Time series applied to finance

Marouane Arab

March 28, 2016

Abstract

This documents is generated by an R script and a brew file (latex file which embeds R code). All sources are available at `https://github.com/arabm/multifractal_model`. This article contains some details about existing financial modelisation, and some examples.

It has :

- modelisation examples on financial stock, exchange rates (ARMA, Markov Switching Model)
- ARMA model presentation

It still misses :

- statistical test explanation (Student, p-value,...)
- deeper details on ARMA model
- GARCH model
- MSM model
- forecasting, backtesting

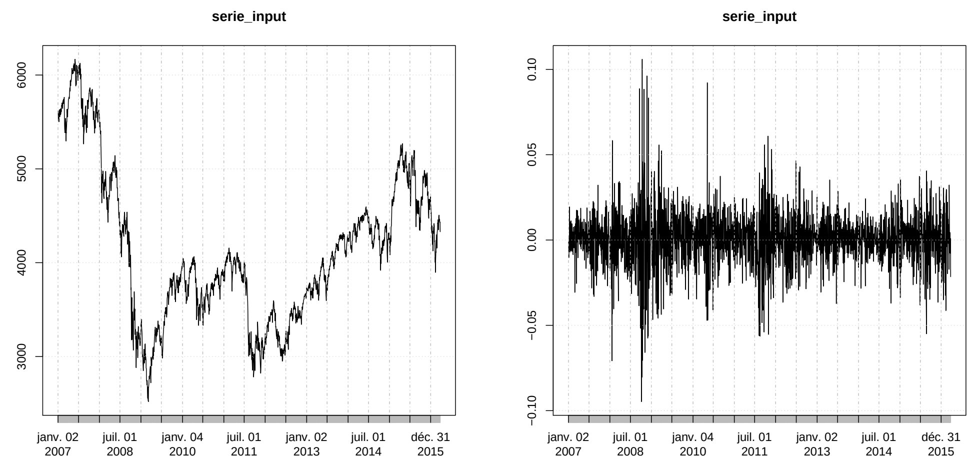
Question to answer :

- Why working on log return ?
- How an ARMA/Garch/MSM model can be accepted/rejected ?
- What are tests that can be done ?
- How validate backtesting ?
- Why doing forecasting ? How much time in prevision is relevant (1 day, 1 month...)?

Contents

1 CAC40

This is the historical close quotation. S_t. A stationarize are the log return. $r_t = \log(\frac{S_t}{S_{t-1}})$



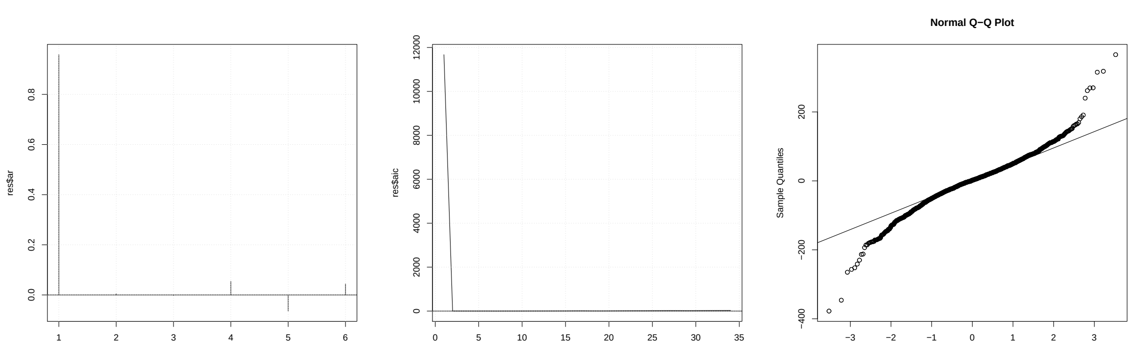
(a) Close level of CAC40

(b) log return CAC40

Looking at original

Is it an AR model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm. Results:

```
Call:
ar(x = value)

Coefficients:
      1        2        3        4        5        6
 0.9586   0.0052  -0.0014   0.0545  -0.0643   0.0442

Order selected 6  sigma^2 estimated as  4526
```



(a) AR coefficients

(b) AIC per number of parameters

(c) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.34186, df = 1, p-value = 0.5588
```

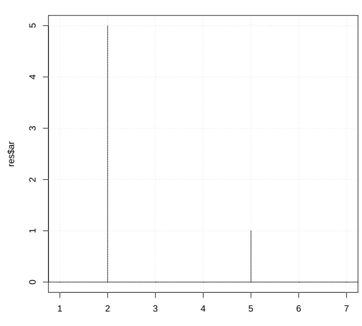
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.9639, p-value < 2.2e-16
```

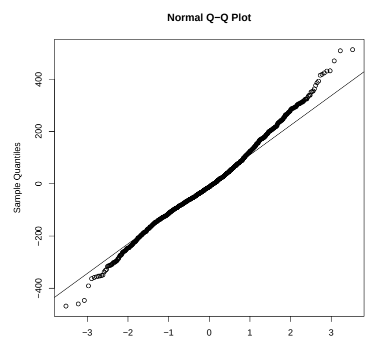
Is it an MA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,5) with non-zero mean

Coefficients:
         ma1     ma2     ma3     ma4     ma5  intercept
      1.7514  2.1009  1.9138  1.3178  0.5338  4139.6223
s.e.  0.0265  0.0439  0.0337  0.0234  0.0180    22.3093

sigma^2 estimated as 15867:  log likelihood=-14782.85
AIC=29579.69   AICc=29579.74   BIC=29620.07
```



(a) MA coefficients



(b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 209.85, df = 1, p-value < 2.2e-16
```

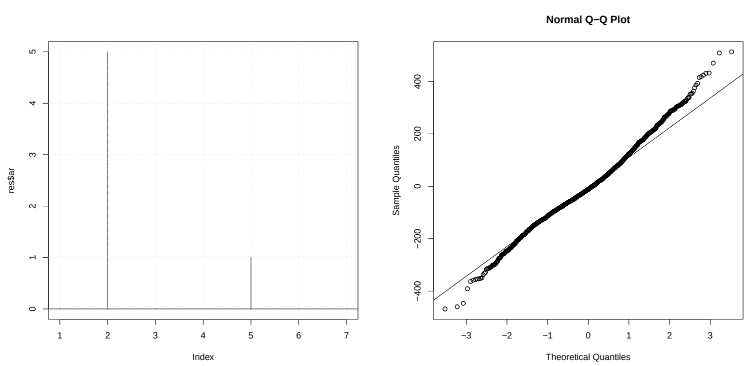
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.98947, p-value = 3.684e-12
```

Is it an ARMA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,5) with non-zero mean

Coefficients:
         ma1     ma2     ma3     ma4     ma5  intercept
      1.7514  2.1009  1.9138  1.3178  0.5338  4139.6223
s.e.  0.0265  0.0439  0.0337  0.0234  0.0180    22.3093

sigma^2 estimated as 15867:  log likelihood=-14782.85
AIC=29579.69   AICc=29579.74   BIC=29620.07
```



(a) MA coefficients

(b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 209.85, df = 1, p-value < 2.2e-16
```

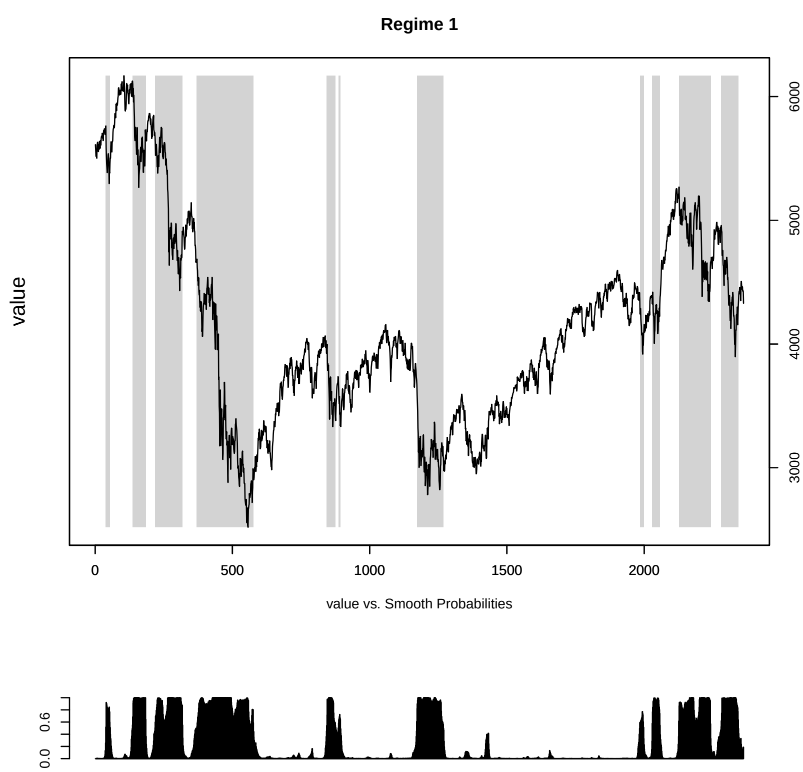
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.98947, p-value = 3.684e-12
```

Markov Switching Model Try with 2 regimes.

Results:

```
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(T, T, T), p = 1, control = list(parallel = F))

       AIC      BIC    logLik
  25609.13 25663.27 -12800.56

Coefficients:
        (Intercept)(S) value_1(S)   Std(S)
Model 1      18.121358  0.9938143 86.46705
Model 2       7.304968  0.9989027 42.64456

Transition probabilities:
          Regime 1  Regime 2
Regime 1 0.9788516 0.0093812
Regime 2 0.0211484 0.9906188
```

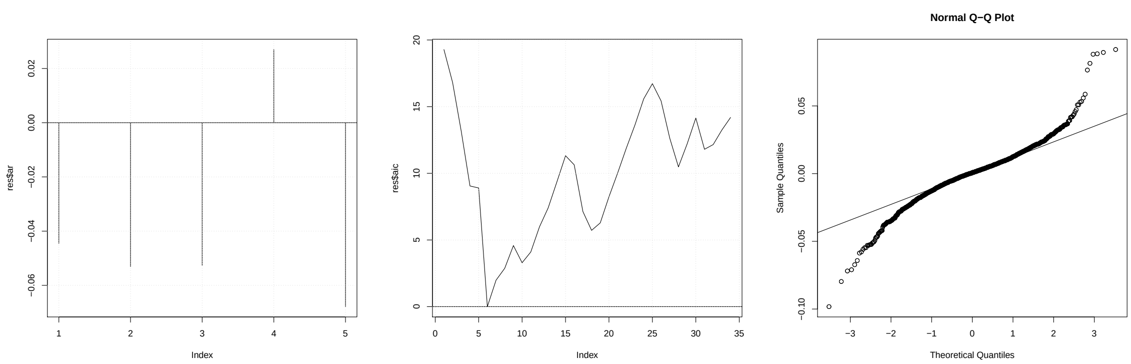


(a) Which 2

Looking at logret

Is it an AR model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm. Results:

```
Call:
ar(x = value)

Coefficients:
      1        2        3        4        5
-0.0446  -0.0531  -0.0526   0.0270  -0.0679

Order selected 5  sigma^2 estimated as  0.0002385
```



(a) AR coefficients (b) AIC per number of parameters (c) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.00024827, df = 1, p-value = 0.9874
```

Test residual normal law with Shapiro:

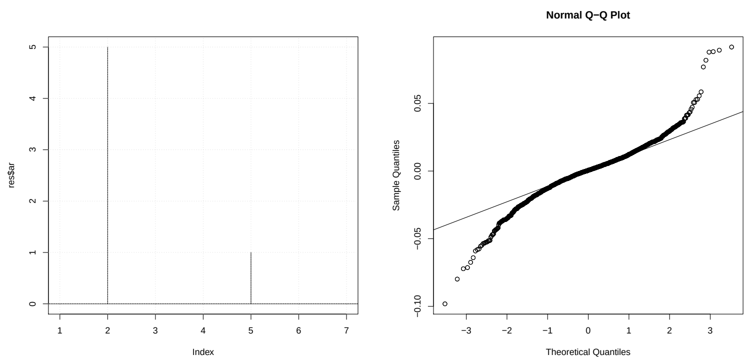
```
Shapiro-Wilk normality test

data:  res
W = 0.94774, p-value < 2.2e-16
```

Is it an MA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,5) with zero mean
```

```
Coefficients:
          ma1      ma2      ma3     ma4      ma5
      -0.0446  -0.0485  -0.0482  0.0307  -0.0627
s.e.   0.0205   0.0207   0.0205  0.0205   0.0211

sigma^2 estimated as 0.000238:  log likelihood=6501.84
AIC=-12991.69   AICc=-12991.65   BIC=-12957.08
```



(a) MA coefficients (b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.00020489, df = 1, p-value = 0.9886
```

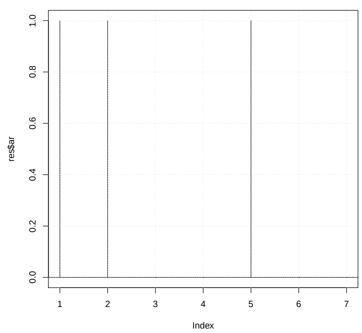
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.94707, p-value < 2.2e-16
```

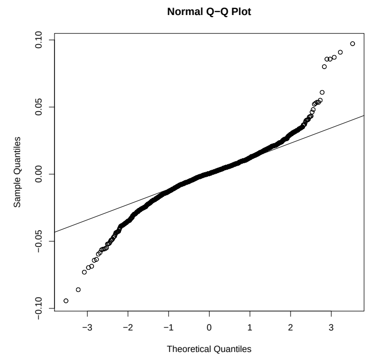
Is it an ARMA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,5) with zero mean

Coefficients:
          ma1      ma2      ma3     ma4      ma5
      -0.0446  -0.0485  -0.0482  0.0307  -0.0627
s.e.   0.0205   0.0207   0.0205  0.0205   0.0211

sigma^2 estimated as 0.000238:  log likelihood=6501.84
AIC=-12991.69   AICc=-12991.65   BIC=-12957.08
```



(a) MA coefficients



(b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.12362, df = 1, p-value = 0.7251
```

