Extreme Risk Analysis of Personal Insurance Claim Based on Block Maxima Method

Tian Yaqiong

Department of Economics, Central University of Finance and Economics, Beijing, China

Email address: tianyq1992@126.com

To cite this article:

Received: July 31, 2018; Accepted: August 22, 2018; Published: September 21, 2018

Abstract: Recent years, the portion of personal insurance, including life insurance, health insurance and accident insurance, were getting larger and larger as the development of insurance market. Besides, the extreme risk of claims always exists in personal insurance. The domestic and foreign personal insurance practices have confirmed that mastery the extreme risk of claims can help insurance company pricing insurance products accurately. Therefore, the paper focused on quantifying the extreme risk of claims in personal insurance. Firstly, the principles of VaR (Value at Risk), extreme value theory, and Block Maxima Method (BM model) were sorted out, and then calculated VaR by theoretically derived. Furthermore, claim amounts of personal insurance in Beijing, Shanghai, Shaanxi Province, Henan Province, Inner Mongolia and Hainan province of China during 2005-2014 were chosen as samples. According to statistical analysis, the claim amounts datum with a same character of sharp peak and fat tail were filtered out, which contained accident insurance in Beijing, Shaanxi Province, Henan Province, Inner Mongolia, and Hainan Province as well as health insurance in Shanghai and Inner Mongolia. Lastly, the different time series of claims data were modeled by GEV distribution respectively, obtained the shape parameter, the position parameter, and the scale parameter, and then measured the extreme risk of each claims data based on BM model to get VaR of corresponding claims. The results show that the extreme risk of claims is more likely to arise in personal accident injury insurance, which exist in most regions. Since the occurrence of accident insurance does not conform to law of large numbers, its risk of claims is difficult to control. However, the extreme claim risk in health insurance has a relatively lower probability, whereas its claim VaR tends to be higher than that of personal accident injury insurance in extreme cases. Therefore, health insurance should be the focus of risk management in insurance company.

Keywords: Personal Insurance, BM Model, GEV Distribution, Claim Amount, VaR

1. Question Raised

Usually, personal insurance is divided into life insurance, health insurance and accident insurance; In the insurance market, life insurance is often separated from the others which are called non-life insurance. In the market, life insurance companies are generally responsible for life insurance, but there are no clear boundaries, sometimes non-life insurance business was also managed by life insurance company. Whether it is a life insurance company or a property insurance company, the health insurance business will be operated accordingly, while accident insurance is often managed by property insurance company.

The life insurance market in China has been recovering for more than 30 years, that has experienced rapid growth for decades. In the last few years, the concentration of personal insurance market has increased year by year. Besides, the phenomenon of oligopoly has become more serious in this market, which has gradually weak the profit of non-oligopoly companies. The extreme risk of claims has led to excessive claims payments, the premium income couldn’t cover huge expenditures either. Therefore, it is extremely important to measure the risk of personal insurance claims accurately, which can not only enlarge the range of this study, but can also promote the insurance companies in their expectation and management of risk.

2. Literature Review

With regard to the risk of personal insurance claims, as well
2.1. Research on Risk of Personal Insurance Claims

Zhang Ruiwu believed that the control of claims risk is directly related to service quality, it is an important indicator of testing the life insurance company’s business management level as well. Meanwhile, Zhang Ruiwu started with the internal and external risk factors of insurance company to analyze various reasons for the formation of claims risk, also, she put forward countermeasures to prevent claims risk from standardizing claim procedures or against insurance fraud. [1] Liu Yuhuan, Fang Rongjun pointed out that the new “Insurance Law” has had a certain impact on life insurance companies, among which the new incontestable clause are likely to increase the moral hazard and adverse selection of policyholders, cause mortality loss—the loss caused by the difference between the scheduled death rate and the actual death rate. It is one of the reasons for life insurance loss—to life insurance companies, and then lead to the rise of loss ratio—the ratio of indemnity to premium income within a certain accounting period. [2] Zhang Yi, Wu Haibo indicated that the claims of health insurance involves multiple links, only when each link is properly handled could the efficient and quality of claims be ensured. [3]

