Joseph Knapp

Data 606 Spring 2020

Deliverable III

Validating Financial Models

Validating Financial Models

In this phase of the project I began to explore some basic models. I have created several models which I will discuss below. The method of evaluation is through Monte Carlo simulations of each model which is compared to two benchmarks.

Each model I have created looks back at historical pricing data and returns the best stock (within the S&P500 Index) to buy on that random day and then tracks the stocks daily returns over the next year (250 trading days). The model is then ran several 100 times (my machine is the limiting factor here) to create a distribution of each models annual returns so the probabilities distributions of each model can be evaluated against each other.

Benchmark Data Models

Two different benchmark models have been creating. The first benchmark model chooses a random day the market was open, picks a random stock, and tracks the stocks return over the next year. This process is repeated several 100 times to create a probability distribution for analysis; see *Fig1* below.

The other benchmark I used to compare stocks was the overall Market Return for which I used the S&P500 Index as a proxy. The Market model chooses a random day the market was open, and tracks the markets return over the next year. This process is repeated several 100 times to create a probability distribution for analysis; see *Fig2* below.

Basic Financial Models – Expected Return

The first financial model is extremely basic. The expected return of stock i is $E[R_i] = \frac{1}{N}\sum R_i$. The Expected Return model chooses a random day the market was open, determines the stock with the greatest expected return based on historical data, and tracks its returns over the next year. The process is repeated several 100 times to create a probability distribution for analysis; see *Fig3* below.

Basic Financial Models – Sharpe-ratio

The Sharpe-ratio is a measure of the risk premium per unit of volatility. The risk premium is the difference in the expected return of an asset and the risk-free return (T-bill rate). Mathematically, the Sharpe-ratio of asset i is $\frac{E[R_i]-r_f}{\sigma_i}$ where σ_i is the standard deviation of asset i's historical returns, also referred to as volatility. The Sharpe-ratio model chooses a random day the market is open, finds the stock with the greatest Sharpe-ratio on that date based on the stocks historical data, and tracks the stocks return over the next year. The process is repeated several hundred times to create a probability distribution for analysis; see *Fig4* below.

Basic Financial Models – Beta

The Beta (β) of a security is a measure of its systematic risk. Non-systematic risk is firm/industry/region/etc. specific and can be diversified away with a well-diversified portfolio, so this risk is not considered here. Systematic risk is risk which is nondiversifiable, market wide, and thus it cannot be eliminated. In theory, the market portfolio contains all tradable assets in the world, and only contains systematic risk. As a proxy for this portfolio, we use the S&P500 market index. Beta of stock i measures systematic risk by calculating the percentage change in the assets return per 1% change in the market; mathematically, $\beta_i = \frac{\text{Cov}[R_i, R_{Mkt}]}{\text{Var}[R_i]}$. I am interested in returns of stocks with large betas, which indicate the security has more systematic risk than the market. Perhaps later I should reconsider other values of β. The Beta model chooses a random day the market is open, finds the stock with the greatest beta on that date based on the stocks historical data, and tracks the stocks return over the next year. The process is repeated several hundred times to create a probability distribution for analysis; see *Fig5* below.

Basic Financial Models – CAPM

The Capital Asset Pricing Model (CAPM) calculates the expected/required return of a securitie based on it's sensitivity to systematic/market risk. Multiplying the beta of the security by the market's risk premium $\left(E[R_{Mkt}] - r_f\right)$ is the sensitivity of the security, which is added to the risk-free rate on that date. The CAPM mathematically is $r_i = r_f + \beta_i(E[R_{Mkt}] - r_f)$. The CAPM model chooses a random day the market is open, finds the stock with the greatest expected/required return based on CAPM on that date using stocks historical data, and then tracks that stocks return over the next year. The process is repeated several hundred times to create a probability distribution for analysis; see *Fig6* below.

Basic Financial Models – ALPA

The Alpha of a stock is the difference in the expected return of an asset and the required return of an asset calculated using CAPM. The expected return of a security can be calculated different ways. Here I will calculate it the same way as the Expected Return model, via the securities historic arithmetic mean, thus the alpha of security i mathematically is $\alpha_i = E[R_i] - r_i$. The Alpha model chooses a random day the market is open, finds the stock with the alpha on that date using stocks historical data, and then tracks that stocks return over the next year. The process is repeated several hundred times to create a probability distribution for analysis; see *Fig7* below.

