UDC 368.021:[519.23](477)

A MATHEMATICAL AND STATISTICAL ANALYSIS OF AN EXPANDED SAMPLE OF BUSINESS INDICATORS FOR UKRAINIAN INSURANCE COMPANIES

A. V. Lytvyn, *Doctoral student, assistant lecturer*

Department of Finance, National University of "Kyiv-Mohyla Academy",
2, Skovorody St., Kyiv, 04655, Ukraine
E-mail: anton.v.lytvyn@gmail.com

In the article, a mathematical-statistical analysis of Ukrainian general insurance companies' business indicators has been performed in order to promote the selection of proper methods of evaluating the financial health of Ukrainian insurers. In particular, descriptive statistics have been investigated, and data characteristics checked against the requirements of probabilistic statistical discriminative methods. Based on the analysis, a list of indicators and methods that can be reasonably used in evaluation of financial condition of Ukrainian insurers have been proposed.

Keywords*: general insurance, financial ratios, mathematical-statistical analysis, descriptive statistics, correlation analysis, multicolinearity, outlier exclusion, normality test.*

INTRODUCTION

As social-economic systems advance in their development, the number, the complexity and the interconnectedness of business operations increase, which may lead to the rise in exposure of economic entities to different kinds of risks. Insurance, as one of the means of dealing with such risks, is becoming increasingly important in everyday activities of both companies and citizen. Yet, insurance business bears some distinct features comparing to other types of business activity. The most important of them is the random nature of insurers' obligations, which means that insurance companies must laboriously manage their financial stability in order to be able to meet both known and any arising obligations. Hence, effective financial risk and crisis management in insurance companies should be among the main corporate governance priorities. On the other hand, the evaluation of insurers' financial health is also an important task for the state regulator responsible for supervising the activities and financial condition of insurance companies.

The typical procedure of analysing any firm's proneness to financial crisis involves the evaluation of its business indicators. However, there are several issues connected with the use of business indicators in financial analysis. Firstly, threshold values must be justified, which greatly depends on the statistical features of data. For example, if the industry data distribution is not symmetrical, the industry mean cannot be regarded as a reference point. Secondly, when analysing a set of indicators, it is implied that interrelations between them are known. However, these interconnections may be not obvious, which may jeopardise the validity of the analysis results. Thirdly, some statistical methods used in financial analysis have limitations and strict requirements to data.

Neglecting the abovementioned issues can compromise the overall objectivity of financial analysis process, as well as decrease the accuracy of the findings. Despite the fact that these issues draw attention of an increasing number of scholars, presently there is a gap in the research on the features of business indicators for Ukrainian insurance companies and their applicability with popular statistical and other financial analysis methods.

LITERATURE REVIEW

There has been significant research dedicated to analysing the activities and financial condition of Ukrainian general insurance companies. Mainly, such analyses are connected with the evaluation of financial health of insurers (Selivestrov, Shevchuk, Gasparian, Okhlovska, Bondarchuk and Kondrat, and others), the assessment of their financial stability and reliability (Shirinian, Biloshpytskiy; Malynych, and others), the prediction of financial distress (bankruptcy) (Matviichuk, Illichevskiy, and others), or studying the financial aspects of insurance company management (Suprun, Tarasova, Klepikova, and others). However, little attention has been given to the analysis of indicators themselves. In considerable number of studies statistical features of indicators are disregarded and method limitations are neglected.

In contrast, foreign research in the field of analysing business indicators includes the consideration of characteristics of data. Moreover, certain works have been targeted specifically at exploring the distributional properties of general microeconomic and financial data (Martikainen; Fuller-Love; Balcaen and Ooghe, and others) and analysing insurers' performance indicators (van der Heijden, Shi, and others). In addition, many studies in the field of insurer financial distress prediction include reasoning on the choice both the appropriate method and the suitable indicators (Ambrose and Seward, BarNiv and McDonald, Browne and Hoyt, and others). Yet, most of the findings are unlikely to be applied directly in Ukrainian practice due the major differences in economic and regulatory conditions of doing business.

