Application and Statistical Inference of Financial Mathematical Models in Portfolio Optimization

Yutong Yang

Shanghai Guanghua Cambridge International School, Shanghai, China

3396498608@qq.com

Keywords: Financial mathematical models; Statistical inference; Portfolio optimization; Risk-reward trade-off

Abstract: Against the backdrop of intensified volatility in global capital markets, this study focuses on the application of financial mathematical models and statistical inference in portfolio optimization to address the imbalance between traditional empirical investment returns and risks. Firstly, this article provides a detailed overview of the basic frameworks of mean variance model, CAPM model, and APT model, analyzes these three types of models, and studies their quantitative roles in determining optimal asset allocation ratios, risk pricing guidance and portfolio adjustment, and multi factor perspective portfolio optimization; Furthermore, this article also elaborates that statistical inference provides reliable data support for model applications through parameter estimation, validity testing, and significance analysis, avoiding decision bias. Research has shown that combining the quantitative framework of financial mathematical models with the validation support of statistical inference can fundamentally enhance and improve the scientific and objective nature of investment decisions. However, it is still worth noting that the limitations of simplifying assumptions and historical data in this model must be recognized. This article hopes to optimize and improve the adaptability of the model through dynamic data updates in the future, so that the model can better adapt to complex markets.

1. Introduction

1.1. Research Background

In recent years, the global capital market has been affected by multiple factors, such as geopolitical conflicts, macroeconomic policy adjustments in major economies, industry cycle changes, and black swan events, as seen in Figure 1. This has led to a significant increase in the volatility of the global capital market, and the uncertainty of asset prices continues to rise. This further increases the complexity of asset allocation, and the limitations of traditional experiential investment models (which rely on investors' subjective experience and historical trend judgments) are becoming increasingly apparent. Its limitations mainly lie in the difficulty of comprehensively capturing various market risks, as well as the imbalance between returns and risks caused by delayed decision-making or emotional interference. Therefore, in extreme market fluctuations, there is a greater probability of significant losses.

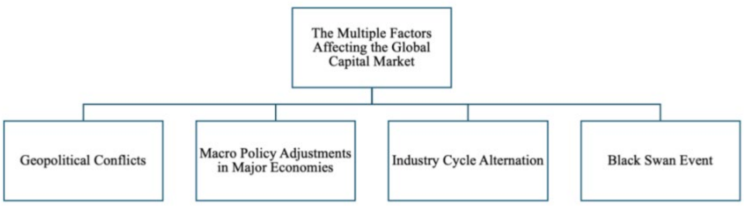

Figure 1: The multiple factors affecting the global capital market.

In the current context, financial mathematical models mainly solve problems through quantitative advantages, and quantification has become the key to solving problems: this model can transform abstract profit goals and risk levels into computable mathematical parameters, thereby outlining the trade-off relationship between profit and risk and becoming clear; Statistical inference, on the other hand, requires the use of historical data to test the rationality of model assumptions and verify the robustness of parameter estimates, thereby further avoiding the disconnection between the model and the actual market. The combination of the two has gradually become an essential tool for various financial institutions to optimize their investment portfolios [1].

The popularity of big data technology is constantly increasing, and algorithmic trading is also rapidly emerging. Therefore, the current market increasingly urgently needs real-time and accurate model driven investment decisions. However, traditional empirical decision-making is no longer suitable for the current needs of massive data processing and dynamic portfolio adjustment, which further promotes the continuous deepening of theoretical research and practical application of financial mathematical models and statistical inference in the field of investment portfolio optimization.

1.2. Research Significance

From a theoretical perspective, this study has a certain promoting effect on the deepening of the connection between financial mathematics and statistical inference. Financial mathematical models, such as mean variance and CAPM models, can provide quantitative frameworks for portfolio optimization, but the rationality of the models still relies on statistical inference and validation; Statistical inference (such as parameter estimation and hypothesis testing) can effectively supplement the limitations of a single model and further improve the theoretical system of quantitative investment, which will provide some basic ideas for subsequent research [2].

