STATS 509 HOMEWORK 9

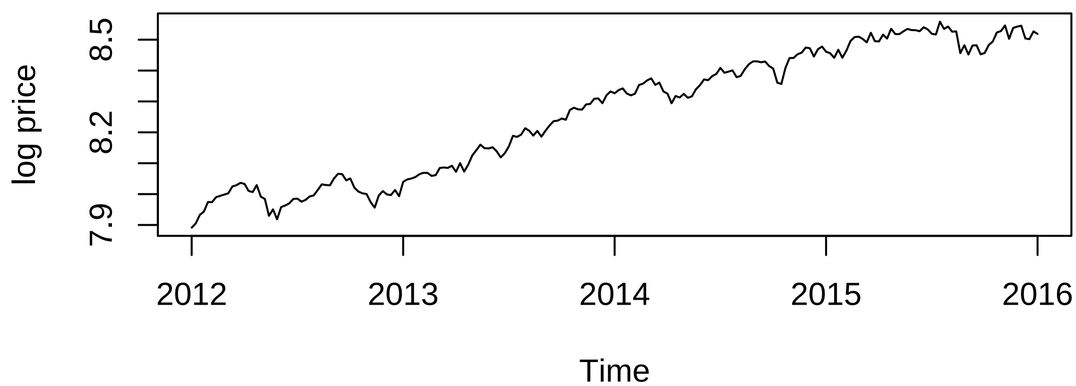
Yuan Yin

3/31/20181

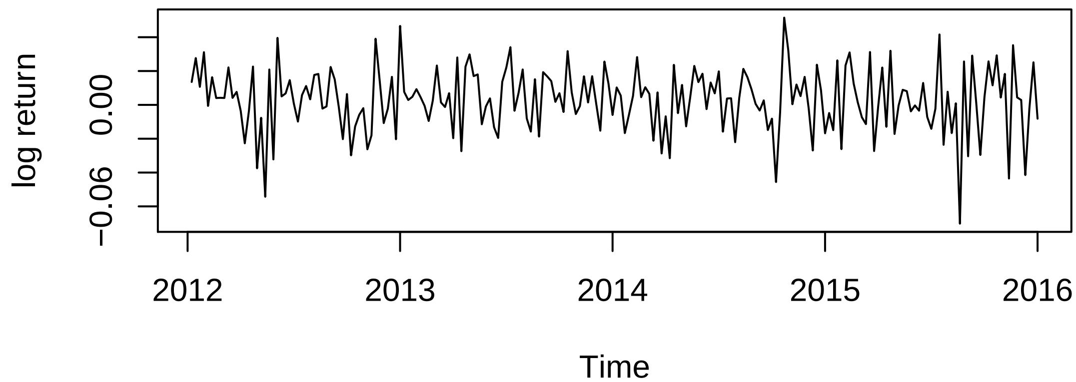
Question 1

(a)

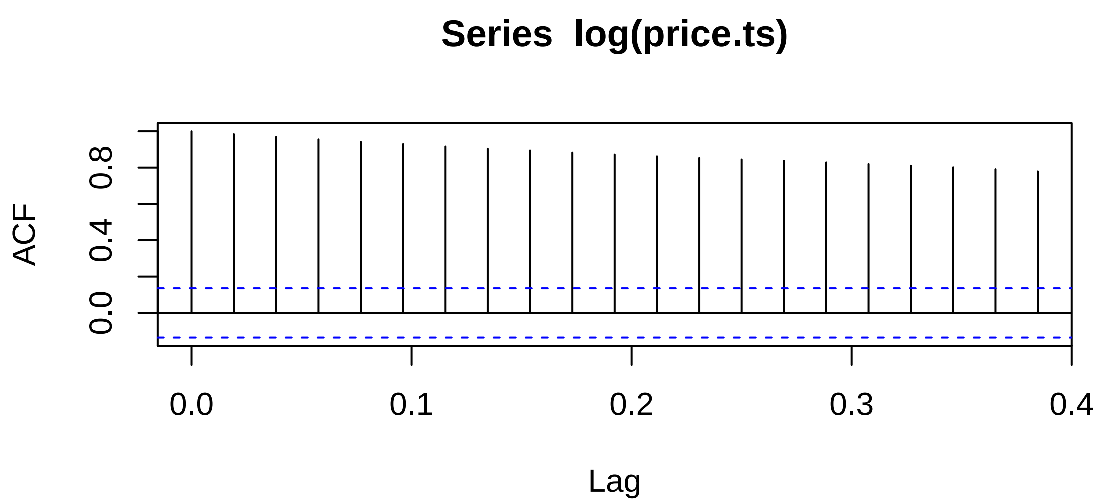
```
data = read.csv("NASDAQ_Wkly_Jan1_2012_Dec31_2015.csv", header = T)
price = rev(data$Adj.Close)
price.ts = ts(data = price, start = c(2012,1), frequency =52)
l_return = diff(log(price.ts))
plot(log(price.ts), type = "l", xlab = "Time", ylab = "log price")
```



