Simulation and Extreme VaR and VaR Confidence Interval Estimation for a Class of Heavy-Tailed Risk Factors

Hao Wang

Department of Mathematics
University of Oregon

Abstract

This paper introduces a calibrated scenario generation method to estimate extreme Value-at-Risk (VaR) and Value-at-Risk confidence interval (VaR CI) of a portfolio with single risk factor which has heavy tailed distribution. It is well known that lot of financial, daily log-return data demonstrate heavy-tailed distribution. This makes all the models with normally, even log-normally distributed assumption become disabled (see [25]). We handle the daily return data with heavy tailed distribution and use a model of log-mixture of normal distributions to calibrate mean, variance, kurtosis, and sixth moment and fit the empirical distribution. An extreme value is a rare event and not easy to be observed. However, once it occurs, it brings disaster to any involved financial institute and financial practitioners. Therefore, undoubtedly how to effectively estimate the portfolio extreme VaR and VaR CI is a primary concern in risk management. In this paper, we will use a non-parametric method to derive portfolio extreme VaR and VaR confidence interval estimates for heavy-tailed distributions based on scenarios which are generated with calibration.

MSC2000 Subject Classification: 62G32, 62P05

Key words and phrases: VaR, VaR CI, heavy tails, Monte Carlo simulation


1 Introduction

With the deficiencies of normal and thin tail distributed models in risk management, heavy tailed phenomena have brought wide attention (see de Haan-Peng [4], de Haan-Pereira [5], Embrechts et al. [6], Feigin-Resnick [7], Resnick [20], Hull-White [12], Huisman et al. [11], Adler et al. eds. [1], Koedijk [15], Ho et al. [10], Muller et al. [18], Sornette et al. [22], to name only a few). Daily log-return financial data, credit risk involved data, energy price data, jumps involved data always demonstrate heavy tails. Empirical results have rejected the hypothesis of normal even log-normal distributions (see Upton-Shannon [25], Tucker et al. [24]). Thus, heavy tailed phenomena have raised lot of new problems and challenges. A famous example is that for heavy tailed cases, Black-Scholes formula is simply deficient. To solve this problem, many new ideas are proposed such as considering stochastic volatility and using GARCH model to estimate the diffusion coefficient (see Heston [8], Ritchken-Trevor [21], Heston-Nandi [9]), calibrating stable processes or Jump-diffusion processes to model the real situation (see Su-Fleisher [23], Tucker et al. [24]). Another challenge is how to develop risk management tools for heavy tailed cases. The primary important tools in risk management, of course, are portfolio VaR (extreme VaR) and VaR CI (extreme VaR CI) estimation. Roughly speaking, VaR is just a quantile in statistic language and extreme VaR is just a high quantile. There are many papers which apply extreme value theory (EVT) to estimate portfolio VaR and extreme VaR. The basic idea in the EVT techniques is to use sample to estimate the tail index of a heavy tailed distribution by tail index estimators such as Hill estimator, Pickand estimator, Deckers-Einmahl-de Haan estimator, etc, then the VaR or extreme VaR is just a function of the tail index. In this method, the optimal choice of the cut-off parameter or the sample fraction is a difficult problem. The recent work of Danielsson et al. ([3]) has proposed good ideas to approach this problem. We know that EVT is only concerning the tail part information of the empirical distribution and ignores the remaining part information. However, extreme values are rare events and the size of extreme observations is often small even though the whole sample size is big enough. This gives troubles for the EVT techniques for the cases that the whole sample size is not big enough. On the other hand, EVT method only calibrates part information of the empirical distribution. Therefore, It can not be directly used to generate scenarios and it can not provide a good estimate of the portfolio VaR CI since portfolio VaR CI is very sensitive to the tail index estimate. The following
intuitive idea leads us to develop another way to estimate portfolio extreme VaR and VaR CI. Since the area under any density curve is always equal to one, the peak part is implicated with the tail part. Therefore, the peak part information should be used. In our model, both the peak part and the tail part information are used for calibration. In the following, we call a distribution temperately-heavy tailed if the kurtosis of the distribution exists and is bigger than 3; we call a distribution intemperately-heavy tailed if the distribution doesn’t have a finite kurtosis. From examples in [6], EVT techniques work pretty well for intemperately-heavy tailed data such as insurance data or other data involving big jumps. However, from simulation results, the daily log-returns of most financial risk factors have symmetric temperately-heavy tailed distributions. The EVT tail index estimators are usually incompetent for this class of distributions. For example, we can generate random numbers of t-distribution with degree equal to 20, where the sample size can be as large as we like. In this case, the value of Hill estimator should be close to 0.05. However, we can find that this is impossible from the Hill plot. In this paper, we calibrate the whole empirical distribution instead of only tail index as in EVT, then we generate scenarios based on the calibrated results. In our model, we use mixture of normal distributions to calibrate and fit the empirical distribution of the daily log-returns. The mixture of normal distributions are of a class of distributions which fit the symmetric temperately-heavy tailed distributions very well (see the following figure 4 and figure 5). Once the model parameters are calibrated, we generate scenarios based on the mixture of normal distributions with calibrated parameters. Finally a non-parametric method is used to estimate portfolio VaR and VaR CI based on the generated scenarios. The scenario size can be as large as we like. This is the advantage which can completely satisfy the requirement of the estimations of non-parametric extreme portfolio VaR and VaR CI. Now let us introduce our model.

