Math 563 Final Report

I-Lun Tu/ 670486091

1. Introduction

My research motivation is to search for a straightforward way to generalize copula to a higher dimension. Currently, much emphasis has been placed upon vine copula when modeling high dimension dependence structure is in consideration. However, the drawbacks of using vine copula constructions are: 1) depending on the selected marginals and coupling methods, there are often a relatively large number of parameters that need to be estimated, 2) regarding the dependence structure, many trimming processes and comparisons across different structures needs to be done, and sometimes there are some ambiguities that call for subjective decisions.

Considering the above drawbacks, I would like to study more about a copula that is implied in a distribution named Skew-T distribution, which is a special case of generalized hyperbolic distribution. The main advantage of using this implicit copula is that it provides a relatively clear way to model dependence structure. Namely, the shape parameter, dispersion matrix, and skewness parameter. In addition, this copula retains asymptotic tail dependency, heavy tail, and, more importantly, asymmetric lower and upper tail. Considering the exchangeability limitation of the Archimedean copula and the symmetry limitation of the elliptical copula, these properties allow more flexibility for modeling financial data, as financial data often show a higher probability of joint lower tail events. Also, previous researchers have shown that the bivariate T distribution is very effective in describing the behavior of pairs of stocks, and this is part of the reason that when selecting marginals for vine copula, T distribution is often included. Arguably, we could expect that a high dimension Skew-T distribution can be used to great effectiveness.

My aim for this research and some possible contributions are finding empirical evidence to support the use of Skew-T distribution and copula in a market risk management context. To the best of my knowledge, the current literature rarely examines the effectiveness of copula modeling from a market risk management angle. For example, when a copula model is established, more consideration is placed upon the method of estimation and the quality of fit, but not on whether the resulting VaR provides adequate coverage or not. Therefore, through this research project, I would like to fill this gap. There will be two main emphases. One of them will be placed on the estimation processes of Skew-T copula, while the other will be placed on verifying the VaR.

Regarding the estimation processes, previous researchers have shown that when modeling the marginal behaviors of publicly available financial data, such as the stock market, the data often does not support the use of skewness parameter, as the loglikelihood test would often select standard T distribution. However, since copula separates the marginal behaviors from dependence structures, the skewness parameter can be more directly estimated as a dependence structure. Therefore, skewness as a dependence structure rather than marginal behavior might be considered.

Regarding the verification of VaR, I intend to emphasize more modern backtesting methods. The criterion of selection is to use a backtesting method that can factor in both the number of exceedance and the severity of the exceedance. As it is known, VaR is effective in setting up a threshold for a certain level of coverage, but it does not provide any information on the possible severity of the violation. Therefore, many other

risk measures must be used in conjunction to provide a more general picture of risks. However, VaR is the most widely used and most accepted method in practice. It would be much more practical to improve or generalize VaR directly through backtesting. One such method I intend to use is the Risk Map backtesting method. If the results do suggest that the effectiveness of VaR is supported by this type of backtesting method, it would be more straightforward to consider the use of Skew-T distribution and copula in high dimension dependence modeling.

2. Literature Review

(1) Skew-T Distribution and Group-T Distribution

Skew-T distribution is a special case of generalized mean-variance mixture distribution and the related implicit copula was introduced in McNeil (2005). There are other types of Skew-T distribution in the literature formed by hidden truncation and have been a very active research topic, but in this project, I mainly focus on the aforementioned type of Skew-T construction. The basic definition was as follows:

$$X = \gamma V^{-1} + V^{-\frac{1}{2}}\mathbf{Z}$$

where V was $G(\frac{\nu}{2}, \frac{\nu}{2})$, γ was skewness parameter vector. When $\gamma = 0$, it reduces to Student-T distribution. When $\nu \to \infty$, it became Normal distribution. Note that the Skew Normal distribution was not nested within. Also, to have finite covariance, $\nu > 4$.

In this paper, the researchers also mentioned the possible construction of Group-T distribution or copula. The basic definition was as follows:

$$X = V^{-\frac{1}{2}}\mathbf{Z}$$

Group-T was constructed by forming subgroups of T distribution that used different ν parameters (different V). The concept was that in modeling different types of financial data, we might have a rough idea of how to group different types of assets. Therefore, Group-T copula could provide a natural structure to couple those different assets together by using different ν while taking in the differences across groups as well as the similarity within groups. However, there was no general guidance or suggestion regarding how to estimate or provide this prior information. Therefore, this grouping structure was subject to users' expertise.