From a practical perspective, research results can bring returns to investors. Both ordinary individual investors and institutional investors can minimize the risk of subjective decision bias through model optimization and statistical verification. At the same time, investors can also develop asset allocation plans with clear risk boundaries, effectively achieving a balance between returns and risks, and further significantly improving the scientificity and operability of investment decisions.

1.3. Definition of Core Concepts

The definition of core concepts includes Financial Mathematical Models, Portfolio Optimization and Statistical Inference, as shown in Figure 2.

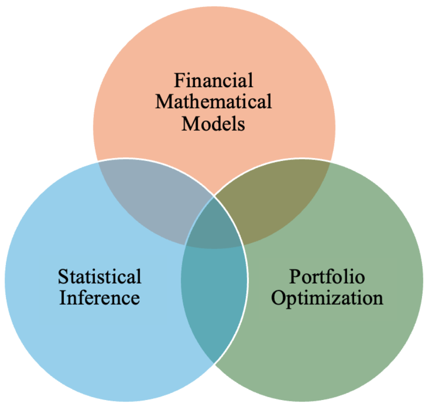

Figure 2: The definition of core concepts.

1.3.1. Financial Mathematical Models

Financial mathematical models, based on mathematical symbols, formulas, and logical frameworks, are important tools for describing the interrelationships between various variables in financial markets, such as asset prices, returns, risks, interest rates, etc. Its core function is to transform abstract financial phenomena into quantifiable and computable mathematical problems, and through the analysis of the inherent laws between variables, provide objective basis for subsequent research on asset value, evaluation of investment risks, and formulation of decision-making plans. Common financial mathematical models include mean variance models, capital asset pricing models (CAPM), arbitrage pricing models (APT), etc. These models simplify the market based on certain assumptions, but still effectively focus on core contradictions, helping users to clearly and accurately grasp the key logic of financial market operation [3].

1.3.2. Portfolio Optimization

Portfolio optimization, fundamentally speaking, refers to the process in which investors screen different types of assets (such as stocks, bonds, funds, etc.) based on their clear investment goals (such as return targets and risk tolerance), analyze and study the return levels, risk characteristics, and correlations between assets of each type, and then adjust and optimize the weight ratios of various assets in the investment portfolio, ultimately achieving a relative balance between returns and risks. Its essence is to balance risk and return, avoid the excessive impact of single asset volatility on the overall portfolio, and effectively achieve the goal of risk diversification and improving investment efficiency [4]. In addition, investors with different risk preferences (such as risk averse and risk prone) tend to choose different asset allocation plans, and they need to develop asset allocation plans that meet their own needs through screening and optimization.

1.3.3. Statistical Inference

Statistical inference is a statistical method based on the principles of probability theory, which infers overall characteristics (such as the overall pattern of future long-term returns of an asset) based on sample data (such as past returns of an asset). The core logic is to analyze local observation data, and then infer the overall data distribution characteristics, parameter sizes, and relationship patterns. It mainly includes two core contents: parameter estimation and hypothesis testing [5]. Parameter estimation is to calculate the estimated values of population parameters (such as mean and variance) based on sample data, and then determine the approximate range of population characteristics; Hypothesis testing is the process of verifying the validity of a hypothesis about the overall data (such as an asset's return following a normal distribution) based on sample data, in order to determine the reliability of the conclusion. In the financial field, statistical inference is mainly used to transform limited historical data into a basis for judging the overall laws of the market, inferring the whole from the small to the big and from the local. This provides favorable data support for financial decision-making and avoids significant errors caused by subjective experience judgments.

2. Basic Types of Financial Mathematical Models

There are three Basic Types of Financial Mathematical Models in this paper, namely Mean-Variance Model, Capital Asset Pricing Model (CAPM), and Arbitrage Pricing Model (APT), as demonstrated in Figure 3.

