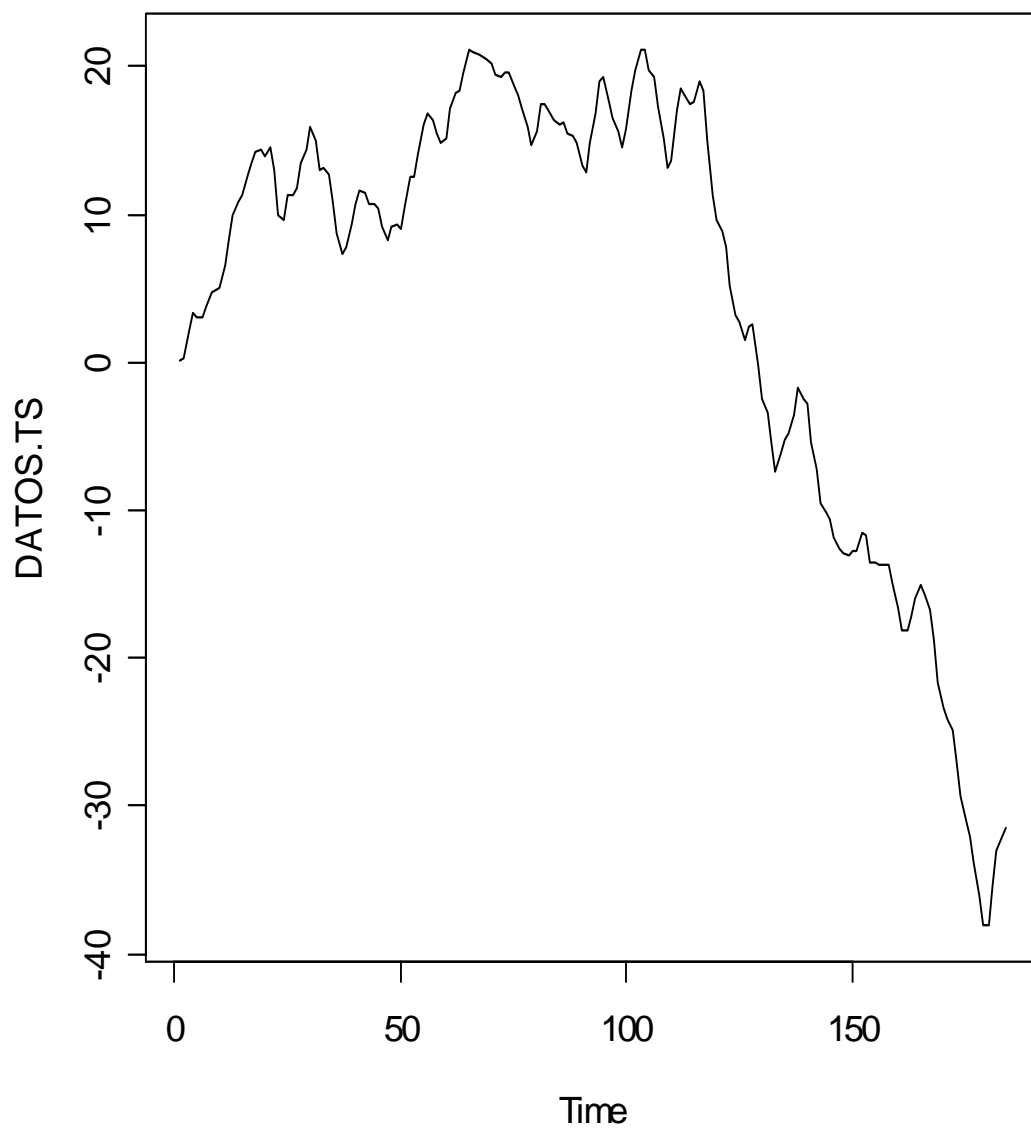
### 3.9 Serie 9

Dada la serie, se procederá a realizar el estudio de análisis de series temporales.

### Análisis descriptivo de una serie temporal.

Todo análisis de serie temporal univariante, comienza con la presentación de un grafico donde se muestra la evolución de la variable a lo largo del tiempo.

Los resultados son los siguientes:



*Figura 9.1*

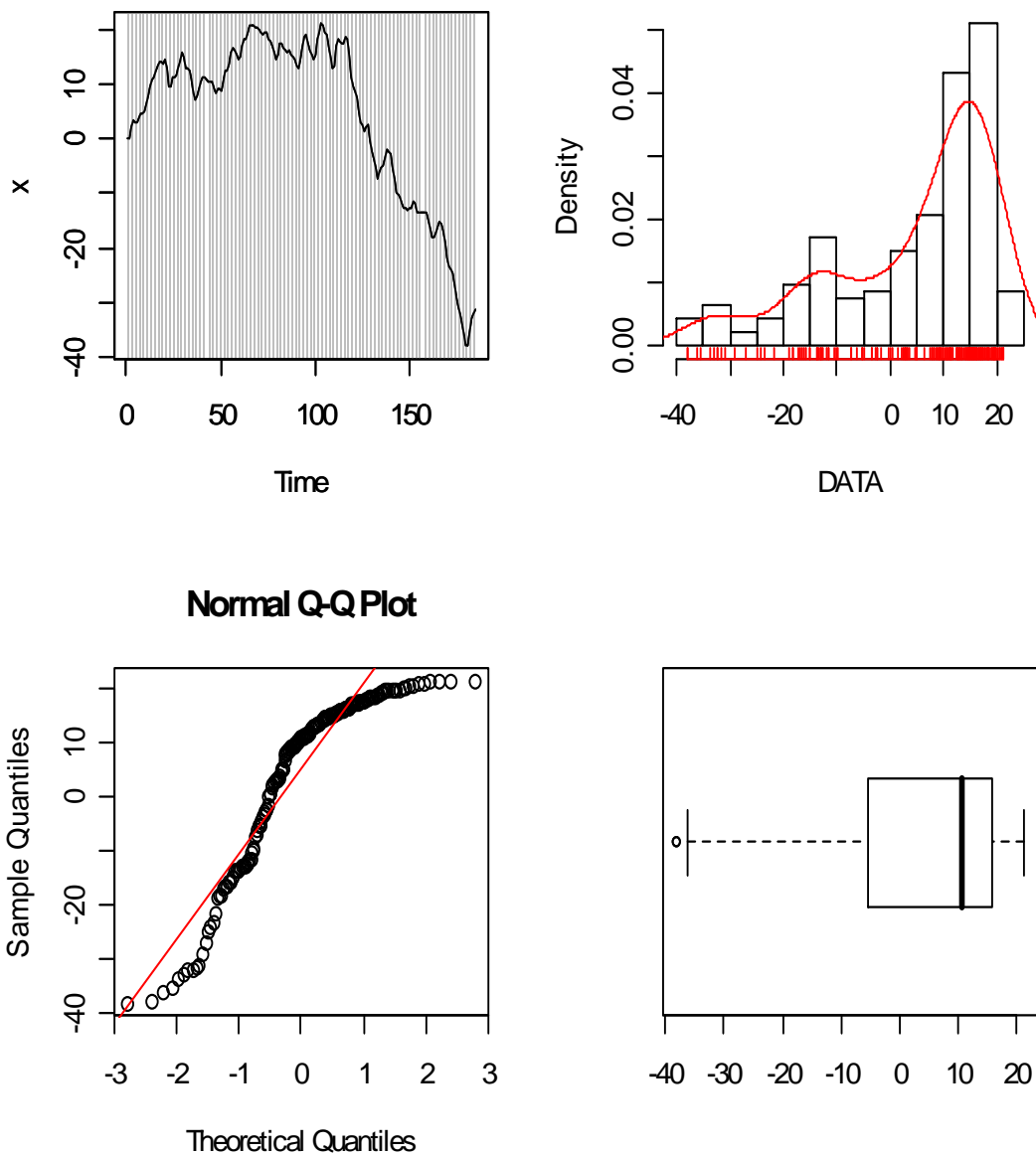


Figura 9.2

[1] "JAB: EDA VIEW 1.1"

The decimal point is 1 digit(s) to the right of the |

```
-3 | 8865
-3 | 432221
-2 | 975
-2 | 432
-1 | 9887776655
-1 | 4444333333222100
-0 | 7765555
-0 | 4333220
```

```

0 | 002223333334
0 | 5555778888999999
1 | 0000011111111111223333333333334444444
1 | 55555555555566666666667777777777888888889999999
2 | 0000000111111