The analysis of business indicators has always been in the centre of attention of researchers and practitioners. Financial ratios are widely used to make inferences on the financial health of companies. Usually, they are compared to some reference values to make a favourable or negative conclusion on a certain sphere of corporate activities. Next, the researcher unites these inferences to come up with a general financial diagnosis for a company. Yet, such an approach is prone to extreme subjectivity stemming from unjustified reference values, overlooking the interactions between indicators, and implicit synthesis of results.

In order to resolve these problems, numerous methods have been proposed, most of which are aimed at objectifying the process of analysing financial indicators and making conclusions based on such analysis. Beaver was among the first to conduct a profound analysis of financial ratios in their relation to corporate financial sustainability. He proposed a method according to which a single indicator had to be calculated based on reporting data, and depending on its value, conclusions on the possibility of bankruptcy could be made. The method allowed for grounded distinction between financially healthy and distressed companies. This is an example of univariate discriminant analysis [1, p. 34]. This approach allowed promoting the objectivity of financial analysis; however, the problem is that the results would vary depending on the choice of the variable.

The method that allowed researchers to alleviate the mentioned problem was the multiple discriminant analysis first applied to predict financial distress by Altman.

Altman developed so called Z-score model, which basically was a linear multivariate model. The author concluded that the model showed superior accuracy in comparison to univariate models; additionally, it was free of the shortcomings found in previous models. The advantages and relative simplicity of multiple discriminant analysis made it one of the mostly applied methods in different field of research, including financial analysis [2].

However, the benefits of multiple discriminant analysis come at a price: the method is rather demanding for data. Firstly, the observations must have multivariate normal distribution; secondly, the variance-covariance matrices must be equal. Violation of these conditions may lead to invalid modelling results, such as biased significance tests or errors, instability of parameter estimates, and low classification accuracy [1, p. 43].

The most successful "competitor" of multiple discriminant analysis in financial distress prediction is logistic analysis, which belongs to conditional probability models. Unlike, classic discriminant analysis, logistic regression does not require normally distributed independent variables and equal variance-covariance matrices. Still, logistic models are susceptible to multicolinearity. Taking into account that financial analysis almost inevitably includes an investigation of financial ratios, which may have identical numerators or denominators, the problem of multicolinearity becomes even more acute [2, p. 69]. Moreover, logistic analysis can be sensitive to extreme cases of non-normality.

There is another important issue connected with applying classic statistical models: the data that is used in the modelling process must be stationary. This implies that the interconnections between variables must remain the same in different periods. Yet, data is frequently unstable, which may lead to low accuracy on new (future) samples.

PROBLEM SPECIFICATION AND GOAL STATEMENT

Despite the abovementioned limitations, these methods remain widely applied in practice [1; 2]. Still, researchers often ignore the importance of analysing the statistical features of data, namely the distributional properties and inter-correlations; on the other hand, even if such analysis is carried out for specific industry data, samples are usually extremely limited and do not represent the general population of companies.

The goal of this study is to determine whether the popular business indicators appearing in financial research literature calculated for Ukrainian general insurance companies can be reasonably applied with the methods commonly utilised in the analysis of financial health. To achieve this goal, the basic characteristics of the selected indicators must be examined, their value allocation tested, and multicoliearity issue investigated.

DATA

The expanded dataset for this research has been formed from the 2010-2011 annual financial reports of Ukrainian general insurance companies using different sources, including the official web-site of Stock market infrastructure development agency of Ukraine, the of Public information database of the National securities and stock market commission of Ukraine, printed editions of the "Ukraine Business Review" journal, and corporate web-sites of insurance companies. Other data such as the number of insurers and their business status have been taken from the

Complex information system of the National commission for regulation of financial services markets of Ukraine. It should be noted that there are two problems connected with estimating the number of insurers that are assumed to have ceased activity (tagged as 'distressed' in table headings). Firstly, the number of distressed insurance companies in each year was determined by matching the list of insurers excluded from the register of financial institutions with the list of insurers that published financial reports in the respective year; hence, discrepancies may occur between the values presented in this research and the official insurance market overviews. Secondly, the absence of archived information regarding the exact date of excluding insurance companies from the register due to reorganisation of State commission for regulation of financial services markets of Ukraine into National commission for regulation of financial services markets of Ukraine did not allow us to form the exact list of companies that were excluded from the register in 2011. Thus, the distressed insurers in year 2011 are actually the companies that were excluded from the register in 2011 and later. The samples gathered for the study, as well as their representation of the general population of Ukrainian non-life insurance companies, are presented in Table 1.

