CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Procedia - Social and Behavioral Sciences 110 (2014) 1148 – 1157
Procedia
Social and Behavioral Sciences

Contemporary Issues in Business, Management and Education 2013

Discriminant analysis as a tool for forecasting company´s financial health

Katarína Kočišová[a]*, Mária Mišanková[a]

[a]*University of Žilina, Faculty of Operation and Economics of Transport and Communications, Department of Economics, Univerzitná 1, 010 26 Žilina, Slovakia*

Abstract

Methods that we can use to find out impending problems in the company are prediction models, i.e. models that are able to assess the financial health of the company and to warn owners of the company and company´s business partners before the threat of bankruptcy. Financial analysis can be ex-post and ex-ante. In the literature we can find several names used for the ex-ante financial analysis, such as financial health prediction, forecast financial health, early warning systems, etc. In the article is used the term *prediction of financial health*. There are many methods on which are based the construction of prediction models and in the article authors further specified *discriminant analysis*.

© 2014 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
Selection and peer-review under responsibility of the Contemporary Issues in Business, Management and Education conference.

Keywords: Discriminant analysis; Financial analysis; canonical discriminant analysis; financial health; prediction

1. Historical development of prediction models

Prediction of company's financial distress and bankruptcy is subject of great interest of financial researchers since the end of the 60th. Beaver (1966) is considered as a pioneer who firstly used a dichotomous classification t-test in one-dimensional framework and he also laid the foundations of prediction models. He used financial ratios of 79 bankrupt and non bankrupt companies operating in the same industry and which had the same amount of assets. He identified a simple ratio – Cash flow / Total debts as the best predictor of bankruptcy. The advantage of a one-dimensional model is mainly its simplicity, which does not require knowledge of statistics.

* Corresponding author. Tel.: +421-41-513-3227.
E-mail address: katarina.kocisova@fpedas.uniza.sk.

1877-0428 © 2014 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
Selection and peer-review under responsibility of the Contemporary Issues in Business, Management and Education conference.
doi:10.1016/j.sbspro.2013.12.961

Beaver´s studies followed Altman (1968) who proposed a multi-dimensional technique known as multivariate discriminant analysis (MDA). This was applied on a sample of 33 bankrupt companies and 33 non bankrupt in the period of years 1946 – 1964. He chose five variables which were the most relevant in predicting bankruptcy of company, namely: Working capital / Total assets, Retained earnings / Total assets, Market value of equity / Book value of total debt, Sales / Total assets. Eisenbeis (1977), Ohlson (1980) and Jones (1987) were critics of MDA and they argued that the results could be biased and did not have sufficient information value.

Ohlson (1980) applied the Logit model which was based on the same assumptions as MDA and became popular for prediction of financial problems of companies. Ohlson used for construction the data of 105 bankrupt companies and 2058 non bankrupt companies for the period of years 1970 – 1976. The results showed that the size, financial structure (total liabilities to total assets), performance and liquidity were important determinants of bankruptcy.

Zmijewski´s Probit model was applied firstly in 1984 for prediction of bankruptcy of the company. The use of Logit models had proved very popular during its history. On the other side Probit models are used for prediction of bankruptcy less than Logit models probably because the use of them requires more calculations.

Econometric problems related to the Logit and Probit models were also discussed by Hillegest in 2004. Shumway (2001) demonstrated that problems of these models may result in bias, inefficiency and inconsistency of the estimated coefficient. He suggested for bankruptcy prediction to use the Hazard model to overcome econometric problems.

At present are also applied for prediction of bankruptcy various types of heuristic algorithms (neural networks and decision trees) and have recorded improvements in prediction of financial distress. Representatives of these studies were Tam & Kiang (1992), Salchenberger *et al.* (1992), Jain & Nag (1988) who provided evidence that neural networks overcome conventional statistical models such as discriminant analysis, Logit models in financial applications including classification and prediction.