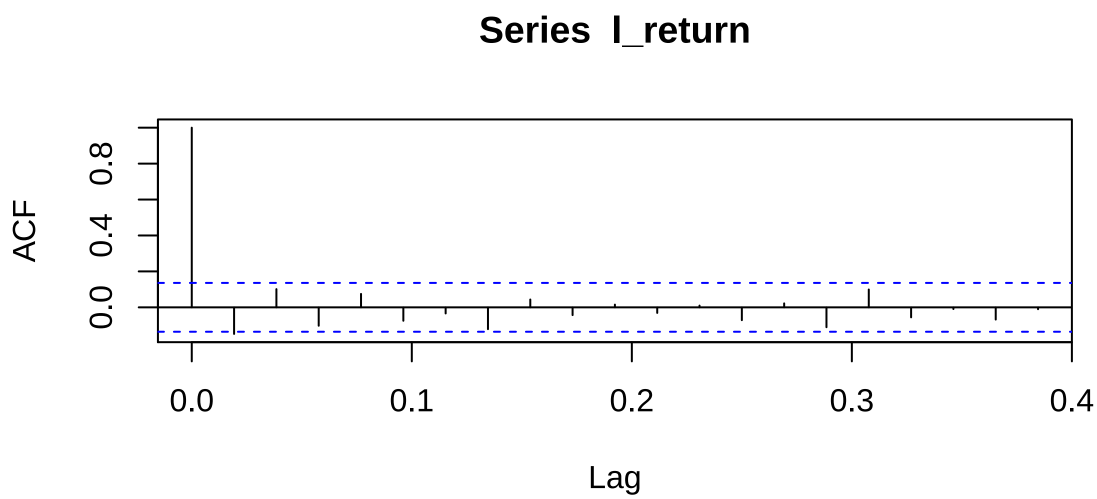

```
plot(l_return, type = "l", xlab = "Time", ylab = "log return")
```




```r
acf(log(price.ts), lag = 20, type = c("correlation"))
```




```r
acf(l_return, lag = 20, type = c("correlation"))
```



- Seeing that for log-adjusted closing price, there is an obvious trend of the data, it increases gradually with time goes by. It's not a stationary process. However, after with plot data of log return, the data seems to be weakly stationary but still further check.
- Autocorrelation plots show that there are strong correlations between log-price, indicating that it's non stationary process, but seems no correlation between log-return.