Table 1 - Ukrainian general insurers' business status in 2010-2011

2010				2011			
Total	Sample			Total	Sample		
	Functioning	Ceased business activity	Ratio (sum to total)		Functioning	Ceased business activity	Ratio (sum to total)
389	308	31	87,15%	378	314	40	93,65%

FINDINGS

Business indicators for this study have been selected based on their appearance in the relevant literature sources. Works on predicting financial distress, insolvency, and bankruptcy in general insurance companies have been reviewed to choose indicators for analysis; additionally works on general bankruptcy prediction were studied. Table 2 contains the list of indicators (most of which are financial ratios) gathered from the sources [3–10]. It should also be noted that the list may not include the whole range of indicators proposed by the respective authors as the variables which could not be calculated based on the available data have been filtered out.

The range of selected business indicators should comprehensively characterise the financial condition of a general insurer as liquidity, profitability, leverage, asset and liability structure, turnover, and other ratios are included. Some ratios have different formulae according to different researchers. The choice of calculation algorithm had to be made in order to hold to the availability of data. For example, gross premiums written have been applied in place of other premium indicators, because other premium figures are not included in the obtainable financial reports. Likewise, loss adjustment expenses, underwriting expenses have been left out due to the same limitations. Other important note is that the values of some ratios could not be computed due to the division by zero; in this case the value is left blank and regarded as a missing value (in any case, the value close to infinity would have been regarded as an outlier for any ratio).

Table 2 - Selected business indicators of general insurance companies

Indicator	Formula
CR	(Incurred Losses + Underwriting Expenses)/Gross Premiums Written[1]
ER	Underwriting Expenses/Gross Premiums Written
LR	Incurred Losses/Gross Premiums Written
LiqR1	Cash and Other Liquid Assets/Current Liabilities
LiqR2	Liquid Assets/Insurance Reserves;
LiqR3	Cash/Current Liabilities
AStrR1	Cash and Other Liquid Assets/Total Assets
AStrR2	Current Assets/Total Assets
GrR1	$(\text{Liquid Assets}_t - \text{Liquid Assets}_{t-1})/\text{Liquid Assets}_{t-1}$
GrR2	$(\text{GPW}_t - \text{GPW}_{t-1})/\text{GPW}_{t-1}$
GrR3	$(\text{Surplus}_t - \text{Surplus}_{t-1})/\text{Surplus}_{t-1}$
GrR4	$(\text{CR}_t - \text{CR}_{t-1})/\text{CR}_{t-1}$
ExpR1	Investment Expense Incurred/Liquid Assets[2]
ExpR2	Operating Expenses/Operating Income
CFR1	Cash Flow from Operations/Gross Premiums Written
CFR2	Cash Flow from Operations/Total Assets
CFR3	Cash Flow from Operations/Total Liabilities
CFR4	Cash Flow from Operations/Sales
CFR5	Cash Flow from Operations/Current Liabilties
CFR6	Net Cash Flow/Total Liabilties
CStrR1	Retained Earnings/Total Assets
CStrR2	Total Liabilities/Total Assets
CStrR3	Current Liabilities/Total Assets
CStrR4	Long term Liabilities/Total Assets
CStrR5	Equity/Total Assets
PrR1	EBIT/Total Assets
PrR2	Net Operating Income/Gross Premiums Written
PrR3	Net Income/Sales
PrR4	Operating Income/Total Assets
PrR5	EBIT/Interest
PrR6	Investment Income/Average Invested Assets
ROA	Net Income/Total Assets
ROE	Net Income/Surplus
TOR1	Sales/Total Assets
TOR2	Sales/Inventory
TOR3	Gross Premium Written/Surplus
CCR1	Total Liabilities/Equity
CCR2	Insurance Reserves/Surplus
OR1	Current Assets/Sales
OR2	Working Capital/Sales
OR3	Quick Assets/Sales
OR4	Working Capital/Equity
OR5	Working Capital/Total Assets
Size	ln(Total Assets)