Analyzing the Basic Financial Models

This analysis involves running each model 500 times, creating probability distributions based on the returns of each model, and comparing them. Below, Fig8 displays the cumulative distribution function of each model. One of the goals of my project was to be able to identifiable stocks which can beat the market consistently. Looking at Fig8 along the x-axis, a value of 1.0 means you broke even and a value of 2.00 means you doubled your money in a year. The y-axis is the probability that the model made less-than the annual return. For example, the Market Return (orange) line has at point (60, 1.1407) can be interpreted as a 60% probability that the Market model will make you can annual return of less than 1.1407, and a 40% chance the return will be greater. But for the same 60% probability, the Alpha model will return up to 1.5319 in a year, with a 40% probability the return will be more.

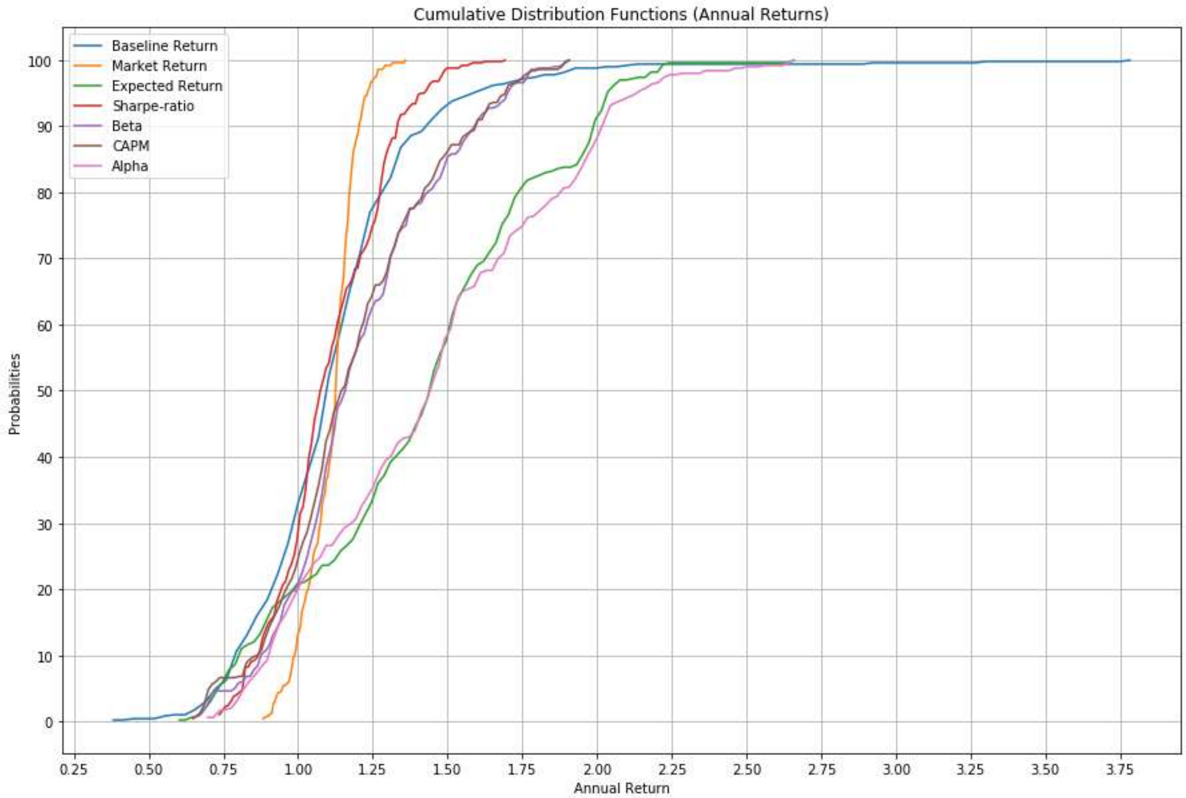

Fig8. Displays the cumulative distribution function of the probabilities of returns of different models tested.

More Advanced Financial Models

I hope to be able to explore more advanced models in the coming weeks. Models I hope to test in the future are the Fama-French-Carhart models, Lognormal Return models, and possibly apply Put-Call Parity and Black-Sholes pricing models to find mispriced stocks.