(b)

```
library(forecast)
```

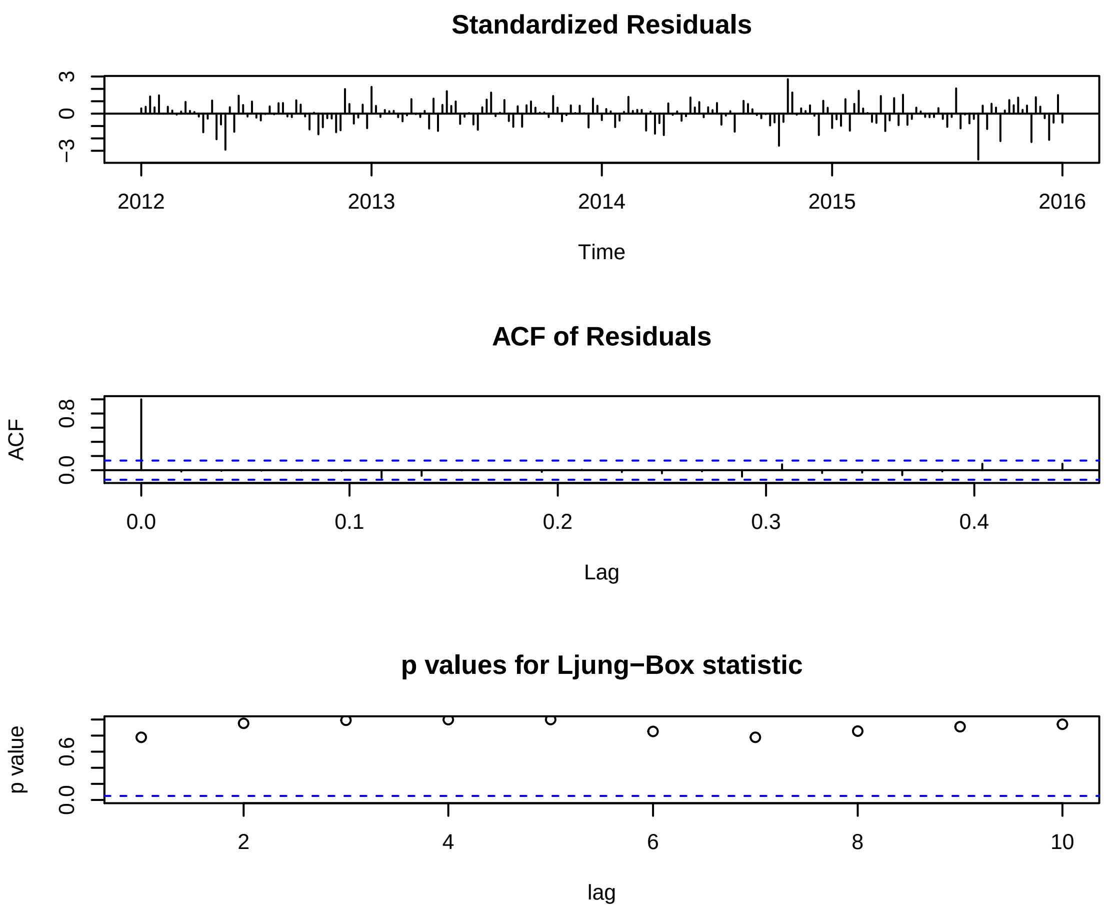
```
## Warning: package 'forecast' was built under R version 3.4.4
```

```
auto.arima(log(price.ts), max.p = 5, max.q = 5, seasonal = FALSE, stepwise = FALSE, ic = c("aic"))
```

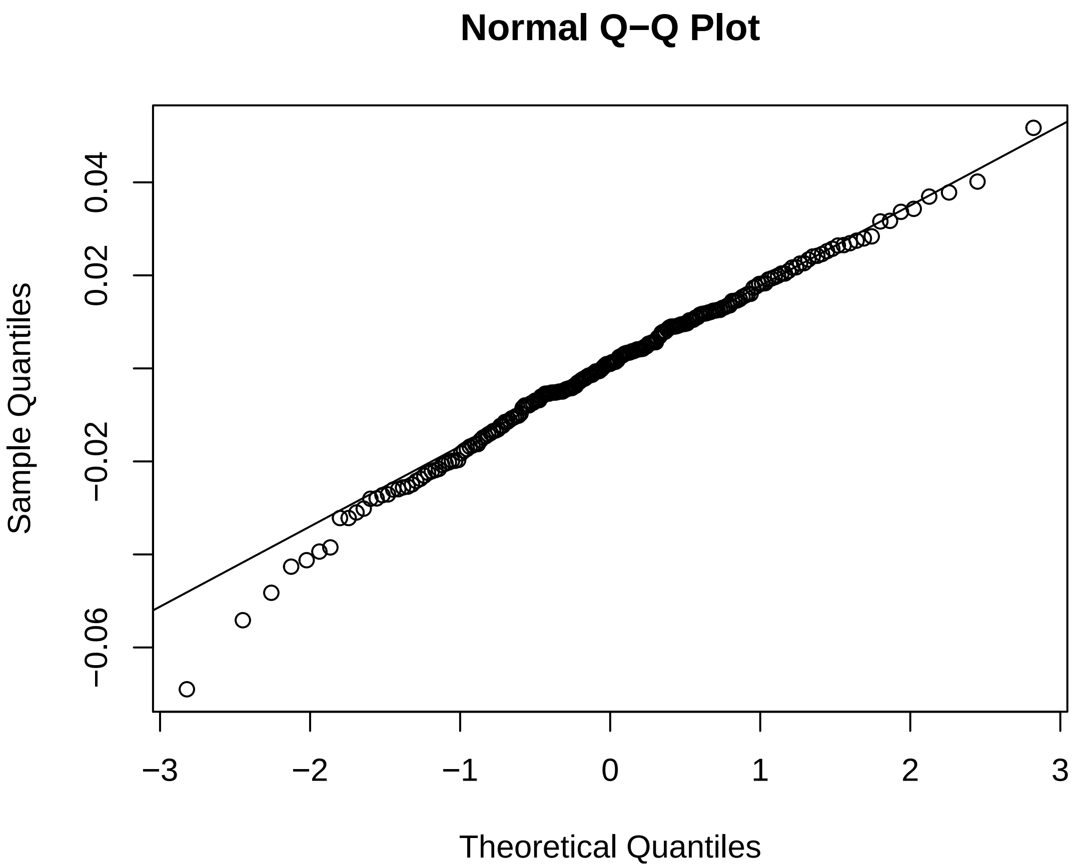
```
## Series: log(price.ts)
## ARIMA(1,1,1) with drift
##
## Coefficients:
##           ar1     ma1   drift
##       -0.8588  0.7577  0.0030
## s.e.   0.1724  0.2248  0.0012
##
## sigma^2 estimated as 0.0003453:  log likelihood=535.41
## AIC=-1062.82   AICc=-1062.62   BIC=-1049.47
```

