DOI: 10.20535/kpisn.2021.4.251684
UDC 004.942:519.216.3

N.V. Kuznietsova[1], O.M. Trofymchuk[2], P.I. Bidyuk[1]*, O.M. Terentiev[2], L.B. Levenchuk[1]

[1] Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine
[2] Institute of Telecommunications and Global Information Sphere at NAS of Ukraine
*corresponding author: pbidyuke_00@ukr.net

BAYESIAN MODELLING OF RISKS OF VARIOUS ORIGIN

Background. Financial as well as many other types of risks are inherent to all types of human activities. The problem is to construct adequate mathematical description for the formal representation of risks selected and to use it for possible loss estimation and forecasting. The loss estimation can be based upon processing available data and relevant expert estimates characterizing history and current state of the processes considered. An appropriate instrumentation for modelling and estimating risks of possible losses provides probabilistic approach including Bayesian techniques known today as Bayesian programming methodology.

Objective. The purpose of the paper is to perform overview of some Bayesian data processing methods providing a possibility for constructing models of financial risks selected. To use statistical data to develop a new model of Bayesian type so that to describe formally operational risk that can occur in the information processing procedures.

Methods. The methods used for data processing and model constructing refer to Bayesian programming methodology. Also Bayes theorem was directly applied to operational risk assessment in its formulation for discrete events and discrete parameters.

Results. The proposed approach to modelling was applied to building a model of operational risk associated with incorrect information processing. To construct and apply the model to risk estimation the risk problem was analysed, appropriate variables were selected, and prior conditional probabilities were estimated. Functioning of the models constructed was demonstrated with illustrative examples.

Conclusions. Modelling and estimating financial and other type of risks is important practical problem that can be solved using the methodology of Bayesian programming providing the possibility for identification and taking into consideration uncertainties of data and expert estimates. The risk model constructed with the methodology proposed illustrates the possibilities of applying the Bayesian methods to solving the risk estimation problems.

Keywords: financial processes; financial risks; Bayesian programming methodology; risk estimation.

Introduction

Repeating financial crises, unfavourable changes of climate, local military conflicts between many countries of the world and fighting terrorists in multiple locations give an evidence for high necessity of solving the problems of analysis and management of growing various type risks in every area of human activity. Together with these highly unfavourable for economic and social developments events the facts are revealed that existing methods of risk analysis, and modelling the situations concentrated on creating their formal description and appropriate risk management procedures usually come with some delays or are inadequate for the quality risk estimating and forecasting in conditions of multiple random external disturbances (risk factors). This is mostly explained by the high dynamics of modern processes (especially financial ones, ecological and climate changes), their high dimensionality, sophisticated vertical and horizontal interconnections at the level of separate companies, economy branches, macro-economy as a whole and at the global level as well [1]. Actually most of the risks met in everyday life, including the risks of natural disasters, can be analysed from the financial point of view, and this

Рекомендуємо цитувати цю статтю так: N.V. Kuznietsova, O.M. Trofymchuk, P.I. Bidyuk, O.M. Terentiev, L.B. Levenchuk, “Bayesian modelling of risks of various origin”, *Наукові вісті КПІ*, no. 4, pp. 7–18, 2021. doi: 10.20535/kpisn.2021.4.251684.

Please cite this article as: N.V. Kuznietsova, O.M. Trofymchuk, P.I. Bidyuk, O.M. Terentiev, and L.B. Levenchuk “Bayesian modelling of risks of various origin”, *KPI Science News*, no. 4, pp. 7–18, 2021. doi: 10.20535/kpisn.2021.4.251684.

way we can find an estimate for possible loss and construct appropriate risk management algorithms.

Sometimes mathematical models available are too complicated for practical use, and the necessity emerges to create simpler formal descriptions of risk, adequacy of which can be substantially different from the ideal ones. Generally speaking any model represents somewhat simplified representation of the world that can lead to incomplete description of situations with uncertainties, inadequacy and wrong forecasts as well as to incorrect decisions based upon the forecasts. That is why one of the most urgent problems that can be often met in the risk management processes is development of appropriate models adequate enough for practical use. The models should be acceptable for practitioners and, when necessary, supplied with extra structural nonlinear elements; they should have a possibility for correcting prior expert estimates of parameters and initial conditions, restrictions, experimental data, and with clearly defined possible practical applications.

A substantial role regarding timely and high quality problem solving to perform the risk modelling and management plays systemic approach. It means that the approach supposes taking into consideration current market factors (including random ones); possibly revealing new deterministic and stochastic factors of influence (including the hidden ones); estimating the scale and frequency of their influence on the processes of interest; identifying and taking into consideration possible structural, statistical and parametric uncertainties met in the processes of model constructing, estimating of forecasts for relevant processes development and estimating respective risk. In most cases it is necessary to perform correct problem stating and solving optimization problems directed towards risk loss minimization. Another problem related to systemic approach of risk analysis is in providing several sets of statistical quality criteria related to analysis of data quality, adequacy of models, quality of forecasts and decision alternatives generated on the basis of the models and forecasts. These sets of criteria provide the possibility for achieving high quality results of computing at each stage of data processing, model constructing and risk estimation [2], [3]. Usually it is more convenient to perform such analysis in the frames of appropriately designed and implemented specialized decision support system (DSS) [3].

Analysis of financial processes development and their internal and external interaction within the last several decades highlights the needs of special attention, from the point of view of financial risk management, for banking system, investment and insurance as well as for large and medium level firms. Especially sophisticated and highly dynamic are financial processes in the area of market processes, insurance and respective risk situations. This is explained by the fact that market and insurance areas are directly connected with many other dynamic processes in international banking system, production, tourism, cargo and passenger transportation, and also with natural and industrial catastrophes etc. Just market and actuarial activities are in the focus of solving complex everyday financial problems at all levels of economy and private activities. The activities require correct practical applications of high quality mathematical models, methods and knowledge of data analysis techniques and procedures. The mathematical methods and data models as well as decision support systems based upon them do not replace professionals making final decisions but they provide the possibility for much deeper analysis and understanding of related processes, for improved data and expert estimates processing, for generating possible objective alternatives, and select objectively the best decision for specific application.

