Homework 1

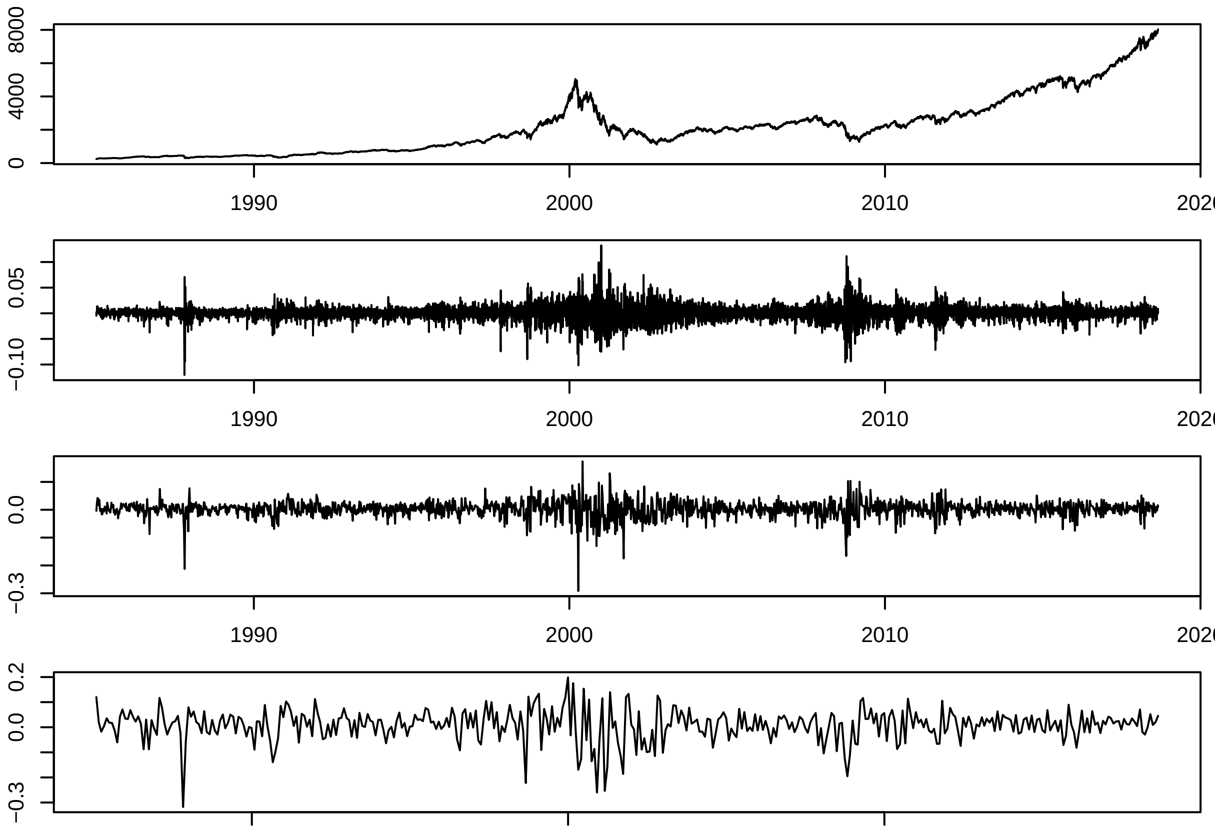
Joyce Yu Cahoon

1.1

Download the daily, weekly and monthly prices for the Nasdaq index and the IBM stock from Yahoo. Reproduce figures 1.3-8 using the Nasdaq index and the IBM stock data.