2 Preliminary and Model

First, let us give some basic definitions. The VaR\(^1\) of a portfolio for a given confidence level \(\theta\) is defined as a \(\theta\) quantile \(Q_\theta\). Formally we can define a \(\theta\) quantile \(Q_\theta\) as follows.

\[
P(V_0 - V_t \geq Q_\theta) = \theta,
\]

\(^1\)In the market or financial literature, the VaR of a portfolio is defined as an amount of money such that with a small probability the portfolio will lose more than that amount over a given period of time. Therefore, the market VaR is just the absolute value of our defined VaR.
where \( V_0 \) is the current portfolio value, \( V_t \) is the portfolio value at future time \( t \), \( \mathbb{P} \) is the market probability. For related terms, reader is referred to [2], [19], [14], [17], [16] [13].

In this paper, we assume that the value of a risk factor is positive. This class of risk factors include interest rates, foreign exchange rates, equity indices, and so on. Generally, a portfolio contains multiply risk factors. To handle this general portfolio, it is not easy to give high dimensional graph of the calibrated density surface and the empirical histogram. In order to simplify our question and easily demonstrate the effectiveness of the model, this paper will concentrate on the case that a portfolio contains only one risk factor. Let \( r_0, \cdots, r_n \) be the observations of consecutive \( n + 1 \) days of the values of one risk factor. Let \( x_i = \log r_{i+1} - \log r_i \). Then, \( x_1, \cdots, x_n \) is called the daily log-return data. The sample kurtosis is defined as follows:

\[
(2.2) \quad \kappa = \frac{\frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})^4}{\left(\frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})^2\right)^2}
\]

By our experience, the daily log-return data always show us with symmetric empirical distributions. Thus, we do not introduce skewness here.

**Definition 2.1** We say that a random variable \( \xi \) has a distribution which is a mixture of two normal distributions (MND) if the density function, \( f_\xi(x) \), of \( \xi \) exists and has the following representation: There exists a positive number \( p \in (0, 1) \) such that

\[
(2.3) \quad f_\xi(x) = pf(\mu_Y, \sigma_Y, x) + (1-p)f(\mu_Z, \sigma_Z, x) \text{ for all } x \in \mathbb{R},
\]

where \( f(\mu_Y, \sigma_Y, x) \) is the density function of a random variable \( Y \) which has a normal distribution function \( N(\mu_Y, \sigma_Y) \) and \( f(\mu_Z, \sigma_Z, x) \) is the density function of a random variable \( Z \) which has a normal distribution function \( N(\mu_Z, \sigma_Z) \), where \( \sigma_Y \) and \( \sigma_Z \) are their standard deviations, respectively.