```

```

      MEAN      MEDIAN      SD  KURTOSIS  SKEWNESS
4.24012891 10.64198449 15.46993304 0.09099091 -1.06650548

```

```
[1] ""
```

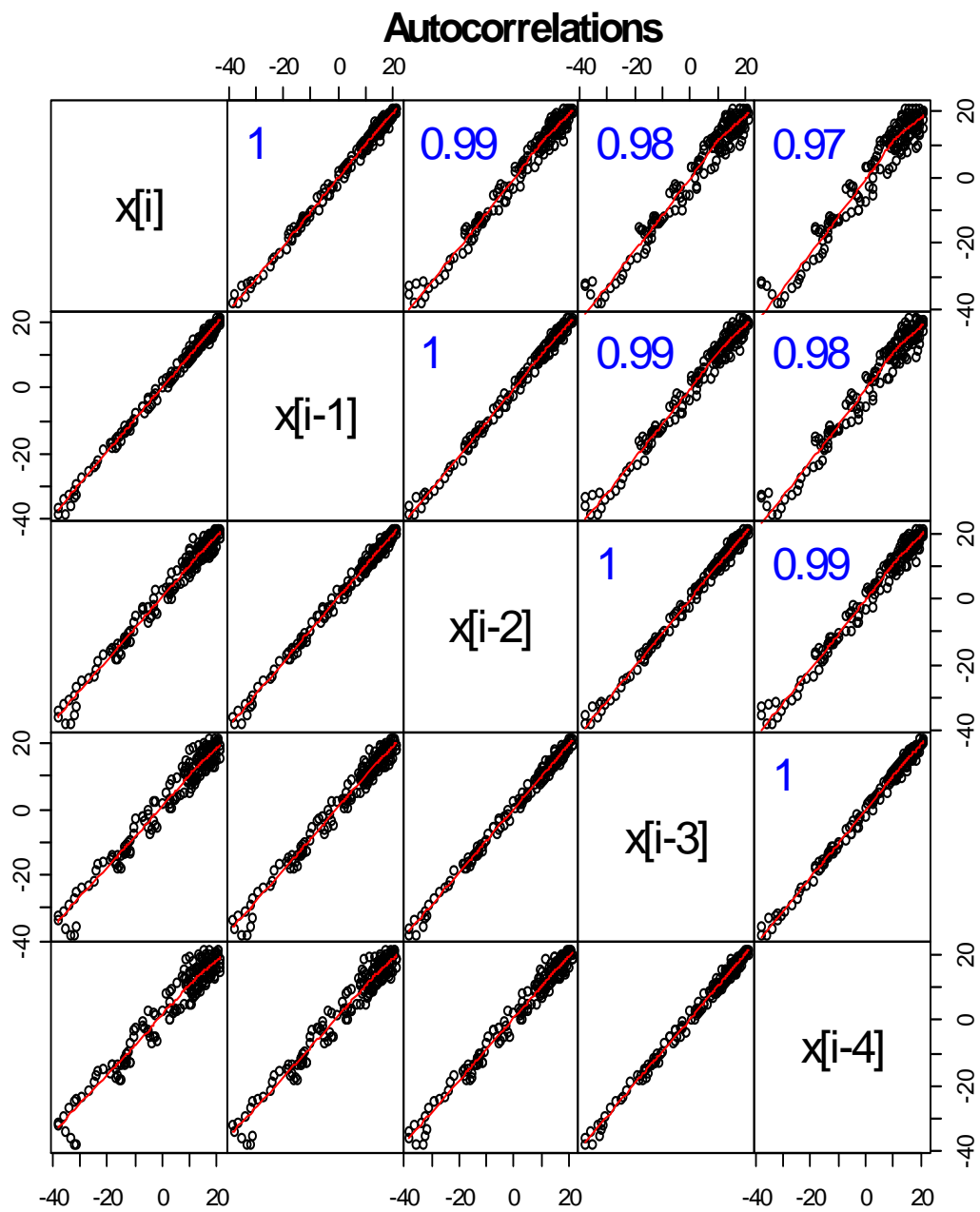
```
[1] "Normal Test, with p-value = 0,05"
```

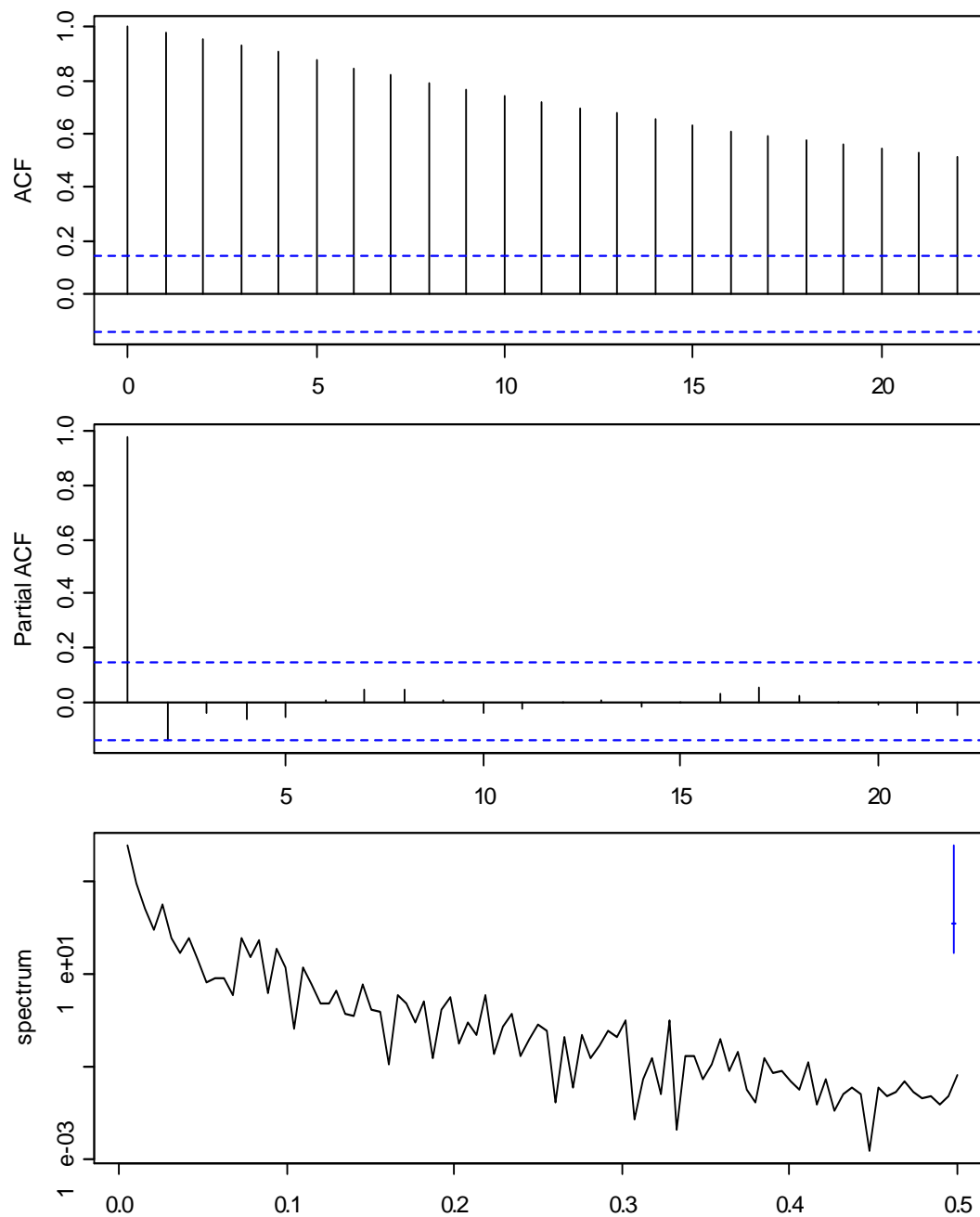
```

      Shapiro-Wilk Anderson-Darling      Cramer-von      Pearson
4.168703e-12  7.446873e-23  1.853906e-09  9.748907e-21
Shapiro-Francia  Jarque-Bera
1.849682e-10  1.882844e-08