Yan Su and Fu Jiatao pointed out that the VaR model, as a popular risk management tool in the world, should be applied to risk management for life insurance companies. They also put forward five specific applications: investment risk, risk quantification of various data in insurance, performance evaluation, information disclosure, and risk limitation management. [4]

2.2. The Study of Introducing Extreme Value Theory into Claims Risk

As for the Extreme Value Theory (EVT), foreign scholars have an early explore on it. Longin proposed the use of EVT on risk management for the first time. As the huge claims in the insurance market continue to climb, some scholars have tried to apply EVT to insurance pricing. [5] McNeil used the threshold model (POT model) to fit the insurance loss distribution and discusses how to apply EVT to estimate the tail size of extreme loss distribution and the high quantile of loss distribution. [6] Guillen M, Prieto F and Sarabia J M used the Pareto strictly stable distribution to stimulate claims data and conducted related research. [7] Xu Haiyan, Bao Haiming, SHI Yongxia through empirical research, proposed that Pareto strictly stable distribution could better fit some insurance claims data. [8] Zhao Zihong, Li Xingxu [9], Ren Jing, Zhang Jiesong [10], Hao Junzhang, Cui Yujie [11] all use the POT model to perform a corresponding empirical to analyse the huge amount of insurance claims data. As for the application of Block Maxima Method (BM model), which is another method of EVT, Qian Yiping, Lin Xiang, Chen Zhiya conducted empirical research on loss data of the operational risk from Chinese commercial banks, pointed out that the BM model has the estimation ability surpass the sample, furthermore, it could obtain a more accurate result with less data. [12] Hua Yongjun, Gao Yuandong, Zhang Zongyi used the BM model to measure extreme risks in Shanghai and Shenzhen stock markets. [13] Liu Fei, Zheng Xiaoya took the advantage of BM model to measure the extreme risks of oil prices. [14]

It can be concluded from the above literature that, on one hand, the POT model has a better forecasting effect, which could be used in catastrophe insurance. On the other hand, the research on the application of BM model is mostly used in bank risk and stock market risk based on its ability to identify periodic characteristics, obviously, this advantage also applies to life insurance claims risk that with periodic characteristics. Therefore, this paper attempts to use BM model to measure the risk of three types of personal insurance, then calculate the claims VaR and comparative the results.

3. The Basic Principle of Claims VaR

This section begin with a brief description about the concept of VaR and EVT, followed by an emphasis on the specific methods and steps for calculating VaR using BM model.

3.1. The Related Theory of VaR

Value at Risk, which are called VaR, usually refers to the maximum loss that financial assets may suffer during a given period of time under the normal market risk and a certain level of confidence. Namely, while the confidence level is \( \alpha \),

\[
P(\Delta V \geq \text{VaR}) = 1 - \alpha
\]

Among them, \( P \) represents probability; \( \Delta V \) represents the loss of assets during the holding period.

There are lots of advantages of using VaR to measure risks, for one thing, the risk status of organizations which depends on a specific value can largely reduce exchange costs among various departments within the organization; For another, not only could VaR reflect the scale of the loss, but also embody the probability of the loss. Through VaR values with different confidence levels, disparate degrees of risks can be grasped. Moreover, as an accurate statistical method, VaR can precisely reflect the risk of the organization compared with other subjective and traditional risk management methods.

3.2. The Basic Principle of EVT

Extreme Value Theory is a model that specifically deals with the tail characteristics of data, which calculated VaR estimates by deriving a certain critical level of risk loss distribution with a certain confidence level, according to these, the extreme risk of corresponding data is estimated and predicted accordingly. The most important feature of EVT is to model the tail of the sample data, only concern the distribution of extreme values of the data, ignore the distribution of the overall sample data. Even if the distribution of the sample population is unknown, the variation...
of the sample extremum could also be gotten according to the sample data.

The Extreme Value Theory includes peaks over threshold (POT model) and block maxima method (BM model), wherein the POT model is a method that models observation data exceeding a threshold value and could effectively use the extreme values of the data; The BM model is to model the maximum value with a group. If you know the claims data of a certain insurance category for several years, you can get the monthly, quarterly, or annually maximum statistics of claims.

