Short Communication

An Empirical Study on the Logarithmic Return of Securities

Xueping Liu¹, Haihua Li¹*, Yingying Cao¹, Xuege Han¹

¹School of Mathematics and Information Science, Henan Polytechnic University, Jiaozuo, Henan Province, China
*Correspondence to: Haihua Li, School of Mathematics and Information Science, Henan Polytechnic University, No.2001 Century Street, Shanyang District, Jiaozuo 454003, Henan Province, China; E-mail: haihli@163.com

Received: May 5, 2023 Revised: June 12, 2023 Accepted: July 12, 2023 Published: August 8, 2023

Abstract

Objective: Previous studies have shown that the distribution of security returns has the characteristics of peak and fat tail, and it is not realistic to use logarithmic normal distribution to describe security returns. This paper mainly studies the specific form of security returns distribution and conducts empirical research.

Methods: We first assume that the security returns follow a logarithmic $t$-distribution. Then, the maximum likelihood estimation and moment estimation are used to explore the degree of freedom of the distribution. Finally, the goodness of fit test is used to conduct an empirical analysis about Shanghai Securities Composite Index.

Results: The specific expression formulas for the estimator of the degree of freedom are obtained. Empirical testing shows that the security returns follow a logarithmic $t$-distribution.

Conclusion: The probability distribution of security returns is unrealistic to assume that the security returns follow the logarithmic normal distribution. It is more practical to use the logarithmic $t$-distribution to characterize the distribution of security returns. Our results provide theoretical assistance for exploring the pricing of financial derivatives.

Keywords: securities returns, maximum likelihood estimation, hypothesis test, goodness of fit

1 INTRODUCTION

Option pricing is one of the hottest topics in the field of financial engineering research. In order to price options reasonably, it is necessary to first explore the distribution of securities returns. Black and Scholes deduced the European option pricing formula under the assumption that the security yield follows the logarithmic normal distribution. However, empirical analysis shows that the distribution of securities returns has a “peak thick tail” characteristic. Compared with
the standard normal distribution, the skewness of the standard normal distribution is 0 and the kurtosis is 3. In the empirical analysis, the financial data does not conform to the normal distribution, that is, the kurtosis ratio is 3, and the tails on both sides are thicker than the normal distribution. Because the logarithmic normal distribution lacks this property, using it to characterize the distribution of securities return rates does not adhere to objective reality.

Currently, there are many literature studies on the distribution of securities returns and other related issues derived from it. Hussain and Li[3] considered how the Shanghai Stock Exchange Composite Index (SSECI)’s excessive daily returns are distributed. According to their findings, the GL distribution and the GEV distribution are better fits for the minima and maxima series of returns for the Chinese stock market, respectively. Gebka and Woha[4] examined the DJIA index returns’ ability to forecast future return distribution when examined at various quantiles of its distribution. Watorek et al.[5] studied the price return distributions of cryptocurrency prices, currency exchange rates, and contracts for differences that reflect commodities, stock indexes, and stock shares. They used power-law, stretched exponential, and q-Gaussian functions to represent the tails of the return distributions at various time scales based on recent data from 2017-2020. By employing an augmented quantile auto-regression method, Jin et al.[6] found that responses to COVID-19 differ across quantile levels of return distributions in any particular country.

The distribution characteristics of the underlying return on assets have an important impact on the pricing of financial derivatives similar to options. Scholars have done a lot of research on various types of options, but most of it is based on the assumption that the return rate of securities follows the logarithmic normal distribution. For example, Aggarwal and Aggarwal[7] found that the security returns of the NYSE, AMEX, and NASDAQ from 1974 to 1988 significantly deviated from the normal distribution, and the deviation of the NYSE and AMEX security return distributions from normality decreases with increasing portfolio size and investment horizon. Basnarkov et al.[8] developed a discrete-time framework for pricing European-style options. The results indicated that the correlation of the pricing program they studied has been empirically verified, and the values calculated by their design scheme matched very well with actual market data. Agrawal and Hu[9] obtained the existence, uniqueness and positivity of solutions of delayed stochastic differential equation with jumps. Then, this equation was applied to the price fluctuation model of risky assets in the financial market, and the Black Scholes formula for European option prices and hedging investment portfolios was obtained. In order to price options more reasonably, it is necessary to propose more reasonable assumptions about the distribution of securities returns. Cassidy et al.[9] proposed the assumption that securities returns follow a logarithmic t-distribution. Based on this, they derived the corresponding option pricing formula. However, they did not demonstrate the rationality of this hypothesis through theoretical or empirical analysis. This paper first assumes that the return rate (e.g., the return rate of SSECI) follows a logarithmic t-distribution, and then calculates the corresponding degree of freedom using the maximum likelihood estimation method and the moment estimation method. Finally, we test the goodness of fit of the overall distribution, and then discuss the rationality of the stock return rate following the logarithmic t-distribution.

