Bayesian Data Analysis in Modeling and Forecasting Nonlinear Nonstationary Financial and Economic Processes

Oleksandr Trofymchuk, Petro Bidyuk,
Tatyana Prosiankina-Zharova and Oleksandr Terentiev

November 28, 2022

УДК 519.246

BAYESIAN DATA ANALYSIS IN MODELING AND FORECASTING NONLINEAR NONSTATIONARY FINANCIAL AND ECONOMIC PROCESSES

Trofymchuk[1] O.M., Bidyuk[1] P.I., Proyankina-Zharova[1] T.I., Terentiev[1] O.M.
[1]Institute of Telecommunications and Global Information Space
E-mail: Trofymchuk@nas.gov.ua, pbidyuke_00@ukr.net, t.pruman@gmail.com, o.terentiev@gmail.com

Abstract

A short review of modern Bayesian methods for data analysis is presented that shows wide applicability of the methods mentioned in many areas of practical human activities. Some application details are provided for generalized linear models, that are popular in analysis of NNP, to highlight their possibilities and specific features. All Bayesian techniques of data analysis are very popular today thanks to their flexibility, high quality of final results, availability of possibilities for adaptation to new data and conditions of functioning. Besides, each of these approaches to data analysis is well supported by appropriate sets of statistical criteria that make possible thorough quality analysis of intermediate and final results. Illustrative example is presented of GLM application providing an insight into some possibilities and special features of the methods. Some considerations are provided regarding development and practical application of specialized intellectual decision support system directed to refinement and quality enhancement of computational results and providing intellectual help to users.

Key words: nonlinear nonstationary processes, Bayesian methods, modeling, forecasting**,** Generalized Linear Models

Introduction

Many studies today are related to modeling and forecasting evolution of processes in various areas; they are mostly touching upon widely spread nonlinear non-stationary processes (NNP). To be more exact, definition of NNP means that such processes exhibit at least one type on non-stationarity (regarding trend or integration, and variance or heteroscedasticity) as well as nonlinearity regarding variables or model parameters. Such processes create majority in ecology, economy, finances, industrial technologies, engineering systems, hydrology, climate studies, in the problems of technical, medical and economic diagnostics, physical experiments of various type etc. For example, many processes in economy show availability of low (first or second) order trend, but transition to second order of integration automatically shifts the process from the class of linear to the nonlinear ones because quadratic and higher order dependence indicates relation to the class of non-linear characteristics.

When we talk about process analysis we mean, as a rule, solving first of all the two most often met problems such as constructing the selected process model and forecasting. The widely used models of NNPs include, at least, the following types: differential equations, nonlinear regression, combination of linear and nonlinear regression, nonlinear autoregression, semi- and non-parametric methods, kernel based models, vector parametric models and methods, vector semi- and non-parametric approaches, state space and frequency-domain models, generalized linear models (GLM), and some others. Today high popularity acquired the methods of intellectual data analysis such as neural networks, Bayesian networks, decision trees and forests, immune, genetic and molecular algorithms. A separate subclass of nonlinear models belongs to models that are nonlinear in parameters such as logit and probit (logistic regression), ecological function, irrational and hyperbolic functions, Tornquist functions, and many others [1, 2]. The models nonlinear in parameters require application of special nonlinear estimation techniques for parameter estimation such as nonlinear LS, maximum likelihood (ML), Monte Carlo for Markov Chains (MCMC) and others.

One of the most widely met in practice subclass of nonlinear processes (and their models) is created by heteroscedastic processes, for example, they are very often considered in financial analysis. This direction of modeling and forecasting is widely known and practically used due to the fact that variance is considered in this case as dynamic variable which is described by dynamic model and its value is predicted to solving many problems. For example, predicted standard deviation (volatility) and variance itself are used for financial risk estimation and short-term forecasting of process volatility. Existing today variety of variance models and their estimation techniques is very high due to wide possibilities for their practical applications [3, 4].