- The model we pick by AIC criteria is ARIMA(1,1,1), which means that first order difference is enough to make the data stationary. This result is corresponding with last problem where we found log return shows stationary. Also we see that the absolute coefficient of AR term is less than 1 and MA term is 1, indicates that the differenced data is only controled by residuals and is always stationary.

```
fit = auto.arima(log(price.ts), max.p = 5, max.q = 5, seasonal = FALSE, stepwise = FALSE)
tsdiag(fit)
```



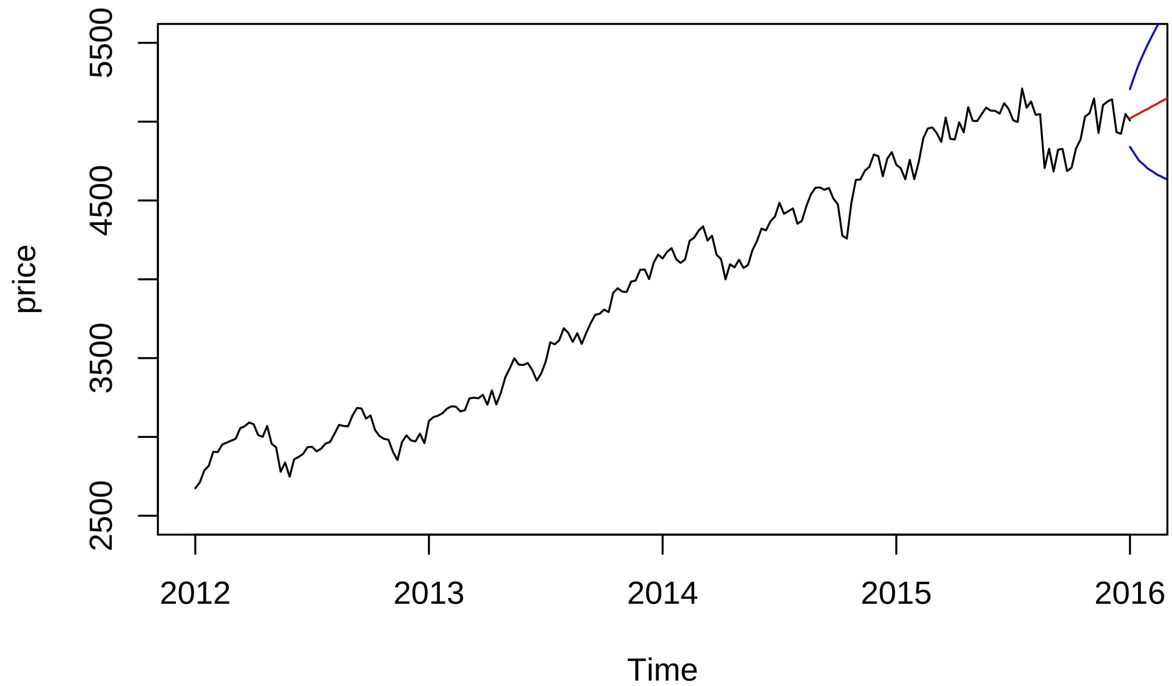

```
qqnorm(fit$resid); qqline(fit$resid)
```



- Diagnostic for residuals is as above. Seeing that the residuals look like a white noise process. The residuals seem mean and variance stationary. Also ACF plot shows there is no autocorrelation for residuals. What's more, Ljung-Box test indicates that there is no correlations between residuals.
- QQ-plot shows that for right tails, it looks like normal distribution very much. But there is heavier tails than normal distribution for loss part.

