Application of ChebyshevInequality and Central Limit Theorem in Financial Field and Comparison of their Advantages and Disadvantages

Yulin Huang

Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, China

1703944074@qq.com

Abstract. This paper discusses the application of Chebyshev inequality and Central limit theorem in the field of finance and compares their advantages and disadvantages. Through theoretical analysis and empirical research, this paper first introduces the basic principles and derivation process of the two statistical methods, and then applies them to the actual financial data analysis. In this study, the data of stock return rate is selected as a sample, which is estimated by Chebyshev inequality and central limit theorem respectively, and compared with the actual distribution. The results show that Chebyshev inequality provides a more conservative but reliable estimate when the distribution of financial data is unknown, while the central limit theorem provides a more accurate approximation when the sample size is large enough. This study provides important theoretical support and practical guidance for financial risk management, investment decision making and pricing of financial derivatives.

Keywords: Chebyshev inequality; Central limit theorem; Financial risk management; Investment decision; Pricing of financial derivatives; Statistical estimation

1 Introduction

In the rapidly evolving financial markets, accurate risk assessment and return prediction are paramount for investment decision-making and risk management. Traditional financial models often rely on distributional assumptions (e.g., normality), which may fail under extreme market conditions or limited data availability. Chebyshev's inequality and the Central Limit Theorem (CLT) offer distinct approaches to address these challenges. Chebyshev's inequality provides distribution-free bounds for deviations from the mean, making it robust for non-normal or heavy-tailed financial data. Conversely, the CLT enables precise asymptotic approximations for large-sample scenarios, facilitating efficient estimation of portfolio returns and risk metrics.

This paper systematically examines the theoretical foundations, empirical applications, and comparative advantages of these two statistical tools in finance. By analyzing S&P 500 stock returns, we demonstrate their efficacy in estimating tail risks, portfolio

P. S. Borah et al. (eds.), *Proceedings of the 2025 5th International Conference on Enterprise Management and Economic Development (ICEMED 2025)*, Advances in Economics, Business and Management Research 346,
https://doi.org/10.2991/978-94-6463-811-0_110

distributions, and derivative pricing. Our findings highlight context-dependent trade-offs: Chebyshev's conservatism suits extreme-risk scenarios, while the CLT excels in large-sample regimes. This study bridges theoretical probability theory with practical financial analytics, offering actionable insights for investors and risk managers.

2 The Significance of the Subject and the Review of Research Status at Home and Abroad

Chebyshev inequality and central limit theorem play an important role in probability estimation of financial problems. These two theorems provide financial analysts and investors with powerful tools for understanding and predicting market behavior, risk management, and investment decisions. Chebyshev's inequality is often used to estimate the probability that a random variable falls within a certain interval near its mean. Specifically, if the expectation of the random variable X is E(X) and the variance is D(X), then for any positive number ε, the inequality states that the probability that X falls outside the interval $[E(X)-\varepsilon, E(X)+\varepsilon]$ is no greater than $D(X)/\varepsilon^2$. In finance, Chebyshev's inequality can be used to estimate the extent to which asset prices, yields, or other financial indicators have deviated from their historical mean or expected value[3]. For example, if the average price and standard deviation of a certain stock are known, the Chebyshev inequality can be used to estimate the probability that the stock price will deviate from a certain range of the average price over a period of time in the future[6]. The central limit theorem is an important theorem in probability theory, which states that when the number of independent and equally distributed random variables is large enough, the distribution of the sum of these random variables will converge to a normal distribution. In finance, this theorem has a wide range of applications because it allows us to estimate the population distribution of a large number of random events using properties of a normal distribution[7][9]. Specifically, in financial analysis, the central limit theorem can be used to estimate the risk and return of an investment portfolio. Assume that a portfolio consists of multiple independent assets, that the return on each asset is a random variable, and that these random variables are independently and identically distributed. According to the central limit theorem, when the number of assets is large enough, the overall return of a portfolio will converge to a normal distribution[12]. This allows investors to use the properties of the normal distribution to calculate key metrics such as a portfolio's confidence interval, value at risk[5], and so on. This paper tries to find out the applicability of chebyshev inequality and Levy central limit theorem in probability estimation of financial problems under known distribution and known expectation and variance scenarios.