```

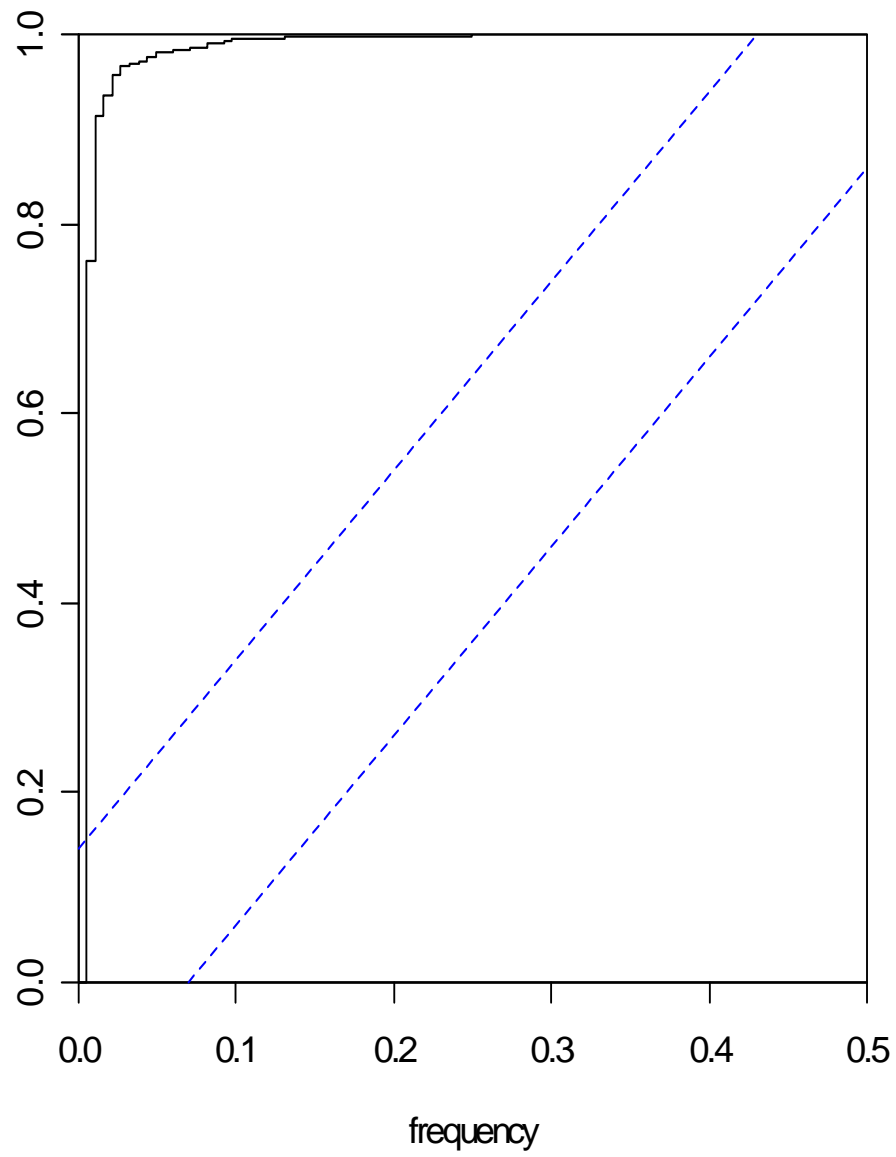
```
[1] "X" "X" "X" "X" "X" "X"
```



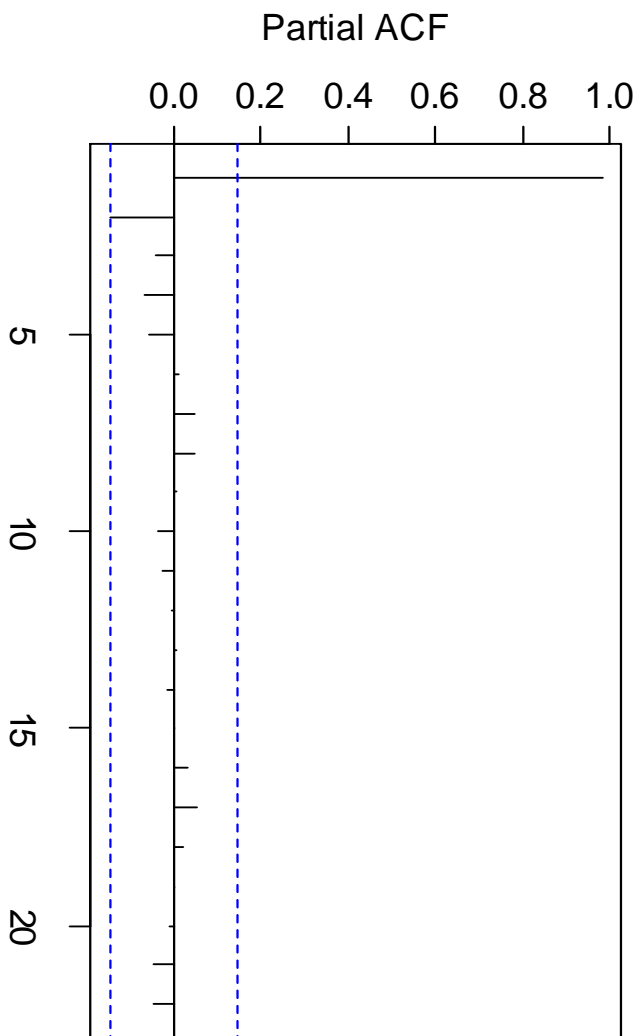
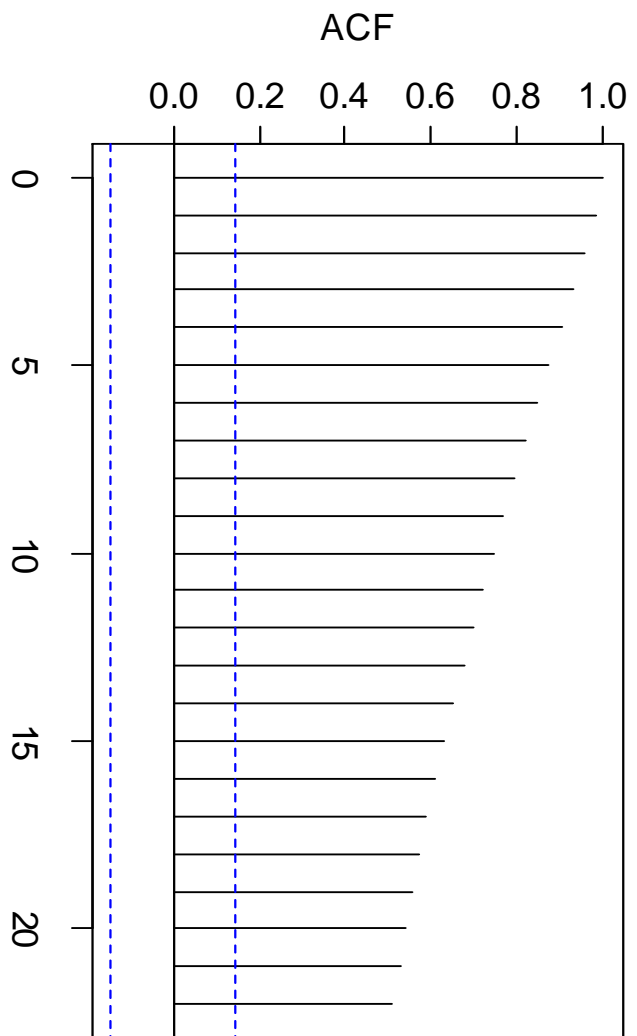


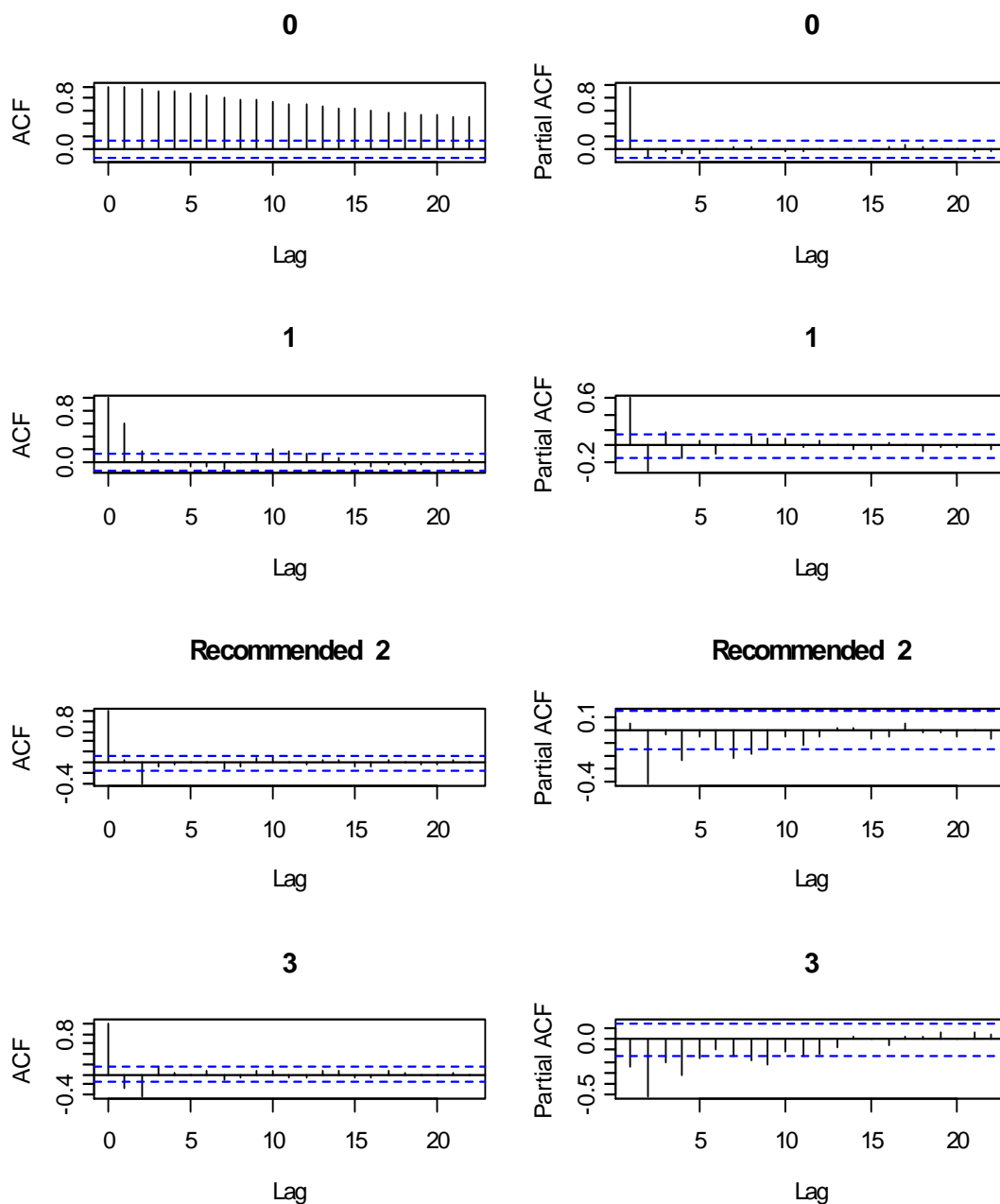
TIPO	FAS	FAP
AR(p)	Muchos coeficientes no nulos	Primeros p no nulos, resto 0
MA(q)	Primeros p no nulos, resto 0	Muchos coeficientes no nulos
ARMA(p,q)	Muchos coeficientes no nulos	Muchos coeficientes no nulos