Very popular approach to modeling NNP today is based upon Bayesian data analysis. Besides well-known static and dynamic Bayesian networks in includes Bayesian regression, structural equation models, probabilistic filtering techniques, complex distribution analysis etc. The Bayesian approach to analysis of data and expert estimates has such positive features as structural flexibility of the models, possibility for taking into consideration uncertainties that are available practically in every area and case of studies, availability of model parameter estimation procedures for the cases of linear and nonlinear modeling, forecasting of complex distributions etc. [5 – 8]. In this study we consider uncertainties as the factors of negative influence to the computational procedures used for model structure and parameter estimation, forecasts and control actions computing etc. Influence of the factors results in lower quality of intermediate and final results of data analysis, i.e. model adequacy, forecast estimates, control actions and decision alternatives.

The purpose of this study is to provide a review of Bayesian methods for modeling NNP, illustrate some applications, and highlight the methods that could be used for reaching high quality results of data analysis. For example, it is often useful to stress development and application of a specialized intellectual decision support system (IDSS) and apply it to solving specific complicated problem.

Problem statement

The main goals of the study are as follows: (1) to provide a review of Bayesian data processing and model constructing methods for their further use in intellectual decision support system for modeling and forecasting nonlinear non-stationary processes in economy and finances; (2) to present illustrative examples of Bayesian techniques application to solving the problems mentioned; (3) to stress the necessity of development intellectual DSS for high quality solving the problems.

Bayesian methods for modeling and forecasting

Today there exist a wide set of Bayesian methods that are often used for preliminary data processing, model constructing, forecasting of future processes evolution, risk estimation, control in various spheres, decision support, classification and solving some other practical problems. Among others, the following Bayesian methods and techniques are actively used in practice, and should be mentioned [7 – 13]:

- generalized linear models (GLM); the set of exponential distribution laws used in the case of GLM application are as follows: normal, Poisson, binomial, Gamma, inverse Gaussian; such approach enhances substantially number of process that can be formally described with GLM;
- structural equation models (SEM); the models of this type make it possible constructing mathematical models for another class of random process that exhibit specific structure that can be adequately described by Bayesian techniques;
- static and dynamic Bayesian networks (BN); BN represent powerful probabilistic instrumentation capable formally describe sophisticated stochastic process and expert estimates, and generate probabilistic inference;
- Bayesian filtering of data and recursive parameter estimation; filtering touches upon the problem of data processing, i.e. reducing influence of noise components spoiling collected observation; parameter estimation relates to parameters of distributions and various models, first of all regression models we have to construct in multiple applications;
- Bayesian maps; trajectory synthesis and control of robotic systems;
- Markov localization models; the problem of robot localization and control is considered;
- multivariate distribution constructing and analysis; forecasting multivariate distributions, parameter estimation;
- decision trees and forests;
- combining Bayesian and statistical techniques into single model; model combining provides a possibility for modeling and forecasting complicated NN processes.

The methods listed above are distinguished with their high flexibility and possibilities for taking into consideration possible data uncertainties and generating alternative final results directed to support of decisions according to specific problem statement. Consider details of some methods.

Generalized Linear Models

Generalized Linear Models (GLM) is a class of regression models that allows for taking into consideration interaction between model factors, specific distribution law of dependent variable and possible nonlinearity [12 – 13]. GLM consists of the three basic components: systematic, stochastic, link function, and can be formally represented as follows:

$$\mu_i = \mathrm{E}[\mathbf{y}_i] = \mathrm{g}^{-1}\left(\sum_j \mathbf{X}_{ij}\beta_j + \xi_i\right), \qquad (1)$$

$$Var[\mathbf{y}_i] = \frac{\phi V(\mu_i)}{\omega_i}$$

where $\mathbf{y}_i$ is a vector of observations for dependent variable; $g(x)$ is a link function; $\mathbf{X}_{ij}$ is a matrix of observations for a model factors;; β_j is a vector of parameters estimated on factor observations; ξ_i is a vector of stochastic residuals; ϕ is a vector of scale parameters for the function of $V(x)$; ω_i – are prior weights of confidence level. Thus, GLM is characterized by the following elements: distribution law for dependent variable, Y; parameters and specific features of a link function $g(\cdot)$; features of the linear predictor, $\eta = \mathbf{X}\beta$.