A crucial role in analysis of risk and estimating possible loss plays probabilistic approach to data analysis, modelling and forecasting. Say, Bayesian paradigm in the form of systemic Bayesian programming creates appropriate probabilistic-and-statistical instrumentation to fight uncertainties and provide appropriate risk analysis results. Practically all the probabilities we have to cope with in solving most practical problems in every area are conditional. These conditions lead to emerging multiple special cases for analysis of specific problems including risk management. That is why application of the Bayesian modelling approach to risk analysis is very important, useful and appropriate what is supported by many available examples from the past developments.

The study is focused on some aspects of modelling and estimation of financial and some other type of risks, application of related mathematical models to solving practical risk estimation and forecasting problems. The study supposes constructing and implementation of appropriate DSS providing for all necessary computational procedures for reaching high quality results at each step of data analysis, risk modelling and estimation.

Problem statement

The purpose of the study is to: determine basic types of risks in different areas of human activities and consider the possibilities for their mathematical description; determine the possibility of hiring

Bayesian approach to constructing mathematical models of financial as well as some other types of risk, and show examples of constructing Bayesian type models for stochastic financial process.

Operational risks

Definition of many types of risk including the financial one is linked to the probability of events that can be accompanied by some material loss, and the level of the loss. An International Standard Organization gave the following definition to risk: "risk is a combination of probability for some event and its consequences [4]". The study [5] formulated the following definition: "risk is a set of possible scenarios, s_i, each of which is characterized by the probability p_i, and the consequence c_i". This definition is very general, robust and can be hired for solving engineering and financial problems. The most practical problems exhibit multivariate risk, i.e. there exist multiple internal and external risk factors that in combination create general situation leading to emergence of risk.

For example, the actuarial activities are characterized by the set of multiple risks with the most known among them are as follows: individual risks; collective risks for one (short) period of time; collective risks for long periods; high distributed risks of loss; operational risks; non-return risk of credit; bankruptcy risk, and other type of risks [6]—[8]. A substantial loss comes to companies and various enterprises today due to availability of operational risk that exists in any organization and can be considered as the "universal" one. It can be viewed as the risk of direct and indirect loss that comes to being due to inappropriate organization of working activities or inappropriate organization of internal processes in a company, incorrect behaviour of a company staff, and/or incorrect functioning of technical equipment, or due to influence of unfavourable external factors.

The operational risk can also be provoked by the absence of appropriate methods and means for management of this type of risk. The operational risk should be analysed qualitatively and quantitatively as well as any other type of risk that requires collecting and thorough analysing appropriate statistical data and expert estimates. As far as emergence of the risk is influenced by the most different events, the problems of collecting necessary data, model constructing, and estimating the volume of possible loss and its probability require substantial efforts of experts in information technology, mathematical modelling, forecasting and decision support systems.

As a possible source of statistical data could serve insurances policies, that contain information regarding possible insurance risks that come to being due to the events resulting in operational loss. However, this is not the best possibility for collecting necessary information because the policies contain confidential information about clients, and the task of processing policies is complicated, time consuming process that is not distinguished with the information completeness regarding the problem stated. As of today, to get mathematical description for all types of the insurance risks the following methods are widely used: applied statistics, probability theory, fuzzy logic, Bayesian theory of data processing, neural networks etc.

In the process of analysing the situations leading to appearance of financial risk it is important to get objective information regarding current state of insurance company from independent sources. It is necessary to establish correctness of functioning of the risk management department, to study the information accessible for the department and the methods of its processing. One of the most important characteristics of this information is its completeness, i.e. does it contain enough data for constructing the model capable to forecast the volume of possible loss? It may turn out rather often that the information is not complete for discovering all types of possible risks that may emerge in a company. The loss discovered may not reflect all possible types of risks met by the company because not all possible risks resulted in the loss and could be ignored.

In such cases additional analyses should be performed directed towards discovering the following events: (1) determining the moment of time when the financial loss occurred as well as establishing the fact of making the decision resulting in the loss; (2) estimation of the possible income that company could get in the case of avoiding the risky situation; (3) distribution of actual financial loss among the existing several possible risk factors when the number of actual risk factors is greater than one; (4) collection of additional information from the company personnel who is related to the risky situation that took place. The purpose of the additional situational analysis is reconstruction of the sequence of events that actually resulted in the loss, discovering the reasons for happening of the events and establishing the possibilities of their avoiding. It would also be logical to establish the reasons why the events were not avoided. The probability always exists that not all the losses were discovered and were taken into account by the risk management department. That is why the interaction with the company personnel can reveal additional information regarding other possible loss that was avoided or it remains actual for further studying.

Generally the procedure for identifying the types of risks and their management can be represented in the form of the following cyclical sequence of actions: (1) establishing the possible types of risk for a company; (2) identification, studying, deeper understanding and description of the situations that are favourable regarding development of the risk factors; (3) a thorough analysis of possible risks with establishing appropriate risk measures, loss estimation and forecasting; (4) development and making appropriate control decisions regarding specific risks management; (5) tracking implementation into life the managerial decisions developed, detecting and analysis of the indicators for emergence of possible risks; (6) making detailed reports on the actions directed towards avoiding, ignoring or active control of the situations with risk emergence. Implementation of the cyclical procedure for risk identification and management of respective situations should be based upon classification and continuous analysis of possible risks for a selected company. Very often the term is used like "*enterprise risk*" that includes all possible risks for a specific enterprise. The enterprise risk is divided into *basic risk for business* and *operational risk*; and both types of risk include their specific components.

Today there exist a set of quantitative approaches to risk analysis that can be applied to solving the problem of deeper understanding the essence and estimation the financial risk level. The methods of this class include the following ones: (1) *statistical estimation* — empirical studies; estimation of maximum possible loss; construction of appropriate probability distribution functions; linear and nonlinear regression analysis; (2) *frequency analysis of loss* — frequency analysis of the level of loss; theory of extreme values; stochastic differential equations; (3) *statistical Bayesian approach* (Bayesian programming techniques) — dynamic Bayesian models for the systems under study; Bayesian belief networks (BBN), and causal models; process development charts; (4) *artificial intelligence systems* — neural nets, neuro-fuzzy models and decision trees for client and enterprises classification as well as estimation of the possible loss risk; (5) *Monte Carlo based models* and the *mode switching models* — generation of the development scenarios, methods of income analysis; strategic investment analysis; (6) *expert estimation and fuzzy logic* — direct likelihood estimation for the development scenarios; Delphi method; the models for capital and price forming; market risk estimation in actuarial studies; (7) *practical approaches to risk estimation and management* — stress testing and scenario analysis; industrial and business scenarios; dynamic financial analysis; market beta-comparison of separate companies in the frames of market sectors.

