

МРНТИ 20.53.21, 20.47.23
УДК 004:85;004.89;004.93

APPLICATION OF THE REGRESSION ANALYSIS METHOD FOR MODELLING THE PROCESSING OF LARGE AMOUNTS OF DATA

Toleugazy R.T.¹, Balakayeva G.T.¹

¹AL-Farabi Kazakh National University, Almaty, Kazakhstan

Abstract

In the modern world, there is often a time when the relevance of actions is not necessary, in addition, there is a need to predict further stages of activity.

All this is possible thanks to the use of regression analysis, which is used in many areas of activity. This article describes the application of this analysis to Bank risk, namely credit risk. Computational and theoretical studies on the processing of big data of banking institutions based on regression analysis, namely the method of multiple regression, have been carried out. In addition, this article provides a forecast based on various economic changes. This is very important for the banking sector, primarily because these studies reflect trends in credit risk.

Keywords: big data, modeling, Bank risk, multiple regression.

Аннотация

Р.Т. Төлеугазы¹, Г.Т. Балакаева¹

ПРИМЕНЕНИЕ РЕГРЕССИОННОГО АНАЛИЗА ДЛЯ МОДЕЛИРОВАНИЯ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ

¹Казахский Национальный университет им. аль-Фараби, г. Алматы, Казахстан

Современном мире часто возникает момент, когда необходима актуальность действий, кроме того возникает необходимость предсказывать дальнейшие этапы деятельности. Все это возможно и благодаря применению регрессионного анализа, который применяется во многих сферах деятельности. В этой статье описывается применение данного анализа в банковском риске, а именно в кредитном риске. Проведены расчетно-теоретические исследования по обработке больших данных банковских учреждений на основе регрессионного анализа, а именно методом множественной регрессии. Кроме того, в данной статье проведено прогнозирование с учетом различных экономических изменений. Для банковского сектора это очень важно, в первую очередь потому, что эти исследования отражают тенденции изменения кредитного риска.

Ключевые слова: большие данные, моделирование, банковский риск, множественная регрессия.

Аңдатпа

Р.Т. Төлеугазы¹, Г.Т. Балакаева¹

ҮЛКЕН КӨЛЕМДІ ДЕРЕКТЕРДІ ӨНДЕУДІ МОДЕЛЬДЕУ ҮШІН РЕГРЕССИЯЛЫҚ ТАЛДАУ ӘДІСТЕРІН ҚОЛДАНУ

¹аль-Фараби атындағы Қазақ Ұлттық Университеті, Алматы қ., Қазақстан

Қазіргі әлемде іс-әрекеттің өзектілігі, сонымен қатар қызметтің одан әрі кезеңдерін болжау қажеттілігі туындайды. Осының бәрі көптеген қызмет салаларында қолданылатын регрессиялық талдауды қолдану арқылы да мүмкін. Бұл мақалада осы талдаудың банк тәуекелінде, атап айтқанда несие тәуекелінде қолданылуы сипатталады. Регрессиялық талдау негізінде, атап айтқанда көптеген регрессия әдісімен банк мекемелерінің үлкен деректерін өңдеу бойынша есептік-теориялық зерттеулер жүргізілді. Бұдан басқа, осы бапта әртүрлі экономикалық өзгерістерді ескере отырып болжау жүргізілді. Банк секторы үшін бұл өте маңызды, себебі бұл зерттеулер кредиттік тәуекелдің өзгеру үрдісін көрсетеді.

Түйін сөздер: үлкен деректер, модельдеу, банк тәуекелі, көпше регрессия.

Introduction

The experience of developed countries underlines the need to develop portfolios of information on all bank customers and a database for information processing as a main way to strengthen the bank's position in its relations with customers, regardless of their size, and as a weapon the prevention and avoidance of credit risk [1]. Bank risk is a phenomenon that occurs during the activity of banking operations and which has a negative impact on this activity: deterioration of business activity or accounting for Bank losses that affect functionality. This maybe caused by an internal or external reasons caused by the competitive environment. Credit risk - non-fulfillment of financial obligations of legal entities (the "debtor") to the supplier of goods or the provider of these services, that is, to the Bank.

By using credit derivatives, banks keep the loan on their balance sheet. Transferring credit risks using derivatives means risks that credit risk transfers with loan sales or securitizations do not have. Banks using these derivatives have to bear associated counterparty, operational, and legal risks [2].

