Data Science and Retirement Planning: Inferential Statistics

Our exploratory data analysis provided sufficient information to conduct an initial hypothesis test. Taking expectations for All Weather Portfolio performance from Tony Robbin's Money: Mastering the Game, 7 simple steps to financial freedom, we can test the following hypothesis:

A well made all weather portfolio should average a return of 9.7% Cumulative Annual Growth Rate and lose 4.3% in its worst year (reinvesting dividends).

This hypothesis was tested by generating a sample space of 5000 random All Weather portfolios. The parameters that varied from portfolio to portfolio included fund names from each of the following categories: Stock Indices, Intermediate Term Bonds, Long Term Bonds, Gold, and Broad Basket Commodities. The amount of money added to the portfolio across the All Weather respective proportions (.30/.40/.15/.075/.075) per pay period was also a randomly generated amount between 100 and 1000 dollars. Lastly, the number of weeks between regular rebalancing of proportions was also randomly selected between 1 and 52 weeks.

Each random portfolio with its random contribution and rebalancing parameters was evaluated using Yahoo Finance's historical record, dating back to mid 2007. CAGR was calculated, along with the worst consecutive 52 week period (signifying worst year loss). Since the stock market crashed in 2008, an alternate timeline was also crafted by reversing all the date from each daily observation. This placed the crash towards the end of the 11 year period, since this would provide another way of evaluating an accumulated portfolio's health under the stress of the 2008 recession. CAGR was calculated given the alternate timeline, and the deeper of the two losses was recorded as the worst year on record.

The Results:

Only 4% of random portfolios had an expected CAGR over 9.7 percent
100% of the worst years calculated had losses exceeding 4.3 percent

The analysis implies that it is extremely unlikely the expected (mean) CAGR and loss values are 9.7 and 4.3 percent, respectively. Based on the analysis, it does not seem likely that this kind of portfolio performs as well as indicated.

Considerations:

1. A different selection of funds within each category can create discrepancies in the final sample space. If a poorly performing fund was selected to contribute to the sample space, it will affect every portfolio in which it is sampled. However, the diversification and approachability for non-seasoned investors was the point of All-Weather.

2. The numbers hypothesized were formed by an analysis which dated back to the inception of the NYSE. While such comprehensive data is available for some of the funds in this data, it is not available for many. This study is limited in that its timeline only goes back as far as the inception of the first Broad Basket Commodities in 2006.

While the initial hypothesis does not seem to be well supported by the data, ths exercise still provides some interesting results:

1. A bootstrap hypothesis test determined that there is no significant difference between those portfolios where the investor contributed $100 a pay period and those where the contribution was $1000. Thus, any investor can take advantage of the kinds of returns an All Weather Portfolio can obtain.

2. Further examination of the results provide an idea of which funds and combinations of funds provide the best returns. Frequencies were examined for which funds appear in those portfolios with CAGR above 9 and which appear in those portfolios which experienced losses in the smallest quartile of all losses. As a result, the following optimally performing portfolio was recommended:

   STOCK: VIGRX
   INTERMED BOND: IEF
   LONG TERM BONG: WHOSX
   GOLD: SGGDX
   COMMODITY: No significant fund is recommended
   Rebalanced every 40 weeks

The recommended portfolio experienced a CAGR of 9.99%, with a worst year's loss of - 9.33%. This is impressive, considering the normal losses experienced by the other All Weather portfolios - 10% to - 21%.

So how did this model fare? The S&P's 10 year average through 2018 was 8.46%, and it's biggest loss in 2008 was 36.55%. The portfolio most recommended by the data analysis delivers on the promise of keeping pace with the market and minimizing losses, albeit not as spectacularly as originally promised. It is a well balanced investment plan for an amatuer investor.

It is also worth noting that the software that has been developed to support this analysis can also be improved upon by considering alternate or additional categories of investment. There's no reason the funding for the stock index portion of the portfolio cannot be split up among stock indices and companies in promising sectors of the economy (like tech for example). Exactly how this should be done, and if such moves can be recommended by time series analysis is a matter for further study.

Tax penalties and fund fees have not been included in this model at this time
Past performance is not a guarantee of future perfromance