Building upon the Group-T structure, later researchers (Luo and Shevchenko 2010) loosened the limitation of grouping structure. As a result, every marginal could have its degree of freedom and thus a prior information on grouping would not be needed. They further demonstrated that this generalized grouping structure could have a major impact on tail distribution and thus would affect the calculation of VaR in a major way. Other researchers (Konrad Banachewicz and Aad van der Vaart 2008) generalized both the Skew-T and Group-T to allow the marginals to have different degrees of freedom as well as different skewness parameters. This flexibility also had a major impact on asymptotic tail dependency.

The above researchers provided a heuristic way to estimate the parameters. To explain, the maximum likelihood estimation for the copula density was suggested. More importantly, the researchers suggested the use of the method of moments by estimating the dispersion matrix through calibrating bivariate Kendall's tau. Although these methods were straightforward, technical difficulties arose when implemented. There were two main difficulties: 1) maximizing likelihood in the marginals involved using the quantile of univariate Skew-T distribution, and since the distribution did not have a close form quantile function, a large number of simulations had to be done. These would be computationally costly especially when the dimension was high. 2) There was no guarantee that the dispersion matrix estimated by Kendall's tau would be invertible.

In light of this limitation, another researcher (Toshinao Yoshiba 2018) provided a partial solution. That is, 1) instead of simulating a large number of realizations to form an empirical distribution function, the researcher suggested an interpolating method, named monotone interpolator, to estimate the distribution function. 2) reparameterize the Cholesky decomposed triangular matrix with trigonometric functions, and this would in effect guarantee the semi-definiteness of the dispersion matrix. In his estimation, the interpolating method was about 400 times faster than the purely simulation-based method and more importantly, the method maintained comparable accuracy, especially for the tail area.

However, there was one drawback to his improvement. That is, the estimation of the skewness parameter was less accurate when using a non-parametric method to estimate the marginals. Although this was not a serious issue, as we could always use the parametric method if we have a particular preconceived model in mind to verify the marginals. However, this did limit the flexibility of copula estimation, since if the goal was to utilize the dependence structure of Skew-T distribution as a method to couple different marginals, the flexibility to allow different estimation methods for the marginals would be important. For example, many tail models in the literature utilized the semi-parametric method, which used a non-parametric method to model the body of the distribution while using parametric methods to model the tails. Therefore, I intended to look deeper into this issue and search for a non-parametric method that was less susceptible to this issue, and also provided more explanation regarding these drawbacks.

(2) VaR Backtesting

There were many backtesting methods designed to verify the effectiveness of VaR (refer to Zhang, Y., & Nadarajah, S. 2017 for a comprehensive list). One of the most frequently used backtesting methods in practice was the unconditional test method. This method emphasized the unconditional coverage of VaR and thus was tested through violation-based tests. One major drawback of the test was that it only factored in the number of exceedance but not the severity of the exceedance. Thus, the effectiveness of VaR would be considered equal if they both adequately control the number of exceedance regardless of the total amount of loss incurred.

A new method of unconditional test, named Risk Map, was introduced by (Gilbert Colletaz 2013). This test overcame the mentioned drawback by establishing two VaR when testing: 1) Standard VaR and 2) Super Exception VaR. The test statistics as usual assume the number of violations followed Binomial distribution with P being the level of the corresponding VaR and Super Exception VaR. Therefore, this test was relatively easy to implement for practitioners in the industry. I would like to use this backtesting framework to verify the effectiveness of copula modeling.

3. Methodology

(1) Data Source

The main focus of this research is to apply copula modeling in market risk modeling. Thus, some possible data sources are provided here. First, publicly available data, 1) stock price 2) stock index 3) futures price 4) exchange rate, would be used to model. Arguably, the ability to model the dependence structure of this data is of paramount importance in so many situations. However, for a more realistic and in-depth analysis of P&L distribution, internal data of banks is required, and the shortage of it will inevitably cause data availability issues.

