CSC 425 Time Series Analysis - Homework 5

Akhil Kumar Ramasagaram

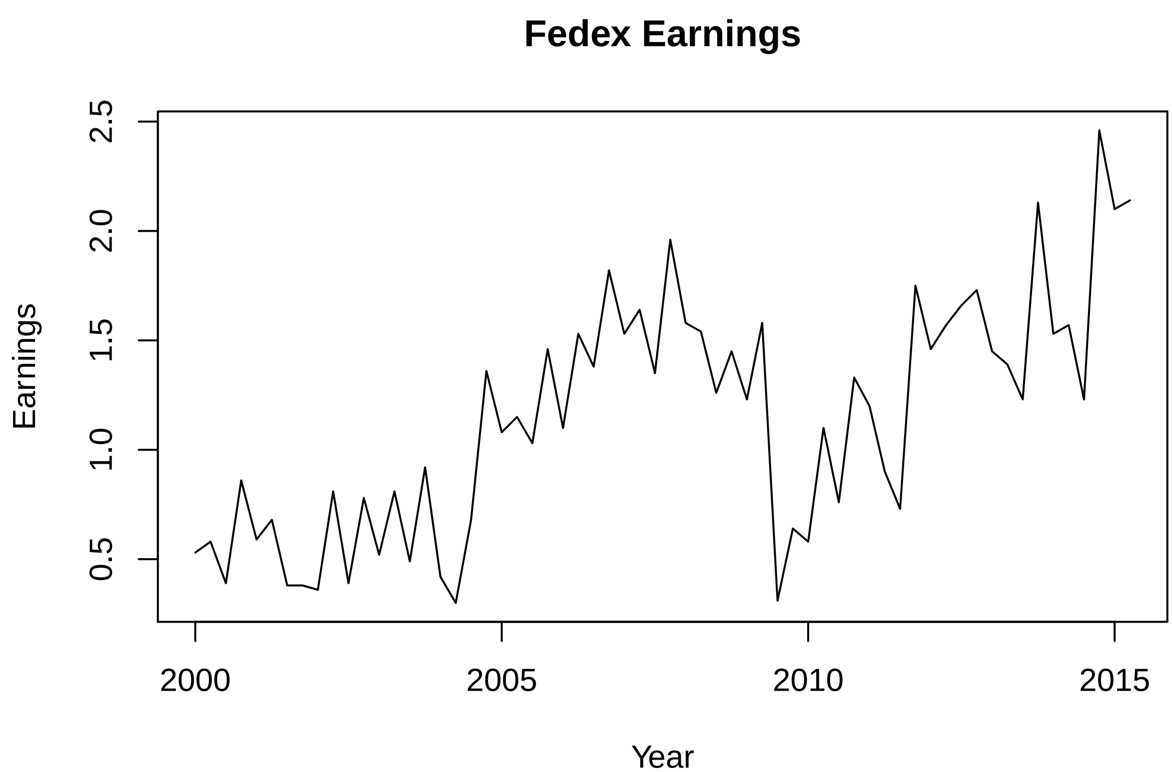
Monday, November 09, 2015

Problem 1

Consider the quarterly earnings per share of the FedEx stock from the first quarter of 2000 to the fourth quarter of 2015 fiscal year (ending in May 2015). The data were obtained from the Fedex website.

a) Import the data and create a time series object

```
fedex=read.table("fedex_earnings2015.txt", header = T)
fedex_ts=ts(fedex$earn,start=c(2000,1), frequency = 4)
plot(fedex_ts, xlab ='Year', ylab = 'Earnings', type ='l', main = "Fedex Earnings")
```