3.3. The Block Maxima Method

The block maxima method, referred to as BM model, is a traditional model of the extremum method. The core idea of which is to divide the time series into several intervals, then select the maxima in each interval to form a new time series. The generalized extreme value distribution is used to model new series composed by this maximum value, the corresponding parameter estimate are obtained by Maximum Likelihood Estimation in turn. Finally, the value of VaR is calculated.

3.3.1. The Generalized Extreme Value Distribution

The generalized extreme value distribution, which referred to as GEV distribution, can be roughly divided into three types: Weibull distribution, Frechet distribution, and Gumbel distribution. These three types of distribution can all be represented by the cumulative distribution function of GEV distribution:

\[
F(x, \mu, \sigma, \xi) = \begin{cases} 
\exp \left[ -\left( \frac{x - \mu}{\sigma} \right)^{\frac{1}{\xi}} \right] & \xi \neq 0 \\
\exp \left[ -\exp \left( -\frac{x - \mu}{\sigma} \right) \right] & \xi = 0
\end{cases}
\tag{1}
\]

From the above formula, \( \xi \) is the shape parameter, \( \mu \) is the position parameter, \( \sigma \) is the scale parameter; \( 1 + \xi \left( \frac{x - \mu}{\sigma} \right) > 0 \), that is, while \( \xi < 0 \), \( x < (\mu - \sigma) / \xi \); while \( \xi > 0 \), \( x > (\mu - \sigma) / \xi \). According to formula (1), the probability distribution density function of GEV distribution can be obtained as follows:

\[
f(x, \mu, \sigma, \xi) = \begin{cases} 
\frac{1}{\sigma} \exp \left[ -\left( \frac{x - \mu}{\sigma} \right)^{(1+\xi)} \right] & \xi \neq 0 \\
\frac{1}{\sigma} \exp \left[ -\exp \left( -\frac{x - \mu}{\sigma} \right) \right] & \xi = 0
\end{cases}
\tag{2}
\]

3.3.2. Estimation of BM Model Parameters

Suppose that the sample time series is \( r_1, r_2, \ldots, r_t \), and the sample is divided into \( h \) mutually disjoint subintervals. There are \( k \) observations in each subinterval, namely, \( t = kh \), then the financial time series can be expressed as \( r_{(i-1)+j} \), where \( i = 1, 2, \ldots, h \), \( j = 1, 2, \ldots, k \). Let \( r_{k,i} \) denote the maximum value in \( i \) subinterval, which is the peak value. Then, the time series composed by the interval maxima can be modeled by the GEV distribution, that is, the set formed by the maximum value of each subinterval is regarded as the \( h \) observations of GEV distribution, let \( \hat{r}_{k,i} \) replace \( x \) in the equation (2), the density function can be obtained.

The likelihood function of the maximum set of subintervals thus obtained as:

\[
L(\hat{r}_{k,1}, \ldots, \hat{r}_{k,h} | \xi, \mu, \sigma) = \prod_{i=1}^{h} f(\hat{r}_{k,i})
\tag{3}
\]

Firstly, take the logarithm of both sides of equation (3), then find the partial derivative of \( \xi \), \( \mu \), and \( \sigma \) respectively, and make each of the partial derivative equal to 0, we can get the maximum likelihood estimate of the shape parameter \( \hat{\xi} \), the position parameter \( \hat{\mu} \), and the scale parameter \( \hat{\sigma} \). This estimation is unbiasedness and effectiveness.

3.3.3. VaR Calculation Based on BM Model

After calculating the maximum likelihood estimation of the parameters of the generalized extreme value distribution, the estimated value of the shape parameter \( \hat{\xi} \), the position parameter \( \hat{\mu} \), and the scale parameter \( \hat{\sigma} \) are substituted into the cumulative distribution function expression of the generalized extreme value distribution, it can be seen in the above formula (1). Then the quantile of GEV distribution at a certain probability level can be obtained. Let \( \tau^* \) be the upper tail probability, so \( F(r) = 1 - \tau^* \), take the logarithm of both sides of the equation, the \((1-\tau^*)\) quantile is as follows:

\[
r_{1-\tau^*} = \left[ \frac{\mu - \sigma \ln \left( -\ln(1 - \tau^*) \right)}{\xi} \right]^{\frac{1}{\xi}} \tag{4}
\]