2. Essence of discriminant analysis

Discriminant analysis is the method which allows company to decide whether an element belongs or does not belong to the advance set group which is not always simple and clear. The term discriminant analysis first appeared in 1936 in works of R. A. Fischer in the article *The Use of Multiple Measurement in Taxonomic Problems*. The nature of discriminant analysis defined by Fisher simply consists of the methods to explore the relationship between a group of independent characters (which we call discriminators) and one qualitative dependent variable – output. The output is in the simplest case a binary variable y that acquires only two values (Cisko & Klieštik, 2009):

- 0 in case that the object is in the first class,
- 1 in case that the object is in the second class.

The classes are known to be clearly distinguishable and each object clearly belongs to one of them. The task can also be identification of features that contribute to the identification process. The purpose is to find a prediction model classifying new objects (for example companies) into classes. New objects are classified into classes based on their high degree of similarity (Cisko & Klieštik, 2009).

The essence of discriminant analysis simply consists of examination of the dependence of one qualitative (classification) variable from several quantitative variables (the specific case of canonical correlation analysis). According to the numbers of variations of qualitative variable we can distinction (Stankovičová & Vojtková, 2007)

- discriminant analysis for two groups,
- discriminant analysis for more groups.

Economics is not the only area in which may be applied discriminant analysis. It is possible to use this mathematical – statistical method in biology, medicine, archeology or in technical fields, etc. Hebák *et al.* (2004) states several examples:

- while checking the quality and reliability is possible in the sample to measure several quantitative variables (size, weight, chemical composition, etc.), then products subjected to stress and watch that they endure or not (which may mean its depreciation). To predict behavior of other products under stress we realize only necessary measurement of quantitative variables,
- applicants for some profession or studies are subjected to various tests, their scores represent observed quantitative variables and then their success or failure is found out in chosen field. Assuming that the test results relate to success of individuals in the sample we can expected future success of other candidates from their results,
- in the sample of patients are for several diagnosed diseases detected results of the various laboratory tests. If there is relationship the doctor may decide about new patient on the basis of the results of test for a specific diagnosis and treatment method.
- the Bank monitors in the sample its client´s way of repaying their loans and some other indicators. Subsequently on this basis bank may evaluated potential clients as more or less credible for a loan.

The main task of discriminant analysis is to find the optimal attributing rules that will minimize the likelihood of erroneous classification of elements, i. e. it will minimize the median of erroneous decision (it may happen that the element actually comes from a particular group, we classified into different group by obtained discriminant analysis). Each element is characterized by several features which reflect its properties. This means that the examined elements (units) are realizations of the random vector $X = (X_1, X_2, ..., Xn)$. Random variables X_i, where i = 1, 2, ..., m, corresponding to measured characteristics. The procedure starts with an analysis of group of elements in which is known relation to a particular group and also values of the random variables. (measured characteristics) – training set. The result of the analysis of the training set is to find discriminant function that determine the likelihood of classification of new still unclassified element to particular group on the basis of measured values (x_1, x_2, ..., x_m) of its characteristics.

Stankovičová & Vojtková (2007) distinguish two basic goals of discriminant analysis:

- ***Descriptive or analytical*** – tries by using Xj (j = 1, ..., k) to find appropriate statistical way to distinguish between groups, it looks for a function that identifies the existence of statistically significant differences between averages of pre-defined groups.
- ***Classification*** – tries to include new statistical unit (object) that is characterized by a vector of k characters to one of the established groups.

3. Goals and assumptions of discriminant analysis

Meloun, Militký, & Hill (2005) state that goals clearly clarify the nature of the discriminant analysis:

a) Determines whether there are statistically significant differences between the profiles of the average score of discriminators for two or more pre-defined classes.
b) Determines which of the discriminator is reflected the most in differential profiles of average score of two or more classes.
c) Determines procedures to include objects (individuals, companies, products, etc.) into classes based on their score in the set of discriminators.
d) Determines the number and composition of the dimensions of discrimination between classes formed by a set of discriminators.

