**Timing The Market Part 1**

**Kyle Kim**

**Abstract**

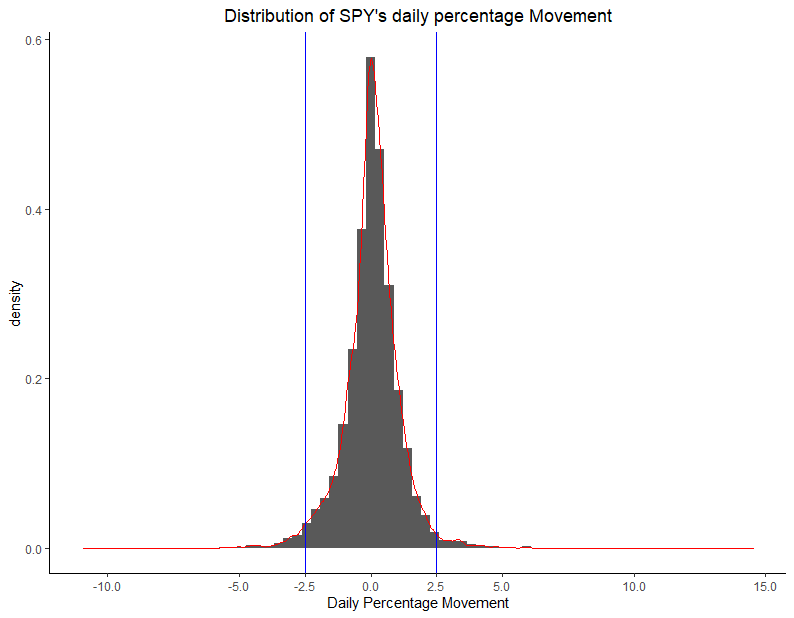
S&P500 index fund has an average of 8% annual return. Although statistics vary, it’s generally believed that index funds beat 80% of actively managed funds. However, for too many people, including myself, an 8% annual return is boring. Additionally, with the possibility of seeing many disastrous days like all of the black x days we have, people want to minimize holding periods but want to maximize return.

Generally known, the stock market usually moves in a specific direction for some time then makes a reversal. Hence, the goal of this research is to identify reversal moments and benefit from movements in a short amount of time. Over the course of several series, various methods of timing the market will be checked. Methods will be chosen based on researching strategies online, and test will be conducted for evaluation both in simulation wise and mathematically.

**Probability Estimation**

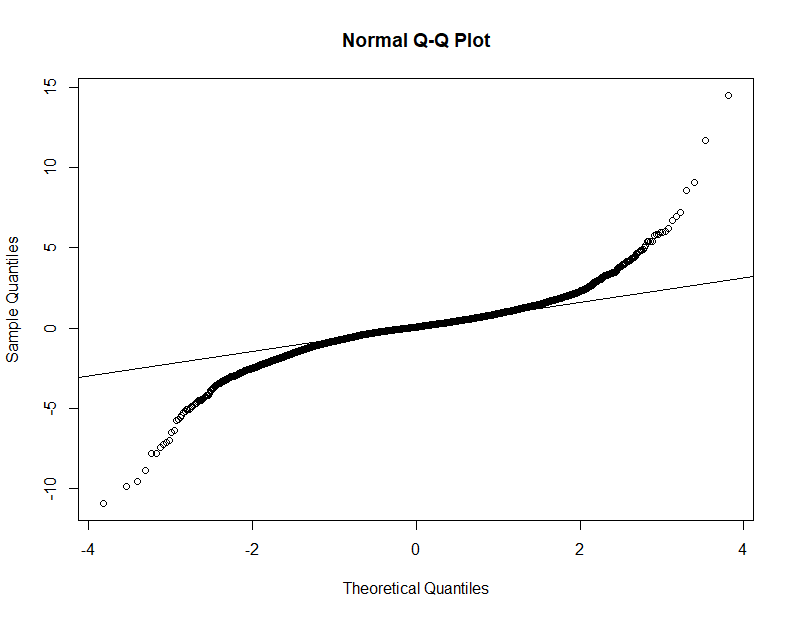
To calculate probability based on Bayes’ Theorem, couple of probabilities are needed. The first is the probability of stock movement. What is the likelihood of seeing 1% movement in the S&P 500 tomorrow? For approximation, 3 methods are used: 1. Histogram Density Estimation 2. Normal Distribution 3. Kernel Density Estimation.

