April 2015, Frankfurt am Main.

DESCRIPTION OF MATHEMATICAL MODELS USED IN RATING ACTIVITIES

This document outlines the principles and key assumptions underlying the rating models and methodologies of Rating-Agentur Expert RA GmbH. Rating methodologies used for assigning credit and reliability ratings are based on the rating models. The rating model for assigning a credit rating represents a regression equation where the explained variable is usually the probability of default (delayed and/or incomplete fulfillment of obligations) of the rating object and the explanatory variables are the quantitative and qualitative (usually in the form of dummy variables) factors, information for which may be received from the rating object or other sources of information available to the Agency.

Common principles of building rating models

The purpose of designing a rating model is identification of factors largely influencing the probability of default (creditworthiness) of the rating object as well as the form of this influence (positive/negative, linear/non-linear dependency).

The rating model is the result of processing the database of source information using the methods of econometrics and mathematic statistics. Building a statistically significant regression equation requires fulfilling a number of assumptions regarding the source data. Key assumptions regarding the source data are provided below:

- The market of the subjects to which the methodology is applied, can be considered as relatively mature, therefore the time series data is available for at least 8 quarters.
- Total number of objects of this type (or their analogues) for which information can be gathered is not less than 20.
- Information is available at least for 5 cases of defaults of the objects of this type (or their analogues).
- Distribution of the source data approximates to a normal distribution (preferably), logistic distribution or Gompertz distribution (for checking the hypothesis about normal distribution, the Pearson criteria is used as the main method. A 0,05 significance level is used on the basis of comparison of theoretical and empirical frequencies).

The collection of information about the objects similar to the rating object is done using both free and fee-based sources of information. With increased volume of available information (in particular, from information provided by the rated entities), improvement of the rating models shall be done by including new factors.

When forming the set of explanatory variables it is compulsory to take into account the experience of preparing rating models for similar markets. In particular, large weight in rating models is usually given to indicators of capital adequacy, quality of assets, profitability and liquidity. Therefore, when starting to work on a new rating model, the set of quantifiable indicators (which can be evaluated) is formed so it can approximate corresponding factors for this market mentioned above.

Modelling is usually based on the probit-regression (model of binary choice: 1 – entity has defaulted within the period, 0 – there was no default), evaluation of its parameters is done on the basis of the most likelihood method. In some cases the logit-regressions are used (source data corresponds to the logistic distribution) or gompit-regression (source data corresponds to the Gompertz distribution).

Determination of the rating methodology

The rating methodology is a set of rules by which a rating object can be assigned to one of the rating classes according to the array of data available for this rating object.

From a mathematical point of view by the methodology of rating assessment the superposition of functions evaluating particular characteristics of the rated object is assumed, converted initially to the probability of default and then to a particular rating class on the basis of a theoretical default matrix.

$$R = Pd(\sum_{i=1}^{n} g_n(\overline{p_n})) \overset{j}{\Leftrightarrow} Md$$

In the above formula:

R – rating class, $R \in Rating\ scale$

Pd – monotone function of transforming the rating score into the probability of default

g_n – function of evaluating the parameter p_n,

$\overline{p_n}$ – parametric vector

Md – defined theoretical default matrix

j – function of displaying the default probability set in form of a rating class

$\overline{p_n}$ represented by three types of data:

1) *Degenerate vector in the form of scalar one-dimensional value*
2) *Usual scalar vector*
3) *Vector of binary values (usually this is how the evaluation of qualitative factors looks like on the basis of so called "check-list")*

g_n is presented by the following types of functions:

1) *Linear function*
2) *Nonlinear function (power transformation, logarithms)*
3) *Piecewise-nonlinear function*

g_n with bifurcation non-linearity (piecewise linear with sharp change of the response) is used for support and stress factors

The theoretical default matrix defines the correspondence of rating classes to the ranges of default probabilities. Usually the lower the rating class the wider the range of default

probabilities which corresponds to it. Dramatic increase of the default probability range usually occurs in the transition from B- class to CCC+ class (according to the international rating scale of the agency).

Common principles of creating rating methodologies

Usually the rating methodology is based on rating models, which are statistically significant at the 0,1 significance level (further referred as alpha). However, for exceptional cases of immature markets, a model with alpha up to 0,2 can be used (in this case the rating methodology includes more conservative requirements).

When creating the rating methodology a whole range of principles shall be taken into account, which are not included by the rating model due to various reasons, including the specifics of the source data (small number of extremely large and extremely small values). The most important principles are provided below:

- **Principle of significance.** Even if the impact on creditworthiness of the corresponding factor within the framework of the rating model is statistically not significant (with alpha less than 0,1) the rating methodology shall include an analysis of all the accounts of the balance sheet and P&L statements on profits and losses exceeding 5% of the assets (10% of the revenue) in order to identify their economic meaning. This is necessary to identify unusual ways of asset stripping or falsification of reports.
- **Principle of reasonable trust.** All significant information, which can be verified, shall be verified. In this regard, the rating methodology shall contain a range of tests to verify mistakes in the provided data and provide penalties for intentional and unintentional mistakes.
- **Principle of grouping the factors.** Due to the fact that increasing the number of explanatory variables (factors) might lead to a decrease in the quality of the rating model (if the above defined threshold is depending on the number of observations in the source data set, this can lead to inability to use some methods of evaluation, in particular the least squares method), indicators close by meaning (for example profitability of assets and profitability of capital) can be grouped before their inclusion in the tested rating model.
- **Principle of strong factors.** Among the analyzed indicators, the factors, which have significant influence on the financial stability of the rated entity (support and stress factors), shall be identified. These factors shall be taken into account separately and have maximum weight in the methodology.
- **Principle of cumulating the risks.** Significance of negative factors increases non-linearly in case of their mutual influence on each other. This mutual influence can be taken into account when forming stress factors in the rating methodology.
- **Principle of supremacy of content over form.** In the process of analysis, priority shall be given to economic and not accounting content of the operations. In relation to that, the rating methodology can take into account the expert's corrections obtained within the framework of the rating model of evaluating the explanatory

variables with compulsory justification and link to the factors not included in the rating model.

After taking into account the principles mentioned above, the obtained rating methodology shall be calibrated by testing on the available sample of defaulted objects; the form of the g_n shall be fixed, coefficients shall be selected so that the value of the function $Pd(\sum_{i=1}^{n} g_n(\overline{p_n}))$ for all defaulted objects was approaching 1 on the horizon of 6 months prior to the default. Then it should be checked that the obtained coefficients are not giving false triggers for the majority of non-defaulted objects.

Back-testing of the rating methodologies

Basic back-testing of the rating methodologies is carried out by constructing the empirical default matrix and comparing it with the theoretical. Serious deviations of the actual default matrix from the theoretical can indicate a reason to review the rating methodology.

Default matrix - snapshot of the transition matrix, which represents a Markov transition matrix with discrete time:

$$\begin{pmatrix} P_{ij} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{pmatrix}$$, where:

P_{ij} – Probability of transition of the rating objects from one rating class to another (usually on the horizon of 15 months). The transition matrix is constructed continuously.

For the default matrix, only the probabilities of transition to “default” class are taken into account.

In-depth back-testing includes an analysis of the whole transition matrix with the purpose of identifying abnormally high probabilities of transitions from one rating class to another. Abnormally high transition probabilities may indicate the necessity of applying more conservative approaches in board line case, which may also require review of the rating methodology.

Frequency of back-testing depends on the specifics of the rated objects. Back-testing of all methodologies shall be done at least once within 2 years and 2-3 times more frequently for dynamically changing markets (including immature) with accumulating statistics of defaults.