In the following, we will use a mixture of normal distributions to calibrate and fit the empirical distribution of the log-return data \( x_1, \cdots, x_n \).

**Definition 2.2** A random variable has a log-mixture distribution if the natural logarithm of the random variable has a mixture of normal distributions.
Therefore, by above definition (2.2) we actually model the daily return of the risk factor by a log-mixture distribution. Now let us consider following question. If we assume that the daily log-return data has a MND, then how can we identify the parameters? In following, without loss of generality, we assume that the distribution of the daily log-return data has zero mean and its variance is denoted by $\sigma^2$. Thus, in above (2.3), we have $\mu_Y = 0 = \mu_Z$. Consider the calibration on second moment, fourth moment, and sixth moment. From (2.3), we have following system of calibration equations.

\[
\begin{align*}
\left\{ \begin{array}{c}
 p\sigma_Y^2 + (1 - p)\sigma_Z^2 &= \sigma^2 \\
 pm_Y^4 + (1 - p)m_Z^4 &= \kappa(\sigma^2)^2 \\
 pm_Y^6 + (1 - p)m_Z^6 &= m_D^6
\end{array} \right.
\]

(2.4)

where $m_D^6$ is the data sixth moment, $\kappa$ is the data kurtosis, $\sigma_Y^2$, $m_Y^4$, and $m_Y^6$ are the second moment, fourth moment, and sixth moment of $Y$, respectively and $\sigma_Z^2$, $m_Z^4$, and $m_Z^6$ are the second moment, fourth moment, and sixth moment of $Z$, respectively. Since $Y$ and $Z$ have normal distributions, we get the following system of equations:

\[
\begin{align*}
\left\{ \begin{array}{c}
 p\sigma_Y^2 + (1 - p)\sigma_Z^2 &= \sigma^2 \\
 3p(\sigma_Y^2)^2 + 3(1 - p)(\sigma_Z^2)^2 &= \kappa(\sigma^2)^2 \\
 15p(\sigma_Y^2)^3 + 15(1 - p)(\sigma_Z^2)^3 &= m_D^6
\end{array} \right.
\]

(2.5)

In order to solve above system of equation, we decompose the task into two stages. In the first stage, we consider $p$ as an given parameter and consider following system of equations:

\[
\begin{align*}
\left\{ \begin{array}{c}
 p\sigma_Y^2 + (1 - p)\sigma_Z^2 &= \sigma^2 \\
 3p(\sigma_Y^2)^2 + 3(1 - p)(\sigma_Z^2)^2 &= \kappa(\sigma^2)^2
\end{array} \right.
\]

(2.6)

Solve above system of equations, we have the following solutions:

Solution-one:

\[
(\sigma_Y^2 = \sigma^2 + \sigma^2 \sqrt{\frac{1 - p}{p} \left( \frac{\kappa}{3} - 1 \right)}, \quad \sigma_Z^2 = \sigma^2 - \sigma^2 \sqrt{\frac{p}{1 - p} \left( \frac{\kappa}{3} - 1 \right)})
\]

for $0 < p < \frac{3}{\kappa}$.

Solution-two:

\[
(\sigma_Y^2 = \sigma^2 - \sigma^2 \sqrt{\frac{1 - p}{p} \left( \frac{\kappa}{3} - 1 \right)}, \quad \sigma_Z^2 = \sigma^2 + \sigma^2 \sqrt{\frac{p}{1 - p} \left( \frac{\kappa}{3} - 1 \right)})
\]

for $1 - \frac{3}{\kappa} < p < 1$. 

