

Study on the Optimization of Default Point of China Listed Company by using Genetic Algorithm KMV Model

Jia Lin and Yongping Gui

School of Economics Shanghai University, Shanghai, China

102857206@qq.com, guiyongping@163.com

Keywords: KMV Model; Genetic Algorithm; Fitness Function; Distance-to-Default

Abstract. Objective of this paper is applying KMV Credit Risk Model to the credit assessment of China listed companies, the KMV Model needs to be modified in combination with the characteristics of listed companies, and setting an accurate default point is crucial. This paper uses several methods including Genetic Algorithm, Double Total independent sample t-test. It introduces Genetic Algorithm into KMV Model, and takes the Double Total independent sample t-test function as the fitness function of genetic algorithm to solve the optimal default point problem of China listed companies. The study shows the conclusion that the KMV model has adaptability in the credit risk assessment of China listed companies, and the optimal default point calculated by genetic algorithm make the KMV model has the strongest distinguishing ability.

Introduction

As the issue of the Basel III in December 2010 and the implementation of Chinese version of the Basel III, risk early warning which take credit risk, market risk and operational risk as three pillars will become the key work of china banking regulatory department and commercial banks in the future. Among these three major risks, the credit risk is critical and also the most important and difficult issues in the risk measurement.

In 1993, America KMV company launched KMV credit risk assessment model, it is one of the key credit risk assessment model in the international community; based on Metron option pricing model, KMV credit risk assessment model evaluates the company's future insolvency and default risk by using the financial data and stock trading data of listed companies as references [1,2].

KMV company believes that when the company's assets are less than or equal to a default point (DP) which is between its short-term liabilities and total liabilities, then the risk of default is greater. Although the default point is given in the model, this point is not necessarily suit the actual situation in China, because the default point is calculated by a large number of historical date based on the American companies. Thus, based on the previous study results, this paper solves the optimization problem of default point of China listed companies by introducing genetic algorithm and Double Total independent sample t-test as the fitness function of genetic algorithm [3,4].

References Review

Research on KMV Model by Domestic Scholars. The domestic scholars' studies on KMV model can be divided into three aspects: In the early stage, some domestic scholars' studies mainly focused on the adaptability of the model in China listed companies. Wei Lu (2003) introduced the KMV model and its role in the credit risk assessment of China listed companies; Feng Xue (2003), Ling Zhang (2004), etc. affirmed the capability of KMV model in the credit risk assessment of China listed companies.

After that, some domestic scholars began to focus on the industry segmentation of KMV model. Xiaohong Chen (2008) introduced the KMV model into the credit risk assessment of small and medium-sized enterprises; Yuanfan Zhou (2009) applied it to insurance company credit risk measurement to strengthen the insurance supervision; Haiyun Zhou (2015), Degong Ma (2015), Xuekai Wang (2015), etc. applied it to local government debt default analysis.

In addition, there are some domestic scholars committed to the KMV improved-model empirical study, also achieved good results. Leining Li(2007), Ruowei Ma(2014), Shubin Jiang(2016), etc. determined the default point of China listed companies by comparing the default situation of ST and non-ST companies under different default point and combining with the statistic tests. Jinghai Feng(2016) solved the optimization problem of default point by introducing genetic algorithm.

Genetic Algorithm. Genetic algorithm is a mechanism calculation model which simulate natural selection and genetic of Darwin's biological evolution theory. It is a kind of self-organizing and adaptive artificial intelligence technology which could solve the extreme value problem. It can find the optimal solution by simulating the natural evolutionary process, and it is widely used in the fields of automatic control, engineering design and social sciences. It is also suitable for solving the nonlinear optimization problem.

Its main idea is: representing gene coding of population by using binary character encoding. After an initial population is randomly generated, successive evolution produce increasingly better approximate solution according to natural selection. In each generation, individuals are selected according to the size of the individual fitness in the problem domain. And the genetic operators in natural genetics are used to produce next generation. Through the selection, crossover and mutation, the population with low adaptability will be phased out, and the one with high adaptability will be left. Finally, the optimal individuals in the last population are decoded and can be used as the approximate optimal solution of the problem

In the genetic algorithm, the selection of fitness function directly affects the convergence rate of genetic algorithm, because the genetic algorithm is based on the fitness to determine the individual fitness for the survival of the fittest. In this paper, the optimization problem of KMV model default point is taking Double Total independent sample t-test as the fitness function of the genetic algorithm to solve the problem of the optimal default point of China listed company by comparing the difference of the total mean difference between the non-ST and the ST companies.