Second-level factors and determination the degree of influence of each factor on the Bank's credit risk level.

1. Macroeconomic

1.1. External Economic. The situation in the financial markets. As shown in recent years, this factor has a particularly strong effect on the change in the Bank's credit risk level. A sharp deterioration in the economic situation and a sharp collapse of markets can lead to a significant increase in the credit risk index. The stabilization of the situation in international markets is gradual and does not cause such strong jumps in the indicator, i.e. this factor usually has a negative impact on the credit risk indicator.

1.2. Domestic economic. Change in the rate of inflation. Inflation causes changes in the structure of banks' credit investments. When there is significant inflation, the terms of Deposit placement are reduced. As a result, the share of medium-and long-term loans decreases and the share of short-term loans increases [2]. This significantly worsens the solvency of enterprises and can subsequently affect the liquidity of banks (in the case of non-return or late repayment of these loans). As a result, credit risk increases significantly. An unexpected jump in inflation can result in a sharp increase in credit risk. A decrease in the rate of inflation has a positive effect on the country's economy and, in particular, on the level of credit risk of banks.

1.3. Change in the exchange rate of the national currency. The depreciation of the state's currency has a negative impact on the country's credit rating, and therefore on the credit rating of banks. A particularly strong increase in the level of credit risk occurs when the national currency is devalued. The stabilization of the national currency rate leads to a decrease in credit risk.

1.4. Regulations The change in reserve requirements. A moderate increase in this indicator causes a decrease in inflationary pressure, contributes to the stability of the national currency exchange rate and, accordingly, affects the Bank's credit risk. An increase in mandatory reserves usually leads to a decrease in credit risk, while a decrease leads to an increase.

1.5. Changes in the current tax system. A sharp increase in taxes can lead to a loss of creditworthiness of some borrowers, and therefore to an increase in credit risk. Lowering the tax rate or introducing tax incentives will improve the creditworthiness of businesses and individuals and reduce the Bank's credit risk.

1.6. State support. Increasing support for certain industries, if loans to these borrowers make up a significant percentage of the loan portfolio, leads to a reduction in credit risk. The state's refusal to Finance industries and the privatization of enterprises leads to a decrease in the reliability of borrowers and, if they make up a significant share in the loan portfolio, to an increase in credit risk.

1.7. Conditions of competition Number of competitors. Increasing the number of competitors in the banking services market and strengthening their position makes the Bank conduct a more risky credit policy, which increases the credit risk. Reducing the number of competitors, their bankruptcy or simply leaving the banking market allows the Bank to lend only to trusted reliable borrowers, which reduces the credit risk.

1.8. Progressive technologies in the credit market. The use of new advanced technologies by competitors in the field of credit risk assessment and management forces the Bank to make new decisions in the field of risk and introduce new management methods. Ultimately, this is expected to reduce the risk. The possible backwardness of the Bank from competitors in this area, the reluctance of management to change existing approaches, conservatism in relation to risk management-all this contributes to an increase in risk. So, in order to predict credit risk, it is necessary to determine the significance of each second-level factor. We suggest that you do this with the help of experts' opinions. To do this, we suggest that each expert (in our case, 10 people) choose 5 of the 15 factors that have the most significant impact on the Bank's credit risk index. Then the experts should rank these five factors by their degree of significance. In this case, the factor that has the greatest impact is assigned 5 points, and the lowest-1 point. Then the points for each factor are added up, and the significance of each factor is determined by dividing the points assigned to it by the total amount of points. Let's assume that the experts' estimates were distributed as follows (Table. 1).

From the results of data analysis the table shows that the most significant changes in credit risk may cause changes of macroeconomic factors (72 %), in particular the situation on financial markets, changing rates of inflation and number of competitors.

Among the internal factors, the most significant are the introduction of new methods and technologies, changes in the organizational structure and personnel structure. Then you need to determine how relevant factors will change over time. Changes in factors can be determined using one of the forecasting methods: statistical methods (taking into account trends), when information about expected changes is received (for example, from the media), or expert opinions.