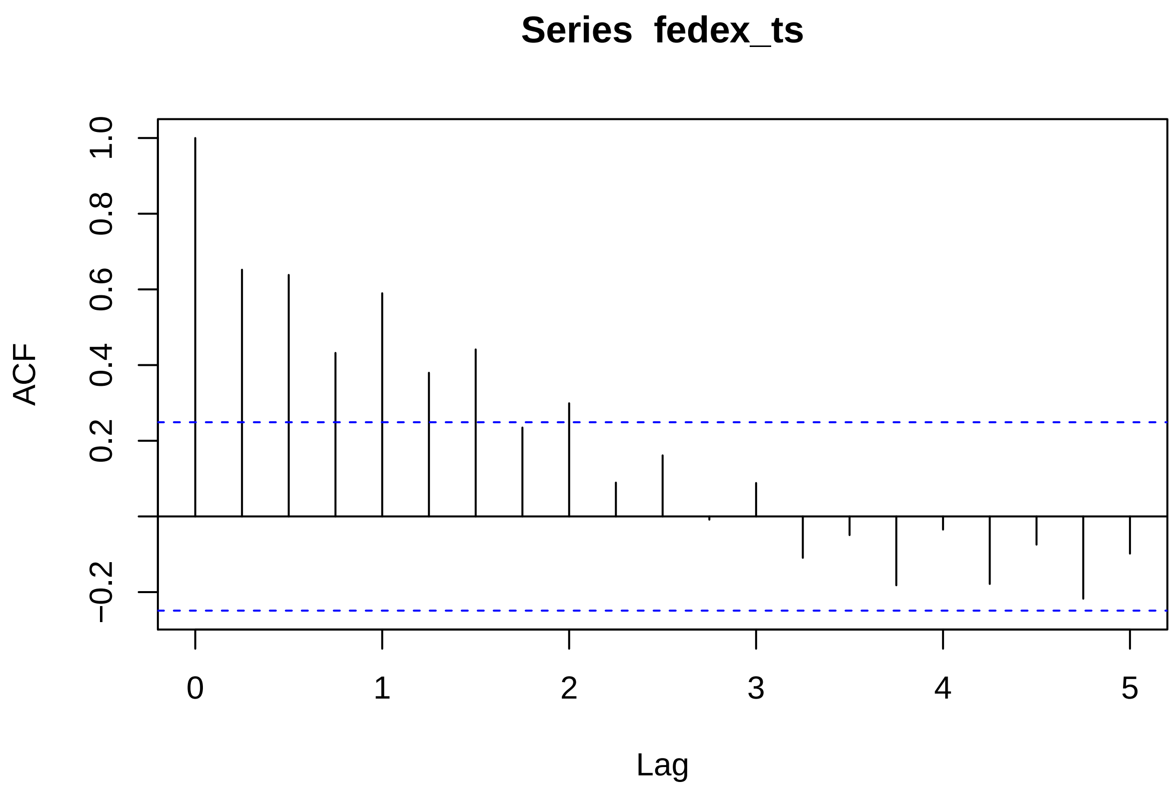


b) Create the time plot for the quarterly earnings, and analyze the trends shown in the plot. Is there any evidence of a seasonal effect in the data? Explain your answer.

Ans: There are two important time periods here, rise at late 2004 & fall at late 2009.

c) Compute the ACF plot of the Fedex earnings time series and discuss the serial correlation in the data shown by the plot.

```r
acf(fedex_ts, lag = 20)
```



It looks like there is a serial correlation from the previous 8 quarters.

d) Compute the Dickey-Fuller test for the quarterly earnings for p=1, 3, and 5. Write down the test statistics and discuss the results of the test.

```r
adfTest(fedex_ts, lags = 1, type = c("c"))
```

```
##
## Title:
##  Augmented Dickey-Fuller Test
##
## Test Results:
##   PARAMETER:
##     Lag Order: 1
##   STATISTIC:
##     Dickey-Fuller: -1.5277
##   P VALUE:
##     0.4899
##
## Description:
##  Mon Nov 09 16:09:44 2015 by user: Akhilkumar
```

```r
adfTest(fedex_ts, lags = 3, type = c("c"))
```

```
##
## Title:
##  Augmented Dickey-Fuller Test
##
## Test Results:
##   PARAMETER:
##     Lag Order: 3
##   STATISTIC:
##     Dickey-Fuller: -0.6019
##   P VALUE:
##     0.8279
##
## Description:
##  Mon Nov 09 16:09:44 2015 by user: Akhilkumar
```