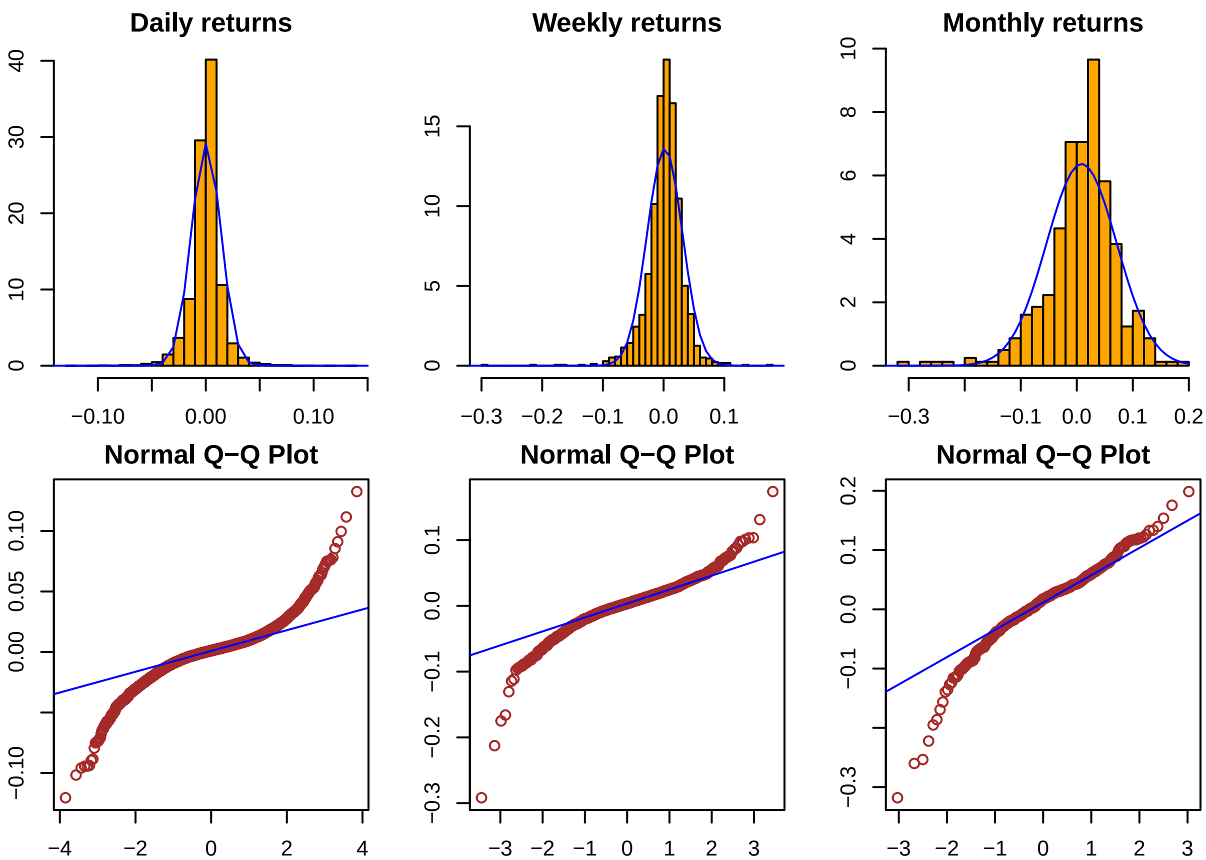
For NASDAQ:

```
# Read in data
IXIC <- readRDS("~/workspace/st790-financial-stats/data/ixic.rds")
IBM <- readRDS("~/workspace/st790-financial-stats/data/ibm.rds")
# Extract
daily <- log(dailyReturn(IXIC)+1)
weekly <- log(weeklyReturn(IXIC)+1)
monthly <- log(monthlyReturn(IXIC)+1)
# Plot
par(mar=c(1,3,2,1),mfrow=c(4,1))
plot(index(IXIC), as.numeric(IXIC$IXIC.Close), type = "l",
     ylab = "daily price")
plot(index(daily), daily, type = "l",
     ylab="daily log return")
plot(index(weekly), weekly, type = "l",
     ylab="weekly log return")
plot(index(monthly), monthly, type = "l",
     ylab="monthly log return")
```



Histograms for NASDAQ:

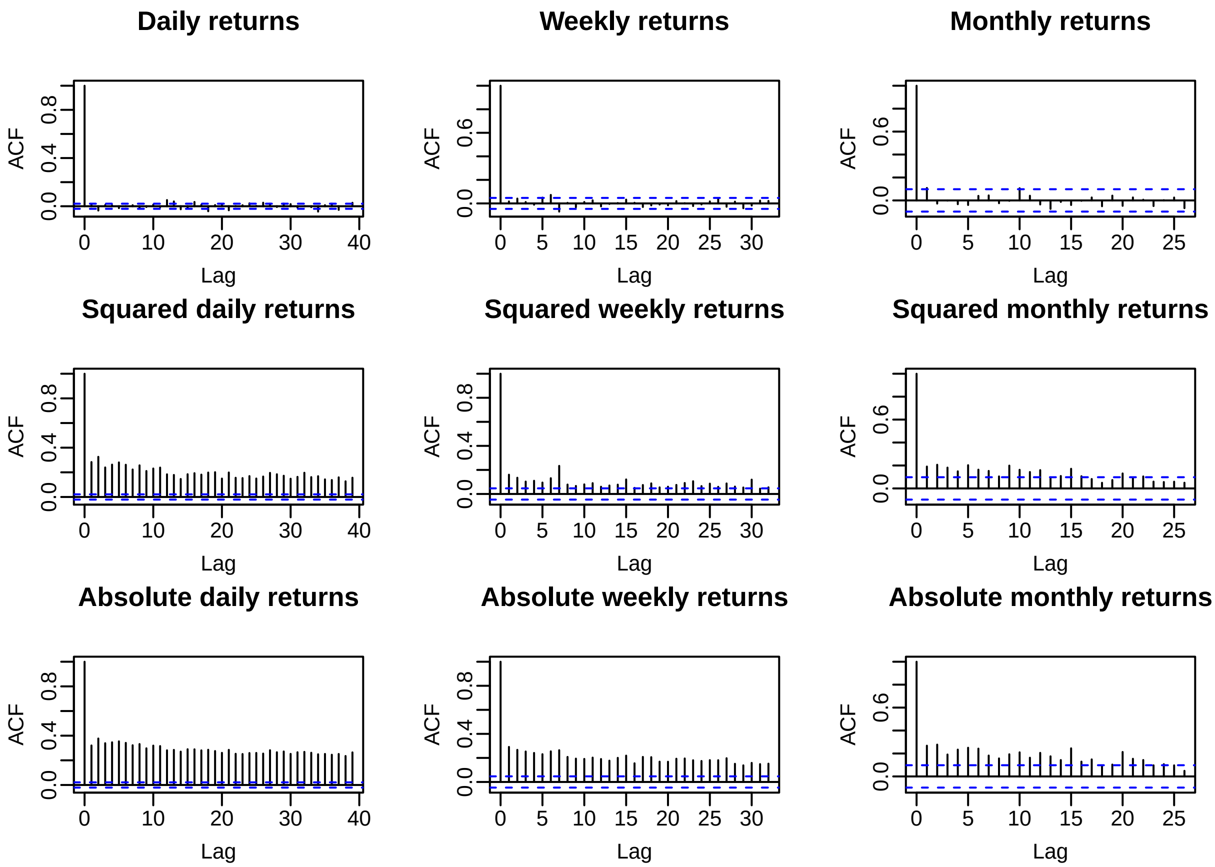
```
par(mar=c(2,3,2,1),mfrow=c(2,3))
hist(daily, probability=T, nclass=35, xlab=’’,
col="orange",main="Daily returns")
x <- seq(-0.5, 0.2, 0.01)
lines(x, dnorm(x, mean(daily), sd(daily)), col="blue")
hist(weekly, probability=T, nclass=40, xlab=’’, col="orange",
main="Weekly returns")
lines(x, dnorm(x, mean(weekly), sd(weekly)), col="blue")
hist(monthly, probability=T, nclass=30, xlab=’’, col="orange", main="Monthly returns")
lines(x, dnorm(x, mean(monthly), sd(monthly)), col="blue")
# now add in qq plots
qqnorm(daily,xlab=’Normal quantile’, ylab=’Quantile of daily returns’,
col="brown")
qqline(daily, col="blue")
qqnorm(weekly,xlab=’Normal quantile’, ylab=’Quantile of weekly returns’,
col="brown")
qqline(weekly, col="blue")
qqnorm(monthly,xlab=’Normal quantile’, ylab=’Quantile of monthly returns’,
col="brown")
qqline(monthly, col="blue")
```