Table 1. Determining the significance of the Bank's credit risk factors

Experts Factors	1	2	3	4	5	6	7	8	9	10	Total number of points	Significance of the factor
1	5	4	1	4	5	3	5	3	2	4	36	0,24 (=36/150)
2	4	1	5	3		5		5	3		26	0,17
3			4				3		4	3	14	0,09
4				2							2	0,01
5	1							1		1	3	0,02
6							1				1	0,01
7		5		5	3		4			5	22	0,15
8				1	4						5	0,03
<i>Total significance of external factors</i>												0,72
1	2		2			2		4	5		15	0,10
2					2	1					3	0,02
3		3				4					7	0,05
4	3	2									5	0,03
5					1		2				3	0,02
6			3					2		2	7	0,05
7									1		1	0,01
<i>Total significance of internal factors</i>												0,28

In our opinion, the expert's opinion is the highest priority in this case. Moreover, as an expert within the Bank, we offer to attract a specialist from the Bank's credit division. An experienced employee can predict changes in these indicators. The specialist of the credit division must determine (taking into account statistical data and incoming information) the impact of the change each factor affects the level of credit risk. Let's assume that the influence indicator can take five values:

+1-there are verified data on negative trends, the impact of this factor will cause a serious increase in the Bank's credit risk

+0,5 — preliminary data on negative trends, the impact of this factor will cause an increase in the credit risk of the Bank;

0 — changes are not expected, this factor will have no impact on the predictive value of credit risk;

-0,5 — have preliminary data about positive trends, the impact factor will reduce the credit risk of the Bank;

-1 — there are serious reasons to believe that expected positive changes, the influence of this factor will cause a serious decline in the credit risk of the Bank.

Having determined the significance of each factor and the impact of its changes on the level of credit risk, a specialist in the credit division can determine the overall impact of all factors on the Bank's forecast credit risk. Thus, the proposed approach to forecasting the Bank's credit risk can be presented in the form of a diagram (Figure 1).

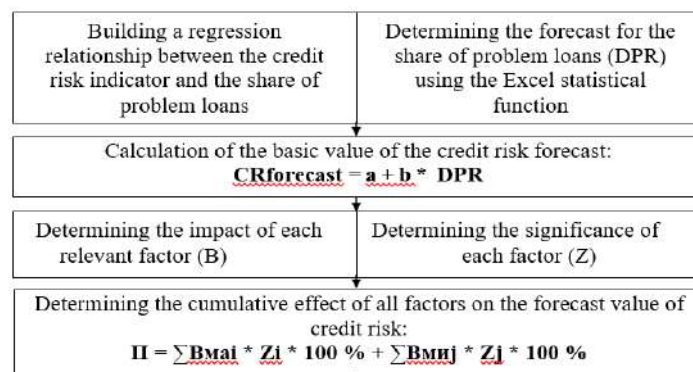


Figure 1. Scheme for determining the forecast value of the Bank's credit risk

Using regression analysis

Using the scoring, the lender can appreciate quickly, objectively and consistently the previous loans, and can calculate the probability that the loan will be repaid according to the contract [3]. Using regression analysis, you can determine the effect of several independent variables on one specific dependent variable. The general formula of the Regression looks like this []:

$$Y_t = a_0 + a_1 * X_{1t} + a_2 * X_{2t} + a_3 * X_{3t} + \dots + a_k * X_{kt} + \epsilon_t,$$

where: $t = 1, 2, \dots, n$ sample observation; Y_t - observation t of the dependent variable; X_j - independent variables, explanatory, $j = 1, 2, 3, \dots, k$; X_{jt} - observation t of independent variables X_j ; a_0 - constant, free term of equation; a_1, \dots, a_k - coefficients of independent variables; ϵ_t - error term of equation.

The coefficient of independent variable reflects how dependent variable Y_t changes when the independent variable, X_j , changes by one unit, while the other independent variables remain constant. If the dependent variable and independent variables are specified in natural logarithms, the coefficients of independent variables can be interpreted as elasticity's. Thus, these coefficients will show the percent change of the dependent variable if the independent variable changes by 1 percent.

To determine the parameters of linear regression (regression coefficients), use the least squares method. To do this, you need to solve the following system of equations:

$$\sum X_i Y_i = a(\sum X_i) + b(\sum X_i^2)$$

where X_i is the percentage of bad loans in the i -th period; y_i is the risk level of the Bank's loan portfolio in the i -th period; n is the number of observation periods, $i = 1..n$.

To determine the relationship between the credit risk indicator and the percentage of bad loans, we suggest using data on the risk of the loan portfolio (Table 2).

