International Journal of Business Marketing and Management (IJBMM)
Volume 7 Issue 5 Sep-Oct 2022, P.P. 112-127
ISSN: 2456-4559
www.ijbmm.com

Double VaR Model and Its Application Research

Xinyi Lu[1], Wanbing Zhang[2] , Peibiao Zhao[3*]
[1]*(School of Mathematics and Statistics, Nanjing University of Science and Technology, Nanjing 210094, Jiangsu, P. R. China. E-mail address: 1519916735@qq.com)*
[2]*(Grain and Oil Assets Management Co. LTD, Jiading District, Shanghai, P. R. China. E-mail address: zvb950831@163.com)*
[3]*(School of Mathematics and Statistics, Nanjing University of Science and Technology, Nanjing 210094, Jiangsu, P. R. China. * Corresponding author. E-mail address: pbzhao@njust.edu.cn)*

Abstract : *This paper intends to construct and study $VaR_\beta\,[VaR_\alpha(x)]$ (as may be called Double VaR, short for DVaR) problem under uncertain distribution based on the best (or worst) criteria,studies and gives the characteristics of DVaR and its application analysis in the portfolio.*

Keywords - *Double VaR, portfolio optimization, uncertainty*

I. Concept of DVaR

VaR as a measure of the size of the risk, its intuitive meaning is the maximum value of the loss of an asset or portfolio at a certain level of confidence. It is a simple and clear concept of risk,which is expressed in simple values, and is therefore widely used. Directly related to the accuracy of the VaR value is the future profit and loss distribution of the asset, which was initially met by the researchers for convenience, often assuming that the return on the asset was normal. Later, some scholars found that the distribution of asset income and normal distribution is not completely consistent, the income of the phenomenon of peak thick tail and fluctuating cluster, so the researchers began to use the distribution with peak thick tail characteristics to study it, such as t distribution, asymmetric Laplace distribution, mixed normal distribution and so on. From the results of the study, the effect is indeed significantly better than the normal distribution. However, the use of actual data to verify that the basis of this kind of distribution calculated VaR there is still some deviation. The future profit and loss distribution of an asset or portfolio is not accurate, and the absolutely reliable information available to researchers is only historical data. However, the fluctuation of the actual profit and loss data is influenced by a variety of market factors in the future, which may vary widely from the historical path, so the distribution function of the direct assumption of profit and loss is very easy to cause the deviation of the calculation result. Based on this consideration, this paper assumes that the collection of all possible profit and loss distributions of assets is $M_+=\{G_1,G_2,\ldots,G_K,\ldots\}$, where $G_1,G_2,\ldots,G_K,\ldots$ represents possible profit and loss distributions of portfolio. Obviously, we can get that the VaR value is a collection $A=\{VaR_1,VaR_2,\ldots,VaR_K,\ldots\}$ from M_+ ,and thus naturally derives the distribution law problem of the VaR value. Unlike the existing dual VaR $\left(\mu,VaR^2\right)$ concept, this paper intends to construct and study $VaR_\beta\left[VaR_\alpha(x)\right]$ (as may be called Double VaR, short for DVaR) problem under uncertain distribution based on the best (or worst) criteria.

For ease of understanding, the concept of DVaR is introduced through a life example. In a high school in a district, each grade is divided into classes, it is assumed that the third grade has twenty classes, each class has 100 people, each class is parallel relationship. If every student at the whole grade takes the college entrance examination, the content of the test paper is uniform, and the teachers in each class can only get the student scores of their own class, assuming that there is no way to know the scores of other classes. At this point, teachers can get statistics on the student achievement of their own teaching class, such as the lowest or bottom 5\% of the student grade of the highest score.If you take down 5% of the scores of all classes and arrange them in a collection of G series from one to twenty classes in class order, G={75, 78, 77, 65, 64, 70, 79, 75, 67, 71, 72, 71, 73, 78, 69, 68, 74, 76, 66, 67}. Take the first number 75 as an example to explain it, that is, a class of students have a 95% probability of higher than 75 points, the latter number is the same. Now on the collection G take down 5% of the digits, get 64, you can get the following conclusion: in the whole grade score in the bottom 5% of students, at least 5% of the students score less than or equal to 64, at least 20% of the student score is greater than or equal to 64, that is, from the school-wide point of view, at least 0.25% of the student score is less than or equal to 64 points, At least 1% of the students scored 64 or more. Therefore, in the absence of teachers' knowledge of other class scores, through the above-mentioned processing methods, we can also obtain a general understanding of the score situation of the whole year. In some areas, college entrance examinations are in the form of volunteering before publishing scores, so students can use this method to know where their scores are in the whole grade by selecting different scale levels. Assuming that, based on past data, about 0.2% of the students in this high school cannot reach the undergraduate line each year, if a student's estimated score is 64 points, it can be predicted that he or she can basically reach the undergraduate line. In the above example, the lower digits are used, if you choose the top number of digits, you can also infer from the previous data whether a classmate can reach the score line of a famous school.

The above-mentioned score 64 is equivalent to the concept of DVaR that we propose, and twenty parallel classes are equivalent to different distributions of asset gains and losses (as mentioned above M_+ is an infinite set, in the case of limited distributions as an example), which is parallel to each other and have no priority relationship. The G collection is equivalent to the VaR value we obtain for different possible distributions, each of which can only get its own information and do not know anything about the other distributions. However, in the processing of the G collection, the DVaR, i.e. 64, in the sample, can obtain a more detailed understanding of the distribution of the tail segment fractions of the whole year scale with very limited information available. In other words, DVaR can provide us with a clearer and more complete picture of the tail risk of a collection of possible k-distributions, which neither VaR nor CVaR can achieve. Therefore, DVaR is an effective tool when dealing with practical problems where the future profit and loss distribution of assets is unknown.

The example of DVaR is not only found in life, but also common in financial markets, as illustrated here by industry rankings. Each year, the Securities Association counts and publishes the ranking of the operating performance of securities companies in the previous year, and will rank the operating data and business performance subdivided indicators, such as net capital return, economic business income, securities revenue as a proportion of revenue. For each sub-index, you get a corresponding ranking of the securities industry. By considering each sub-index, you can get an assessment of the overall ranking of any securities company in the securities industry. Here by the ranking of each sub-index of the overall ranking can be regarded as DVaR, on the basis of the original indicators, the various indicators are considered comprehensively, using different indicators for multiple screening, in order to achieve the overall evaluation of brokers, in line with the basic idea of DVaR.

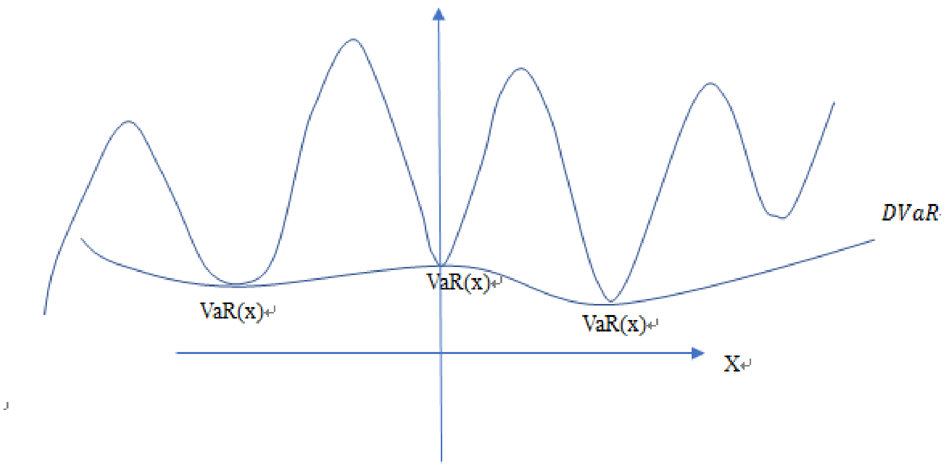

Figure 1. DVaR and VaR Relationship Map

The relationship between DVaR and VaR is as seen from Figure 1. The intuitive implication of VaR is that the maximum loss that an investor may face at a given confidence level is that, as can be seen from Figure 1, VaR itself is extremely valuable, represents a very small value of the return, and can also be understood as a loss that may be possible in the worst case scenario. DVaR is based on the uncertainty of the profit and loss distribution of assets or portfolios, and is based on the best (or worst) criterion by selecting a very small (or very large) value to DVaR for a collection of VaR values derived from different distributions. Therefore, DVaR itself has extreme value, using its extreme value to filter the VaR collection, in order to reduce the risk.

