FE 542- HW1

Problem 1

An urn contains three type A coins and two type B coin. When a type A coin is ipped, it comes up heads with probability 1/3, whereas a type B coin is ipped, it comes up heads with robability 2/3. A coin is randomly chosen from the urn and ipped. Given that theip landed on heads, what is the probability that it was a type B coin?

Solution:

0.5714286

Prob(B|heads)= P(heads|B).P(B)/(P(heads))

P(heads|B)=if the coin is B prob of getting head is 2/3 which is given in the question

P(heads|A)=if the coin is A prob of getting head is 1/3 which is given in the question

P(B)=2/5 There are totally 5 coins and 2 of them are B coins so the probability of chosing B coin is 2/5

P(A)=3/5 There are totally 5 coins and 3 of them are A coins so the probability of chosing A coin is 3/5

P(heads)= P(heads|B).P(B)+P(heads|A).P(A)

Prob(B|heads)= P(heads|B).P(B)/(P(heads))

- Calculating denominator of the equation

```
denominator=2/3*2/5 + 1/3*3/5
denominator

## [1] 0.4666667
```

- Calculating numerator of the equation

```
numerator=2/3*2/5
numerator

## [1] 0.2666667
```

- Showing the result

```
result=numerator/denominator

result

## [1] 0.5714286
```

Problem 2

(i) Generate 100 observations from a normal distribution with mean 2 and variance 7.

```
sample=rnorm(100,mean=2,sd=sqrt(7))

sample

##   [1] -4.39319271  2.14485098  0.16782802  1.24999814  1.95085494
3.37695288
##   [7] -0.09983850  3.51536906  0.25479209 -3.41737617 -2.83300008
1.26784929
##  [13]  6.21828415  4.93559201  2.92908836  3.68103218 -0.70468085
5.52113811
##  [19] -2.68118838  5.70004080  5.38311197  2.63408977 -0.54217118 -
0.02963278
##  [25]  3.32986502 -1.71698019  6.18417602  7.74089906 -1.34983370
1.16928112
##  [31]  6.36333014  6.09995627  7.82872342 -0.05117579 -0.99696227
1.70894251
##  [37]  6.69355490  1.96133355  4.69851664  3.07492388  3.44747303
1.49423940
##  [43]  1.16253392  4.86653623  2.84427852  0.24793108 -2.60061588
7.07477571
##  [49] -0.09511843  0.64983504  3.45096968 -0.32563908  2.09012501
0.78329327
##  [55] -0.73704266  1.82728500 -0.54922955  3.73543878 -1.14264523 -
0.04276004
##  [61] -0.39335280  4.85599984 -1.40847510 -0.97624015  2.58707994
1.70742788
##  [67]  2.24554188  1.45597011 -3.08437790 -0.61904290  1.97440317
9.51518446
##  [73] -1.75449114  5.15955868  1.12404805 -0.97147011  1.43355924
3.22375405
##  [79]  5.48369089  6.93110660 -2.29646795 -3.49811820  0.84573388
7.27476165
##  [85]  0.08175841  3.85296690 -1.83233339  2.80884243  1.50156079 -
2.05071184
##  [91]  1.05228487  5.93271220  1.94415859 -0.26729167  2.45244755
3.08377865
##  [97] -2.05149199  4.90560814  2.68053165  2.97145635
```

(ii) Compute the sample mean, standard deviation, skewness and kurtosis (if excess kurtosis state that clearly).

```
mean(sample)

## [1] 1.850321

sd(sample)

## [1] 2.977175
```

- Measure of assymetry
- if skewness is positive, the data is skewed right
- if skewness is negative, the data is skewed left
- if skewness is between −1 and −½ or between +½ and +1, the distribution is moderately skewed.
- If skewness is between −½ and +½, the distribution is approximately symmetric.
- If skewness is less than −1 or greater than +1, the distribution is highly skewed.