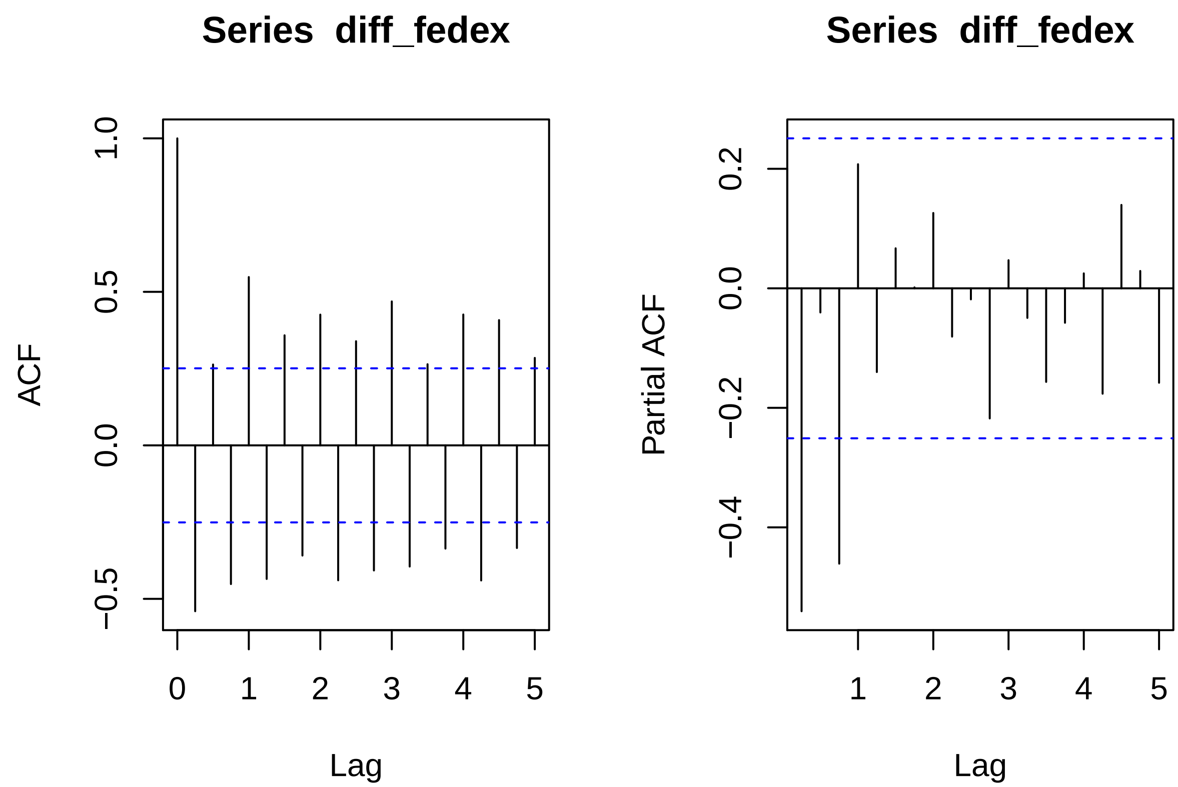
```r
adfTest(fedex_ts, lags = 5, type = c("c"))
```

```
##
## Title:
##  Augmented Dickey-Fuller Test
##
## Test Results:
##   PARAMETER:
##     Lag Order: 5
##   STATISTIC:
##     Dickey-Fuller: -0.798
##   P VALUE:
##     0.7563
##
## Description:
##  Mon Nov 09 16:09:44 2015 by user: Akhilkumar
```

Dickey fuller test confirmsthat there is non stationarity in the unit root as p-values are > 0.05

e) Take the first difference of the Fedex earnings, and analyze its autocorrelation function. Discuss if the ACF plot displays evidence of seasonal behavior.

```r
par(mfrow = c(1,2))
diff_fedex = diff(fedex_ts)
acf(diff_fedex, lag = 20)
pacf(diff_fedex,lag = 20)
```



We can observe seasonality.