II. Definition of DVaR

The concept of VaR based on the indeterminate distribution is represented by VaR(x), and in order to facilitate differentiation and avoid conceptual confusion, the concept of DVaR presented in this paper is expressed as $VaR_\beta[VaR_\alpha(x)]$.

As can be seen from the example in the previous section, the definition of $VaR_\alpha(x)$ is the same as that of a traditional VaR by definition, except that because of the uncertainty surrounding the profit and loss distribution of the portfolio mentioned in the previous section of this chapter, this article no longer solves it by assuming the profit and loss distribution of the portfolio. Instead, use only the accurate information available to preserve its uncertainty. x represents the price of an asset or portfolio, assuming that the profit and loss distribution of the asset or portfolio is F(x), then F(x) is uncertain, resulting in a random variable $VaR_\alpha(x)$, calculated based on F(x) at a given confidence level, $VaR_\alpha(x)$ contains a collection of VaR values that correspond to different unknown distributions. Based on this, $VaR_\alpha(x)$ must obey a probability distribution G(VaR), and $VaR_\beta[VaR_\alpha(x)]$ is a VaR value for the random variable $VaR_\alpha(x)$. Therefore, by comparing $VaR_\alpha(x)$ with $VaR_\beta[VaR_\alpha(x)]$ and finding that the essence is to calculate its quantite for a random variable, except that $VaR_\alpha(x)$ is calculated as the profit and loss distribution F(x) of the portfolio, $VaR_\beta[VaR_\alpha(x)]$ the computing object is the probability distribution G(VaR) as a random variable presented in this paper.

When we consider DVaR, i.e. $VaR_\beta[VaR_\alpha(x)]$, it is very similar in form to the form of a composite function, but it is fundamentally different, unlike the mapping of the compound function from real to real. As noted earlier, the value of VaR is uncertain due to the existence of distribution uncertainty, so you can think of $VaR_\alpha(x)$ as a random variable (or an indeterminate variable), considering that it is subject to a distribution and is recorded as g(VaR). With the definition of VaR, DVaR can be defined as follows:

$$DVaR_\alpha(g(VaR)) = min\left\{\eta \in R : \int_{VaR:g(VaR),, \eta} p(VaR)dVaR..\alpha\right\} \tag{2.1}$$

That is, the density function of VaR is integrated, and the minimum value of the integral value greater than or equal to the alpha is DVaR.

Note 2.1. The problem of DVaR studied in this paper is essentially based on the corresponding problem under the uncertainty of distribution rather than the random meaning, and for the sake of simplicity, only the DVaR problem under the random meaning is considered.

III. VaR and DVaR Studies Based on Distribution Uncertainty

VaR's calculations and many portfolio optimization issues depend on the future profit and loss distribution of assets, but we cannot know the future price of the asset in advance and can only be estimated through historical data or contextual simulations. The assumption that historical data are used to fit a particular economic model, such as the GARCH model, is that past market behavior will recur, but changes in the future price of assets are influenced by a variety of market factors and are likely not to follow the historical trajectory. The future price estimated based on simulation may be largely influenced by individual subjective judgment. Therefore, with only historical observations available, a future-oriented strategy is needed to deal with possible uncertainties, which is the best way to deal with future uncertainties.

In the case of a completely unknown future profit and loss distribution of the portfolio, assume that the collection of all possible probability distributions is M_+:

$$M_+ = \{G_1, G_2, \ldots, G_K, \ldots\}$$

Among them, $G_1, G_2, \ldots, G_K, \ldots$ are different profit and loss distribution functions that represent the future of the portfolio, i.e. the elements within the collection M_+ are likely to become real profit and loss functions in the future of the portfolio, and the collection M_+ can be a finite set or an infinite set, and its size cannot be determined.

The uncertainty of the probability distribution set M_+ makes it impossible to obtain an accurate distribution of the portfolio profit and loss, and therefore the exact calculation of VaR is not possible. Therefore, the uncertainty of M_+ needs to be addressed and the optimal strategy in uncertain situations should be developed.

First, the sample obtained, i.e. historical data, is based on a more explicit revision of the M_+ to narrow the range of possible probability distribution. Drawing on the definition of fuzzy set by Kang and Li[24], Zhongfei Li improved the fuzzy set based on the previous research, and defined the uncertainty of the Ellipsoidal-ball type:

$$D_F(\gamma_1, \gamma_2) = \begin{cases} P(\xi \in \Omega) = 1 \\ P \in M_+ : (E_P(\xi) - \hat{\mu})^T \cdot \hat{\Sigma} \cdot (E_P(\xi) - \hat{\mu}) \leq \gamma_1 \\ \| Cov_p(\xi) = \hat{\Sigma} \|_F \leq \gamma_2, Cov_p(\xi) \succ 0 \end{cases}$$

Where the $\|X\|_F$ represents the Frobenius model number, the parameters $\gamma_1 \gamma_2$ represent the fuzzy level, the sample mean is $\hat{\mu}$, the covariance is $\hat{\Sigma}$, $\Omega \in R^n$ is the closed convex set containing the random vector , M_+ represents the set of all probability distributions .

By defining fuzzy sets to deal with uncertainty, it avoids the model setting error that directly assumes the profit and loss distribution of the portfolio, and only uses the available accurate information to give the possible distribution range of the assets, thus determining the specific range of the future profit and loss distribution of the portfolio and laying the foundation for subsequent processing.

Then develop a reasonable and effective strategy, from the fuzzy set to choose the required distribution, to eliminate the uncertainty of distribution. The change of confidence level under any normal distribution can be obtained by using the density control function, and this is taken as the selection criterion. Assuming given the confidence level α , the VaR of the standard normal distribution is written as χ , and the density function of the standard normal distribution is recorded as $\Phi(\cdot)$, as to the arbitrary normal distribution $f(\xi)$:

$$\int_{-\infty}^{x} f(\xi) d\xi = \int_{-\infty}^{x} [f(\xi) - \Phi(\xi) + \Phi(\xi)] d\xi = \int_{-\infty}^{x} [f(\xi) - \Phi(\xi)] d\xi + \int_{-\infty}^{x} \Phi(\xi) d\xi \tag{3.1}$$

$$\int_{-\infty}^{x} [f(\xi) - \Phi(\xi)] d\xi = \int_{-\infty}^{x} |f(\xi) - \Phi(\xi)| d\xi = \varepsilon$$

The value of $|f(\xi) - \Phi(\xi)|$ is recorded as $M(\xi)$, the value of $\int_{-\infty}^{x} \Phi(\xi) d\xi = \alpha$ is recorded as ε , because $\int_{-\infty}^{x} \Phi(\xi) d\xi = \alpha$, the amount of change in the confidence level α in given x is indicated as ε . That is, under the standard normal distribution, the probability of the portfolio loss not exceeding x is α , and under the partial normal distribution $f(\xi)$ the probability that the portfolio loss does not exceed x is α .Given x, the investor wants the value of $\alpha \pm \varepsilon$ to be as large as possible, but is not sure of the true distribution of ε , and when it is assumed that the investor is a risk aversion, develop the following target strategy:

$$\inf_{f(\xi) \in D_F} \left\{ \int_{-\infty}^{x} [f(\xi) - \Phi(\xi)] d\xi \right\} \triangleq \varepsilon_0 \tag{3.2}$$

That is, of the possible values of all the ε , select the partial distribution with the least change in the confidence level of the fuzzy set, according to which the selection will result in the worst-case distribution, and this strategy is applied when the investor is a risk aversion.