The mathematical models that are hired for the formal description of processes related to risk estimation can be classified as deterministic and stochastic. It is known that such models represent simplified formal pictures of possible consequences for the future uncertain events. The uncertainty is represented in this case with the time of emergence of the events and their consequences, as possible loss that in most cases can be related to the financial one. Though it should be stressed that when mathematical model is hired for a study the forecast estimate can be computed in the form of quite definite value but this "certainty" is based upon the assumptions that are uncertain by their nature. It is useful to remember that the data available practically always contain deterministic and stochastic components, and this fact requires including at least one stochastic variable into deterministic model.

This is related to the probability distributions of random variables used for analysis, selection of the methods for model structure and parameter estimation, determining the type and parameters for stochastic disturbances, selecting the technique for forecast estimation etc. If the assumptions accepted in the process of model constructing correspond to actual behaviour of the process being studied and the future developments of the process correspond to the forecast estimates generated by the model then such estimates could be used for decision making. Probabilistic models are directed to forecasting probabilities of the events (processes) using the historical information regarding former behaviour of the processes. The models are based upon the hypothesis and assumptions that are logically coordinated with the probable development of the events in the future taking into consideration possible uncertainties. It should be noted here that the forecasts computed with the models cannot be completely deterministic (certain) because their complete certainty would actually mean practical uselessness of the model.

Bayesian programming methods

Today there exist a set of Bayesian methods of filtering, modelling, forecasting and decision making known under the general name of Bayesian programming or Bayesian methodology. The methodology provides the possibility for solving the following problems:

– constructing probabilistic and statistical models (model structure and parameters estimation) using statistical (experimental) data and expert estimates;

– computing final results on the basis of the models created according to the specific problem statement: estimates of forecasts; control actions; estimates of variables and parameters using filters; image recognition; making managerial decisions regarding the process and systems under investigation, and many other tasks;

– analysis of quality of the results received at each stage of data processing by making use of appropriate sets of quality criteria.

Some methods related to the Bayesian programming methodology include the techniques in short mentioned below.

1. Recursive Bayesian estimation: filtering, forecasting, and smoothing the data. The basic equation of estimation has the following form:

$$P(S(k)\,|\,O(0)...O(k)) = P(O(k)\,|\,S(k))\times$$
$$\times\sum_{S(k-1)}\begin{bmatrix}(P(S(k)\,|\,S(k-1))\times \\ \times(P(S(k-1)\,|\,O(0)...O(k-1)\end{bmatrix},$$

where, $S(0),\ldots, S(k)$, is time series of state variables; $O(0),\ldots, O(k)$ is time series of observations; $P(S(k)\,|\,S(k-1))$ is system or state transition model; $P(O(k)\,|\,S(k))$ is conditional model for observations that shows what would be observation at the moment k, if the system is in the state $S(k)$.

This model can be used for computing future distribution of states $P(S(k+l)\,|\,O(0)\ldots O(k))$ at the moment $k+l$ on the basis of the available observations $O(0),\ldots, O(k)$.

When $l=0$, the filtering procedure is implemented; if $l>0$, then forecast estimate is generated, and when $l<0$, the data is being smoothed. Smoothing means restoring of former state on the basis of the observations received before or after the moment of smoothing.

2. Hidden Markov models (HMM) – is modification of Bayesian filter which supposes that data is discrete; the state transition models and observations are defined by the probability matrices or conditional probability tables. If the variables observed are continuous such models are called semi-continuous HMM.

3. Optimal recursive Kalman filters (KF) can process continuous or discrete data. The state transition and observation models are constructed in this case with the use of Gaussian processes for describing external random disturbances and measurement noise. If nonlinear models are hired then Tailor expansion is used what results in application of local linear models. For simultaneous estimation of parameters and states an Extended Kalman Filter (EKF) can be applied.

4. Particle (granular) filters (PF) are used for data processing using the following distribution model: $P(S(k-1)\,|\,O(0),\ldots,O(k-1))$, where $S(\cdot)$ is matrix of states, and $O(\cdot)$ is observation matrix. The observations are approximated by the set of particles with weighting coefficients proportional to the probabilities of their occurring. The state probabilities are renewed by recursive procedure.

5. Static Bayesian networks (BN) are probabilistic and statistical models used to describe formally available data and expert estimates in conditions of uncertainty. To the net variables are imposed practically no restrictions, and no special semantics is used for their description. Thus, definite freedom exists for selecting the variables and constructing the network model.

Graphically BN is represented by directed acyclic graph the vertices of which are variables of the network, and the arcs indicate existing conditional relations between the variables. Formally the model is described by the triple: $BN = \{V, G, T\}$, where V represents the set of variables for constructing the net (data base); G is directed acyclic graph constructed with application of optimization procedures; T is conditional probability tables associated with the graph vertices (for parent vertices these probabilities are unconditional).

The model parameters are created by conditional probabilities in the conditional probability tables. Continuous variables are represented by the specific distributions. If BN contains both discrete and continuous variables it is called a hybrid one. Continuous variables in such cases are transformed into discrete ones using known procedures.

The result of constructing and application of BN is provided in the form probabilistic inference like following: $P(X^i\,|\,Known)$, where *Known* means the set of other network variables the probabilities of states for which are known. Generally the probabilistic inference generated by BN is in propagation of probabilities and parameters of Gaussian distribution laws over the whole network and depends on availability of evidences – additional information about the network states. The process of forming (computing) the inference is based upon rather complicated mathematical algorithms. One of the simplest approaches is based upon Bayes rule.