(2) Data Preprocess

Since financial data is always time-dependent, I would make use of the time series model, such as ARIMA and GARCH models, to filter the log return and take the resulting residuals for subsequent modeling. This way the risk measurements could be more dynamic and realistic, as volatility clustering, as well as other important shifting of parameters during stressful market periods, are taken into account. In this project, filtered and unfiltered returns will both be studied and compared

(3) Copula Estimation and Comparison

The standard way to estimate copula in the literature (McNeil 2015) involves two-stage estimation processes. The first stage is to transform the marginal random variable into a pseudo copula. The second stage is to maximize the likelihood. For the first stage, a non-parametric method as shown below is chosen. This method will map the random variable to be within the boundary of the hyper-cube of the copula. For the second stage, I will utilize the recently introduced method (Toshinao Yoshiba 2018) to estimate the copula.

$$\frac{1}{n+1} \sum_{t=1}^{n} I_{(X_{t,i <= x})}$$

After the estimation, the result will be compared with a one-stage EM estimation. Through the comparison, we could observe how the marginal would affect the overall dependence structure. In addition, I aim to explore semi-parametric methods when estimating the marginals. It has been suggested in the literature (McNeil 2015) that using empirical function would often underestimate the tail risks. One solution to this issue is to model the tail using generalized Pareto distribution while estimating the body of distribution through empirical distribution function.

(4) Simulation Process

After copula models are fitted, the simulation method would be straightforward. As previously defined, Skew-T distribution can be sampled through inverse Gaussian distribution. With the sampling method in place, VaR can be calculated accordingly. However, as previously mentioned, to have a more sophisticated analysis, realistic P&L data will be required. One way to simulate such data might be to use the Markov Switching GARCH model (Bauwens et al., 2010), which allows conditional mean and volatility to change over time. This may alleviate some of the issues of data availability problems.

(5) Backtesting Method

The emphasis would be placed on using Risk Map (Gilbert Colletaz 2013) as the backtesting method. Due to its design, this method is especially suitable for tail risk models. The test statistics is as follows (Zhang, Y., & Nadarajah, S. 2017):

$$LR = -2 * ln \left[\frac{(1-\alpha)^{N_0} (\alpha - \alpha')^{N_1} (\alpha')^{N_2}}{(\frac{N_0}{n})^{N_0} (\frac{N_1}{n})^{N_1} (\frac{N_2}{n})^{N_2}} \right]$$

where $0 < \alpha' < \alpha < 1$ and N_2 denotes the number of violations for $VaR_{\alpha'}$, N_1 denotes the number of occurrences of a loss between VaR_{α} and $VaR_{\alpha'}$, and N_0 denotes the number of occurrences of a loss lower than $-VaR_{\alpha}$. The asymptotic null distribution of the statistic is the chi-squared distribution with two degrees of freedom.

4. Preliminary Results and Discussion

(1) Establish VaR for stock portfolio on financial institution

The data was weekly data sourced from Yahoo Finance for the last 5 years. For preliminary results, I aimed to model dependence structure for a portfolio of 15 stocks. There are stocks from consumer finance, commercial banking, brokerage and investment management firms. The effectiveness of resulting VaR was examined through simple coverage test.

(2) Estimation Process

The psuedo copula was formed through the aforementioned non-parametric method and the recently introduced interpolating method was used to calculate the quantile of univariate Skew-T distribution in the marginals.