Chebyshev inequality was first proposed by the Russian mathematician Chebyshev in the 19th century. As an important inequality in probability theory, Chebyshev inequality provides a powerful tool for understanding and estimating the probability distribution of random variables. In the field of finance, the application of Chebyshev's inequality was initially limited to theoretical research and academic discussion. Financial scholars and mathematicians began to recognize the potential applications of the inequality in risk assessment, asset pricing, etc., and began to try to introduce it into financial models and analysis. In finance, it is often used in aspects such as risk management, portfolio optimization, and asset pricing. In risk management, chebyshev's inequality can help financial institutions and investors estimate the likelihood that asset prices or yields will deviate from their mean (Lepetit Laetitia et al., 2020)[6][17]. By setting different degrees of deviation (i.e., k values), it is possible to calculate the probability that prices or yields fall within different ranges, thereby quantifying risk. This approach allows financial institutions to more precisely assess the level of risk in their portfolios and develop risk management strategies accordingly[14]. In terms of portfolio optimization, when constructing a portfolio, investors need to consider the correlations and risks between different assets. Chebyshev's inequality can be used to estimate the overall risk level of a portfolio (Hannan Amoozad Mahdiraji et al., 2018) [10][18], which is the probability that a portfolio's return will deviate from its expected value[8]. By optimizing the weighting of various assets in a portfolio, investors can pursue higher returns while controlling risk[11] [15]. In terms of asset pricing, asset pricing is a core issue in financial markets. Chebyshev inequality can be used to estimate the volatility range of asset prices[13] [16], thus providing reference for asset pricing. For example, in option pricing, Chebyshev's inequality can be used to estimate the range in which the price of the underlying asset will fluctuate at some point in the future, and thus calculate a reasonable price for the option.

The central limit theorem is an important theorem in probability theory, which states that when the sample size is large enough, the distribution of the sample mean will converge to a normal distribution. This theorem has a wide range of applications in the financial field. In terms of portfolio return distribution, in portfolio management, the return of a portfolio is often affected by many factors, such as market fluctuations and industry trends. According to the central limit theorem, when the number of assets in a portfolio is large enough, its return distribution will converge to a normal distribution. This allows investors to use the properties of the normal distribution to estimate the future return distribution of a portfolio and formulate an investment strategy accordingly. In terms of Risk measurement, VaR (Value at Risk) is a commonly used risk measurement tool in financial risk management (Festus Fatai Adedoyin et al., 2024)[1]. It represents the maximum loss a portfolio can suffer in a given confidence level and time frame. According to the central limit theorem, the VaR value of a portfolio can be estimated by calculating the mean and standard deviation of its return (Hoga Yannick, 2019) [7]. This approach allows financial institutions to more precisely assess the level of risk in their portfolios. In terms of asset pricing, the central limit theorem also plays an important role in the field of asset pricing. For example, in the CAPM (Capital Asset Pricing Model) model (Anna Rutkowska-Ziarko et al., 2024) [2], it is assumed that all investors hold the optimal proportion of the market portfolio and that the return distribution of the market portfolio follows a normal distribution. This allows the CAPM model to use the central limit theorem to estimate the equilibrium rate of return and risk premium of an asset.

As two important tools in probability theory, Chebyshev inequality and central limit theorem play an important role in probability estimation and risk management in the financial field. By deeply studying the application status and potential value of these two theories, we can provide financial institutions and investors with more accurate and effective risk management tools and decision support. This study will try to further expand the application fields of these two theories and optimize the relevant models to better serve the stability and development of the financial market.

3 Theoretical Basis of Chebyshev's Inequality and Central Limit Theorem

3.1 Proof and Generalization of Chebyshev's Inequality

Chebyshev's inequality can be expressed as follows: $P(|Y-EY|\geq x) \leq \frac{VarY}{X^2}$

The Chebyshev inequality in the general case can be expressed as follows: $P(|Y|\geq x) \leq \frac{E|Y|^r}{x^r}$

In fact, we have the following promotion: g(x) is a non-degenerate, non-negative function on $[0, +\infty)$. If for a random variable Y, there is $Eg(|Y|) < \infty$, Then for any $a > 0$ such that $g(a) > 0$, all have

$$P(|Y|\geq a) \leq \frac{Eg(|Y|)}{g(a)}$$

By $P(|Y|\geq a) = EI(|Y|\geq a) \leq EI(g(|Y|)\geq g(a)) \leq E\{\frac{g(|Y|)}{g(a)} I(g(|Y|)\geq g(a))\} \leq \frac{Eg(|Y|)}{g(a)}$

Proof conclusion[4].