[1] Gross Premiums Written may further be abbreviated as GPW;
[2] The ratios in which the numerator presents a state indicator and the denominator presents a flow indicator (and vice versa) imply that the state indicator is taken at its mean value to ensure unit consistency

Having calculated the selected indicators, the basic statistical analysis of the sample has been performed. The descriptive statistics of the whole sample for 2011 are presented in Table 3.

Table 3 - Descriptive statistics for the entire sample

Indicator	Mean	Min	Max	Coef. Var.	Skewness	Kurtosis	Grubbs Test	p-value
CR	0.569	-25.9	64	732.76	10.728	170.069	15.215	0.000
ER	0.562	-18.3	64	719.41	12.329	189.008	15.699	0.000
LR	0.007	-7.6	1	6133.50	-15.152	259.886	17.124	0.000
LiqR1	138.310	0.0	10784	632.66	10.694	123.317	12.166	0.000
LiqR2	3228.951	0.0	676477	1296.39	14.485	219.963	16.083	0.000
LiqR3	40.085	0.0	5388	763.80	15.909	273.525	17.467	0.000
AStrR1	0.330	0.0	1	96.50	0.761	-0.740	2.102	1.000
AStrR2	0.671	0.0	209	1656.62	18.811	353.891	18.760	0.000
GrR1	9.035	-1.0	1285	915.68	12.825	179.704	15.428	0.000
GrR2	7.549	-324.8	2281	1687.17	17.382	312.071	17.846	0.000
GrR3	15.855	-1.0	2741	1157.25	13.163	179.535	14.850	0.000
GrR4	15.651	-169.4	2886	1181.80	14.986	231.914	15.517	0.000
ExpR1	46.362	-0.9	15575	1785.48	18.806	353.783	18.759	0.000
ExpR2	0.872	-4276.7	2399	31165.73	-8.991	194.541	15.737	0.000
CFR1	-0.503	-112.8	4	-1311.81	-15.527	256.278	17.005	0.000
CFR2	0.007	-0.7	2	1450.14	10.928	184.633	15.805	0.000
CFR3	1.052	-17.6	192	1152.23	13.102	190.507	15.768	0.000
CFR4	-0.503	-112.8	4	-1311.81	-15.527	256.278	17.005	0.000
CFR5	1.049	-17.6	192	1155.78	13.101	190.483	15.768	0.000
CFR6	-0.113	-48.8	25	-2757.27	-9.783	184.501	15.646	0.000
CStrR1	-2.138	-649.1	1	-1654.75	-17.623	320.001	18.286	0.000
CStrR2	2870.898	0.0	1016051	1881.03	18.815	354.000	18.762	0.000
CStrR3	0.671	0.0	209	1656.62	18.811	353.891	18.760	0.000
CStrR4	0.010	0.0	1	582.35	9.432	107.335	13.464	0.000
CStrR5	0.032	-268.6	1	44492.31	-18.809	353.839	18.760	0.000
PrR1	0.075	-2.0	2	293.50	-0.290	31.854	9.381	0.000
PrR2	-2.901	-387.0	221	-1143.30	-5.779	79.104	11.580	0.000
PrR3	31.391	-3089.0	11434	2084.27	15.653	279.799	17.427	0.000
PrR4	0.059	-0.6	1	258.06	1.990	7.867	5.427	0.000
PrR5	2205.258	-175.3	128469	616.34	9.026	84.336	9.290	0.000
PrR6	-131.916	-46723.6	12	-1882.52	-18.815	354.000	18.762	0.000
ROA	0.487	-1.4	19	441.15	6.195	41.458	8.776	0.000
ROE	0.563	-1.8	22	422.95	5.991	39.962	9.074	0.000
TOR1	0.252	-0.2	2	134.66	2.321	7.508	6.326	0.000
TOR2	3282.261	-1806.1	221843	525.77	9.624	107.848	12.665	0.000
TOR3	0.436	-0.2	5	159.70	2.728	9.717	6.770	0.000
CCR1	0.249	-0.8	18	462.18	12.667	186.067	15.772	0.000
CCR2	0.274	-0.2	3	165.46	2.965	10.330	5.490	0.000
OR1	776.538	-6800.8	71913	778.36	9.121	88.715	11.769	0.000
OR2	552.833	-6735.8	71902	904.61	11.214	140.506	14.267	0.000
OR3	768.931	-6800.8	71913	782.61	9.196	90.138	11.822	0.000
OR4	-0.170	-209.3	1	-6570.18	-18.786	353.285	18.752	0.000
OR5	0.575	-8.0	3	129.08	-4.167	50.811	11.566	0.000
Size	10.715	-4.6	15	14.68	-2.654	25.912	9.743	0.000