In reality, investors are not always risk-averse, such as gambling investors may be risk-averse, want to trade high risk for high returns, so for risk appetite investors developed the following strategies:

$$\sup_{f(\xi)\in D_F}\left\{\int_{-\infty}^{x}[f(\xi)-\Phi(\xi)]d\xi\right\}\geq\varepsilon_0 \tag{3.3}$$

That is, from the fuzzy set to choose the most optimistic situation, the use of loss of not more than x of the probability of the largest distribution as a real distribution, for subsequent processing.

Both of these selection criteria are considered for extreme situations, assuming that the investor is completely risk-averse or completely risk-averse, but in reality there will be risk-neutral investors, in other words, different investors may have different attitudes towards risk. Therefore, in order to develop a universal distribution selection strategy, the risk aversion coefficient λ is introduced to describe the level of risk aversion. In conjunction with the above, the following strategies are given:

$$(1-\lambda)\sup_{f(\xi)\in D_F}\left\{\int_{-\infty}^{x}[f(\xi)-\Phi(\xi)]d\xi\right\}+\lambda\inf_{f(\xi)\in D_F}\left\{\int_{-\infty}^{x}[f(\xi)-\Phi(\xi)]d\xi\right\}\geq\varepsilon_0 \tag{3.4}$$

$\xi\in R^n$ represents the acceptable set, risk aversion coefficient $\lambda\in[0,1]$, when $\lambda=0$,the second strategy is the same as the above, the investor is considered to be risk-preferred, when $\lambda=1$, the first strategy is the same as the above,, the investor is considered to be risk-averse, and when $\lambda=\frac{1}{2}$, the investor is considered to be risk-neutral. When $\lambda\in(0,\frac{1}{2})$, investors are perceived as inclined to take risks because they give a higher weight to optimism, and when $\lambda\in(\frac{1}{2},1)$, they think that investors tend to avoid risk and give a higher weight to the worst-case scenario.The value of λ determines the best case scenario for selection, and the worst case scenario is in between.

Finally, with a defined fuzzy set, given the selection strategy, VaR can be calculated directly from a defined basis, with the profit and loss distribution function determined

$$VaR_\alpha(x,P)=min\left\{\eta\in R:\int_{\xi:l(x,\xi)\leq\eta}p(\xi)d\xi\geq\alpha\right\}$$

The VaR here can be understood as the α quantile of the distribution function P, which is written as:

$$VaR_\alpha(x,P)=P_x^{-1}(\alpha)$$

$$P(x)=\int_{\xi:l(x,\xi)\leq\eta}p(\xi)d\xi$$

The value of VaR is obtained if the distribution function or probability density function of the asset is known.

This section refers to Zhongfei Li's method of dealing with uncertainty, and deals with the uncertainty of VaR as mentioned earlier. In the event that the future profit and loss of the asset is unknown, provide an effective strategy for investors with different levels of risk aversion to help them choose a reasonable distribution to estimate the maximum possible loss of the portfolio. This strategy solves the VaR calculation deviation caused by the incorrect model setting and individual subjectivity, and it is of great practical significance to deal with the complex attitude of different investors in the face of risk. In addition to being an effective strategy for dealing with VaR uncertainty, it can be further applied to other financial issues with uncertainties, such as portfolio optimization.

The following examples illustrate how the above strategies can be applied to portfolio optimization, and the mean-VaR model is optimized using the strategies proposed by the author. The D_G of the set is defined first, D_G consisting of a distribution set that filters the fuzzy set D_F by a utilization formula (3.5)

$$D_F(\gamma_1,\gamma_2)=\begin{cases} P(\xi\in\Omega)=1 \\ P\in M_+:(E_P(\xi)-\hat{\mu})^T\cdot\hat{\Sigma}\cdot(E_P(\xi)-\hat{\mu})\leq\gamma_1 \\ \|Cov_p(\xi)=\hat{\Sigma}\|_F\leq\gamma_2, Cov_p(\xi)\succ 0 \end{cases}$$

Next, the optimized mean-VaR model is given:

$$\begin{aligned} &\min_{P\in D_G} VaR_\alpha(x,P) \\ &s.t.\quad E_P(\xi)^T x .. \rho \\ &\qquad x\in\chi \end{aligned} \tag{3.5}$$

χ represents a collection of acceptable portfolios, $E_P(\xi)^T x$ represents the expectations of the portfolio.

Compared with the traditional mean-VaR model, new constraints and the introduction of fuzzy sets are added. The definition of fuzzy set is driven by sample data, eliminating the deviation of model setting skewed and simulating distribution function skewed using other methods, and using only accurate sample information to define the range of asset income distribution. The new conditions are used to further constrain the distribution of fuzzy concentration, and the nether boundary is set by using the strategy proposed in this paper to set the nether, requiring that the confidence level of the $f(\xi)$ obtained by the new strategy greater than or equal to $\alpha+\varepsilon_0$, in order to improve the confidence level, and the use of VaR for portfolio optimization under this constraint condition. And the new constraints can be changed according to the different attitude of investors to risk, this paper is only to take this as an example, the introduction of risk aversion parameters can also meet the needs of investors with different levels of risk aversion. The new model takes into account the investor's attitude towards risk on the basis of reducing deviation, which has strong practical significance.

Note 3.1. This section uses the constraints formed by density control functions to filter the fuzzy set D_F generated by the sample driver, thus establishing the distribution set D_G, and thus defining and processing the indeterminate distribution. Therefore, the distribution in the D_G can be used to calculate VaR, and a collection of VaRs is obtained $A\{VaR\,|\,VaR: D_G\to R\}$. If VaR is a continuous form, the distribution law of VaR can be obtained by the collection A, according to the distribution law of VaR, the distribution law of DVaR can be obtained, and if VaR is a discrete form, the value of DVaR can be obtained directly to the collection A.

Note 3.2. Further, depending on the investor's different level of risk aversion, ε_0 or λ can be set different , you can get different D_G, resulting in different sets of A. Different VaR distribution laws can be obtained by different sets ad, so that different DVaR values can be obtained according to the distribution of VaR. Because different investors' aversion to risk and tolerance are different, the selection of the ε_0 and λ will vary, so the law of the VaR distribution inferred from the set A is uncertain. Based on this, the following two sections use the typical distribution to discuss the distribution law of VaR, and draw the resolution of DVaR.

IV. Combination selection questions based on mean-DvaR

4.1 Sample Selection

The Shanghai 50 Index is based on a scientific and objective method, selecting the Shanghai stock market large-scale, liquidity of the most representative 50 stocks to form a sample stock, in order to reflect the Shanghai stock market the most market impact of a number of leading enterprises in the overall situation. The Shanghai 50 Index has been officially released since January 2, 2004. The goal is to establish an investment index that is active in trading, large-scale, mainly as the basis of derivative financial instruments. The SSE 50 Index adjusts its constituent stocks every six months according to the principle of combining sample stability and dynamic tracking, and the adjustment time is consistent with that of the SSE 180 Index. Temporary adjustments to the sample may also be made in exceptional cases. In order to be more representative of the empirical research, this paper selected 10 component stocks from the Shanghai Stock Exchange as a sample for analysis. The basis for selection is to cover as many different industries as possible, and not suspended during sample selection. The data for this article comes from the Cathay Ehn database. The final selection of 10 stocks is shown in Table 4.1.

