Stat 461 HW5

Roshan Poduval

12/8/2020

```
#getSymbols("AAPL",from="2010-07-31",to="2020-07-31") # this won't work for me, so I am just goi
ng to download the csv from yahoo for the correct date range
AAPL = read.table("AAPL.csv", sep=",", header=TRUE)
adj_price = AAPL$Adj.Close
log_rt = diff(log(adj_price))
```

```
mean(log_rt)
```

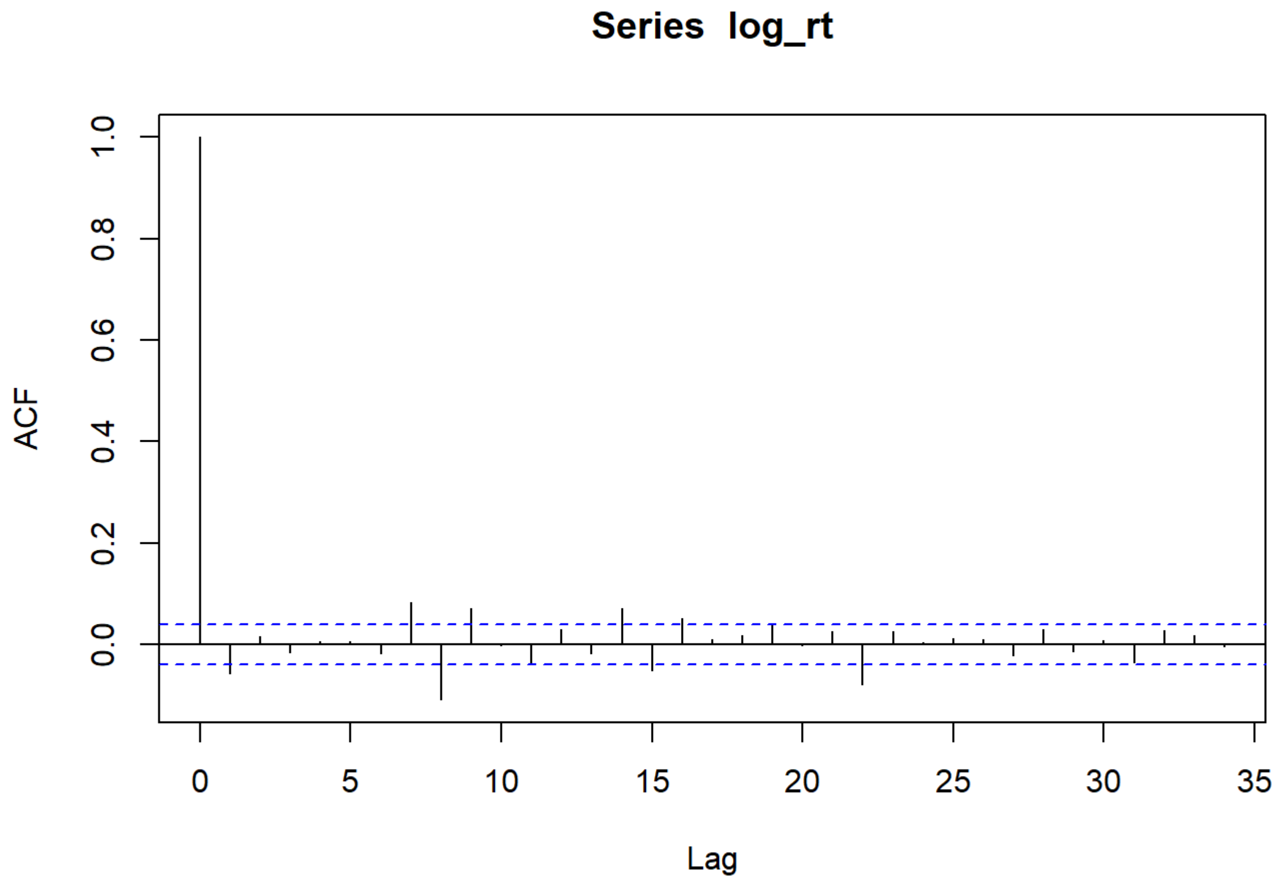
```
## [1] 0.0009869365
```

```
t.test(log_rt, mu=0, alternative="two.sided")
```