Due to the data omissions in ratios, the valid number of observations for different ratios varies (for example, total sample contains 354 observations, whereas the indicator Pr5 has only 92 valid cases). It may itself pose a problem if the data were to be used in bankruptcy prediction.

The analysis of variance coefficient shows that all variables are volatile. Moreover, the distribution of indicator values is skewed for all variables: the negative skew can be observed for LR, ExpR2, CFR1, CFR4, CFR6, CStrR1, CStrR5, PrR1, PrR2, PrR6, OR4, OR5, and Size; all other ratios are positively skewed. Such results denote that mean values are not representative for the sample and cannot be used as reference points during financial analysis. Likewise, the kurtosis of indicators is not consistent: only AStrR1 has a kurtosis close to zero; all other indicators have large kurtosis values, which indicates that the distributions of data must have pointed centres and long flat tails. Such descriptive statistics could point on the presence of outliers and non-normality of variable sample distribution. Indeed, the Grubbs' test indicates that only one variable (LqR1) does not include outliers.

As many statistical methods used in discriminating between financially healthy and distressed firms rely on mean values and dispersion to determine maximally distinct groups, the normality of data is necessary to ensure that mean and variance values are representative. For the further checking of the variables' distributions, Chen-Shapiro test has been applied due to its relative superiority when applied for relatively small to medium samples. The results of the application of the test are presented in Table 4.

Table 4 - Normality tests for the entire sample

Indicator	Statistic	10% critical value	5% critical value	Decision at level (5%)
1	2	3	4	5
CR	11.3462	-0.0100	0.0010	Reject normality
ER	11.5697	-0.0100	0.0010	Reject normality
LR	11.8858	-0.0100	0.0010	Reject normality
LiqR1	11.6631	-0.0101	0.0008	Reject normality
LiqR2	13.8600	-0.0099	0.0013	Reject normality
LiqR3	12.8309	-0.0101	0.0008	Reject normality
AStrR1	1.2400	-0.0102	0.0006	Reject normality
AStrR2	15.4254	-0.0102	0.0006	Reject normality
GrR1	13.0665	-0.0101	0.0007	Reject normality
GrR2	13.8851	-0.0100	0.0012	Reject normality
GrR3	13.4394	-0.0099	0.0012	Reject normality
GrR4	11.6897	-0.0091	0.0035	Reject normality
ExpR1	15.4897	-0.0102	0.0006	Reject normality
ExpR2	12.8004	-0.0101	0.0007	Reject normality
CFR1	13.5099	-0.0100	0.0010	Reject normality
CFR2	11.8913	-0.0102	0.0006	Reject normality
CFR3	13.0626	-0.0101	0.0006	Reject normality
CFR4	13.5099	-0.0100	0.0010	Reject normality
CFR5	13.0525	-0.0101	0.0006	Reject normality
CFR6	12.3687	-0.0101	0.0006	Reject normality
CStrR1	14.9406	-0.0102	0.0006	Reject normality
CStrR2	15.6001	-0.0102	0.0006	Reject normality
CStrR3	15.4254	-0.0102	0.0006	Reject normality
CStrR4	11.0548	-0.0102	0.0006	Reject normality
CStrR5	15.3435	-0.0102	0.0006	Reject normality
PrR1	3.6836	-0.0102	0.0006	Reject normality
PrR2	9.9054	-0.0100	0.0010	Reject normality
PrR3	13.5489	-0.0100	0.0010	Reject normality
PrR4	2.3356	-0.0102	0.0006	Reject normality
PrR5	5.8425	-0.0026	0.0176	Reject normality