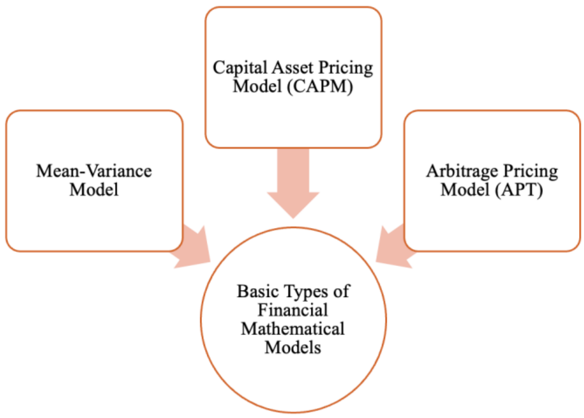

Figure 3: Basic types of financial mathematical models.

2.1. Mean-Variance Model

The mean variance model was proposed by Markowitz and is mainly applied in the field of portfolio optimization. It is one of the widely used basic models in this field. The core idea is to use the mean to determine the expected return of an investment portfolio, which is the weighted average of the returns of all assets within the portfolio, and can be used to reflect the average return level of investors in the future; This model uses variance (or standard deviation) to measure the risk of an investment portfolio, which is determined by the degree of volatility of returns around the mean, that is, the greater the volatility, the higher the risk [6].

This model assumes in advance that investors are risk averse, meaning they are more inclined to choose combinations with lower risk when facing the same returns, or to choose combinations with higher returns when facing the same level of risk. The core logic of this model is to construct an effective frontier by adjusting and optimizing the weights of various assets within the portfolio. This curve encompasses all the optimal combinations of risk and return, which can determine the combination with the highest return under a given risk and the lowest risk under a given return. Although the model is based on simplified assumptions such as returns following a normal distribution and rational investor decision-making, it is the first to combine returns with risk quantification, laying a theoretical framework for subsequent quantitative investments. At present, this model is still one of the important tools for building basic investment portfolios [7].

2.2. Capital Asset Pricing Model (CAPM)

Sharp, Lintner, and others analyzed the mean variance model and proposed the capital asset pricing model. The core of this model is to solve the problem of how assets are priced based on risk. This model divides asset risk into two categories: one is non systematic risk, which refers to the risk unique to a single asset (such as company operational risk), which can be eliminated through diversified investment; The other type is systemic risk, which refers to the risk that affects the entire market (such as economic cycles, interest rate changes). This type of risk cannot be solved through diversified investment, but it can bring additional returns compensation to investors.

The key indicator of this model is the β coefficient, which is used to measure the sensitivity of an asset to systemic risk: when $\beta=1$, the asset risk is the same as the overall market risk; β> At 1 o'clock, the asset risk is higher than the market, and the returns obtained by investors fluctuate greatly; When $\beta<1$, asset risk is lower than the market, and investors receive more stable returns. The core formula of this model is asset expected return=risk-free return+β×(market expected return risk-free return), where "market expected return risk-free return" is referred to as the market risk premium. Through this model, investors can determine whether asset pricing is reasonable (if expected returns are higher than the model's calculated value, the asset is undervalued), and can also choose assets with appropriate β coefficients for investment based on their own risk tolerance [8].

2.3. Arbitrage Pricing Model (APT)

The arbitrage pricing model was proposed by Ross, which is based on and extends the capital asset pricing model. The core logic of this model is that asset returns are determined by multiple independent macro or industry factors, rather than being solely determined by the overall market risk (such as the single factor of CAPM).

This model does not rely on strict assumptions such as normal distribution of returns and investor rationality, and is more in line with the complex and ever-changing real financial market. The core step is to first identify the key factors that affect asset returns, which typically include macroeconomic indicators such as interest rates, inflation rates, GDP growth rates, industry characteristics such as industry prosperity and policy support; Then use statistical methods to calculate the sensitivity of assets to each influencing factor (i.e. factor load), that is, the higher the sensitivity, the greater the impact of that factor on asset returns [9].

The return formula of this model can be simply understood as asset expected return=risk-free return+factor 1 sensitivity × factor 1 expected return+factor 2 sensitivity × factor 2 expected return+...+error term. Compared to CAPM's single factor logic, APT is more flexible and can analyze the influencing factors of investment risk from multiple dimensions, thereby helping investors optimize their investment portfolios more targetedly. For example, if interest rates are expected to rise, investors can reduce their holdings of assets with high sensitivity to interest rates (such as long-term bonds) and choose to increase their holdings of assets with low sensitivity (such as short-term monetary instruments), in order to more accurately control and reduce risks, and improve investment returns.

