DS 613 HW 5

David Saff

DS 413/613 Homework 5

Instructions: Generate an Rmarkdown File and a Word File that shows all Rcode, script, and output. Push your files to GitHub and email me the link for your GitHub repository so that I can view your files.

Statistics Review  
An investor sued his broker for lack of diversification. Below are the rates of return (in percent) for the investor’s portfolio over 39 months (data from Moore, McCabe, and Craig (2017)). The average of the S&P 500 stock index for the same period was 0.95%. Does the broker perform worse than average? Copy and paste the data below into R

Stockreturns <- c(-8.36, 1.63, -2.27, -2.93, -2.70,   
 -2.93, -9.14, -2.64, 6.82, -2.35,   
 -3.58, 6.13, 7.00, -15.25, -8.66,  
 -1.03, -9.16, -1.25, -1.22, -10.27,  
 -5.11, -0.80, -1.44, 1.28, -0.65,  
 4.34, 12.22, -7.21, -0.09, 7.34,   
 5.04, -7.24, -2.14, -1.01, -1.41,   
 12.03, -2.53, 4.33, 1.35)  
 Stockreturns

## [1] -8.36 1.63 -2.27 -2.93 -2.70 -2.93 -9.14 -2.64 6.82 -2.35  
## [11] -3.58 6.13 7.00 -15.25 -8.66 -1.03 -9.16 -1.25 -1.22 -10.27  
## [21] -5.11 -0.80 -1.44 1.28 -0.65 4.34 12.22 -7.21 -0.09 7.34  
## [31] 5.04 -7.24 -2.14 -1.01 -1.41 12.03 -2.53 4.33 1.35

1. Use and show R code to find the mean of the sample data.

mean(Stockreturns)

## [1] -1.124615

1. Use and show R code to find the standard deviation of the sample data

sd(Stockreturns)

## [1] 5.977673

1. Assuming a normal distribution, use and show R code to find the proportion of returns that are less than -1.5.

library(tidyverse)

## -- Attaching packages ------------------------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.2 v purrr 0.3.4  
## v tibble 3.0.3 v dplyr 1.0.2  
## v tidyr 1.1.2 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.5.0

## Warning: package 'forcats' was built under R version 4.0.3

## -- Conflicts ---------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

pnorm(q = -1.5, mean = mean(Stockreturns), sd= sd(Stockreturns))

## [1] 0.4749638

1. Assuming a normal distribution, use and show R code to find return value that is above 70% of the returns

pnorm(.70, mean= mean(Stockreturns), sd= sd(Stockreturns))

## [1] 0.6199077

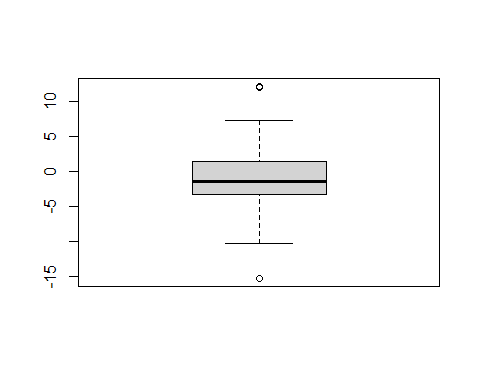
1. Use and show R code to find Q1 for the data.

qnorm(p = 0.25, mean= mean(Stockreturns), sd= sd(Stockreturns))

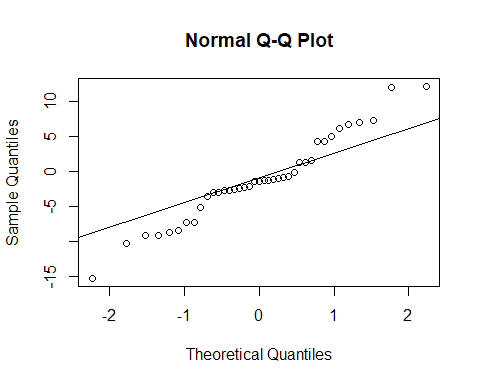
## [1] -5.156495

1. Explore the data by producing and examining a boxplot and checking for normality.

boxplot(Stockreturns)



qqnorm(Stockreturns)  
qqline(Stockreturns)



According to the plots above, the data does not seem entirely normal since the QQ plot shows some points that are not on the diagonal line.

1. State the appropriate null and alternative hypothesis required for the appropriate t test

H0: The true mean is .95 HA: The true mean is less than .95

1. Use and show R code that will output the needed p value and confidence interval to determine if the null hypothesis should be rejected.

t.test(Stockreturns , mu=0.95, alternative = "less", conf.level = .95)

##   
## One Sample t-test  
##   
## data: Stockreturns  
## t = -2.1674, df = 38, p-value = 0.01827  
## alternative hypothesis: true mean is less than 0.95  
## 95 percent confidence interval:  
## -Inf 0.4891698  
## sample estimates:  
## mean of x   
## -1.124615

1. Now answer the question originally stated. Does the broker perform worse than average? (Explain or justify in two or three sentences)

Because the t-test yielded a p-value that was less than .05, we reject the null hypothesis. This means the broker perfromed worse than average since the true mean is less than .95.