We can say that sample is approximately symmetric since the skewness value is between −½ and +½

```
library(fBasics)

## Loading required package: timeDate

## Loading required package: timeSeries

skewness(sample)

## [1] 0.2393729
## attr(,"method")
## [1] "moment"
```

- Measure of tail thickness
- Negative excess kurtosis has short tails Platykurtic
- Positive excess kurtosis have heavy tails Leptokurtic

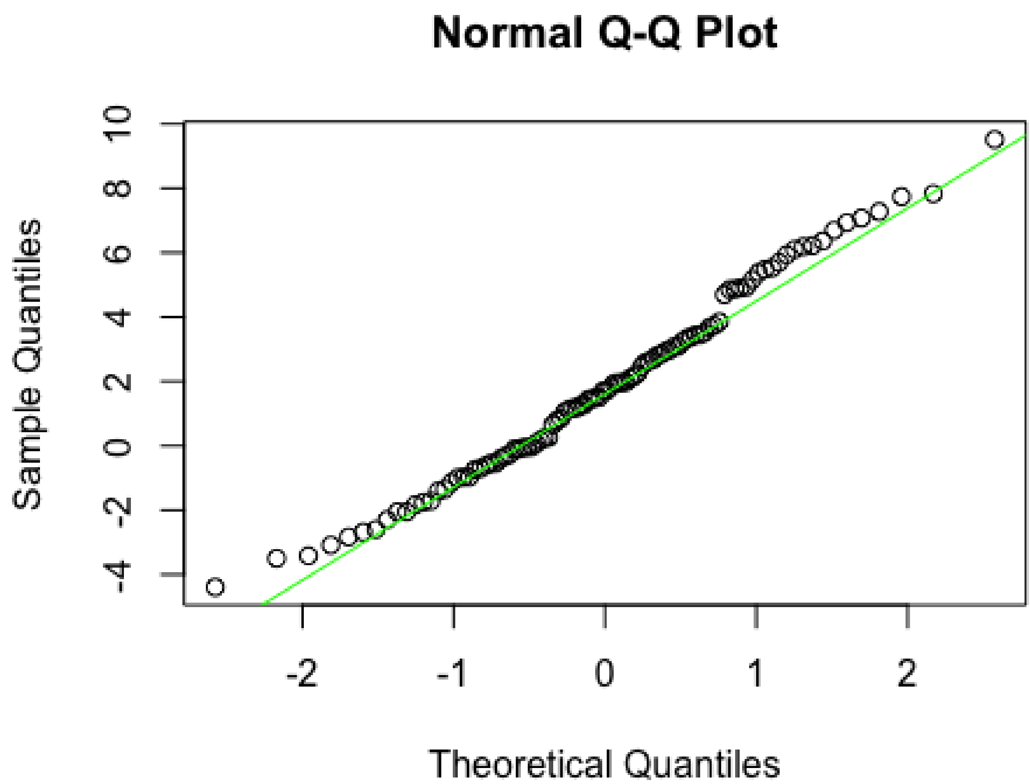
We can say that sample is Platykurtic(tails are light) since excess kurtosis value is negative

```
kurtosis(sample)

## [1] -0.5862402
## attr(,"method")
## [1] "excess"
```

(iii) Generate a Q-Q plot of the observations from (i) versus the standard normal distribution. Describe your plot in no more than 3 sentences.

- Data is fairly normal distributed since majority of the data points in line with the 45 degree normal line.

```
qqnorm(sample)
qqline(sample,col="green")
```