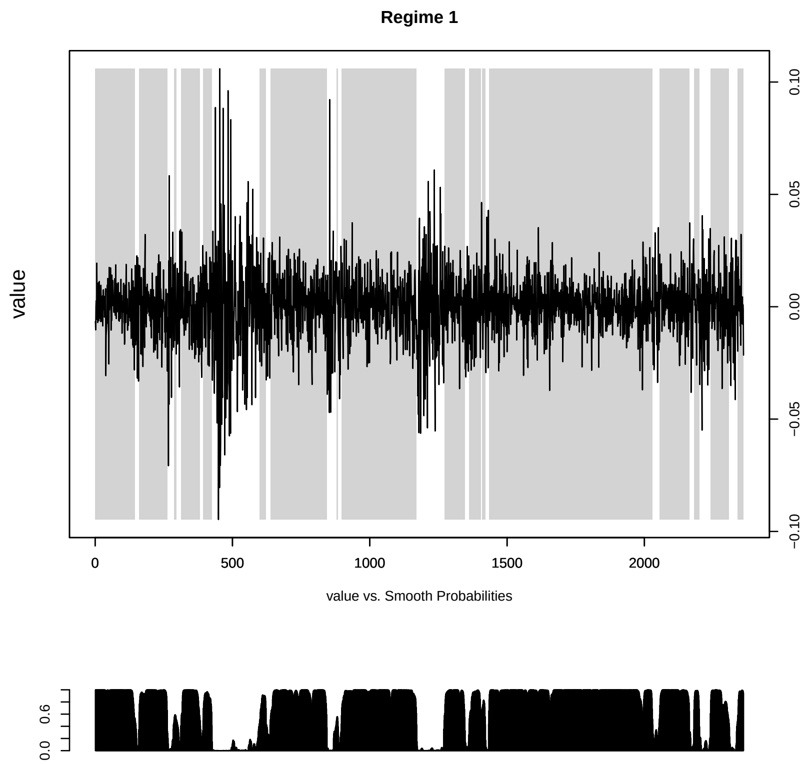
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.94391, p-value < 2.2e-16
```

Markov Switching Model Try with 2 regimes.

Results:

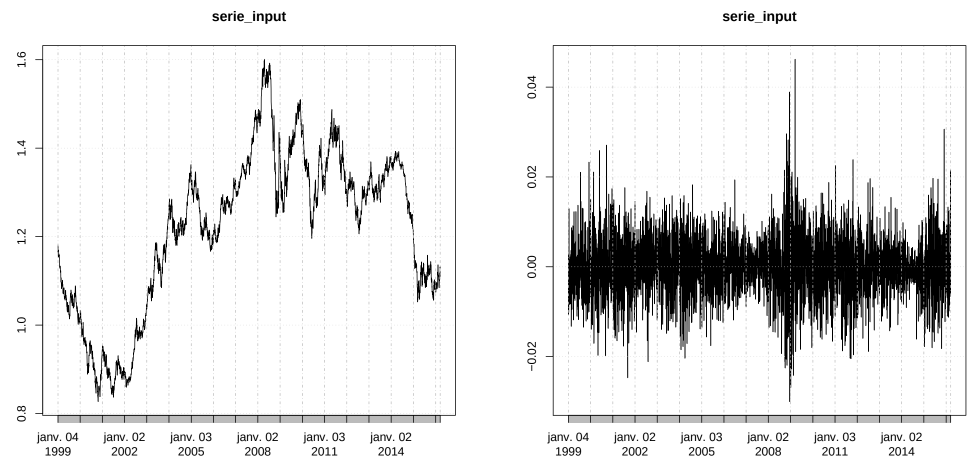
```
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(T, T, T), p = 1, control = list(parallel = F))

        AIC       BIC   logLik
  -13538.51 -13484.38 6773.255

Coefficients:
        (Intercept)(S)  value_1(S)     Std(S)
Model 1   0.0005173361 -0.03039309 0.01084260
Model 2  -0.0021582307 -0.06141297 0.02489934

Transition probabilities:
           Regime 1   Regime 2
Regime 1 0.98836266 0.03654694
Regime 2 0.01163734 0.96345306
```


Regime 1
value
0.10
0.05
0.00
−0.05
−0.10
0
500
1000
1500
2000
value vs. Smooth Probabilities
0.6
0.0

(a) Which 2

2 EURUSD

This is the historical close quotation. S_t. A stationarize are the log return. $r_t = \log(\frac{S_t}{S_{t-1}})$



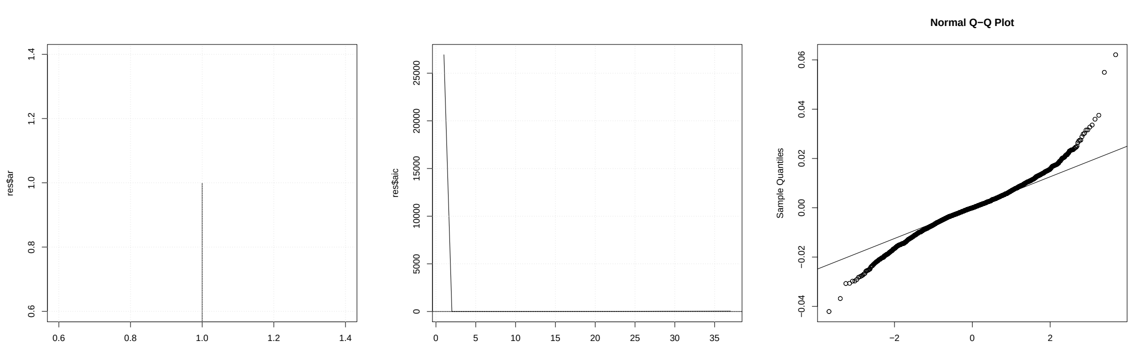
(a) Close level of EURUSD

(b) log return EURUSD

Looking at original

Is it an AR model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm. Results:

```
Call:
ar(x = value)

Coefficients:
    1
0.999

Order selected 1  sigma^2 estimated as  6.313e-05
```



(a) AR coefficients

(b) AIC per number of parameters

(c) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.17242, df = 1, p-value = 0.678
```

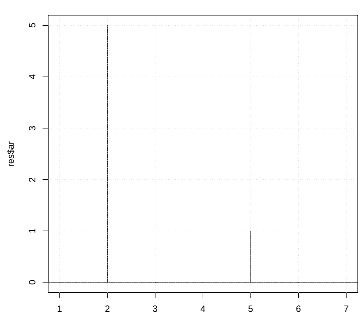
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.97606, p-value < 2.2e-16
```

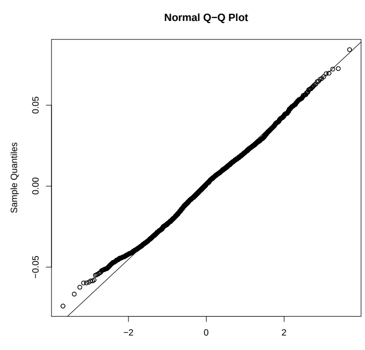
Is it an MA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,5) with non-zero mean

Coefficients:
         ma1     ma2     ma3     ma4     ma5  intercept
      2.1483  2.8344  2.6013  1.6428  0.5836     1.2171
s.e.  0.0182  0.0336  0.0296  0.0182  0.0116     0.0036

sigma^2 estimated as 0.0004708:  log likelihood=10429.22
AIC=-20844.43   AICc=-20844.4   BIC=-20799.82
```



(a) MA coefficients



(b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 479.06, df = 1, p-value < 2.2e-16
```

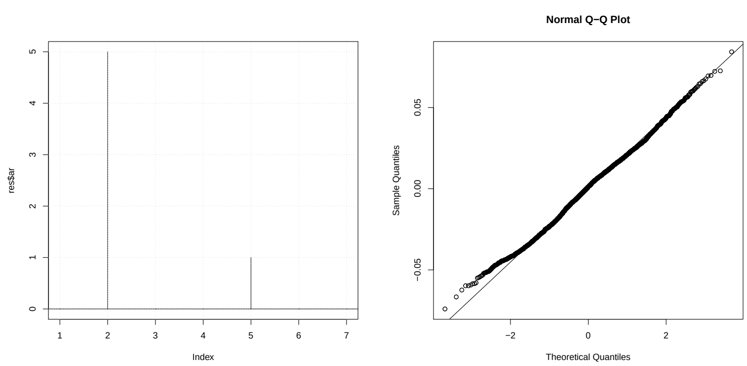
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.99702, p-value = 1.523e-07
```

Is it an ARMA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,5) with non-zero mean

Coefficients:
         ma1     ma2     ma3     ma4     ma5  intercept
      2.1483  2.8344  2.6013  1.6428  0.5836     1.2171
s.e.  0.0182  0.0336  0.0296  0.0182  0.0116     0.0036

sigma^2 estimated as 0.0004708:  log likelihood=10429.22
AIC=-20844.43   AICc=-20844.4   BIC=-20799.82
```



(a) MA coefficients

(b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 479.06, df = 1, p-value < 2.2e-16
```

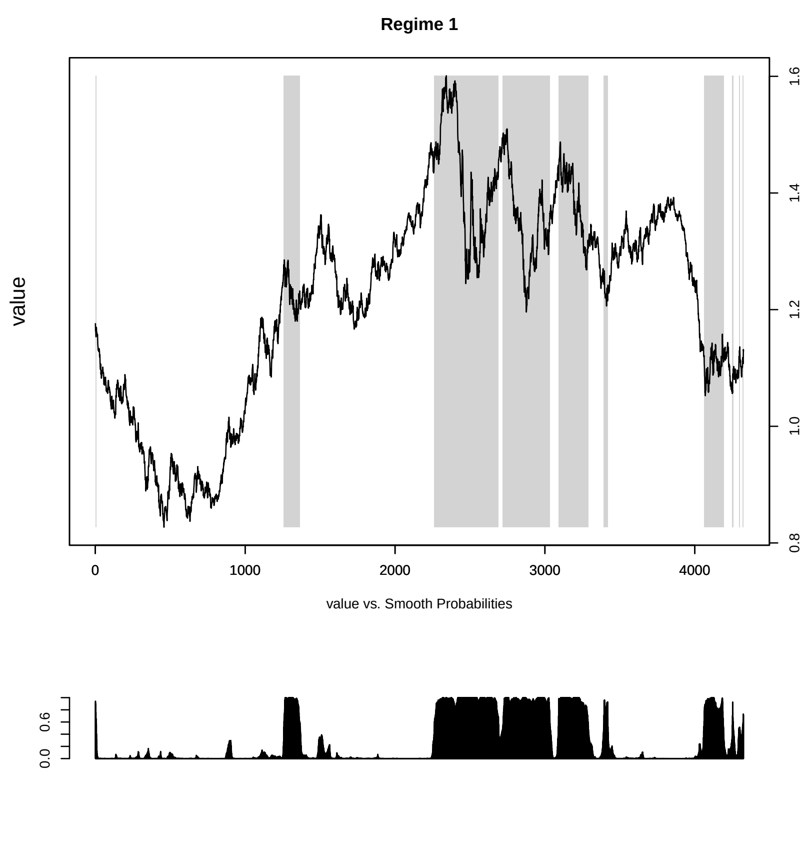
Test residual normal law with Shapiro:


```
Shapiro-Wilk normality test

data:  res
W = 0.99702, p-value = 1.523e-07
```

Markov Switching Model Try with 2 regimes.

Results:

```
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(T, T, T), p = 1, control = list(parallel = F))

        AIC      BIC   logLik
  -30176.57 -30117.6 15092.29

Coefficients:
        (Intercept)(S) value_1(S)      Std(S)
Model 1   0.0051391138  0.9959467 0.010649831
Model 2  -0.0001810602  1.0002282 0.006159484

Transition probabilities:
           Regime 1    Regime 2
Regime 1 0.98867243 0.005043856
Regime 2 0.01132757 0.994956144
```

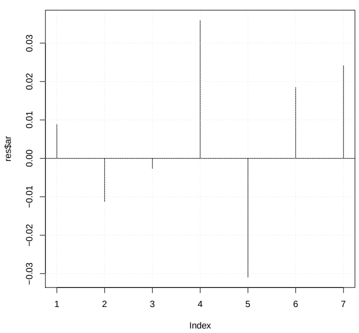


(a) Which 2

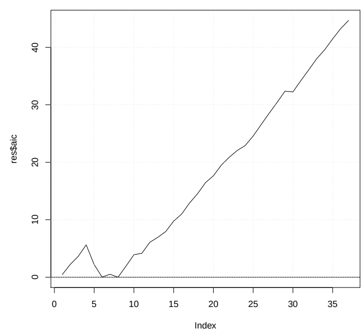
Looking at logret

Is it an AR model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm. Results:

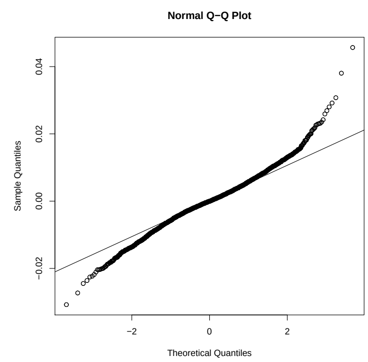
```
Call:
ar(x = value)

Coefficients:
      1        2        3        4        5        6        7
 0.0088  -0.0113  -0.0026   0.0359  -0.0309   0.0185   0.0241

Order selected 7  sigma^2 estimated as  4.069e-05
```



(a) AR coefficients



(b) AIC per number of parameters



(c) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 1.1456e-05, df = 1, p-value = 0.9973
```

Test residual normal law with Shapiro:

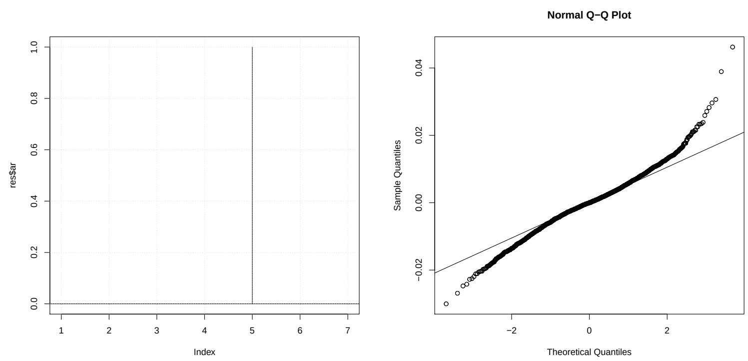
```
Shapiro-Wilk normality test

data:  res
W = 0.98309, p-value < 2.2e-16
```

Is it an MA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

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