Stankovičová & Vojtková (2007) indicate that the application of discriminant analysis requires fulfillment of following assumptions:

a) Multivariate normal distribution

To confirm this assumption is necessary to use tests of significance of individual discriminatory variables and discriminatory functions. If the data do not come from a multi-dimensional normal distribution, these tests are not valid and also results of classification are incorrect. However, the total error of classification is not violated by non-performance of that assumption because the error of classification in one group may be overvalued and in the other group undervalued (Sharma, 1996).

The data must be asymptotically multi-dimensional normally distributed in the groups in the case that the aim of discriminant analysis is to derive a linear discriminant analysis or quadratic discriminant analysis. If this assumption is not fulfilled there are also non parametric discriminant analysis methods or also can be used logistic regression analysis which does not demand fulfillment of this assumption.

b) Strictly defined groups of statistical units (objects)

These groups can be defined in the following ways:

- by objective factors (breakdown by SK NACE, division places by districts, countries, etc.),
- by expert assessment (based on the assessment by experts in the field),
- by statistical methods (breakdown of industrial companies in terms of efficiency of their management through multicriteria evaluation methods and cluster analysis).

c) The significance of selected discriminatory variables

Before derivation of discriminant functions should be tested:

- the conformity of variances, it means intra – covariance matrix – in order to derive the Fisher discriminant function it is necessary to be the covariance matrix by groups the same. In the case of the intra-covariance matrices are not the same is suitable to use derive quadratic discriminant function for classification. The degree in which a violation of assumption affects significance of tests and estimation of classification error depends on the number of discriminatory variables and groups sizes. It is recommended about twenty objects for each variable, although this ratio is difficult to adhere. It is necessary to realized that if this ratio is too small the results of discriminant analysis are unstable (minimum size of sample should be five objects on one discriminant variable). And there is also the rule about the size of the group: the smallest group must be greater than the number of discriminatory variables (Meloun & Militký, 2004).
- the conformity of medians in each groups – selected discriminant variables vary by extent in which are differences between the analyzed groups. Between the averages of the variables should be statistically significant differences and therefore on the basis of tests the variables can be excluded from the analysis, if these variables do not distinguish existing groups.

4. Canonical discriminant analysis

Canonical discriminant analysis belongs to the descriptive task of discriminant analysis. The objective of descriptive role of discriminant analysis consists in the derivation of the discriminant function or functions that statistically significant separates two or more groups. It means that the discriminant function can be obtained by two alternative ways of calculating:

4.1. Canonical discriminant analysis

Canonical discriminant analysis by Hebák *et al.* (2004) allows following relationship between objects in space between canonical axes, i. e. axes through which is defined canonical space. The goal is to find such a linear combination of p observed variables $x_1, x_2, x_3, \ldots, x_p$:

$$Y = b^T x \tag{1}$$

Where b^T is equal to $b_1, b_2, \ldots, b_p$ and it is such a vector of parameters which helps to separated considered groups in the way that the intra – group variability will be the smallest and inter – group variability will be the largest. This vector of parameters is better than any other linear combination.

If the total variability of the original observed variables is expressed by matrix:

$$T = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \left(x_{ih} - \overline{x}\right)\left(x_{ih} - \overline{x}\right)^T \tag{2}$$

where:

- x_{ih} = vector of values of i statistical unit in h group,
- $\overline{x}$ = average of the whole group.

And if it is spread to the sum of the matrix E expressing intra – group variability,

$$E = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \left(x_{ih} - \overline{x}_h\right)\left(x_{ih} - \overline{x}_h\right)^T \tag{3}$$

where:

$\overline{x}_h$ = vector of averages in h group.

And matrix B expressing the inter – group variability,

$$B = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \left(\overline{x}_h - \overline{x}\right)\left(\overline{x}_h - \overline{x}\right)^T = \sum_{h=1}^{H} n_h \left(\overline{x}_h - \overline{x}\right)\left(\overline{x}_h - \overline{x}\right)^T \tag{4}$$

where:

n_h = number of statistical units in h group.