Autocorrelations for NASDAQ:

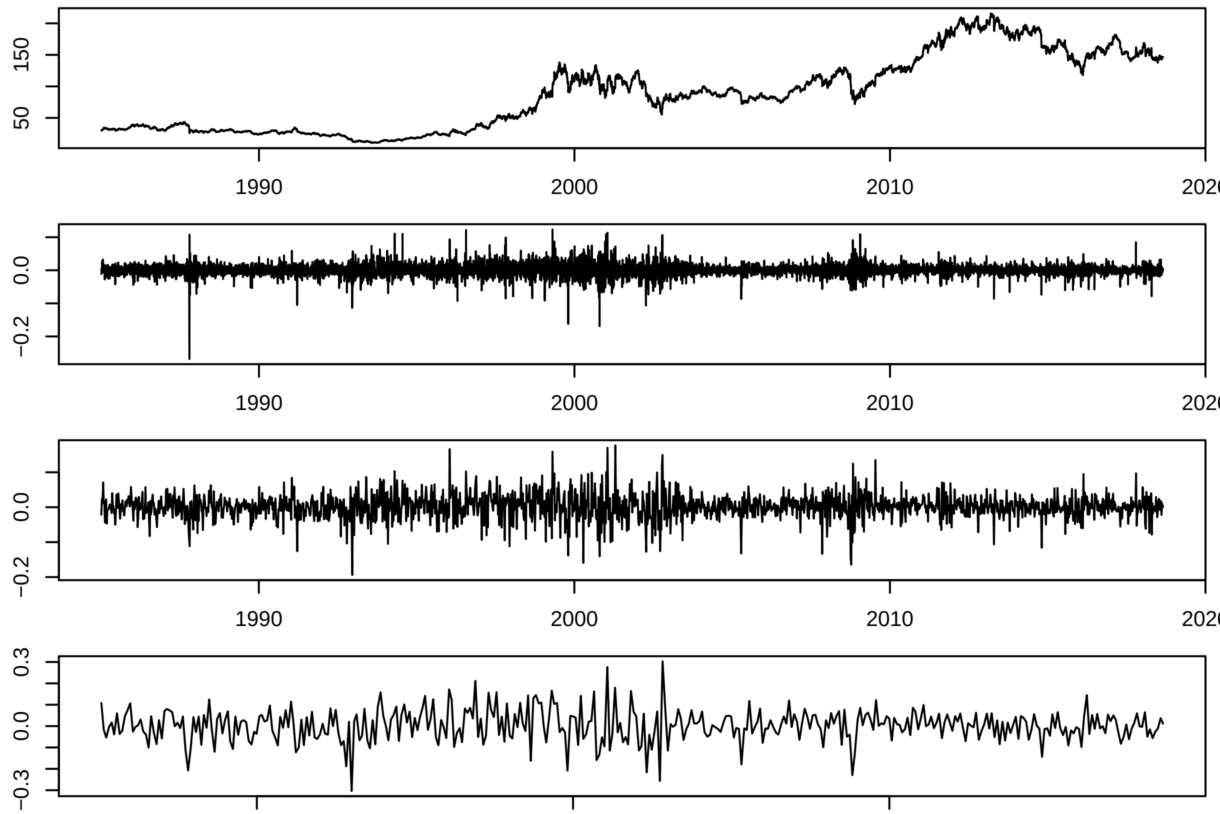
```
par(mar=c(3,4,3,1),mfrow=c(3,3), mgp=c(1.8,0.5,0))

acf(daily,  main='Daily returns')
acf(weekly, main='Weekly returns')
acf(monthly, main='Monthly returns')
acf(daily*daily, main='Squared daily returns')
acf(weekly*weekly, main='Squared weekly returns')
acf(monthly*monthly, main='Squared monthly returns')
acf(abs(daily), main='Absolute daily returns')
acf(abs(weekly),  main='Absolute weekly returns')
acf(abs(monthly), main='Absolute monthly returns')
```