Implementation Moving Forward

My goal is to use an ensemble method and try to combine the more successful models I test. I would also like to identify the stock which perform well in several models and see if I can uncover any commonalities through various statistical analysis.

For now, my biggest problem is computation time. So any help with making my code more *pytonic* would be greatly appreciated.

Figures/Diagrams

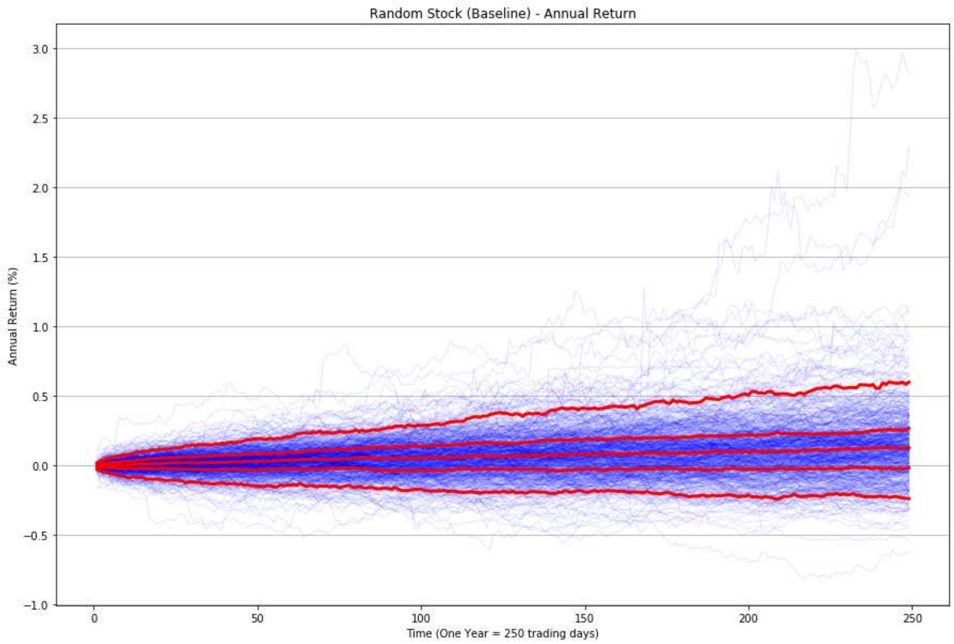

Fig1. The yearly return of a random stock chosen on a random day. The blue lines are timeseries return of each individual stock. The red lines are the 5%, 25%, 50%, 75% and 95% confidence intervals of the returns.