6. Dynamic Bayesian networks (DBN) are models of the same type as static BN but the difference is that they take into account dynamics of the processes being modelled (i.e. their evolution in time) as well as possible stochastic influences to the

basic variables. Actually, DBN represents further evolution of static networks directed towards taking into consideration evolution of the processes considered in time. First static network is constructed for the variables selected and data available, that is considered as time invariant one. Then this model structure is repeated for each next moment of time with coming new observations (measurements).

7. Part of the model graph that corresponds to the specific moment of time t_k, or simply k, is called time cross-section. If hypothesis is accepted that the current system state (being modelled) depends on the previous state only then this assumption is called *first order Markov assumption*. And if the model structure remains unchanged for all the time cross-sections then such DBN is called stationary one. In such cases model that corresponds to one moment of time is defined as local and time-invariant or homogeneous.

8. Markov localization (ML) models: these are models of the Bayesian filter type that additionally include control variables, $\mathbf{u}(0),\ldots,\mathbf{u}(k-1)$. They also have another name of *hidden Markov input-output models*. Such models provide a possibility for improving state estimates using control variables as follows: $P(S(k)\mid \mathbf{u}(k-1), S(k-1))$. Sometimes the model is called action model. Generally these models are represented in matrix form similar to those used in particle filters under the name of Monte Carlo with Markov localization (MCML). The model constructed should give an answer to the question: what is the probability of current state of a system under study:

$$P(S(k)\mid \mathbf{u}(0),\ldots,\mathbf{u}(k-1), O(0),\ldots,O(k)),$$

using the information regarding former control actions and observations of previous states. The term "localization" is linked to robotic system applications, i.e. localization of a robot in space. The basic state equation is similar to the basic filtering equation given above.

9. *Bayesian maps*: the name that means generalization of the Markov localization models that also were originated in the sphere of robotic control systems.

10. Bayesian approach to data processing and decision making using hierarchical models.

11. Bayesian regression in the form of generalized linear models (GLM).

Bayesian approach to modelling and decision making

The Bayesian approach to modelling and computing probabilistic inference supposes fulfilling the following stages given in Fig. 1. According to the approach the initial information is coming from two sources: prior information from researcher and statistical data received as a result of performing appropriate experiments.

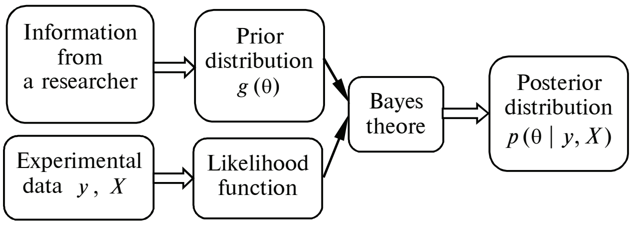

Fig. 1 Information flows in the data processing system based upon Bayes theorem

The prior information from a researcher is based upon the results of his former studies; theoretical information regarding the processes under study; informal data received from various sources. Generally this is extra information that is added to the experimental one. The Bayes rule (BR) is in the center of the methodology, and it can be written as follows:

$$p(\theta\mid y,\mathbf{X}) = \frac{g(\theta)p(y\mid\theta,\mathbf{X})}{p(y)} \mu g(\theta)p(y\mid\theta,\mathbf{X}), \quad (1)$$

where y is observation vector for dependent variable for the system under study; $\mathbf{X}$ is observation matrix for independent variables (regressors) that influence behaviour of the dependent one; θ is vector of random parameters that together with matrix $\mathbf{X}$ define probability density function (PDF) for y (if we refer to classic regression, θ is vector of the regression); $p(\cdot)$ is PDF; $p(\theta)$ is prior density for random parameter θ, that is based upon prior knowledge of a researcher (y, $\mathbf{X}$) before applying experimental results; $p(\theta\mid y, \mathbf{X})$ is conditional PDF for the data y given θ and $\mathbf{X}$, in other words, his is likelihood function for data y; $p(y)$ is marginal probability for y after excluding influence of independent variables and parameters. Here $p(\theta\mid y, \mathbf{X})$ is posterior PDF for parameter, θ, based upon renovation of the prior distribution by the data (y, X).

Thus BR shows the way how to combine prior distribution with experimental data to correct the prior PDF. Actually the BR provides the possibility not only for combining information coming from two sources but any other way of combining will violate logical (and mathematical) essence of the rules for operating with PDF.

The expression (1) contains two variants of representing the Bayes rule. The first one is given in the form of equality due to availability of denominator $p(y)$ (unconditional density for, y) that plays the role of normalizing constant providing condition that posterior conditional density for θ is proper and integrates to 1 over the parameter definition range. The second variant of the rule representation is given without normalizing constant what simplifies the expression. In practice, first the nominator is computed, and then the normalizing constant if necessary. In some cases only ratio of results is of interest and the normalizing constant is cancelled.

Application of Bayesian approach to risk modelling and estimation has the following advantages:

– this approach is convenient for taking into consideration various uncertainties that are related to statistical data like imputation of missing observations, noise and short samples processing; appropriate state and parameter estimation procedures; application of different data formats; hiring of expert estimates in modelling procedures; possibilities for application of high dimension of data and models; the use of static and dynamic approaches to modelling various systems; availability of alternative procedures for model structure and parameter estimation; possibilities for model adaptation and constructing simulation procedures etc.;

– possibilities for constructing complex hierarchical models;

– possibilities for constructing combined models that include linear and nonlinear regression equations, Bayesian and neural networks, fuzzy logic etc.;

– taking into consideration probabilistic uncertainties of the type: “will some event happen or not”, and “under what conditions it will happen”;

– elimination of amplitude-like uncertainties by replacing missing measurements with estimated or generated values;

– availability of linear and nonlinear filtering procedures;

– possibilities for application of forecasting distributions;

– availability of possibilities for application of Bayesian optimization.