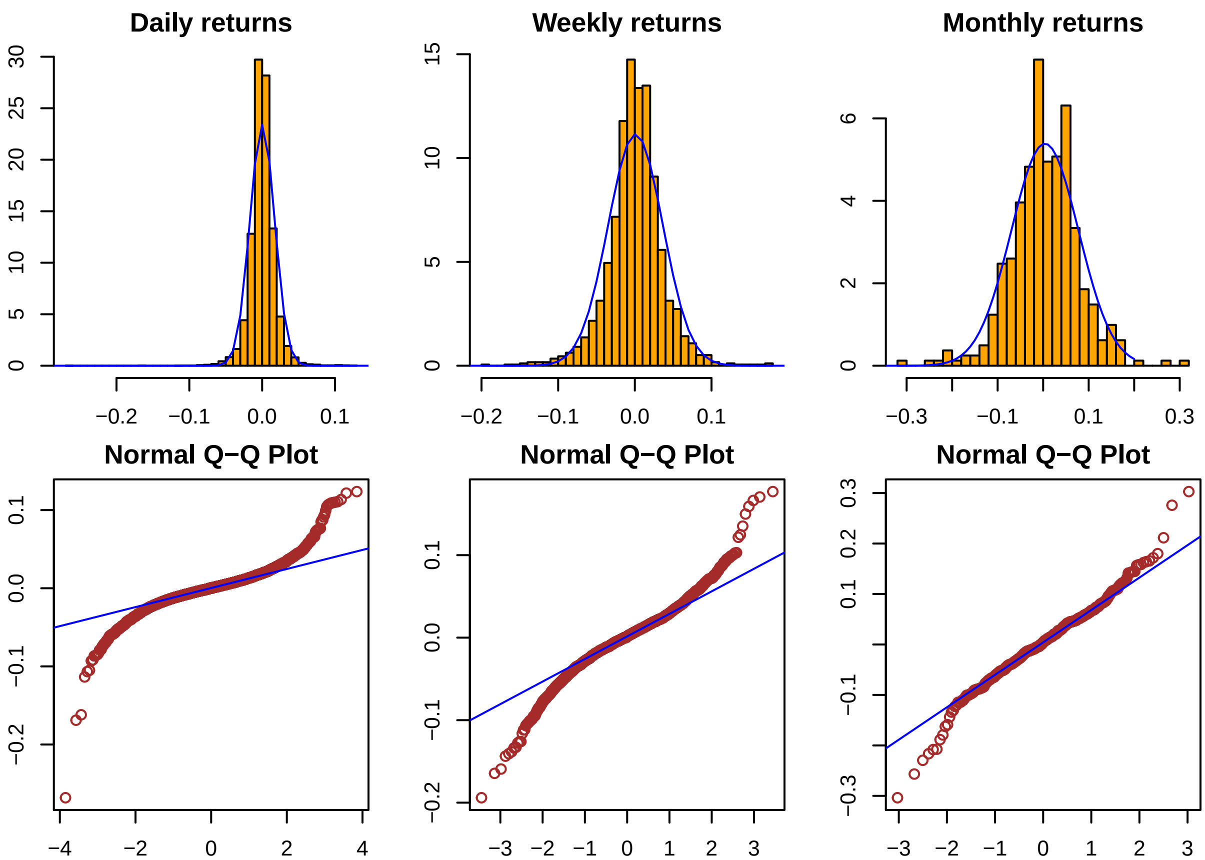
For IBM:

```
# Extract
daily <- log(dailyReturn(IBM)+1)
weekly <- log(weeklyReturn(IBM)+1)
monthly <- log(monthlyReturn(IBM)+1)
# Plot
par(mar=c(1,3,2,1),mfrow=c(4,1))
plot(index(IBM), as.numeric(IBM$IBM.Close), type = "l",
     ylab = "daily price")
plot(index(daily), daily, type = "l",
     ylab="daily log return")
plot(index(weekly), weekly, type = "l",
     ylab="weekly log return")
plot(index(monthly), monthly, type = "l",
     ylab="monthly log return")
```