KMV Model and Parameter Setting

Calculation Step. According to American scholars Merton and Scholes's Black-Scholes-Merton option pricing theory, KMV model see listed company's equity value as a call option, liabilities as a put option, and assume that the company's asset value is subject to normal distribution for some time in the future. Generally, the KMV model can be divided into three steps:

Step 1: Calculate the company's asset value(V_A)and asset value volatility (σ_A). According to the Black-Scholes-Merton option pricing model, the relevant conclusions shown as below:

$$\left\{ \begin{array}{l} V_E = V_A N(d_1) - e^{-rT} DN(d_2) \\ \sigma_E = \frac{N(d_1)V_A\sigma_A}{V_E} \\ d_1 = \frac{\ln \frac{V_A}{D} + (r + \frac{\sigma_A^2}{2})T}{\sigma_A \sqrt{T}} \\ d_2 = d_1 - \sigma_A \sqrt{T} \end{array} \right. \quad (1)$$

where V_E is company equity value; σ_E is equity value volatility; D is face value of company's debt, which can be observed; r is risk-free rate; T is debt maturity. V_E 、 σ_E 、 D 、 r 、 T can be obtained through the listed companies' stock circulation and parameter setting, then calculate V_A 、 σ_A using simultaneous Eq.1 .

Step 2: Calculate Distance-to-Default (DD). KMV company found that when the company's asset value fell to the book value of the debt, the default event does not necessarily occur, because

the long-term debt provides a buffer space to company. KMV company gets a default point between short-term debt and total debt based on the United States historical data. That is, in the United States, the most frequent default point in the company's asset value is less than or equal to short-term debt plus 50% long-term liabilities,

$$DP = SD + 0.5LD \quad (2)$$

KMV defines the Distance-to-Default as:

$$DD = \frac{E(V_A^T) - DP}{E(V_A^T)\sigma_A} \quad (3)$$

According to the hypothesis of the Black-Scholes-Merton model, the firm's asset value is subject to a normal distribution or a lognormal distribution, the firm asset value at future point T should satisfy:

$$V_A^T = V_A \exp[(\mu - \sigma_A^2/2)T + \sigma_A \sqrt{T} \varepsilon] \quad (4)$$

where $E(V_A^T)$ is the expected asset value in the future point T; μ is expected growth rate of firm's value; ε is the stochastic factors that affect the yield, which subject to the standard normal distribution, $\varepsilon \sim N(0,1)$.

Mathematical expectation for both sides of Eq.4 (here does not give the derivation process due to the limited space):

$$E(V_A^T) = V_A e^{\mu T} \quad (5)$$

thus, the DD is calculated as follow:

$$DD = \frac{V_A e^{\mu T} - DP}{V_A e^{\mu T} \sigma_A} \quad (6)$$

Step 3: Calculate Expect Default Frequency (EDF). After lots of studies, KMV company found that there is a mapping relationship between the DD and EDF , EDF can be calculated by the following equation:

$$EDF = \Pr(E(V_A^T) < DP) = N(-DD) = 1 - N(DD) \quad (7)$$

Parameter Setting. In the KMV model, the parameters which need to be pre-calculated and set are: market value of firm's equity (V_E), firm's equity value volatility (σ_E), the face value of company's debt (D), the risk-free rate (r), the maturity of debt (T) and the expected growth rate of firm's value (μ).