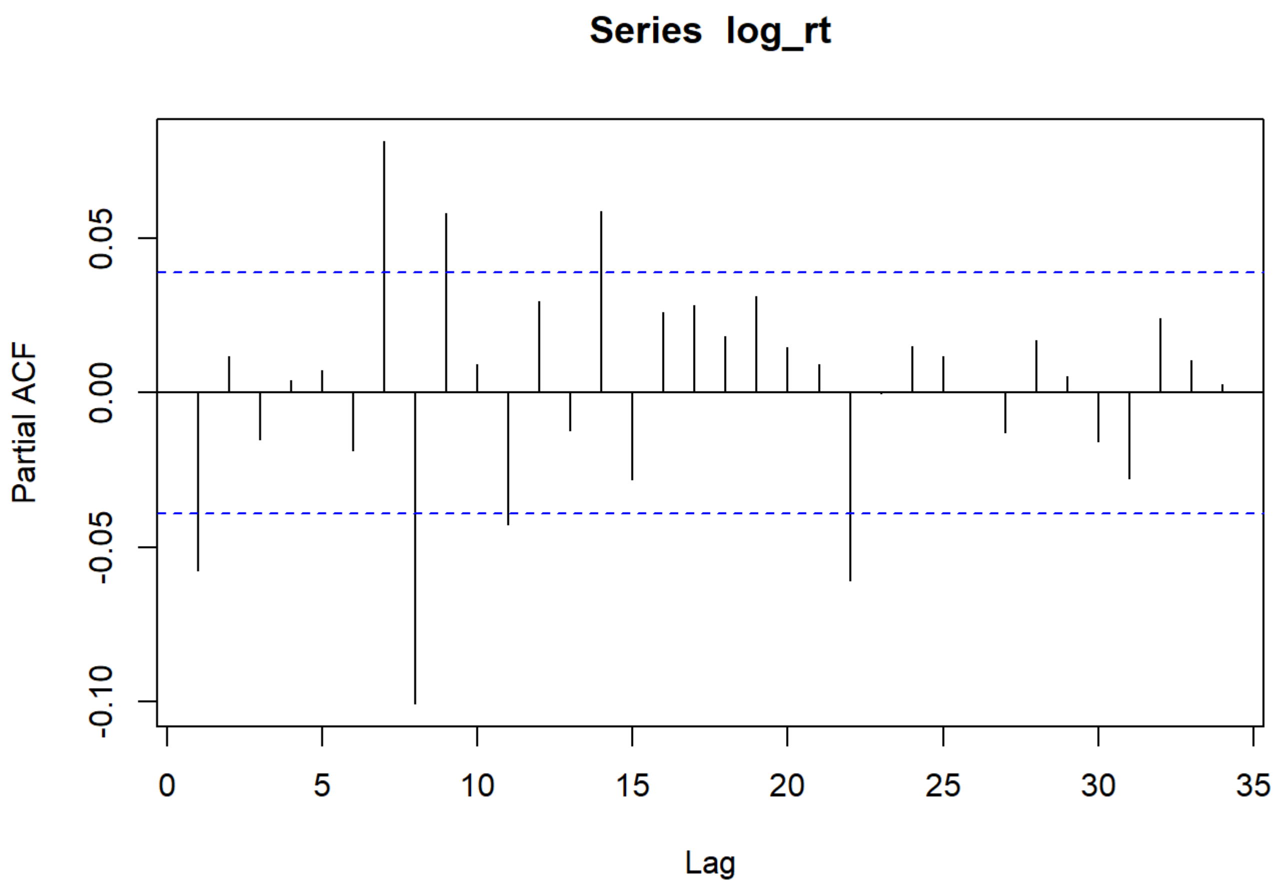
```
##
##  One Sample t-test
##
## data:  log_rt
## t = 2.8581, df = 2515, p-value = 0.004297
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.000309816 0.001664057
## sample estimates:
##    mean of x
## 0.0009869365
```

2a) It appears that the expected value of r_t is different from 0, so we should implement MA in our model.

```
acf(log_rt)
```




```
pacf(log_rt)
```




```
## This section takes a long time to run, so I have commented it out and written the important r
esults down
#lags = 10
#aic_m = matrix(0, nrow=lags, ncol=lags)
#bic_m = matrix(0, nrow=lags, ncol=lags)

#for (i in 1:lags) {
#  for (j in 1:lags) {
#    aic_m[i,j] = AIC(arima(log_rt, order=c((i-1),0,(j-1))))
#    bic_m[i,j] = BIC(arima(log_rt, order=c((i-1),0,(j-1))))
#  }
#}
#which(aic_m==min(aic_m), arr.ind=TRUE) # [10, 7]
#which(bic_m==min(bic_m), arr.ind=TRUE) # [6, 3]
```