```
Coefficients:
      intercept
          0e+00
s.e.      1e-04

sigma^2 estimated as 4.075e-05:  log likelihood=15721.53
AIC=-31439.06   AICc=-31439.06   BIC=-31426.32
```



(a) MA coefficients (b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.24259, df = 1, p-value = 0.6223
```

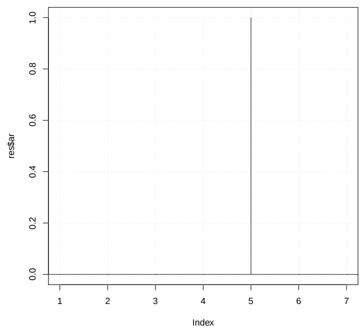
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.98276, p-value < 2.2e-16
```

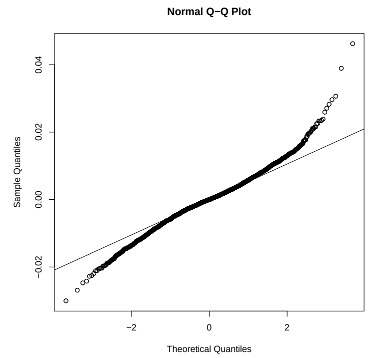
Is it an ARMA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

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

Coefficients:
      intercept
          0e+00
s.e.      1e-04

sigma^2 estimated as 4.075e-05:  log likelihood=15721.53
AIC=-31439.06   AICc=-31439.06   BIC=-31426.32
```



(a) MA coefficients



(b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.24259, df = 1, p-value = 0.6223
```

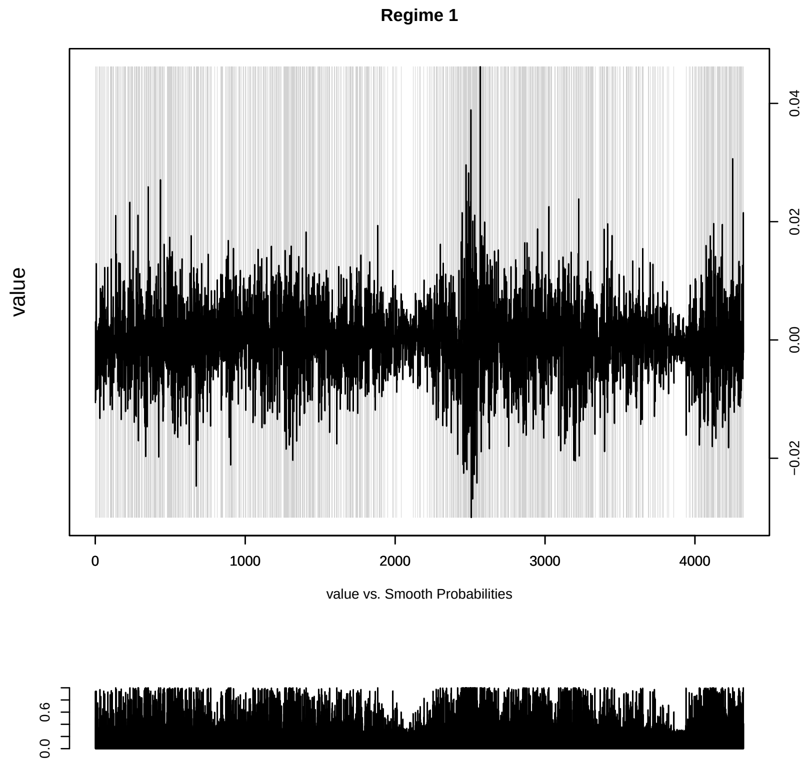
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.98276, p-value < 2.2e-16
```

Markov Switching Model Try with 2 regimes.

Results:

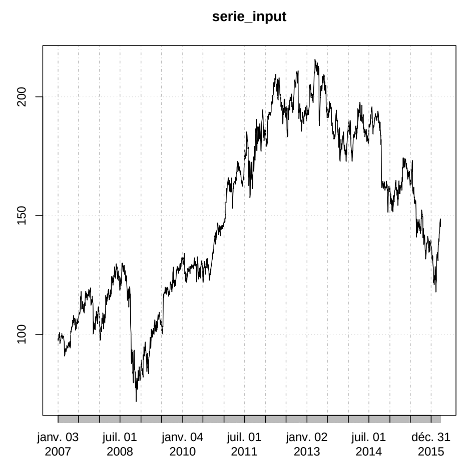
```
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(T, T, T), p = 1, control = list(parallel = F))

        AIC      BIC   logLik
  -31659.28 -31600.3 15833.64

Coefficients:
        (Intercept)(S)  value_1(S)      Std(S)
Model 1  -0.0002184246  0.06645888 0.008243233
Model 2   0.0001969037 -0.06868736 0.003960878

Transition probabilities:
          Regime 1  Regime 2
Regime 1 0.6089057 0.3545759
Regime 2 0.3910943 0.6454241
```


Regime 1
value
0.04
0.02
0.00
−0.02
0
1000
2000
3000
4000
value vs. Smooth Probabilities
0.6
0.0

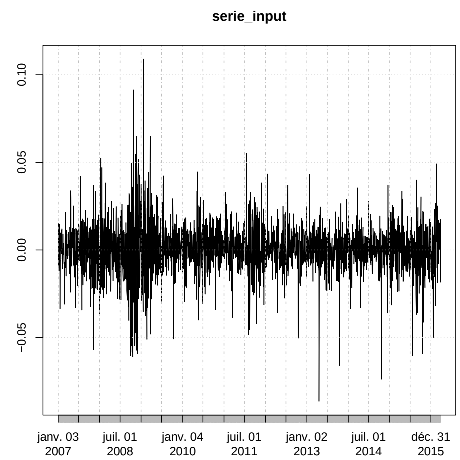
(a) Which 2

3 IBM

This is the historical close quotation. S_t. A stationarize are the log return. $r_t = \log(\frac{S_t}{S_{t-1}})$



(a) Close level of IBM

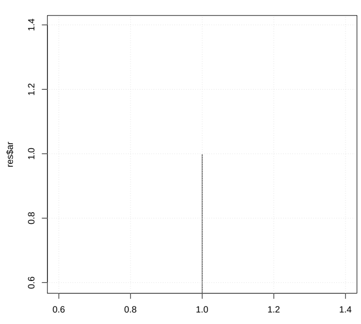


(b) log return IBM

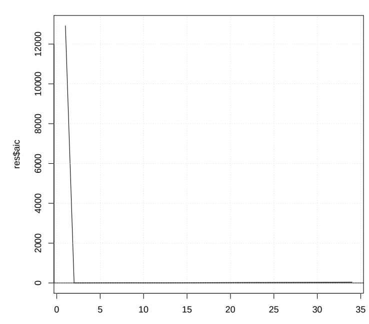
Looking at original

Is it an AR model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm. Results:

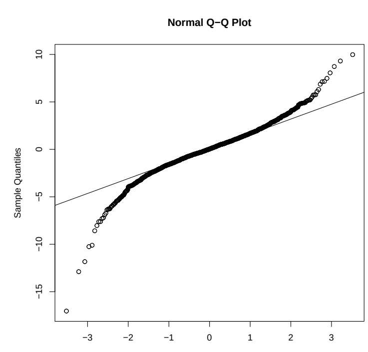
```
Call:
ar(x = value)

Coefficients:
     1
0.9981

Order selected 1  sigma^2 estimated as  5.151
```



(a) AR coefficients



(b) AIC per number of parameters



(c) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.098621, df = 1, p-value = 0.7535
```

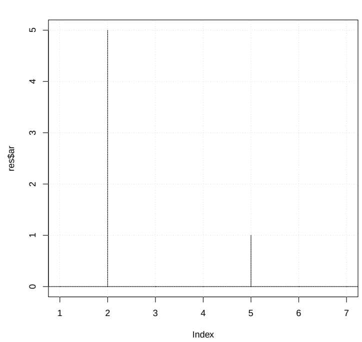
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.9508, p-value < 2.2e-16
```

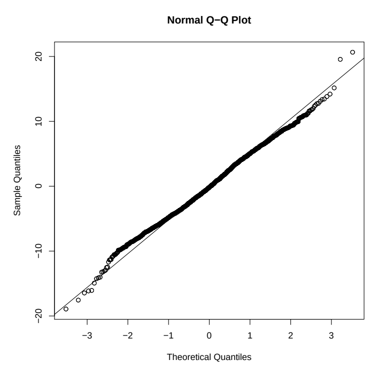
Is it an MA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,5) with non-zero mean

Coefficients:
         ma1     ma2     ma3     ma4    ma5  intercept
      2.0207  2.6273  2.4116  1.5390  0.568   150.1698
s.e.  0.0216  0.0375  0.0352  0.0242  0.015     1.0180

sigma^2 estimated as 23.34:  log likelihood=-6958.04
AIC=13930.08   AICc=13930.13   BIC=13970.33
```



(a) MA coefficients



(b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 213.87, df = 1, p-value < 2.2e-16
```

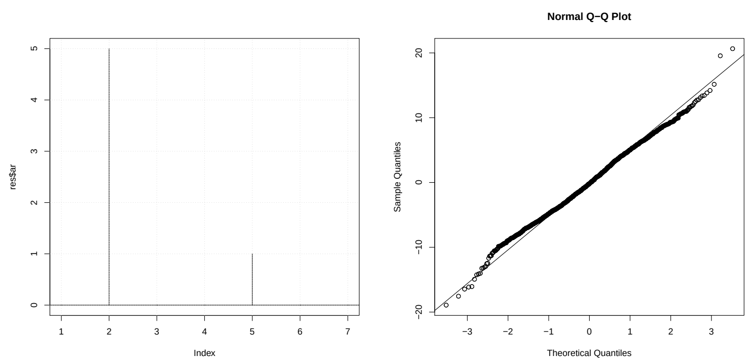
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.99703, p-value = 0.0001803
```

Is it an ARMA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,5) with non-zero mean

Coefficients:
         ma1     ma2     ma3     ma4    ma5  intercept
      2.0207  2.6273  2.4116  1.5390  0.568   150.1698
s.e.  0.0216  0.0375  0.0352  0.0242  0.015     1.0180

sigma^2 estimated as 23.34:  log likelihood=-6958.04
AIC=13930.08   AICc=13930.13   BIC=13970.33
```



(a) MA coefficients (b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 213.87, df = 1, p-value < 2.2e-16
```

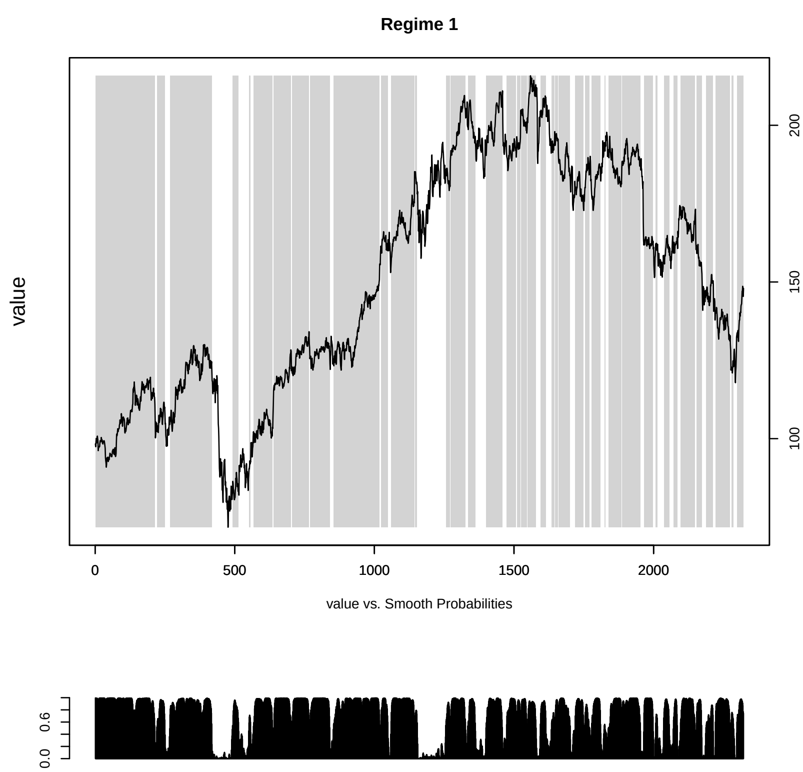
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.99703, p-value = 0.0001803
```

Markov Switching Model Try with 2 regimes.