1. Firm's Equity Value (V_E). In the domestic research, people agree on the stock price behalf of the equity price; In the handling of non-tradable shares, although most of the China listed companies have completed the split share structure reform, but in the data sampling process, still found some non-tradable shares exist; so in the calculation of equity value, the number of tradable shares and non-tradable shares should be all taken into consideration, the value of the firm's equity calculation using the following equation:

$V_E = \text{Number of share outstanding} \times \text{Closing price} + \text{Number of non-tradable shares} \times \text{Net assets per share}$

2. Firm's Equity Value Volatility (σ_E). The volatility calculation is calculated from the historical data volatility and the GARCH (1,1) model. The former predicts the stock which base on the historical closing price of the sample company, the latter considers the variance of time volatility. This paper uses the historical data to calculate the equity value volatility. Under the

assumption that the stock price is subject to the normal distribution, the daily return rate of the stock can be expressed as:

$$u_i = \ln(S_{i+1}/S_i) \quad (8)$$

where S_i is closing price of day i ; S_{i+1} is closing price of day $i+1$; The standard deviation of the stock daily return is:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (u_i - \bar{u})^2} \quad (9)$$

where $\bar{u} = \frac{1}{n} \sum_{i=1}^n u_i$; the annual stock trading day is about 245 days, the relationship between the equity value volatility and the standard deviation of the daily income is:

$$\sigma_E = \sqrt{245} \sigma \quad (10)$$

3. Company's Debt (D). The book value of the Company's debts is the sum of the long-term liabilities and short-term liabilities which disclosed in the annual report of the listed company.

4. Risk-free Rate (r) and maturity of debt (T). Domestic scholars are consistent with those two parameters. The risk-free rate is set as the benchmark one-year deposit rate announced by the People's Bank. In this paper, we use the benchmark interest rate of 2016 as the risk-free rate, which equal to 1.5%, and the maturity of debt is 1 year.

5. Expected growth rate of firm's value (μ). Leining Ning (2007) used the arithmetic mean of the company's net income growth rate in the past three years to show the expected growth rate of firm's value. Yongsheng Yang (2010) compared four existing estimation method of expected growth rate of firm's value. To simplify, we assume the expected growth rate of firm's value is 0.

Optimization design of Default Point

In the KMV Model, setting default point is critical. The default point used in KMV company is based on the U.S historical database, which is more suitable for capital market in western developed countries. This point is not necessarily consistent with the actual situation in China, so it should be modified when introduction of the model.

In the design of the default point $DP = SD + \alpha LD$, some domestic scholars use step size method in the (0,1) interval to evaluate the long-term debt coefficient α . Comparing the mean differences of DD between non-ST sample company and ST sample company, scholars use the maximum value of mean differences as long-term debt coefficient. However, this method is not reasonable for two reasons. Firstly, only by comparing the sample mean difference can not represent the difference between the sample population; secondly, the value of α can only be the maximum or minimum number which is set by author, because the expression of sample's mean differences is a linear function, shown in Eq.11.

$$M = \frac{1}{n} \sum_{i=1}^n \frac{E(V_{NST_i-A}^T) - (SD_{NST_i} + \alpha LD_{NST_i})}{E(V_{NST_i-A}^T) \sigma_{NST_i-A}} - \frac{1}{m} \sum_{j=1}^m \frac{E(V_{ST_j-A}^T) - (SD_{ST_j} + \alpha LD_{ST_j})}{E(V_{ST_j-A}^T) \sigma_{ST_j-A}} = C + B\alpha \quad (11)$$

where M is the sample mean difference between the non-ST company and the ST company, and it is a constant.

Therefore, we introduce statistical test method to solve the problem of default point. This paper uses Double Total independent sample t-test method to determine the sample population mean differences' significance of non-ST and ST companies DD . $DP = SD + \alpha LD$, where $\alpha \in (0,1)$, the equation shown as below:

$$\left\{ \begin{aligned}
 t &= \frac{\overline{DD}_{NST} - \overline{DD}_{ST}}{\sqrt{\frac{(n-1)S_{NST}^2 + (m-1)S_{ST}^2}{n+m-2} \left(\frac{1}{n} + \frac{1}{m}\right)}} \\
 \overline{DD}_{NST} &= \frac{1}{n} \sum_{i=1}^n \frac{E(V_{NST_i-A}^T) - (SD_{NST_i} + \alpha LD_{NST_i})}{E(V_{NST_i-A}^T) \sigma_{NST_i-A}} \\
 \overline{DD}_{ST} &= \frac{1}{m} \sum_{j=1}^m \frac{E(V_{ST_j-A}^T) - (SD_{ST_j} + \alpha LD_{ST_j})}{E(V_{ST_j-A}^T) \sigma_{ST_j-A}}
 \end{aligned} \right. \quad (12)$$

where \overline{DD}_{NST} 、 \overline{DD}_{ST} is the sample mean of distance-to-default of non-ST company and ST company. S_{NST}^2 、 S_{ST}^2 is the sample variance of non-ST company and ST company. According to t-test, the greater the value of t , the more significant of the sample population mean differences of non-ST company and ST company. Therefore, the optimization problem of the default point can be regarded as the problem about choosing maximum value t in the internal of $\alpha \in (0,1)$.