\( r_{1-\tau^*} \) is the \((1-\tau^*)\) quantile of the subsample maximum set, namely VaR. If each observation in each subinterval is irrelevant or weakly correlated, the following result can be got,

\[
1 - \tau^* = \Pr(r_{k,i} \leq r_{1-\tau^*}) = \left[ \Pr(r_t \leq r_{1-\tau^*}) \right]^k
\tag{5}
\]

For the time series \( r_1, r_2, \ldots, r_t \), let \( \tau \) be the upper tail probability, the \((1-\tau)\) quantile of \( r_t \) is \( r_{1-\tau^*} \), that is, \( 1 - \tau^* = \Pr(r_{k,i} \leq r_{1-\tau^*}) \), from equation (5), the VaR of the time series \( r_1, r_2, \ldots, r_t \) can be found as follows,

\[
VaR_{1-\tau^*} = \left[ \frac{\mu - \sigma \ln \left( -\ln(1 - \tau) \right)}{\xi} \right]^{\frac{1}{\xi}} \tag{6}
\]

For example, when the upper tail probability is equal to 0.05, while \( \xi \neq 0 \), \( VaR_{0.95} = \mu - \sigma \left( 1 - \left( -\ln 0.05 \right)^2 \right) / \xi \).

In general, data with periodic features is more suitable for BM model.
4. The Basic Principle of Claims VaR Metrics

In order to obtain the VaR of personal insurance claims, this part mainly empirically analyzes the data based on the BM model of EVT theory. Firstly, use SPSS to describe the data, to which the periodic distribution and thick tail characteristics of the data can be obtained. Secondly, use R to fit the data by GEV distribution so that the fitting parameters of the model can be found. Then judge the fitting condition and calculate the corresponding VaR. Finally, make a comparative analysis of the measures of personal insurance claims risk in different provinces and draw the conclusion.

4.1. Data Selection and Analysis

This part of the paper collects the data of monthly claims for personal insurance in six regions of China the year from 2005-2014. Personal insurance includes health insurance, accident insurance and life insurance. Therefore, the analysis is based on the difference between the three types of personal insurance. Table 1 shows the statistical analysis of health insurance, accident insurance and life insurance data from eight regions.

Table 1. Descriptive statistics of personal insurance claims by region (Unit: 10000 yuan).

<table>
<thead>
<tr>
<th>Region</th>
<th>Health Insurance</th>
<th>Accident Insurance</th>
<th>Life Insurance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing</td>
<td>1.29E5</td>
<td>2.54E4</td>
<td>1.44E5</td>
</tr>
<tr>
<td>Shanghai</td>
<td>5.11E5</td>
<td>1.66E4</td>
<td>5.11E5</td>
</tr>
<tr>
<td>Shaanxi Province</td>
<td>7536.4</td>
<td>1.59E5</td>
<td>4.58E4</td>
</tr>
<tr>
<td>Henan Province</td>
<td>1.71E4</td>
<td>4.24E5</td>
<td>1.21E3</td>
</tr>
<tr>
<td>Inner Mongolia</td>
<td>6.16E4</td>
<td>1.59E5</td>
<td>7.67E4</td>
</tr>
<tr>
<td>Hainan Province</td>
<td>1.21E3</td>
<td>1.59E5</td>
<td>1.82E4</td>
</tr>
</tbody>
</table>

Table 1 covers the descriptive statistics of three types of personal insurance claims data in Beijing, Shanghai, Shaanxi Province, Henan Province, Inner Mongolia and Hainan province. From the statistical results, by comparing the average and the median of various types of personal insurance claims, it can be judged that the claims data of the three types of personal insurance in each region are right-skewed, also, the skewness and kurtosis are both deviated from the normal distribution, which are 0 and 3 respectively. Among them, some kurtosis is greater than 3, the data exhibits the characteristic of “sharp peak”, while other claims data does not have this characteristic. After screening, the claims data with “sharp peak” contains the following data: the accident insurance in Beijing, Shaanxi Province, Henan Province, Inner Mongolia Autonomous Region, Hainan Province, and the health insurance in Shanghai as well as Inner Mongolia Autonomous Region, the kurtosis of these seven types of data are both significantly larger than 3. Continue to test the thick tail characteristics of the data, and use Q-Q plot to detect its tail. It can be seen that the sample points of the following sample data index Q-Q plot all have convex features.