Bayesian approach to the formal description of risk

Consider possible approach to formal description of a risk using Bayesian approach. Let $\mathbf{M}=\{M_1, M_2, ..., M_n\}$ is a set of risk describing models; $p(x \mid M_i, \theta_i)$ is a likelihood function for the model $M_i, i=1,...,N$ with parameters θ_i and available data, x; and $p(\theta_i \mid M_i)$ is prior distribution for parameter vector θ_i related to selected model M_i. According to Bayes rule posterior distribution of vector θ_i given known model structure M_i and data x can be written in the form:

$$p(\theta_i \mid M_i, x) = c_i p(x \mid M_i, \theta_i) p(\theta_i \mid M_i),$$

where c_i is normalizing constant. Various model structures may differ with the number of parameters and better models are considered those that have lower number of parameters, according to the economy principle, but provide the possibility for acceptable quality (adequacy) of formal description for the process under study. To select the best model it is possible to use appropriately modified Akaike information criterion (AIC) [9]:

$$\log \tilde{p}(x \mid M_i, \theta_i) = \log(x \mid M_i, \theta_i) - A(k_i),$$

where k_i is dimension of vector θ_i; $A(k_i) = k_i$ is increasing function of k_i. This criterion can be represented in the form that is minimized for the best model, i.e. $-2\log p(x \mid M_i, \theta_i) + 2k_i$. An alternative for the AIC is Bayes-Schwartz criterion, where $A(k_i) = 0.5k_i \log N$, and where N is sample power for data $\{x\}$. Some other modifications of the criteria are hired for analysis of regression models of various structures.

After selection of the model structure, $M^* = M_{i^*}$, it is necessary maximize posterior density $p(\theta_{i^*} \mid M_{i^*}, x)$ for θ_{i^*} to find better estimate of the parameter vector $\hat{\theta}_i$. If the power of sample $\{x\}$ is large enough, and the prior distribution $p(\theta_{i^*} \mid M_{i^*})$ is of diffusion type then the posterior maximum can be replaced by the maximum likelihood estimate. The model structure, M_{i^*}, and parameters, θ_{i^*}, found this way can be considered as acceptable ones. Now consider the possibility of introducing uncertainties into the model and parameters.

Let $p_r(M)$ is prior probability for estimating model structure M, and $p(\theta_i \mid M_i)$ be conditional prior distribution for parameters θ with known structure M. According to the Bayes rule it can be written:

$$p(\theta, M \mid x) = p(\theta \mid m, x) p_r(M \mid x) =$$
$$p(\theta \mid m, x) \times c \times p_r(M) p(x \mid M),$$

where $p(x \mid M) = \int p(x \mid \theta, M) p(\theta \mid M) d\theta$; and c is normalizing constant. Thus, joint posterior density for (θ, M) is determined by the product of posterior density for θ, with condition that model M is adequate, and posterior probability of determining adequate structure M for the data $\{x\}$:

$$p(\theta, M \mid x) =$$
$$c\left(c_i p(x \mid \theta, M) p(\theta \mid M)\right) \times p(x \mid M) p_r(M),$$

where c_i is normalizing constant for $p(\theta \mid M, x)$, and c is normalizing constant for the whole posterior distribution.

Suppose it is necessary to determine the probability of bankruptcy, y, that is determined via parameters of model M. Conditional probability for y can be written as follows:

$$p(y \mid x) = \sum_i p(y \mid x, M_i) p_r(M_i \mid x),$$

where $p(y \mid x, M_i) = \int p(y \mid x, M_i, \theta_i) p(\theta_i \mid M_i, x) d\theta_i$;

$$p(\theta_i \mid M_i, x) = c_i \, p(x \mid M_i, \theta_i) \, p(\theta_i \mid M_i).$$

An asymptotic analysis of the conditional distribution $p(\theta_i \mid M_i, x)$ says [10], [11], that

$$\log p(\theta_i \mid M_i, x) =$$
$$c - \frac{1}{2}(\theta_i - \hat{\theta}_i)^T H_i (\theta_i - \hat{\theta}_i) + o(|\theta_i - \hat{\theta}_i|^2),$$

where $\hat{\theta}_i$ is weighted mean of maximum likelihood estimate for $\hat{\theta}_i$, and of the prior distribution mode for $\hat{\theta}_i$; $\mathbf{H}_i$ is a sum of hessians in corresponding maxima of likelihood function and prior density.

If the prior distribution is of diffusion type and power of sample N is large enough then parameter estimate computing can be performed using appropriate approximation. For example, matrix $\mathbf{H}_i$ can be approximated by the product, $N\hat{\mathbf{B}}_i$, where $\hat{\mathbf{B}}_i = \mathbf{B}_i(\hat{\theta}_i)$ is information matrix for one observation with condition that $\hat{\theta}_i$ is actual value for $\hat{\theta}_i$. The estimated $\hat{\theta}_i$ can also be computed with maximum likelihood technique.

Now consider posterior density for the model M_i:

$$p_r(M_i \mid x) = c p_r(M_i) p(x \mid M_i),$$

where $p(x \mid M_1), p(x \mid M_2), \ldots$ are Bayes factors approximated by the scaling constant. Somewhat simplified computing of the factors can be performed using the expression [12], [17]:

$$\log p(x \mid M_i) = c + \frac{1}{2} k_i \log 2\pi - \frac{1}{2} \log\left|\hat{I}_i\right| +$$
$$\log p(x \mid \hat{\theta}_i, M_i) + \log p(\hat{\theta}_i \mid M_i) + O(N^{-1}), \quad (1)$$

where $\hat{\mathbf{I}}_i$ is information matrix for the data $\{x\}$, that is described by the model having parameters $\hat{\theta}_i$; k_i is dimension of the parameter vector $\hat{\theta}_i$; $p(\theta_i \mid M_i)$ is a prior distribution for $\hat{\theta}_i$. As far as the observations are independent as assumed, then, $\hat{\mathbf{I}}_i = N\hat{\mathbf{B}}_i$, where $\hat{\mathbf{B}}_i$ is information matrix that corresponds to a single observation when the model M_i with parameters $\hat{\theta}_i$ is used. Thus, it can be written:

$$\log\left|\hat{\mathbf{I}}_i\right| = k_i \log N + \log\left|\hat{\mathbf{B}}_i\right|.$$