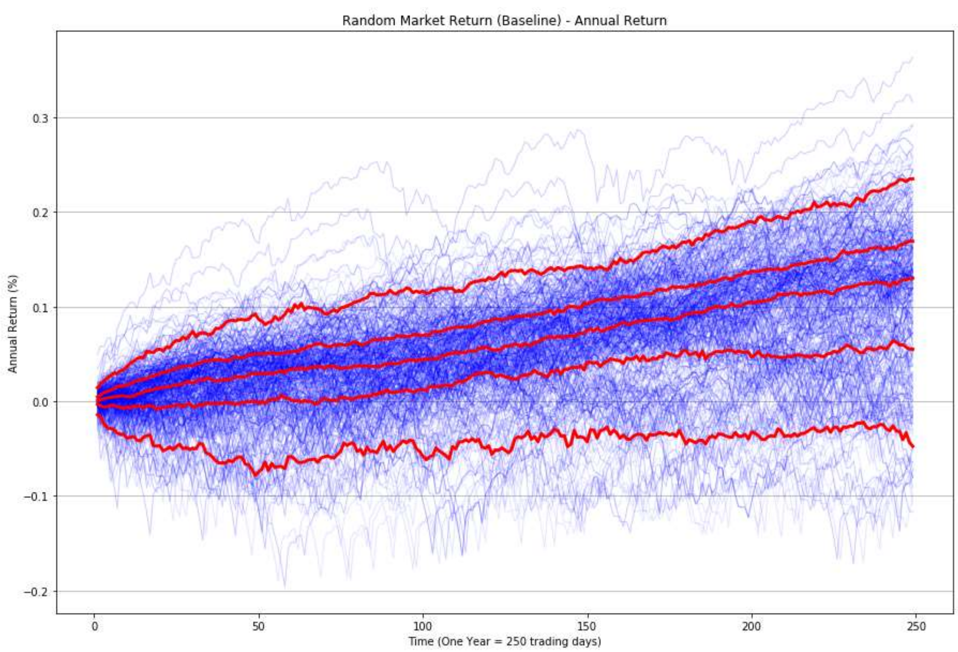

Fig2. The yearly return of investing in the Market Index chosen on a random day. The blue lines are timeseries return of the Market. The red lines are the 5%, 25%, 50%, 75% and 95% confidence intervals of the returns.

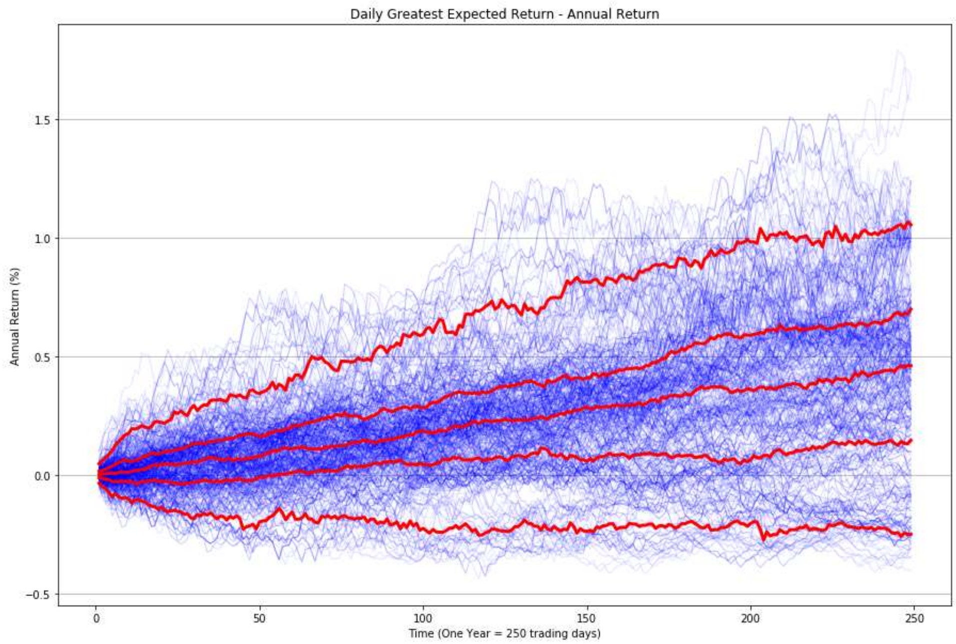

Fig3. The yearly return of a stock chosen on a random day based on its Expected Return. The blue lines are timeseries return of each individual stock. The red lines are the 5%, 25%, 50%, 75% and 95% confidence intervals of the returns.