3. The Application of Financial Mathematical Models in Portfolio Optimization

The Application of Financial Mathematical Models in Portfolio Optimization includes Mean Variance Model, CAPM Model and APT Model, as seen in Figure 4.

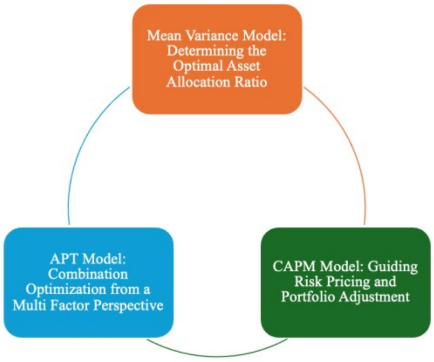

Figure 4: The application of financial mathematical models in portfolio optimization.

3.1. Mean Variance Model: Determining the Optimal Asset Allocation Ratio

The core role of the mean variance model in portfolio optimization is to calculate the optimal weight ratio of various assets by quantifying returns and risks. Therefore, when applying this model, firstly, it is necessary to collect historical return data of candidate assets within the portfolio, and determine the expected return of each asset class by calculating the sample mean. It is also necessary to calculate the sample variance to determine the risk of a single asset class. At the same time, the covariance (or correlation coefficient) between assets is also calculated. A smaller covariance suggests a stronger correlation between changes in asset returns and a better benefit of risk diversification.

The model will build a return risk goal function using the aforementioned data, and by solving the function, it will determine the effective frontier. Specifically, each point on the efficient frontier corresponds to a set of asset weights, representing the highest return at a specific risk level or the lowest risk at a specific return level. Investors can choose suitable investment allocation plans based on their own risk tolerance (such as risk averse individuals tending towards low volatility and risk averse individuals tending towards high returns) through effective frontier analysis, ultimately determining the optimal proportion of various assets such as stocks, bonds, and funds, and effectively achieving a balance between returns and risks.

3.2. CAPM Model: Guiding Risk Pricing and Portfolio Adjustment

In portfolio optimization, the application of CAPM model is mainly for risk pricing judgment and portfolio structure adjustment. Firstly, the model uses the β coefficient to help investors screen for assets that are suitable for themselves. For investors with lower risk tolerance (such as conservative individual investors), they can prioritize assets with β less than 1 (such as utility stocks and short-term bonds), which are less affected by market fluctuations and have more stable returns; For investors with high risk tolerance (such as institutional investors), it is advisable to allocate assets with a β greater than 1 (such as growth stocks and industry ETFs) appropriately to obtain higher risk returns.

This model can be used to determine the rationality of asset pricing: the theoretical expected return of an asset is calculated through the core formula of the model: if the actual market return of an asset is higher than the theoretical value, it indicates that the asset is undervalued, and investors can increase their holdings of the asset; If the actual return is lower than the theoretical value, it indicates that the asset is overvalued, and investors need to reduce their holdings of the asset. Investors can also adjust and optimize the overall β coefficient of the investment portfolio by adjusting the weights of each asset in the portfolio (such as reducing the proportion of high β assets and increasing the proportion of low β assets to reduce portfolio risk), thereby ensuring that the risk of the investment portfolio is in line with their own goals.

3.3. APT Model: Combination Optimization from a Multi Factor Perspective

The main advantage of the APT model lies in analyzing multiple factors, which can provide investment portfolio optimization with adjustment and optimization ideas that are more in line with the current actual market. When applying this model, the first step is to identify the key influencing factors of asset returns within the portfolio. Common factors include macroeconomic factors (such as interest rate changes, inflation rates, GDP growth rates) and industry factors (such as policy support, industry supply and demand relationships), all of which directly affect the trend of future asset returns.