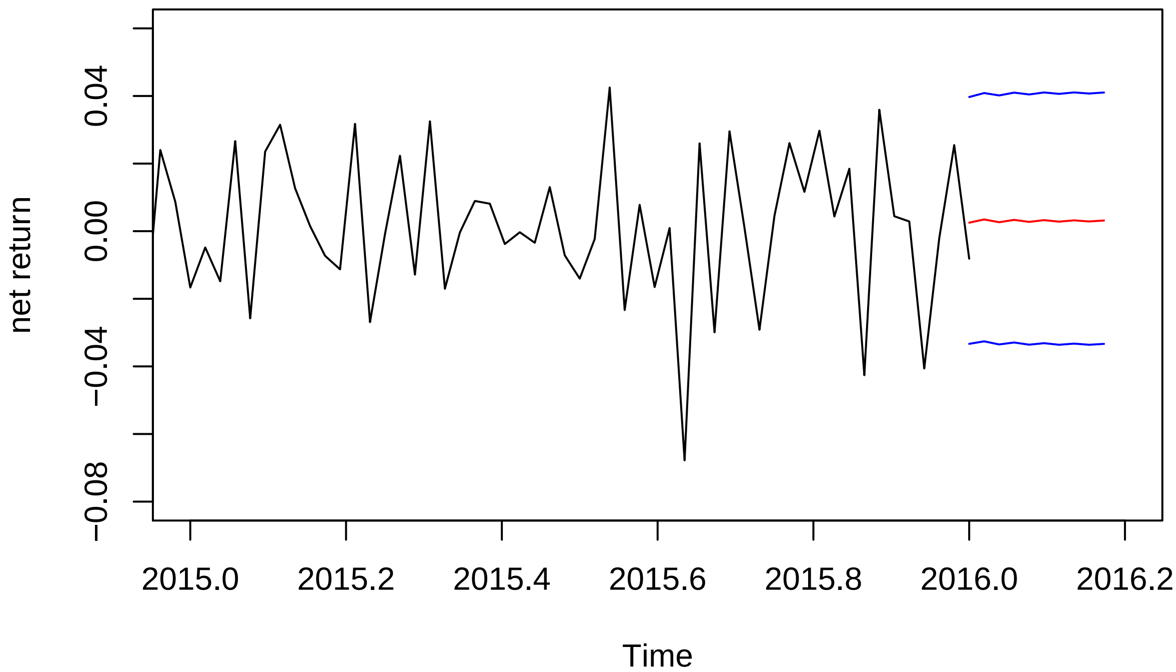
(c)

```
forecasts = forecast(fit, h = 10)
plot(price.ts, type = "l", xlab = "Time", ylab = "price", xlim = c(2012,2016), ylim = c(2500,5500))
lines(seq(from = 2016, by = 1/52, length = 10), exp(forecasts$mean), col = "red")
lines(seq(from = 2016, by = 1/52, length = 10), exp(forecasts$upper)[,2], col = "blue")
lines(seq(from = 2016, by = 1/52, length = 10), exp(forecasts$lower)[,2], col = "blue")
```



(d)

```
n_return = exp(l_return) - 1
fit2 = auto.arima(l_return, max.p = 5, max.q = 5, seasonal = FALSE, stepwise = FALSE)
forecasts2 = forecast(fit2, h = 10)
plot(n_return, type = "l", xlab = "Time", ylab = "net return",
     xlim = c(2015,2016.2), ylim = c(-0.08,0.06))
lines(seq(from = 2016, by = 1/52, length = 10),
      exp(forecasts2$mean)-1, col = "red")
lines(seq(from = 2016, by = 1/52, length = 10),
      exp(forecasts2$upper[,2])-1, col = "blue")
lines(seq(from = 2016, by = 1/52, length = 10),
      exp(forecasts2$lower[,2])-1, col = "blue")
```



(e)

```
N = 100000
set.seed(000)
l_return2 = rep(0,N)
for (i in 1:N){
  l_price.sim = arima.sim(list(order = c(1,1,1), ar = -0.8588, ma = 0.7577),
                          n = 2, sd = sqrt(0.0003453)) + 0.003
  l_return2[i] = l_price.sim[3]-l_price.sim[2]
}
VaR = -(exp(quantile(l_return2, 0.005))-1)
```

- Finding that the relative VaR is 0.0476 at $\alpha = 0.005$. But from the diagnostic in (b), we know that left part of residuals is not normal distribution, but a heavier tail than normal distribution. Thus the true value of VaR should be larger than our result.