2 METHODS

2.1 Estimation of Degrees of Freedom

Let $S_t$ represent the price of the underlying asset at time $t$, and $X$ represents the logarithmic return rate of the underlying asset, then

$$X = \ln \frac{S_{t+1}}{S_t}. \quad (1)$$

We assume that $X$ follows the t-distribution with degrees of freedom $n$, i.e., $X \sim t(n)$, and the probability density function of the $t(n)$-distribution is

$$f(x, n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \Gamma\left(\frac{n}{2}\right)} (1 + \frac{x^2}{n})^{-\frac{n+1}{2}}, \quad -\infty < x < \infty. \quad (2)$$

$X_1, X_2, \ldots, X_m$ is a sample of the population $X \sim t(n)$, and $x_1, x_2, \ldots, x_m$ is its sample value. Now use the maximum likelihood estimation method to estimate the parameter $n$. Let $L(x_1, x_2, \ldots, x_m, n)$ be a likelihood function, then

$$L(x_1, x_2, \ldots, x_m, n) = \prod_{i=1}^{m} f(x_i, n) = \prod_{i=1}^{m} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \Gamma\left(\frac{n}{2}\right)} (1 + \frac{x_i^2}{n})^{-\frac{n+1}{2}}, \quad (3)$$

where, $\Gamma(x) = \int_{1}^{\infty} e^{-xt}t^{x-1}dt$. Taking the logarithm of Equation (3) yields

$$\ln L(x_1, x_2, \ldots, x_m, n) = \sum_{i=1}^{m} \ln f(x_i, n) = m\ln\left(\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \Gamma\left(\frac{n}{2}\right)}\right) - \frac{n+1}{2} \sum_{i=1}^{m} \ln (1 + \frac{x_i^2}{n}). \quad (4)$$

Taking the derivative of Equation (4) with respect to $n$ and making this derivative equal to zero, we obtain the following equation:

$$\frac{d\ln L(x_1, x_2, \ldots, x_m, n)}{dn} = 0. \quad (5)$$

By using the MATLAB program to solve Equation (5)
(see Appendix 1), the value of \( n \) can be obtained. However, in practice, when the sample size is large, we have not yet found a fast and effective MATLAB program to calculate \( n \). When the sample size is large, we can evaluate the degree of freedom by adopting the moment estimation method. Compared to maximum likelihood estimation, moment estimation is more convenient in calculation. In the following, we adopt the moment estimation method to estimate the degree of freedom \( n \). Since \( E(X) = 0 \), we use the second-order moment for estimation, namely,

\[
E(X^2) = \frac{1}{m} \sum_{i=1}^{m} X_i^2. \tag{6}
\]

After computing the second-order moment of population \( X \) from Equation (6) we can the moment estimation of the degree of freedom as follows:

\[
\hat{n} = 2 \frac{\sum_{i=1}^{m} X_i^2}{\sum_{i=1}^{m} X_i^2 - m}. \tag{7}
\]

2.2 Hypothesis Testing of Population Distribution

Suppose the distribution function \( F(x) \) of the population \( X \) is unknown and \( X_1, X_2, \ldots, X_n \) are the sample of \( X \) and \( x_1, x_2, \ldots, x_n \) are the sample observations against which to test the hypothesis about the overall distribution:

\[
H_0: F(x) = F_0(x), \quad H_1: F(x) \neq F_0(x). \tag{8}
\]

Where \( F_0(x) \) is the distribution function with known distribution type but may contain unknown parameters. Divide the interval \([M, N]\) into \( k \) disjoint intervals, where \( t_0, t_1, t_2, \ldots, t_k \) denote the endpoints of each interval and \( t_0 < t_1 < t_2 < \ldots < t_k \). \( N \) and \( M \) are the upper and lower bounds of the logarithmic return of the security, respectively.

\[
M = t_0 < t_1 < t_2 < \ldots < t_k = N. \tag{9}
\]

For financial markets without price limits, \( M=\infty \) and \( N=+\infty \). We can replace \( M \) and \( N \) in Equation (9) with sufficiently large and sufficiently small numbers, respectively. Let \( f_i \) be the frequency that the sample observation values \( x_1, x_2, \ldots, x_n \) fall into the interval \([t_i, t_{i+1}]\). Then the frequency of falling into the interval \([t_i, t_{i+1}]\) is \( \frac{f_i}{m} \), \( i=1,2,\ldots,k \). For the hypothesis testing problem of Equation (8), we construct the test statistic as follows:

\[
x^2 = \sum_{i=1}^{k} \frac{f_i - mp_i^2}{mp_i}, \tag{10}
\]

where \( p_i = F_0(t_i) - F_0(t_{i-1}) \). According to the result in the work of Sheng et al.\cite{19}, if the sample size \( n \) is sufficiently large, then when \( H_0 \) is true, the statistic (Equation (10)) approximately follows a \( \chi^2 (k-v-1) \) distribution, where \( v \) is the number of estimated parameters. Therefore, an approximate distribution of statistic (Equation (10)) can be obtained.

