World Applied Sciences Journal 30 (8): 938-942, 2014
ISSN 1818-4952

DOI: 10.5829/idosi.wasj.2014.30.08.265

The Evaluation of Credit Scoring Models Parameters Using Roc Curve Analysis

Dmitriy A. Garanin, Nikita S. Lukashevich and Sergey V. Salkutsan

St. Petersburg State Polytechnical University; St. Petersburg, Russia

Abstract: The possibility of the ROC curve analysis application in the estimation of credit scoring models parameters is considered. The possible parameters of such models are presented. Several criteria for determination of the optimal threshold value of credit rating are proposed. ROC curve analysis approbation using the factual data for model formalization is conducted. Recommendations for the application of ROC curve analysis in banking practice are given.

Key words: Credit scoring · Logistic regression · ROC curve · Sensitivity · Specificity

INTRODUCTION

In modern conditions the problem of credit risk management is becoming increasingly important. The requirements for the reliability of the banking system, imposed by the various regulatory bodies, credit terms and the number of credit operations, success of which directly depend on the economic situation of the borrowers, are constantly growing. In accordance with the Basel Capital Accord, known as Basel III, it is recommended for the estimation of credit quality to use an approach based on the internal banking ratings and according to which it is required to develop the mathematical models to estimate the probability of default. The analyst can use the abbreviated, structural and credit scoring models that have the greatest practical interest to allow estimation of the borrowers' credit rating [1].

Each credit scoring model can be summarized as follows:

$< I_0\ (G, L, \Phi, A) >$,

where I_0 – credit rating as a measure of creditworthiness of the borrower; G - a set of factors of the borrower's creditworthiness; L – a set of estimates for each factor from the set G; Φ - a set of weights defining the significance of each factor from the set G; A – a method for calculation I_0.

The article [1] describes a variety of approaches to the development of credit scoring models, among which the statistical and neural network methods that are traditionally used in practice and implemented in most modern banking software products. All recommendations of how to choose an approach are detailed in the article [1].

The practical credit scoring models, developed on the basis of the statistical, neural networks or fuzzy sets methods and the comprehensive interpretation of the peculiarities of their application for the purpose of credit risk analysis are presented in the papers [1, 2].

Regardless of the chosen approach, an important prerequisite for the effective implementation of credit scoring models is the reasonable choice of their parameters, required for decision making on crediting, as well as the estimation of the predictive capability of the models, that defines the classification accuracy of the borrowers. To resolve this problem it is possible to use ROC curve analysis [3].

The research objective is testing the application of ROC curve analysis to estimate the parameters and predictive capability of credit scoring models. As the information base for research an impersonal sample of the individual borrowers was captured. Based on the sample and using logistic regression as the traditional statistical tool to estimate the probability of default, a credit scoring model was designed for testing ROC curve technique.

Corresponding Author: Dmitriy A. Garanin, St. Petersburg State Polytechnical University; St. Petersburg, Russia.

MATERIALS AND METHODS

In signal detection theory, a receiver operating characteristic (ROC), or simply ROC curve, is a graphical plot which illustrates the performance of a binary classifier system as its discrimination threshold is varied. ROC curve analysis is widely used in various fields such as the theory of signal detection [4], the diagnostic tests in medicine [5], a comparison of models and algorithms in the theory of management decisions [6, 7].

Despite the fact that the approach focuses mainly on the application in medicine and technology, there is experience with the ROC curve analysis application in domestic banking practice. In the paper [1] the author examines the important problem of variables selection in the scorecard using logistic regression. The author's presented approach to variables selection depends on the calculated values of the area under the ROC curve. The ROC analysis algorithm is introduced in some new software products for the automation of credit risk management, for example, «Scorto™ Model Maestro» and «SAS Credit Scoring Solution», actively used in banks.

Any binary classifier can be obtained by logistic regression, neural networks, classification trees or using other classification techniques. The ROC curve allows us to construct the dependence of the number of correctly classified positive examples on the number of incorrectly classified negative examples [3].

The Main Parameters of the ROC Curve Analysis: Let us characterize the main parameters of credit scoring models from the viewpoint of ROC analysis. Each binary classifier involves two classes, one of them is a class with the positive outcomes and the second is with the negative outcomes. In the context of the current tasks, the positive outcome is a successful repayment of the loan (trustworthy borrower) and the negative one is credit default (unreliable borrower). The share of the true positive outcomes *TPR* (true positives rate), the share of the false positive outcomes *FPR* (false positives rate), the share of the true negative outcomes *TNR* (true negative rate) and the share of the false negative outcomes *FNR* (false negative rate) are calculated accordingly as follows:

$$TPR = \frac{TP}{TP + FN},$$

$$FPR = \frac{FP}{TN + FP},$$

$$TNR = \frac{TN}{TN + FP},$$

$$FNR = \frac{FN}{FN + TP},$$

where *TP* (true positives) – the true classified positive outcomes (true positive outcomes); *TN* (true negatives) – the true classified negative outcomes (true negative outcomes); *FN* (false negatives) – the positive outcomes classified as the negative one (false negative outcomes); *FP* (false positives) – the negative outcomes classified as the positive one (false positive outcomes).