Continuation of Table 4

1	2	3	4	5
PrR6	15.5980	-0.0102	0.0006	Reject normality
ROA	8.8653	-0.0102	0.0006	Reject normality
ROE	8.4728	-0.0102	0.0006	Reject normality
TOR1	2.4921	-0.0102	0.0006	Reject normality
TOR2	9.2941	-0.0092	0.0032	Reject normality
TOR3	3.4180	-0.0102	0.0006	Reject normality
CCR1	10.6922	-0.0101	0.0006	Reject normality
CCR2	3.7512	-0.0101	0.0006	Reject normality
OR1	11.7987	-0.0100	0.0010	Reject normality
OR2	12.2527	-0.0100	0.0010	Reject normality
OR3	11.8199	-0.0100	0.0010	Reject normality
OR4	15.0317	-0.0102	0.0006	Reject normality
OR5	2.7221	-0.0101	0.0006	Reject normality
Size	1.6878	-0.0102	0.0006	Reject normality

Apparently, none of the variables follows the normal distribution. This means that the data for 2011 violates one of the assumptions of classic discriminant analysis. Moreover, the results of most parametric statistical tests will be unreliable if applied to such data. The results also confirm that averaging should not be applied when analysing the selected financial indicators.

It was previously concluded that there must be outliers in most variables. Exclusion of extreme values can usually help to enhance the normality of data. However, in the studied case the application of percentile filtering (10% two-sided crop) to the data drastically reduces the number of valid observations (to the minimum number of 74; see Table 5). The cause of such a reduction is that different ratios have varying indexes of extreme observations, which, when juxtaposed, leave unaffected only a limited number of observations. Moreover, the test on the normality of distributions for the reduced sample did not show much improvement with all variables remaining non-normally distributed.

Table 5 - Indicators with considerable number of missing values and excluded outliers

Indicators	Missing values	Outliers	Valid observations
PrR5	262	18	74
GrR4	100	28	226
TOR2	92	52	210
CFR1	19	n/a*	335*
CFR4	19	n/a*	335*
CFR3	2	n/a*	352*
CFR5	2	n/a*	352*
CFR6	2	n/a*	352*
CFR2	0	n/a*	354*
* Outliers could not be determined due to a large number of zero values			

It should be also noted that in types of financial analysis such as business failure prediction using discriminating methods, the recoding of outliers is not an option because the financial condition particularities of each firm would be blurred. Additionally, despite the fact that the reduced combined sample is still relatively big, the outlier exclusion has critically limited the distressed insurer group, which impedes the application of classic statistical discrimination methods to the purged sample.

The next step is to analyse whether problem of multicolinearity is present in the studied sample. As no regression is built in the study, the variance inflation factor has not been used to spot multicoliearity; instead, correlation matrix analysis in performed to examine the linear correlations between the indicators.

The numbers of significant highly correlated variable pairs for both groups and the entire sample (prior and after outliers exclusion) are presented in Table 6.

Table 6 - Number of significant ($p \le 0.05$) correlations between indicators correlations $\ge |0.4|$ counted)

Sample/group	Number of correlated pairs
Entire sample	90
Non-distressed insurers	88
Distressed insurers	99
Entire sample (outliers excluded)	71

It can be concluded that considerable multicolinearity is present both within the groups of distressed and non-distressed insurance companies, and in the entire sample. Multicolinearity is more evident in the distressed group mainly due to its extremely small size. The results indicate that the studied sample can be used neither in discriminant analysis, nor in logistic analysis, which is less demanding to normality yet susceptible to multicolinearity.

