Olha Kozmenko (Ukraine), Olha Kuzmenko (Ukraine)

# The formation of ratings as a multidimensional function. Express-ratings and time ratings based on the Bayes theorem


## Abstract

It is suggested in the article to calculate express rating and time rating of the study objects according to the theory of chances – Bayes theorem. As a random event is used the achievement of certain level of reliability by the study object. It is proposed to estimate express rating as one-dimensional assessment of input data and time rating as multidimensional assessment of input data.


**Keywords:** express rating, time rating, Bayes formula of hypotheses probability correction, safety level, binary indicator, multidimensional assessment.

## Introduction

The solution of problems related to the economic categories such as reliability and stability of functioning characterizing all objects of study, has recently become a relevant issue. In order to make informed management decisions, that is, the decisions that correspond to the chosen ratio of profitability and risk, it is important for the objects of study to possess the objective information about the current state of their partners. For this purpose a rating is conducted, which allows any user to carry out a comparative assessment of different objects. Therefore, imperfect methods of ranking lead to the biased assessment of the current situation, the waste of time required to get accurate ratings and the losses from transactions with unreliable partners. The problem of improving the methods of ranking the study objects has led to the development of scientific and methodological approach to compiling the rankings based on the express assessment and Bayesian analysis.

The method of express assessment provides the incremental analysis, the implementation of which involves the stages presented in Figure 1 (see Appendix).

During the first stage, a set of indicators that can signal potentially negative aspects of the study object is determined.

During the second stage one conducts the assessment of acceptable (boundary) values for the identified characteristics (formation of a “corridor”of acceptable values).

The third stage is devoted to the formation of binary indicators that primarily depend on the previously obtained boundary values: if the value of an indicator belongs to the “corridor” of acceptable values the corresponding binary indicator assumes the value “0”, and in the opposite case the value – “1”.

During the fourth stage the sum of binary indicators that received the value “1” is calculated. The obtained sum of binary indicators performs the role of express assessment for the future rating of the observed objects.

The characteristic feature of the proposed methods of analysis is economic transparency, that is, substantiation of economic causes for giving a particular score. This can be achieved through a specific type of information, which is the basis for rating assessments. In fact, the state of each object of study is characterized by a set of binary characteristics, which assume the values “1” / “0” (*“0” – if the corresponding characteristic is within the range of acceptable values, and “1” in the opposite case*). The values of “0” indicate the positive and the values of “1” the negative aspects. This makes it possible to obtain the total numerical rating assessment.

In order to obtain binary characteristics we have to determine the boundaries for the “corridor” of the study’s problematic objects. Thus, to determine the values of binary characteristics of each of the $K_j$ indicators we use the following formula for each study object:

$$K_j bin \begin{cases} = 1, K_j \geq K_j av \\ = 0, K_j < K_j av \end{cases}, \qquad (1)$$

where *Kjbin* are the binary characteristics for each study object;

$K_j av$ is the average value of the indicator $K_j$.

In addition, before the calculation of average values for each of the proposed $K_j$ indicators it is necessary to carry out data normalization as the absolute values are not comparable values. Normalization is conducted according to the following formula:

$$Kjnorm \begin{cases} = (Kj - \min(Kj)) / (\max(Kj) - \min(Kj)),\ \min(Kj) < Kj < \max(Kj), \\ = 1;\ Kj = \max(Kj), \\ = 0;\ Kj = \min(Kj) \end{cases} \qquad (2)$$

where $K_j$ is the indicator's initial value; *Kjnorm* are the indicators' normalized values; min (*Kj*) is the minimal initial value for each of the *Kj* indicators; max (*Kj*) is the maximal initial value for each of the *Kj* indicators.

After carrying out the normalization we will determine the average values of the normalized $K_j$ indicators according to the arithmetic mean.

$$Kjav = \frac{\sum Kj}{n}, \tag{3}$$

where *Kjav* is the average value of the $K_j$ indicator; $\sum K_j$ is the sum of normalized values of the $K_j$ indicators; $n$ is the number of the observation objects.

When the empirical standards have been determined each of the study objects will be characterized by a set of binary characteristics $B = (B_1, B_2,\ldots, B_n)$, where $B_k$ assumes the value 0 if the corresponding standard is performed, and the value 1 in the opposite case. The sum of binary indicators for each of the study objects presents coded information about the existing problematic aspects of their functioning.