Results:

```
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(T, T, T), p = 1, control = list(parallel = F))

       AIC      BIC    logLik
  9388.168 9442.169 -4690.084

Coefficients:
        (Intercept)(S) value_1(S)   Std(S)
Model 1      0.2648856  0.9988069 1.368896
Model 2      0.3165829  0.9969839 3.033226

Transition probabilities:
           Regime 1   Regime 2
Regime 1 0.96502377 0.09019822
Regime 2 0.03497623 0.90980178
```

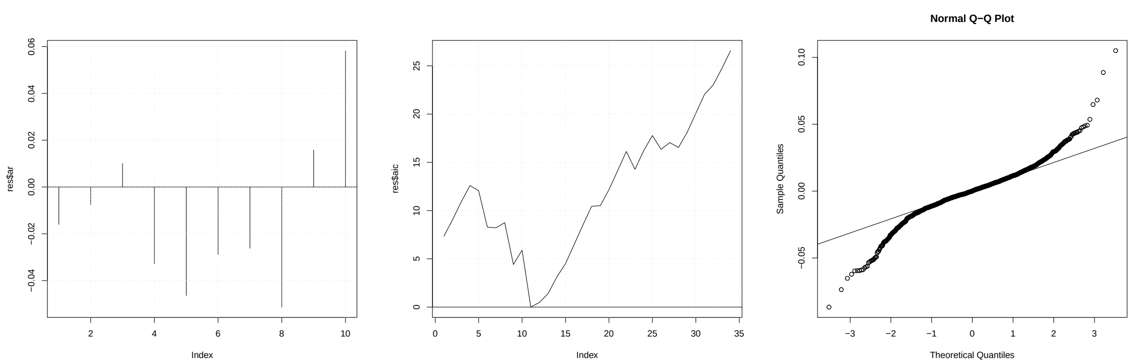


(a) Which 2

Looking at logret

Is it an AR model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm. Results:

```
Call:
ar(x = value)

Coefficients:
      1        2        3        4        5        6        7        8
-0.0160  -0.0076   0.0100  -0.0327  -0.0464  -0.0288  -0.0262  -0.0513
      9       10
 0.0159   0.0583

Order selected 10  sigma^2 estimated as  0.0002069
```



(a) AR coefficients (b) AIC per number of parameters (c) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.0038622, df = 1, p-value = 0.9504
```

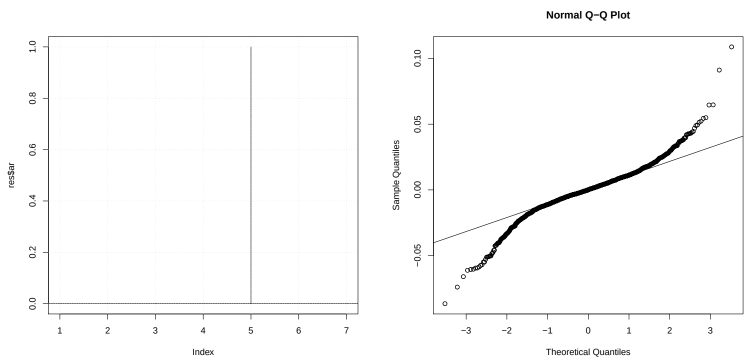
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.94224, p-value < 2.2e-16
```

Is it an MA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

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

Coefficients:
      intercept
          2e-04
s.e.      3e-04

sigma^2 estimated as 0.0002084:  log likelihood=6545.89
AIC=-13087.77   AICc=-13087.77   BIC=-13076.27
```



(a) MA coefficients (b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.28676, df = 1, p-value = 0.5923
```

Test residual normal law with Shapiro:

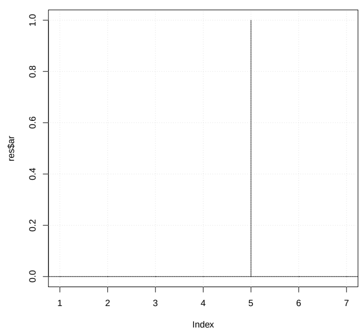
```
Shapiro-Wilk normality test

data:  res
W = 0.94027, p-value < 2.2e-16
```

Is it an ARMA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

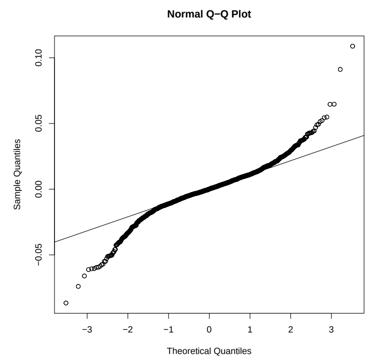
Results:

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

Coefficients:
      intercept
          2e-04
```

```
s.e.      3e-04

sigma^2 estimated as 0.0002084:  log likelihood=6545.89
AIC=-13087.77   AICc=-13087.77   BIC=-13076.27
```



(a) MA coefficients



(b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.28676, df = 1, p-value = 0.5923
```

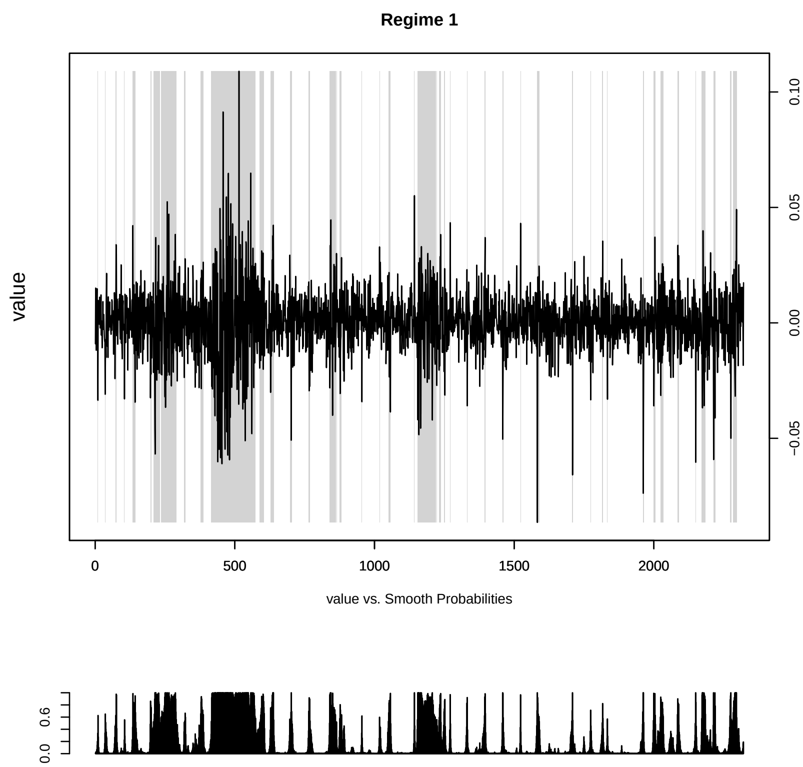
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.94027, p-value < 2.2e-16
```

Markov Switching Model Try with 2 regimes.

Results:

```
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(T, T, T), p = 1, control = list(parallel = F))

        AIC       BIC   logLik
  -13687.27 -13633.27 6847.635

Coefficients:
        (Intercept)(S)  value_1(S)      Std(S)
Model 1  -0.0007904372 -0.00119046 0.023807260
Model 2   0.0005225907 -0.03309368 0.009267872

Transition probabilities:
           Regime 1   Regime 2
Regime 1 0.90554306 0.03205306
Regime 2 0.09445694 0.96794694
```

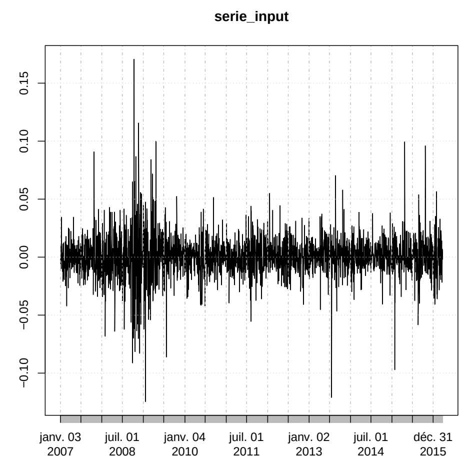

Regime 1
value
0.10
0.05
0.00
−0.05
0
500
1000
1500
2000
value vs. Smooth Probabilities
0.0
0.6

(a) Which 2

4 MSFT

This is the historical close quotation. S_t. A stationarize are the log return. $r_t = \log(\frac{S_t}{S_{t-1}})$

(a) Close level of MSFT

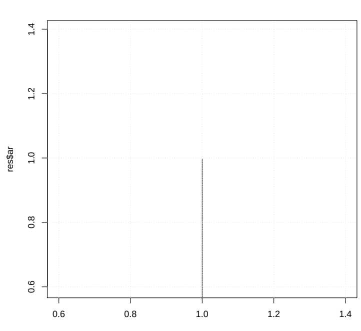


(b) log return MSFT

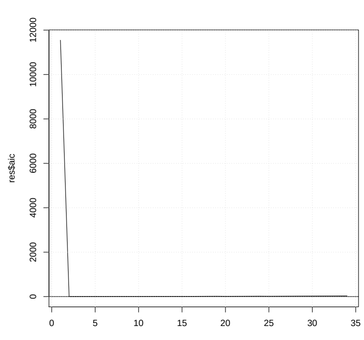
Looking at original

Is it an AR model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm. Results:

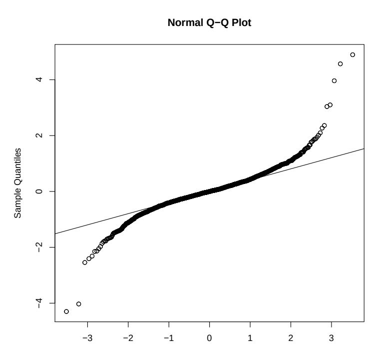
```
Call:
ar(x = value)

Coefficients:
     1
0.9965

Order selected 1  sigma^2 estimated as  0.5186
```



(a) AR coefficients



(b) AIC per number of parameters



(c) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.23417, df = 1, p-value = 0.6284
```

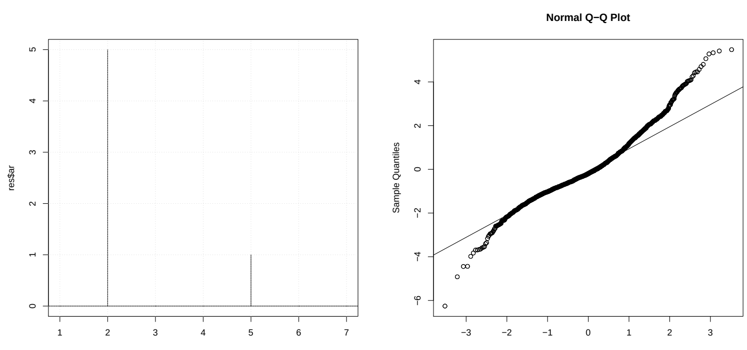
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.91461, p-value < 2.2e-16
```

Is it an MA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,5) with non-zero mean

Coefficients:
         ma1     ma2     ma3     ma4     ma5  intercept
      1.8856  2.3315  2.1797  1.4642  0.5639    32.1564
s.e.  0.0188  0.0280  0.0285  0.0251  0.0143     0.2431

sigma^2 estimated as 1.549:  log likelihood=-3807.32
AIC=7628.63   AICc=7628.68   BIC=7668.89
```



(a) MA coefficients (b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 210.06, df = 1, p-value < 2.2e-16
```

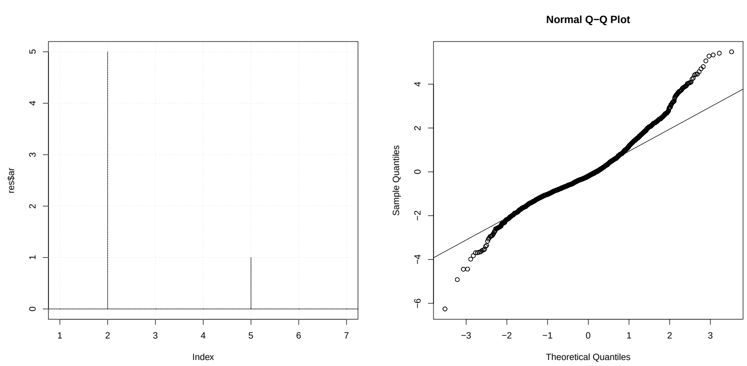
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.96066, p-value < 2.2e-16
```

Is it an ARMA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,5) with non-zero mean

Coefficients:
         ma1     ma2     ma3     ma4     ma5  intercept
      1.8856  2.3315  2.1797  1.4642  0.5639    32.1564
s.e.  0.0188  0.0280  0.0285  0.0251  0.0143     0.2431

sigma^2 estimated as 1.549:  log likelihood=-3807.32
AIC=7628.63   AICc=7628.68   BIC=7668.89
```



(a) MA coefficients (b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 210.06, df = 1, p-value < 2.2e-16
```

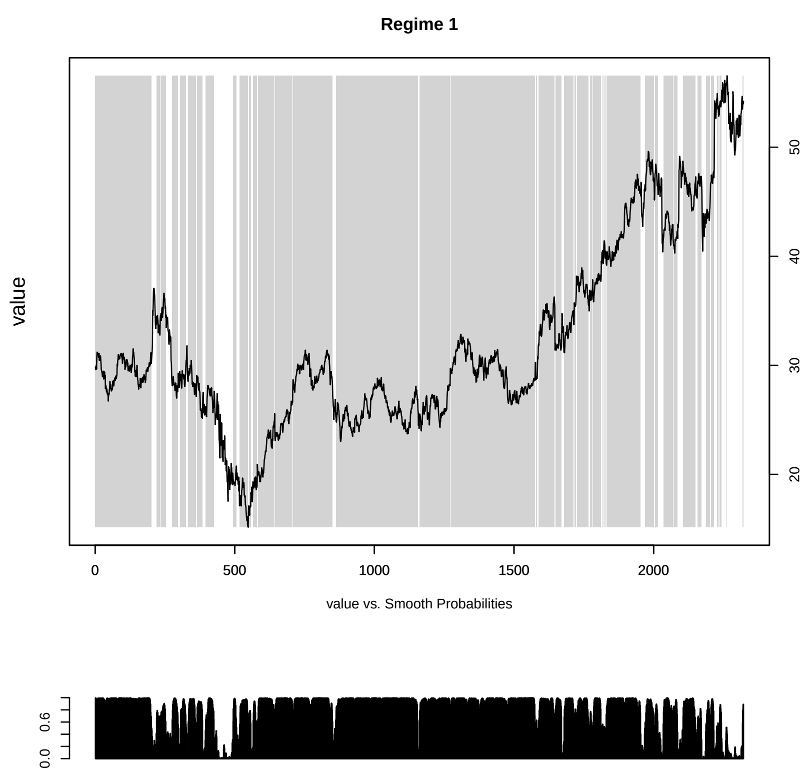
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.96066, p-value < 2.2e-16
```

Markov Switching Model Try with 2 regimes.

Results:

```
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(T, T, T), p = 1, control = list(parallel = F))

       AIC      BIC    logLik
  3252.634 3306.636 -1622.317

Coefficients:
        (Intercept)(S) value_1(S)    Std(S)
Model 1     0.03245657  0.9993361 0.3726641
Model 2     0.02607107  0.9994247 0.9869924

Transition probabilities:
           Regime 1  Regime 2
Regime 1 0.96872469 0.1242708
Regime 2 0.03127531 0.8757292
```

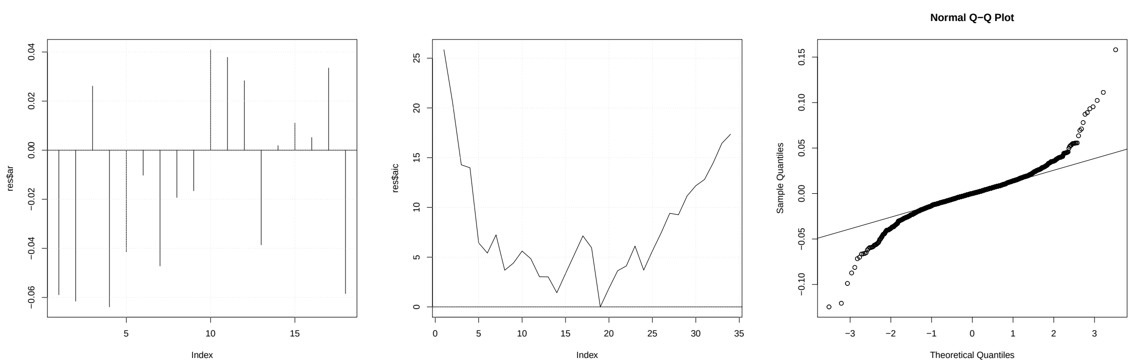


(a) Which 2

Looking at logret

Is it an AR model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm. Results:

```
Call:
ar(x = value)

Coefficients:
      1        2        3        4        5        6        7        8
-0.0589  -0.0615   0.0261  -0.0638  -0.0414  -0.0102  -0.0472  -0.0193
      9       10       11       12       13       14       15       16
-0.0166   0.0409   0.0378   0.0283  -0.0386   0.0019   0.0110   0.0052
     17       18
 0.0335  -0.0584

Order selected 18  sigma^2 estimated as  0.000324
```



(a) AR coefficients (b) AIC per number of parameters (c) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.00018124, df = 1, p-value = 0.9893
```

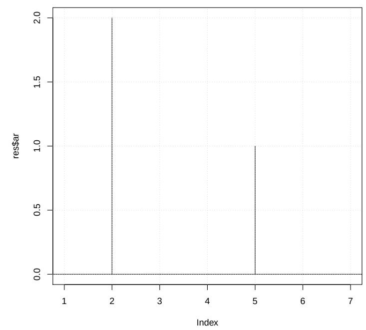
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.91986, p-value < 2.2e-16
```

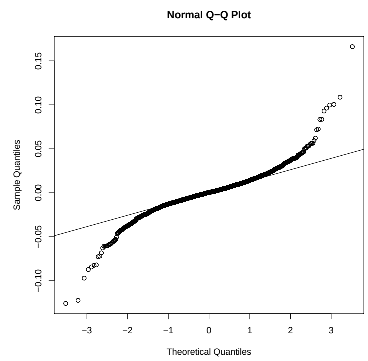
Is it an MA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,2) with zero mean

Coefficients:
          ma1      ma2
      -0.0560  -0.0613
s.e.   0.0209   0.0224

sigma^2 estimated as 0.0003278:  log likelihood=6019.9
AIC=-12033.8   AICc=-12033.79   BIC=-12016.55
```



(a) MA coefficients



(b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.0083701, df = 1, p-value = 0.9271
```

Test residual normal law with Shapiro:

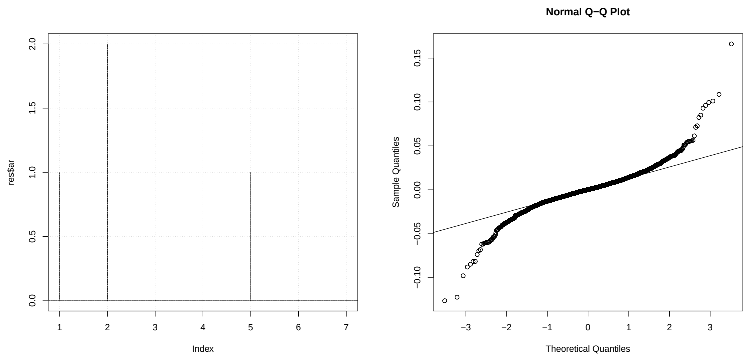
```
Shapiro-Wilk normality test

data:  res
W = 0.91432, p-value < 2.2e-16
```

Is it an ARMA model ? We look for the AR coefficient, and minimize it with the AIC indicator. The AIC is a ratio between a good calibration and the number of parameters. The AR model implies that residual should be white noise (iid normal law). It's tested with the QQnorm.

Results:

```
Series: value
ARIMA(0,0,2) with zero mean

Coefficients:
          ma1      ma2
      -0.0560  -0.0613
```

```
s.e.   0.0209   0.0224

sigma^2 estimated as 0.0003278:  log likelihood=6019.9
AIC=-12033.8   AICc=-12033.79   BIC=-12016.55
```



(a) MA coefficients (b) QQnorm of residual

Test residual independences with ljung:

```
Box-Ljung test

data:  res
X-squared = 0.0017872, df = 1, p-value = 0.9663
```

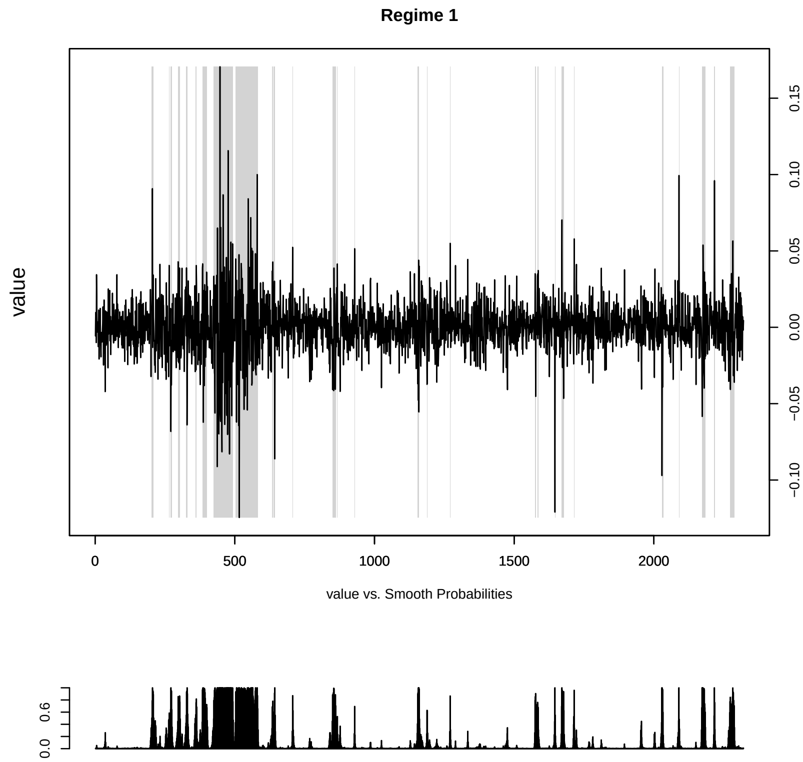
Test residual normal law with Shapiro:

```
Shapiro-Wilk normality test

data:  res
W = 0.91427, p-value < 2.2e-16
```

Markov Switching Model Try with 2 regimes.

Results:

```
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(T, T, T), p = 1, control = list(parallel = F))

        AIC       BIC   logLik
  -12733.08 -12679.08 6370.541

Coefficients:
        (Intercept)(S)  value_1(S)     Std(S)
Model 1   -0.001015905 -0.09460918 0.03640935
Model 2    0.000462657 -0.01414423 0.01256221

Transition probabilities:
          Regime 1   Regime 2
Regime 1 0.8903983 0.01873014
Regime 2 0.1096017 0.98126986
```

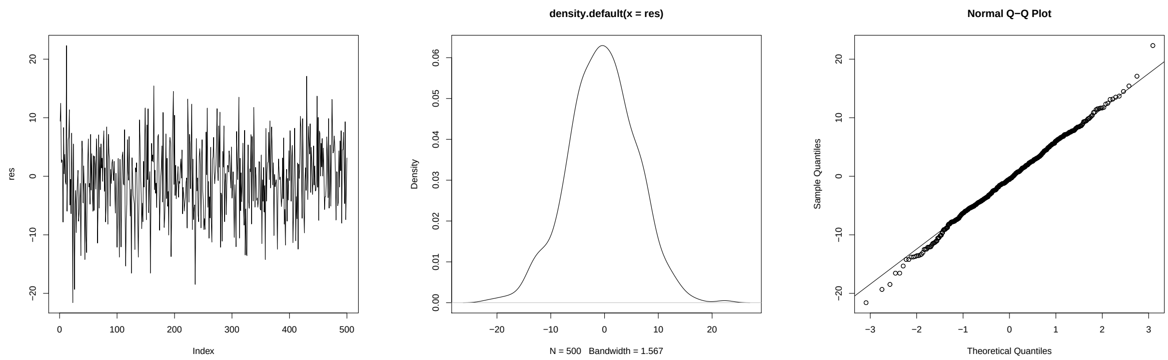

Regime 1
value
0.15
0.10
0.05
0.00
−0.05
−0.10
0
500
1000
1500
2000
value vs. Smooth Probabilities
0.6
0.0

(a) Which 2

A To determine white noise

What is it ? Be ϵ_t a time serie. It's a white noise, if it's independ through time t and got the same law. We will consider a normal law $\mathcal{N}(0, \sigma)$ for white noise.

Example Below example of noise

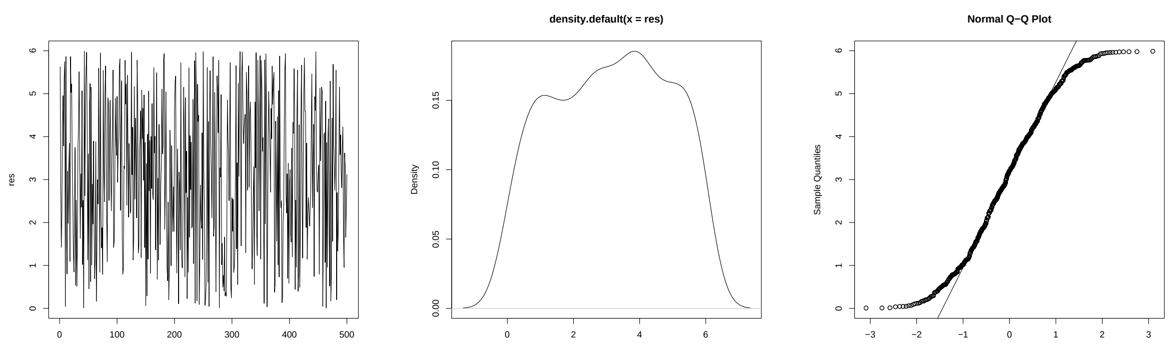


(a) Normal process (b) Density function (c) QQnorm

Figure 37: Example of $\mathcal{N}$(0,6)

Test of Box-Ljung for independance : Test of Shapiro for Normal law :

```
Shapiro-Wilk normality test

data:  res
W = 0.99763, p-value = 0.7064
```



(a) Normal process (b) Density function (c) QQnorm

Figure 38: Example of $\mathcal{N}$(0,6)

Test of Box-Ljung for independance : Test of Shapiro for Normal law :