Table 2

№ n/n	Percentage of bad loans %	Credit Risk, %
1	17,0	25,41
2	1,7	3,86
3	0,0	1,34
4	0,0	1,03
5	0,0	0,91
6	0,0	2,03
7	4,3	5,56
8	12,5	13,46
9	1,2	3,56
10	0,0	1,26
11	0,0	0,94
12	0,0	0,85
13	0,0	1,09
14	6,3	7,06
15	0,1	1,11
16	5,4	6,46
17	0,5	1,34
18	0,2	2,19
19	0,1	1,28
20	26,6	27,38
21	2,0	2,83

We calculate the regression coefficients (in this case, the number of observations is 21), and we present the data for calculating the parameters of the regression equation in the (Table 3). As a result, the following values of the constant and variable regression coefficients are obtained: $a = 1.28$; $b = 1.08$.

Table 3. Calculation of derived data for regression analysis

№ n/n	X	Y	X*Y	X^2	Y^2	Y(x)
1	17	25,41	431,95	289,0	645,62	19,64
2	1,7	3,86	6,56	2,9	14,90	3,12
3	0	1,34	0,00	0,0	1,78	1,28
4	0	1,03	0,00	0,0	1,07	1,28
5	0	0,91	0,00	0,0	0,82	1,28
6	0	2,03	0,00	0,0	4,10	1,28
7	4,3	5,56	23,92	18,5	30,94	5,92
8	12,5	13,46	168,25	156,3	181,17	14,78
9	1,2	3,56	4,27	1,4	12,67	2,58
10	0	1,26	0,00	0,0	1,58	1,28
11	0	0,94	0,00	0,0	0,88	1,28
12	0	0,85	0,00	0,0	0,73	1,28
13	0	1,09	0,00	0,0	1,18	1,28
14	6,3	7,06	44,45	39,7	49,79	8,08
15	0,1	1,11	0,11	0,0	1,23	1,39
16	5,4	6,46	34,91	29,2	41,79	7,11
17	0,5	1,34	0,67	0,3	1,80	1,82
18	0,2	2,19	0,44	0,0	4,81	1,50
19	0,1	1,28	0,13	0,0	1,64	1,39
20	26,6	27,38	728,23	707,6	749,51	30,01
21	2	2,83	5,65	4,0	7,99	3,44
Sum	77,9	110,9	1449,6	1248,8	1756,0	

Therefore, the relationship equation describing the dependence of the risk level of the Bank's loan portfolio on the percentage of bad loans has the following form:

$$y = 1,28 + 1,08x$$

Graphically, we express the results as follows (Figure 2).

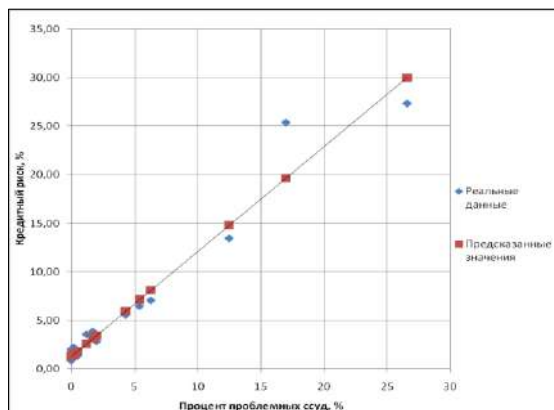


Figure 2. Dependence of the Bank's credit risk level percentage of bad loans in the loan portfolio

Conclusion

The resulting forecast reflects the trend of credit risk changes taking into account various economic changes. For the Bank, this indicator is very important, especially the forecast of changes in the credit risk indicator compared to its current value – its possible growth or decline. Based on these data, the Bank will adjust its risk management tactics, determine the riskiness of each new credit transaction, and decide whether to issue a loan or refuse to lend. Using the proposed approach to credit risk forecasting will allow the Bank to conduct a more balanced credit policy.

This approach takes into account various factors of the external environment, so its application will help make the credit institution more responsive to various changes, including crisis phenomena.

