Financial Modeling
High-Frequency Trading Strategy Using Statistical Arbitrage

Authors: Divya Krishnan (ESE ’99)
Danish Murir (CIS ’99)

Advisors: Dr. Abraham J. Wyner (Dept. of Statistics, Wharton School), Dr. Insup Lee (Dept. of Computer & Information Science)

Abstract
Due to the rapid maturity of trading engines and the level sophistication of communications technology, high-frequency trading represents the ability to place trades that are only feasible by automated computer-based systems on high-speed electronic networks. The profitability of such trading strategies arises in the intraday options of stock prices.

The trading system created in this project analyzed the profitability of trading frequency throughout the day and comparing the results with retail market performance. With a user-defined minimum time interval, the algorithm monitors the movement of prices of the stocks with each time increment, it is determined if the stocks diverge from its equal-weighted position, and profitable trades are identified using the next price.

Given a user-defined portfolio, it was shown that the returns of a long-short strategy of two portfolios—one high-frequency traded portfolio and the second a low-frequency portfolio—outperform the second market. Furthermore, this strategy takes advantage of the scaling of portfolio statistics. It was shown that correlation between assets decreases to close to 0 as the minimum time interval between trades decreases. Thus, the returns of this strategy, or the mean, are as substantial as market returns but with a lower variance, or risk.

System Methodology
The fundamental lesson learned in modern portfolio theory is that diversification of assets can lead to greater returns at lower risk. Thus, a portfolio, P, is simply a linear model where the weights, pi, represent the amount invested in each asset X_i such that the returns for the overall portfolio is

\[ r = \sum_{i=1}^{n} r_i P_i \]  

High-frequency data incorporates observations of these returns on a finer time scale, from daily down to every tenth of a second. \( T \) is the time period interval, and \( n \) is the basic unit of time within \( T \), such that \( nT = T \), then it can be shown with basic statistical theory that the product of the individual returns is equivalent to the return over the entire time period \( T \). This is known as the first moment.

\[ \bar{r} = \frac{1}{n} \sum_{i=1}^{n} r_i \]  

A similar approach is used to find equivalency in the second moment. \( \bar{r} \). Unlike the first moment, it can be shown that the second moments are not equivalent to the effects of covariance are present, the term that explains the relationship between two assets. Many studies have been conducted that showing evidence that correlation and covariance increase as the scale of the time interval increases. This provides an arbitrage opportunity in the stock markets, where an investor can take advantage of price differentials that occur within assets.

The strategy employed here exploits the observation that returns diminish with longer time intervals between trades as the correlation increases. Two portfolios are held by an investor, each with identical assets. One portfolio is frequently traded, or rebalanced, within a given day. The second portfolio is rebalanced on a longer time interval such as every stock, or month. The less-frequently rebalanced portfolio is held in a long position.

Rebalancing Methodology

Performance
Using a sample strategy of two 2-asset portfolios each containing shares of different stocks, two different years were analyzed. The first portfolio contained of Alcoa Incorporated (NYSE: AA) and The Walt Disney Company (NYSE: DIS) for the year 1993. The second portfolio contained American Express Company (NYSE: AXP) and The Boeing Company (NYSE: BA).

Figure 2: The graph shows the relationship between the length of the minimum time interval to trade (in seconds, on the log scale) between the returns of the portfolios and the correlation between the two assets. Both portfolios, during the two different decades show this inverse relationship. From this graph, we can now show the profitability of higher frequency trading. Low correlation between assets is a desired trait in any investment strategy, using the statistics to negate the bad performance of one with the good performance of the other.

The Sharp ratio for the 2005 portfolios was 2.88 and 3.06 respectively for Strategy 1 and 2. In comparison, the Dow Jones Industrial Average, of whom all four stocks are a component, had a Sharpe ratio of 1.51 across the same time period, while the S&P had a ratio of 1.47. Similarly, calculating the Sharpe ratio for 1993 portfolios showed that Strategy 1 and Strategy 2 each had ratios of 4.25 and 5.02 respectively. Other calculations were also made including the information ratio, to show that Strategy 2 would always be able to achieve higher returns due to higher effect of leveraging.