- The below psuedo copula revealed a clear dependence structure


Banking Portfolio
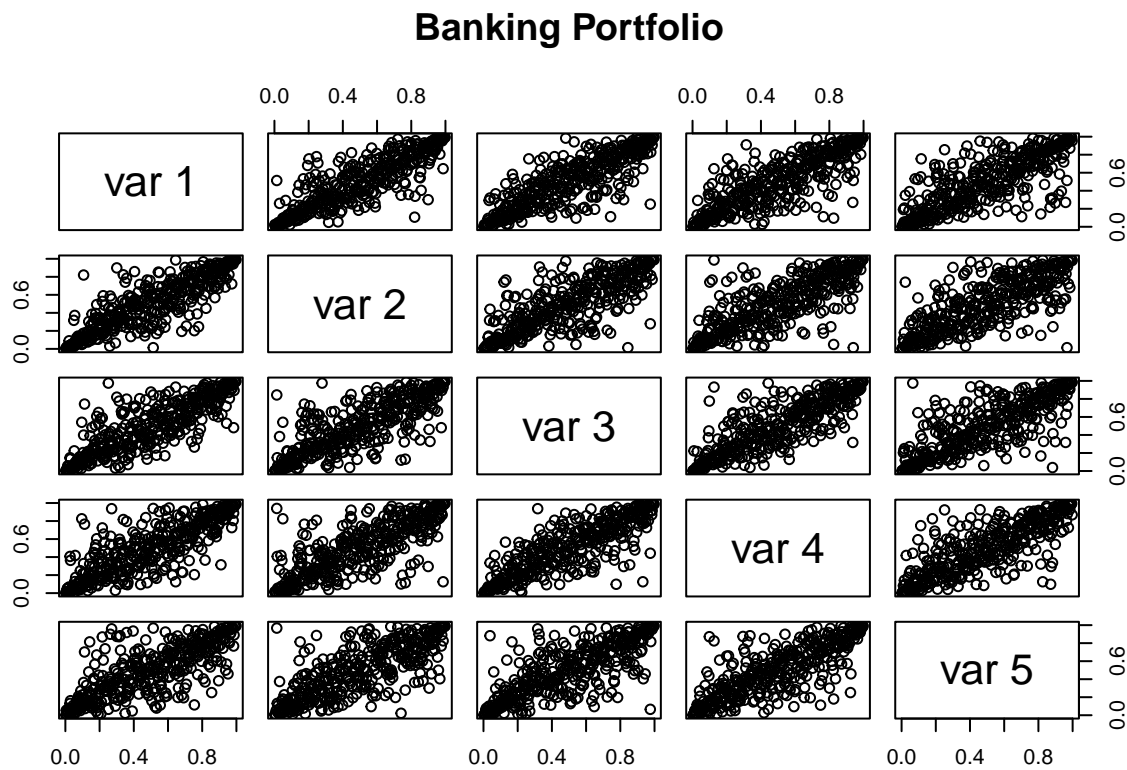
- The below table summarized the relative fit of standard-T and Skew-T copula. The likelihood improved a lot, suggesting the incorporation of skewness. Here, equal skewness was assumed, so there was only one additional parameter.

	T-copula	Skew-T copula
nu	5.658878	5.9137642
gamma	NA	-0.2259792
log_lik	734.429860	851.7734165
AIC	-1256.859720	-1489.5468330
BIC	-843.470601	-1072.2578162

- However, from below graph, we could observe the dependence structure differed from the original dependence structure. This copula behaved more like independent copula.