*Histogram Density Estimation*

Density is estimated by plotting a histogram, and the probability can be computed by multiplying density X binwidth because the probability is the area under the curve. Because histogram is heavily depended on binwidth, selection must be carefully made. In this analysis, one of widely formula, , where n is the sample size, and sd is the standard deviation of the data.  
  
The histogram above shows that the majority of daily percentage movement lies between -2.5% and 2.5%. Such interval accounts about 94.72% of the movement.

*Normal Distribution*

It’s hard to state that the distribution of the daily percentage SPY movement follows normal distribution. The histogram above shows that the distribution has long tails and may look normal if restricted with limitations. For more exact normality test, Q-Q plot is created.

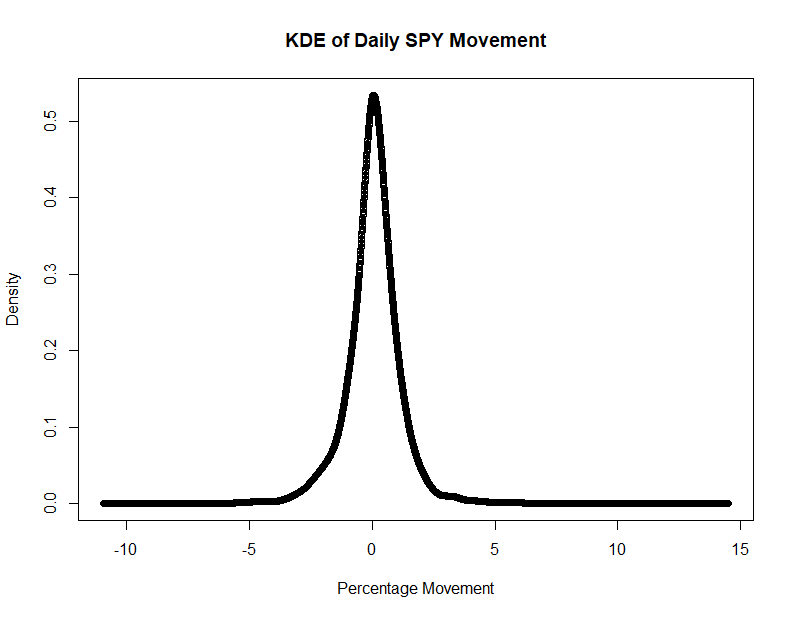


The Q-Q plot above verifies what was assumed in the histogram. In conclusion, the movement only follows normal distribution in limited conditions. Hence, the probability generated by using normal distribution can’t be accurate theoretically. The probability of return between -2.5% and 2.5% is 96.55%.

A reason why normal distribution is used despite not meeting the assumption is because of the use of Bollinger Bands in technical analysis. Bollinger Bands assumes that movements in the price range within the 2 standard deviations in both directions from the center, which is the widely known characteristic of normal distribution. The center could be any measurement. The most frequently used value is 20 simple moving average.

*Kernel Density Estimation*

Kernel Density Estimation (KDE) is the most widely used nonparametric density estimation, and, similar to histogram, it is heavily affected by the bandwidth. The assumptions are univariate and independent and identically distributed. In a timeseries data, like SPY’s daily movement, movements between days are dependent, so KDE may not provide an accurate measurement of probability.



