Programming Assignment in R

TIME SERIES ANALYSIS (MAL7430)



Submitted To: Prof. Ganesh Manjhi

Submitted by: Vandita Agarwal (M20MA208)

Vinod Kumariya (M21MA010)

Dataset

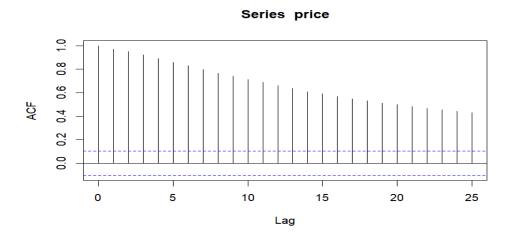
Firstly, Download the dataset from Investing.com and import in the dataset.We have shown the portion of our dataset which is "DOGECOIN in US Dollars". We have taken daily prices of length of one year(From 2022-03-25 to 2023-03-24). From which we have done model fitting on the closing price column, which is named as "Price" in the data below.

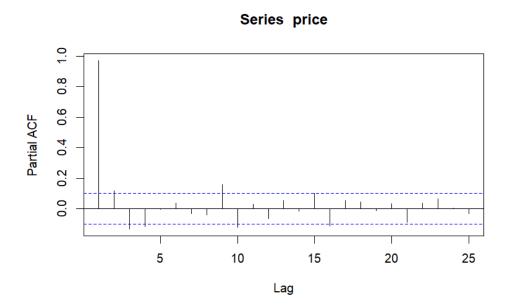
> View(dataset)

•	ïDate ‡	Price [‡]	Open [‡]	High [‡]	Low [‡]	Vol.	Change
1	Mar 24, 2023	0.074573	0.077387	0.077498	0.073103	1.15B	-3.64%
2	Mar 23, 2023	0.077387	0.073843	0.078919	0.072986	2.18B	4.79%
3	Mar 22, 2023	0.073851	0.076685	0.077821	0.071648	2.53B	-3.70%
4	Mar 21, 2023	0.076685	0.071342	0.077779	0.070404	1.97B	7.49%
5	Mar 20, 2023	0.071345	0.074541	0.075558	0.070528	1.37B	-4.19%
6	Mar 19, 2023	0.074462	0.072959	0.076498	0.072959	1.04B	2.06%
7	Mar 18, 2023	0.072959	0.076443	0.078775	0.072531	1.91B	-4.53%
8	Mar 17, 2023	0.076421	0.072588	0.077228	0.071528	1.66B	5.31%
9	Mar 16, 2023	0.072567	0.069681	0.074004	0.068919	1.25B	4.14%
10	Mar 15, 2023	0.069682	0.074640	0.075918	0.067488	1.91B	-6.67%

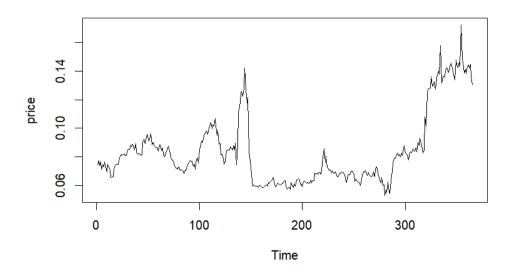
ARIMA and its variations

1) Firstly, We have plotted acf and pacf of Price





2) Then to check stationarity, we have plotted price vs time. Clearly, the plotted graph shows that the price is not stationary.



3) Done acf , pp test to check stationarity.> adf.test(price)

Augmented Dickey-Fuller Test

data: price

Dickey-Fuller = -2.2189, Lag order = 7, p-value = 0.4845

alternative hypothesis: stationary

Since the value of p is more than 5%, then the price is not stationary.

> PP.test(price)

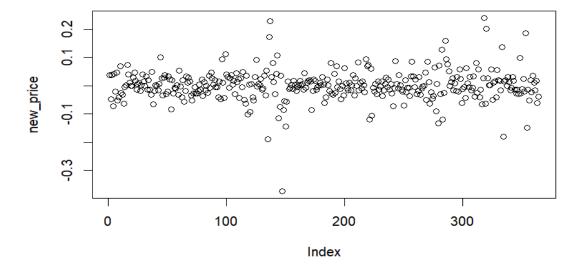
Phillips-Perron Unit Root Test

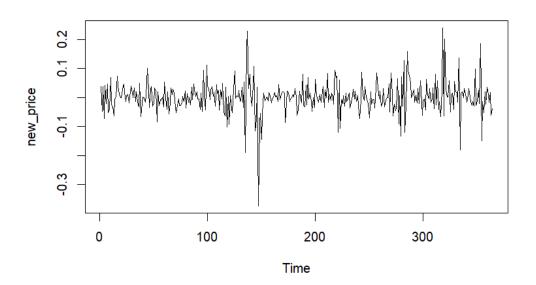
data: price

Dickey-Fuller = -1.9941, Truncation lag parameter = 5, p-value = 0.5794

Since the value of p is more than 5%, then the price is not stationary.