**Series: DATA**









Call:  
`ar(x = DATOS.TS, aic = TRUE, order.max = 30)`

Coefficients:  
 1 2  
 1.1253 -0.1465

Order selected 2 sigma^2 estimated as 8.672

Series: DATOS.TS  
 ARIMA(0,2,3) model

0 1 2 3 4 5 6

608.487540	1.989651	0.000000	1.636949	2.792993	4.168163	6.157941
7	8	9	10	11	12	13
7.758947	9.398573	11.397169	13.123087	14.986914	16.981597	18.973361
14	15	16	17	18	19	20
20.924445	22.923396	24.753789	26.286378	28.213212	30.213046	32.188295
21	22	23	24	25	26	27
33.815650	35.420130	37.210982	39.210943	41.068717	43.016983	45.016637
28	29	30				
46.787496	48.669210	50.663531				

Series: DATOS.TS  
ARIMA(0,2,3) model

Coefficients:

	ma1	ma2	ma3
	-0.1083	-0.6686	-0.1696
s.e.	0.0738	0.0507	0.0691

sigma<sup>2</sup> estimated as 0.9476: log likelihood = -254.56, aic = 517.11

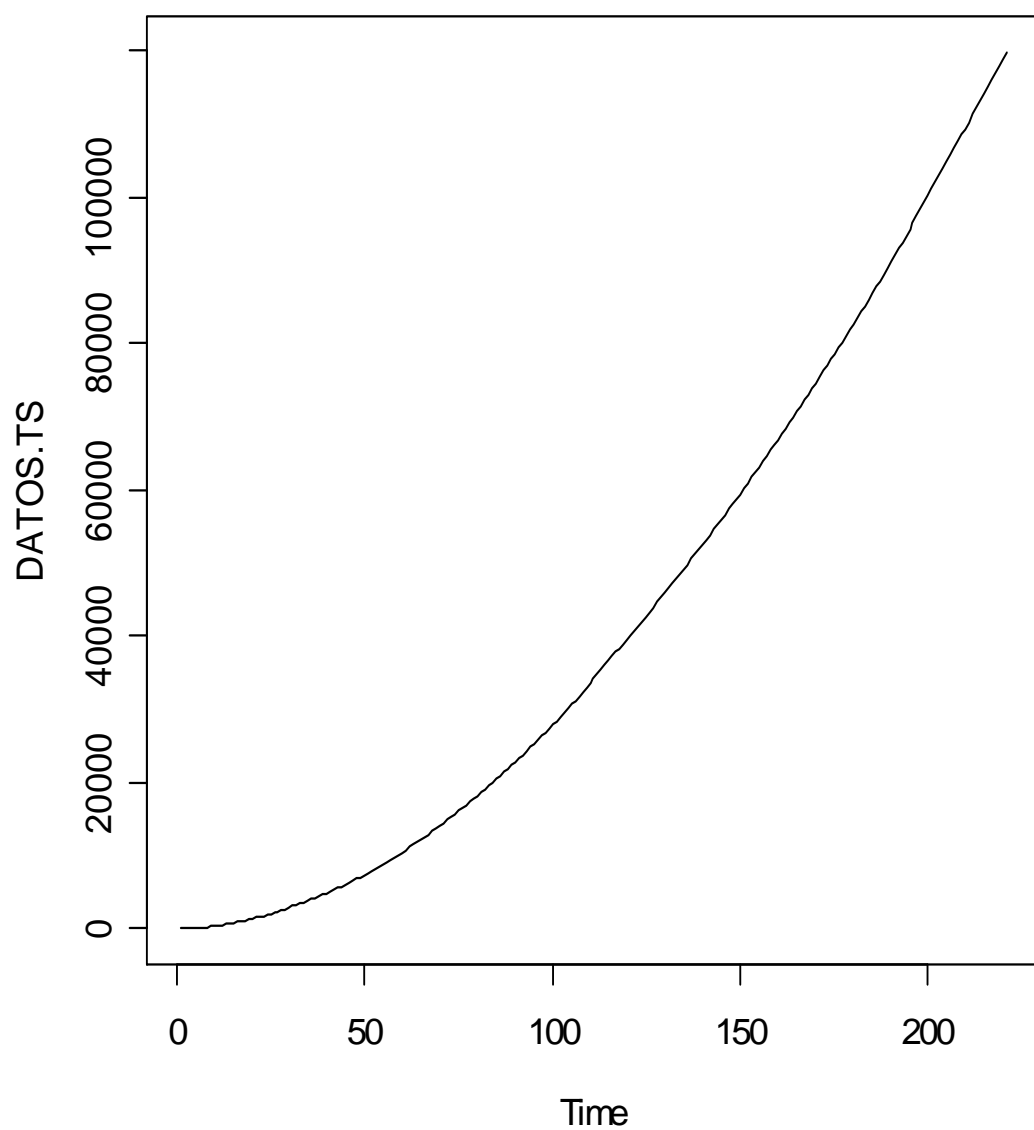
### 3.10 Serie 10

Dada la serie, se procederá a realizar el estudio de análisis de series temporales.

#### Análisis descriptivo de una serie temporal.

Todo análisis de serie temporal univariante, comienza con la presentación de un grafico donde se muestra la evolución de la variable a lo largo del tiempo.

Los resultados son los siguientes:



*Figura 10.1*

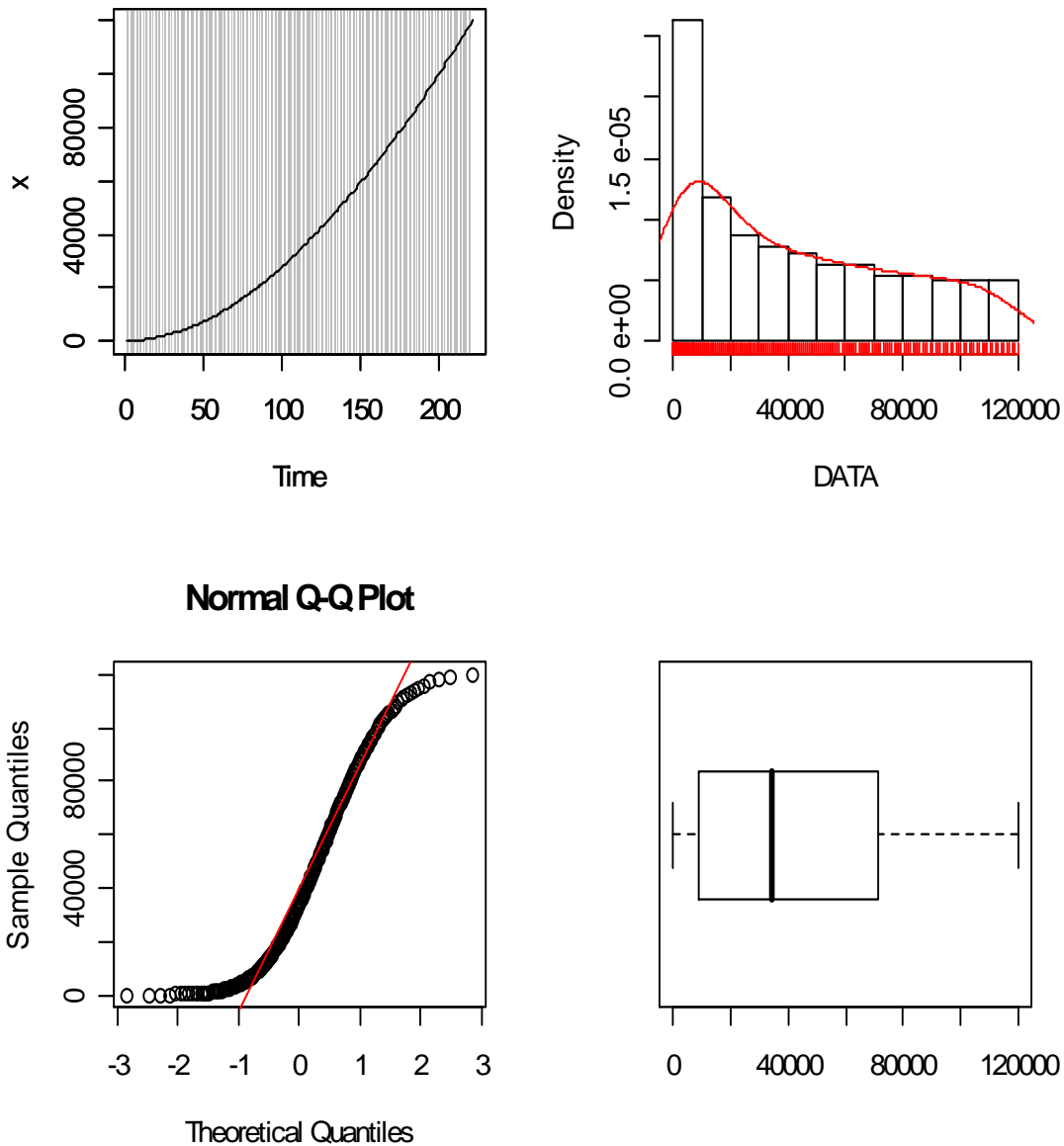


Figura 10.2

[1] "JAB: EDA VIEW 1.1"