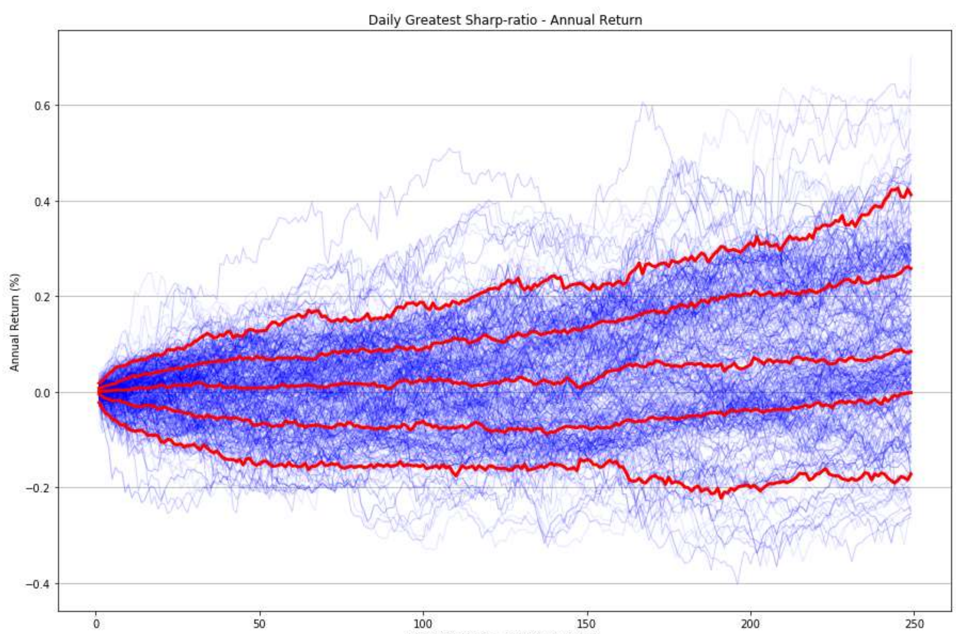

Fig4: The yearly return of a stock chosen on a random day based on its Sharpe-ratio. The blue lines are timeseries return of each individual stock. The red lines are the 5%, 25%, 50%, 75% and 95% confidence intervals of the returns.

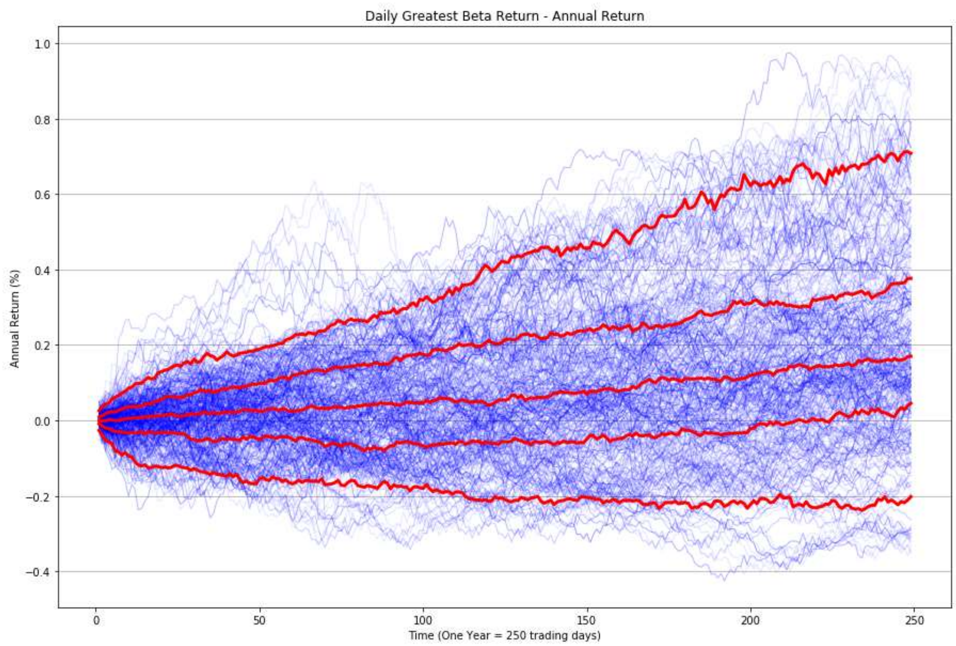

Fig5. The yearly return of a stock chosen on a random day based on its Beta. The blue lines are timeseries return of each individual stock. The red lines are the 5%, 25%, 50%, 75% and 95% confidence intervals of the returns.