The plot above shows the density estimation using gaussian kernel function. Triangular and uniform kernel functions are also checked to identify the most effective kernel function, but all three functions generated almost identical plot. Based on the KDE, the probability of return between -2.5% and 2.5% is about 95.82%.

Overall, these three methods estimated that SPY movements between -2.5% and 2.5% on daily basis for about 95% of the time. Therefore, it’s convincing if SPY moves outside of the interval then significant change in the market has occurred and likely to cause a movement.

**Modelling and Evaluation**

|  |  |
| --- | --- |
| Pr(daily % ∆ >0) = 53.61 | Pr(daily % ∆ <0) = 45.55 |

Note: Pr( daily % ∆ = 0) = 0.84

In the previous section, it was discussed that about 95% daily percentage movement lies between -2.5% and 2.5%, and any movements outside of this interval can be interpreted as key moments that reflect serious changes in the market psychology. The hypothesis is does a price movement outside of 95% confidence interval (key moments) cause steep movements in a direction where market participants can realize significant gains in a short amount of time?

The table below is returns after holding for a certain amount of day after buying at the close price of positive key moments of at least 2.5%, separated in three different levels. Higher percentage movement is assumed as a stronger movement. The test is conducted on the past 3 years, 2019-2021 historical prices of SPY. The same percentage is used for the exit strategy.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Percentage Movement | Empirical Occurrence | Estimated Occurrence  (order of hist,normal,kde) | Average Return on holding for 5 days | Average Return on holding for 10 days | Average Return on holding for 20days |
| ≥2.5% | 2.11% | 2.33%,1.86%, 1.92% | - 0.86% | -1.21% | -0.94% |
| ≥3.0% | 1.59% | 1.3%, 0.61%,1.29% | -0.79% | -1.1% | -0.93% |
| ≥3.5% | 1.06% | 0.99%, 0.17%,0.82% | -1.02% | -1.35% | -1.02% |

On average, none of buying on the date of positive key moments led to meaningful return that a person can capitalize on. One interesting fact is that normal distribution provided much lower estimated occurrence which proves that SPY’s movement is not based on the normal distribution. Furthermore, SPY is a more stable ticker most of tickers since it is accounting companies listed on the S&P 500 index, so other tickers should have more volatility thus cannot be estimated using the normal distribution.

*Reverse Approach*

From the previous analysis, it seems like the strong positive movement represent top of the market rather than a continuous strong movement or no movement at all since average returns are extremely low across all levels. Does negative movements result in the same conclusion as the positive movement?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Percentage Movement | Empirical Occurrence | Estimated Occurrence  (order of hist,normal,kde) | Average Return on holding for 5 days | Average Return on holding for 10 days | Average Return on holding for 20days |
| ≤-2.5% | 3.30% | 2.95%,1.58%, 2.35% | -0.93% | -1.09% | -1.12% |
| ≤-3.0% | 2.11% | 1.34%, 0.5%,1.32% | -0.67% | -0.86% | -1.32% |
| ≤-3.5% | 1.06% | 1.19%, 0.14%,0.8% | -0.67% | -0.76% | -1.33% |

The result of negative movements aligns with that of the positive movement. Hence, it’s evident to state that the price movement alone doesn’t indicate market direction for at most the next 20 days.

**Limitations**

1. Ticker Selection  
   SPY was selected for convenience because it automatically leaves out unhealthy stocks or manipulated stocks.

Future Work: 1. Select 10 random tickers and test them out. 2. Compare healthy stocks based on financial metrics with unhealthy stocks.

1. Better use of probability

It seemed like projecting certain percentage movement almost seems impossible because the numbers are so small so figuring out a better way to handle this to obtain meaningful probability would enhance the quality of the report.

Future Work: Create a reasonable percentage group.

1. New Entry and Exit strategies

It’s proven that the percentage alone and holding for x number of days don’t result in meaningful returns, so figuring out different entry and exit strategies could improve the success rate.

Future Work: Check crossovers and doji candles