On the basis of the obtained sums of binary indicators (∑BI) the express rating score is determined for each of the study object:

- if $0 \le \sum BI < 6$ normal level;

- if $6 \le \sum BI < 12$, increased level;

- if $12 \le \sum BI \le 18$, high level.

If, according to the results of the conducted express rating assessment some problematic aspects in the functioning of each of the study objects were discovered it is necessary to carry out a more detailed complex analysis with the use of probability approach.

Thus, one of the promising methods of the rating assessment is the Bayesian analysis. The essence of the proposed method is the obtaining of information about the compliance of the activity of the study objects with the set of standards and the calculation of the possible emergence of certain problematic aspects in their functioning, which is the most informative general indicator of the rating assessment. This can be achieved by using the Bayesian approach in the analysis of information.

The Bayesian analysis as a method of rating assessment based on the use of the Bayes theorem. The Bayes theorem **calculates the probability of the hypothesis correctness under the conditions when on the basis of the observations only some of the information about the event is known**. In other words, the Bayes theorem helps to make more accurate calculations of the probability taking into account both the already available information and the data from the new observations.

The Bayesian approach intends to determine how a priori expectations of a certain phenomenon can be specified and how the observed data can be integrated with such a priori predictions to achieve improved a posteriori expectations of an event.

Let us consider an example of the Bayesian approach: an observer carries out the consistent observations of the study object to identify the problematic aspects of its functioning. The results of each observation or test can be compared with a priori knowledge about the object of study and the expectations regarding the validity of the corresponding assessment. The goal is to obtain the final rating assessment of the object of study, which from the observers' viewpoint is correct with a certain degree of reliability.

The Bayesian formula has two components: an indicator that **characterizes the data of observations and an indicator of the degree of our confidence in the hypothesis validity.** The first indicator is often called the Bayesian criterion or the likelihood ratio. The Bayesian criterion is separate from the subjective component of the formula. This criterion is also called the relative chances criterion, and in the logarithmic form – the proof. The difference between the proved data and the probability of error is discovered if the Bayesian criterion is presented as a coefficient that reflects the degree in the change of probability of the hypothesis correctness after obtaining all the observation's data. The Bayes formula is presented in the following way:

$$P_{an} \times P_{bc} = P_{apr}, \qquad P_{bc} = \frac{P_n}{P_{al}}, \tag{4}$$

where $P_{an}$ are a priori chances the zero hypothesis; $P_{bc}$ is the Bayesian criterion; $P_{apr}$ – the posterior odds of the zero hypothesis; $P_n$ – the probability of obtaining data if the zero hypothesis is correct; $P_{al}$ – the probability of obtaining data if an alternative hypothesis is valid.

The Bayesian criterion demonstrates how each of the two hypotheses corresponds to the obtained data. The hypothesis that best describes the data is considered to be proven.

The proposed rating assessment is carried out in 3 stages.

**The first stage** determines the numerical characteristics of the study objects that make it **possible to conduct rating assessments.**

**During the second stage** the table of the characteristics' acceptable values is filled. If the corresponding characteristic is in the "corridor" of acceptable values it assumes the value of 0, and in the opposite case the value 1.

**The third stage** includes the rating assessments of risks according to the Bayesian analysis using one of the nomograph scales while using another scale to determine the score that characterizes the problematic aspects of functioning.

A distinctive feature of the proposed method is its economic transparency, that is, the validity of economic reasons for giving a particular rating score. It is achieved through the specific type of information determining the rating assessment.

In fact, the state of each study object is characterized by a set of binary characteristics that assume the values "yes" / "no" (*"no" - if the corresponding characteristic is within the range of acceptable values, and "yes" in the opposite case*). The "no" values indicate the positive and "yes" values the negative aspects in the functioning of certain objects of study. It makes it possible to obtain a single numerical rating assessment.

In order to obtain binary characteristics we need to determine the boundaries for the "corridor" of the observations' problematic objects.