Problem 3

(i) Download daily price data for January 1, 2017 through December 31, 2020 of Microsoft stock from Yahoo Finance. You may use the quantmod package in R for this purpose.

```
library(quantmod)

## Loading required package: xts

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following object is masked from 'package:timeSeries':
##
##     time<-

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Loading required package: TTR
```

```
##
## Attaching package: 'TTR'

## The following object is masked from 'package:fBasics':
##
##     volatility

## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

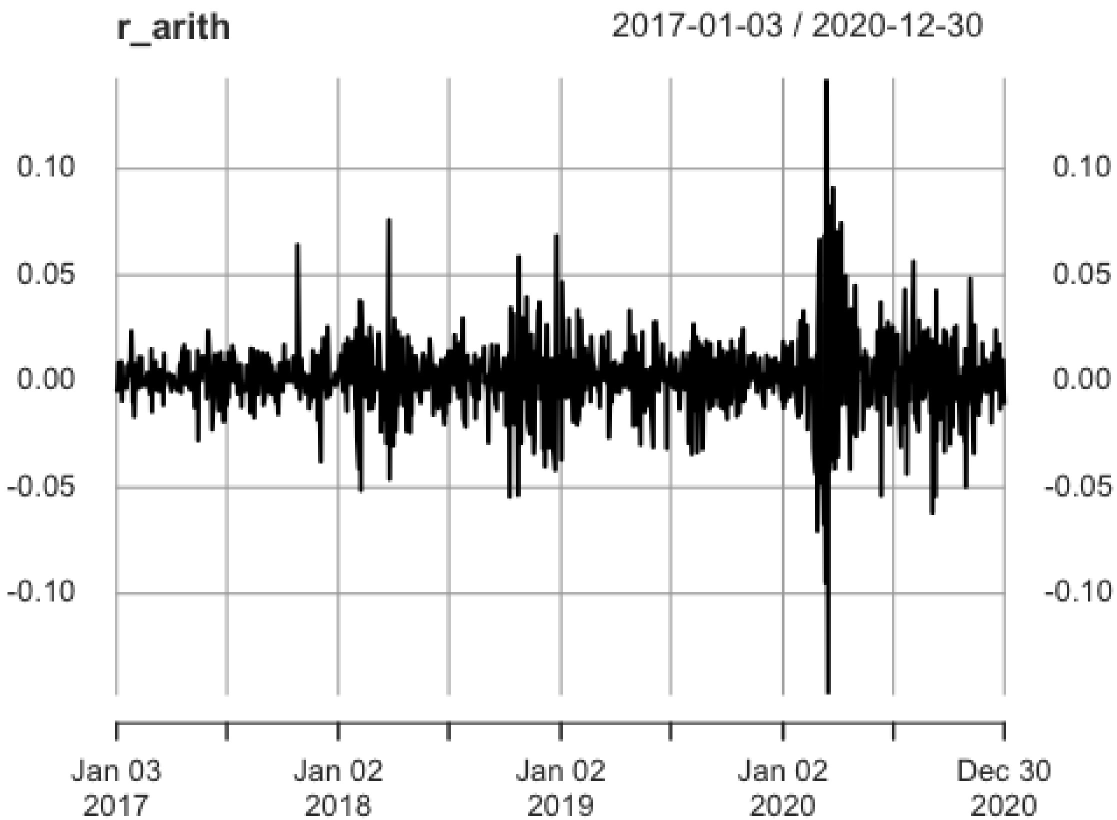
```
getSymbols('MSFT',src='yahoo',from='2017-01-01',to='2020-12-31')
```

```
## 'getSymbols' currently uses auto.assign=TRUE by default, but will
## use auto.assign=FALSE in 0.5-0. You will still be able to use
## 'loadSymbols' to automatically load data. getOption("getSymbols.env")
## and getOption("getSymbols.auto.assign") will still be checked for
## alternate defaults.
##
## This message is shown once per session and may be disabled by setting
## options("getSymbols.warning4.0"=FALSE). See ?getSymbols for details.