f) Use the auto.arima(xts) function to compute the model selected by the BIC criterion.

```
auto.arima(fedex_ts,ic = c('bic'))
```

```
## Series: fedex_ts
## ARIMA(1,0,0)(0,1,1)[4] with drift
##
## Coefficients:
##          ar1     sma1   drift
##       0.7093  -0.9287  0.0213
## s.e.  0.0999   0.2440  0.0061
##
## sigma^2 estimated as 0.07264:  log likelihood=-7.47
## AIC=22.94   AICc=23.7   BIC=31.18
```

```
m1 = Arima(fedex_ts, order = c(1,0,0), seasonal = list(order=c(0,1,1), period = 4), include.drift = T, 
```

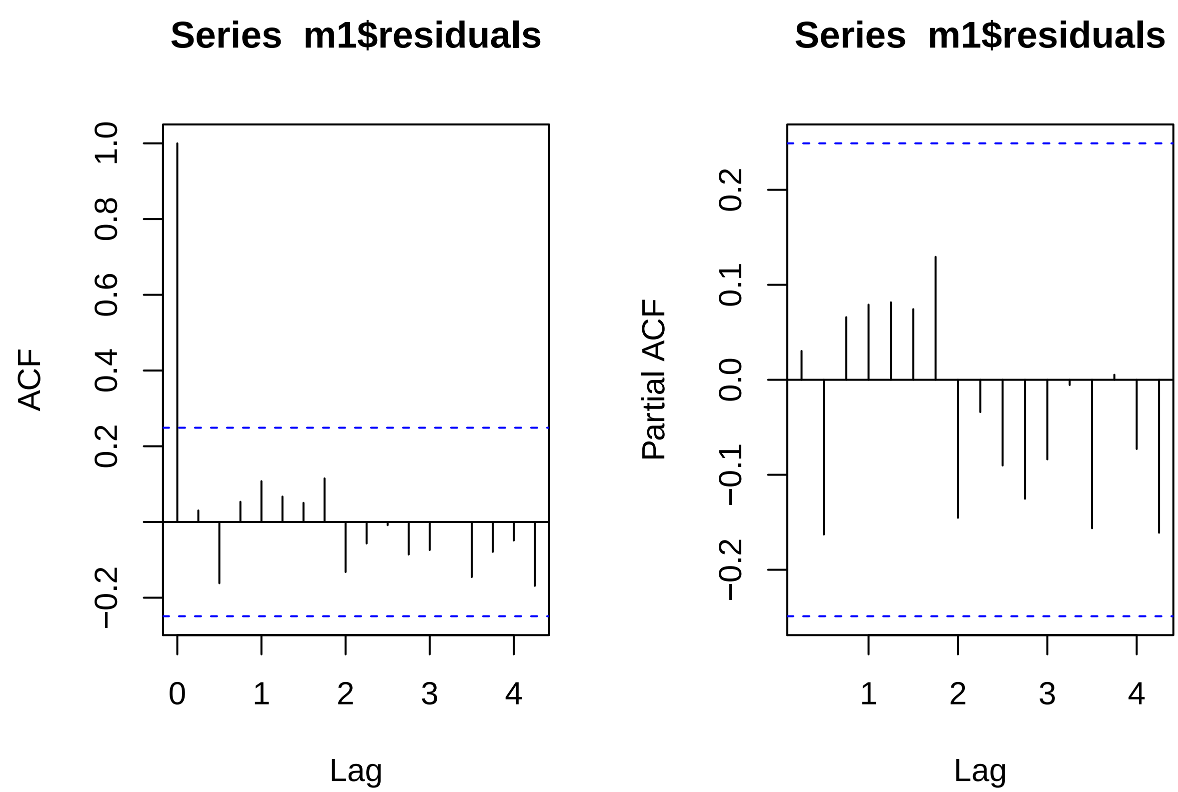
g) Fit the model found by the auto.arima() function to model the Fedex quarterly earnings.

```
coeftest(m1)
```

```
##
```

```
## z test of coefficients:
##
##         Estimate Std. Error z value  Pr(>|z|)
## ar1    0.7092878  0.0998973  7.1002 1.246e-12 ***
## sma1  -0.9286242  0.2435822 -3.8124 0.0001376 ***
## drift  0.0212657  0.0061399  3.4635 0.0005331 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
par(mfrow = c(1,2))
acf(m1$residuals)
pacf(m1$residuals)
```