Above solution-one means that based on the data kurtosis \( \kappa \), once given a mixture parameter \( p \in (0, 3/\kappa) \), we can find the variances \( \sigma_Y^2 \) and \( \sigma_Z^2 \) of the normal distributions \( N(0, \sigma_Y) \) and \( N(0, \sigma_Z) \), respectively. Solution-two has similar meaning. So far we have calibrated the variances of two normal distributions, but they depend on the mixture parameter \( p \). What is the optimal choice of the mixture parameter \( p \) for our model? It should be the solution of sixth moment equation, but it is not easy to find its closed form, algebraic solution from the system of equations (2.5). Now we use a plot method to find an optimal solution of the mixture parameter. First, from the third equation (the sixth moment equation) of (2.5) we can derive the following equation:

\[
p^* = \frac{m_6^D/15 - (\sigma_Z^2)^3}{[(\sigma_Y^2)^3 - (\sigma_Z^2)^3]}.
\]

Based on data variance \( \sigma^2 \), data kurtosis \( \kappa \), given a value of \( p \in (0, 3/\kappa) \), through (2.7) we can find the values of \( \sigma_Y^2 \) and \( \sigma_Z^2 \), then based on data sixth moment \( m_6^D \), through the equation (2.9) we can find \( p^* \). In this way \( p^* \) is a function of \( p \). Plot function \( p^* \) with variable \( p \). Thus, the intersection of the diagonal line and the curve of \( p^*(p) \) gives the solution of \( p \) for the system of equations 2.5, which is the optimal choice of the mixture parameter in our model. The solutions of \( p \) may not be unique. Anyway we can find at least one solution. If there are more than one solution, we just use the minimal solution. From several examples the graphs of \( p^*(p) \) are monotone curves, so, it has a unique solution. Thus, we got all the parameters we want. Once we find the mixture of normal distributions, we can generate scenarios. From generated scenarios, we can estimate the portfolio extreme VaR and VaR CI by non-parametric method. Now let us discuss the non-parametric method.

## 3 Non-parametric Method

Suppose that the generated scenarios \( \{\bar{x}_i : 1 \leq i \leq m\} \) has distribution \( \tilde{F}_\eta(x) \). For a given \( \theta \in (0, 1) \), let \( q_\theta \) be the \( \theta \) quantile of \( \tilde{F}_\eta \) which satisfies

\[
P(\eta \geq q_\theta) = 1 - \theta.
\]

In order to find the VaR CI, define

\[
\Lambda = \#\{i : \bar{x}_i \leq q_\theta\},
\]
Then $\Lambda$ has a binomial distribution $b(m, \theta)$ which has mean $m\theta$ and standard deviation $\sqrt{m\theta(1-\theta)}$ and the distribution of $(\Lambda - m\theta)/\sqrt{m\theta(1-\theta)}$ converges to the standard normal distribution. That is

$$P(-Z_{\alpha/2} < \frac{\Lambda - m\theta}{\sqrt{m\theta(1-\theta)}} < Z_{\alpha/2}) \approx 1 - \alpha$$

for $m$ big enough. After a little adjustment, we have

$$P(\pi \leq \Lambda \leq \Pi) \geq 1 - \alpha \quad \text{for } m \text{ big enough},$$

where

$$\pi = \lfloor m\theta - (Z_{\alpha/2} + 0.05)\sqrt{m\theta(1-\theta)} \rfloor,$$

$$\Pi = \lfloor m\theta + (Z_{\alpha/2} + 0.05)\sqrt{m\theta(1-\theta)} \rfloor + 1,$$

$\lfloor a \rfloor$ is the maximum integer less than $a$.

Let $Y^1 \leq Y^2 \leq \cdots \leq Y^n$ be the order statistics constructed from the scenarios $\bar{x}_1, \cdots, \bar{x}_m$. Then the above inequality (3.13) is equivalent to

$$P(Y^\pi \leq q_\theta \leq Y^\Pi) \geq 1 - \alpha, \quad \text{for } m \text{ big enough}$$

This gives the non-parametric VaR confidence interval. To find the VaR, define

$$\mathcal{R} = \min\{i : 0 \leq i \leq m, Y^i \geq q_\theta\}.$$

Then, $Y^\mathcal{R}$ is the estimated VaR with probability level $\theta$.