Take the significance level as \( \alpha \). We can get that the rejection domain under the \( \chi^2 \) goodness of fit test is

\[
\left\{ x^2 = \sum_{i=1}^{k} \frac{f_i - mp_i^2}{mp_i} \geq x_{\alpha}^2 (k - v - 1) \right\}. \tag{11}
\]

If falling into the rejection domain, we will reject \( H_0 \); Otherwise, there is no reason to reject the original assumption, and \( H_0 \) should be accepted.

3 RESULTS AND DISCUSSION

3.1 Empirical Analysis

Based on the trading data of the SSECI, this paper conducts an empirical study on the return rate of securities. It tests the return rate of the SSECI through the normal test (Figure 1 and 2). The results show that the return rate does not obey the logarithmic normal distribution. However, we use the goodness of fit test to explore the rationality of adopting the logarithmic t-distribution.

We first conduct a normal test on the logarithmic return of the SSECI from 2016 to 2019. Here, in order to make our test more convincing, we use the normal test Q-Q chart and the normal test P-P chart for the test. The principle of Q-Q chart is that if the data is normal, the assumed normal quantile will be basically consistent with the actual data. According to the P-P chart’s principle, the cumulative percentage of the data is essentially consistent with the cumulative proportion of the normal distribution if the data is normal. If the data are subject to a normal distribution, the scatter plot should roughly resemble a diagonal straight line whether it’s tested through a P-P test or a Q-Q test. The data probably do not exhibit normal characteristics if the scatter plot does not seem like a straight line. Observing Figures 1 and 2, we find that in most cases (e.g., Figure 1A, C, D and Figure 2B, C, D), the deviation of scatter points from the straight line is more obvious in the scatter plot, except for Figure 1B and Figure 2B. Therefore, in general, using a logarithmic normal distribution to describe the probability distribution of security returns is not quite in line with the actual situation.

Now we test the logarithmic return of the SSECI. The statistical analysis data in the following are the closing prices of the SSECI from January 4, 2002 to December 31, 2013, sourced from the Zhongyuan Securities System. Assuming that its return \( X \) follows the t-distribution of degrees of freedom \( n \). Using the moment
estimation, we can estimate $n$. Taking sample size $m=2902$, we obtain that the upper and lower bounds of the logarithmic return are 0.988458387 and 1.0128305, respectively. Taking $k=10$, then the interval length of $[t_i, t_{i+1}]$ is 0.002437211.

Based on the data in Table 1, it can be calculated that

$$x^2 = \sum_{i=1}^{k} \frac{t_i - mp_i^2}{mp_i} = 10.59722703 \approx 10.5972. \quad (12')$$

From Table 1, it can be seen that when the significance levels of $\alpha$ are taken as 0.005, 0.01, 0.025, 0.05, 0.01, respectively, the rejection domain that did not fall into $H_0$ is

$$\left\{ x^2 = \sum_{i=1}^{k} \frac{t_i - mp_i^2}{mp_i} \geq x^2_\alpha (k - v - 1) \right\}. \quad (13)$$

Therefore, it is reasonable to accept the original assumption that the return on securities follows a logarithmic $t$-distribution.

4 CONCLUSION

The security return is an important parameter in the pricing research of financial derivatives, and its distribution is also a hot topic in the field of financial engineering. This paper has mainly explored the rationality of the distribution of security returns proposed in the work of Cassidy et al. [9]. We first assumed that the logarithmic return rate of the underlying asset follows the logarithmic $t$-distribution with a degree of freedom of $n$, and then calculates the corresponding degree of freedom $n$ by using the point estimation method. Subsequently, we offered the basic principle and method of goodness of fit test, and obtained the rejection domain that did not fall into the original hypothesis of goodness of fit test. Finally, based on the trading data of the SSECI, this paper conducted an empirical study on the return rate of securities. The results showed that the return rate did not obey the logarithmic normal distribution, and the security return rate following the logarithmic $t$-distribution was reasonable.

The data in this paper are mainly based on the returns of China’s securities market, without fully considering the characteristic distribution of returns in other countries’ securities markets. Additionally, we did not analyze other financial products, but only the return on stocks. In future work, we will compare and analyze different financial products in different markets to gain more insights into management or economics.

**Acknowledgements**

The authors thank the anonymous referee’s comments leading to a much more improved presentation. This
work was supported by the Doctoral Fund of Henan Polytechnic University (No. B2020-36).

Conflicts of Interest
We confirm that the submitted manuscript contains original unpublished work and has not been submitted for publication elsewhere at the same time. And we also confirm that there are no relevant financial or non-financial competing interests to report.

Author Contribution
Liu X and Li H completed the theoretical analysis. Cao Y and Han X designed the numerical experiment. Liu X and Li H supervised the work. Han X performed the data analysis. Liu X, Li H and Cao Y drafted the manuscript. All authors contributed to the writing of the article, read and approved the article for submission.

Abbreviation List
SSECI, Shanghai stock exchange composite index

References