The decimal point is 4 digit(s) to the right of the |

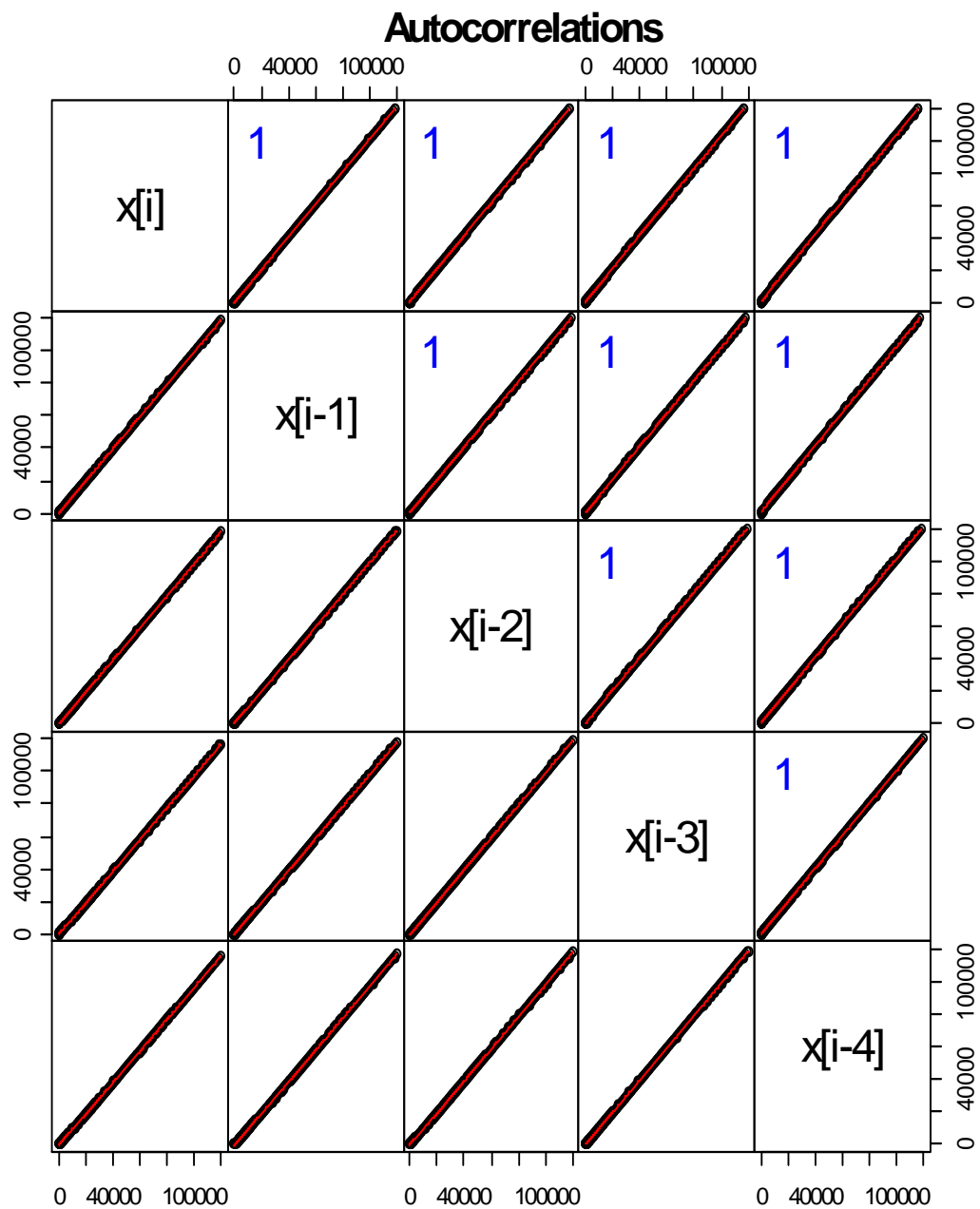
```

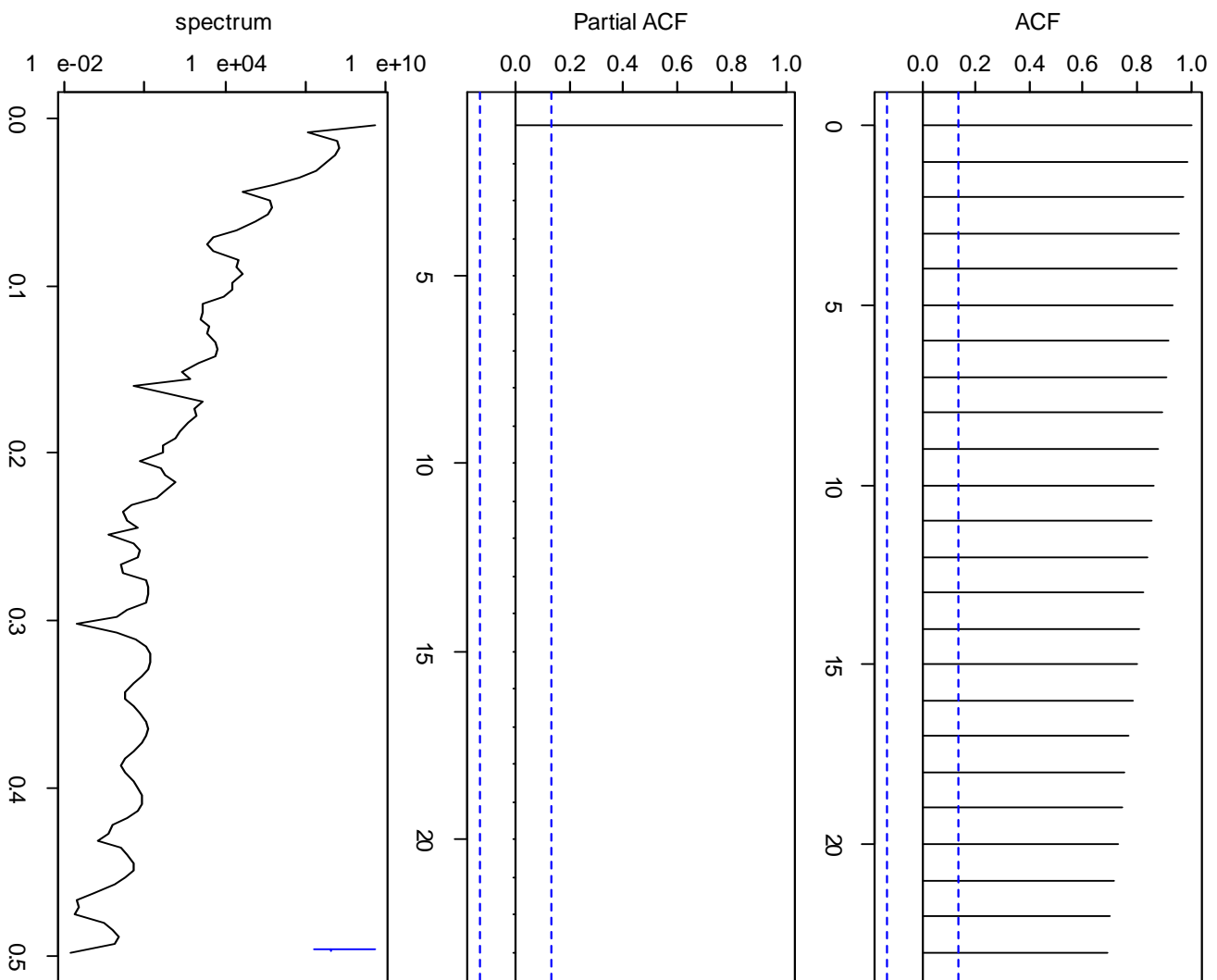
0 | 00000000000111111112222222333333444445555666677778888999
1 | 0001112233344455667778899
2 | 00011223344556677889
3 | 00112234455677889
4 | 0011233455677889
5 | 001233455677899
6 | 0122345567889
7 | 0122345567899

```

```
8 | 01233456789
9 | 001234567789
10 | 01234567889
11 | 0123456789
12 | 0
```

