PS11

Conrad Polkosnik

April 2018

1 Introduction

What is the stock market?

Why is it important?

Who can be a part of the stock market? Who is involved in the markets?

Trading and investing

2 Literature Review

History/Background of stock price prediction

What have poeple done to predict stock prices?

What techniques have they used?

Did these techniques work?

What methods were good? What methods were bad?

What is technical analysis? How is it used?

Implications of stock prediction

3 Data

Where did I get the data from?

Is the data trustworthy/unbiased?

What data did I use specifically?

Differences in prediction based on different time periods of historical data

Differences in prediction based on number of simulations

4 Methods

Quantmod for data

Visualization of overall trend

Fundamental Technical Analysis

Analyzation of distribution (log values, mean, standard deviation)

Random Walk

Monte Carlo Simulation

5 Findings

Technical Analysis

Log returns

Probabilities and likelihood

Daily average increase

Random Walk (parameters, organization, visualization)

Monte Carlo Simulation

6 Conclusions

Summarize overall introduction and literature review in regard to the benefits from doing this project.

Summarize data, methods, and findings.

What was the ultimate goal of the project? Was it met? Does it make sense?

References

Biermann, Luke. 2006. *An Investigation Into Stock Market Predictions Using Neural Networks Applied To Fundamental Financial Data.* Ph.D. thesis. URL http://www.cs.bath.ac.uk/ mdv/courses/CM30082/projects.bho/2005-6/biermann-l-dissertation-

business science.io. 2016. "Quantitative Stock Analysis Tutorial: Screening the Returns for Every SP500 Stock in Less than 5 Minutes." URL http://www.business-science.io/investments/2016/10/23/SP500$_A$*nalysis.html*.

Falinouss, Pegah. 2007. *Stock Trend Prediction Using News Articles.* Ph.D. thesis. URL http://www.diva-portal.org/smash/get/diva2:1019373/FULLTEXT01.pdf.

Irding, Martin and Joacim Lyden. 2007. *Is Historical Data A Good Estimate Of The Future Risk of Funds?* Ph.D. thesis. URL https://gupea.ub.gu.se/bitstream/2077/9525/1/0708.4.pdf.

Xu, Selene Yue. ???? *Stock Price Forecasting Using Information from Yahoo Finance and Google Trend.* Ph.D. thesis. URL https://www.econ.berkeley.edu/sites/default/files/Selene Yue Xu.pdf.

7 Code

```
---
title: "Final Project"
author: "Conrad Polkosnik"
date: "3/6/2018"
output: word_document
---

Loading Packages
```{r}
library(quantmod)   # get stock prices & useful stock analysis functions
library(xts)        # extensible time series
library(tidyverse)  # ggplot2, purrr, dplyr, tidyr, readr, tibble
library(stringr)    # working with strings
```


Downloading Stock Data
```{r}
# Calling quantmod
library("quantmod")

# Gathering historical stock data for Exxon Mobil since 1970
getSymbols("XOM", src='yahoo', from="1970-01-01")

```


```{r}
# Understanding data pulled
library(magrittr) #for pipe operator

XOM %>% class()

XOM %>% str()
```
```

```
XOM %>% head()
```

Visualize Trend

```{r}
XOM %>% Ad() %>% chartSeries()
```

Fundamental Technical Analysis

```{r}
XOM %>% chartSeries(TA='addBBands();
                    addBBands(draw="p");
                    addVo();
                    addMACD()',
                    subset='1970-2018',
                    theme="white"
                    )
```

Gathering log daily returns for XOM

```{r}
XOM %>% Ad() %>% dailyReturn(type = 'log') %>% head()
```

Using log returns to analyze distribution

```{r}
XOMLogReturns <- XOM %>% Ad() %>% dailyReturn(type = "log")

names(XOMLogReturns) <- "XOM.Log.Returns"

# Plotting the log returns using a histogram
library(ggplot2)
XOMLogReturns %>% ggplot(aes(x = XOM.Log.Returns)) + geom_histogram(bins = 100) + geom_densi
```

Setting probabilities for study of liklihood

```{r}
Probabilities <- c(.005, .025, .25, .5, .75, .975, .995)

DistributionalLogReturns <- XOMLogReturns %>% quantile(probs = Probabilities, na.rm = TRUE)

DistributionalLogReturns
```

```{r}
MeanLogReturns <- mean(XOMLogReturns, na.rm = TRUE)
StDevLogReturns <- sd(XOMLogReturns, na.rm = TRUE)

MeanLogReturns
StDevLogReturns
```

Finding daily average increase or decrease in return
```{r}
MeanLogReturns %>% exp()
```
Analysis: On average, the investor should expect 0.0657% increase daily in return.

Setting Parameters for Random Walk
```{r}
# Parameters
N <- 252
mu <- MeanLogReturns
sigma <- StDevLogReturns
day <- 1:N

InitialPrices <- XOM$XOM.Adjusted[[nrow(XOM$XOM.Adjusted)]]

set.seed(386) # for reproducibility
Price  <- c(InitialPrices, rep(NA, N-1))
for(i in 2:N) {
    Price[i] <- Price[i-1] * exp(rnorm(1, mu, sigma))
              }

library(tidyverse)
PriceSimulation <- cbind(day, Price) %>% as_tibble()

# Visualize price simulation
PriceSimulation %>%
    ggplot(aes(day, Price)) +
    geom_line() +
    ggtitle(str_c("XOM: Simulated Prices for ", N," Trading Days"))
```

Monte Carlo Simulation for Random Walk
```{r}
# Parameters
```

```
N     <- 252 # Number of Stock Price Simulations
M     <- 100  # Number of Monte Carlo Simulations
mu    <- MeanLogReturns
sigma <- StDevLogReturns
day <- 1:N
InitialPrices <- XOM$XOM.Adjusted[[nrow(XOM$XOM.Adjusted)]]

# Simulate prices using for loop and Monte Carlo Matrix
set.seed(386)
MonteCarloMatrix <- matrix(nrow = N, ncol = M)
for (j in 1:M) {
    MonteCarloMatrix[[1, j]] <- InitialPrices
    for(i in 2:N) {
        MonteCarloMatrix[[i, j]] <- MonteCarloMatrix[[i - 1, j]] * exp(rnorm(1, mu, sigma))
                }
              }

# Organizing fata frame and format
PriceSimulation <- cbind(day, MonteCarloMatrix) %>% as_tibble()
nm <- str_c("Sim.", seq(1, M))
nm <- c("Day", nm)
names(PriceSimulation) <- nm
PriceSimulation <- PriceSimulation %>%
    gather(key = "Simulation", value = "Stock.Price", -(Day))

# Visualization
PriceSimulation %>%
    ggplot(aes(x = Day, y = Stock.Price, Group = Simulation)) +
    geom_line(alpha = 0.1) +
    ggtitle(str_c("XOM: ", M,
                  " Monte Carlo Simulations for Prices Over ", N,
                  " Trading Days"))
```

Stock price estimation and likelihood

```{r}
EstimatedStockPrices <- PriceSimulation %>% filter(Day == max(Day))

Probabilities <- c(.005, .025, .25, .5, .75, .975, .995)

DistributionalEstimatedStockPrices <- quantile(EstimatedStockPrices$Stock.Price, probs = Pro

DistributionalEstimatedStockPrices %>% round(2)
```