When the empirical standards have been determined each of the study objects will be characterized by a set of binary characteristics $B = (B_1, B_2, \ldots, B_n)$, where $B_k$ assumes the value 0 if the corresponding standard is performed, and the value 1 in the opposite case. A series of zeros and ones containing the coded information about the level of the functioning's problematic aspects make it possible to determine the probability ($p_B$ $(H1)$) that the analyzed object of study is problematic subject to the availability of information $B$. Thus, according to the Bayes' formula the following equation holds true.

$$P_B(H1) = \frac{P(H1)\cdot P_{H1}(B)}{P(B)} = \frac{P(H1)\cdot p_{H1}(B)}{\sum_{i=1}^{2} P(Hi)\cdot P_{Hi}(B)} = \frac{P(H1)\cdot P_{H1}(B)}{p(H1)\cdot p_{H1}(B) + p(H2)\cdot p_{H2}(B)} = \frac{1}{1+\dfrac{P(H2)\cdot P_{H2}(B)}{P(H1)\cdot P_{H1}(B)}}. \quad (5)$$

The probabilities $P(H1)$, $P(H2)$ in the Bayesian approach are called a priori, and their values must be determined prior to the analysis. The probability $P(H1) = y$ is the likelihood that due to the lack of a posteriori information the study object is problematic. Accordingly, the probability $P(H2)$ is the likelihood that due to the lack of a posteriori information the analyzed object is not problematic.

The probability $p_{H1}(B)$ is the likelihood that for the problematic object a priori information $B$ will be obtained. Accordingly, the probability $p_{H2}(B)$ is the likelihood that for the non-problematic object a priori information $B$ will be obtained.

In making the assumption about the independence of binary characteristics it is possible to use the formula of probabilities, which leads to the following equation:

$$\frac{P(H2)\cdot p_{H2}(B)}{P(H1)\cdot p_{H1}(B)} = \frac{P(H2)}{P(H1)}\cdot\frac{\prod_{k=1}^{n} P_{H2}(Bk)}{\prod_{k=1}^{n} P_{H1}(Bk)} = \frac{P(H2)}{P(H1)}\cdot\prod_{k=1}^{n}\frac{P_{H2}(Bk)}{P_{H1}(Bk)} = \frac{P(H2)}{P(H1)}\prod_{k=1}^{n}\left(\frac{b_k}{g_k}\right)^{Bk}\left(\frac{1-b_k}{1-g_k}\right)^{1-Bk}, \quad (6)$$

where $b_k$ is the event probability $B_k = 0$, and $g_k$ is the event probability $B_k = 1$ for each object of the study.

Thus, the general formula (6) linking the rating assessment with the available information leads to the following simple equation:

$$p_B(H1) = \frac{1}{1+e^{\{\lambda_0 + L\}}},$$

$$L = \sum_{k=1}^{n} \lambda_k b_k, \quad (7)$$

$$\lambda_k = \ln\left(\frac{b_k(1-g_k)}{g_k(1-b_k)}\right), k = 1,\ldots,n,$$

$$\lambda_0 = \ln\left(\frac{p(H2)}{p(H1)}\right) + \sum_{k=1}^{n}\ln\left(\frac{1-b_k}{1-g_k}\right),$$

where $L$ is the integral indicator (weighted sum) of binary characteristics $B$ (the available information about the state of the study object based on the values of analytical indicators).

To determine the rating score of a particular study object we calculate the value $b_k$ – the probability of the event $B_k = 0$, and $g_k$ – the probability of the event $B_k = 1$ for all indicators $k = 1 \div n$ according to the following formula:

$$g_k = \frac{\sum_k B_k}{n}, \quad (8)$$

$$b_k = 1 - g_k.$$

After determining $b_k$ – the probability of the event $B_k = 0$, and $g_k$ – the probability of the event $B_k = 1$

for all indicators of each study object $k = 1 \div n$ we calculate the parameters $\lambda_k$ та $\lambda_0$ according to the formulas (8) and then determine the value $L$ – an integral indicator (weighted sum) of binary characteristics (the available information about the state of the study object based on the values of analytical indicators) and substitute them in the general formula (7), which shows the value of the rating assessment.

The above mentioned method tested by the National Bank of Ukraine helped conduct the express assessment [7], including the assessment of insurance companies based on the statistical data [4, 5].