```
Shapiro-Wilk normality test

data:  res
```

```
W = 0.95801, p-value = 9.744e-11
```

B Notions

Strongly Stationary Be X_t a time series. X_t is strictly stationary process or strongly stationary process if :

$$\forall h, \forall n, X_{t_1}...X_{t_n} \text{ has the same law than } X_{t_1+h}...X_{t_n+h}$$

Weakly stationary X_t is a weak or wide-sense stationary process if :

(i) $\forall t, \mathbb{E}\left\{X_t^2\right\} < \infty$, finite variance

(ii) $\forall t, \mathbb{E}\left\{X_t\right\} = m$, expectation doesn't depend of time t

(iii) $\forall t, \operatorname{cov}(X_t, X_{t+h}) = \gamma(h)$, covariance doesn't depend of time t and only on the lag h.

In practice strongly stationary process are hard to demonstrate. In models below we'll demonstrate weak stationarity.

Covariance Be X, Y two random variables. We note $\operatorname{cov}(X, Y)$ the covariance between X and Y

$$\operatorname{cov}(X, Y) = \mathbb{E}\left\{(X - \mathbb{E}\left\{X\right\})(Y - \mathbb{E}\left\{Y\right\})\right\}$$

Correlation Be X, Y two random variables. We note $\rho_(X, Y)$ the correlation between X and Y

$$\rho_{X,Y} = \frac{\operatorname{cov}(X, Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}$$

Autocorrelation Be X_t a time serie. $X_1, X_2, ...$ are random variables. We note γ_l the covariance between X_t and X_{t-l}

$$\gamma_l = \frac{\operatorname{cov}(X_t, X_{t-l})}{\sqrt{\operatorname{Var}(X_t)\operatorname{Var}(X_{t-l})}}$$

Note that $\gamma_0 = 1$

Partial autocorrelation (PACF) We note $\pi(k)$ the partial autocorrelation :

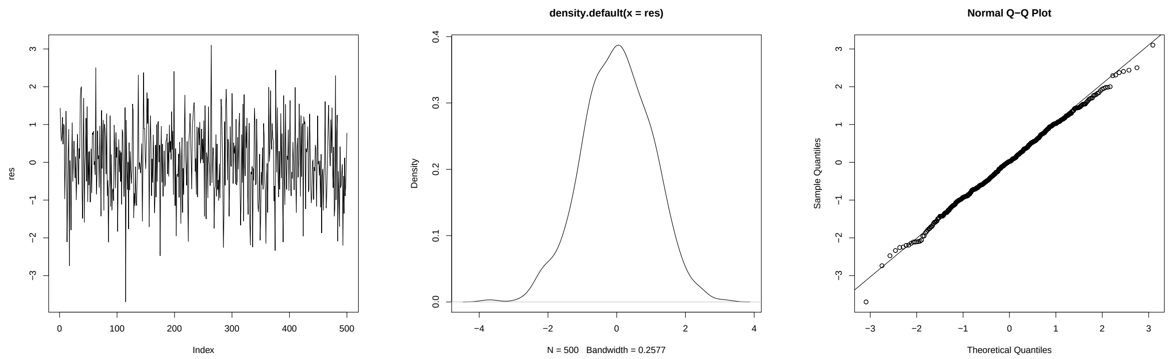
$$\pi(k) = \operatorname{corr}(X_t - \mathbb{E}\left\{X_t | X_{t-1}...X_{t-k+1}\right\}, X_{t-k} - \mathbb{E}\left\{X_{t-k} | X_{t-1}...X_{t-k+1}\right\}) =$$

Akaike Information Criterion (AIC)

$$AIC = \frac{-2}{T}\ln(\text{likelihood}) + \frac{2}{T}(\text{number of parameters})$$

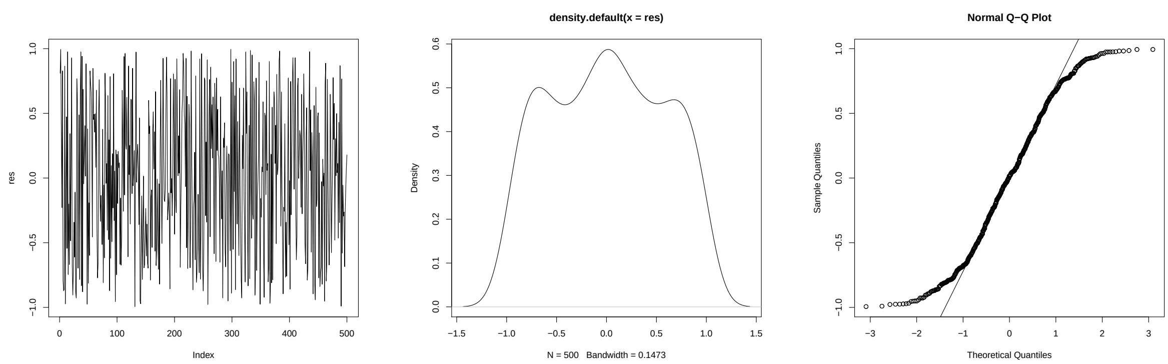
Normal law $\mathcal{N}(\mu, \sigma)$ X follow a normal law $\mathcal{N}(\mu, \sigma)$. The density function is :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



(a) Normal process (b) Density function (c) QQnorm

Figure 39: Example of $\mathcal{N}$(0,1)



(a) Normal process (b) Density function (c) QQnorm

Figure 40: Example of $\mathcal{U}$(-1,1)

C AR Model

AR(1)

Definition Be $c, \beta \in \mathbb{R}^2$, X_t is an AR(1) process if

$$\begin{aligned} X_t &= c + \beta X_{t-1} + \epsilon_t \\ &= c + \beta^t X_0 + \sum_{i=0}^{t} \beta^i \epsilon_{t-i} \end{aligned}$$

Stationary condition

(i) X_t has a finite variance if $|\beta| < 1$:

$$\begin{aligned} \mathrm{Var}(X_t) &= \mathrm{Var}(\sum_{i=0}^{t} \beta^i \epsilon_{t-i}) \\ &= \sum_{i=0}^{t} \beta^{2i} \sigma^2 \\ &= \frac{1-\beta^{2t}}{1-\beta} \sigma^2 \end{aligned}$$

When $|\beta| > 1$, $\mathrm{Var}(X_t) \xrightarrow[t\to\infty]{} \infty$.
When $|\beta| < 1$, $\mathrm{Var}(X_t) \xrightarrow[t\to\infty]{} \frac{1}{1-\beta}\sigma^2$

(ii) Expectation doesn't depend of t implies $\mathbb{E}\left\{X_t\right\} = \mathbb{E}\left\{X_{t-1}\right\} = \mu$

$$\begin{aligned} \mathbb{E}\left\{X_t\right\} &= c + \beta \mathbb{E}\left\{X_{t-1}\right\} \\ \Leftrightarrow \mu &= \frac{c}{1-\beta} \end{aligned}$$