In order to reduce the number of redundant indicators, such sequence of excluding variables is followed: first, the indicators with many missing values and outlier-exclusion problems (see Table 5) are eliminated; second, indicators which can be found in the largest number of correlated pairs are excluded; third, multiple ratios that duplicate the sphere of an insurer's activities are filtered out.

Having followed the described algorithm, a non-correlated subset of variables has been formed, which includes such indicators: CR, GrR1, GrR2, GrR3, PrR3, PrR6, ROA or ROE, TOR1, OR1, OR4, Size, ExpR2, AStrR2, LiqR3. The reduced sample, in which the indications of multicolinearity have been eliminated, can be applied in conditional probability modelling. Moreover, the analysis of such a set of indicators should still grant comprehensive insight on the different sides of an insurer's activities, financial performance and condition.

Similar analysis has been performed on the business indicators for the previous years (based on the available financials reports), and the results are analogous: the statistical features indicate that the value allocation is skewed, kurtosis is considerable; none of the indicators corresponds to normal distribution; the exclusion of numerous outliers does not yield improvement of statistical features; and multicolinearity is evident in the initial sample, but can be reduced via elimination of redundant variables.

CONCLUSIONS AND IMPLICATIONS FOR FURTHER RESEARCH

The conducted analysis has shown that the popular business indicators calculated for Ukrainian general insurers cannot be directly used in common financial analysis algorithms and most classic statistical models such as discriminant analysis and logistic analysis. Firstly, the values of selected indicators are asymmetrically distributed, have high positive kurtoses, showing biased mean and less informative standard deviation values. Moreover, most variables have numerous missing data points due to the typical division by zero operation in many financial ratios; some of the ratios have almost constant zero values. Additionally, the statistical tests show that outliers are present in most variables. The exclusion of extreme values with the aim of improving the characteristics of the variables has yielded discouraging results: the samples are drastically reduced, whereas resemblance of normality is still not achieved, which means that all indicators violate one of the main conditions of applying classic multiple discriminant analysis to the data. Next, the investigation of correlation matrices has shown that considerable multicolinearity can be observed in both non-distressed and distressed insurer groups and the entire sample (initial and with outliers excluded). The multicolinearity problem has been resolved by reducing the number of indicators in the sample, which enables the application of data in conditional probability models for predicting corporate financial distress. Yet, application of classic discriminant analysis as well as common ratio analysis, which are based on the assumptions of the representativeness of mean and variance, should give invalid results.

There are several possible ways to overcome the discovered limitations of the studied business indicators for Ukrainian general insurance companies. On the one hand, the researcher should seek robust and effective methods of filtering data in order to alleviate at least some of the problems. Secondly, the researcher might want to include other variables or further limit those which have been previously used with a view to improving the features of the data. Besides, the updated, less demanding classical statistical methods can be applied to the data. Alternatively, the researcher may decide to turn to non-parametric methods of discrimination between financially healthy and financially distressed insurance companies. Such methods, in particular, include neural networks, decision trees, support vector machines, rough sets, case-based reasoning, and other.

РЕЗЮМЕ

МАТЕМАТИКО-СТАТИСТИЧНИЙ АНАЛІЗ РОЗШИРЕНОЇ ВИБІРКИ ПОКАЗНИКІВ ДІЯЛЬНОСТІ УКРАЇНСЬКИХ СТРАХОВИХ КОМПАНІЙ

А. В. Литвин, *аспірант, асистент*

Національний університет «Києво-Могилянська академія»,
вул. Г. Сковороди, 2, м. Київ, 04655, Україна
E-mail: anton.v.lytvyn@gmail.com

У статті проведено математико-статистичний аналіз бізнес-індикаторів українських компаній з ризикового страхування з метою сприяння вибору релевантних методів оцінки фінансового здоров'я українських страховиків. Зокрема досліджено описову статистику обраних показників, перевірено відповідність характеристик даних основним припущенням ймовірнісних статистичних дискримінаційних методів. На основі проведеного аналізу запропоновано перелік змінних та методів, що можуть бути обґрунтовано використані в оцінці фінансового стану українських страховиків.