Banking Portfolio
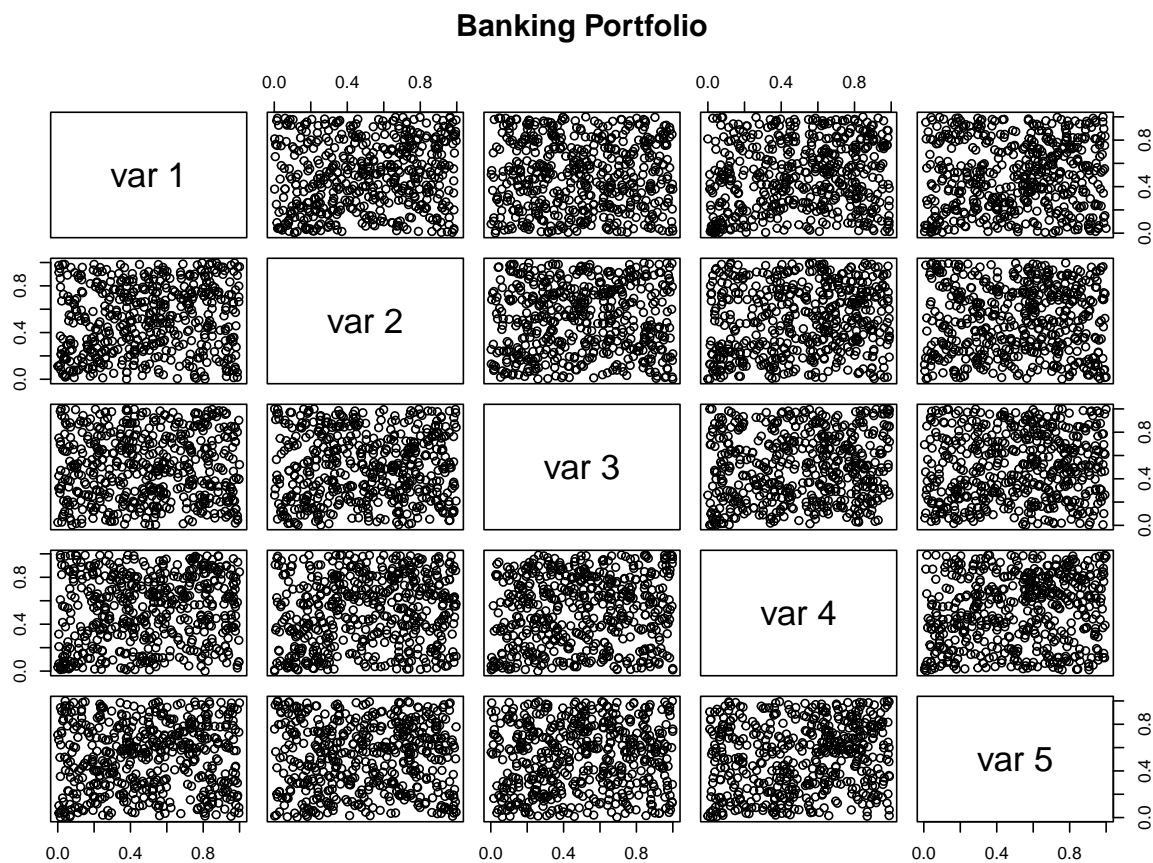
- The below graph showed the aggregate loss. We could observe that the model was not a good fit for the data. As a result, from the table, we could see that the VaR did not provide adequate coverage. (Note that the data was scaled and centered)


N = 100000   Bandwidth = 0.02256
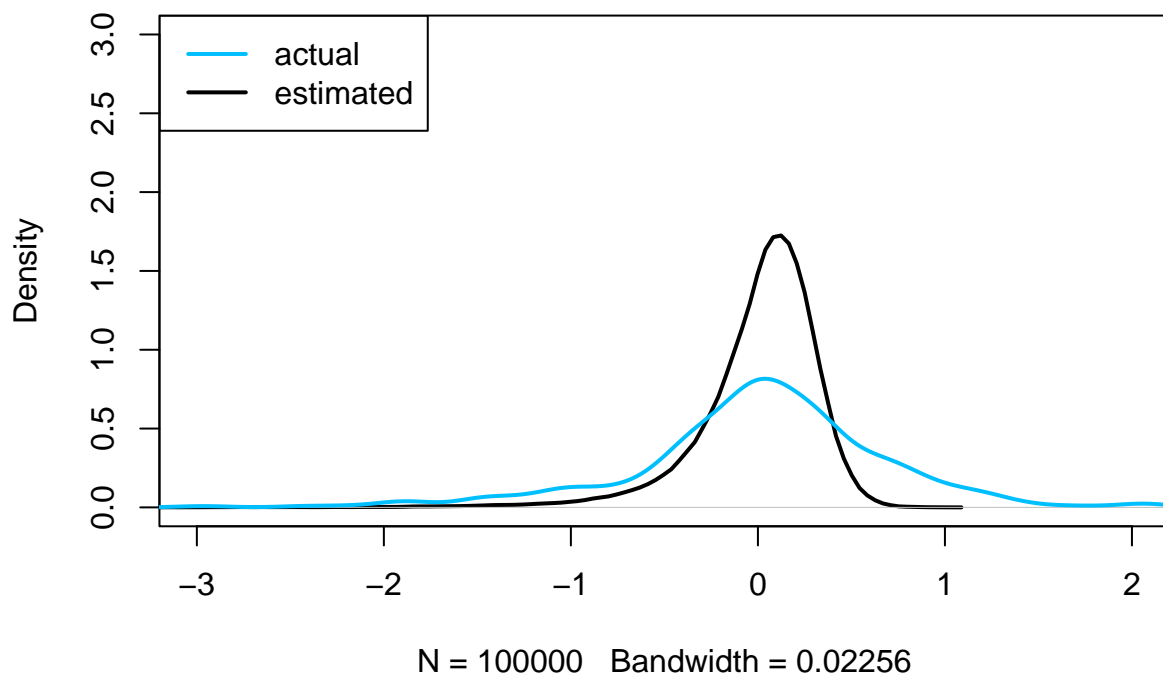
	99%	95%	90%	85%
% VaR	-1.23	-0.57	-0.37	-0.26
Empirical Violation Percentage	0.05	0.15	0.21	0.28