In order to solve the value of t , we introduce the genetic algorithm to solve the optimal solution of the nonlinear function. We define the fitness function of the genetic algorithm as follows:

$$GA(\alpha) = - \frac{\overline{DD}_{NST} - \overline{DD}_{ST}}{\sqrt{\frac{(n-1)S_{NST}^2 + (m-1)S_{ST}^2}{n+m-2} \left(\frac{1}{n} + \frac{1}{m}\right)}} \quad (13)$$

Main idea: Calculate sample company's asset value and asset value volatility through the sample data. Substitute V_A 、 σ_A and DP (α unknown) into appropriate function, then find the concrete expression of the nonlinear function $GA(\alpha)$. Using the genetic algorithm method, we calculate α and the minimum value of $GA(\alpha)$ (the maximum value of t) in the internal of $\alpha \in (0,1)$, then get the optimal default point. Fig.1 is the genetic algorithm flow chart.

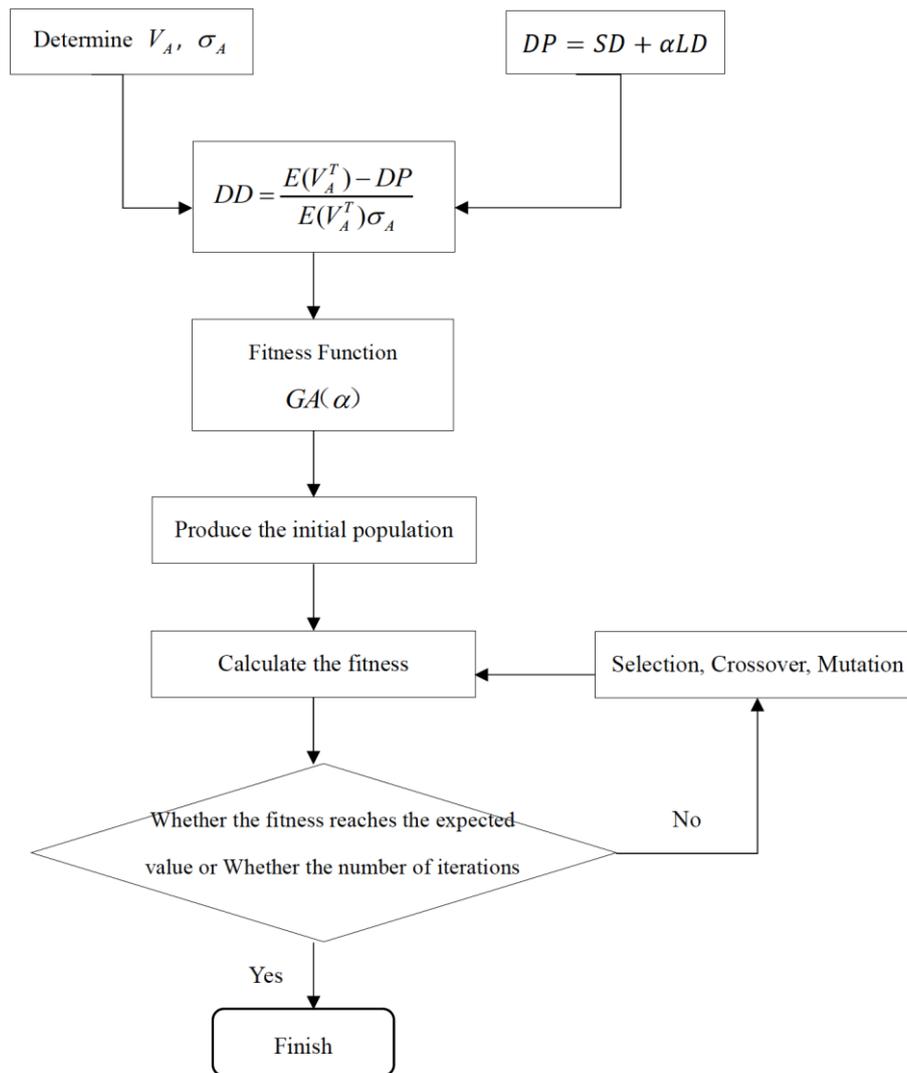


Figure 1. Genetic algorithm flow chart

Empirical Research

Sample selection and data source. Taking the availability of data into account, this paper uses ST company as a sign of default, and randomly selects 20 listed companies in Shanghai and Shenzhen stock market from January 2016 to December 31, 2016. Also taking industry distribution and sample matching into account, we randomly select 20 non-ST companies to ensure that ST and non-ST companies distribution is corresponding. The data comes from the annual report of listed companies, flush software, Oriental wealth network, etc.