Histograms for IBM:

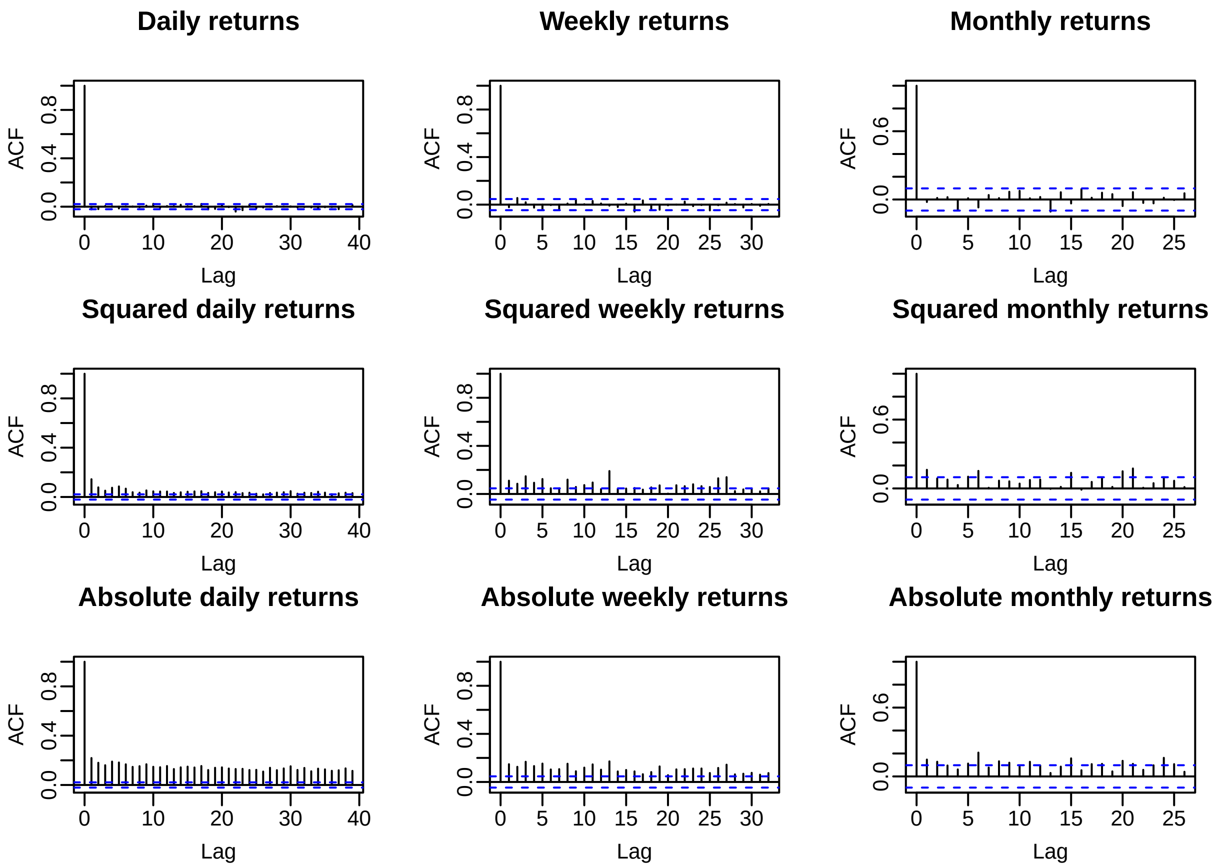
```
par(mar=c(2,3,2,1),mfrow=c(2,3))
hist(daily, probability=T, nclass=35, xlab='',
col="orange",main="Daily returns")
x <- seq(-0.5, 0.2, 0.01)
lines(x, dnorm(x, mean(daily), sd(daily)), col="blue")
hist(weekly, probability=T, nclass=40, xlab='', col="orange",
main="Weekly returns")
lines(x, dnorm(x, mean(weekly), sd(weekly)), col="blue")
hist(monthly, probability=T, nclass=30, xlab='', col="orange", main="Monthly returns")
lines(x, dnorm(x, mean(monthly), sd(monthly)), col="blue")
# now add in qq plots
qqnorm(daily,xlab='Normal quantile', ylab='Quantile of daily returns',
col="brown")
qqline(daily, col="blue")
qqnorm(weekly,xlab='Normal quantile', ylab='Quantile of weekly returns',
col="brown")
qqline(weekly, col="blue")
qqnorm(monthly,xlab='Normal quantile', ylab='Quantile of monthly returns',
col="brown")
qqline(monthly, col="blue")
```