From the above conclusion, the claims of both health insurance and accident insurance may be subject to extreme risks. As for life insurance, it is difficult to suffer extreme claims due to its own characteristic that have the long-term insurance period and the relative stability of claims. The BM model can better handle the tail of the data to estimate the trend of extreme risks. Therefore, only the above seven types data with characteristics of “sharp peak and heavy tails” will be fitted by BM model in next part, and the claims risk of life insurance will be ignored.
Figure 2. QQ-plot of health insurance in Shanghai.

Figure 3. QQ-plot of accident insurance in Shaanxi.

Figure 4. QQ-plot of accident insurance in Henan.

Figure 5. QQ-plot of health insurance in Inner Mongolia.

Figure 6. QQ-plot of accident insurance in Inner Mongolia.

Figure 7. QQ-plot of accident insurance in Hainan.
In order to understand the trend of claims more intuitively, figure 8 and 9 show a line chart based on seven types of personal insurance. Obviously, the monthly claims of personal insurance have periodic characteristic, and the maximum value in each cycle is increasing in recent years. Furthermore, these claims data all have the extreme claims. Therefore, in order to get a quantifiable result, it’s vital to determine the specific claims risk (VaR) for each type of personal insurance.

---

Due to the large value of the extreme claims in Beijing, Shanghai, and Henan Province, if they are placed in a figure with the claims of Inner Mongolia, Shanxi province, and Hainan Province, the data characteristics of the latter will not displayed, so they are divided into two figures.
4.2. Risk Assessment of Personal Insurance Claims Based on BM Model

Because the monthly data of personal insurance claims in each region is characterized by periodic distribution, according to which we choose the length of subsample interval k=12, and use the maximum likelihood estimation to obtain parameters of personal insurance. Select the 95% and 99% confidence levels respectively, then calculate the VaR of monthly claim at the confidence level. The results are shown in table 2.

Table 2. Measurement results of personal insurance claims risk (Unit:10000 yuan).

<table>
<thead>
<tr>
<th></th>
<th>$\xi$</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>VaR$_{95%}$</th>
<th>VaR$_{99%}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accident insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beijing</td>
<td>9.42</td>
<td>29223.23</td>
<td>146112.6</td>
<td>1515236</td>
<td>6.98e+12</td>
</tr>
<tr>
<td>Shaanxi Province</td>
<td>0.07</td>
<td>54736.13</td>
<td>22061.2</td>
<td>68048</td>
<td>102577</td>
</tr>
<tr>
<td>Henan Province</td>
<td>1.70</td>
<td>38352.15</td>
<td>43156.69</td>
<td>70857</td>
<td>932448</td>
</tr>
<tr>
<td>Inner Mongolia</td>
<td>0.73</td>
<td>9983.98</td>
<td>3852.4</td>
<td>1228.9</td>
<td>29488</td>
</tr>
<tr>
<td>Autonomous Region</td>
<td>0.72</td>
<td>1587.06</td>
<td>597.12</td>
<td>1934</td>
<td>4551</td>
</tr>
<tr>
<td>Hainan Province</td>
<td>0.25</td>
<td>38647.46</td>
<td>18300.99</td>
<td>48088</td>
<td>89617</td>
</tr>
<tr>
<td>Shanghai</td>
<td>1.18</td>
<td>23220.87</td>
<td>18077.71</td>
<td>35059</td>
<td>193535</td>
</tr>
<tr>
<td>Health insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inner Mongolia</td>
<td>1.85</td>
<td>35059</td>
<td>48088</td>
<td>932448</td>
<td>6.98e+12</td>
</tr>
</tbody>
</table>