Table 4.1. Sample stock code and name

600016	Minsheng Bank	600028	Sinopec
600029	Southern Airlines	600276	Hengrui Pharmaceuticals
600837	Haitong Securities	601166	Xingye Bank
601186	China Railway Construction	601288	Abc
600031	San Yi Heavy Industries	601857	China Petroleum

The sample was selected for a total of 1,339 trading days from March 31, 2014 to September 18, 2019, with daily earnings calculated based on the daily closing price of each share. Assuming that p_t the price of the t-day asset, p_{t-1} is the price of the t-1 day asset, there are generally two ways to calculate the return on financial assets, one is a simple rate of return: $E = \frac{p_t - p_{t-1}}{p_{t-1}}$, and the other is the yield on numbers: $E = \ln \frac{p_t}{p_{t-1}}$. Simple yield is suitable for the situation of single period, the positive yield is suitable for the calculation of multiple periods, because of its more accurate and good statistical nature, so the asset pricing is now generally used in the aspect of arithin yield. This paper also uses the yield of the arithion to calculate the daily earnings of individual stocks. For stocks on the day of suspension, select the closing price of the previous trading day as the closing price for the day.

4.2 Sample Analysis

Due to the large sample size, select 4 stocks to draw the yield chart, as shown in Figure 2-5.

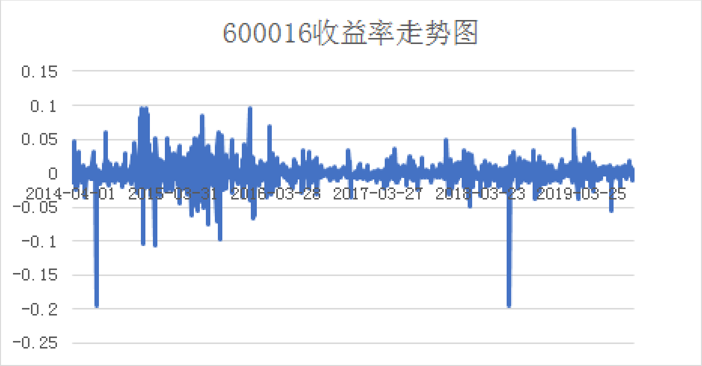

Figure 2. 600016 yield Chart

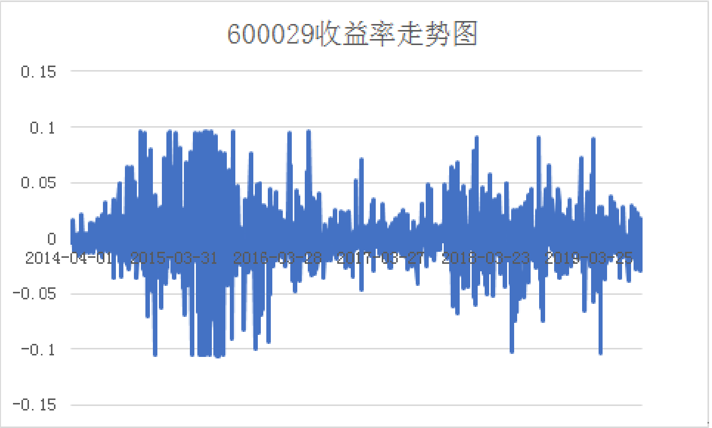

Figure 3. 600029 yield Chart

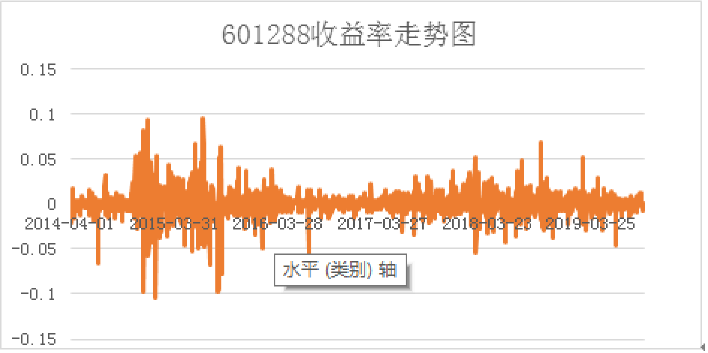

Figure4. 601288 yield Chart

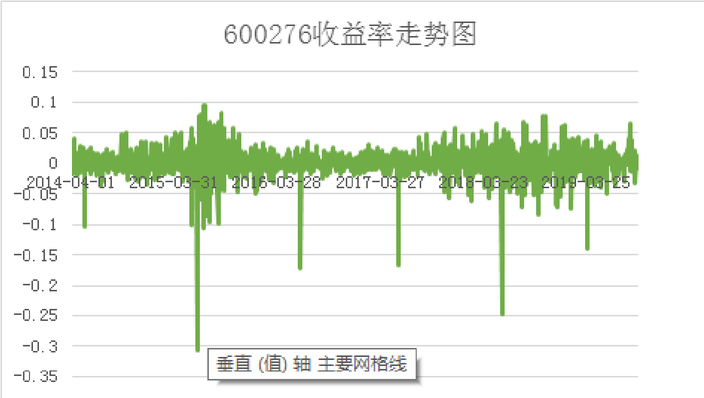

Figure 5. 600276 yield Chart

A comparison of Figure 2-5 shows that the range of 600029 yield fluctuations is small but frequent, 601288 yield volatility range is small, the most stable, the two stocks are basically concentrated in the range of the volatility of the two stocks on the range of [-0.1,0.1]. The range of fluctuations in 60016 is concentrated on the range of [-0.2,0.1], and the yield of -0.2 has occurred twice. The range of 60276 volatility is concentrated in the range of [-0.3,0.1], and the yield has twice fallen below -0.2. The above yield chart gives a visual chart of the fluctuations of the yield of 4 representative stocks, and the following table shows the specific values of the average yield and variance of the selected sample stocks during the observation period, see Table 4.2, and the covariance matrix and related coefficient matrix of the sample stocks, see Table 4.3.

Table 4.2. Sample Stock Expected Yield and Variance

Stock code	Expectation	Variance	Stock code	Expectation	Variance
600016	-0.00017	0.000335	600028	0.00000886	0.000331464
600029	0.00074	0.000873	600276	0.000652	0.000656
600837	0.000379	0.000636	601166	0.000453	0.00032
601186	0.00064	0.000804	601288	0.000271	0.00025
600031	0.000713	0.000649	601857	-0.00013	0.000306

The data in Table 4.2 is represented by a scatter plot for comparative analysis, as shown in Figure 6.

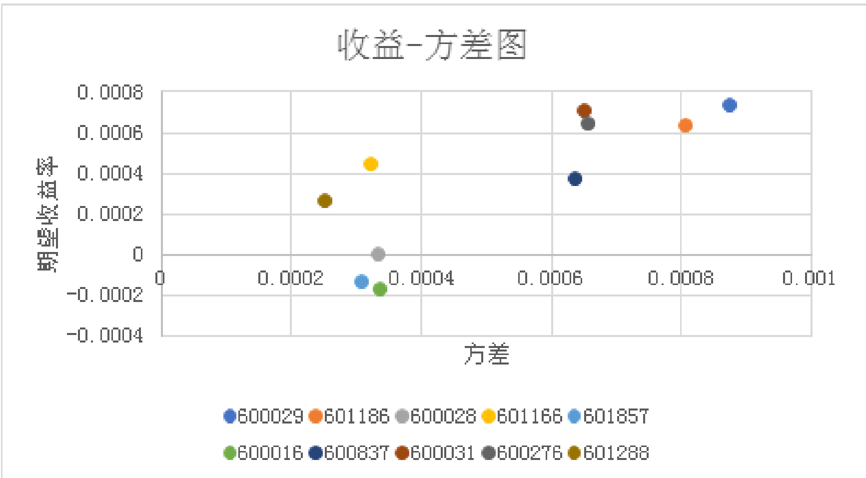

Figure 6. Variance and returns for sample shares

Available from Figure 6, the smallest variance is 601288, whose yield is in the middle of the sample, suitable for risk-averse investors; The largest variance is 600029, at the same time its yield is also the highest, belongs to the typical high-risk high-yield investment choice; There are two stocks with negative expected returns, 601857 and 600016, of which MinSheng Bank's yield is the lowest of all samples; 600028's expected return rate far less than the yield of other stocks, in the higher-yielding stocks, 600031 and 600276 risk is relatively low, compared with 601186, higher yield, lower risk, is a relatively high-quality stock.