It should be noted here that with growing N the second member in the RHS will be approximately constant for each model. In a case when k_i accepts the same value for each model the only $\log\left|\hat{\mathbf{B}}_i\right|$ will be varying. An influence of the member $\log p(\hat{\theta}_i \mid M_i)$ that is linked to prior distribution is negligible, especially when the diffusion distribution is used. After introducing the notation of $\hat{l}_i = \log p(x \mid \hat{\theta}_i, M_i)$, and eliminating the member $\log p(x \mid \hat{\theta}_i, M_i)$, the criterion (1) is transformed into the following simplified form:

$$\log p(x \mid M_i) \approx c + \frac{1}{2} k_i \log 2\pi +$$
$$\hat{l}_i - \frac{1}{2} k_i \log N - \frac{1}{2} \log\left|\hat{\mathbf{B}}_i\right|,$$

or

$$\log p_r(M_i \mid x) \approx \log p_r(M_i) + \frac{1}{2} k_i \log 2\pi +$$
$$\hat{l}_i - \frac{1}{2} k_i \log N - \frac{1}{2} \log\left|\hat{\mathbf{B}}_i\right| + c,$$

Where c is a normalizing constant that provides for holding the equality: $\sum_i p_r(M_i \mid x) = 1$. Now consider examples of application of the Bayesian approach to financial risk estimation using appropriate statistical data.

Example 1. **The model of randomly incoming payments**. Let $\{x(k)\}$ is stochastic process of incoming payments, where k denotes discrete moments of time. Thus the total sum of payments for the first moment of time can be represented as $\exp(x(1))$, and for arbitrary moment k the stored capital will amount to $\exp(x(1)+x(2)+...+x(k))$. For the sake of convenience denote the sum under exponent as follows: $y(k)=\sum_{i=1}^{k}x(i)$. It is necessary to determine the types of distributions for $y(k)$, and $F(k)=\exp(y(k))$. One of the simplest models that could be hired to formally describe the financial process is low order autoregression, say AR(1):

$$x(k)=a_0+a_1x(k-1)+\varepsilon(k), \qquad (2)$$

where $\varepsilon(k)$ is stochastic process including random disturbances, measurement errors as well as model structure inadequacies. Usually the process is considered as a normal one due to growing number of measurements and multiple ingredients of the noise. Rewrite equation (2) in a more convenient for further analysis form:

$$x(k)-\mu=a(x(k-1)-\mu)+\sigma z(k), \qquad (3)$$

where μ is sample mean for the data $\{x(k)\}$; a, σ are model parameters; $\{z(k)\}\sim Norm(0,1)$ is a standard normal distribution sequence. Now find expressions for the model parameter estimates using maximum likelihood technique.

The conditional likelihood function for the sequence of payments $\{x(k)\}$ is as follows:

$$f(x\mid\mu,\sigma^2,a)=\prod_{i=-N+2}^{0}\left\{\begin{array}{l}(2\pi\sigma^2)^{-\frac{1}{2}}\times\\ \times\exp\left[-\dfrac{\left(x(i)-\mu-a(x(i-1)-\mu)\right)^2}{2\sigma^2}\right]\end{array}\right\}. \qquad (4)$$

Using the function the following expressions for model (3) parameter estimates can be written as follows:

$$\hat{a}=\frac{\sum_{i=-N+2}^{0}\left(x(i)-\hat{\mu}\right)\left(x(i-1)-\hat{\mu}\right)}{\sum_{i=-N+2}^{0}\left(x(i-1)-\hat{\mu}\right)^2};$$

$$\hat{\mu}=\frac{1}{N-1}\left(\sum_{i=-N+1}^{-1}x(i)+\frac{x(0)-x(-N+1)}{1-\hat{a}}\right);$$

$$\sigma^2=\frac{1}{N-1}\sum_{i=-N+2}^{0}\left[x(i)-\hat{\mu}-\hat{a}(x(i-1)-\hat{\mu})\right]^2.$$

Now write approximate expression for the conditional likelihood function (4) using the Tailor expansion:

$$f(x\mid\mu,\sigma^2,a)\propto$$

$$\propto(\sigma^2)^{-\frac{N-1}{2}}\exp\left\{-\frac{1}{2\sigma^2}\left[\begin{array}{l}\varphi_1+\varphi_2(\mu-\hat{\mu})^2+\\ +\varphi_3(a-\hat{a})^2+\\ +\varphi_4(\mu-\hat{\mu})(a-\hat{a})+\\ +\varphi_5(\mu-\hat{\mu})(a-\hat{a})^2+\\ +\varphi_6(\mu-\hat{\mu})^2(a-\hat{a})+\\ +\varphi_7(\mu-\hat{\mu})^2(a-\hat{a})^2\end{array}\right]\right\},$$

where "$\propto$" stands for proportionality; $\varphi_1=(N-1)\hat{\sigma}^2$; $\varphi_2=(N-1)(1-\hat{a})^2$; $\varphi_3\approx(N-1)\hat{\sigma}^2/(1-\hat{a}^2)$; $\varphi_4=2(x(-N+1)-x(0))\approx 0$; $\varphi_5=\varphi_4/(1-\hat{a})\approx 0$; $\varphi_6=-2(N-1)(1-\hat{a})$; $\varphi_7=(N-1)$.

To solve the problem of modelling risk in the problem statement given above the following prior parameter distribution is hired [13], [14]:

$$f(\mu,\sigma^2,a)=\sigma^{-3}(1-a)^{1/2}(1+a)^{-1/2},$$

$$\text{or } f(\mu,\sigma^2,a)=\sigma^{-3}(1-a)^{3/2}(1+a)^{1/2},$$

that can be used when $a\neq\pm 1$. Now the conditional posterior distribution for the process parameters can be written as follows:

$$(\mu\mid x,\sigma^2,a)\sim Norm\left(\hat{\mu},\frac{\sigma^2/(N-1)}{(1-a)^2}\right);$$

$$f(a\mid x)=d(\hat{a},N)\left[1-\hat{a}^2+(a-\hat{a})^2\right]^{-\frac{N-1}{2}}(1-a^2)^{-\frac{1}{2}},$$

$$-1<a<1.$$

Now, taking into consideration the expressions for parameter μ, σ^2, a distributions given data, $\{x(k)\}$, return to equation (3):

$$x(i)-\mu=a(x(i-1)-\mu)+\sigma z(i)=$$

$$\sigma(z(i)+az(i-1)+...+a^{s-1}z(1))+a^s(x(0)-\mu).$$

The results achieved provide the possibility for writing expression for the process $y(k)$ in the form:

$$y(k)\mid\mu,\sigma^2,a,x=$$

$$\mu k+(x(0)-\mu)M(a,k)+\sigma(V(a,k))^{\frac{1}{2}}Z,$$

where, $M(a,k)=a(1-a^k)/(1-a)$;

$$V(a,k)=\frac{1}{(1-a)^2}\left(k-\frac{2a(1-a^k)}{1-a}+\frac{a^2(1-a^{2k})}{1-a^2}\right).$$