- However, if we changed the estimation process to one stage EM estimation. We could observe from the below simulation that the dependence structure was properly preserved.
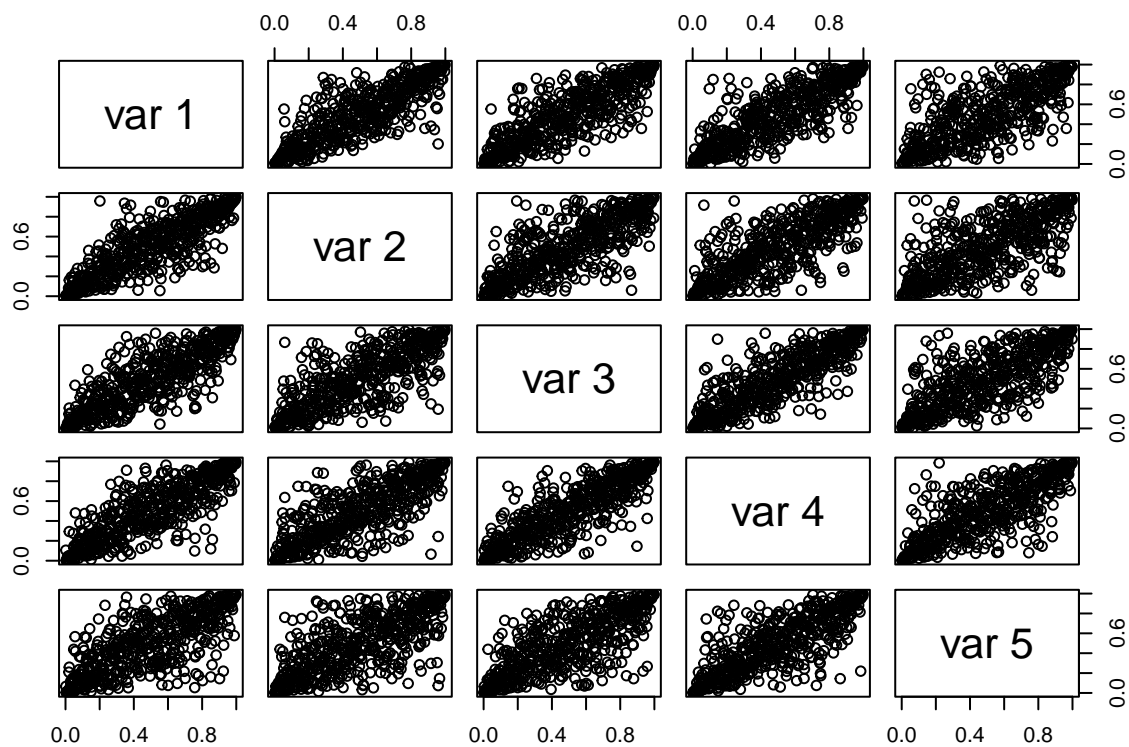
- From the below table, we could see that the VaR performed reasonably well (100,000 simulated data points).