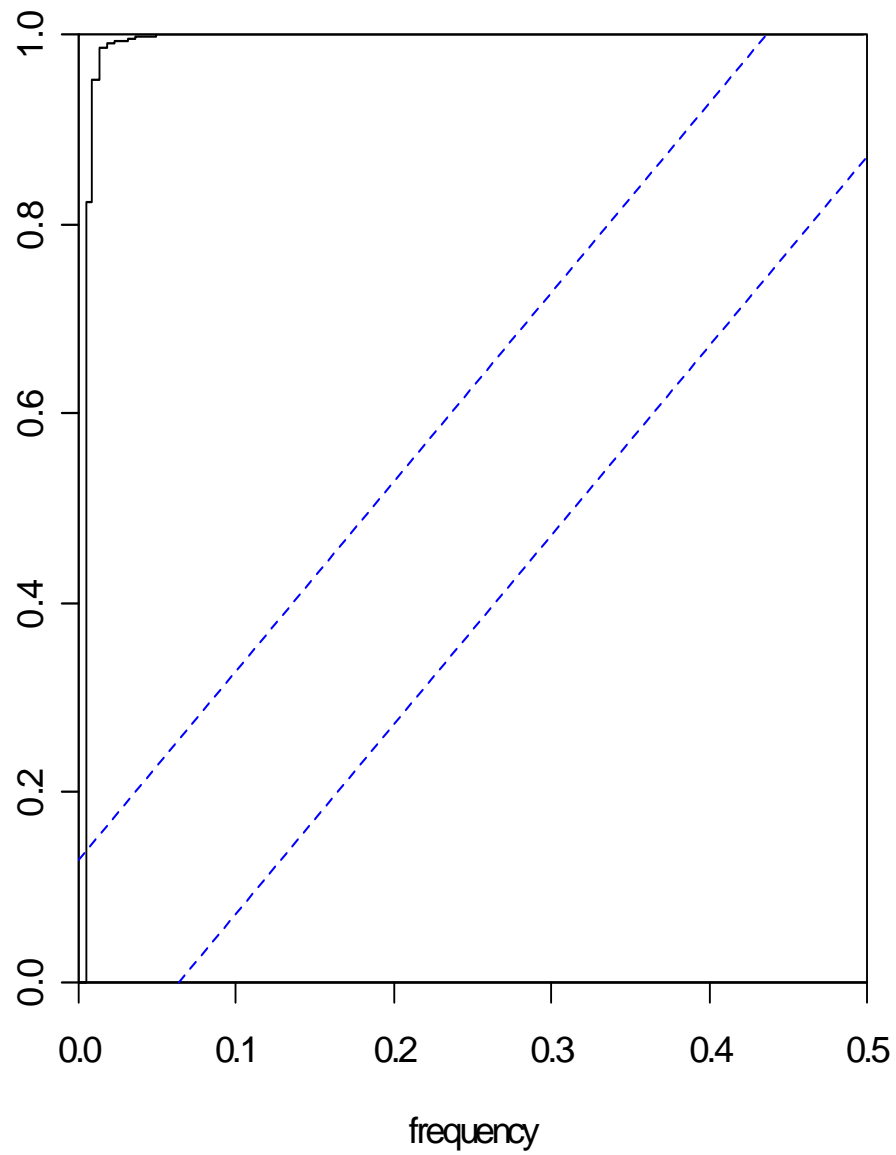
```
      MEAN      MEDIAN      SD      KURTOSIS      SKEWNESS
42669.236878 34120.862300 36389.907073 -1.008645  0.534839
[1] ""
[1] "Normal Test, with p-value = 0,05"
      Shapiro-Wilk Anderson-Darling      Cramer-von      Pearson
1.406287e-10  8.314536e-16  1.785345e-09  1.978966e-27
Shapiro-Francia  Jarque-Bera
4.443703e-09  5.235769e-05
[1] "X" "X" "X" "X" "X" "X"
```

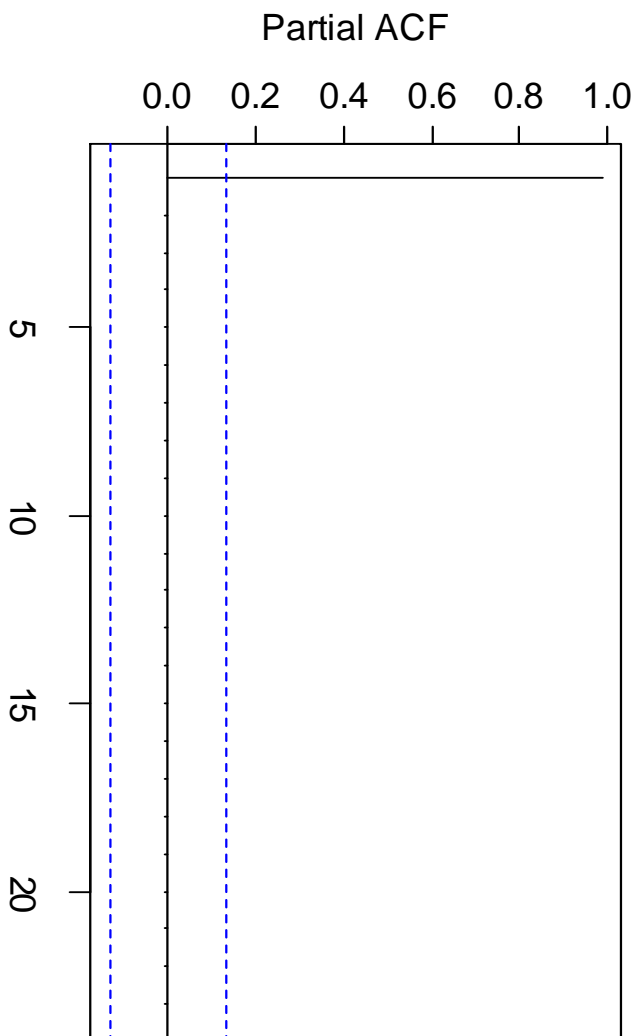
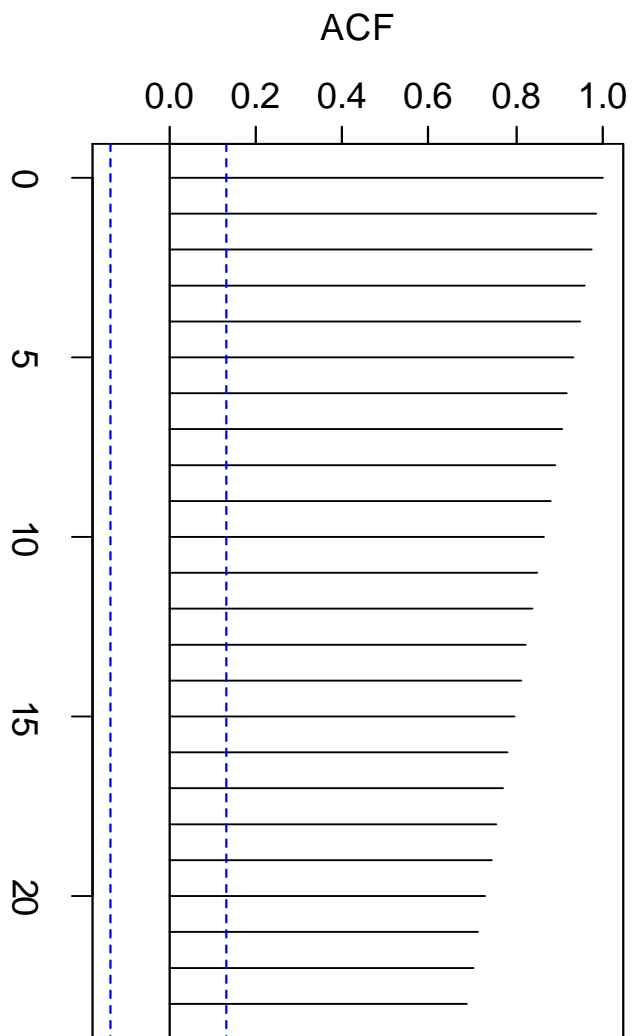