Table 4.3. Sample Share Covariance Matrix $(\times 10^{-4})$

	600016	600028	600029	600276	600837	601166	601186	601288	600031	601857
600016	3.35	1.70	1.68	1.04	2.15	2.27	2.16	1.90	1.58	1.67
600028	1.70	3.31	2.34	1.18	2.31	2.01	2.53	1.83	2.11	2.65
600029	1.68	2.34	8.73	2.27	3.34	2.32	3.89	1.82	3.88	1.89
600276	1.04	1.18	2.27	6.56	2.05	1.25	1.62	0.68	2.34	1.03
600837	2.15	2.31	3.34	2.05	6.36	2.64	3.44	1.90	3.36	2.17
601166	2.27	2.01	2.32	1.25	2.64	3.20	2.43	2.11	2.23	1.79
601186	2.16	2.53	3.89	1.62	3.44	2.43	8.04	2.02	4.05	2.29
601288	1.90	1.83	1.82	0.68	1.90	2.11	2.02	2.50	1.47	1.70
600031	1.58	2.11	3.88	2.34	3.36	2.23	4.05	1.47	6.49	1.77
601857	1.67	2.65	1.89	1.03	2.17	1.79	2.29	1.70	1.77	3.06

As can be seen from Table 4.3, if the value solder in the covariance matrix is greater than 0, there is a positive correlation between any two stocks, i.e. the trend of change in any two stocks is consistent. However, since the values are very small, it is not possible to draw more conclusions.

Table 4.4. Sample stock correlation coefficient matrix

	600016	600028	600029	600276	600837	601166	601186	601288	600031	601857
600016	1.00	0.51	0.31	0.22	0.47	0.69	0.42	0.66	0.34	0.52
600028	0.51	1.00	0.43	0.25	0.50	0.62	0.49	0.63	0.45	0.83
600029	0.31	0.43	1.00	0.30	0.45	0.44	0.46	0.39	0.52	0.37
600276	0.22	0.25	0.30	1.00	0.32	0.27	0.22	0.17	0.36	0.23
600837	0.47	0.50	0.45	0.32	1.00	0.58	0.48	0.48	0.52	0.49
601166	0.69	0.62	0.44	0.27	0.58	1.00	0.48	0.75	0.49	0.57
601186	0.42	0.49	0.46	0.22	0.48	0.48	1.00	0.45	0.56	0.46
601288	0.66	0.63	0.39	0.17	0.48	0.75	0.45	1.00	0.37	0.61
600031	0.34	0.45	0.52	0.36	0.52	0.49	0.56	0.37	1.00	0.40
601857	0.52	0.83	0.37	0.23	0.49	0.57	0.46	0.61	0.40	1.00

Correlation coefficients are statistical indicators that reflect the correlation between variables. Correlation coefficients can also be seen as covariance: a special covariance that excludes the effect of two variable series and is standardized, which eliminates the effect of the magnitude of the two variables and simply reflects the similarity between the two variables in each unit. As can be seen from Table 4.4, the stock correlation coefficients of the same industry are generally relatively high, more than 0.5, there is a strong correlation.

Next, Table 4.5 shows the bias and peakness of the ten stocks.

Table 4.5. Bias and peakness of sample stocks

	Bias	Peakness
600016	-1.52114	24.65873
600028	-0.60419	7.126034
600029	-0.05474	2.93764
600276	-2.58896	26.20004
600837	0.076353	4.975348
601166	0.272293	6.975624
601186	0.148014	4.029943

601288	-0.20295	9.265186
600031	-0.27642	4.009454
601857	-0.07902	9.348355

Bias is a measure of the skewed direction and degree of statistical data, the normal distribution of the bias is 0, when a set of data bias is greater than 0, we think it is normal, less than 0, is considered to be negative bias. If the absolute value of the bias is greater than 0.5, the data bias is obvious. As can be seen from the table above, more than half of the stock bias value is less than 0, the performance is on the image is left- and the rest is right-side, there is no completely symmetrical distribution. Most of the stocks are less than 0.5, only 3 stocks have a significant bias, respectively, 600016, 600028, 600276. The kurtosis is a statistic that describes the steepness of the data distribution and is used financially to describe whether there are spikes in the data. The kurtosis of normal distribution is 3, and if the kurtosis of the data exceeds 3, there is a peak phenomenon. From Table 4.5, it can be seen that, in addition to 600029' peak value of 2.94 close to 3, the peak value of other stocks are significantly more than 3, there is a significant peak phenomenon. From this, the stock market is a peak phenomenon.

The peak phenomenon of the stock market is proved by the peak and bias of the stock market, and the following is further observed by drawing the income histogram. First, the 10 stocks are recorded as A, B... J, draw a histogram of its revenue distribution, as shown in Figure 7 below

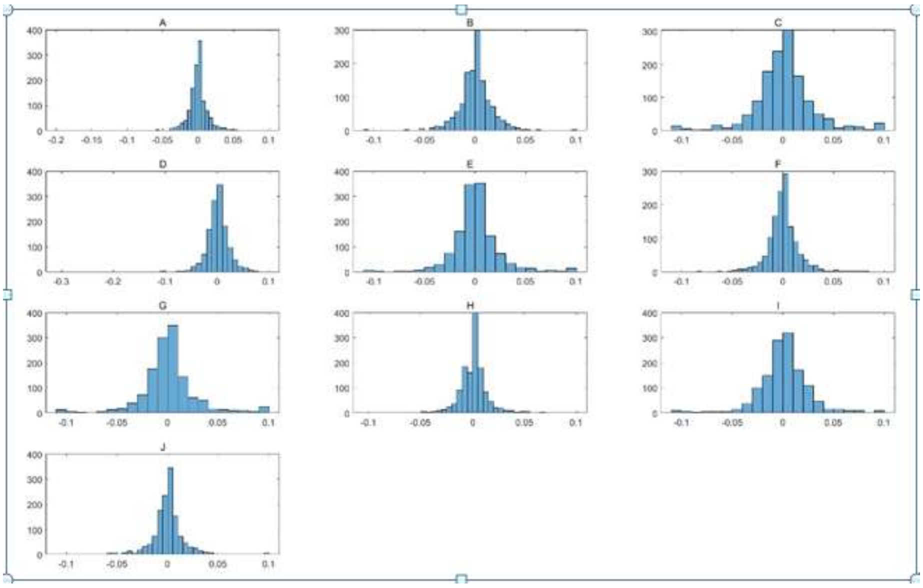

Figure 7. The histogram of the distribution of returns on sample shares

As can be seen from Figure 7, the distribution of 600837 Haitong Securities is closest to normal, other samples of earnings have different degrees of peak phenomenon; The yield distribution histogram of the sample stock is a good proof that there is a clear peak and tail phenomenon in the earnings distribution of stocks, so the normal distribution assumption is unreasonable.