## [1] "MSFT"
```

(ii) Generate a plot of the closing prices and of daily log returns.

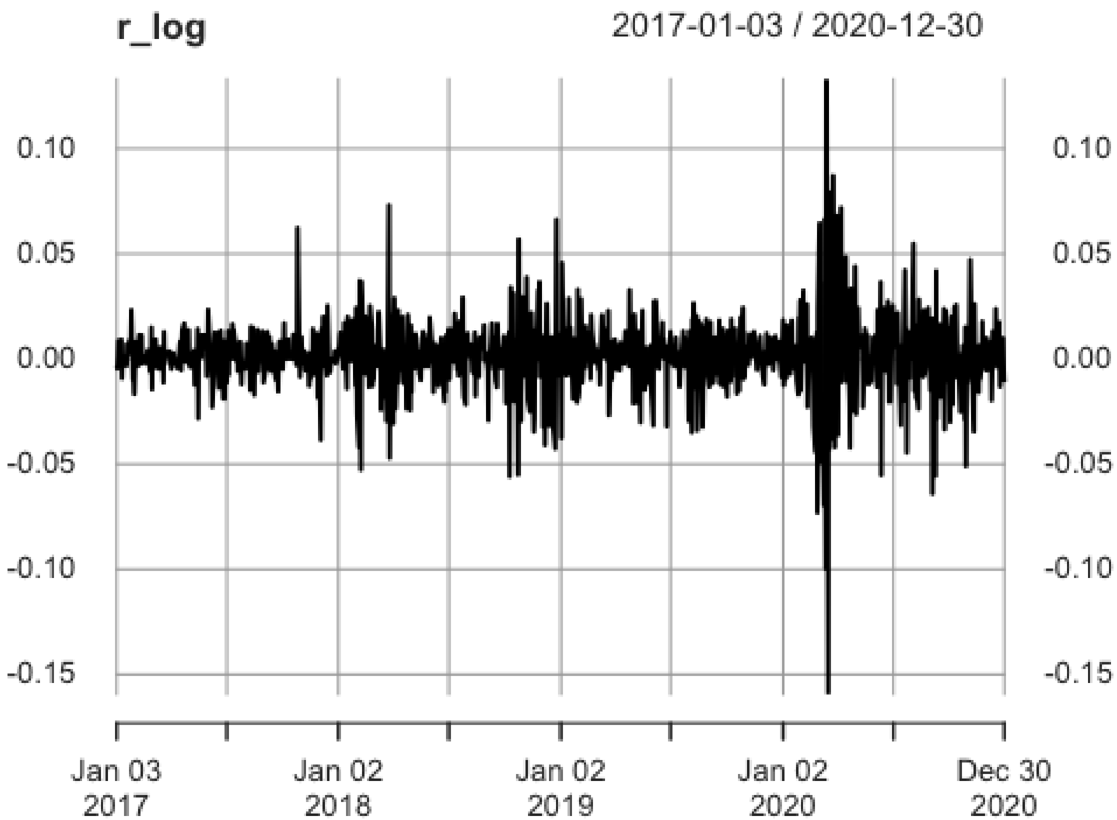
- plot of the daily return

```
r_arith=dailyReturn(MSFT)

plot(r_arith)
```



- plot of daily log return

```
r_log = dailyReturn(MSFT,type='log')

plot(r_log)
```



(iii) Compute the sample mean, standard deviation, skewness and kurtosis (if excess kurtosis state that clearly) of the daily log returns.

```
mean(r_log)

## [1] 0.001253916

sd(r_log)

## [1] 0.01822633

skewness(r_log)

## [1] -0.3733297
## attr(,"method")
## [1] "moment"

kurtosis(r_log)

## [1] 11.62756
## attr(,"method")
## [1] "excess"
```

(iv) Perform the Jarque-Bera Normality test (with 5% significance level) for daily log returns and interpret your results. Compare your result a visualization from a Q-Q plot.

Ho: (Normal Distribution) Ha: (Non-Normal Distribution)

- p value of Jarque - Bera Normalality Test Normalality Test is very small and smaller than significance level(0.05) so we can reject null hypothesis. Hence, the data is not normally distributed