It means E + B = T is a sum of squares $Q_B(Y)$ a $Q_E(Y)$, which represent the extent of intra and inter – group variability for new variable Y, which can be written as:

$$Q_B\left(Y\right) = b^{\mathbf{T}} Bb \text{ and } Q_E\left(Y\right) = b^{\mathbf{T}} Eb \tag{5}$$

The largest inter – group and the smallest intra – group variability of variable Y can be achieved at the maximum ratio:

$$F = \frac{Q_B\left(Y\right)}{Q_E\left(Y\right)} = \frac{b^T Bb}{b^T Eb} \tag{6}$$

Which is known as the Fisher´s discriminatory criterion. To determine the value Y = $b^T x$, which would determine differences between groups, it is necessary to specify the elements of the vector b to maximize the discriminatory criterion (6). Maximizing task can be solved as:

$$\left(B - \lambda E\right)b = 0 \tag{7}$$

respectively

$$\left(BE^{-1} - \lambda I\right)b = 0 \text{ at } \left(E\right(\neq 0 \tag{8}$$

The characteristic equation $\left(BE^{-1} - \lambda I\right) = 0$ has r solutions, which are characteristic numbers $\lambda_1, \lambda_2, ..., \lambda_r$ of BE^{-1} matrix. To the largest of these characteristic numbers λ_1 corresponds the characteristic vector b_1, which maximize the discriminatory criterion F. The characteristic equation does not determine the vector b_1 clearly, but only sets the ratio between its elements. It means coefficients for searched linear combination should be chosen as:

$$\frac{1}{n-H} b_1^T Eb = 1 \tag{9}$$

thus by dividing the vector $\mathbf{b}_1$ by expression:

$$\sqrt{\frac{b_1^T Eb}{n-H}} \tag{10}$$

Then intra – group variability of variable $Y = b_1^T x$ expresses unitary dispersion and criterion F is possible to write as:

$$F = \frac{1}{n-H} b_1^T Bb \tag{11}$$

And the characteristic number λ_1 expressing inter – group variability rate of variable Y_1.

If the file of units described by *p* variables is sorted into two groups, for expressing variability of original variables is sufficient only one discriminant (Hebák *et al.*, 2004). In the case of sorting into more than two groups is possible through one discriminant express only a part of the variability of the original variables. Using others of characteristic numbers $\lambda_2, \lambda_3, ..., \lambda_r$ and therefore characteristic vectors b_2, b_3, ..., b_r we obtained another canonical variables $Y_j = b_j^T x$, j = 2, 3, ..., r. These discriminates are mutually independent and their maximum number is given by r = min (p, H – 1).

The elements of the vector $b_j, b_j^T = \left[b_{j1}, b_{j2}, ..., b_{jp}\right]$ are coefficients of *j* canonical variable. If we substitute into the canonical variable ($Y_j = b_j^T x$) for each unit obtained values of variables X_1, X_2, …, X_p, we obtain the discriminatory score. In case we use constant cj when determining the score.

$$c_j = -b_j^T \overline{x} \tag{12}$$

Then the discriminatory score of discriminant is zero. For *i* unit, i = 1, 2, …, n_h, in *h* group (h = 1, 2, …, H), we determine *j* discrimination score, j = 1, 2, …, r, as:

$$y_{ijh} = c_j + \sum_{k=1}^{p} b_{jk} x_{ihk} \tag{13}$$

To make an idea of how in terms of the *j* canonical variable groups differ from each other it is better to use the calculation of vectors of average values of discriminant in groups (group centroids).

$$\overline{y}_{jh} = c_j + \sum_{k=1}^{p} b_{jk} \overline{x}_{ik} \tag{14}$$

If we mark diagonal matrix with the letter F with the square roots of diagonal elements of the matrix E, then we get standardized coefficients:

$$b^* = \frac{1}{\sqrt{n-H}} Fb \tag{15}$$

An alternative might be interpretation by using correlation coefficients between the canonical variables and the original variables, so-called structural variables. Their high positive or negative value says that the monitored variable is characteristic for given discriminant. According to sign (+ / –) we can determine whether the values of the original variable lead to an increase or decrease of value of discriminatory score. Vector of these correlations coefficient for *j* discriminant can be determined as:

$$a_j = \frac{1}{\sqrt{n-H}} F^{-1} Eb \tag{16}$$

Discriminatory score can be used to classify *n* objects into *H* groups. It means that the object will be included into the group which is the closest in terms of distance of units from the centroid of the group. This can be expressed by Mahalanobis and also Euclidean (because the canonical variables are uncorrelated) distance.