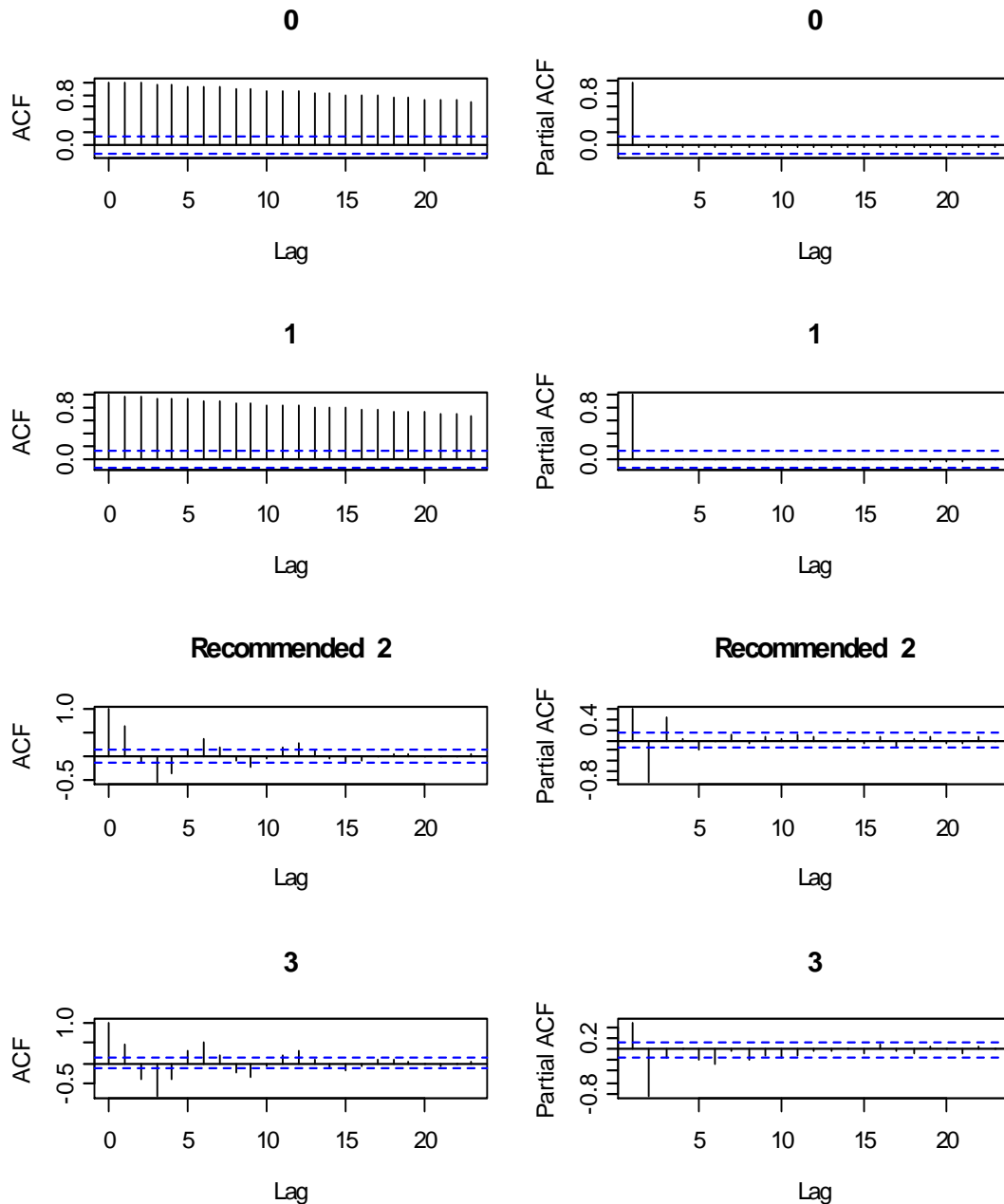




**Series: DATA**







## 5. Conclusiones

## 6. Bibliografía.

La bibliografía empleada es la siguiente:

1. Análisis preliminar de los datos de una serie temporal. A. Mendez
2. Análisis de series temporales. Daniel Peña

3. Time Series with R, Part 1, Walter Zucchini, Oleg Nenadić
4. Econometrics with R, Grant V. Farnsworth
5. Introduction to R for Times Series Analysis, G.P. Nason
6. Time Series, [http://zoonek2.free.fr/UNIX/48\\_R/15.html](http://zoonek2.free.fr/UNIX/48_R/15.html), Vincent Zoonekynd
7. TimeSeries, <http://cran.r-project.org/doc/contrib/Ricci-refcard-ts.pdf> , Vito Ricci
8. <http://cran.r-project.org/src/contrib/PACKAGES.html>
9. <http://cran.r-project.org/doc/packages/dyn.pdf>
10. <http://cran.r-project.org/doc/packages/ArDec.pdf>
11. <http://cran.r-project.org/doc/packages/urca.pdf>
12. <http://cran.r-project.org/doc/packages/forecasting.pdf>
13. <http://cran.r-project.org/doc/packages/fSeries.pdf>
14. JGR, <http://www.rosuda.org/jgr/>
15. KPSS test, <http://www.newyorkfed.org/research/economists/hobijn/kpsstest.pdf>

## 7. Anexos.

### 7.1 Scripts desarrollados para el desarrollo de los análisis

Los Scripts empleados en el análisis de series temporales han sido programados en el lenguaje R, un dialecto del lenguaje S, empleado en el motor estadístico SPLUS.

#### Script de Carga de datos y representación visual de la serie

```
#Estableciendo el directorio de trabajo
setwd("C:/DATOS/DOCUMENTACION/2006/DOCTORADO/STAT_METHODS/TIME_SERIES/SERIES/Series/");
#Cargando libreria de analisis de normalidad
library("nortest");
#Cargando libreria de EDA, desarrollada por Juan Antonio Breña Moral
source("JAB.EDA.txt");
#Cargando los datos en memorial
Serie <- read.table("Serie_10.dat",header=F,sep="\t", quote="");
#Generando una serie temporal
DATOS.TS <- ts(Serie$V1,start=start(Serie$V1), frequency = frequency(Serie$V1));

JAB.EDA(DATOS.TS)
```

#### Análisis EDA de una serie

```
#FROM: http://tolstoy.newcastle.edu.au/R/help/99b/0171.html
JAB.EDA.KURTOSIS <- function(x) {
  m4 <- mean((x-mean(x))^4)
  kurt <- m4/(sd(x)^4)-3
  kurt
}

JAB.EDA.SKEWNESS <- function(x) {
  m3 <- mean((x-mean(x))^3)
  skew <- m3/(sd(x)^3)
  skew
}

#http://zoonek2.free.fr/UNIX/48_R/15.html
plot.band <- function (x, ...) {
  plot(x, ...)
  a <- time(x)
  i1 <- floor(min(a))
  i2 <- ceiling(max(a))
  y1 <- par('usr')[3]
  y2 <- par('usr')[4]
```

```

if( par("ylog") ){
  y1 <- 10^y1
  y2 <- 10^y2
}
for (i in seq(from=i1, to=i2-1, by=2)) {
  polygon( c(i,i+1,i+1,i),
           c(y1,y1,y2,y2),
           col = 'grey',
           border = NA )
}
par(new=T)
plot(x, ...)
}

JAB.EDA.GRAPH <- function(DATA){
  par(mfrow= c(2,2));
  plot.band(DATA);
  #plot(DATA,type="l");
  hist(DATA, probability =T, main='');
  rug(x, quiet = getOption("warn") < 0, col=2);
  lines(density(DATA), col="red");
  qqnorm(DATA);
  qqline(DATA, col = 2)
  boxplot(DATA, horizontal=T);
  par(mfrow= c(1,1));
  stem(DATA)
}

JAB.EDA.TABLE <- function(DATA){
  COLUMN_NAMES <- c("MEAN", "MEDIAN", "SD", "KURTOSIS", "SKEWNESS");
  DATA_TABLE <- c();
  DATA_TABLE[1] <- mean(DATA);
  DATA_TABLE[2] <- median(DATA);
  DATA_TABLE[3] <- sd(DATA);
  DATA_TABLE[4] <- JAB.EDA.KURTOSIS(DATA);
  DATA_TABLE[5] <- JAB.EDA.SKEWNESS(DATA);
  names(DATA_TABLE) <- COLUMN_NAMES;
  print(DATA_TABLE);
}

JAB.EDA.NORMAL_TEST <- function(DATA){
  library("nortest");
  PVALUE <- 0.05;
  print("Normal Test, with p-value = 0,05");
  COLUMN_NAMES <- c("Shapiro-Wilk", "Anderson-Darling", "Cramer-von", "Pearson", "Shapiro-
Francia");
  DATA_TABLE <- c();
  DATA_TABLE_KPI <- c();
  SHAPIRO <- shapiro.test(DATA);
  AD <- ad.test(DATA);
  KRAMER <- cvm.test(DATA);
  PEARSON <- pearson.test(DATA);
  SF <- sf.test(DATA);
  DATA_TABLE[1] <- SHAPIRO$p.value;
  DATA_TABLE[2] <- AD$p.value;
  DATA_TABLE[3] <- KRAMER$p.value;
  DATA_TABLE[4] <- PEARSON$p.value;
  DATA_TABLE[5] <- SF$p.value;
  names(DATA_TABLE) <- COLUMN_NAMES;
  print(DATA_TABLE);
  DATA_TABLE_KPI[1] <- if(PVALUE >= SHAPIRO$p.value){ "OK" }else{ "X" };
  DATA_TABLE_KPI[2] <- if(PVALUE >= AD$p.value){ "OK" }else{ "X" };
  DATA_TABLE_KPI[3] <- if(PVALUE >= KRAMER$p.value){ "OK" }else{ "X" };
  DATA_TABLE_KPI[4] <- if(PVALUE >= PEARSON$p.value){ "OK" }else{ "X" };
  DATA_TABLE_KPI[5] <- if(PVALUE >= SF$p.value){ "OK" }else{ "X" };
  print(DATA_TABLE_KPI);
}

JAB.EDA <- function(DATA){
  print("JAB: EDA VIEW 1.2");
  JAB.EDA.GRAPH(DATA);
  JAB.EDA.TABLE(DATA);
  JAB.EDA.NORMAL_TEST(DATA);
}