Optimal default point calculation. Firstly, according to the collected sample data, the asset value and asset value volatility of ST and non-ST companies are calculated by using the *fsolve()* simultaneous equation of function in MATLAB tool, as shown in Table 1 and Table 2.

Table 1 σ_E , V_A , σ_A statistics of ST companies

Stock Name	Code	Long-term Debt (in 10,000)	Short-term Debt (in 10,000)	Equity value Volatility	Asset Value (in 10,000)	Asset value Volatility
*ST Hongsheng	600817	2464.16	1679.16	36.14%	3.2806e+005	35.69%
*ST Yunsheng	600767	231.08	46608.83	53.86%	5.7626e+005	49.55%
*ST Fantai	000629	86799.25	514562.55	50.05%	1.8846e+006	34.33%
*ST Jizhuang	000856	11020	118744.87	47.00%	4.8354e+005	34.58%
*ST Zhengmei	600121	42880.93	654830.19	43.81%	1.2406e+006	19.56%
ST Shanshui	600234	11694.35	26511.75	58.62%	4.0832e+005	53.22%
ST Jinggu	600265	280	30921.93	37.92%	4.2935e+005	35.21%
S*ST Qianfeng	600733	4632.06	9030.8	49.62%	4.0429e+005	47.97%
*ST Tianyi	000710	14824.35	28572.55	88.39%	7.5958e+005	83.42%
*ST Xinmei	600732	0	25853.35	61.04%	3.6784e+005	56.81%
*ST Shunchuan	002608	840453.26	1462137.04	59.50%	4.2255e+006	27.86%
*ST Sanwei	000755	34009.89	428458.70	60.14%	8.4188e+005	27.92%
*ST Pitu	600696	22407.02	27855.54	59.69%	4.1698e+005	52.61%
*ST Jinyu	000803	6186.36	40232.46	72.28%	2.5464e+005	59.35%
*ST Jiadian	000922	6279.19	131031.18	54.89%	7.431e+005	44.90%
*ST Huke	600608	10135.53	9332.03	51.37%	5.5372e+005	49.59%
*ST Hehua	000953	207.59	87219.86	74.82%	5.7713e+005	63.69%
*ST Dayou	600403	93230.53	699453.1191	48.72%	2.1436e+006	30.99%
*ST Changjiu	600228	3488.28	31903.07	66.34%	5.0906e+005	61.80%
*ST Aifu	600636	10074.08	99352.06	60.82%	7.0996e+005	51.59%