N = 100000   Bandwidth = 0.00203
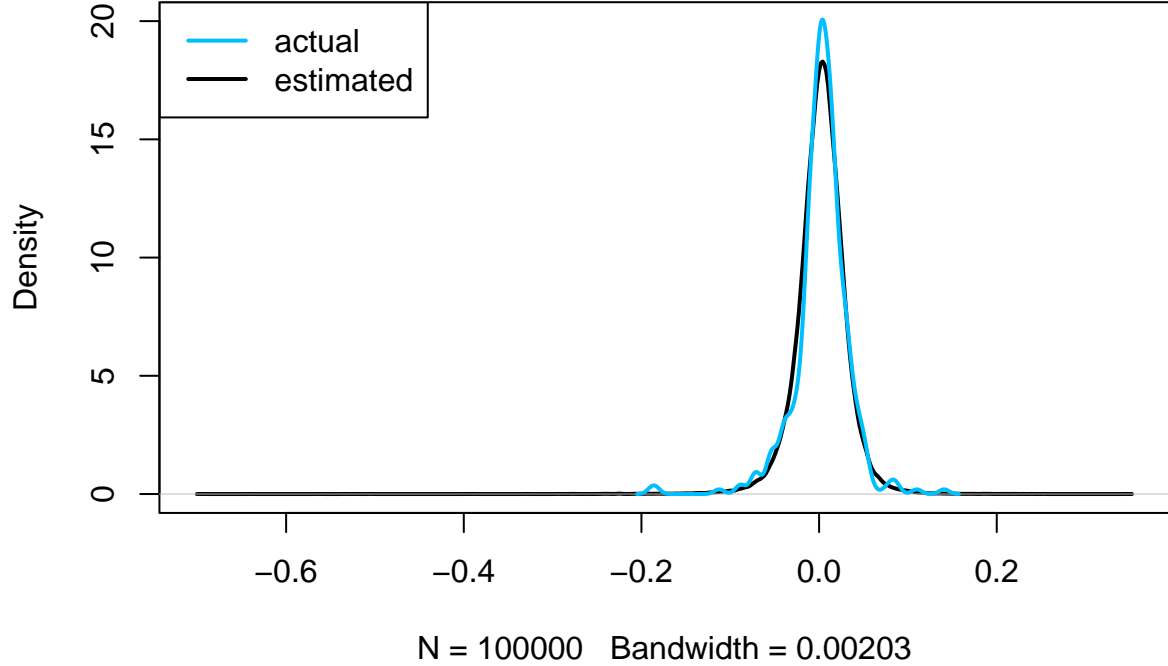
	99%	95%	90%	85%
% VaR	-0.08	-0.04	-0.03	-0.02
Empirical Violation Percentage	0.01	0.06	0.11	0.14

(3) Discussion

The poor performance of copula-based estimation might stem from the fact that non-parametric method did not take into account the skewness information, yet we understood that skewness parameter affected both dependence structure and marginal behaviors. This issue aligned with the result from previous researcher. Another possible issue might be that the kurtosis of this particular empirical data was close to the boundary of 4. We could see that in the EM estimation there was still some underfitting of kurtosis.

5. Conclusion

With empirical implementation, the speed and accuracy were confirmed. However, the problem of estimating skewness parameters were met. To solve this issue, I would like to search for more sophisticated non-parametric methods. To relax the restriction on the degree of freedom, I intend to look into other construction of Skew-T distribution, such as the one introduced in Azzalini and Capitanio (2003). Moreover, the scope and depth of the initial analysis would be improved as described in the methodology section. Namely, different data will be studied, P&L will be simulated through a more realist way, and Risk Map will be used to verify the effectivenss of the model.

6. Reference

- Azzalini, A. and A. Capitanio (2003) "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution," Journal of the Royal Statistical Society Series B,65(2), 367–389.
- Bauwens, L., Preminger, A., Rombouts, J.V.K., 2010. Theory and inference for a Markov switching GARCH model. Econometrics Journal 13, 218–244.
- Colletaz, G., Hurlin, C. and Perignon, C. (2013). The risk map: A new tool for validating risk models. Journal of Banking and Finance, 37, 3843-3854
- Demarta, S. and A. J. McNeil (2005) "The t copula and related copulas," International Statistical Review, 73(1), 111–129.
- Konrad Banachewicz, Aad van der Vaart (2008), Tail dependence of skewed grouped t-distributions, Statistics and Probability Letters 78 (2008) 2388–2399
- McNeil, A. J., R. Frey, and P. Embrechts (2015) Quantitative Risk Management: Concepts, Techniques, and Tools, Princeton University Press, revised ed
- Toshinao Yoshiba Maximum likelihood estimation of skew- t copulas with its applications to stock returns May 2018 Journal of Statistical Computation and Simulation 88(2):1-18
- Xiaolin Luo, Pavel V. Shevchenko (2009) The t copula with Multiple Parameters of Degrees of Freedom: Bivariate Characteristics and Application to Risk Management, Quantitative Finance. November 2009
- Zhang, Y., & Nadarajah, S. (2017). A review of backtesting for value at risk. Communications in Statistics - Theory and Methods, 1-24.