$$d_{ih}^2 = \sum_{j=1}^{s} \left(y_{ij} - \bar{y}_{jh} \right)^2 = \sum_{j=1}^{s} \left[b_j^T \left(x_i - \bar{x}_h \right) \right]^2 \tag{17}$$

4.2. Stepwise discriminant analysis

In stepwise discriminant analysis into the analysis enter a larger number of variables and with a sequence of steps are selected variables which discriminate the best and from them is created discriminant function. We can distinguish by the criteria how the stepwise discriminant analysis looks at selection of these variables (Stankovičová & Vojtková, 2004):

- ***Forward selection*** – Variables enter into the discriminant function gradually and always is chosen the one that has the greatest benefit in terms of discrimination. If this marginal benefit is not statistically significant, no new variable enter into the function.
- ***Backward selection*** – Into the discriminant function enter all variables and progressively are discarded those whose removal does not case a statistically significant decrease rate of discrimination. This process is completed only when any other discard would mean significant decrease in discrimination between groups.
- ***Stepwise selection*** – This selection is combination of the two previous. It means that into discriminant function enter new variables gradually and it is always chosen one with the greatest benefit in terms of discrimination, while in each step is verified the possibility whether the variable would be excluded and if excluded variable does not have significant impact on decrease rate of discrimination.

These methods achieve same results but condition is that the input data have to be mutually uncorrelated. Otherwise if between input variables exist statistically significant correlation, it is suitable to select *Stepwise selection*, where initially selected variable may be excluded in further steps because it is only correlation of other variables in the model.

Criteria for making decision about enter of variable into the model or its removal from the model serve following statistics (Meloun, Militký, & Hill, 2005):

Wilks λ

Wilks λ statistic expresses the ratio of intra – group variability to the total variability. At each step is selected the variable that achieves the minimum value of this statistic. The significance of changes of Wilks criteria λ after the introduction of discriminators into the model or removal from the model is based on test criterion F. It means that statistical significance of each discriminator is calculated by F-test. Value of F for change of Wilks criteria λ while adding discriminator into the model so that the model contains p discriminators is calculated according to the formula:

$$F_{zmeny} = \frac{n-g-p}{g-1}\left(\frac{\frac{1-\lambda_{p+1}}{\lambda_p}}{\frac{\lambda_{p+1}}{\lambda_p}}\right) \tag{18}$$

where:

- p = number of discriminators in the model,
- n = total number of objects,
- g = number of classes,
- λ_p = Wilks criterion λ before adding discriminators,
- λ_{p+1} = Wilks criterion λ after adding discriminators into the model.

Thus into the model is in each step included the discriminator which causes the least value of Wilks λ criteria.

Rao V

It is based on the Mahalanobis distance. It means that is selected variable that maximizes the distance between the two closest groups. It is also known as *Lawley-Hotelling track* and is defined as:

$$V = (n-g)\sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij} \sum_{k=1}^{g} (\bar{x}_{ik} - \bar{x}_i)(\bar{x}_{jk} - \bar{x}_j) \tag{19}$$

where:

- n = number of discriminators in the model,
- g = number of classes,
- $\bar{x}_{ik}$ = average of i discriminator in the k class,
- $\bar{x}_i$ = average of i discriminator for all classes combined together,
- w_{ij} = the element of the inverse covariance matrix between classes.

The greater are differences between medians (usually averages) of classes, the greater is the value of Rao V.