Then use the historical simulation method to simulate the distribution of stock earnings, set the number of simulations 2000, produce a random sample, get the income distribution histogram as shown in Figure 8 below

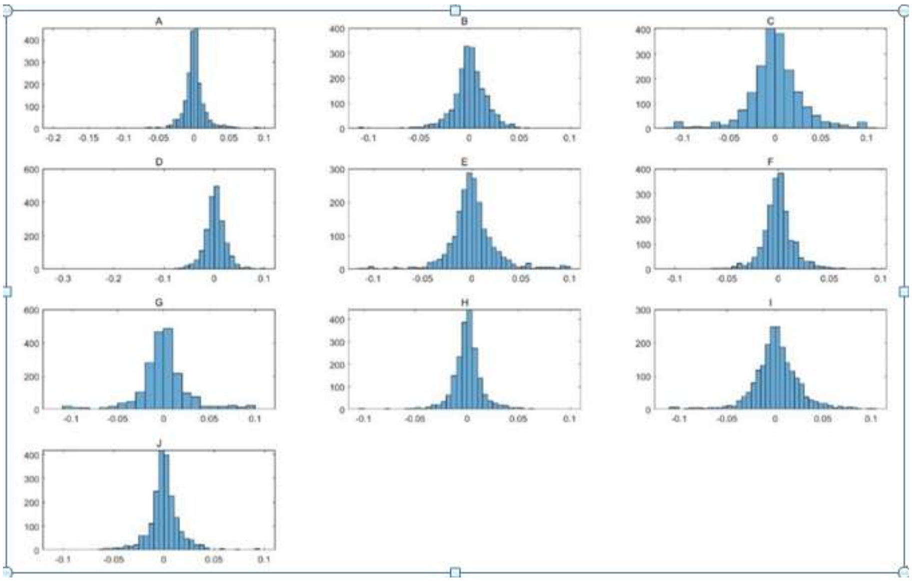

Figure 8. The histogram of benefit distributions using empirical distribution simulations

Next, the profit and loss distribution histogram of the portfolio containing 10 sample stocks is plotted using matlab, and the resulting profit and loss histogram is fitted using the normal distribution and the common spike thick-tailed distribution, as shown in Figure 9 below. As can be seen from the following figure, the yield distribution histogram is fitted using Logistic distribution, stable distribution, normal distribution, non-parametric distribution and t distribution respectively, and the best fitting effect is non-parameter distribution, followed by t-distribution, stable distribution is also basically in line with the characteristics of yield data peak thick tail, and the coincident degree is high. As can be seen from the graph, the normal distribution and Logistic distribution fitting effect is not ideal, mainly at the peak of the value is small, does not meet the peak characteristics of financial data.

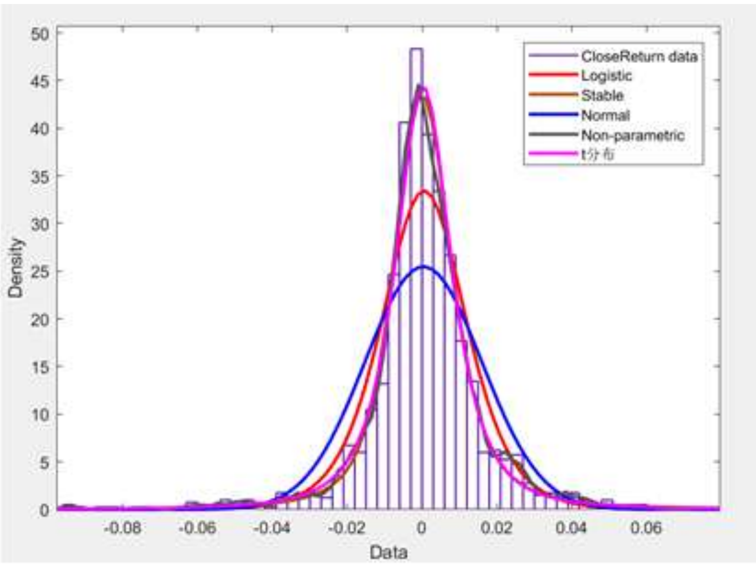

Figure 9. Portfolio profit and loss distribution fit comparison chart

4.3 Portfolio selection based on the Mean-DVaR

4.3.1 Mean-DVaR model building

The concept of DVaR and the characteristics combined with the uncertain distribution situation in chapter 3 of this paper can be seen, based on the best (distributed) scenario of the mean-DVaR model can be equivalent to the mathematical model representation as follows:

$$\min_{x\in R^n} \inf_{P\in D_G} VaR_\alpha(x,P)$$

$$s.t.\quad \begin{cases} E_P(\xi)^T x..\rho \\ xe^T = 1 \\ x\in\chi \end{cases} \tag{4.1}$$

Of these, $M_+ = \{G_1, G_2, \ldots, G_K, \ldots\}$,where $G_1, G_2, \ldots, G_K, \ldots$ represents the different profit and loss distributions that the portfolio may have in the future, D_G is a fuzzy set defined above, $E_P(\xi)^T x$ represents the expectations of the portfolio, χ represents a collection of acceptable portfolios, and $\inf VaR_\alpha(x,P)$ represents the very small value of VaR taken from all possible distributions.

4.3.2 Effective frontier comparison of VaR models in different distributions

As can be seen from the previous, the DVaR is proposed based on the uncertainty of distribution or changes in the confidence level that cause VaR to be uncertain, so the application of DVaR to the actual need to consider the impact of uncertainties. The empirical part of this paper expects to obtain a large number of portfolio income distribution through simulation method, so as to draw the effective frontier of the portfolio. The selection of simulation methods will inevitably have a direct impact on the results, and the impact of distribution uncertainty on the effective frontiers of the portfolio is considered here by selecting two different methods, historical simulation and Monte Carlo simulation. The previous section has simulated the distribution of stock earnings using historical simulation method, and has drawn a histogram of the income distribution, so that the results can be used directly to draw an effective frontier. Next, the author also uses the Monte Carlo simulation to get the portfolio, drawing an effective frontier. Figure 10 is a Monte Carlo simulation of the portfolio yield risk scatter plot, where the number of simulations set to 2500 times, i.e. 2500 sets of portfolio weight vectors and the corresponding return-VaR value.

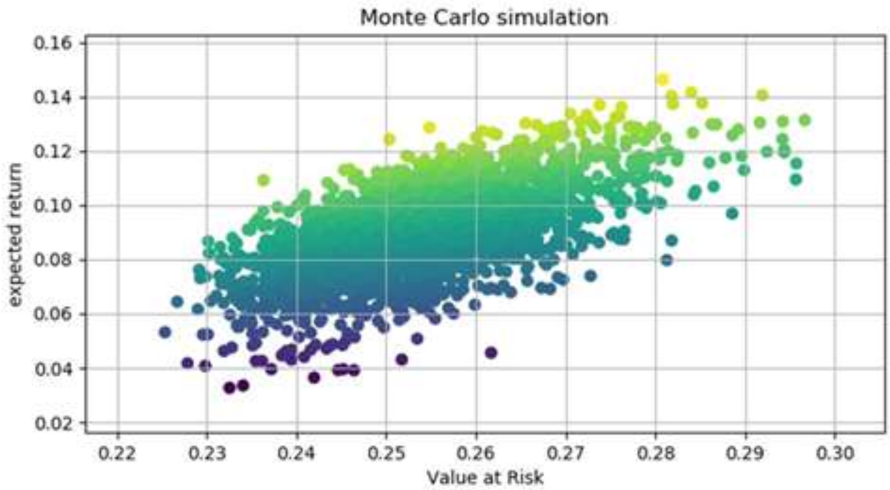

Figure10. Monte Carlo Simulation Scatter Plot

Next, using the results obtained by the historical simulation method and the Monte Carlo simulation method, a comparison of the portfolio's effective frontiers is drawn, as shown in Figure 11 below

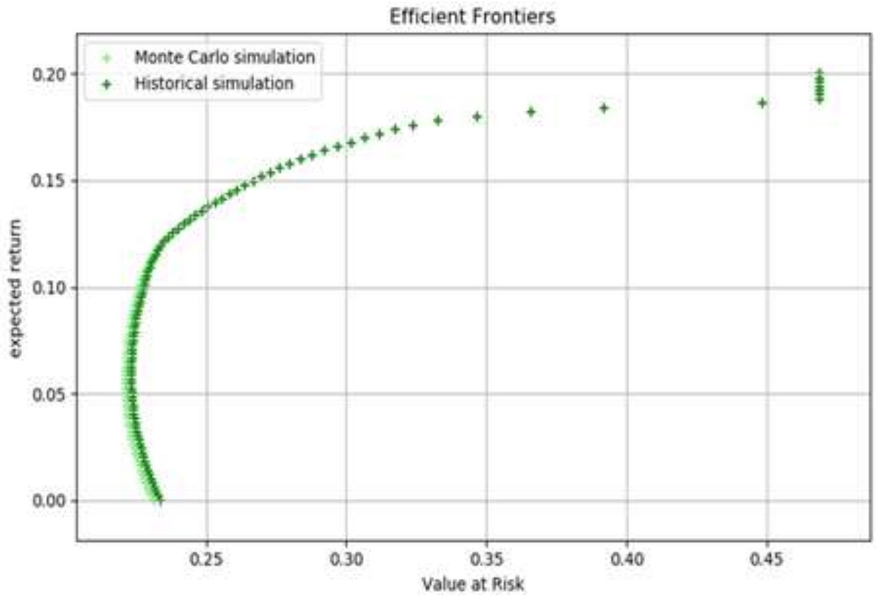

Figure11. Monte Carlo Simulation and Historical Simulation Effective Frontier Comparison