## 4 VaR Transformation

In the financial markets, risk managers and financial market practitioners always use money VaR to make decisions. Therefore, we need to transform our model VaR and VaR CI into money VaR and money VaR CI. Essentially this is just the inverse transformation of our model. For example, we consider the case that the risk factor is a foreign exchange rate. Let $q_v$ denote the VaR value in our model and let $v_t$ and $v_{t+1}$ denote today’s and tomorrow’s exchange rates, respectively. Then, tomorrow’s exchange rate at VaR position is given by:

$$v_{t+1} = v_t e^{q_v}.$$
Based on our assumption that the portfolio has only one risk factor, if the amount of money invested today is denoted by $m_t$, then the portfolio money VaR, denoted by $M_{\text{var}}$, is given by:

\[(4.19)\]

\[M_{\text{var}} = m_t(v_{t+1} - v_t).\]

Similarly, if we replace $q_v$ by the lower bond or upper bond of the model VaR CI, we can find the money VaR CI lower bond and upper bond. Now let us verify the effectiveness of our model.

5 Model Validation and Data Analysis

In this section, we will use quantile-quantile plot (qq-plot) method, graph match, and quantile comparison to test our model. Although we have tested different data sets and different risk factors, here briefly we only calibrate one data set in this study. The data set consists of 1000 observations of daily log-returns of JPY/USD (Japanese Yen/U.S. Dollar) exchange rates, from March 21, 1996, to January 20, 2000. We use our model to calibrate the desired parameters. From the Glivenko-Cantelli Theorem (cf. Example 2.1.4. of [6]), if the mixture of normal distributions with found parameters well fits the empirical distribution of the data, it must demonstrate approximate linearity of the qq-plot. In order to demonstrate the degree of the calibrated MND fitting to the market data and the simulation error produced by scenario generation, for comparison, we will give two qq-plots. One is for the scenarios generated by the mixture of normal distributions with found parameters vs. the market daily log-return data. The second is for the scenarios generated by the mixture of normal distributions with found parameters vs. the scenarios generated by the mixture of normal distributions with found parameters. The second qq-plot is just showing the simulation error by computer. Then, we give the graphs of the $p^*(p)$ and the diagonal line and show the optimal value of the mixture parameter for our example. To get a clear view of the density curves, we illustrate and compare our model density curve, the found density curve of the mixture of normal distributions with the data histogram and a normal density curve which has mean and variance same as the data histogram. Finally, we use tables to compare data and scenario VaR’s and VaR CI’s with different confidence levels.
**Figure 1:** qq-plot for the market daily log-return data of exchange rates of JPY/USD vs. the scenarios generated by the mixture of normal distributions with calibrated parameters. It demonstrates approximate linearity.
Figure 2: qq-plot for the scenarios generated by the mixture of normal distributions with calibrated parameters vs. the scenarios generated by the mixture of normal distributions with calibrated parameters. This means that two groups of scenarios are generated by the same distribution. It demonstrates approximate linearity. Comparison of this qq-plot and the qq-plot in figure 1, we may realize that the calibrated MND fits the empirical distribution very well.
Figure 3: Plot of the function $p^*(p)$ and the diagonal line. The intersection gives us the optimal choice of $p$. 
Figure 4: Comparison of histogram of 1000 historical observations of daily log-return of JPY/USD exchange rates, calibrated density curve of mixture of normal distributions, and the normal density curve with same mean and variance as that of the histogram.