Mahalanobis distance

This is a generalized measure of distance between two classes 1 and 2 and it is expressed as:

$$D_{1,2}^2 = (n-g)\sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij} (\bar{x}_{i1} - \bar{x}_{i2})(\bar{x}_{j1} - \bar{x}_{j2}) \tag{20}$$

where:

- m = number of discriminators in the model,
- $\bar{x}_{i1}$ = average of i discriminator in class 1,
- w_{ij} = the element of the inverse covariance matrix.

Cause the Mahalanobis distance is criterion for selection of discriminators is calculated at first and it is criterion of all pairs of classes. The discriminator which had the greatest value for the two classes, it means which had the smallest value, is included into the model.

Inter – group F – values

When calculating the Mahalanobis distance all groups are equally important. This deficiency can be overcome by using inter – group F – values. When calculating these values larger groups have greater weight than smaller ones.

There is no exact rule for selection of a specific rule. The analyst chooses the best method after assessment of individual results. The most commonly used is Wilks λ criterion.

If a large number of independent variables enter into the discriminant function and can´t be determined the size of their distinctive ability, it is appropriate to use a stepwise discriminant analysis.

5. Conclusion

Discriminant analysis is a broad term that includes a several similar statistical techniques used for testing hypotheses. These methods allow studying differences between two or more groups of objects characterized by several features. We can divided these techniques into those used for interpretation of the differences between pre-defined groups of objects and those used for classification of objects into groups. Procedures used for interpretation of differences (canonical discriminant analysis) allow answering questions:

- whether and to what extent it is possible to distinguish set groups of objects based on the available characters,
- which of the features contribute to that distinguish in greatest extent.

The second group of techniques (classification discriminant analysis) allows deriving one or more equations for identification of objects.

The goal of canonical discriminant function is the maximum separation of defined groups while the classification functions minimize the number of „wrong“ set cases (the minimum value of „error rate“). Discriminant analysis is used for classification of objects into groups. The use of discriminant analysis may be to determine variables that have the highest ability to distinguish the groups to which the object belongs. Verbal formulation of the role of discriminant analysis is relatively simple and practical richness of usage is indisputable. On the other hand, mathematical and statistical apparatus of solution of the task is for common users rather complex. We can divide discriminant analysis into univariate and multiple discriminant analysis (MDA). Their use has advantages and also disadvantages.

An important ***advantage*** of the univariate failure prediction model is its simplicity. The application of a univariate model does not require any statistical knowledge: for each ratio, one simply compares the ratio value for the firm with a cut – off point and decides on the classification accordingly. Although the simplicity of the univariate model is appealing, this method also shows some important ***disadvantages***. Firstly, firm classification can only occur for one ratio at a time, which may give inconsistent and confusing classifications results for different ratios on the same firm (Altman, 1968). This problem is called the *"inconsistency problem"*. Secondly, when using financial accounting ratios in a univariate model, it is difficult to assess the importance of any of the ratios in isolation, because most variables are highly correlated (Cybinski, 1998). In the same context, the univariate model contradicts with reality in that the financial status of a company is a complex, multidimensional concept, which cannot be analyzed by one single ratio. Finally, the optimal cut-off points for the variables are chosen by 'trial and error' and on an 'ex post' basis, which means that the actual failure status of the companies in the sample is known (Bilderbeek, 1973). Consequently, the cut-off points may be sample specific and it is possible that the classification accuracy of the univariate model is (much) lower when the model is used in a predictive context (i.e. 'ex ante').

Although MDA is the most frequently used modeling technique in failure prediction, it has some serious ***disadvantages***, additional to the problems related to the violation of the basic assumptions. Firstly, MDA requires that the *classification rule* is linear, which means that a discriminant scores above or below a certain cut-off point automatically signals a good or a poor financial health. In the same respect, the MDA classification rule intuitively contradicts with the fact that some variables do not show a linear relationship to financial health: some variables indicate financial problems both when they have a very low and a very high value. Secondly, we should bear in mind that the discriminant scores are only *ordinal* measures, which allow for a relative (ordinal) ranking between firms. MDA can also generate failure probabilities, but this requires a subjective and possibly inaccurate assessment of the probabilities associated with particular discriminant scores (Zavgren, 1985). Thirdly, although MDA is very similar to the technique of multiple regression analysis, it is computationally not equivalent. The estimation method of least-squares is not suitable when estimating a linear relation with a binary dependent variable (Bilderbeek, 1978; Bilderbeek, 1979).