The second step is to calculate the sensitivity (factor load) of various assets to different factors through statistical methods: if the sensitivity of an asset to the rising interest rate factor is negative, it indicates that the asset's return may decrease when the interest rate rises, and special attention should be paid to the asset; If the sensitivity of an asset to favorable industry policies is positive, it indicates that the asset's returns may increase with policy support, and the asset can be selected as a key allocation target [10].

Investors adjust asset weights based on their judgment of the future trends of various factors: if they judge that expected interest rates will rise, they will reduce their holdings of interest rate sensitive assets (such as long-term bonds) and increase their holdings of less sensitive assets (such as money market funds); If it is expected that a certain industry policy will be favorable, the proportion of highly sensitive assets in that industry will be increased. By making multidimensional adjustments through judgment, investors can not only respond to various risks in a targeted manner, but also capture the profit opportunities brought by different factors, comprehensively improving the accuracy of portfolio optimization.

4. The Application of Statistical Inference in Portfolio Optimization

4.1. Application of Statistical Inference in Asset Allocation Parameter Estimation

In the asset allocation stage of portfolio optimization, the core role of statistical inference is to accurately estimate key parameters through sample data, thereby providing reliable data support for tools such as mean variance models. The key parameters that asset allocation relies on mainly include expected returns, risks, and asset correlations. These parameters cannot directly obtain overall data and need to be inferred through historical return samples, such as estimating the long-term expected returns of assets using sample means, estimating the risk level of assets themselves using sample variances, and estimating the degree of correlation between return fluctuations of different assets using sample variances.

The confidence interval method in statistical inference can further improve the reliability of parameters: by calculating the confidence interval of parameters (such as 95% confidence interval), the reasonable fluctuation range of parameters can be delineated to avoid configuration bias caused by accidental fluctuations in a single sample data. For example, if the average return of a stock sample is 8% and the 95% confidence interval is 6%-10%, investors can make a rough judgment on the fluctuation range of the stock's return based on this, and will not overly rely on the single point estimate of 8% when allocating. This parameter estimation based on statistical inference can make asset allocation more objective and fair, reducing portfolio risk caused by data bias.

4.2. Application of Statistical Inference in Validity Testing of Risk Pricing

Statistical inference may be used to verify the validity of risk pricing tools like CAPM models and subsequently change the investment portfolio based on the results. Hypothesis testing and goodness of fit analysis are two of the main ways that statistical inference works. For example, hypothesis testing may check the importance of crucial risk pricing indicators, such whether the beta coefficient in CAPM is substantially non-zero. If the p-value of the beta coefficient is less than 0.05 after t-test, it indicates that the systematic risk of the asset has a significant impact on returns, and also demonstrates the effectiveness of the pricing logic of the model; If the P value is greater than 0.05, it indicates that the β coefficient cannot explain the changes in returns, and the model needs to be adjusted or assets replaced.

The goodness of fit analysis (such as R^2 index) can determine the explanatory power of the model on actual returns: the higher the R^2, the stronger the explanatory power of risk factors in the model (such as market risk in CAPM) on returns, and the higher the credibility of pricing results; If R^2 is too low, it indicates the presence of uncaptured risk factors and requires a reassessment of the asset situation. Through these tests, investors can screen out assets that are priced effectively, eliminate assets that cannot be explained by the model, and ensure that the risk pricing of the portfolio can rely on reliable logic to avoid the loss of returns caused by the allocation of assets with pricing failures.

4.3. Application of Statistical Inference in Significant Analysis of Multi-factor Influence

Statistical inference, which employs significance testing, may be utilized to discover components that are helpful for the multi-factor optimization of the APT model. This method prevents factors that are not of significance from interfering with the selection of a combination. To begin, determine the significance of each component by doing a t-test. As an example, you may get the t-statistic for each element that has an impact on the result, such as the rates of inflation or the rates of interest. If the absolute value of the t-statistic is greater than the critical value, which is 1.96 in this case, it indicates that the factor has a significant impact on asset returns and need to be included in the model. If it is lower than the critical value, it indicates that the relationship between components and returns is tenuous, and you might remove it in order to make the model more straightforward to comprehend.