Autocorrelations for IBM:

```
par(mar=c(3,4,3,1),mfrow=c(3,3), mgp=c(1.8,0.5,0))

acf(daily,  main='Daily returns')
acf(weekly, main='Weekly returns')
acf(monthly, main='Monthly returns')
acf(daily*daily, main='Squared daily returns')
acf(weekly*weekly, main='Squared weekly returns')
acf(monthly*monthly, main='Squared monthly returns')
acf(abs(daily), main='Absolute daily returns')
acf(abs(weekly),  main='Absolute weekly returns')
acf(abs(monthly), main='Absolute monthly returns')
```



1.3

Consider the following quote from Eugene Fama who was Myron Scholes' thesis adviser: "If the population of prices changes is strictly normal, on the average for any stock… an observation more than 5 standard deviations from the mean should be observed once every 7000 years. In fact such observations seem to occur about once every 3 or 5 years." For $X \sim N(\mu, \sigma^2), P(|X - \mu| > 5\sigma) = 5.733 \times 10^{-7}$, deduce how many observations per year Fama was implicitly assuming to be made. If a year is defined as 252 trading days and daily returns are normal, how many years is it expected to take to get a 5 standard deviation event? How does this answwer to the last question change when the daily returns follow the t-distribution with 4 degrees of freedom?

```
cat("The probability of getting this shock is 1 out of ",
    1/(5.7333*10e-7), " days or ", 1/(5.7333*10e-7)/252, " years", "\n")
```

```
## The probability of getting this shock is 1 out of  174419.6  days or  692.1413  years
```

To get $P(Z < -5)$:

```
p <- pnorm(-5)
cat("The probability of getting this shock is 1 out of ",
    1/p, " days or ", 1/p/252, " years", "\n")
```

```
## The probability of getting this shock is 1 out of  3488556  days or  13843.48  years
```

When the daily returns follow a t-distribution with 4 df:

```
p <- pt(-5, df = 4)
cat("The probability of getting this shock is 1 out of ",
    1/p, " days or ", 1/p/252, " years", "\n")
```

```
## The probability of getting this shock is 1 out of  267.0072  days or  1.059552  years
```

1.8

According to the efficient market hypothesis, is the return of a portfolio predictable? Is the volatility of a portfolio predictable? State the most appropriate mathematical form of the efficient market hypothesis.