2b) There appears to be some serial correlation (and the mean of returns is not 0). The AIC tells me that I should use ARMA(9,6), but the BIC suggests to use ARMA(5,2), which is simpler. So I will use ARMA(5,2).

```
ar = 5
ma = 2
m1 = arima(log_rt, c(ar,0,ma))
m1
```

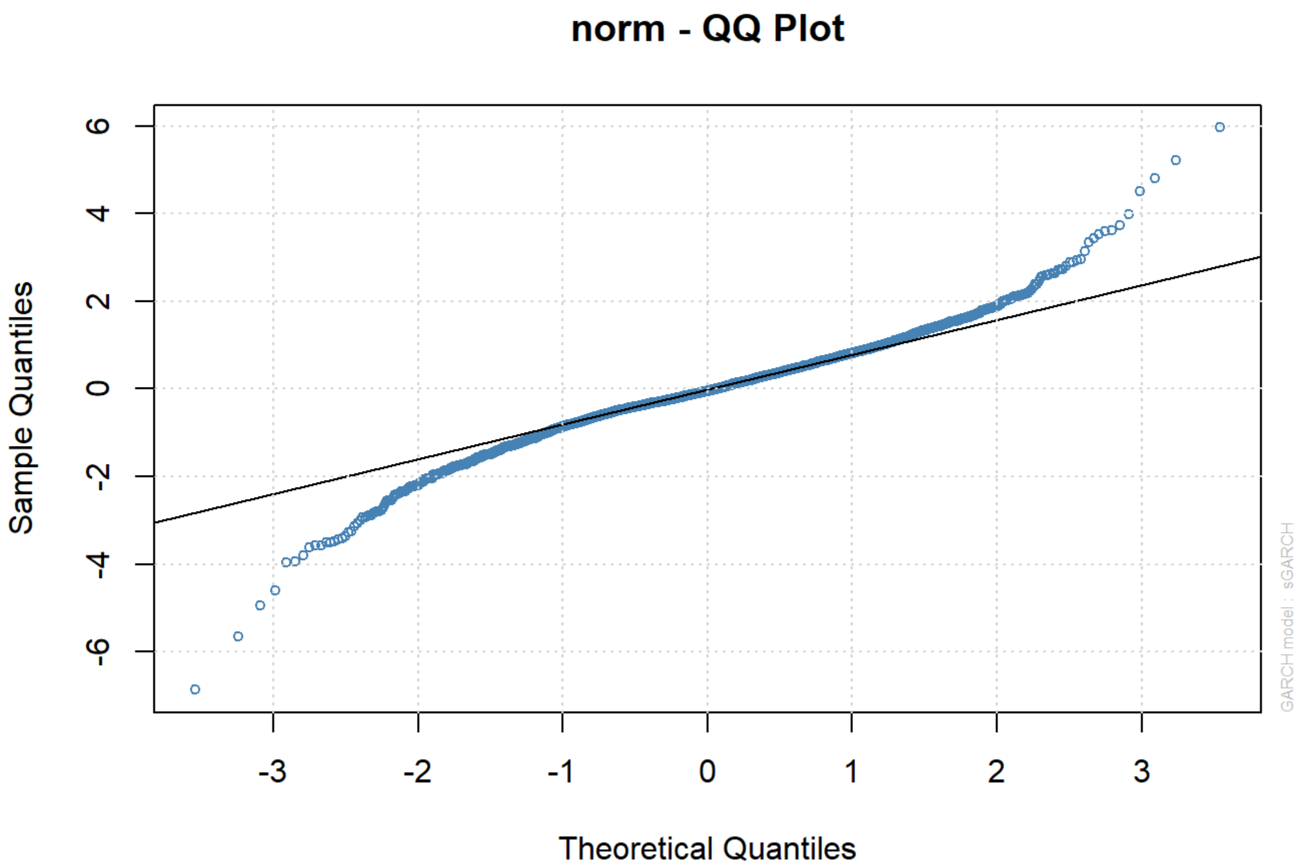
```
##
## Call:
## arima(x = log_rt, order = c(ar, 0, ma))
##
## Coefficients:
##           ar1      ar2      ar3     ar4     ar5     ma1     ma2  intercept
##       -1.6231  -0.8721  -0.0425  0.0116  0.0488  1.5848  0.8122      1e-03
## s.e.   0.0566   0.0592   0.0418  0.0399  0.0239  0.0535  0.0434      3e-04
##
## sigma^2 estimated as 0.0002924:  log likelihood = 6666.51,  aic = -13315.01
```

```
Box.test(m1$residuals^2, lag=8, fitdf=7, type="Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  m1$residuals^2
## X-squared = 496.42, df = 1, p-value < 2.2e-16
```