But we could rewrite $\mathbb{E}\left\{X_t\right\}$ as :

$$\begin{aligned} \mathbb{E}\left\{X_t\right\} &= c + \mathbb{E}\left\{\beta X_{t-1} + \epsilon_t\right\} \\ &= c + \mathbb{E}\left\{\sum_{i=1}^{t} \beta^i (c + \epsilon_{t-i})\right\} \\ &= c\frac{1-\beta^t}{1-\beta} \end{aligned}$$

(iii) Covariance doesn't depend of t ?

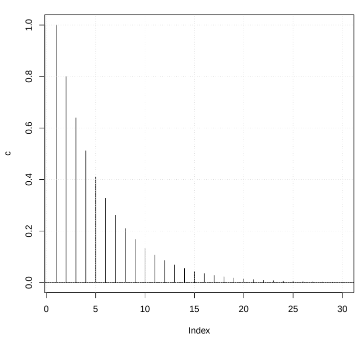
$$\begin{aligned} \mathrm{cov}(X_t, X_{t-1}) &= \mathbb{E}\left\{(c + \beta X_{t-1} + \epsilon_t) X_{t-1}\right\} - c^2 \\ &= c\mathbb{E}\left\{X_{t-1}\right\} + \beta \mathbb{E}\left\{X_{t-1}^2\right\} - c^2 \\ &= c\frac{1-\beta^{t-1}}{1-\beta} + \beta \frac{1-\beta^{2(t-1)}}{1-\beta} \sigma^2 - c^2 \end{aligned}$$

$$\begin{aligned} \gamma(1) &= \mathbb{E}\left\{(X_t - c)(X_{t-1} - c)\right\} \\ &= \mathbb{E}\left\{(\beta X_{t-1} + \epsilon_t)(X_{t-1} - c)\right\} \\ &= \mathbb{E}\left\{\beta X_{t-1}\right\} \end{aligned}$$

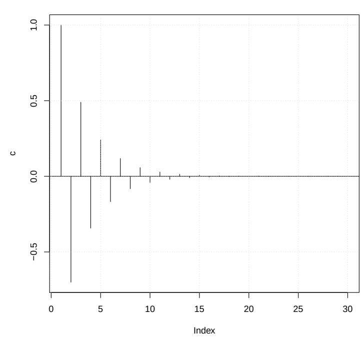
ACF of AR(1) process

Demonstration There is a linear relation between t and $t-1$.

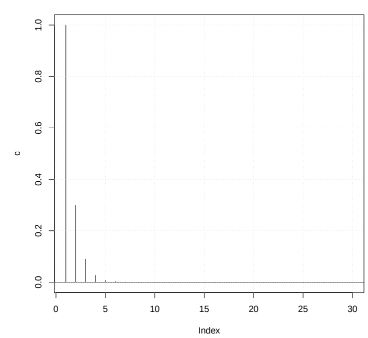
$$
\begin{aligned}
\gamma_l &= \operatorname{cov}(X_t, X_{t-l}) \\
&= \mathbb{E}\left\{(X_t-\mu)(X_{t-l}-\mu)\right\} \\
&= \mathbb{E}\left\{(\epsilon_t + \sum_{i=1}^{\infty}\beta^i\epsilon_{t-i})(\sum_{i=1}^{\infty}\beta^i\epsilon_{t-l-i})\right\} \\
&= \mathbb{E}\left\{\sum_{i,j=1}^{\infty}\beta^i\epsilon_{t-i}\beta^j\epsilon_{t-l-j}\right\} \\
&= \mathbb{E}\left\{\sum_{i=1}^{\infty}\beta^{i+l}\epsilon_{t-l-i}\beta^i\epsilon_{t-l-i}\right\} \\
&= \sigma^2\sum_{i=1}^{\infty}\beta^{i+l}\beta^i \\
&= \sigma^2\frac{\beta^l}{1-\beta^2} \\
&= \beta\gamma_{l-1}
\end{aligned}
$$



(a) β =0.8



(b) β =-0.7

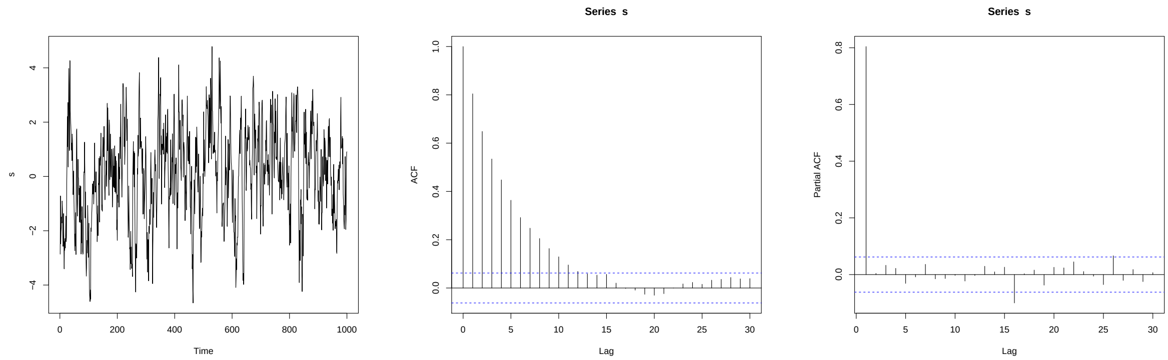


(c) β =0.3

Figure 41: Expected ACF for AR(1) process

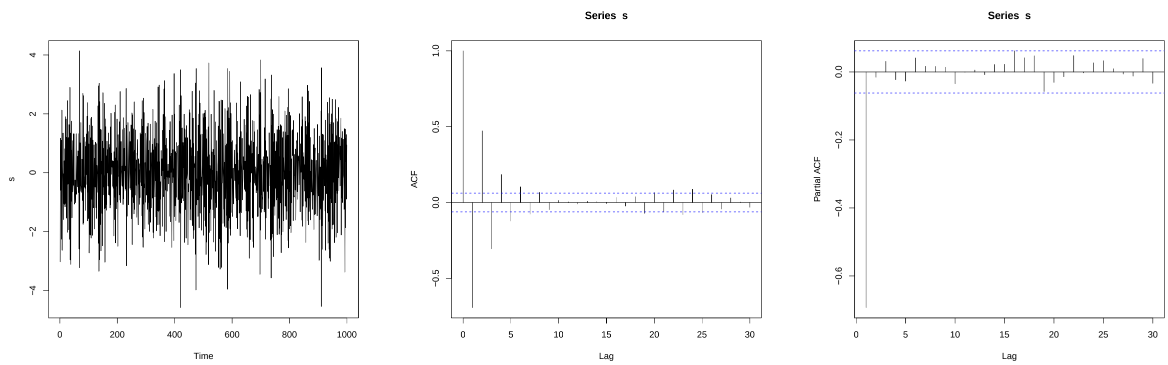
Examples

Examples See below.



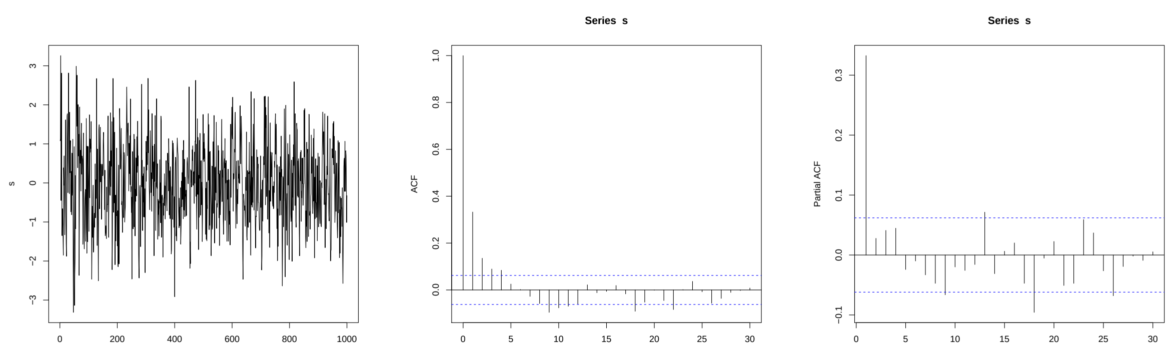
(a) Process (b) ACF (c) PACF

Figure 42: Example of AR(1) process, coefficient AR(0.8) (n=1000)



(a) Process (b) ACF (c) PACF

Figure 43: Example of AR(1) process, coefficient AR(-0.7) (n=1000)



(a) Process (b) ACF (c) PACF

Figure 44: Example of AR(1) process, coefficient AR(0.3) (n=1000)

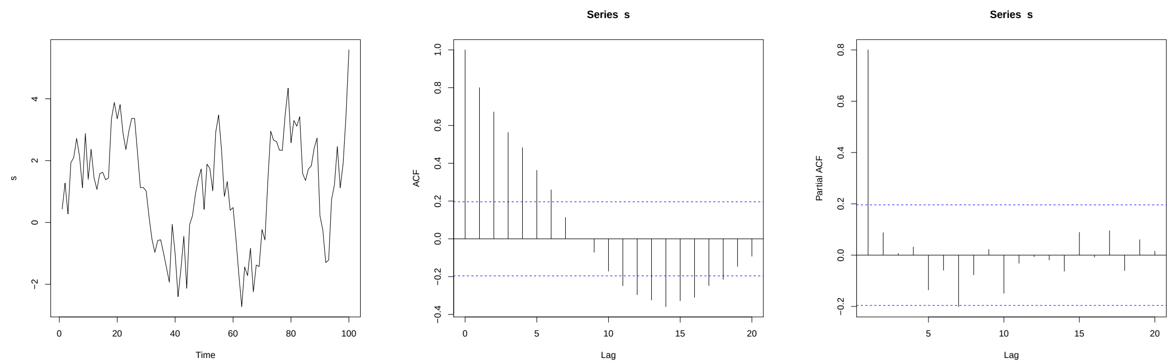
AR(p)

Definition Be $c, \beta \in \mathbb{R}^2$, X_t is an AR(p) process if

$$\begin{aligned} X_t &= c + \sum_{i=1}^{p} \beta_i X_{t-i} + \epsilon_t \\ &= c + \sum_{i=1}^{p} \beta_i \end{aligned}$$

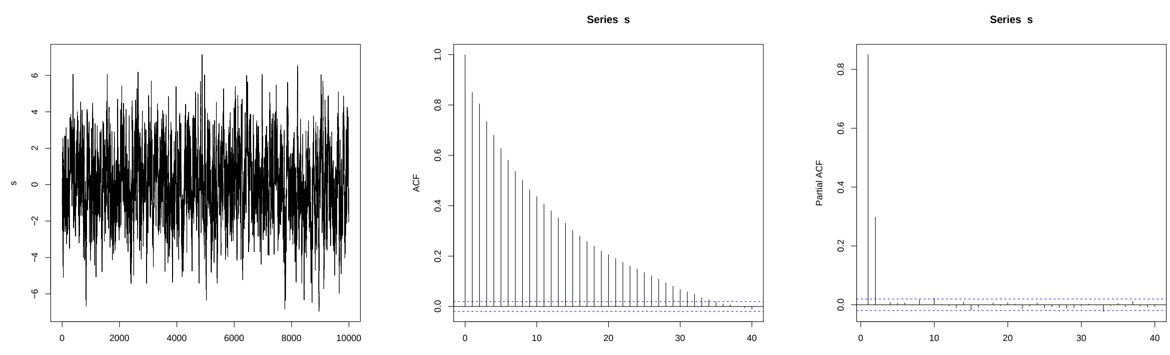
Stationary condition To be stationary, roots of the polynom $z^p - \sum_{i=1}^{p} \beta_i z^{p-i}$ must be within the unit circle, $|z_i| < 1$

Examples See below.



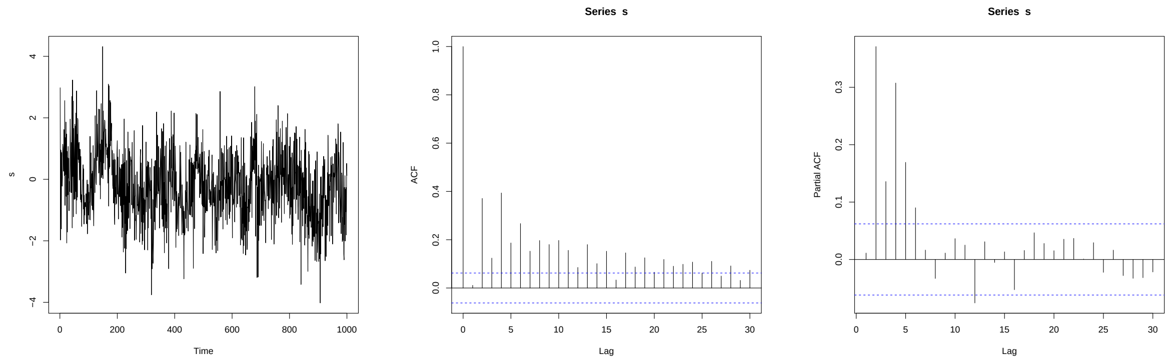
(a) Process (b) ACF (c) PACF

Figure 45: AR(2) process with coefficient AR(0.6,0.3) (n=100)



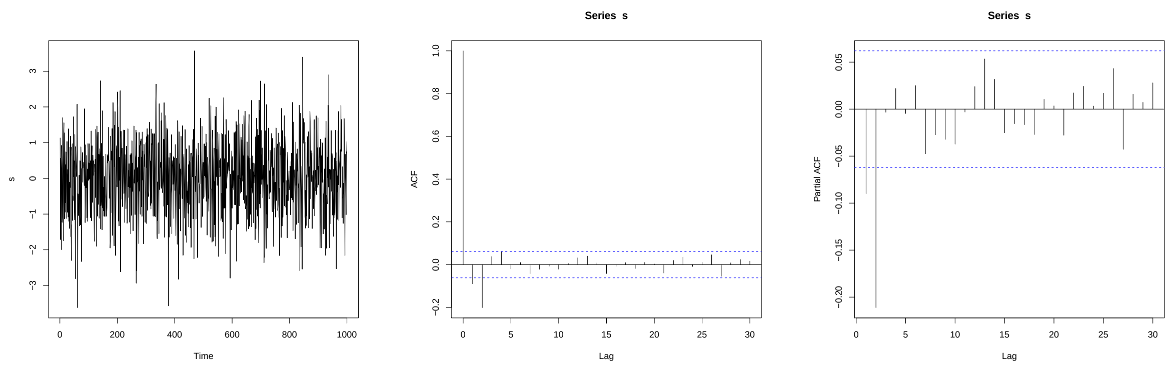
(a) Process (b) ACF (c) PACF

Figure 46: AR(2) process with coefficient AR(0.6,0.3) (n=10000)



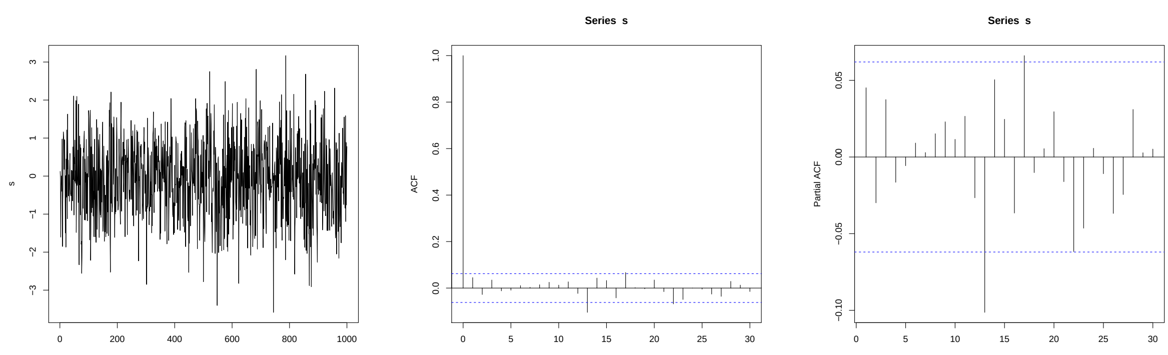
(a) Process (b) ACF (c) PACF

Figure 47: AR(6) process with coefficient AR(-0.2,0.2,0.1,0.3,0.2,0.1) (n=1000)



(a) Process (b) ACF (c) PACF

Figure 48: AR(2) process with coefficient AR(-0.1,-0.2) (n=1000)



(a) Process (b) ACF (c) PACF