By EMH, return of a portfolio is not predictable. The volatility is predictable since on average it is assumed that the average change in log price reaches the expectation μ_t. In other words, our asset return process can be expressed as:

$$r_t = \mu_t + \epsilon_t \quad \epsilon_t \sim (0, \sigma_t^2)$$

where μ_t is the rational expectation of the return r_t at time $t-1$ and ϵ_t is the return due to unpredictable news that arrives between $t-1$ and t. The EMH holds when:

$$E(r_t | r_{t-1}, r_{t-2}, \dots) = E(r_t) \quad \text{almost surely}$$

In other words, $r_t - \mu_t$ is a martingale difference with respect the information available up to $t-1$.

1.9

If the Ljung-Box test is employed to test the efficient market hypothesis, what null hypothesis is to be tested? If the autocorrelation for the first 4 lags of the monthly log-returns of the S&P500 is:

$$\hat{\rho}(1) = .2, \quad \hat{\rho}(2) = -0.15, \quad \hat{\rho}(3) = 0.25, \quad \hat{\rho}(4) = 0.12$$

based on the last 5 years of data, is the efficient market hypothesis reasonable?

The null hypothesis tested is whether r_t is a white noise process. The efficient market hypothesis is reasonable as shown below; we cannot reject the white noise hypothesis for the log return data.

```
Qm <- function(t, rho, m){
  temp <- 0
  for(j in 1:m){
    temp <- temp + (1/(t-j))*rho[j]^2
  }
  return(t(t+2)*temp)
}
# feed in different m, find the P(Q > Qm)
cat("When m = 1, the P( Q > Qm) is :",
    1-pchisq(Qm(60, c(.2, -0.15, .25, .12), 1), df = 1), "\n")
```

```
## When m = 1, the P( Q > Qm) is : 0.8375552
```

```r
cat("When m = 2, the P( Q > Qm) is :",
    1-pchisq(Qm(60, c(.2, -0.15, .25, .12), 2), df = 2), "\n")
```

```
## When m = 2, the P( Q > Qm) is : 0.9674971
```

```r
cat("When m = 3, the P( Q > Qm) is :",
    1-pchisq(Qm(60, c(.2, -0.15, .25, .12), 3), df = 3), "\n")
```

```
## When m = 3, the P( Q > Qm) is : 0.9874569
```

```r
cat("When m = 4, the P( Q > Qm) is :",
    1-pchisq(Qm(60, c(.2, -0.15, .25, .12), 4), df = 4), "\n")
```

```
## When m = 4, the P( Q > Qm) is : 0.9973239
```

1.13

Let S_t be the price of an asset at time t. One version of the EMH assumes that the prices of any asset form a martingale process in the sense that:

$$E(S_{t+1}|S_t, S_{t-1}, \ldots) = S_t \quad \forall t$$

To understand the implication of this assumption, we consider the following simple investment strategy. With initial capital C_0 dollars, at the time t we hold α_t dollars in cash and β_t shares of an asset at the price S_t. Hence the value of our investment at time t is $C_t = \alpha_t + \beta_t S_t$. Suppose that our investment is self-financing in the sense that

$$C_{t+1} = \alpha_t + \beta_t S_{t+1} = \alpha_{t+1} + \beta_{t+1} S_{t+1}$$

and our investment strategy is entirely determined by the asset prices up to the time t. Show that if S_t is a martingale process, there exist no strategies such that $C_{t+1} > C_t$ with probability 1.

Given that $\alpha_t + \beta_t S_{t+1} = \alpha_{t+1} + \beta_{t+1} S_{t+1}$ must hold and the fact that this portfolio is self-financing, we can assume that $C_t = \alpha_t + \beta_t S_t$. Thus, when the portfolio is updated:

$$C_{t+1} - C_t = \beta_t(S_{t+1} - St)$$

and

$$E_t[C_{t+1} - C_t] = \beta_t[E(S_{t+1}) - S_t] = 0$$

Since, $E_t(S_{t+1}) = S_t$ is a martingale process. Therefore, it is not possible to find a strategy ensuring a strictly positive net return with probability 1.