2c) The result of the Ljung-Box test tells us that we reject $H_0 : \rho_1 = \rho_2 = \ldots\ldots = \rho_8 = 0$, so there appears to still be correlation with the squared term.

```
# using arma(0,0) per TA's recommendation
garch_spec_norm = ugarchspec(
  mean.model = list(armaOrder = c(0,0)),
  variance.model = list(model = "sGARCH", garchOrder = c(1,1)),
  distribution.model = "norm")
m2 = ugarchfit(log_rt, spec=garch_spec_norm)
plot(m2, which=9)
```

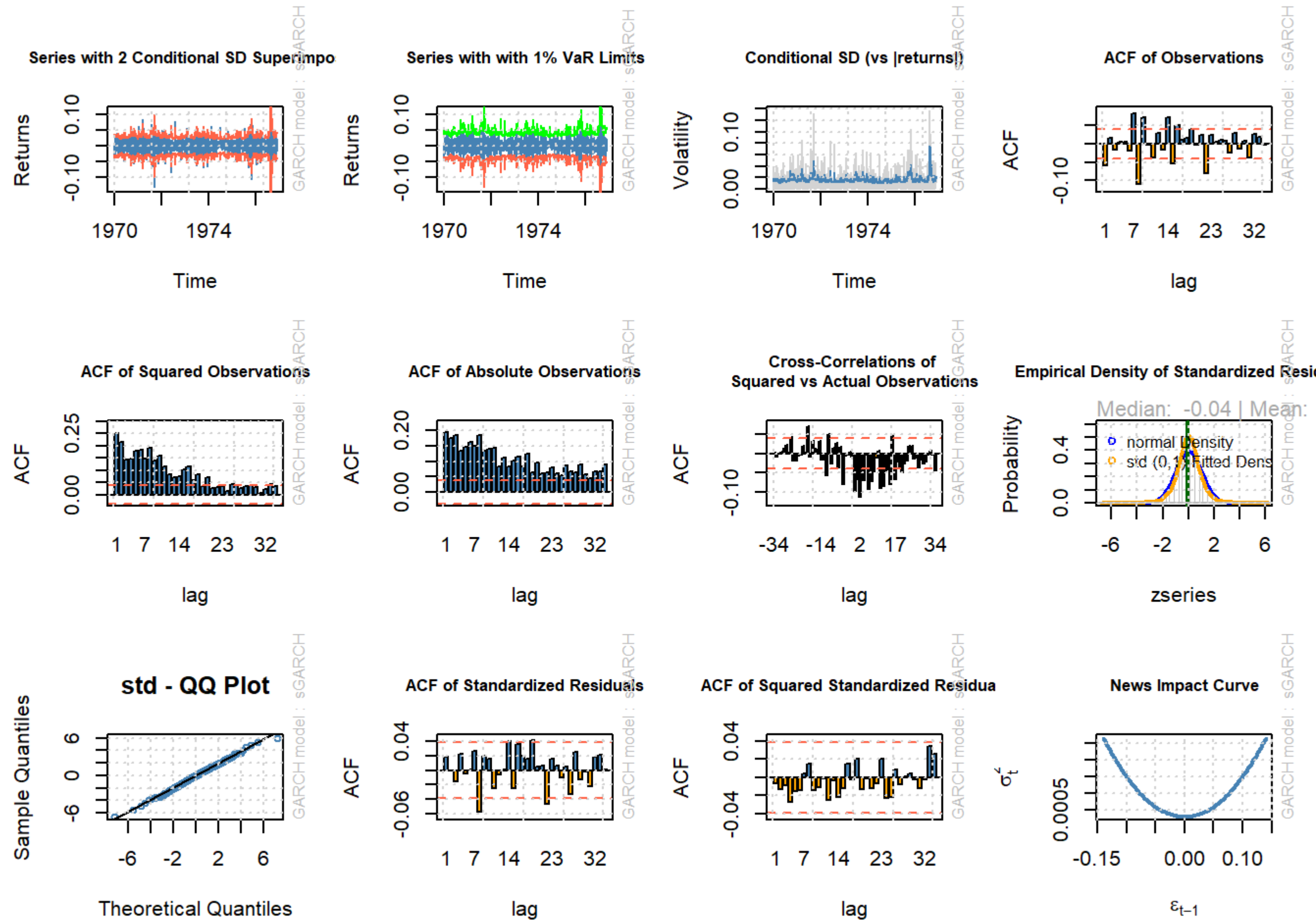


2d)

No, the model is not adequate. We can see from the norm-QQ Plot that the tails of the normal distribution are not heavy enough.