```
normalTest(r_log,method="jb")
```

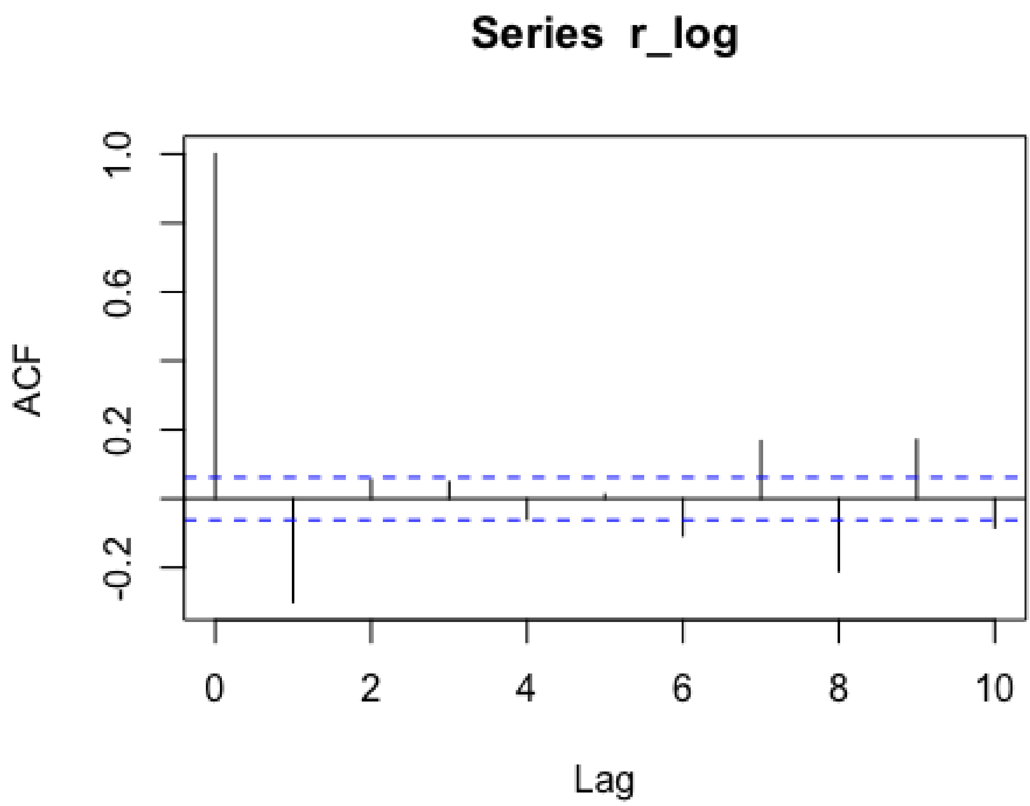
```
## Warning in if (class(x) == "fREG") x = residuals(x): the condition has
length >
## 1 and only the first element will be used
```

```
##
## Title:
##  Jarque - Bera Normalality Test
##
## Test Results:
##   STATISTIC:
##     X-squared: 5719.0022
##   P VALUE:
##     Asymptotic p Value: < 2.2e-16
##
## Description:
##  Sat Feb 20 06:30:24 2021 by user:
```

(v) Compute and plot the lag-1 through lag-10 autocorrelation of the daily log returns.

- plot the autocorrelation
- calculate the autocorrelation