To perform computational experiment the data regarding incoming payments for a selected insurance company was gathered within five year period. As a result the following parameter estimates were found:

$$\hat{a} = -0.2587;\ \hat{\mu} = 0.2165;\ \hat{s}^2 = 0.0873.$$

The prior probability for the model was selected equal to 0.5, what is quite logical in conditions of absence of extra information regarding the estimate. The posterior probability for the model let to the value of, $p_r(M) = 0.59$. Thus the probabilistic approach demonstrated its advantage: the parameters of forecasting posterior distribution in conditions of parametric and statistical uncertainties correspond to actual ones. This way we demonstrated how parametric uncertainty could be taken into account.

It is clear that achieved posterior probability for the model generally is not high enough what could be explained by the use of simplified initial model AR(1) and approximate computing of the conditional likelihood function. Also, it is necessary to perform additional study of influence of prior distribution on the final result. The additional deeper study of the possibility for application of probabilistic models of the type considered requires application of several alternative techniques to model building with subsequent comparison of the results achieved for each of them.

Example 2. **Quality analysis of automatized client service (operational risk analysis)**: the case of discrete data and discrete parameters. Insurance company introduced automatized service for clients that provides for automatic registration of clients insuring their cars. The number of clients, users of the service, reaches several thousand per month. Consider the problem of estimating operational error θ after servicing *n* clients.

To simplify the problem statement suppose that θ can accept the following three values: 0.25 is good result; 0.50 is acceptable result; 0.75 is bad result that cannot be accepted. Available statistic shows that within former two years the company provided the automatized client service with the following quality:

– within 60% of time the probability of service error was at the level of $\theta = 0.25$ (good result);

– within 30% of time the probability of service error was at the level of $\theta = 0.50$ (acceptable result result);

– within 10% of time the probability of service error was at the level of $\theta = 0.75$ (result that cannot be accepted).

These results were used as prior probabilities so that to forecast the level of service in the future. This distribution is given in the table 1.

Table 1. Prior probabilities for parameter θ

	Service quality		
	good	acceptable	unacceptable
Probability of service errors (θ)	0.25	0.50	0.75
Density for the errors $\theta(p(\theta))$	0.60	0.30	0.10

After 10000 cases of client services the company decided to perform the service quality control. The control showed that out of ten randomly selected service cases two of them contained errors. What conclusion regarding the service quality should be made in this case? I.e., in other words, what is actual posterior distribution for the parameter, θ?

In this case the data has discrete form and first it is necessary to determine (say, on the bases of previous experience) type of distribution for data. On the basis of former experience we can suppose that it has binomial distribution with parameter, θ:

$$f(\theta, n, r) = \binom{n}{r}\theta^r(1-\theta)^{n-r} = C_n^r\theta^r(1-\theta)^{n-r},$$

where, $\binom{n}{r} = \dfrac{n!}{r!(n-r)!}$. The number of successful events is equal in this case to $r = 2$; successful are events linked to emergence of errors in the client service out of 10 possible, i.e. $n = 10$. Thus, likelihood function for data in this case has the following form:

$$L(\theta) = \binom{n}{r}\theta^r(1-\theta)^{n-r} = C_{10}^2\theta^2(1-\theta)^{10-2},$$

where $\theta = [0.25;\ 0.5;\ 0.75]$ is distribution of possible events linked to the number of service errors. The nominator of the Bayes rule in this case is as follows:

$$h(\theta \mid r, n) \propto L(r \mid \theta, n)g(\theta).$$

Now compute necessary likelihood values, $L(\theta)$:

– if $\theta = 0.25$, then

$$L(\theta) = \frac{n!}{r!(n-r)!}(0.25)^2(1-0.25)^8 = 0.30300;$$

– if $\theta = 0.50$, then

$$L(\theta) = \frac{n!}{r!(n-r)!}(0.50)^2(1-0.50)^8 = 0.04405;$$

– if $\theta = 0.75$, then

$$L(\theta) = \frac{n!}{r!(n-r)!}(0.75)^2(1-0.75)^8 = 0.00039.$$

The prior and posterior densities for *h* are given in table 2.

Table 2. Posterior probabilities

θ	Prior probabilities for θ	Likelihood $L(\theta)$	*h* = posterior density for θ
0.25	0.60	0.30300	0.955
0.50	0.30	0.04405	0.034
0.75	0.10	0.00039	0010
Total sum:			0.999

Table 2 shows that the most probable value of θ, estimated on the basis of analysis of service quality in the case of automated servicing is the value: $\theta = 0.25$. It means that quality of service is on the "good" level (posterior probability for θ is 0.955). Before making final conclusion regarding quality of client service using automatized servicing system it is necessary to analyse larger number of cases than, 10; say 500 or 1000.

Conclusions

Thus there exists a wide sector of financial and other type of risks that require analytical studies with application of mathematical, statistical and probabilistic models of various structures and complexity. Some methods of quantitative analysis were determined that can be used for solving the problem of thorough understanding of the origin and estimating the level of possible financial loss. For estimating the level and probability of loss the probabilistic type of models can be successfully hired thanks to the fact they provide the possibility for taking into consideration parametric and statistical uncertainties of the processes under study. Here we considered a special case of Bayesian approach to formal description of financial risks and the procedure was proposed for probabilistic model constructing directed towards forecasting the incoming payments. It should be noted that Bayesian approach to risk analysis is a natural way for their formalization and receiving highly useful results of estimation and forecasting possible loss as well as its probability.