- All coefficients are significant. - Residual analysis shows that model is a good fit for given time series, since residuals are white noise and follow a normal distribution

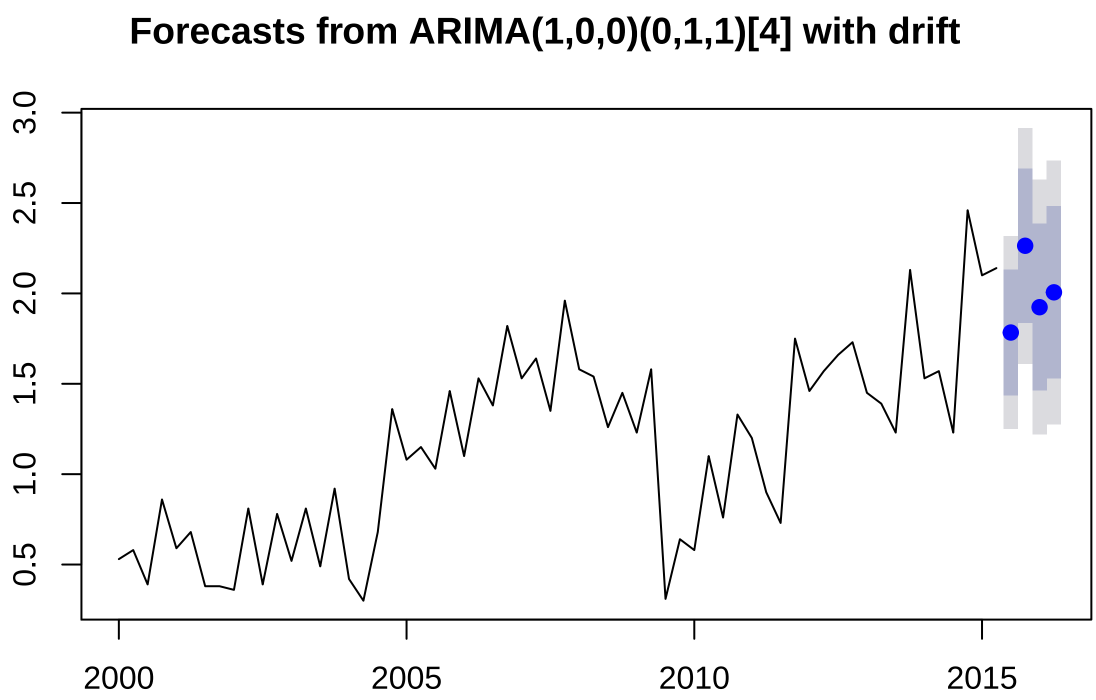
h) Compute the next 4 forecasts for Fedex quarterly earnings.

```r
forecast(m1,h=4)
```

```
##         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 2015 Q3       1.783437 1.435230 2.131644 1.250900 2.315974
## 2015 Q4       2.263633 1.836955 2.690311 1.611085 2.916181
## 2016 Q1       1.923727 1.463009 2.384445 1.219119 2.628335
## 2016 Q2       2.005964 1.528930 2.482998 1.276403 2.735525
```

i) Create a plot of forecasts and discuss if forecasts are consistent with the observed trend.

```
par(mfrow = c(1,1))
plot(forecast(m1,h=4))
```