Figure 49: AR(2) process with coefficient AR(0.01,0.02) (n=1000)

D MA Model

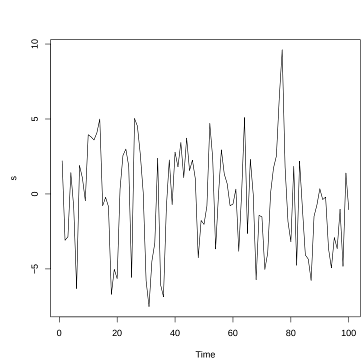
MA(1)

Definition Be $\theta \in \mathbb{R}$, X_t is an MA(1) process if

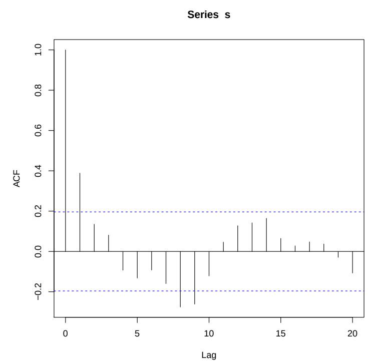
$$
\begin{aligned}
X_t &= c + \theta \epsilon_t \\
&= c + \beta^t X_0 + \sum_{i=0}^{t} \beta^i \epsilon_{t-i}
\end{aligned}
$$

Stationary condition X_t has a finite variance : $\mathrm{Var}(X_t) = \theta^2 \sigma^2$

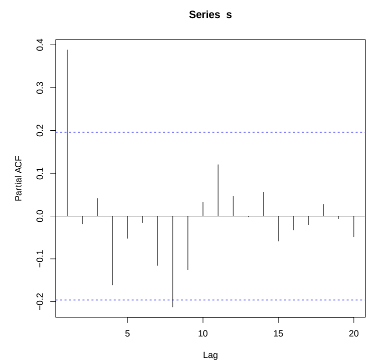
Examples See below.



(a) Process

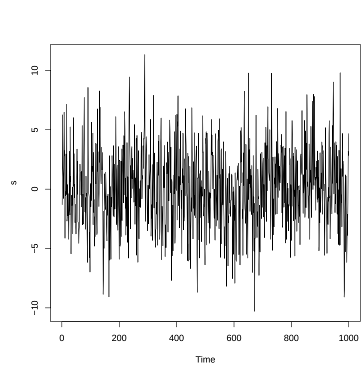


(b) ACF

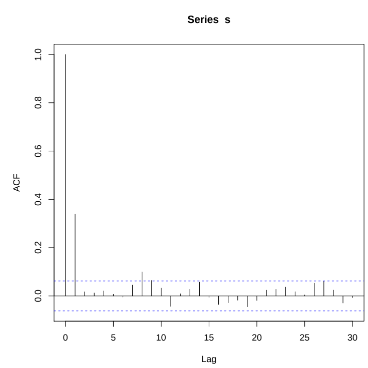


(c) PACF

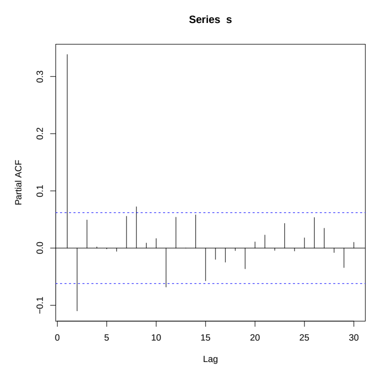
Figure 50: MA(1) process with coefficient MA(3) (n=100)



(a) Process

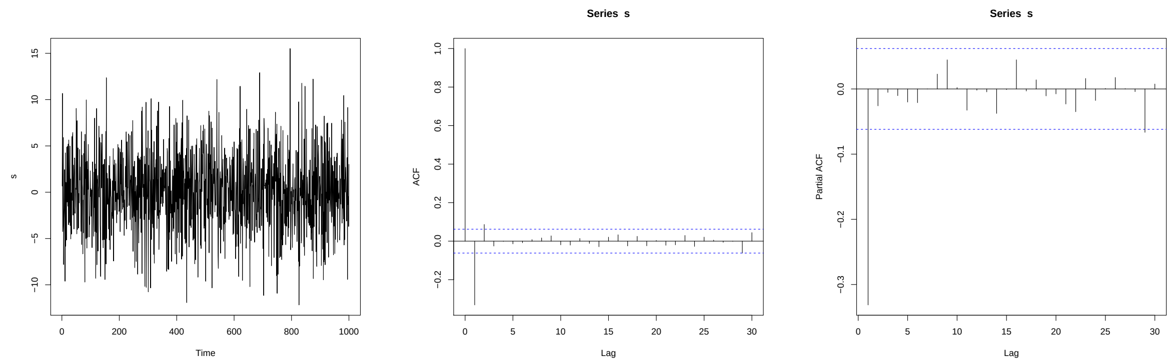


(b) ACF



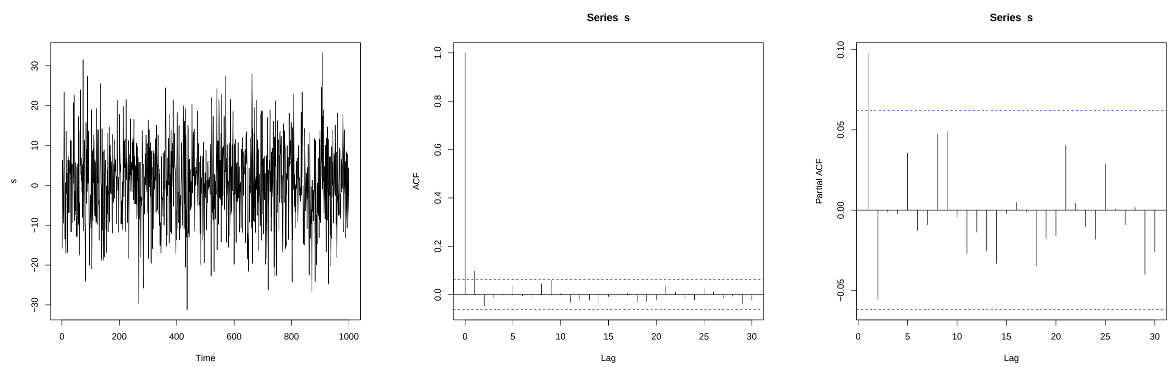
(c) PACF

Figure 51: MA(1) process with coefficient MA(3) (n=1000)



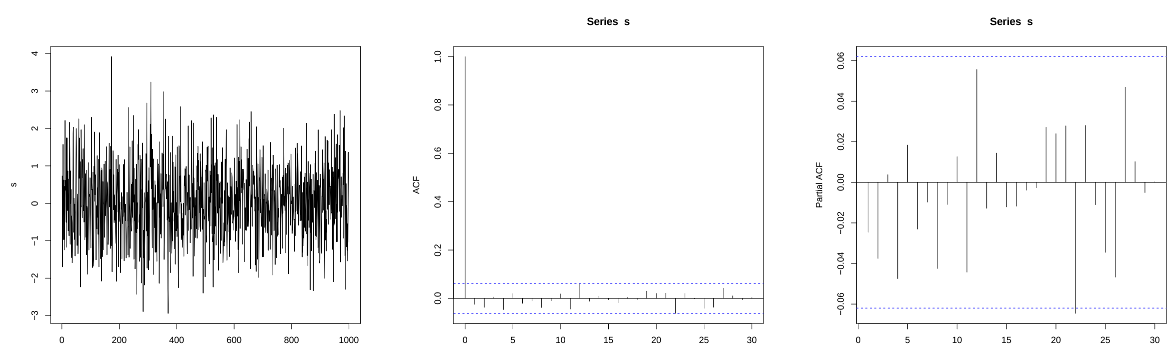
(a) Process (b) ACF (c) PACF

Figure 52: MA(1) process with coefficient MA(-4) (n=1000)



(a) Process (b) ACF (c) PACF

Figure 53: MA(1) process with coefficient MA(10) (n=1000)



(a) Process (b) ACF (c) PACF

Figure 54: MA(1) process with coefficient MA(0.01) (n=1000)

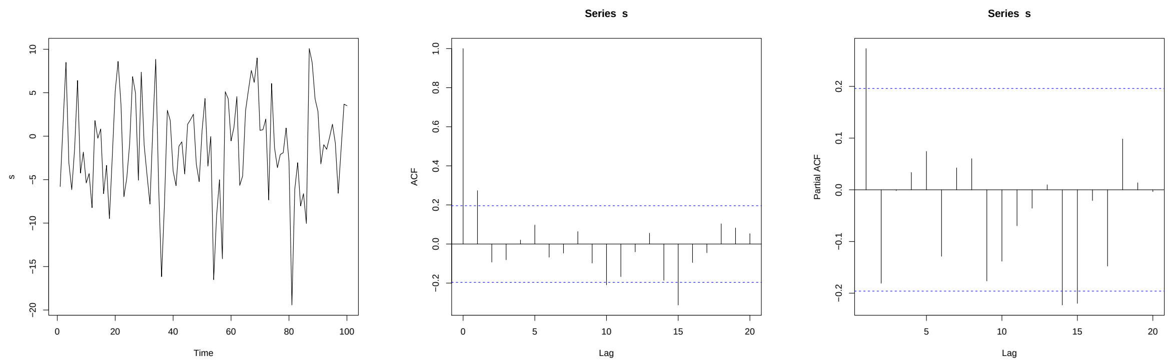
MA(q)

Definition Be $\theta \in \mathbb{R}$, X_t is an MA(q) process if

$$X_t = c + \sum_{i=0}^{q} \theta_i \epsilon_{t-i}$$

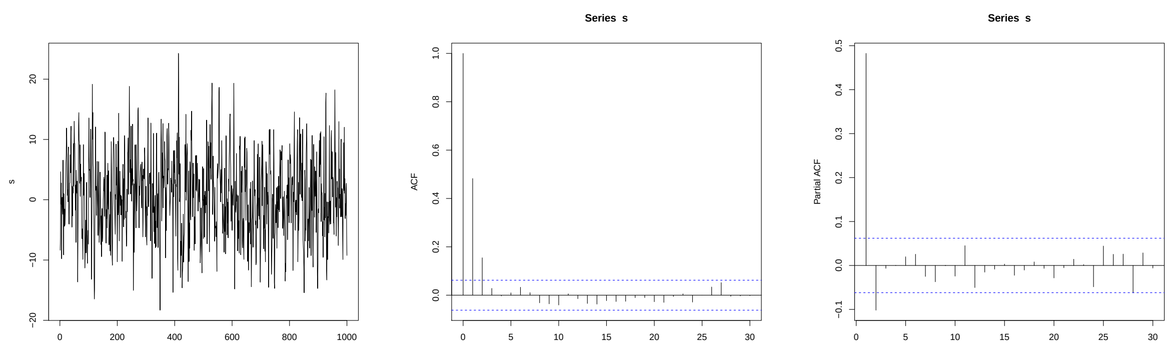
Stationary condition X_t has a finite variance : $\mathrm{Var}(X_t) = \sum_{i=0}^{q} \theta_i^2 \sigma^2$

Examples See below.



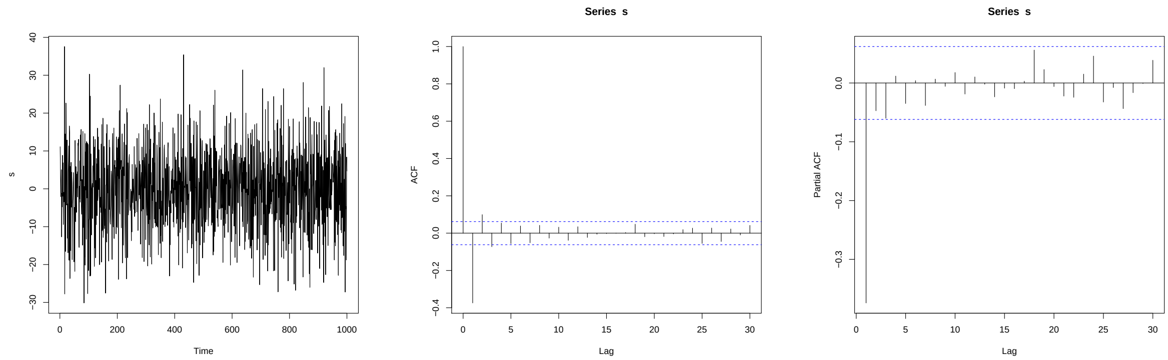
(a) Process (b) ACF (c) PACF

Figure 55: MA(2) process with coefficient MA(3,6) (n=100)



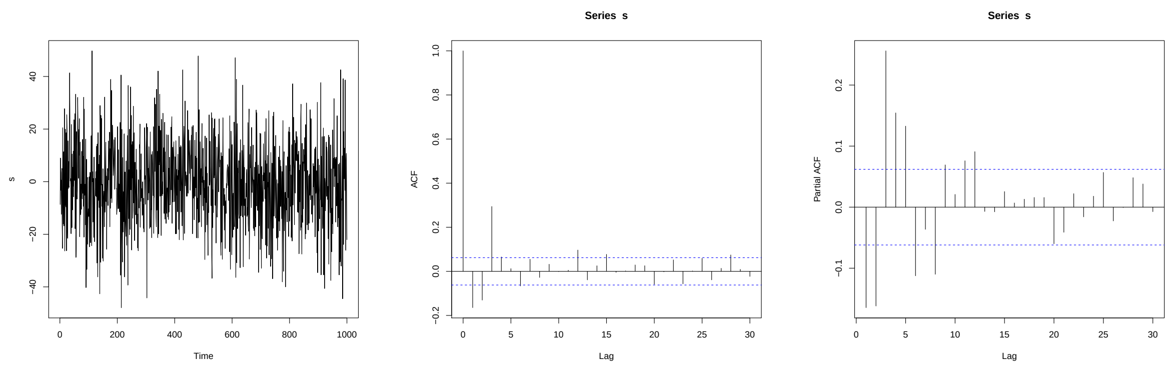
(a) Process (b) ACF (c) PACF

Figure 56: MA(2) process with coefficient MA(3,6) (n=1000)



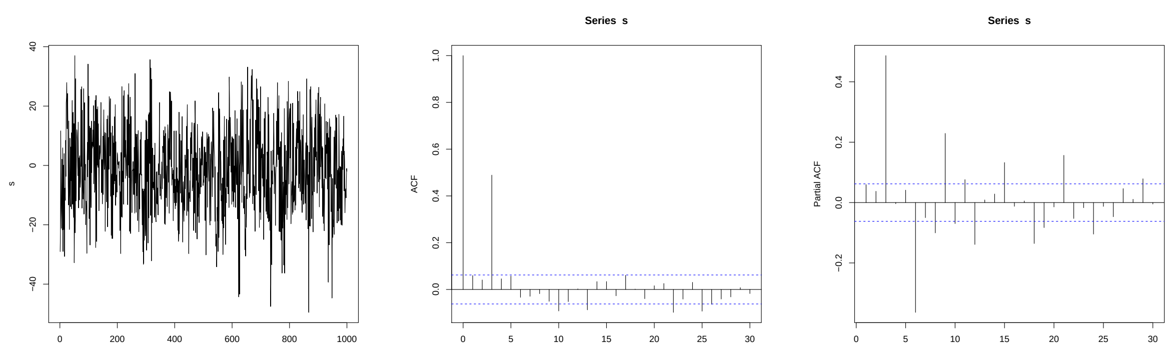
(a) Process (b) ACF (c) PACF

Figure 57: MA(2) process with coefficient MA(-4,10) (n=1000)



(a) Process (b) ACF (c) PACF

Figure 58: MA(4) process with coefficient MA(10,4,-7,9) (n=1000)



(a) Process (b) ACF (c) PACF

Figure 59: MA(4) process with coefficient MA(10,0,0,9) (n=1000)

E ARMA(p,q)

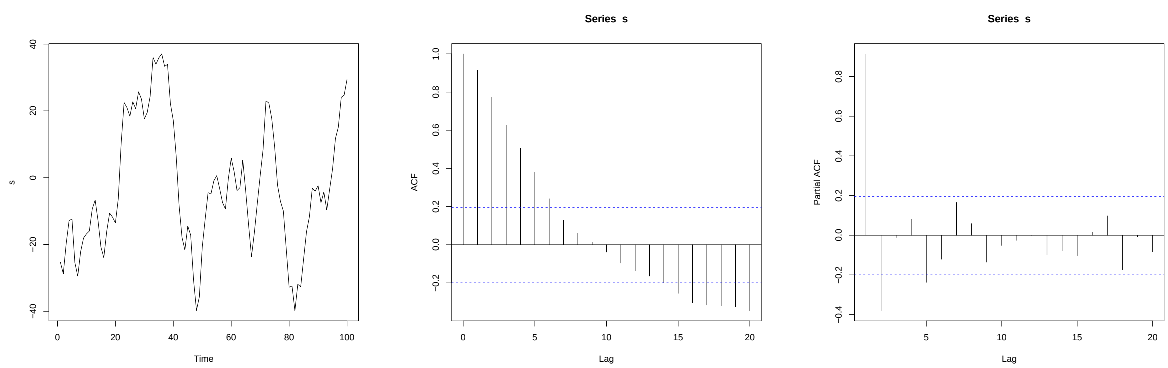
Definition Be $\theta \in \mathbb{R}$, X_t is an MA(q) process if

$$X_t = c + \sum_{i=0}^{q} \theta_i \epsilon_{t-i} + \sum_{i=1}^{p} \beta_i X_{t-i}$$

stationary To see if X_t is stationnary we look at its variance.

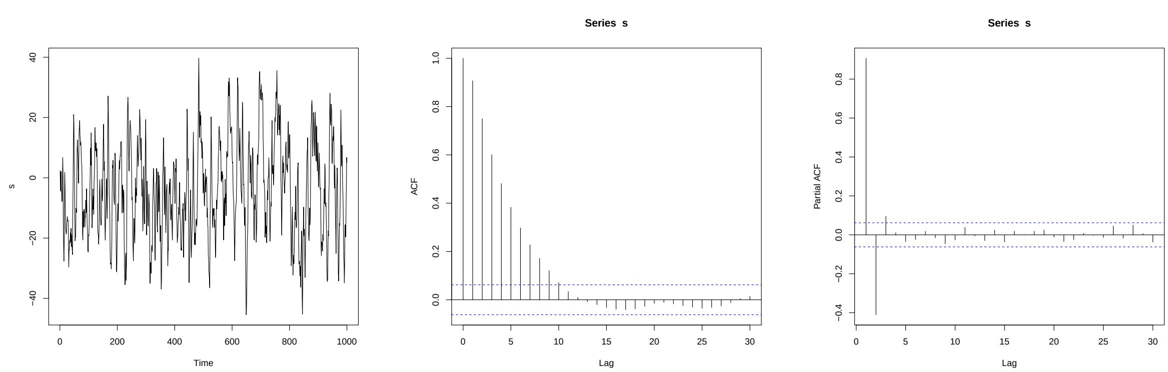
$$\mathrm{Var}(X_t) = \sum_{i=0}^{q} \theta_i^2 \sigma^2$$

Examples See below various example of ARMA(p,q)



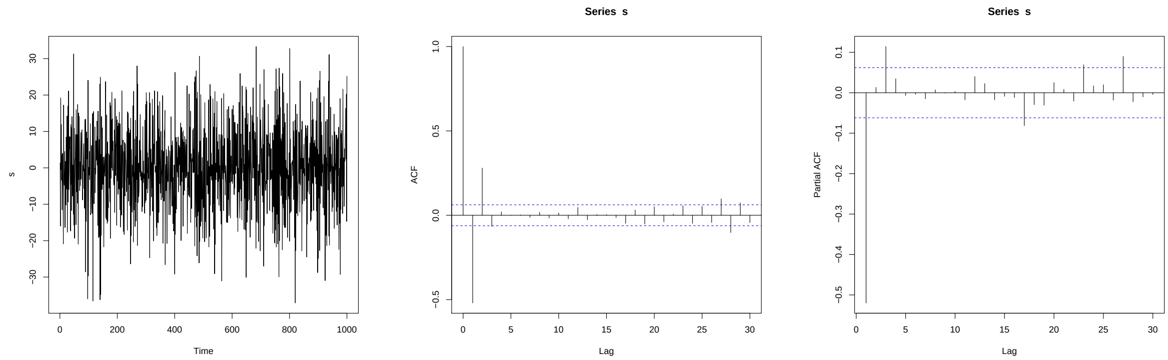
(a) Process (b) ACF (c) PACF

Figure 60: ARMA(1,2) process with coefficient AR(0.8), MA(3,6) (n=100)



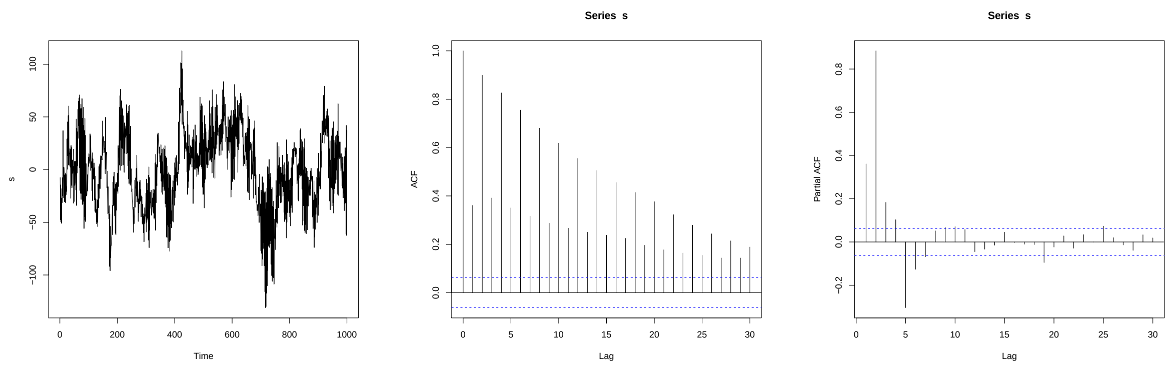
(a) Process (b) ACF (c) PACF

Figure 61: ARMA(1,2) process with coefficient AR(0.8), MA(3,6) (n=1000)



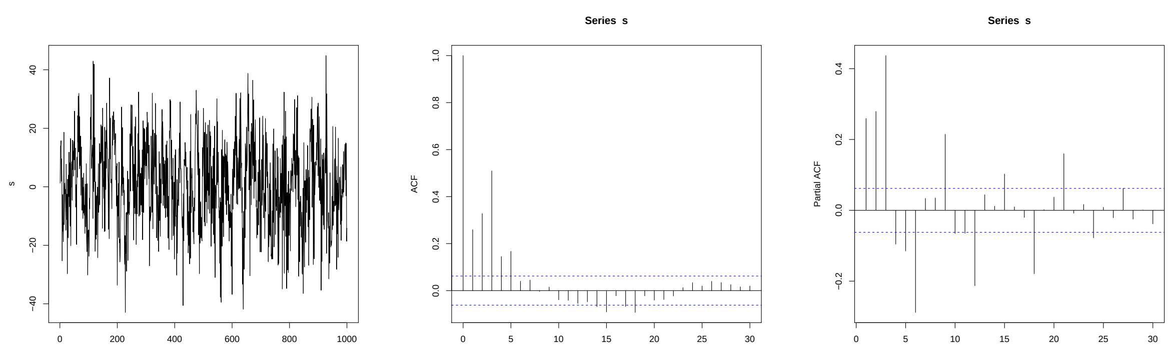
(a) Process (b) ACF (c) PACF

Figure 62: ARMA(3,2) process with coefficient AR(-0.1,0.2,0.1), MA(-4,10) (n=1000)



(a) Process (b) ACF (c) PACF

Figure 63: ARMA(2,4) process with coefficient AR(0,0.9), MA(10,4,-7,9) (n=1000)



(a) Process (b) ACF (c) PACF

Figure 64: ARMA(2,4) process with coefficient AR(0,0.3), MA(10,0,0,9) (n=1000)