4.3.3 Effective Frontier Establishment and Analysis of DVaR

The effective front of VaR corresponding to different simulation methods can be obtained from Figure 11, and the effective front line of DVaR is drawn by envelope analysis technology. The essence of DVaR is to use extreme value to filter VaR in uncertain situations to reduce risk. Therefore, for the valid boundaries obtained by different simulation methods, If, as shown in Figure 11, the Monte Carlo simulation has a slight left shift over the effective boundary of the historical simulation method, the effective edge of DVaR should be formed by including all feasible sets; This means that the effective boundary formed by DVaR can envelope all feasible sets of the mean-VaR under different distributions, i.e. as shown in Figure 12.

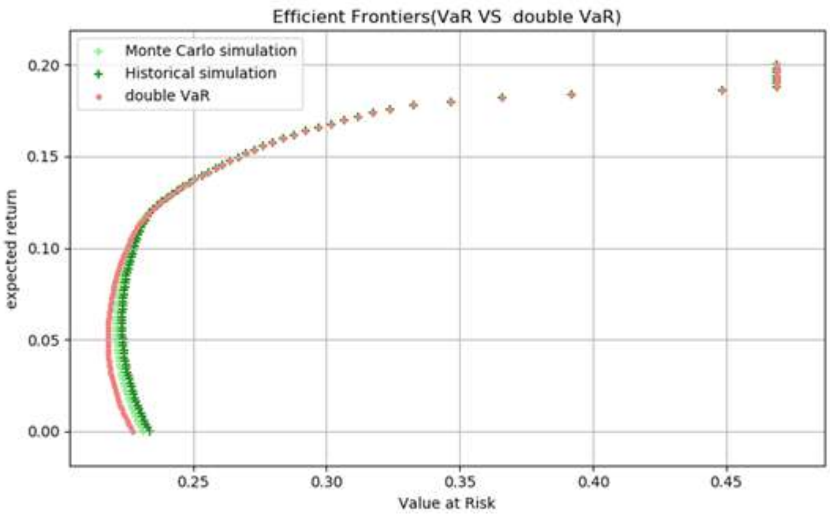

Figure12. DVaR and VaR Effective Frontier Comparison Chart

This paper uses different simulation methods to consider the impact of uncertainties on VaR, plots the effective boundary of mean-VaR, and uses envelope analysis to obtain the effective boundary of DVaR, and applies the proposed DVaR concept to the combination selection model. It can be seen that dVaR is used to deal with uncertainty and maximize portfolio efficiency for the effective boundary deviation of the portfolio caused by uncertainty.

As can be seen from Figure 12, the feasible set outlined by DVaR Effective Boundary covers all feasible sets obtained by different simulation methods, and optimizes the mean-VaR model in terms of efficiency. All points on the right side of the red envelope curve are a viable set of DVaRs, and for any point on both green curves, when the yield is given, a lower risk point is always obtained on the red curve. The application of DVaR makes the portfolio effectively boundary left and the portfolio efficiency improves.

When the author uses Monte Carlo simulation to generate a portfolio, it is found that the difference in the number of simulations also leads to the deviation of the effective boundary, so the number of simulations of 2500 and 5000 times, respectively, and the mapping of scattered points is drawn for comparison, as shown in Figure 13.

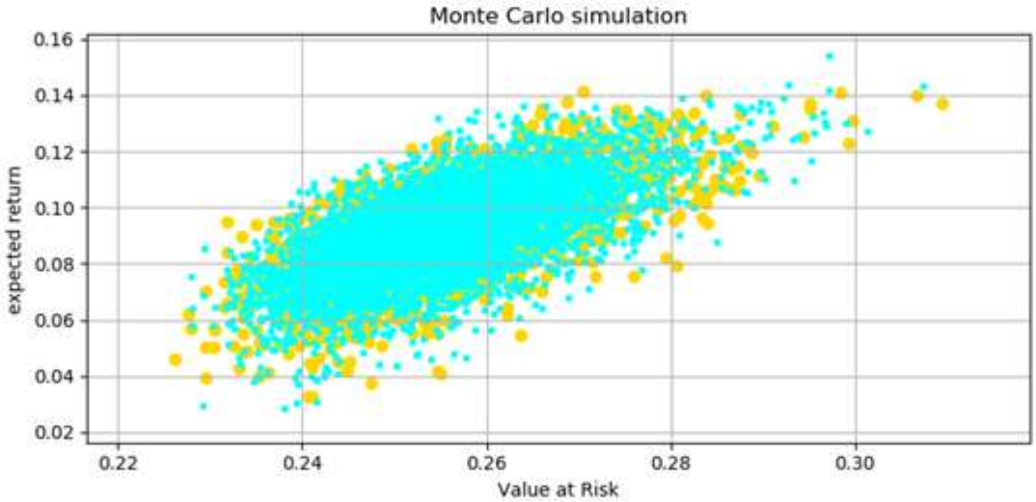

Figure13. 2500 and 5,000 Monte Carlo simulations

Among them, the yellow scatter was the result of 5000 Monte Carlo simulations and the blue scatter point was the result of 2500 Monte Carlo simulations. The effective frontier produced by the 5000 Monte Carlo simulations, which is readily available in Figure 13, is located on the effective front produced by 2500 simulations, and the effective front of DVaR at this time should be made up of the benefit-risk points from the two simulations, as shown in Figure 14

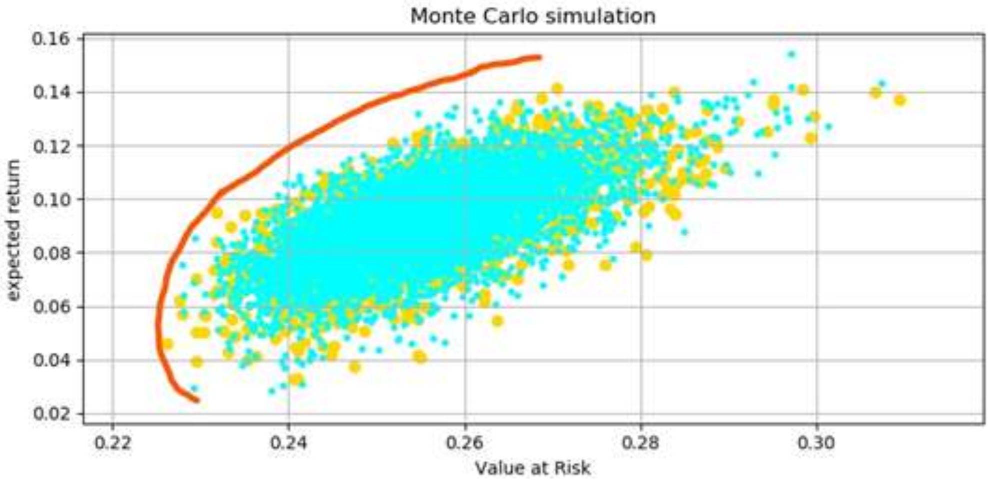

Figure14. DVaR Effective Frontier

Compared with the above analysis, the application of DVaR makes the effective boundary of the portfolio shift to the left, i.e. reduces the risk of the portfolio under the condition that the portfolio return is unchanged, so it is of practical significance. In this paper, the influence of the uncertainty of the simulation method on the portfolio, as well as the effect and optimization effect of DVaR in this case, is considered mainly from the perspective of the profit and loss distribution of the portfolio, and the empirical analysis shows that the effective frontier of the portfolio has been optimized.

