

STATS 509 HW1

Author: Chongdan Pan(pandapcd)

1. Problem 1

(a)

Since X is double exponential with $\mu = 0$ and $\sigma = 2$, then we get

$$\frac{\sqrt{2}}{\lambda} = 2$$
$$f_X(x) = \frac{\lambda}{2} e^{-\lambda|x|}$$

The 0.01-quantile should be smaller than μ , therefore, we get

$$F_X(x) = \frac{1}{2} e^{\frac{\sqrt{2}}{2}x} = 0.2$$

Then we can get the quantile of X and transform it to Y

$$x_{0.1} = \sqrt{2} \ln 0.1 \rightarrow y_{0.1} \approx -4.276$$

(b)

Since $Z = \frac{1}{(X-2)^2}$, then we can get the $F_Z(x) = P[Z \leq x]$, therefore

$$F_z(x) = P\left[\frac{1}{(X-2)^2} \leq x\right] = P\left[(X-2) \geq \frac{1}{\sqrt{x}} \cup (X-2) \leq -\frac{1}{\sqrt{x}}\right]$$

if $x \leq 0$, then $F_z(x) = 0$, otherwise

$$F_Z(x) = F_x\left(2 - \frac{1}{\sqrt{x}}\right) + \left(1 - F_x\left(2 + \frac{1}{\sqrt{x}}\right)\right)$$

When $x \leq \frac{1}{4} \rightarrow 2 - \frac{1}{\sqrt{x}} \leq 0$, we have

$$F_z(x) = \frac{1}{2} e^{-\frac{\sqrt{2}}{2}(2 - \frac{1}{\sqrt{x}})} + \frac{1}{2} e^{-\frac{\sqrt{2}}{2}(2 + \frac{1}{\sqrt{x}})}$$

Since $F_z(\frac{1}{4}) \geq \frac{1}{2} > 0.1$, we just need to consider $F_z(x) = 0.1$, so we get

$$x_{0.1} \approx 0.05267$$

2. Problem 2

(a)

Since $r \sim N(0, 0.025)$ and $r = \ln(R+1)$, we have $\frac{\ln(R+1)}{0.025} \sim N(0, 1)$, therefore

$$R_{total} = 1e8 \cdot R_{0.002} = 1e8 \cdot (e^{0.025\Phi(0.002)} - 1) \approx -6942634$$

(b)

- When $\nu = 0.5$, the quantile can be computed by the following R command

```
res <- (exp(qged(0.025, 0, 0.025, 0.5))-1)*100000000
```