4) To make price stationary, we have changed it into new_price=diff(log(price))





Clearly the above graph has constant mean along the horizontal line, so the price has become stationary.

> adf.test(new_price)

Augmented Dickey-Fuller Test

data: new_price

Dickey-Fuller = -7.4326, Lag order = 7, p-value = 0.01

alternative hypothesis: stationary

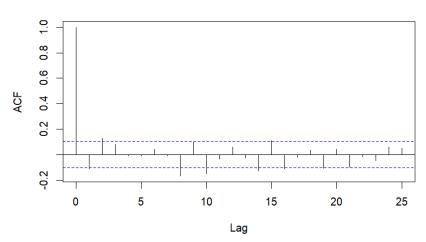
> PP.test(new_price)

Phillips-Perron Unit Root Test

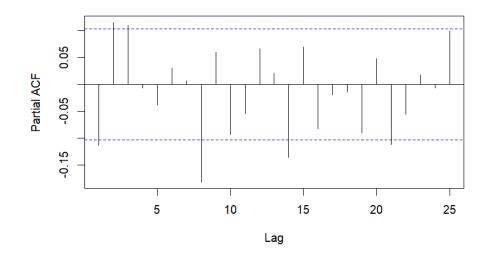
data: new_price
Dickey-Fuller = -21.149, Truncation lag parameter = 5, p-value = 0.01

Since, the value of p is less than 5% in both tests, so the new_price has become stationary. And we will use this price to fit the ARIMA model.

Series new_price



Series new_price



Also from the graphs of acf and pacf, the new_price is stationary.

5) Using auto.arima, I get the best model of ARIMA. According to this the best model is with order = c(1,0,3).

```
> auto.arima(new_price)
Series: new_price
ARIMA(1,0,3) with zero mean
Coefficients:
     ar1 ma1 ma2 ma3
   -0.8911 0.7868 0.0408 0.1887
s.e. 0.0615 0.0776 0.0649 0.0538
sigma^2 = 0.002913: log likelihood = 548.02
AIC=-1086.03 AICc=-1085.87 BIC=-1066.55
   6) This is the model fitted.
      > model1
      Call:
      arima(x = new_price, order = c(1, 0, 3))
      Coefficients:
            ar1 ma1 ma2 ma3 intercept
          -0.8911 0.7864 0.040 0.1882
                                          0.0015
      s.e. 0.0615 0.0776 0.065 0.0539 0.0030
      sigma<sup>2</sup> estimated as 0.002879: log likelihood = 548.14, aic = -1084.28
   7) Used some tests to check the model fitted.
      > print(box_pierce_test)
             Box-Pierce test
      data: res1
      X-squared = 15.254, df = 10, p-value = 0.1231
      > print(ljung_box_test)
             Box-Ljung test
      data: res1
      X-squared = 15.686, df = 10, p-value = 0.109
```

Using these tests, we see as the p value is more than 5%, our model is fitted very accurately.

8) Summary of the model1 and its evaluation

```
> summary(model1)
```

Call:

```
arima(x = new_price, order = c(1, 0, 3))
```

Coefficients:

```
ar1 ma1 ma2 ma3 intercept
-0.8911 0.7864 0.040 0.1882 0.0015
s.e. 0.0615 0.0776 0.065 0.0539 0.0030
```

sigma^2 estimated as 0.002879: log likelihood = 548.14, aic = -1084.28

Training set error measures:

```
ME RMSE MAE MPE MAPE MASE ACF1
Training set 2.284797e-06 0.05365766 0.03591419 104.7944 190.7375 0.6309907 0.005007664
```

By looking at these MAE and MPE, we can say that our model is fitted so well.