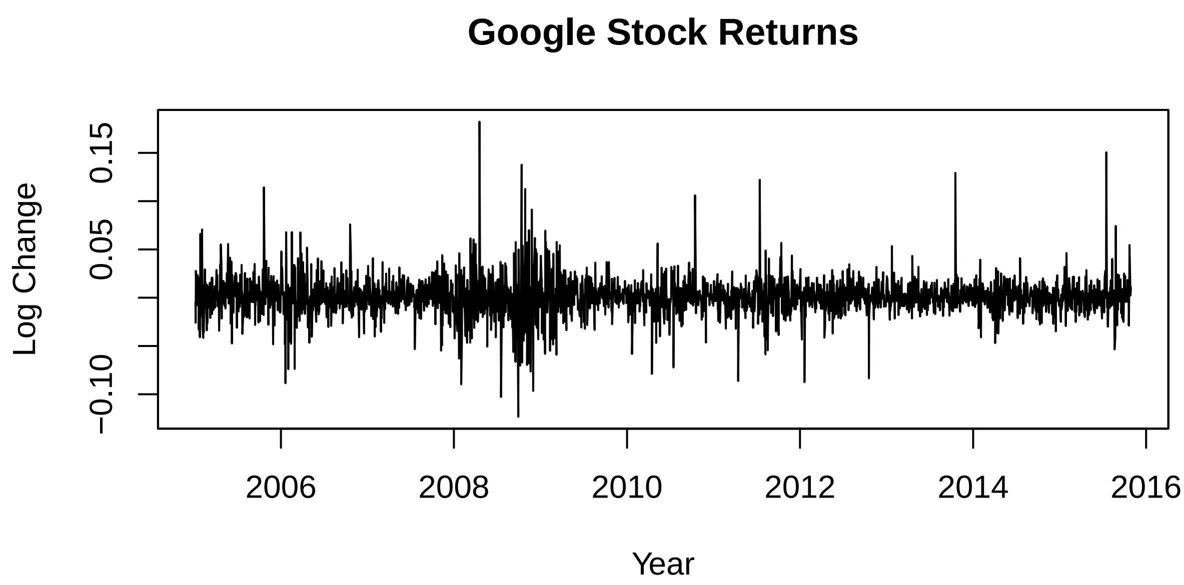


Problem 2

Use the datafile google.csv that contains the Google (AMZN) stock daily returns from 1/4/2005 to 10/28/2015. The data file contains dates (Date), Prices (price). You can use the code and the analysis of the IBM returns used in week 8 lectures as your reference for the analysis of this data. Analyze the stock log returns following the steps below.

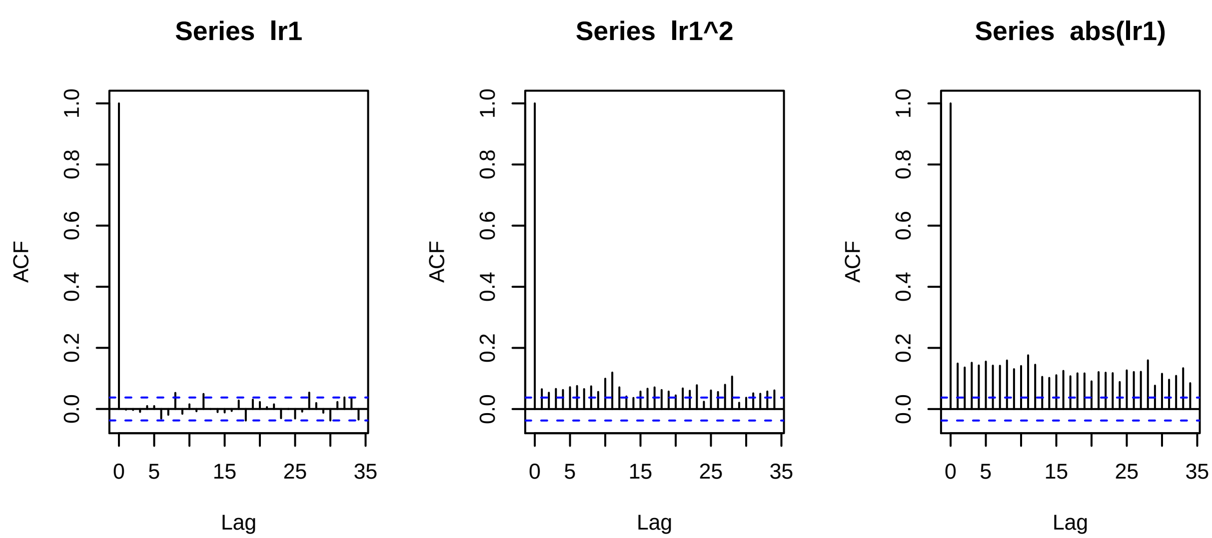
1) Create the time plot for the stock returns and analyze the time plot.

```
google= read.table('google.csv', header=T, sep=',')
google_ts = zoo(google$Price, as.Date(as.character(google$Date), format=c("%m/%d/%Y")))
lr = log(google_ts/lag(google_ts, -1))
plot(lr, xlab = "Year", ylab = "Log Change", main = "Google Stock Returns")
```