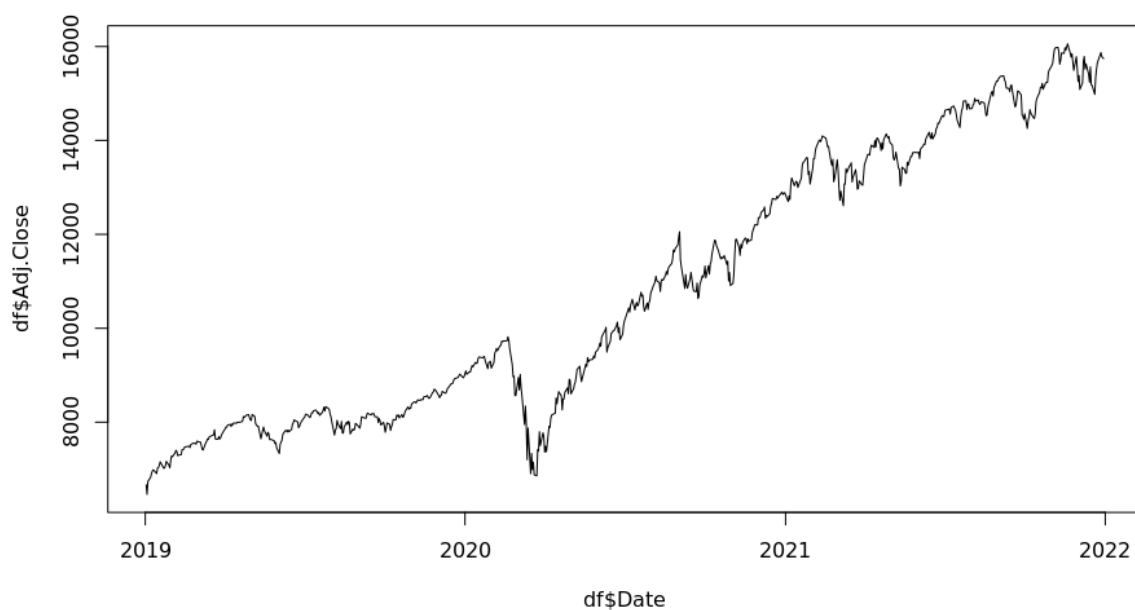
So the result is -5006212

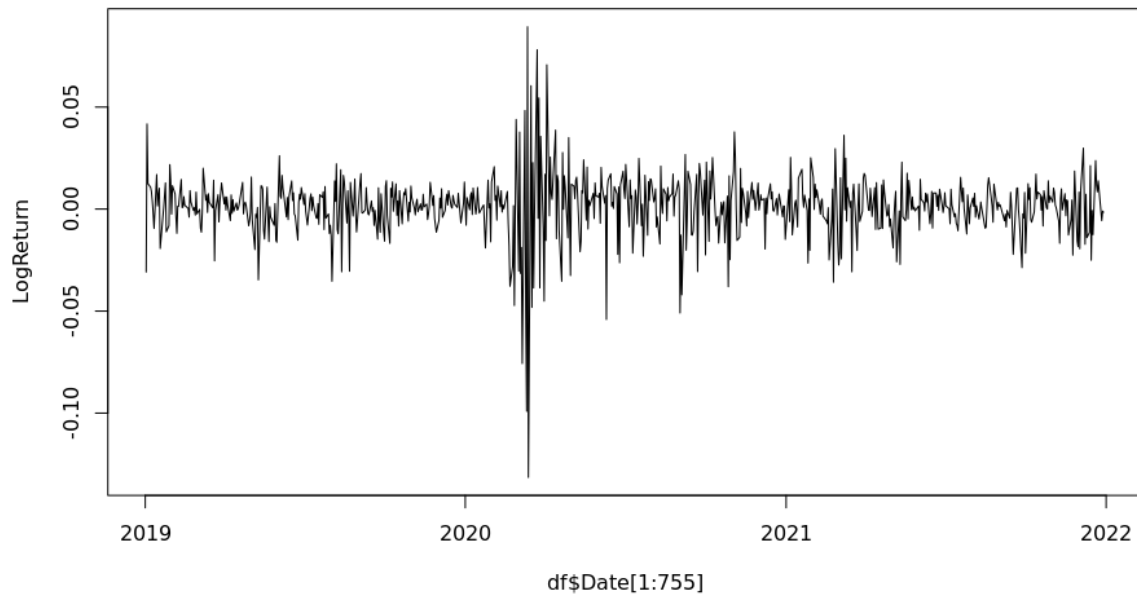
- When $\nu = 0.9$, the quantile can be computed in the same way and the result is -5189489
- When $\nu = 1.4$, the quantile can be computed in the same way and the result is -4995869

3. Problem 3

(a)

```
df <- read.csv("Nasdaq_daily_Jan1_2019-Dec31_2021.csv", header=TRUE)
df <- read.csv("Data.csv", header=TRUE)
df$Date <- as.Date(df$Date)
LogReturn <- diff(log(df$Adj.Close))
plot(df$Date, df$Adj.Close, type="l")
plot(df$Date[1:755], LogReturn, type="l")
```





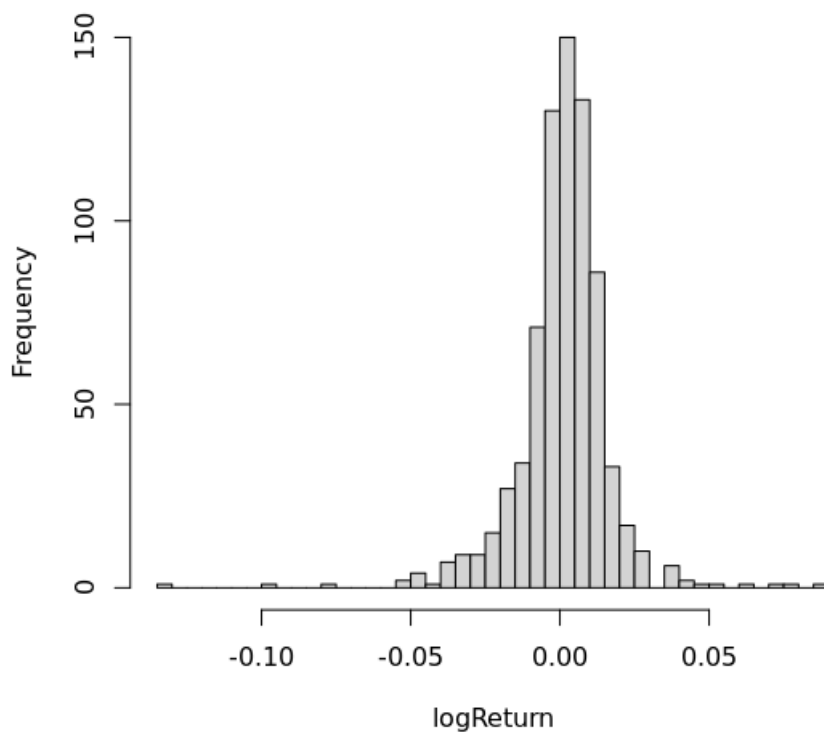
The plot shows that there is a rising trend in the adjusted close, and the mean of the log return is close to 0. The log return is typically in the interval of $[-0.05, 0.05]$, but it will have some extreme value when there is high volatility of the adjusted close price.