The use of Bayesian analysis for the ranking of study objects is an effective economic and mathematical method for improving the quality of supervision over their operations, making it possible to group these objects according to their reliability, obtaining strategic and tactical ratings. The obtained results can be used for practical calculation of ratings in order to make informed management decisions.

The proposed algorithm provides an opportunity to carry out a quick assessment through the use of express ratings (univariate analysis) of economic entities, which is important for the conclusion of agreements. The making of ratings through the use of the Bayes' formula makes it possible to obtain the time estimates of the study objects with a single number that assumes the values from zero to one.

This approach is an innovative technique in which a rating is regarded as a multidimensional value and which involves a series of steps leading to the optimal investigation of the process. Moreover, the advantage of the proposed method is its dimensionless form and the possibility to apply it for any objects of study.

## Conclusion

Bayesian analysis usage for determination of the research object reliability (stability) is an effective economic and mathematical method of enhancing supervision considered by the market. This approach allows to reveal hidden defects in the operation of business entities, to hold the grouping reliability, and to get numeric reliability characteristics of the population considered in contrast to traditional methods, which give only a descriptive characteristic. But there is a necessity for continual adjustment of this method in accordance with the needs of the current economic situation.

## References


1. Basel II: International Convergence of Capital Measurement and Capital Standards: a Revised Framework [Electronic resource]: Basel Committee on Banking Supervision. Access mode: http://www.bis.org/publ/bcbs107.pdf?noframes=1.
2. Core Principles for Effective Banking Supervision [Electronic resource]: Basle Committee on Banking Supervision. Access mode: http://www.bis.org/publ/bcbs30a.pdf?noframes=1.
3. Guidance for Money Service Businesses – Risk-Based Approach [Electronic resource]: FATF, July 2009. Access mode: http://www.fatf-gafi.org/dataoecd/45/1/43249256.pdf.
4. Ivchenko, I.Yu. (2004). *Ekonomichni ryzyky* [Economic risks], Kiev: Center of scientific literature, 2004, pp. 60-61.
5. Kozmenko, O.V. (2008). *Reityngguvannia strakhovych kompaniy i rozrakhunok strakhovykh taryfiv na bazi vykorystannia ekonomiko-matematychnykh metodiv* [Rating of insurance companies and calculation of insurance tariffs on the basis of economic and mathematical methods] [Text]: monography / O.V. Kozmenko. – Sumy: UAB NBU, 95 p.
6. Kozmenko, O.V., Kozmenko, S.N. (2008). Ekonomiko-matematycheskiye metody v strachovanii ekologicheskich i katastroficheskich riskov // Upravleniye proizvodstvom: problemy teorii i praktiki. Monografiya [Economic and mathematical methods in insurance of ecologic and catastrophe risks // Production management: problems of theories and practice. Monography] / E.V. Martiakova (editorship). – Donetsk: Donetsk National Technical University, pp. 403-410.
7. Lukyanova, V.V. (2008). *Ekonomichny ryzyk* [Economic risk] / Lukyanova V.V., Holovach T.V. – Kiev: Akademvidav, 2007, 464 p.
8. *Modeliuvannia otsinku ryzykiv vykorystannia poslug bankiv abo yich strukturnych odynyts' z metoyu legalizatsii kryminal'nych dochodiv abo finansuvannia terorismu: monografiia* [Modeling of evaluation of risks of using bank services or its structural units in order to legalize criminal revenues or terrorism financing: monography] / S.O. Dmytrov, O.V. Merenkova, L.G. Levchenko, T.A. Medvid', O.M. Berezhny (editorship). – Sumy: UAB NBU, 2008. – 75 p.
9. Pravyla organizatsii statystychnoyi zvitnosti scho podayet'sia do Natsional'nogo banku Ukrainy [Regulations of organization of statistical reporting that is given to the National Bank of Ukraine] [Electronic resource]: Resolution of the National Bank of Ukraine Board from 19.03.2003 No. 124 as of January 1, 2008. Access mode: http://zakon.nau.ua/doc/?uid=1030.1286.0.
10. System of risks evaluation [Electronic resource] / guidelines on inspection on banks of the NBU Board. Access mode: http://zakon.nau.ua/doc/?uid=1045.5669.1&nobreak=1.
11. Vitlins'ky, V.V. (2004). *Ryzykologiia v ekonomitsi ta pidpryiemstvi* [Riskology in economics and enterprises] [Text]: monography / Vitlins'ky V.V., Velykoivanenko G.I. – Kiev: KNEU, 480 p.