3.2 Proof of the Central Limit Theorem

Levy Central Limit theorem: Let {X, N} be a X_n, n $\in$sequence of i.i.d random variables, then when $X\in L_2$, EX=a, $0<VarX=\sigma^2<\infty$, there is

$\frac{S_n-na}{\sqrt{n}\sigma} \xrightarrow{d} N(0,1)$

Proof: Remember $f_n(t)=E\exp\{it\frac{S_n-na}{\sqrt{n}\sigma}\}$,then

$f_n(t)=\prod_{k=1}^{n} E\exp\{it\frac{X_k-na}{\sqrt{n}\sigma}\}=(E\exp\{it\frac{X-a}{\sqrt{n}\sigma}\})^n$

Since $X\in L_2$, $E\frac{X-a}{\sqrt{n}\sigma}=0, E(\frac{X-a}{\sqrt{n}\sigma})^2 = \frac{VarX}{n\sigma^2}=\frac{1}{n}$

Therefore, by the Taylor expansion of the eigenfunction at t=0, for any $t\in R$, there is

$$E\exp\{it\frac{X-a}{\sqrt{n}\sigma}\}=1-\frac{t^2}{2n}+o(\frac{t^2}{n}), n\rightarrow\infty$$

Thus, $f_n(t)=(1-\frac{t^2}{2n}+o(\frac{t^2}{n}))^n \rightarrow \exp\{-\frac{t^2}{2}\}$, $n\rightarrow\infty$,

In fact, for independent different distributions, we have the Lindeberg central limit theorem, which is not elaborated here.

4 Research Methods and Data Collection

This research selects the daily return rate data of the constituent stocks of the US S&P 500 index from 2010 to 2020 as the research sample.

The data is sourced from Yahoo Finance and Bloomberg terminals and includes a total of about 2,500 trading days. In order to ensure that the data is representative and reliable, we have excluded stocks that were removed from the index during this period and processed the missing data appropriately.

The data processing process includes: first, calculating the daily logarithmic return rate of each stock; Then, the yield data of all stocks are classified by industry into 11 industry groups; Finally, a descriptive statistical analysis was carried out on the data of each industry group to calculate the sample mean and standard deviation of each group. In order to compare the performance under different market conditions, we also divided the data into two sub-samples, high volatility and low volatility, based on market volatility.

In our empirical analysis, we first estimate the extreme volatility of stock returns using Chebyshev's inequality. Take the information technology industry as an example, which has a sample mean of 0.0008 and a standard deviation of 0.015. Let's say we focus on situations where yields deviate from the mean by more than 3 standard deviations, according to Chebyshev's inequality:

$P(| X-\mu | \geq 3\sigma) \leq 1/3^2 = 0.1111$

This means that there is no more than an 11.11% chance that the yield will deviate from the mean by more than 4.5%. The observed probability of 8.76% is slightly lower than the upper bound given by Chebyshev's inequality, verifying the validity of the inequality in financial risk management.

Next, we apply the central limit theorem to estimate the average rate of return of the industry. Taking the financial industry as an example, suppose we randomly select 100 stocks to calculate their average return. According to the central limit theorem, the distribution of the sample mean approximates a normal distribution with a standard deviation of 0.0018. We can use this property to calculate the probability within a particular interval. For example, the probability that the sample mean falls within the [0.0005, 0.0011] interval is about 68.27%, which is very close to the actually observed 69.12%, showing the utility of the central limit theorem in financial data analysis.

In order to further compare the pros and cons of the two methods, we perform the following analysis: First, Chebyshev's inequality provides a conservative but reliable estimate when the distribution is unknown, while the central limit theorem requires a larger sample size to provide an accurate approximation. Second, the central limit theorem usually provides a more accurate estimate when the distribution and expectation and variance are known, while the Chebyshev inequality may be too conservative. Finally, we observe that Chebyshev's inequality performs better than the central limit theorem during periods of high volatility, while the two perform equally well during periods of low volatility.