```

## Extraer partes de una muestra

```
DATOS_1 <- Serie_01[1:100,];  
DATOS_1.TS = ts(DATOS_1);  
plot(DATOS_1.TS)
```

## Descomposición de una serie

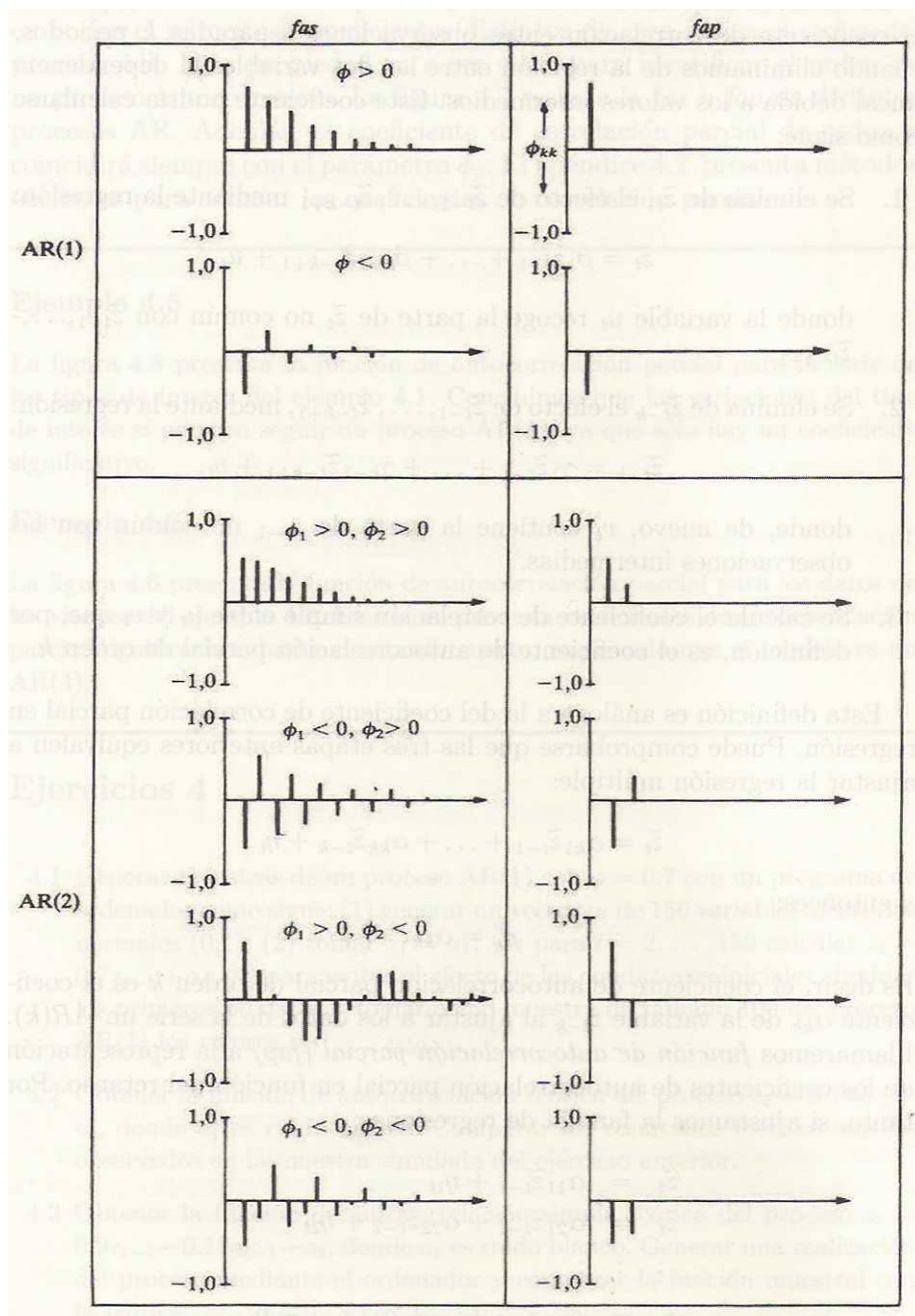
```
DATOS_2<- Serie_01[101:150,];  
DATOS_2.TS <- ts(DATOS_2, start = 1, frequency= 17);  
plot(stl(DATOS_2.TS, s.window="periodic"));
```

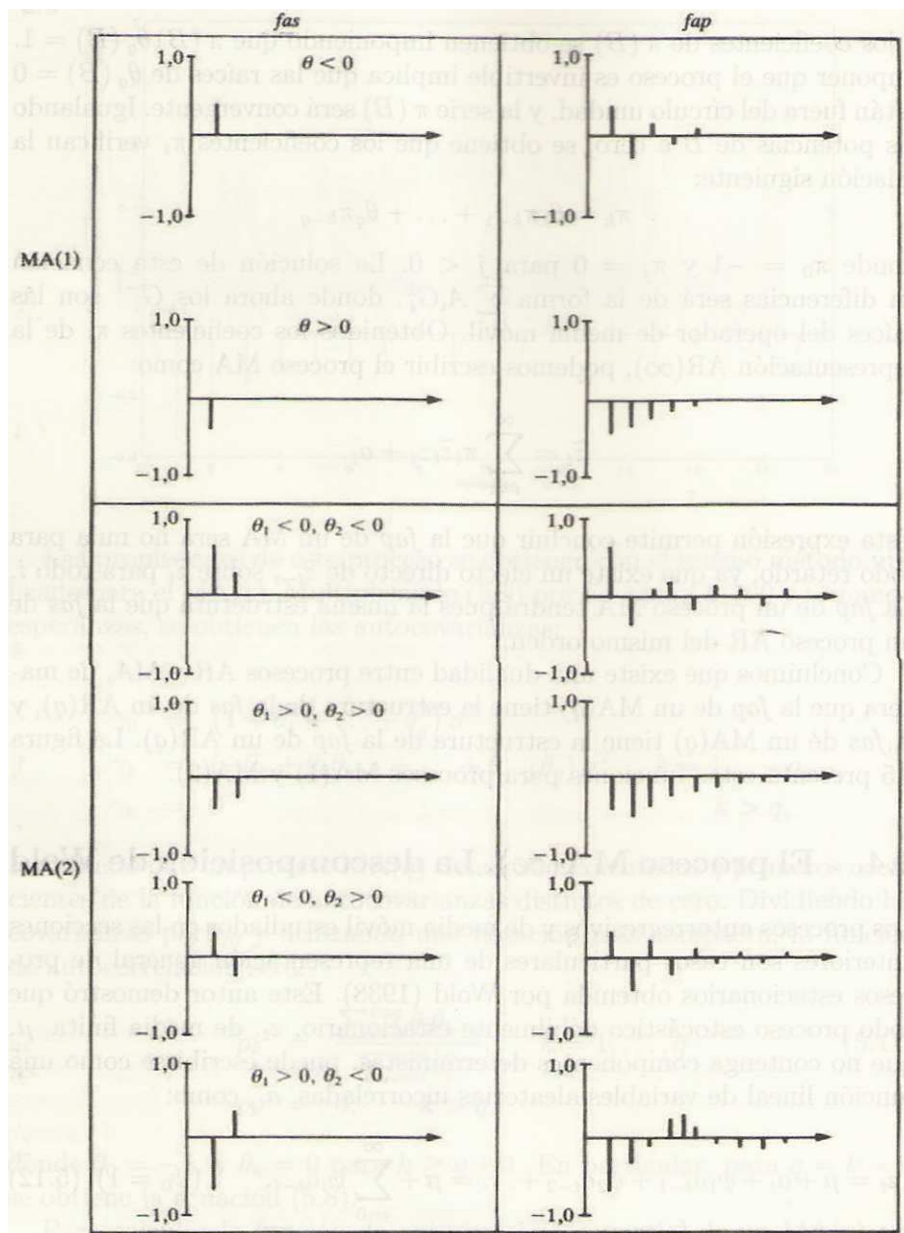
## Inspección de los datos desde el punto de vista de serie temporal

```
JAB.TS.VIEW <- function(DATA){  
  op <- par(mfrow = c(3,1), mar = c(2,4,1,2)+.1);  
  acf(DATA, xlab="");  
  pacf(DATA, xlab="");  
  spectrum(DATA, xlab="", main="");  
  par(op);  
}  
  
JAB.TS.DIFFERENCES <- function(DATA){  
  N <- ndiffs(DATA);  
  if(N < 1){  
    k <- 2;  
  }else{  
    k <- N+1;  
  }  
  op <- par(mfrow=c(k,2));  
  for(i in 1:k) {  
    if(N == i-1){  
      acf(DATA, main=paste("Recommended ",i-1));  
      pacf(DATA, main=paste("Recommended ",i-1));  
    }else{  
      acf(DATA, main=i-1);  
      pacf(DATA, main=i-1);  
    }  
    DATA <- diff(DATA);  
  }  
  par(op);  
}
```

## 7.2 Modelos AR y MA Teóricos

A continuación se muestran los modelos teóricos de AR MA





TIPO	FAS	FAP
AR(p)	Muchos coeficientes no nulos	Primeros p no nulos, resto 0
MA(q)	Primeros p no nulos, resto 0	Muchos coeficientes no nulos
ARMA(p,q)	Muchos coeficientes no nulos	Muchos coeficientes no nulos