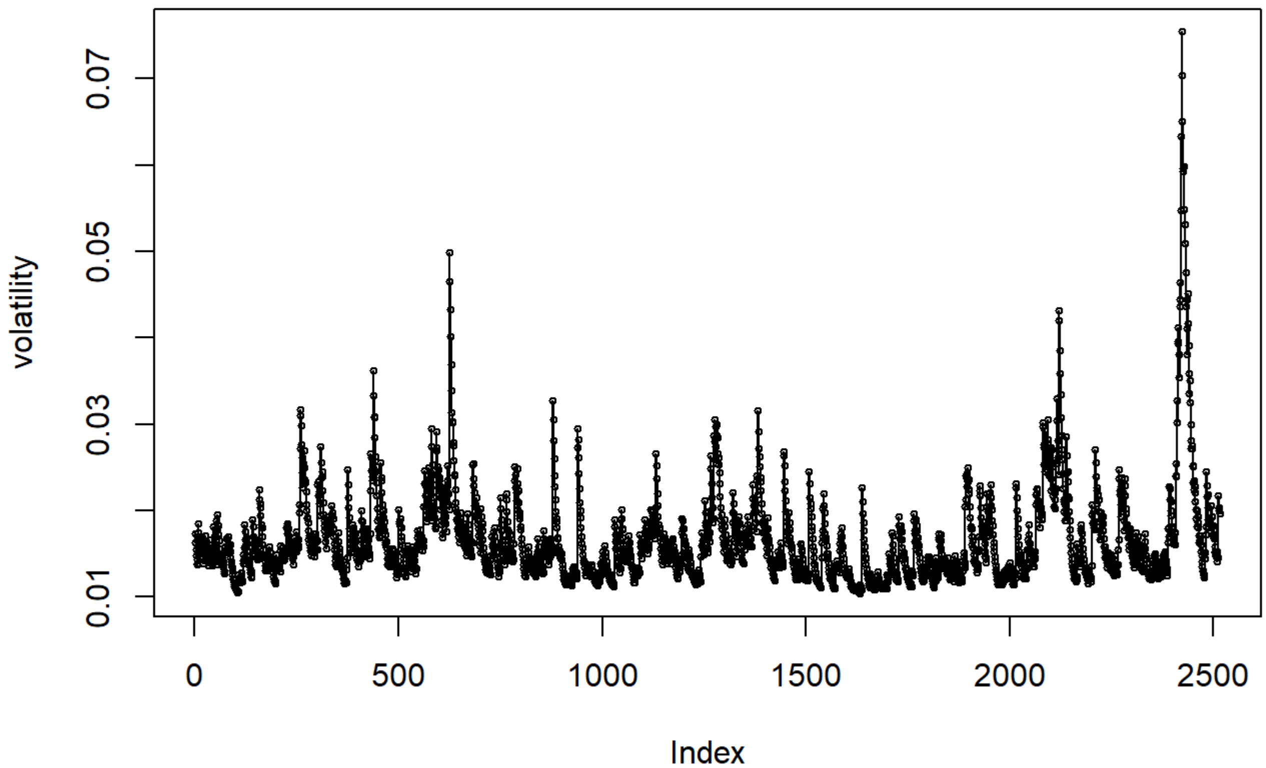
```
garch_spec_std = ugarchspec(
  mean.model = list(armaOrder = c(0,0)),
  variance.model = list(model = "sGARCH", garchOrder = c(1,1)),
  distribution.model = "std")
m3 = ugarchfit(log_rt, spec=garch_spec_std)
plot(m3, which="all")
```

```
##
## please wait...calculating quantiles...
```



2e) It looks like this model is adequate. The QQ Plot looks good, and the ACF plots of the standardized residuals and squared standardized residuals seem to indicate that the model has fit the data relatively well.

```
plot(as.numeric(sigma(m3)), main="", type="o", ylab="volatility", cex=0.5)
```




```
forecast = ugarchforecast(m3, n.ahead = 5)
forecast
```

```
##
## *------------------------------------*
## *       GARCH Model Forecast         *
## *------------------------------------*
## Model: sGARCH
## Horizon: 5
## Roll Steps: 0
## Out of Sample: 0
##
## 0-roll forecast [T0=1976-11-20 18:00:00]:
##       Series   Sigma
## T+1 0.001506 0.01877
## T+2 0.001506 0.01876
## T+3 0.001506 0.01876
## T+4 0.001506 0.01875
## T+5 0.001506 0.01874
```