9) Check forecasted prices

> forecasted prices

```
Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

365 -0.011576717 -0.08034177 0.05718834 -0.11674380 0.09359036

366 0.002002794 -0.06713849 0.07114407 -0.10373967 0.10774526

367 -0.010278794 -0.08002514 0.05946755 -0.11694662 0.09638903

368 0.011981136 -0.05792814 0.08189041 -0.09493588 0.11889815

369 -0.007854984 -0.07789338 0.06218341 -0.11496946 0.09925949

370 0.009821246 -0.06031951 0.07996200 -0.09744978 0.11709227

371 -0.005930278 -0.07615220 0.06429165 -0.11332545 0.10146489

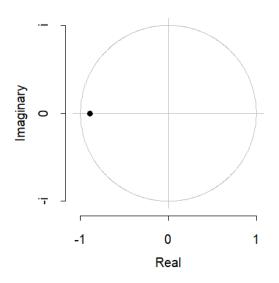
372 0.008106115 -0.06218020 0.07839243 -0.09938753 0.11559976

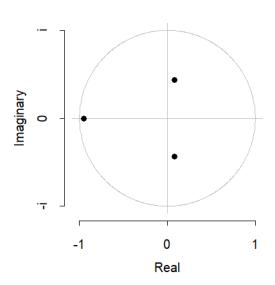
373 -0.004401901 -0.07473931 0.06593551 -0.11197369 0.10316988

374 0.006744158 -0.06363380 0.07712211 -0.10088963 0.11437795
```

Inverse AR roots

Inverse MA roots





Variation of ARIMA with order = c(1,1,2)

```
1) Model fitted> model2
```

Call:

 $arima(x = new_price, order = c(1, 1, 2))$

Coefficients:

ar1 ma1 ma2 -0.4418 -0.6750 -0.325 s.e. 0.2277 0.2373 0.237

sigma 2 estimated as 0.002976: log likelihood = 537.73, aic = -1067.46

2) Evaluation with Box, L jung tests> print(box_pierce_test2)

Box-Pierce test

```
data: res2
X-squared = 28.205, df = 10, p-value = 0.001674
```

Using these tests, we see as the p value is less than 5%, our model is not fitted accurately with this variation.

3) Summary of this model

```
> summary(model2)
Call:
arima(x = new price, order = c(1, 1, 2))
Coefficients:
     ar1
            ma1
                   ma2
   -0.4418 -0.6750 -0.325
s.e. 0.2277 0.2373 0.237
sigma<sup>2</sup> estimated as 0.002976: log likelihood = 537.73, aic = -1067.46
Training set error measures:
                               MAE
                                       MPE
                                               MAPE
                                                        MASE
             ME
                    RMSE
                                                                   ACF1
Training set 0.001012948 0.05447365 0.03602601 100.4219 141.9838 0.6329554
```

In this case, the MAE and MPE error is comparatively higher than model1. So, this is not a good model.

4) Forecast

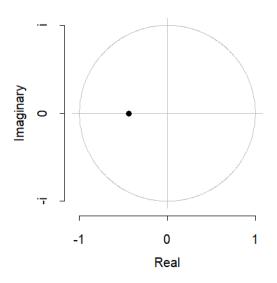
0.01500424

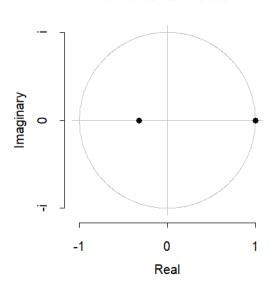
```
> forecasted_prices2
Point Forecast Lo 80 Hi 80
```

```
Point Forecast Lo 80 Hi 80 Lo 95 Hi 95
365 0.0036371468 -0.06636560 0.07363989 -0.1034228 0.1106971
366 0.0006067545 -0.06984962 0.07106313 -0.1071470 0.1083605
367 0.0019456545 -0.06861195 0.07250326 -0.1059629 0.1098542
368 0.0013540964 -0.06921750 0.07192569 -0.1065758 0.1092840
369 0.0016154609 -0.06896141 0.07219233 -0.1063225 0.1095535
370 0.0014999838 -0.06907679 0.07207676 -0.1064379 0.1094378
371 0.0015510044 -0.06902625 0.07212826 -0.1063876 0.1094896
372 0.0015284623 -0.06904867 0.07210559 -0.1064099 0.1094669
373 0.0015384219 -0.06903878 0.07211562 -0.1064001 0.1094769
374 0.0015340215 -0.06904315 0.07211119 -0.1064044 0.1094725
```

Inverse AR roots

Inverse MA roots





Variation of ARIMA with order = c(1,1,2)

1) Model Fitted

> model3

Call:

 $arima(x = new_price, order = c(2, 0, 2))$

Coefficients:

ar1 ar2 ma1 ma2 intercept 0.3486 -0.0472 -0.4606 0.2331 0.0015 s.e. 0.4082 0.2481 0.4021 0.2255 0.0031

sigma² estimated as 0.00291: log likelihood = 546.29, aic = -1080.58

2) Evaluation

> print(box_pierce_test3)

Box-Pierce test

data: res3

X-squared = 18.548, df = 10, p-value = 0.0464

> print(ljung_box_test3)

Box-Ljung test

```
data: res3
X-squared = 19.099, df = 10, p-value = 0.03902
```

Using these tests, we see as the p value is less than 5%, our model is fitted very accurately. But this is better than the previous variation.

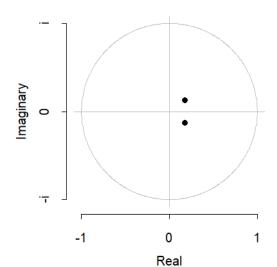
```
3) Summary of the model
   > summary(model3)
   Call:
   arima(x = new price, order = c(2, 0, 2))
   Coefficients:
        ar1
               ar2
                            ma2 intercept
                     ma1
      0.3486 -0.0472 -0.4606 0.2331
                                        0.0015
   s.e. 0.4082 0.2481 0.4021 0.2255
                                         0.0031
   sigma<sup>2</sup> estimated as 0.00291: log likelihood = 546.29, aic = -1080.58
   Training set error measures:
                                   MAE
                                           MPE
                                                  MAPE
                ME
                        RMSE
                                                            MASE
   Training set 4.066745e-06 0.05394287 0.03597446 115.9026 174.8021
   0.6320497
               ACF1
   Training set 0.0004684999
```

In this case the error of MAE and MPE is higher than the 1st model but lesser than the previous variation. So, this variation is better than the previous variation.

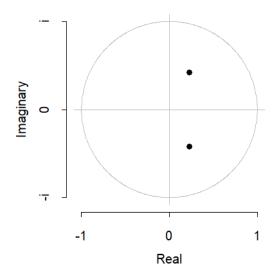
4) Forecast

```
> forecasted_prices3
Point Forecast Lo 80 Hi 80 Lo 95 Hi 95
365 -0.0016075758 -0.07073815 0.06752300 -0.1073337 0.1041185
366 -0.0089427256 -0.07850586 0.06062040 -0.1153304 0.0974449
367 -0.0019987530 -0.07229858 0.06830108 -0.1095131 0.1055156
368 0.0007678573 -0.06964028 0.07117600 -0.1069121 0.1084478
369 0.0014047979 -0.06900887 0.07181846 -0.1062836 0.1090932
370 0.0014963537 -0.06891742 0.07191013 -0.1061922 0.1091849
371 0.0014982300 -0.06891554 0.07191200 -0.1061903 0.1091868
372 0.0014945661 -0.06891921 0.07190834 -0.1061940 0.1091831
373 0.0014932003 -0.06892057 0.07190697 -0.1061954 0.1091818
374 0.0014928970 -0.06892088 0.07190667 -0.1061957 0.1091815
```





Inverse MA roots



GARCH and its variations

- 1) Since the stationarity is checked for new_price, we will directly check if there is any arch effect or not.
- > ArchTest(new_price)

ARCH LM-test; Null hypothesis: no ARCH effects

```
data: new_price
Chi-squared = 29.955, df = 12, p-value = 0.002836
```

Since p value is less than 5%, we conclude that there is no arch effect present. So, we can use garch model.

2) Model fitted with order = c(1,1)

```
*-----*
* GARCH Model Fit
* *
```

Conditional Variance Dynamics

GARCH Model : sGARCH(1,1) Mean Model : ARFIMA(1,0,1)

Distribution : norm

Optimal Parameters

Estimate Std. Error t value Pr(>|t|)
mu 0.003626 0.001864 1.9452 0.051757
ar1 -0.516039 0.169928 -3.0368 0.002391
ma1 0.336119 0.186252 1.8047 0.071130
omega 0.000537 0.000127 4.2217 0.000024
alpha1 0.418530 0.113450 3.6891 0.000225
beta1 0.475112 0.075884 6.2610 0.000000

Robust Standard Errors:

Estimate Std. Error t value Pr(>|t|)
mu 0.003626 0.002097 1.7289 0.083826
ar1 -0.516039 0.178154 -2.8966 0.003772
ma1 0.336119 0.176163 1.9080 0.056392
omega 0.000537 0.000191 2.8142 0.004891
alpha1 0.418530 0.146349 2.8598 0.004239
beta1 0.475112 0.096712 4.9127 0.000001

LogLikelihood: 585.2012

Information Criteria

Akaike -3.1824 Bayes -3.1182 Shibata -3.1830 Hannan-Quinn -3.1569

Weighted Ljung-Box Test on Standardized Residuals

statistic p-value Lag[1] 1.226 0.2682

d.o.f=2

H0: No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals

statistic p-value

Lag[1] 0.504 0.4778 Lag[2*(p+q)+(p+q)-1][5] 1.033 0.8521 Lag[4*(p+q)+(p+q)-1][9] 1.555 0.9510 d.o.f=2

Weighted ARCH LM Tests

Nyblom stability test

Joint Statistic: 1.4126 Individual Statistics: mu 0.34180 ar1 0.06084 ma1 0.13426 omega 0.37240 alpha1 0.09509 beta1 0.12617

Asymptotic Critical Values (10% 5% 1%)

Joint Statistic: 1.49 1.68 2.12

Individual Statistic: 0.35 0.47 0.75

Sign Bias Test

t-value prob sig
Sign Bias 0.5506 0.5823
Negative Sign Bias 0.7291 0.4664
Positive Sign Bias 0.6136 0.5399
Joint Effect 1.4905 0.6845

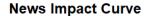
Adjusted Pearson Goodness-of-Fit Test:

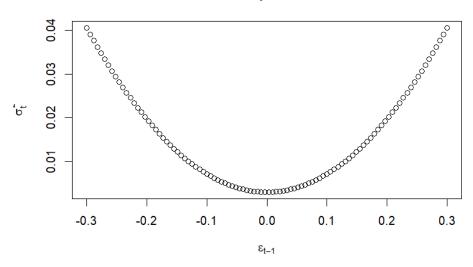
group statistic p-value(g-1)

1 20 37.87 0.006168 2 30 51.38 0.006379 3 40 56.66 0.033492 4 50 62.37 0.095010

Elapsed time: 0.1376641

3) News impact curve





As it is symmetric then our model is fitted well.

4) Evaluated our model

> print(box_pierce_test4)

Box-Pierce test

data: res4

X-squared = 29.2, df = 10, p-value = 0.001156

> print(ljung_box_test4)

Box-Ljung test

data: res4

X-squared = 29.898, df = 10, p-value = 0.0008902

From the p value, it shows that the model is fitted well.

5) Forecast

* GARCH Model Forecast

Model: sGARCH Horizon: 10

Roll Steps: 0 Out of Sample: 0

0-roll forecast [T0=1970-12-31 05:30:00]:

Series Sigma

T+1 0.006987 0.05631

T+2 0.001891 0.05806

T+3 0.004521 0.05957

T+4 0.003164 0.06090

T+5 0.003864 0.06206

T+6 0.003503 0.06307

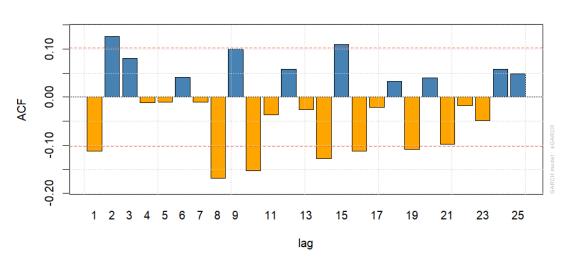
T+7 0.003689 0.06397

T+8 0.003593 0.06476

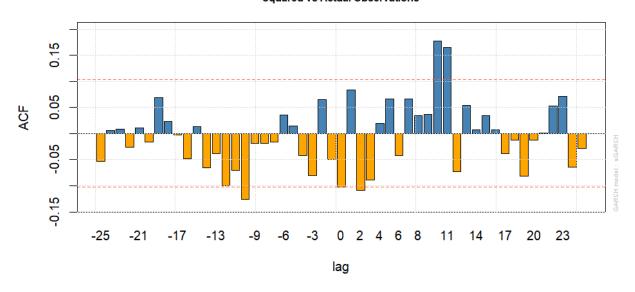
T+9 0.003643 0.06546

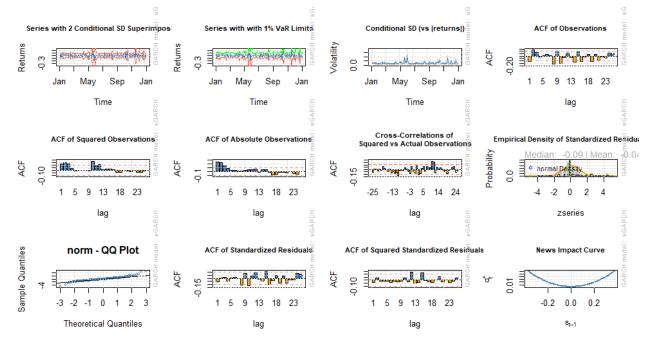
T+10 0.003617 0.06608

ACF of Observations



Cross-Correlations of **Squared vs Actual Observations**





eGARCH

1) Model fitted with order = c(1,1)

GARCH Model Fit

Conditional Variance Dynamics

 $\begin{array}{ll} \mathsf{GARCH}\;\mathsf{Model} & : \mathsf{eGARCH}(1,1) \\ \mathsf{Mean}\;\mathsf{Model} & : \mathsf{ARFIMA}(1,0,1) \end{array}$

Distribution : norm

Optimal Parameters

Estimate Std. Error t value Pr(>|t|)
mu 0.001377 0.001867 0.73755 0.460787
ar1 -0.579751 0.101139 -5.73220 0.000000
ma1 0.405483 0.115593 3.50786 0.000452
omega -1.288585 0.273829 -4.70580 0.000003
alpha1 -0.167905 0.078153 -2.14841 0.031681
beta1 0.778874 0.045280 17.20136 0.000000
gamma1 0.630077 0.101068 6.23421 0.000000

Robust Standard Errors:

Estimate Std. Error t value Pr(>|t|)
mu 0.001377 0.001842 0.74751 0.454756
ar1 -0.579751 0.078248 -7.40910 0.000000
ma1 0.405483 0.075311 5.38408 0.000000
omega -1.288585 0.415396 -3.10206 0.001922
alpha1 -0.167905 0.086287 -1.94589 0.051668
beta1 0.778874 0.063647 12.23733 0.000000
gamma1 0.630077 0.103738 6.07373 0.000000

LogLikelihood: 590.3509

Information Criteria

Akaike -3.2052 Bayes -3.1303 Shibata -3.2059 Hannan-Quinn -3.1754

Weighted Ljung-Box Test on Standardized Residuals

statistic p-value

Lag[1] 1.513 0.2187 Lag[2*(p+q)+(p+q)-1][5] 2.006 0.9569 Lag[4*(p+q)+(p+q)-1][9] 3.853 0.7251

d.o.f=2

H0: No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals

statistic p-value

Lag[1] 0.2579 0.6116 Lag[2*(p+q)+(p+q)-1][5] 0.8514 0.8923 Lag[4*(p+q)+(p+q)-1][9] 1.3338 0.9680

d.o.f=2

Weighted ARCH LM Tests

Statistic Shape Scale P-Value

ARCH Lag[3] 0.06924 0.500 2.000 0.7924

ARCH Lag[5] 0.26213 1.440 1.667 0.9502

ARCH Lag[7] 0.61665 2.315 1.543 0.9666

Nyblom stability test

Joint Statistic: 1.2476 Individual Statistics: mu 0.27537 ar1 0.02897 ma1 0.05199 omega 0.18004 alpha1 0.05097 beta1 0.20826 gamma1 0.05328

Asymptotic Critical Values (10% 5% 1%)

Joint Statistic: 1.69 1.9 2.35

Individual Statistic: 0.35 0.47 0.75

Sign Bias Test

t-value prob sig
Sign Bias 0.5705 0.5687
Negative Sign Bias 0.9071 0.3650
Positive Sign Bias 0.1151 0.9084
Joint Effect 2.3624 0.5007

Adjusted Pearson Goodness-of-Fit Test:

group statistic p-value(g-1)

1 20 25.23 0.1531 2 30 38.53 0.1110 3 40 48.53 0.1410 4 50 56.33 0.2198

Elapsed time: 0.1413

2) Evaluating model

> print(box_pierce_test5)

Box-Pierce test

data: res5

$$X$$
-squared = 29.2, df = 10, p-value = 0.001156

> print(ljung_box_test5)

Box-Ljung test

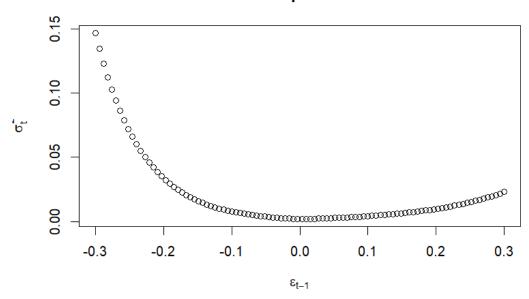
data: res5

X-squared = 29.898, df = 10, p-value = 0.0008902

From looking at the p value, it is clear that this model is not fitted well.