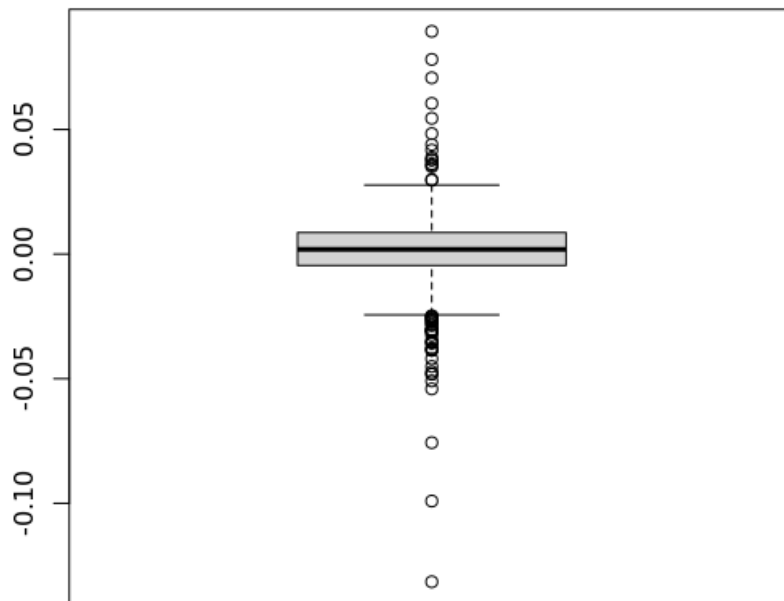
(b)

```
skewness(logReturn)
kurtosis(logReturn)
hist(logReturn, breaks=50)
boxplot(logReturn)
```

The skewness is -1.047 and the kurtosis is 12.38

Histogram of logReturn





The log return's distribution is not fully symmetric, and it's a bit skewed to right, meaning the expectation of log return is positive. The kurtosis of the log return is extremely high, implying that usually the value is very close to 0.

(c)

```
mu <- mean(logReturn)
sigma <- sd(logReturn)
estimated_q <- qdexp(0.004, mu, 2^0.5 / sigma)
q <- quantile(logReturn, 0.004)
```

The estimated mean is 1.138E-3 , and the estimated standard deviation is 1.566E-2

The estimated_quantile is -0.052 and the sample quantile is -0.054. Therefore, the quantile is a bit smaller than the estimated one, but they're very close.

4. Problem 4

(a)

The 20-period log return $r_t(20) = \ln \frac{100}{97} = \sum_{i=1}^{20} r_{t+i}$

Since r_{t+i} are i.i.d variables following $N \sim (0.0002, 0.03^2)$, $r_t(20)$ should follow $N \sim (0.004, 0.018)$

Then we can compute the probability with R

```
prob <- 1 - pnorm(log(100/97), 0.004, 0.018^0.5)
```

Therefore, we get the probability equals to 0.4218

(b)

The n -period log return $r_t(n) \geq e^2$, and it should follow $N \sim (0.0005n, 0.012\sqrt{n})$

We can use binary search to solve this problem

```
l <- 1
r <- 10000
while(l < r){
  m = floor((l + r) / 2)
  prob <- 1 - pnorm(log(2), 0.0005 * m, 0.012 * m ^ 0.5)
  if(prob < 0.9){
    l <- m + 1
  }
  else{
    r <- m
  }
}
```

Then we get $l=3099$, which means we need 3099 days.