```
acf(r_log, lag.max=10, plot=TRUE)
```



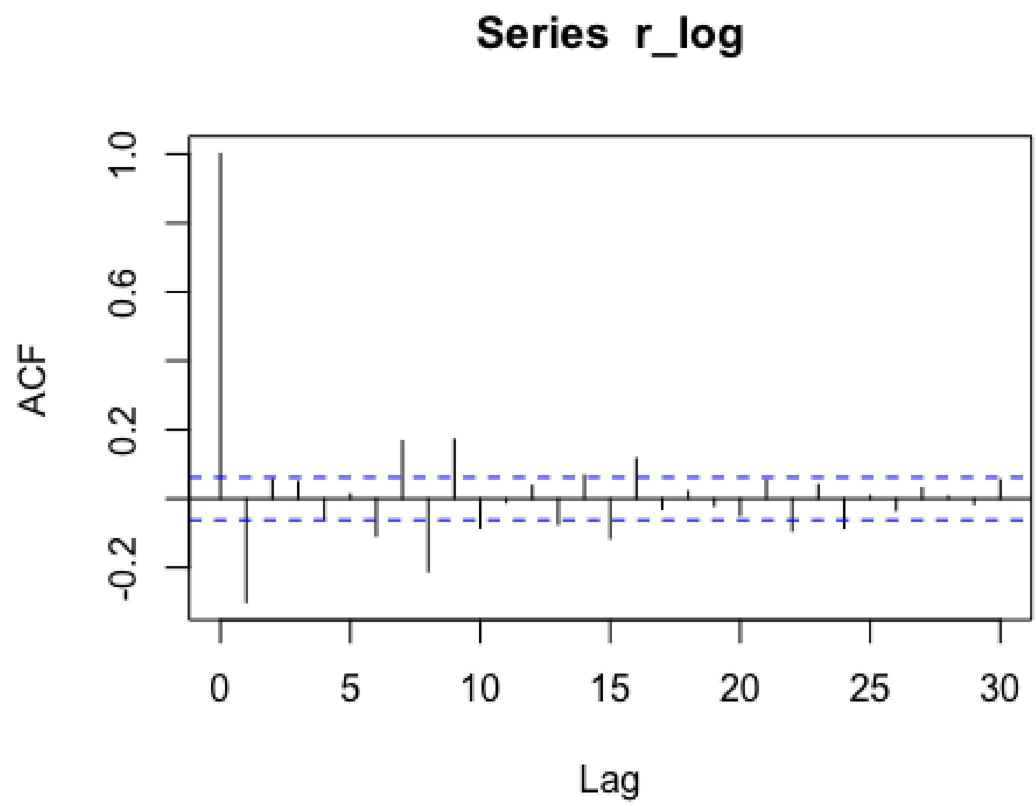

```
acf(r_log, lag.max=10,plot="False")

##
## Autocorrelations of series 'r_log', by lag
##
##      0      1      2      3      4      5      6      7      8      9
10
##  1.000 -0.300  0.055  0.049 -0.059  0.013 -0.107  0.167 -0.211  0.171 -
0.084
```

(vi) Test if the lag-5 autocorrelation is non-zero (with 5% significance level). Interpret your results.

H0 : ρl = 0 Ha : ρl not equal 0 if |t ratio|>Zα/2, we reject null hypothesis

- According to individual ACF test, our t ratio is not bigger than qnorm value. Hence, we fail to reject null.

```
ac = acf(r_log,plot=TRUE)
```




```
T = length(r_log)
ratio = sqrt(T)*ac$acf[6]

significance = 0.05
abs(ratio) > qnorm(1-significance/2) #FALSE = FAIL TO REJECT NULL

## [1] FALSE
```

(vi) Test if the lag-5 autocorrelation is non-zero (with 5% significance level). Interpret your results.

(vii) Perform the Ljung-Box test to test if the first 2, first 5, and first 8 lagged autocorrelations are non-zero (with 5% significance level). Interpret your results.

H0: The data are independently distributed(no auto correlation) Ha: The data are not independently distributed; they exhibit serial correlation.

-First two autocorrelations are non zero because p- value of the Ljung-Box Tests is too small with given 5% significance level. Hence, we can reject null hypothesis(no auto correlation)

```
Box.test(r_log,lag=2,type="Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  r_log
## X-squared = 93.936, df = 2, p-value < 2.2e-16
```

- First 5 autocorrelations are non zero because p- value of the Ljung-Box Tests is too small in comparison to given 5% significance level. Hence, we can reject null hypothesis(no auto correlation)

```
Box.test(r_log,lag=5,type="Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  r_log
## X-squared = 100.04, df = 5, p-value < 2.2e-16
```

- First 8 autocorrelations are non zero because p- value of the Ljung-Box Tests is too small in comparison to given 5% significance level. Hence, we can reject null hypothesis(no auto correlation)

```
Box.test(r_log,lag=8,type="Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  r_log
## X-squared = 185.33, df = 8, p-value < 2.2e-16
```