3) News Impact Curve

News Impact Curve



Since it is not symmetric, the model is not very good.

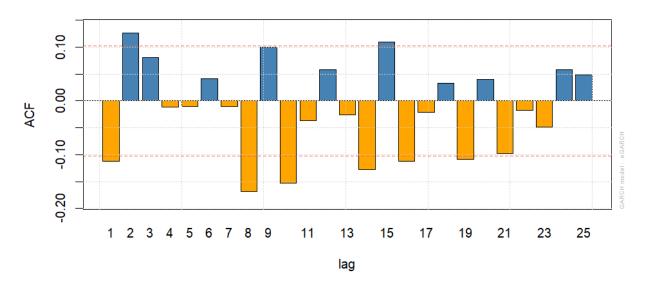
4) Forecast

* GARCH Model Forecast

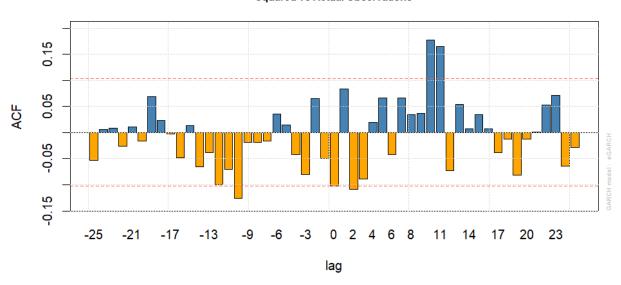
Model: eGARCH Horizon: 10 Roll Steps: 0 Out of Sample: 0

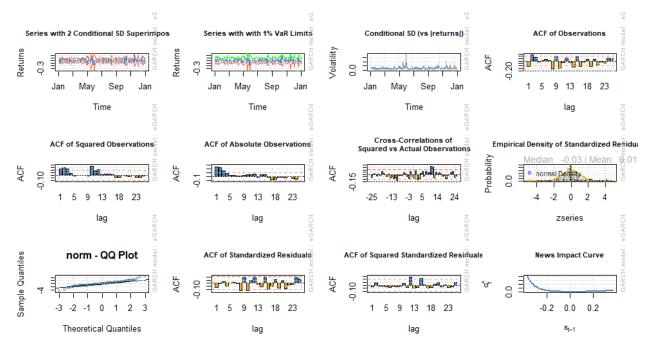
0-roll forecast [T0=1970-12-31 05:30:00]: Series Sigma T+1 0.0033754 0.06491
T+2 0.0002185 0.06239
T+3 0.0020487 0.06050
T+4 0.0009876 0.05906
T+5 0.0016028 0.05797
T+6 0.0012462 0.05713
T+7 0.0014529 0.05649
T+8 0.0013331 0.05599
T+9 0.0014026 0.05561
T+10 0.0013623 0.05531

ACF of Observations



Cross-Correlations of Squared vs Actual Observations





gjrGARCH

1) Model fitted with order = c(1,1)

* GARCH Model Fit

* *

Conditional Variance Dynamics

GARCH Model : gjrGARCH(1,1) Mean Model : ARFIMA(1,0,1)

Distribution : norm

Optimal Parameters

Estimate Std. Error t value Pr(>|t|) mu 0.002316 0.001761 1.3149 0.188545 ar1 -0.550397 0.167346 -3.2890 0.001006 ma1 0.344477 0.191305 1.8007 0.071755 omega 0.000547 0.000113 4.8403 0.000001 alpha1 0.219058 0.076159 2.8763 0.004023 beta1 0.402340 0.066271 6.0712 0.000000 gamma1 0.711742 0.280479 2.5376 0.011162

Robust Standard Errors:

Estimate Std. Error t value Pr(>|t|) 0.002316 0.001806 1.2826 0.199615 ar1 -0.550397 0.202311 -2.7205 0.006517 ma1 0.344477 0.227254 1.5158 0.129564 omega 0.000547 0.000154 3.5575 0.000374 alpha1 0.219058 0.074346 2.9465 0.003214 beta1 0.402340 0.064403 6.2472 0.000000 gamma1 0.711742 0.390089 1.8246 0.068067