The parameter *TPR* determines the sensitivity of the model. If a model possesses high sensitivity, it offers a greater probability of the correct recognition for the positive outcomes. The parameter *TNR* determines the specificity of the model. A model with high specificity provides a greater probability of the correct recognition for the negative outcomes. Briefly summarized, a model with high specificity corresponds to a conservative credit policy (a high level of rejected credit applications) and a model with high sensitivity corresponds to a risky credit policy (a high level of approved credit applications). In the first case, the losses from credit risk are minimized and in the second case the loss of economic benefit is minimized. The last important parameter of credit scoring models is the threshold (limit) value *C* (cutoff point). This value is essential in order to apply the model in practice and classify the new outcomes. Choosing the threshold value, the analyst can control the probability of the correct recognition of the positive and negative outcomes. When reducing the threshold value, the probability of the erroneous recognition of the positive outcomes (false positive outcomes) increases and conversely, when maximizing, the probability of the incorrect recognition of the negative outcomes increases (false negative outcomes).

ROC Curve: The ROC curve represents a set of coordinates, specified by *TPR* and *(1 - TNR)* at different values of *C*. For the perfect classifier, the graph for the ROC curve passes through the upper left corner, where the share of the false positive outcomes is equal to zero. Therefore, the closer the curve to the upper left corner, the higher the predictive capability of the model. The diagonal line (the so-called line of no-discrimination or random guess) corresponds to the "bad" classifier. Parameter *AUC* is calculated as the area under the ROC curve using for example, trapezoid rule [7] and takes values in the interval [0; 1]. The high value for *AUC* is evidence of the high quality of the model in terms of its predictive capability.

The Possible Criteria for Cut-off Point Determination: The key problem in the ROC curve analysis is to determine the acceptable threshold value on the basis of the formalized ROC curve. The possible criteria for determining the acceptable threshold value among *k* possible values are presented below:

- Ensuring the minimum allowable value of the model sensitivity TPR_{min} (criterion K_1):

$$TPR_k = TPR_{min}.$$

- Ensuring the minimum allowable value of the model specificity TNR_{min} (criterion K_2):

$$TNR_k = TNR_{min}.$$

- Ensuring the maximum value of total sensitivity and specificity of the model (criterion K_3):

$$max \{(TNR_k + TPR_k)\}.$$

- Ensuring a balance between sensitivity and specificity of the model (criterion K_4):

$$min \{|TPR_k - TNR_k|\}.$$

- Ensuring the maximum value of Youden's index (criterion K_5) [8]:

$$max \{(TPR_k + TNR_k - 1)\}.$$

- Ensuring the maximum value of reliability index (criterion K_6)

$$\max \{(\frac{TN_k + TP_k}{TN_k + TP_k + FN_k + FP_k})\}.$$

- Ensuring the minimum sum of losses from classification errors (criterion K_7):

$$min \{(S_{FP} \bullet FP_k + S_{FN} \bullet Fn_k)\},$$

where S_{FP} – cost of the false positive outcome; S_{FN} - cost of the false negative outcomes.

The greatest practical interest provides the last criterion. On the one hand, it allows linking classification errors with economic indicators, but on the other hand, the determination of the false outcomes cost is a difficult problem, requiring special research, that significantly limits the application of this criterion in practice. The analyst can roughly calculate the cost of classification errors for each false outcome on the basis of data on overdue debt and credit conditions.

Classifier Construction: Two credit scoring models based on logistic regression were defined during the statistical processing. Due to correlation between predictors, the parameters of the model may be inaccurate, resulting in a significant number of false outcomes. The matrix of pair correlation coefficients is formed and presented in Table 1. A conclusion about partial multicollinearity can be made. In this case, it is formally possible to obtain estimates of the model parameters and their exact values, but they will not be stable and will affect the predictive accuracy of the models. Considering that the research objective is testing the application of the ROC curve analysis in banking practice rather than getting the adequate practical credit scoring models, parameters of the models were found.