The computational experiments were performed aiming at parameter estimation of forecasting distribution. The results achieved are very close to the ones computed by the method of moments applied to actual statistical data. It was estimated that posterior probability for the model created achieved the following value: $p_r(M) = 0.59$. The achieved posterior probability for the model generally is not high enough what could be explained by the usage of simplified initial model AR(1) and approximate computing of the conditional likelihood function.

Another example of the Bayesian approach application showed that Bayes rule can be applied directly to solving the problem of analysing risky situations leading to emergence of operational risk. In this case we estimated the state of automatized servicing system and the risk of erroneous service that may result in operational loss. Thus, general conclusion regarding possible application of the Bayesian approach to risk estimation is positive what coincides with numerous other available examples of application the theory.

In the future studies it will be necessary to apply a set of alternative probabilistic candidate models aiming selection of the best one for high quality forecasting of possible loss. To estimate effective forecasting models it will be perspective to apply Monte Carlo procedures for Markov chains. In this case Markov chains will represent unknown model parameter estimates. Such numerical approach to model constructing provides a possibility for substantial increase of a number of parameters and to develop more complex model structures. Application of the Monte Carlo method will also provide the possibility for constructing simulation models that will explain away many hidden interactions of the risk factors involved.

References

[1] P.L. Bernstein, *Against the Gods: The Remarkable Story of Risk*, NY, USA, Wiley, 1996.

[2] A.J. McNeil, R. Frey and P. Embrechts, *Quantitative Risk Management.* Princeton, NJ, USA, Princeton Univ. Press, 2005.

[3] C.W. Holsapple and A.B. Winston, *Decision Support Systems.* Eagan, MN, USA: West Publishing Company, 1996.

[4] T. Aven and O. Renn, "On risk defined as an event where the outcome is uncertain", *Journal of Risk Research*, no. 12, pp. 1–11, 2009.

[5] S. Kaplan, "The words of risk analysis", *Risk Analysis*, vol. 17, pp. 407–417, 1997.
[6] N.L. Bowers, H.U. Gerber, J.C. Hickman, D.A. Jones and C.J. Nesbitt, *Actuarial Mathematics*. Itaxa, IL, USA, The Society of Actuaries, 1986.
[7] P. De Jong and G.Z. Heller, *Generalized Linear Models for Insurance Data*. Cambridge, MA, USA: Cambridge Univ. Press, 2008.
[8] Ph. Jorion, *Financial risk Manager Handbook*. NY, USA, Wiley, 2003.
[9] Sik-Yum Lee, *Structural Equation Modeling*. NY, USA, Wiley, 2007.
[10] J.M. Bernardo and A.F.M. Smith, *Bayesian Theory*. NY, USA, Wiley, 2001.
[11] A. Gelman, J.B. Carlin, H.S. Stern and D.B. Rubin, *Bayesian Data Analysis*. Boca Raton, FL, USA, Chapman and Hall/CRC, 2004.
[12] D. Draper, "Assessment and propagation of model uncertainty", *Journal of the Royal Statistical Society*, Ser. B, vol. 57, pp. 45–97, 1995.
[13] S. Klugman, *Bayesian Statistics in Actuarial Science*. Amsterdam, NL: Kluwer, 1992.
[14] P.E. Rossi, G.M. Allenby and R. McCulloch, *Bayesian Statistics and Marketing*. NY, USA: Wiley, 2005.
[15] C.P. Robert, *The Bayesian Choice,* NY, USA: Springer, 2007.
[16] N.V. Kuznietsova and P.I. Bidyuk, *Theory and Practice of Analysis Financial Risks*. Kyiv, UA: Lyra-K, 2020.
[17] M.Z. Zgurovsky, P.I. Bidyuk, O.M. Terentiev, and Prosyankina-Zharova, *Bayesian Networks in Decision Support Systems*. Charkov, UA: Edelweis, 2015.

Н.В. Кузнєцова, О.М. Трофимчук, П.І. Бідюк, О.М. Терентьєв, Л.Б. Левенчук

ЙМОВІРНІСНЕ МОДЕЛЮВАННЯ РИЗИКІВ РІЗНОЇ ПРИРОДИ

Проблематика. Усім видам людської діяльності притаманні певні ризики, зокрема фінансові. Відповідно, існує проблема оцінювання і прогнозування можливих втрат, для розв'язання якої необхідно створити адекватний математичний опис для формального представлення обраних ризиків. Оцінювання можливих втрат може ґрунтуватись на обробленні наявних даних та експертних оцінок, що характеризують історію та поточний стан процесів, які аналізуються. Належний інструментарій для моделювання та оцінювання ризиків можливих втрат забезпечується використанням ймовірнісного підходу, який включає баєсові методи, відомі на сьогодні як методологія баєсового програмування.

Мета дослідження. Зробити огляд деяких методів баєсового аналізу даних, які забезпечують можливість побудови моделей вибраних ризиків. Зокрема, використати статистичні дані для формального опису операційного ризику, що може з'явитись у процедурах обробки інформації.

Методика реалізації. Для обробки даних і побудови моделей використовується методологія баєсового програмування. Для оцінювання операційного ризику також застосовано теорему Баєса у формулюванні для дискретних подій та дискретних параметрів.

Результати дослідження. На основі запропонованого підходу побудовано модель операційного ризику, пов'язаного з некоректною обробкою інформації. Для того, щоб побудувати і застосувати модель для оцінювання ризику, проаналізовано задачу оцінювання ризику, вибрано змінні та оцінено умовні апріорні ймовірності. Функціонування побудованих моделей продемонстровано на ілюстративних прикладах.

Висновки. Практично важлива задача моделювання та оцінювання ризиків різних типів, зокрема фінансових, може бути вирішена методами баєсового програмування, які дають можливість ідентифікувати і врахувати невизначеності даних та експертних оцінок. Модель ризику, побудована за допомогою запропонованого методу, ілюструє можливості застосування баєсових методів для розв'язання задачі оцінювання ризиків.

Ключові слова: фінансові процеси; фінансові ризики; методологія баєсового програмування; оцінювання ризику.

Рекомендована Радою
Навчально-наукового інституту прикладного системного аналізу
КПІ ім. Ігоря Сікорського

Надійшла до редакції
18 листопада 2020

Прийнята до публікації
14 лютого 2022 року