LogLikelihood: 590.6057

Information Criteria

Akaike -3.2066 Bayes -3.1317 Shibata -3.2073 Hannan-Quinn -3.1768

Weighted Ljung-Box Test on Standardized Residuals

statistic p-value

Lag[1] 1.320 0.2507 Lag[2*(p+q)+(p+q)-1][5] 1.871 0.9775 Lag[4*(p+q)+(p+q)-1][9] 3.906 0.7129

d.o.f=2

H0: No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals

statistic p-value

0.1607 0.6885 Lag[1] Lag[2*(p+q)+(p+q)-1][5] 0.9800 0.8642 Lag[4*(p+q)+(p+q)-1][9] 1.4253 0.9615 d.o.f=2

Weighted ARCH LM Tests

Statistic Shape Scale P-Value

ARCH Lag[3] 0.3362 0.500 2.000 0.5621

ARCH Lag[5] 0.4672 1.440 1.667 0.8933

ARCH Lag[7] 0.8056 2.315 1.543 0.9430

Nyblom stability test

Joint Statistic: 1.3964 Individual Statistics:

mu 0.3617

ar1 0.1349

ma1 0.2017

omega 0.2650

alpha1 0.1757

beta1 0.1471

gamma1 0.1677

Asymptotic Critical Values (10% 5% 1%)

Joint Statistic: 1.69 1.9 2.35

Individual Statistic: 0.35 0.47 0.75

Sign Bias Test

t-value prob sig

Sign Bias 0.7246 0.4692

Negative Sign Bias 1.0578 0.2909 Positive Sign Bias 0.1665 0.8678

Joint Effect 3.4138 0.3321

Adjusted Pearson Goodness-of-Fit Test:

group statistic p-value(g-1)

1 20 26.22 0.1242

2 30 35.73 0.1817 3 40 44.35 0.2562

4 50 50.84 0.4012

Elapsed time: 0.3211601

2) Evaluating model

> print(box_pierce_test6)

Box-Pierce test

data: res6

X-squared = 30.745, df = 10, p-value = 0.0006465

> print(ljung_box_test6)

Box-Ljung test

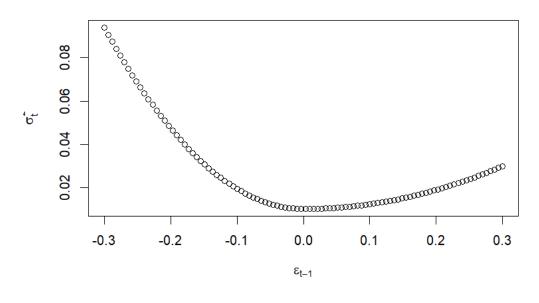
data: res6

X-squared = 31.449, df = 10, p-value = 0.0004944

From looking at the p value, it is clear that this model is not fitted well.

3) News Impact Curve

News Impact Curve



Again, it is not symmetric. So, the model is not good.

4) Forecast

* GARCH Model Forecast

* * ***

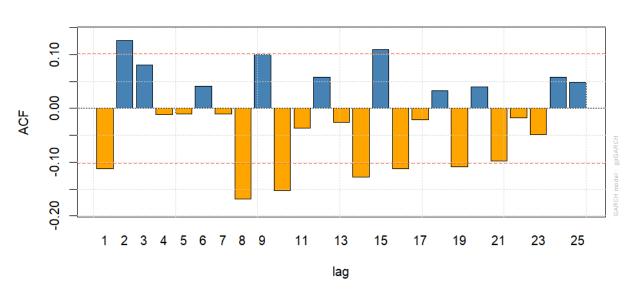
Model: gjrGARCH

Horizon: 10 Roll Steps: 0 Out of Sample: 0

0-roll forecast [T0=1970-12-31 05:30:00]:

Series Sigma T+1 0.0058244 0.07274 T+2 0.0003848 0.07561
T+3 0.0033787 0.07832
T+4 0.0017309 0.08088
T+5 0.0026379 0.08330
T+6 0.0021387 0.08560
T+7 0.0024134 0.08779
T+8 0.0022622 0.08988
T+9 0.0023454 0.09188
T+10 0.0022996 0.09379

ACF of Observations



Cross-Correlations of Squared vs Actual Observations