The MDA technique has the advantage of considering an entire profile of characteristics common to the relevant firms, as well as the interaction of these properties. A univariate study, on the other hand, can only consider the measurements used for group assignments one at a time.

References

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, *23(4)*, 589–609. http://dx.doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, E. I. (1993). Corporate financial distress and bankruptcy, 2. edition, New York: John Wiley & Sons, Inc.

Altman, E. I., & Hotchkiss, E. (2006). Corporate Financial Distress and Bankruptcy: Predict and Avoid Bankruptcy, Analyze and Invest in Distressed Debt, 3. vyd., New Jersey, John Wiley & Sons, Inc.

Beaver, W. (1996). Financial ratios as predictors of failure, *Empirical Research in Accounting*, selected studies supplement to the *Journal of Accounting Research, 4(1),* 71–127.

Bilderbeek, J. (1979). An empirical study of the predictive ability of financial ratios in the Netherlands. *Zeitschrift Fur Betriebswirtschaft, 5,* 388–407.

Bílková, D., Budinský, P., & Vohánka, V. (2009). Pravděpodobnost a statistika, Plzeň: Aleš Čeněk.

Cisko, Š., & Kliještik, T. (2010). Finančný manažment podniku I, Žilina: Edis.

Cisko, Š., & Kliještik, T. (2013). Finančný manažment podniku II, Žilina: Edis.

Eisenbeis, R. (1977). Pitfalls in the application of discriminant analysis in business, finance and economics. *Journal of Finance, 32,* 875–900. http://dx.doi.org/10.1111/j.1540-6261.1977.tb01995.x

Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems, Annals of Eugenics, London.

Gilbert, L. R., Menon, K., Schwartz, K. B. (1990). Predicting Bankruptcy for Firms in Financial Distress. *Journal of Business Finance & Accounting, 17(1)*, 161–171. http://dx.doi.org/10.1111/j.1468-5957.1990.tb00555.x

Herbák, P., Hustopecký, J., Jarošová, E., & Pecáková, I. (2004). Vícerozměrné statistické metódy, Praha: Informatorium.

Hindls, R., Holman, R., & Hronová, S. *et al.* (2003). Ekonomický slovník, Praha: C. H. Beck.

Hillegeist, S. A., Keating, E. K., Cram, D. P., & Lundstedt, K. G. (2004). Assessing the probability of Bankruptcy Review of Accounting Studies.

Jones, F. L. (1987). Current techniques in bankruptcy prediction. *Journal of Accounting Literature, 6,* 131–164.

Kliještik, T., & Dengov, V. (2011). Predikcia na báze viacrozmernej diskriminačnej analýzy, *Společenská odpovědnost firem – transfer vědeckých poznatků do praxe*, Olomouc.

Meloun, M., Militký, J., & Hill, M. (2005). Počítačová analýza vícerozměrných dát v příkladech, Praha: Academia.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18(1),* 109–131. http://dx.doi.org/10.2307/2490395

Rimančík, M. (2007). Štatistika pre prax, vydané nákladom vlastným.

Sekera, B. (2002). Matematické a statistické metody ve financováni, cenných papírech a pojištění, Praha: Profess Consulting s.r.o.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business, 74(1),* 101–124. http://dx.doi.org/10.1086/209665

Stankovičová, I., & Vojtková, M. (2007). Viacrozmerné štatistické metódy s aplikáciami, Bratislava, Iura Edition.

Zavgren, C. V. (1985). Assessing the vulnerability to failure of American industrial firms: A logistic analysis. *Journal of Business Finance and Accounting, 12(1),* 19–45. http://dx.doi.org/10.1111/j.1468-5957.1985.tb00077.x

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research, 22(1),* 59–82. http://dx.doi.org/10.2307/2490859