References:

- 1 Dimitriu M, Oprea IA (2009) Credit risk management in financial crisis, Review of International Comparative Management, 2:987-994.
- 2 Minton BA, Williamson RW (2005) How much do banks use credit derivatives to reduce risk?, National Bureau of Economic Research, Working Paper 11579, p. 12.
- 3 URL: <https://towardsdatascience.com/speech-emotion-recognition-with-convolution-neural-network-1e6bb7130-ce3>.
- 4 Balakayeva G.T., Phillips C., Darkenbayev D.K., Turdaliyev M. Using NoSQL for Processing Unstructured Big Data // News of the National Academy of Sciences of the Republic of Kazakhstan, Series of Geology and Technical Sciences. - 2019 - Volume 6, Number 438 – p.12-21.
- 5 Ключников М.В. Методы построения моделей прогноза основных показателей деятельности коммерческих банков // Финансы и кредит. 2004. № 3(141). С. 15–19.
- 6 Управление рисками и эффективностью бизнеса в условиях кризиса [Электронный ресурс]. Режим доступа: <http://www.iso.ru/rus>.
- 7 Финансовые риски – 2009. Работа над ошибками [Электронный ресурс]. Режим доступа: <http://risk.rcb.ru>.

МРНТИ 20.19.29
УДК 025.3/4:(084+086)

Б.Қ. Тульбасова¹, А.Н. Салыкова¹

¹ Абай атындағы Қазақ Ұлттық педагогикалық университеті, Алматы қ., Қазақстан

ЦИФРЛЫҚ БІЛІМ РЕСУРСТАРЫН ОРТА МЕКТЕПТЕ ҚОЛДАНУ ЕРЕКШЕЛІКТЕРІ

Аңдатпа

Мақалада цифрлық білім ресурстарын орта мектепте қолдану компоненттері анықталды, атап айтқанда, цифрлық оқытудың уәждемелік-мақсатты компоненті; цифрлық оқытудың мазмұнды компоненті; цифрлық оқытудың амалдық-іс-әрекеттік компоненті; цифрлық оқытудың бағалау-нәтижелі компоненті анықталды. Олардың мазмұны және оқу үрдісінде айқындалатын мәселелері анықталды.

Жаратылыстану-математика бағытындағы информатика пәнін оқыту мақсатында электронды оқыту жобасы аясында бірнеше республикаларда қолданыста бар цифрлы білім ресурстарына талдау жасалды. Microsoft Office, Macromedia Flash, Corel Draw, Adobe Photoshop, Ulead GifAnimator, Adobe Premiere серияларының бағдарламалық өнімдері, HTML-құжаттарын Macromedia DreamWeaver, Microsoft FrontPage сияқты редакциялау құралдары және MS Power Point, Prezi Desktop, Sway, SlideRocket, Google Docs, 280 Slides, Photodex ProShow Producer, ПромоШОУ т.б. цифрлық білім ресурстарын қолдану ерекшеліктері қарастырылды.

Түйін сөздер: цифрлық білім ресурсы, оқу үрдісі, MS Power Point, Prezi Desktop, Sway, SlideRocket, Google Docs, 280 Slides, Photodex ProShow Producer, ПромоШОУ, информатика пәні.

Аннотация

Б.К. Тульбасова¹, А.Н. Салыкова¹

¹ Казахский Национальный педагогический университет им. Абая, г. Алматы., Қазақстан

ОСОБЕННОСТИ ИСПОЛЬЗОВАНИЯ ЦИФРОВЫХ ОБРАЗОВАТЕЛЬНЫХ РЕСУРСОВ В СРЕДНЕЙ ШКОЛЕ

В статье определены компоненты использования цифровых образовательных ресурсов в общеобразовательной школе, в частности, мотивационно-целевой, содержательный и операционально-деятельностный и оценочно-продуктивный компоненты цифрового обучения. Определены их содержание и задачи, которые необходимо решать в учебном процессе.

Для преподавания информатики естественно-математического направления при электронном обучении был проведен анализ цифровых образовательных ресурсов, использующихся в нескольких республиках. Microsoft Office, Macromedia Flash, Corel Draw, Adobe Photoshop, Ulead GifAnimator, Ulead GifAnimator, Adobe Premiere серии программных продуктов, HTML-документов, Macromedia DreamWeaver, в Microsoft FrontPage и как инструменты редактирования MS Power Point, Prezi Desktop, Sway, SlideRocket, Google Docs, 280 Slides, Photodex ProShow Producer, ПромоШОУ и т.д. были рассмотрены особенности применения цифровых образовательных ресурсов.

Ключевые слова: цифровой образовательный ресурс, учебный процесс, MS Power Point, Prezi Desktop, Sway, SlideRocket, Google Docs, 280 Slides, Photodex ProShow Producer, ПромоШОУ, предмет информатики.