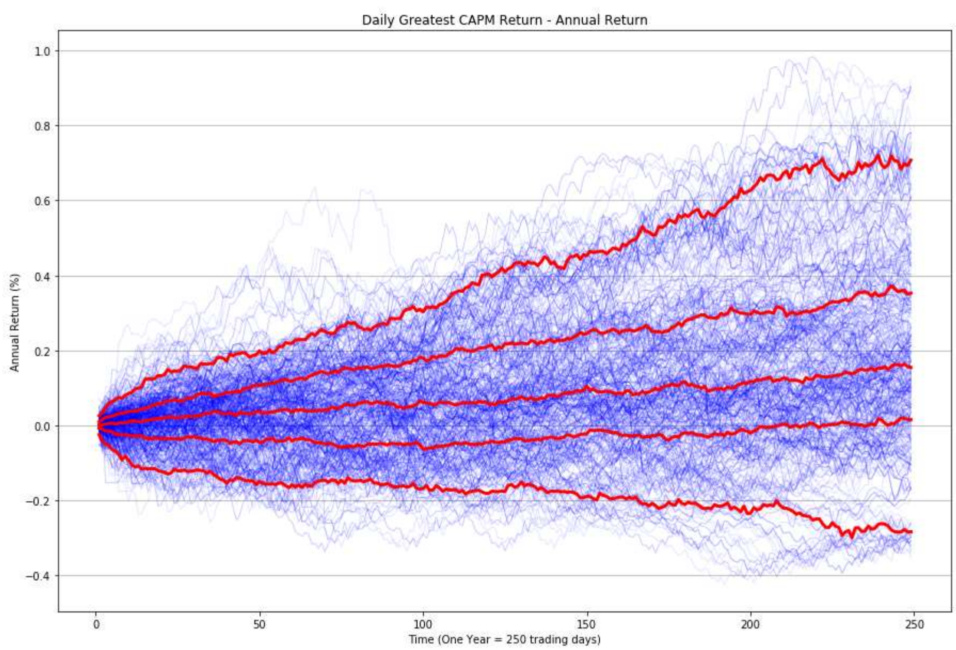

Fig6. The yearly return of a stock chosen on a random day based on its expected/required return calculated with CAPM. The blue lines are timeseries return of each individual stock. The red lines are the 5%, 25%, 50%, 75% and 95% confidence intervals of the returns.

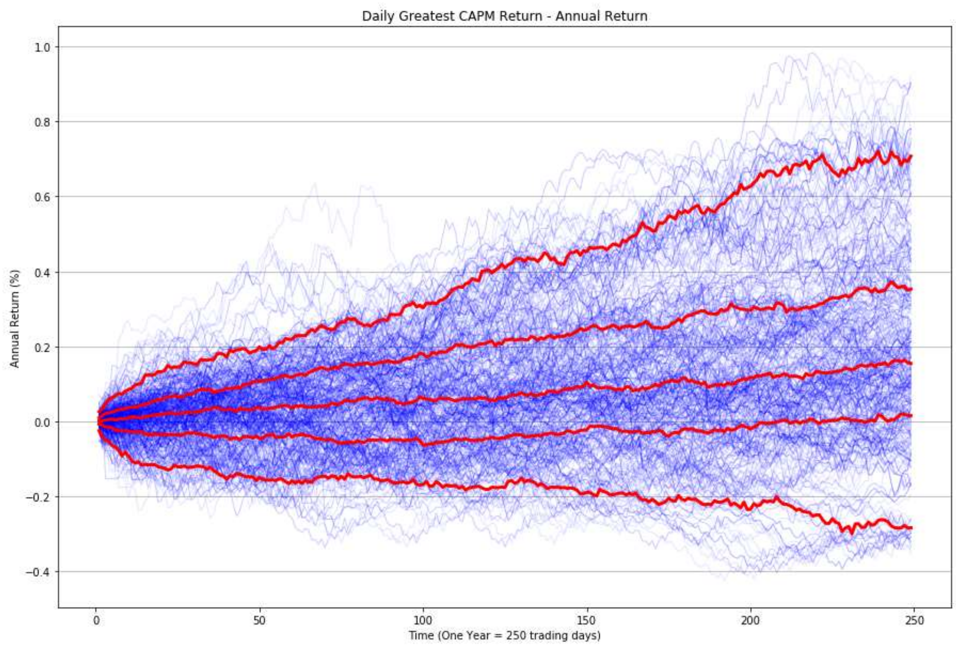

Fig7. The yearly return of a stock chosen on a random day based on its expected/required return calculated with CAPM. The blue lines are timeseries return of each individual stock. The red lines are the 5%, 25%, 50%, 75% and 95% confidence intervals of the returns.

Reference:

1. Corporate Finance (Fourth Edition), 2017, by Berk, J. and DeMarzo, P., Pearson
2. Derivatives Markets (Third Edition), 2013, by McDonald, R.L.
3. "Best Practices in Estimating the Cost of Capital: Survey and Synthesis," Financial Practice and Education 8 (1998) https://www.hbs.edu/faculty/Publication%20Files/Best%20Practices%20in%20Estimating%20the%20Cost%20of%20Capital%20Survey%20and%20Synethesis_e59fb55c-eeac-4abe-9ae9-04c4c623e8c3.pdf