Table 1: The matrix of pair correlation coefficients (highlighted significant coefficients)

Q_i	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9	Q_{10}
Q_1	1,000	0,146	0,314	-0,085	-0,189	0,017	-0,050	0,182	0,071	-0,053
Q_2	0,146	1,000	-0,231	0,204	-0,143	0,230	-0,154	0,026	0,021	0,083
Q_3	0,314	-0,231	1,000	-0,280	0,147	-0,117	-0,189	0,199	-0,015	-0,080
Q_4	-0,085	0,204	-0,280	1,000	-0,259	0,092	0,136	-0,115	-0,150	0,054
Q_5	-0,189	-0,143	0,147	-0,259	1,000	0,039	-0,001	0,163	-0,021	-0,084
Q_6	0,017	0,230	-0,117	0,092	0,039	1,000	-0,111	0,025	0,007	-0,075
Q_7	-0,050	-0,154	-0,189	0,136	-0,001	-0,111	1,000	0,081	-0,073	-0,230
Q_8	0,182	0,026	0,199	-0,115	0,163	0,025	0,081	1,000	0,253	-0,179
Q_9	0,071	0,021	-0,015	-0,150	-0,021	0,007	-0,073	0,253	1,000	-0,283
Q_{10}	-0,052	0,082	-0,080	0,053	-0,084	-0,075	-0,230	-0,179	-0,283	1,000

Table 2: The ROC curve analysis results

		The threshold value, C																				
Model	Parameters	0	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	1
Z_1	TP	35	35	35	35	35	32	31	28	25	24	23	22	22	18	15	14	12	9	8	5	0
	TN	0	0	2	5	9	11	17	19	23	24	28	29	30	30	30	31	33	34	35	35	35
	FN	0	0	0	0	0	2	4	7	9	11	12	13	13	17	20	22	23	26	27	30	35
	FP	35	35	33	30	26	25	18	16	13	8	7	6	5	5	5	3	2	1	0	0	0
	TPR	1,00	1,00	0,95	0,88	0,80	0,74	0,65	0,60	0,52	0,50	0,45	0,43	0,42	0,38	0,33	0,31	0,27	0,21	0,19	0,13	0,00
	FPR	1,00	1,00	0,94	0,86	0,74	0,69	0,51	0,46	0,36	0,25	0,20	0,17	0,14	0,14	0,14	0,09	0,06	0,03	0,00	0,00	0,00
	TNR	0,00	0,00	0,06	0,14	0,26	0,31	0,49	0,54	0,64	0,75	0,80	0,83	0,86	0,86	0,86	0,91	0,94	0,97	1,00	1,00	1,00
	FNR	0,00	0,00	0,00	0,00	0,00	0,06	0,11	0,20	0,26	0,31	0,34	0,37	0,37	0,49	0,57	0,61	0,66	0,74	0,77	0,86	1,00
	$\hat{E}_1$	1,00	1,00	0,95	0,88	0,80	0,74	0,65	0,60	0,52	0,50	0,45	0,43	0,42	0,38	0,33	0,31	0,27	0,21	0,19	0,13	0,00
	$\hat{E}_2$	0,00	0,00	0,06	0,14	0,26	0,31	0,49	0,54	0,64	0,75	0,80	0,83	0,86	0,86	0,86	0,91	0,94	0,97	1,00	1,00	1,00
	$\hat{E}_3$	1,00	1,00	1,00	1,02	1,05	1,05	1,13	1,14	1,16	1,25	1,25	1,26	<u>1,28</u>	1,23	1,19	1,22	1,21	1,18	1,19	1,13	1,00
	$\hat{E}_4$	1,00	1,00	0,89	0,73	0,54	0,44	0,16	<u>0,05</u>	0,12	0,25	0,35	0,40	0,43	0,48	0,52	0,60	0,68	0,76	0,81	0,88	1,00
	$\hat{E}_5$	0,00	0,00	0,00	0,02	0,05	0,05	0,13	0,14	0,16	0,25	0,25	0,26	<u>0,28</u>	0,23	0,19	0,22	0,21	0,18	0,19	0,13	0,00
	$\hat{E}_6$	0,50	0,50	0,53	0,57	0,63	0,61	0,69	0,67	0,69	0,72	0,73	0,73	<u>0,74</u>	0,69	0,64	0,64	0,64	0,61	0,61	0,57	0,50
	AUC	0,00	0,06	0,08	0,09	0,04	0,12	0,03	0,05	0,05	0,02	0,01	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00