## Appendix

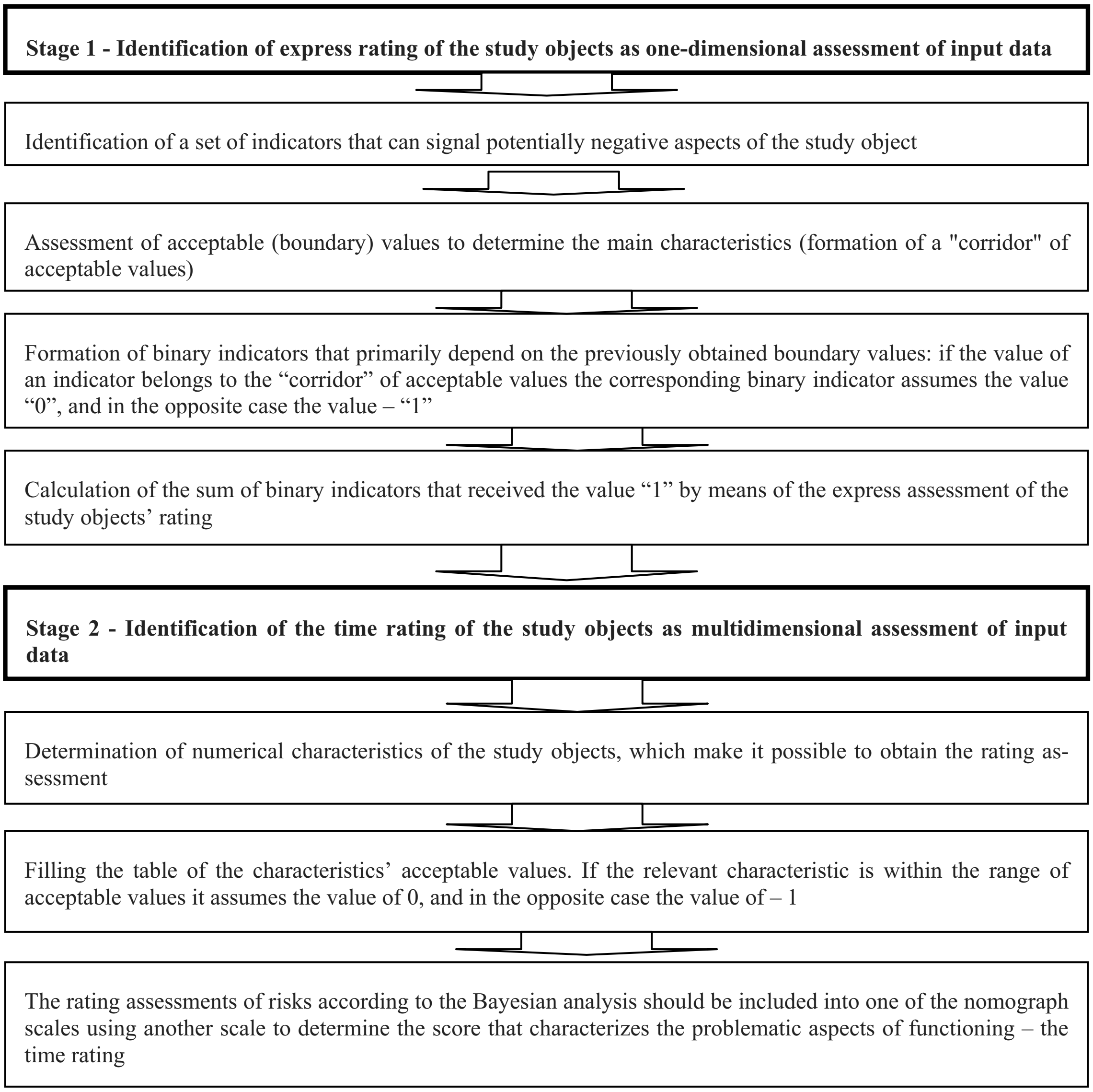


**Fig. 1. Generalized scheme for the formation of the express rating and the time rating based on the Bayes’ formula**