Ключові слова*: ризикове страхування, відносні фінансові показники, математико-статистичний аналіз, описова статистика, кореляційний аналіз, мультиколінеарність, елімінація викидів, тест на нормальність розподілу.*

РЕЗЮМЕ

МАТЕМАТИКО-СТАТИСТИЧЕСКИЙ АНАЛИЗ РАСШИРЕННОЙ ВЫБОРКИ ПОКАЗАТЕЛЕЙ ДЕЯТЕЛЬНОСТИ УКРАИНСКИХ СТРАХОВЫХ КОМПАНИЙ

А. В. Литвин, *аспирант, ассистент*
Национальный университет «Киево-Могилянская академия»,
ул. Г. Сковороды, 2, г. Киев, 04655, Украина
E-mail: anton.v.lytvyn@gmail.com

В статье проведен математико-статистический анализ бизнес-индикаторов украинских компаний по рисковому страхованию с целью способствования выбору релевантных методов оценки финансового здоровья украинских страховщиков. В частности исследована описательная статистика выбранных показателей, проверено соответствие характеристик данных основным предположениям вероятностных статистических дискриминационных методов. На основании проведенного анализа предложен перечень переменных и методов, которые могут быть обоснованно использованы в оценке финансового состояния украинских страховщиков.

Ключевые слова*: рисковое страхование, относительные финансовые показатели, математико-статистический анализ, описательная статистика, корреляционный анализ, мультиколлинеарность, тест на нормальность распределения.*

REFERENCES

1. Sun J. Predicting financial distress and corporate failure: a review from the state-of-the-art definitions, modeling, sampling, and featuring approaches / J. Sun, H. Li, Q.-H. Huang, K.-Y. He // Knowledge-Based Systems. – 2014. – № 57. – P. 41-56.
2. Balcaen S. 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems / S. Balcaen, H. Ooghe // The British Accounting Review. – 2006. – № 38. – P. 63-93.
3. Van der Heijden H. Distributional properties of some financial ratios in insurance // Paper presented at the British Accounting and Finance Association (BAFA) Annual Meeting. – Birmingham, UK, 2011. – P. 1-13.
4. Lee S. Analysis and prediction of insolvency in the property-liability insurance industry: a comparison of logit and hazard models / S. Lee, J. Urrutia // The Journal of Risk and Insurance. – 2006. – № 63(1). – P. 121-130.
5. Cummins J. Regulatory solvency prediction in property-liability Insurance: risk-based capital, audit ratios, and cash-flow simulation / J. Cummins, M. Grace, R. Phillips // The Journal of Risk and Insurance. – 1999. – № 66(3). – P. 417-458.
6. Brockett P. A comparison of neural network, statistical methods, and variable choice for life insurers' financial distress prediction / P. Brockett, L. Golden, J. Jang, Ch. Yang // The Journal of Risk and Insurance. – 2006. – № 73(3). – P. 397-419.
7. Sharpe I. Financial distress in Australian general insurers / I. Sharpe, A. Stadnik // The Journal of Risk and Insurance. – 2007. – № 74(2). – P. 377-399.
8. Insolvency prediction model using artificial neural network for Malaysian general insurers // N. Chiet, S. Jaaman, N. Ismail, S. Shamsuddin / Proceedings of IEEE World congress on Nature and Biologically Inspired Computing (NaBIC). – Coimbatore, India, 2009. – P. 584-589.
9. Kleffner A. An examination of property & casualty insurer solvency in Canada / A. Kleffner, R. Lee // Journal of Insurance Issues. – 2009. – № 32(1). – P. 52-77.
10. Bellovary J. A review of bankruptcy prediction studies: 1930-present / J. Bellovary, D. Giacomino, M. Akers // Journal of Financial Education. – 2007. – № 33. – P. 1-42.

Надійшла до редакції 12 березня 2015 р.