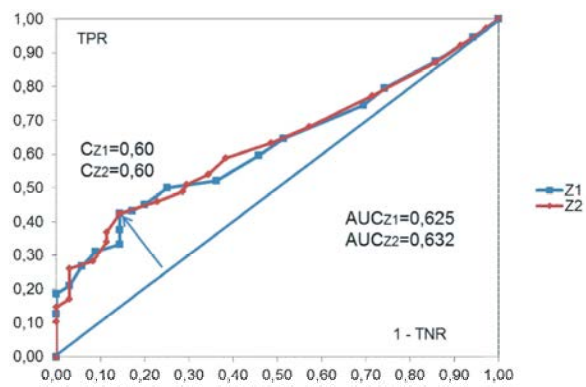

Fig. 1: The constructed ROC curves for both models

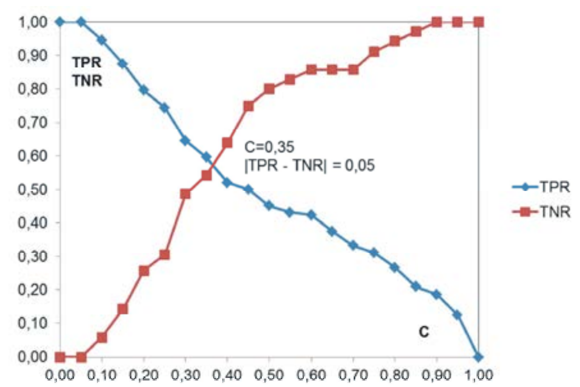

Fig. 2: The balance between sensitivity and specificity for the model Z_1

To build the first model

$$Z_1=-0,17Q_1-0,04Q_2+1,9Q_3+0,5Q_4+0,3Q_5+0,58Q_6+1,7Q_7+4,8Q_8+0,9Q_9+0,21Q_{10}-7,2$$

based on logistic regression the method of step-by-step inclusion with Wald test is used.

For the second model

$$Z_2= 1,79\ Q_3 + 1,53\ Q_7+4,9\ Q_8 - 6,89.$$

the same settings for logistic regression are used, but with the forced inclusion of all factors.

The models include the following factors: Z – default ("yes" or "no"), Q_1 – gender, Q_2 – age, Q_3 – marital status, Q_4 – record of service, Q_5 – type of employer, Q_6 – credit history, Q_7 – savings Q_8 - the ratio of income to expenses, Q_9 - income variation, Q_{10}–security for credit.

ROC Curve Analysis Results: On the basis of the formalized logistic regression models the main parameters and criteria (K_1-K_6) were calculated to conduct the ROC curve analysis. The results of calculations only for the first model Z_1 are presented in Table. 2. The calculated parameters allowed making the ROC curves for both models, presented in Fig. 1 and define the rational threshold value C.

Despite the various parameters and methods of logistic regression construction, the predictive accuracy of both models is the same because of the similar values of AUC, obtained by summing the figures in the corresponding row in Table 2. This fact can be explained by the sufficient correlation between factors. The curves are closer to the diagonal line of random guess that confirms the fact of the correspondence between both models and "bad" classifier. The rational threshold value was found using criteria K_3, K_5 and K_6 and equal to 0.60 for all criteria (see the underlined figures in Table 2).

The balance between sensitivity and specificity for the model Z_1 is achieved at the threshold value 0.35, as shown in Fig. 2.

CONCLUSION

Summing up, we can say that the ROC curve analysis can be applied to solve the following tasks in credit risk management:

- Estimation and comparison of the predictive accuracy, sensitivity and specificity of credit scoring models.
- Determination of the rational threshold values for credit scoring models.
- Parameters of credit scoring models assessed by the ROC curve analysis may be used as the indicators showing the need for adjusting the model (classifier). The lower sensitivity of the model, increase in the number of the false positive outcomes are some examples of such indicators.

Thus, the research shows the possibility of application of the ROC curve analysis in solving practical problems of credit risk and predictive capability estimation. The area for the further research can be, first of all, consideration of the ROC curve analysis in terms of the economic indicators, for example, the economic benefits and losses from the true and false classified credit applications. This fact takes into account the results of the bank's financial activity. Secondly, it is very important to discuss the influence of the adjustable model parameters on *AUC* that will provide sufficient grounds for recommendations how to configure classifiers with the best predictive capability. Finally, the priority task for the future research is to develop an approach of the ROC curve analysis application for the situation of more than two classes of the borrowers.

REFERENCES

1. Lukashevich, N.S., 2011. Comparison of the neural and statistical methods of credit risk assessment // Finance and credit, 1(433). C. 32-41.
2. Dubolazov, V.A. and N.S. Lukashevich, 2009. Fuzzy - multiple approach to the estimation of the individual creditworthiness // Finance and credit, 13(349): C. 35-45.
3. Drummond, C. and R. Holte, 2004. What ROC curves can't do (and cost curves can) // ROCAI, pp: 19-26.
4. Egan, J.P., 1975. Signal detection theory and ROC analysis. New York: Acad. Press, pp: 386.
5. Swets, J.A., 1988. Measuring the accuracy of diagnostic systems // Science, 240: 1285-1292.
6. Fawcett, T., 2004. ROC Graphs: Notes and Practical Considerations for Researchers. Kluwer Acad. Publ., pp: 38.
7. Hand, D.J. and R.J. Till, 2001. A simple generalization of the area under the ROC curve to multiple class classification problems // Machine Learning, 45(2): 171-186.
8. Youden, W.J., 1950. Index for rating diagnostic tests // Cancer, 3: 32-35.