2 & 3) Is there evidence of serial correlations for the stock returns? Use autocorrelations and 5% significance level to answer the question.

```
par(mfrow=c(1,3))
lr1 <- coredata(lr)
acf(lr1)
acf(lr1^2)
acf(abs(lr1))
```



The ACF plots show that the log returns of google stock are not correlated, indicating a constant mean. Both the square root and the absolute returns time series show large autocorrelations. We conclude that the log returns process has a strong non-linear dependence. This is called GARCH effect, where the volatility is non constant and is affected by past shocks.

```r
Box.test(coredata(lr^2),lag=6,type='Ljung')
```

```
##
##  Box-Ljung test
##
## data:  coredata(lr^2)
## X-squared = 71.1488, df = 6, p-value = 2.376e-13
```

```r
Box.test(coredata(lr^2),lag=12,type='Ljung')
```

```
##
##  Box-Ljung test
##
## data:  coredata(lr^2)
## X-squared = 186.3212, df = 12, p-value < 2.2e-16
```

Based on the p-value we can conclude that the squared returns are autocorrelated.

4) Fit an AR(0)- GARCH(1,1) model to model the return volatility using a t- distribution for the error terms.

```r
g1 = ugarchspec(variance.model = list(garchOrder = c(1,1)), mean.model = list(armaOrder = c(0,0)), dist
g1 = ugarchfit(spec = g1, data=lr)
g1
```

```
##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : sGARCH(1,1)
## Mean Model   : ARFIMA(0,0,0)
## Distribution : std
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      0.000884    0.000267   3.3115 0.000928
## omega   0.000003    0.000001   2.4396 0.014702
## alpha1  0.030195    0.003162   9.5501 0.000000
## beta1   0.961009    0.002445 393.0284 0.000000
## shape   4.029353    0.326314  12.3481 0.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error   t value Pr(>|t|)
## mu      0.000884    0.000259   3.40705 0.000657
## omega   0.000003    0.000003   0.92435 0.355303
## alpha1  0.030195    0.003243   9.31103 0.000000
## beta1   0.961009    0.002956 325.15839 0.000000
```

```
## shape   4.029353    0.486720   8.27859 0.000000
##
## LogLikelihood : 7338.955
##
## Information Criteria
## ------------------------------------
##
## Akaike       -5.3847
## Bayes        -5.3738
## Shibata      -5.3847
## Hannan-Quinn -5.3808
##
## Weighted Ljung-Box Test on Standardized Residuals
## ------------------------------------
##                         statistic p-value
## Lag[1]                      1.827  0.1765
## Lag[2*(p+q)+(p+q)-1][2]     1.914  0.2774
## Lag[4*(p+q)+(p+q)-1][5]     2.300  0.5495
## d.o.f=0
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ------------------------------------
##                         statistic p-value
## Lag[1]                     0.1852  0.6669
## Lag[2*(p+q)+(p+q)-1][5]    0.2890  0.9848
## Lag[4*(p+q)+(p+q)-1][9]    0.5304  0.9981
## d.o.f=2
##
## Weighted ARCH LM Tests
## ------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[3]   0.02546 0.500 2.000  0.8732
## ARCH Lag[5]   0.03951 1.440 1.667  0.9965
## ARCH Lag[7]   0.18099 2.315 1.543  0.9976
##
## Nyblom stability test
## ------------------------------------
## Joint Statistic:  41.5254
## Individual Statistics:
## mu     0.07748
## omega  3.57444
## alpha1 1.16958
## beta1  0.62582
## shape  0.78465
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:          1.28 1.47 1.88
## Individual Statistic:     0.35 0.47 0.75
##
## Sign Bias Test
## ------------------------------------
##                    t-value   prob sig
## Sign Bias           0.2686 0.7883
```

```
## Negative Sign Bias  1.0633 0.2877
## Positive Sign Bias  0.4580 0.6470
## Joint Effect        1.9191 0.5894
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## ------------------------------------
##   group statistic p-value(g-1)
## 1    20     17.17       0.5780
## 2    30     35.54       0.1874
## 3    40     37.44       0.5411
## 4    50     64.25       0.0707
##
##
## Elapsed time : 0.7504978
```