Firstly, in comparison with others, the accident insurance is more likely to have extreme claims, and its risk cannot be ignored. For the six regions which were selected in this paper, the VaR of accident insurance claims is quite different. At the confidence level of 95%, the VaR of accident insurance claims in Beijing have reached 15 billion yuan in the past ten years, while the lowest value of VaR in Hainan province is less than 20 million yuan, the former is 750 times than the latter. The VaR of accident insurance claims in Inner Mongolia Autonomous Region, Shaanxi Province and Henan Province are at the level which between 100 million to 1 billion. So, the scale of the region has an important influence on that, especially the regional GDP and population density. For example, the population density of Beijing had reached 1311 people per square kilometer in 2014, and the growth of GDP was 213.3 billion yuan, which would have a great influence on the high claims for accident insurance.

Secondly, there is a small probability that health insurance experience the extreme risk. According to the descriptive analysis of selected sample, only two areas have experienced extreme risks, which were fitted by BM model and get the VaR of claims. From the results, it can be seen that under the confidence level of 95%, the VaR of health insurance claims in Shanghai is about 500 million yuan, while Inner Mongolia is about 350 million yuan. Claims data of accident insurance and health insurance in Inner Mongolia are characterized by sharp peak and fat tails. Compare them and the result can be got that the VaR of health insurance is significantly greater than accident insurance, this is because the claims data of health insurance are generally larger than that of accident insurance. However, health insurance has a lower probability of extreme claims than accident insurance.

In addition, under the confidence level of 99%, the claims VaR of each region is obviously higher than that of 95%. This phenomenon indicates that, with higher confidence coefficient, the tail features of the claims data can be better captured so that the risk of them can be measured more accurately.

5. Conclusion

In this paper, Claims amount of personal insurance were chosen as samples among 6 regions of China during 2005-2014. According to statistical analysis, the claim amounts datum with a same character of sharp peak and fat tail were filtered out, which were then modeled by GEV distribution and risk measured based on BM model to get VaR of corresponding claims. From the analysis above, some basic conclusions can be realized in the following aspects:

Above all, as can be seen from the selected samples, the probability of extreme risk in life insurance is much smaller than that of health insurance and accident insurance, while accident insurance is most likely to have extreme risks. And, through the comparison of the results, it can be found that the probability of extreme risk in health insurance is much smaller than accident insurance. However, when it happens, the VaR of health insurance claims is likely to be greater than that of accident insurance.

Furthermore, the second point indicates that the risk of large claims in health insurance cannot be ignored, for it may affect the profitability of insurance companies and even hinder their development. But the occurrence of accident insurance does not conform to law of large numbers, so its risk of claims is difficult to control. Among these three types of insurance, the health insurance’s loss ratio shows a large volatility, and health insurance should become the focus of claims risk management.

Additionally, one of the most serious problems in the insurance market is moral hazard, which is basically unavoidable. In comparison, the probability of moral hazard in health insurance is far greater than that of life insurance and accident insurance, this is one of the reasons why the VaR of health insurance will be larger. Studies have proved that moral hazard in the health insurance market not only exists, but also the impact of moral hazard in ordinary health insurance is more serious than that of dread disease insurance. Therefore, in order to reduce the waste of resources caused by moral hazard as much as possible, insurance companies can appropriately improve the medical insurance for dread
diseases of people, and the level of protection for general daily diseases can be relatively low.

The health insurance market in China is still in its infancy, the supply of it is obviously insufficient, and the most health insurance is operated by non-professional companies. Some studies suggest that insurance companies lack the motivation to operate health insurance because it’s business does not significantly promote the performance of the insurance companies. The analysis of this paper verifies this statement to some extent, that is, the VaR with large health insurance claims will have a very fatal impact on insurance companies. In view of this, the government’s appropriate tax support for health insurance to a certain extent is also an effective strategy to promote the development of their business.

Fund Project

This article is the research and innovation project “Extreme Risk Analysis of Personal Insurance Claim Based on BMM Model”(project number: 201624) of Central University of Finance and Economics, funded by the Graduate Research Innovation Fund of Central University of Finance and Economics.

References