The next stage involves the use of the F-test in order to determine if the multi-factor combination is significant when seen as a whole. This entails determining whether or not assembling all of the components has a significant impact on returns. The combination of variables that you selected is capable of providing a satisfactory explanation for the variations in returns in the event that the F-test P value is less than 0.05. This demonstrates that the model functions as a whole entity. The AIC, BIC, and other information that is utilized in statistical inference may be used to determine the combination of elements that is most optimal. As an example, you are able to determine which combination of factors has the lowest value by using the AIC or BIC values of various combinations of factors. This will assist you in ensuring that the model is a good match for the data while also preventing overfitting. The APT model is able to concentrate on the most significant components of the system as a result of this significance analysis, which makes multidimensional combinatorial optimization much more precise and reduces the amount of decision interference that is generated by aspects that are not effective.

5. Conclusion

The use of financial mathematical models and statistical inference in the process of portfolio optimization is the main emphasis of this research project. The following conclusions are the most important ones: The mean variance model, the CAPM model, and the APT model are quantitative frameworks that provide solutions to the problem of traditional empirical decision-making in that it is difficult to accurately control risk and return. These models optimize portfolios from the perspectives of return and risk balance, single factor risk pricing, and multi factor impact analysis, respectively. Statistical inference provides reliable support for model application through parameter estimation, validity testing, and significance analysis, ensuring reasonable asset allocation parameters, effective risk pricing, and accurate multi factor screening. The scientific and objective quality of investment choices is much improved when the two are used in conjunction with one another.

However, it is also important to consider the limits of research. For instance, the simplified assumptions that are made by financial mathematical models, such as the assumption that returns follow a normal distribution, may not always reflect the reality of the market. In addition, statistical inference depends on data from the past, which is not always an accurate predictor of volatility in the future. In the future, the combination of dynamic data updates and developing technologies will be able to be further coupled to maximize model adaptability, which will allow quantitative tools to better adapt to the complex and ever-changing capital markets, and it will also provide investors with more accurate portfolio optimization options.

References

[1] Taniguchi, M., Hirukawa, J., & Tamaki, K. (2007). Optimal Statistical Inference in Financial Engineering. Chapman & Hall/CRC.

[2] Cai, T. T., Guo, Z., & Ma, R. (2023). Statistical inference for high-dimensional generalized linear models with binary outcomes. Journal of the American statistical association, 118(542), 1319-1332.

[3] Bao, Z., Ding, X., Wang, J., & Wang, K. (2022). Statistical inference for principal components of spiked covariance matrices. The Annals of Statistics, 50(2), 1144-1169.

[4] Han, R., & Zhang, A. R. (2023). Rungang Han and Anru R. Zhang's contribution to the Discussion of 'Vintage factor analysis with varimax performs statistical inference' by Rohe & Zeng. Journal of the Royal Statistical Society Series B: Statistical Methodology, 85(4), 1069-1070.

[5] Chen, E. Y., & Fan, J. (2023). Statistical inference for high-dimensional matrix-variate factor models. Journal of the American Statistical Association, 118(542), 1038-1055.

[6] Aydınhan, A. O., Kolm, P. N., Mulvey, J. M., & Shu, Y. (2024). Identifying patterns in financial markets: extending the statistical jump model for regime identification. Annals of Operations Research, 1-37.

[7] Chen, X., Lai, Z., Li, H., & Zhang, Y. (2024). Online statistical inference for stochastic optimization via kiefer-wolfowitz methods. Journal of the American Statistical Association, 119(548), 2972-2982.

[8] Pham, M., Simar, L., & Zelenyuk, V. (2024). Statistical inference for aggregation of Malmquist productivity indices. Operations Research, 72(4), 1615-1629.

[9] Wasserbacher, H., & Spindler, M. (2022). Machine learning for financial forecasting, planning and analysis: recent developments and pitfalls. Digital Finance, 4(1), 63-88.

[10] Strielkowski, W., Guliyeva, A., Rzayeva, U., Korneeva, E., & Sherstobitova, A. (2021). Mathematical modeling of intellectual capital and business efficiency of small and medium enterprises. Mathematics, 9(18), 2305.