Select 10 portfolios from the effective fronts of VaR and DVaR to draw their portfolio ratios, as shown in Figures 15 and 16 below

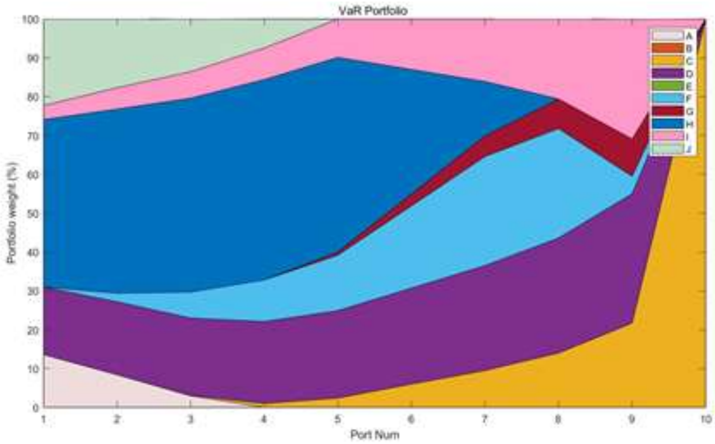

Figure15. Asset weights under the VaR model

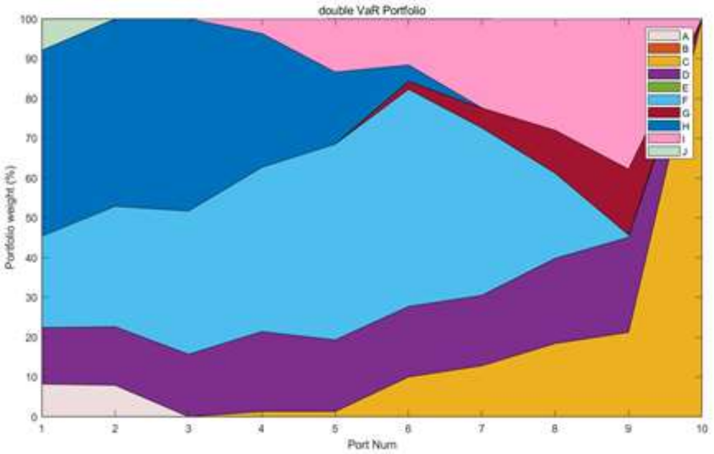

Figure16. Asset weights under the VaR model

From the above-mentioned asset proportion chart, it can be seen that 600028 Sinopec and 600837 in the portfolio of the proportion of 0, in the DVaR model, 601857 and 601166 accounted for a significant increase in the proportion, The corresponding proportion of 600031 decreased significantly, combined with the yield variance of the sample shares, it can be seen that 601857 and 601166 are low-risk assets, 600031 risk is higher. It can be seen that in the effective boundary of DVaR, the proportion of low-risk assets has been increased and the proportion of high-risk assets decreased, so the overall risk of the portfolio is lower than the mean-VaR model, which is consistent with the conclusion of the effective front shift of the portfolio reached in the previous text.

4.4 Summary of this chapter

Through the above empirical analysis, the author will put forward the DVaR reality, combined with the portfolio-related theory, made a exploration of its practical application. In this paper, the extreme value of DVaR is used to optimize the portfolio, considering the influence of distribution uncertainty on the portfolio, selecting different simulation methods as the variable factors, and using DVaR to maximize the effective cutting-edge efficiency of the portfolio under the influence of uncertainty. Comparing the effective front of DVaR with the mean-VaR model, the results show that the proposed DVaR risk measurement and the combination selection based on the mean-DVaR effectively improve the combined investment effect of the mean-VaR model.

References

[1]. A Gaivoronski and G. C. Pflug, Value-at-risk in Portfolio Optimization: Properties and Computational Approach, Social Science Electronic Publishing, 2005, 7(2): 2004-2005.
[2]. A. J. Patton, J. F. Ziegel and R. Chen, Dynamic semiparametric models for expected shortfall (and value-at-risk), Journal of Econometrics, 2019, 211(2): 388-413
[3]. B. Sebastian, Combining Value-at-Risk forecasts using penalized quantile regressions, Econometrics and Statistics, 2018, 8: 56-77
[4]. C. Acerbi and D. Tasche, On the coherence of expected shortfall, Journal of Banking and Finance, 2002, 26: 1487-1503
[5]. C. Martins-Filho, F. Yao and M. Torero, Nonparametric estimation of conditional value at-risk and expected shortfall based on extreme value theory. Econometric Theory, 2018, 34(1): 23-67
[6]. C. Guermat and R.D.D. Harris, Robust conditional variance estimation and Value-at-Risk, Risk, 2001, 4: 25-41
[7]. E. S. Manganelli, CAVaR: Conditional Autoregressive Value at Risk by Regression Quantiles, Journal of Business & Economic Statistics, 2004, 22(4):367-381
[8]. F. Jamshidian and Y. Zhu, Scenario simulation: Theory and methodology, Finance and stochastics, 1996, 1(1): 43-67
[9]. G. Geenens and R. Dunn, A nonparametric copula approach to conditional Value-at-Risk, arXiv:1712.05527, 2017
[10]. J. Kˇret´ınsk´y and T. Meggendorfer, Conditional value-at-risk for reachability and mean payoff in Markov decision processes, Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science. ACM, 2018: 609-618
[11]. L. Lin, A theoretical and Empirical Study on multi VaR-mean portfolio optimization, Times Finance, 2017, 3: 25-26
[12]. M. Bottone, M. Bernardi and L. Petrella, Unified Bayesian Conditional Autoregressive Risk Measures using the Skew Exponential Power Distribution, arXiv:1902.03982, 2019
[13]. M. B. Girach, S. Oberoi and S. P. Chakrabarty, Is being Robust beneficial: A perspective from the Indian market, arXiv:1908.05002, 2019
[14]. P. Artzner, F. Delbaen, J. M. Eber, and D. Heath, Coherent measures of risk, Mathematical Finance, 1999, 9(3): 203-228
[15]. P. Jorion, Value at Risk: the benchmark for controlling market risk. Chicago: Irwin Professional, 2000, 14(2): 47-56
[16]. R. Gerlach, Z. Lu and H. Huang, Exponentially Smoothing the Skewed Laplace Distribution for Value-at-Risk Forecasting, Journal of Forecasting, 2013, 32(6): 534-550
[17]. R. Gerlach and C. Wang, Forecasting risk via realized GARCH, incorporating the realized range, Quantitative Finance, 2016, 16(4): 501-511
[18]. R. T. Rockafellar and S. Uryasev, Optimization of conditional Value at Risk, Journal of Risk, 2000, 2:21-42
[19]. R. T. Rockafellar and S. Uryasev, Conditional Value at Risk for general loss distributions, Journal of Banking and Finance, 2002, 26: 1443-1471
[20]. T. S. Beder, VAR: Seductive but dangerous, Financial Analysis Journal, 1995, 51(5):12-24
[21]. W. Zhang, Double VaR Model and Its Application Research, Master's thesis, Nanjing University of Science and Technology, 2019
[22]. W. Zhang, S. Zhang and P. Zhao, On Double Value at Risk, Risks, 2019, 7(1): 1-22
[23]. Y. Zhang and S. Nadarajah, A review of backtesting for value at risk, Communications in statistics-theory and methods, 2018, 47(15): 3619-3639
[24]. Z. Kang, X. Li and Z. Li, Data-driven robust mean-CVaR portfolio selection under distribution ambiguity, Quantitative Finance, 2019, 19(1): 105-121