5 Conclusion

Through theoretical analysis and empirical research, this paper discusses the application of Chebyshev inequality and Central limit theorem in the financial field and compares their advantages and disadvantages. The results show that these two statistical methods are of great value in financial data analysis and risk management. Because of its universality and conservatism, Chebyshev's inequality performs well when the distribution is unknown or the sample size is small, and is especially suitable for the estimation of extreme risk events. The central limit theorem can provide a more accurate approximation under large sample conditions, which is suitable for routine financial data analysis and risk assessment of large-scale investment portfolios.

In practical application, financial practitioners should choose the appropriate method according to the specific problem and data characteristics. For conservative investors or risk managers, Chebyshev's inequality may be more appropriate; While for quants or high-frequency traders, the central limit theorem may be more advantageous. Future research could further explore the application of these two methods in more complex financial environments, such as considering factors such as market interactivity and asymmetric distribution, to improve the accuracy of financial risk management and investment decisions.

References

1. Adedoyin F F, Bekun V F. Retraction Note: Modelling the interaction between tourism, energy consumption, pollutant emissions and urbanization: renewed evidence from panel VAR. [J]. Environmental science and pollution research international, 2024,
2. Ziarko – R A, Markowski L, Abdou A H. Conditional CAPM relationships in standard and accounting risk approaches [J]. North American Journal of Economics and Finance, 2024, 72, 102123.
3. Reza F, Vadim P A Statistical Proof of Chebyshev's Sum Inequality[J]. The American Mathematical Monthly, 2024, 131 (3): 224-224.
4. Chen H, Zhou Y, Mei K, et al. An Improved Density Peak Clustering Algorithm Based on Chebyshev Inequality and Differential Privacy [J]. Applied Sciences, 2023, 13 (15).
5. Andreas B Interest rate risk in the banking book: A closed-form solution for non-maturity deposits [J]. Journal of Banking and Finance, 2021, 125.
6. Laetitia L, Frank S, Ha T T An alternative Z-score measure for downside bank insolvency risk [J]. Applied Economics Letters, 2021, 28 (2): 137-142.
7. Yannick H Limit Theory for Forecasts of Extreme Distortion Risk Measures and Expectiles* [J]. Journal of Financial Econometrics, 2019, 20 (1): 18-44.
8. Science Operations Science; Study Data from University of Tehran Update Understanding of Operations Science (Bi-objective mean-variance method based on Chebyshev inequality bounds for multi-objective stochastic problems)[J]. Science Letter, 2019.
9. Sirignano A J, Giesecke K. Risk Analysis for Large Pools of Loans. [J]. Management Science, 2019, 65 (1): 107-121.
10. Mahdiraji A H, Hajiagha R H S, Hashemi S S, et al. Bi-objective mean–variance method based on Chebyshev inequality bounds for multi-objective stochastic problems [J]. RAIRO - Operations Research, 2018, 52 (4-5): 1201-1217.

11. Agahi H. Multivariate Chebyshev inequality in generalized measure space based on Choquet integral [J]. Communications in Statistics - Theory and Methods, 2017, 46 (19): 9625-9628.
12. Stellato B, Parys V G P B, Goulart J P. Multivariate Chebyshev Inequality With Estimated Mean and Variance [J]. The American Statistician, 2017, 71 (2): 123-127.
13. Budny K. An extension of the multivariate Chebyshev's inequality to a random vector with a singular covariance matrix [J]. Communications in Statistics - Theory and Methods, 2016, 45 (17): 5220-5223.
14. Statistics; New Findings on Statistics from Columbia University Summarized (Markov's Inequality and Chebyshev's Inequality for Tail Probabilities: A Sharper Image) [J]. Journal of Technology & Science, 2015.
15. Johnson T. On the support of the simple branching random walk [J]. Statistics and Probability Letters, 2014, 91 107-109.
16. Bi T, Zhang B, Wu H. Measuring Downside Risk Using High-Frequency Data: Realized Downside Risk Measure [J]. Communications in Statistics - Simulation and Computation, 2013, 42 (4): 741-754.
17. Yadav D, Singh S, Kumari R. Application of minimax distribution free procedure and Chebyshev inequality for backorder discount inventory model with effective investment to reduce lead-time and defuzzification by signed distance method [J]. Int. J. of Operational Research, 2012, 15 (4): 371-390.
18. Sellke M T, Sellke H S. Chebyshev Inequalities for Unimodal Distributions [J]. The American Statistician, 2012, 51 (1): 34-40.