The parameters of the histogram and the different density curves in the figure 4 are listed in the following table 1 and table 2:

<table>
<thead>
<tr>
<th>MEAN</th>
<th>VARIANCE</th>
<th>KURTOSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histogram</td>
<td>0.0006</td>
<td>0.7</td>
</tr>
<tr>
<td>Normal</td>
<td>0.0006</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Table 1: Parameters of histogram and the normal curve.

<table>
<thead>
<tr>
<th>$\mu_Y$</th>
<th>$\mu_Z$</th>
<th>$\sigma_Y$</th>
<th>$\sigma_Z$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixture of Normals</td>
<td>0.0006</td>
<td>0.0006</td>
<td>6.5513</td>
<td>0.4338</td>
</tr>
</tbody>
</table>

Table 2: Parameters of the density curve of the mixture of normal distributions.
Then, we use our model to generate scenarios and we illustrate scenario histogram and compare the model density curve of the mixture of normal distributions to the histogram of the generated scenarios and to the normal density curve with same mean and variance as that of the data histogram.

![Figure 5](image)

**Figure 5:** Comparison of histogram of 10000 scenarios generated by the calibrated mixture of normal distributions, the calibrated density curve of the mixture of normal distributions, and the normal density curve with same mean and variance as that of the histogram.

The parameters of the histogram of the 10000 scenarios in the figure 5 are listed in the following table 3:

<table>
<thead>
<tr>
<th>MEAN</th>
<th>VARIANCE</th>
<th>KURTOSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0066</td>
<td>0.7138</td>
<td>12.9192</td>
</tr>
</tbody>
</table>

**Table 3:** Parameters of the histogram of the 10000 scenarios.
Now we use table to compare data and scenario VaR’s and VaR CI’s.

<table>
<thead>
<tr>
<th>SIZE</th>
<th>VaR(0.1)</th>
<th>VaRCILB(0.1)</th>
<th>VaRCIUB(0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>1000</td>
<td>0.8961</td>
<td>0.8508</td>
</tr>
<tr>
<td>scen</td>
<td>10000</td>
<td>0.8760</td>
<td>0.8589</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SIZE</th>
<th>VaR(0.05)</th>
<th>VaRCILB(0.05)</th>
<th>VaRCIUB(0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>1000</td>
<td>1.1678</td>
<td>1.0911</td>
</tr>
<tr>
<td>scen</td>
<td>10000</td>
<td>1.1638</td>
<td>1.1386</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SIZE</th>
<th>VaR(0.01)</th>
<th>VaRCILB(0.01)</th>
<th>VaRCIUB(0.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>1000</td>
<td>1.6868</td>
<td>1.5798</td>
</tr>
<tr>
<td>scen</td>
<td>10000</td>
<td>2.0571</td>
<td>1.9197</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SIZE</th>
<th>VaR(0.001)</th>
<th>VaRCILB(0.001)</th>
<th>VaRCIUB(0.001)</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>1000</td>
<td>2.7601</td>
<td>2.7601</td>
</tr>
<tr>
<td>scen</td>
<td>10000</td>
<td>4.9962</td>
<td>4.6179</td>
</tr>
</tbody>
</table>

**Table 4:** VaR(α) means VaR with confidence level α(100%). VaRCILB(α) means the 95% VaR CI lower bound of the VaR with confidence level α(100%) and VaRCIUB(α) means the 95% VaR CI upper bound of the VaR with confidence level α(100%). In the table, the * means the value is not available.

As a conclusion of this section, we give some comments on the simulation results. First, the illustration demonstrates the well fitting of the calibrated density curve of mixture of normal distributions with the histogram of the data of 1000 historical observations of daily log-return of JPY/USD exchange rates. Second, for the 10% and 5% confidence levels, the data and the scenario VaR’s and VaR CI’s are very close to each other. For the 1% and 0.1% confidence levels, since the data only has 1000 observations, it gives either no information or underestimate due to out-of-sample or lack-of-sample. For the generated scenarios, we can have any desired size. Combining the non-parametric method, this gives us a powerful tool for the risk management.
References