Usually it is reasonable to make the following suggestions regarding GLM:

– all the components of dependent variable, Y, are independent, and their distribution law belongs to the family of exponential distributions;

– the suggestion regarding systematic feature of a model is treated as follows: p predictors are combined into single "linear predictor" η;

– the suggestion regarding link function: mutual dependence between suggestions of stochastic and systematic features is expressed by the link function that is supposed to be differentiable and monotonic and has an inverse:

$$\mathrm{E}[\mathbf{y}] = \mu = g^{-1}(\eta). \qquad (2)$$

The set of possible distribution laws for GLM and parameters of the distributions for dependent variable are presented in Table 1.

Table 1. The distribution laws and their parameters

Compo-nent	Normal distribution	Poisson	Binomial	Gamma-distribution	Inverse Gaussian
	$N(\mu,\sigma^2)$	$P(\mu)$	$\frac{B(n,\pi)}{n}$	$G(\mu,\nu)$	$IG(\mu,\sigma^2)$
Variance param., ϕ	$\phi=\sigma^2$	1	$\frac{1}{n}$	$\phi=\nu^{-1}$	$\phi=\sigma^2$
Cumulant function,	$\frac{\theta^2}{2}$	$\exp(\theta)$	$\log(1+e^\theta)$	$-\log(-\theta)$	$-(-2\theta)^{1/2}$

b(θ)					
c(y, ϕ)	$-\frac{1}{2}\left\{\frac{y^2}{\sigma^2}+\log(2\pi\sigma^2)\right\}$	$-\log(y!)$	$\log\frac{n}{y}$	$\nu\log(\nu y)-\log(y)-\log(\Gamma(\nu))$	$-\frac{1}{2}\left\{\log(2\pi\phi y^3)+\frac{1}{\phi y}\right\}$
μ(θ)= E(Y,θ)	θ	$\exp(\theta)$	$\frac{e^\theta}{1+e^\theta}$	$-\frac{1}{\theta}$	$(-2\theta)^{1/2}$
Canonic link, θ(μ)	identity	log	logit	reciprocal	$\frac{1}{\mu^2}$
Variance func, V(μ)	1	M	$\mu(1-\mu)$	μ^2	μ^3
Var (μ)	σ^2	$n\mu(1-\mu)$	μ	$\frac{\mu^2}{\nu}$	$\frac{\sigma^2}{\mu^3}$

The link function links linear predictor, η, to the estimate, μ, related to *Y*. In a classic linear model mean the value of dependent variable and linear predictor are identical, and identity link (η and μ) are selected arbitrarily but from the set of real numbers. GLM is distinguished from linear model of general form (special case of such model is multiple regression) by the following:

- distribution of dependent variable (system reaction) can be non-Gaussian and not necessarily continuous, for example, binomial;
- dependent variable predictions are computed as linear combination of predictors that are linked to the dependent variable with a link function.

Dependently on the distribution law of dependent variable and type of the link function there exist types of GLM presented in Table 2.

Table 2. Types of GLM

Model type	Link function	Dependent variable distribution
1. Linear model of general type	$g(\mu)=\mu$	Normal distribution
2. log-linear model	$g(\mu)=\ln(\mu)$	Poisson distribution
3. Logit-model	$g(\mu)=\ln\left(\mu/(1-\mu)\right)$	Binomial distribution
4. Probit-model	$g(\mu)=\Phi^{-1}\mu$	Binomial distribution
5. “Survival” analysis	$g(\mu)=\mu^{-1}$	Gamma distribution, Exponential distribution

Thus, GLM represents quite general class of statistical models that includes linear regression, variance and covariance analysis, Log-linear models for analysis of random tables, logit/probit models, Poisson regression and many others.