- The model can be writtern as $rt = 0.0009 + at$, $at = \sigma(t)e(t)$

$$\sigma^2(t) = 0.000003 + 0.03a^2(t-1) + 0.96\sigma^2(t-1)$$

- Since the sum of aplha and beta is less than 1, it is GARCH model.
- The persistence for this time series is 0.9912035 which tells the shock is not permanent. The halflife value for this time series is 78.4507226 which tells us that it takes almost 79 observations to recover from the shock.
- The ACF plot shows there is white noise in the time series. [plot]:C:/Users/Akhilkumar/Desktop/depaul/CSC 425/hw5.png

5 & 6) Fit and EGARCH(1,1) model for the log returns using a t- distribution for the error terms. Perform model checking and write down the fitted model.

```r
e1=ugarchspec(variance.model=list(model = "eGARCH", garchOrder=c(1,1)), mean.model=list(armaOrder=c(0,0
e1=ugarchfit(spec=e1, data=lr)
e1
```

```
##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : eGARCH(1,1)
## Mean Model   : ARFIMA(0,0,0)
## Distribution : std
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error   t value Pr(>|t|)
## mu      0.000771    0.000295    2.6167 0.008879
## omega  -0.075455    0.003907  -19.3121 0.000000
## alpha1 -0.039321    0.009427   -4.1709 0.000030
## beta1   0.990750    0.000462 2145.6244 0.000000
## gamma1  0.100730    0.015274    6.5948 0.000000
```

```
## shape   4.136856    0.334265   12.3760 0.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error   t value Pr(>|t|)
## mu      0.000771    0.000328    2.3533 0.018608
## omega  -0.075455    0.002393  -31.5378 0.000000
## alpha1 -0.039321    0.011524   -3.4119 0.000645
## beta1   0.990750    0.000251 3946.5142 0.000000
## gamma1  0.100730    0.018307    5.5023 0.000000
## shape   4.136856    0.376319   10.9929 0.000000
##
## LogLikelihood : 7362.719
##
## Information Criteria
## ------------------------------------
##
## Akaike       -5.4014
## Bayes        -5.3884
## Shibata      -5.4014
## Hannan-Quinn -5.3967
##
## Weighted Ljung-Box Test on Standardized Residuals
## ------------------------------------
##                         statistic p-value
## Lag[1]                      3.158 0.07554
## Lag[2*(p+q)+(p+q)-1][2]     3.275 0.11820
## Lag[4*(p+q)+(p+q)-1][5]     3.688 0.29569
## d.o.f=0
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ------------------------------------
##                         statistic p-value
## Lag[1]                     0.3368  0.5617
## Lag[2*(p+q)+(p+q)-1][5]    0.4387  0.9667
## Lag[4*(p+q)+(p+q)-1][9]    0.7370  0.9946
## d.o.f=2
##
## Weighted ARCH LM Tests
## ------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[3] 6.789e-05 0.500 2.000  0.9934
## ARCH Lag[5] 2.363e-02 1.440 1.667  0.9983
## ARCH Lag[7] 2.334e-01 2.315 1.543  0.9958
##
## Nyblom stability test
## ------------------------------------
## Joint Statistic:  1.8046
## Individual Statistics:
## mu     0.2922
## omega  0.5773
## alpha1 0.5131
## beta1  0.5569
## gamma1 0.4978
```

```
## shape  0.3075
##
## 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           0.4205 0.6741
## Negative Sign Bias  0.3252 0.7451
## Positive Sign Bias  0.6014 0.5476
## Joint Effect        0.6283 0.8899
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## ------------------------------------
##   group statistic p-value(g-1)
## 1    20     18.98       0.4581
## 2    30     35.54       0.1874
## 3    40     37.79       0.5249
## 4    50     48.06       0.5111
##
##
## Elapsed time : 1.541025
```

Now since alpha < 1, there is significant leverage effect.