Table 2 σ_E 、 V_A 、 σ_A statistics of non-ST companies

Stock Name	Code	Long-term Debt (in 10,000)	Short-term Debt (in 10,000)	Equity value Volatility	Asset Value (in 10,000)	Asset value Volatility
Quanxinhao	000007	45.13	7668.44	46.04%	4.2542e+005	45.39%
Lvjing	000502	0	37144.86	50.21%	2.7804e+006	45.89%
Hanggang	600126	80308.42	976510.53	48.34%	6.5506e+006	30%
Xugong	000425	541031.15	3704815.33	31.98%	9.9624e+005	11.56%
Shanghai Energy	600508	178233.18	36331.9	32.73%	6.6378e+005	25.78%
Dasheng Cultural	600892	675.48	79838.90	53.52%	5.9539e+005	47.13%
Shennong Gene	300189	8996.06	105485	50.46%	6.0319e+006	40.90%
Greatown	600094	1794588.08	2489400.99	43.67%	5.8471e+005	13.16%
Xiyi	002265	1813.00	27476.40	63.17%	1.6958e+006	60.05%
Vantone	600246	378522.71	356220.27	45.59%	2.7886e+007	26.14%
Jiuding Investment	600053	54500.72	196309.76	53.55%	9.0138e+005	53.08%
Black Cat	002068	18278.26	380425.30	48.48%	7.0709e+006	27.38%
Leshi Internet	300104	926880.33	1248327	37.96%	1.4205e+006	26.46%
Cgn Nuclear	000881	124549.82	463604.49	50.72%	5.3226e+006	30.06%
Chint Electrics	601877	901667.26	1228475.16	45.99%	1.6609e+006	27.64%
Greatstar	002444	1113.75	134160.86	46.06%	1.2401e+007	42.36%
Salt Lake	000792	3480969.99	2176465.73	42.04%	6.062e+006	23.15%
Xishan	000983	1243193.37	2206873.92	53.83%	6.0223e+006	23.80%
Wanhua	600309	982967.19	404512.50	35.06%	1.2234e+006	27.11%
Huayi Electric	600290	8341.26	274337.17	39.56%	4.2542e+005	30.56%

Then, substitute the assets value, the asset value volatility and default point of the sample companies into the Genetic algorithm fitness function, and the final expression of $GA(\alpha)$ by using MATLAB is shown as below:

$$GA(\alpha) = -\frac{3.3286-1.5885\alpha}{\sqrt{0.8075-0.8336\alpha+0.3523\alpha^2}} \quad (14)$$

Finally, use MATLAB genetic algorithm toolbox *gatool* to find the optimal solution of the fitness function. When $\alpha=0.20497$, $GA(\alpha)$ has minimum number -3.7207; when $\alpha=0.2$, ($DP=SD+0.2LD$), the overall mean difference between ST and non-ST samples is the most significant. Fig.2 shows the best fitness of the genetic algorithm and the value of the α by using MATLAB.

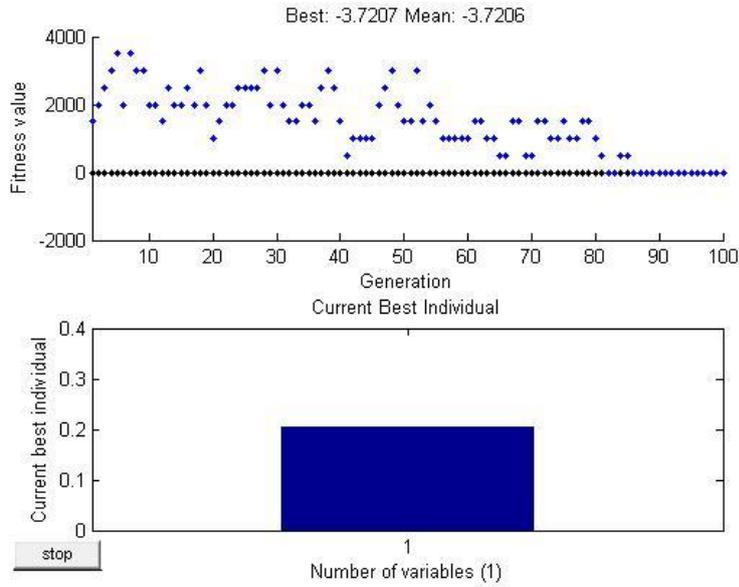


Figure 2. Genetic algorithm best fitness, value of α

Analysis. In order to verify the reliability of the optimal default point obtained by genetic algorithm method, we substitute the default point coefficient $\alpha=0.2$, $\alpha=0.5$, and $\alpha=0.1$, $\alpha=0.7$ which is randomly chose into sample; and then compare the sample mean differences between the ST and non-ST companies, the data shown in Table 3.

Table 3 Sample mean difference and t test for non-ST and ST companies

Company Type	# of sample	$\overline{DD1}$ ($\alpha = 0.7$)	$\overline{DD2}$ ($\alpha = 0.5$)	$\overline{DD3}$ ($\alpha = 0.2$)	$\overline{DD4}$ ($\alpha = 0.1$)
Non-ST company	20	2.354	2.4424	2.5751	2.6193
ST company	20	1.8584	1.8758	1.9018	1.9105
Sample mean difference		0.4956	0.5666	0.6733	0.7088
t test		3.5171	3.661	3.72	3.7155
P value		0.11%	0.076%	0.064%	0.065%

From Table 3, we can get the following three conclusions: 1. Regardless of the value α , non-ST company's distance-to-default is significantly greater than that of ST company, which indicate that the KMV model is adaptable in predicting the credit risk of China listed companies. 2. With the decrease of α , the difference of sample mean is increasing, because there is a linear relationship between the sample mean difference and α . But the difference of the sample mean is not representative of the significance of the sample population. 3. When $\alpha=0.2$, t is maximum; in other words, when $DP=SD+0.2LD$, the mean difference between non-ST and ST companies is most significant.

Expected default frequency prediction. Substitute $DP=SD+0.2LD$ into sample data, predict 40 sample companies' probability of default by using MATLAB. The results are shown in Table 4.

Table 4 Distance-to-default and expected default frequency of ST and non- ST companies

Stock Name	DD	EDF	Stock Name	DD	EDF
*ST Hongsheng	2.7833	0.27%	Quanxinhao	2.1719	1.49%
*ST Yunsheng	1.8548	3.18%	Lvjing	1.9888	2.34%
*ST Fantai	2.0911	1.83%	Hanggang	2.1429	1.61%
*ST Jizhuang	2.1688	1.505%	Xugong	3.6151	0.02%
*ST Zhengmei	2.3789	0.868%	Shanghai Energy	3.5984	0.016%
ST Shanshui	1.7462	4.039%	Dasheng Cultural	1.8662	3.1%
ST Jinggu	2.6355	0.42%	Shennong Gene	2.0044	2.25%
S*ST Qianfeng	2.0333	2.1%	Greatown	4.0114	0.003%
*ST Tianyi	1.149	12.53%	Xiyi	1.5859	5.64%
*ST Xinmei	1.6365	5.09%	Vantone	2.8508	0.22%
*ST Shunchuan	2.2047	1.37%	Jiuding Investment	1.8701	3.07%
*ST Sanwei	1.7297	4.18%	Black Cat	2.0959	1.8%
*ST Pitu	1.7536	3.98%	Leshi Internet	3.0135	0.13%
*ST Jinyu	1.4105	7.92%	Cgn Nuclear	2.1827	1.45%
*ST Jiadian	1.8307	3.36%	Chint Electrics	2.66	0.39%
*ST Huke	1.9751	2.41%	Greatstar	2.1696	1.5%
*ST Hehua	1.3327	9.13%	Salt Lake	3.319	0.045%
*ST Dayou	2.1462	1.59%	Xishan	2.4995	0.62%
*ST Changjiu	1.5145	6.49%	Wanhua	3.321	0.045%
*ST Aifu	1.6617	4.83%	Huayi Electric	2.5343	0.56%

Conclusion

The main point of this paper is setting right default point of KMV model, and solve the optimal default point problem of China listed companies by introducing the genetic algorithm and using Double Total independent sample t-test as the fitness function of genetic algorithm. Compared with the traditional method, the default point obtained by the genetic algorithm is more accurate. Due to the limited time, this paper is limited in sample selection. And by increasing the sample size, the default point will be more in line with the actual default situation of China listed companies.

References

- [1] Altman E I. Financial ratios discriminant analysis and the prediction of corporate bankruptcy[J]. *Journal of Finance*, 1968: 589-609.
- [2] Altman E I, Haldeman R, Narayanan P. Zcta analysis:A new model to identify bankruptcy risk of corporations[J]. *Journal of Banking and Finance*, 1977: 29-54.
- [3] Lee Wo-chiang. Redefinition of the KMV model's optimal default point based on genetic algorithms-Eidence from Taiwan [J]. *Expert Systems with Applications*, 2011, 38(8):10107-10113.
- [4] Kurbat M, Korablev I. Methodology for testing the level of the EDFTM credit measure [R]. White Paper, Moody's KMV Revised August, 2002.