Each distribution has its specific link function for each exists substantiated equality statistic regarding parameter vector β of linear predictor, $\eta = \sum x_j \beta_j$. Such canonic link comes to being when, $\theta = \eta$, where θ is canonic parameter which is determined when likelihood function is introduced. Specific distribution type corresponding to each link function is given in Table 2, and generalized form of distribution is given below:

$$f(y, \theta, \phi) = \exp\left[\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right], \tag{3}$$

where *a*, *b*, *c* – are functions that correspond to specific distribution law; y is dependent variable; θ is canonic parameter or a function of some parameter of a specific distribution; ϕ is variance parameter.

Function, $b(\cdot)$ has a special meaning in generalized linear models because it described relation between mean value and variance in a distribution. When ϕ is known then we have exponential model with canonic parameter θ. When ϕ – is unknown, then exponential distribution can be of two-parameter type. Thus, canonic link for a set of exponential distributions has the form presented in Table 3.

Table 3. Canonical forms for different distributions

№	Distribution type	Canonic link
1	Normal distribution	$\eta = \mu$
2	Poisson	$\eta = \log(\mu)$
3	Binomial distribution	$n = \log\left\{\frac{\pi}{1-\pi}\right\}$
4	Gamma distribution	$\eta = \mu^{-1}$
5	Inverse Gaussian distribution	$\eta = \mu^{-2}$

A substantiated statistic for canonical forms is the vector, $\mathbf{X}^T \mathbf{y}$:

$$\sum_i x_{ij} y_i, \; j = 1 \ldots p, \tag{4}$$

It should be noted that canonical link results in desirable features of a model, especially for short samples, for example existence of systematic effects. And here are approaching the problem of selecting statistic criteria for model estimating. As a result of analysis of various criteria for model selection the conclusion was made that Akaike information criterion fits better to solving the problem together with the maximum likelihood approach to estimation.

Quality estimation for generalized linear models

Usually in the process of analyzing quality (adequacy) of a model several different criteria are applied. In some cases these criteria coincide with the same statistical criteria used in constructing linear and/or nonlinear regression. When we use the models on the purpose of classification, among them are the following: (1) common accuracy of a model; I-st and II-nd type errors; ROC-curve, and Gini index.

Common accuracy (CA) statistic is determined by the expression:

$$CA = \frac{Correct\ \ Forecast}{N}, \quad (5)$$

where, *Correct Forecast* is a number of correctly forecasted cases (examples); *N* is a total number of cases under investigation. This criterion is somewhat subjective measure of a model quality because it depends on a number of defaults in the dataset, and the level of the cut-off threshold. Different levels of the threshold result in different values of common accuracy.

ROC-curve (Receiver Operation Characteristic) shows dependence of a number of correctly classified positive examples on the number of incorrectly classified negative examples. The first set of examples is called true positive, and the second set are referred to as false positive ones. Here it is suggested that the classifier has some parameter that can be varied to reach necessary division into classes. This parameter is called cut-off point or threshold. Depending on the parameter value different values of the I-st and II-nd type errors will be achieved.

Most often the following statistics (in percentages) are used for determining quality of a model:

- a part of true positive examples (True Positives Rate):

$$TPR = \frac{TP}{TP + FN};$$

- a part of false positive examples (False Positives Rate):

$$FPR = \frac{FP}{TN + FP}.$$

Usually for completeness of the analysis two more characteristics are applied: sensitivity and specificity.

Model sensitivity is defined as a part of true positive cases:

$$Se = TPR = \frac{TP}{TP + FN}. \quad (6)$$

Model specificity is defined as a part of true negative cases that were correctly classified by a model:

$$Sp = \frac{TN}{TN + FP}. \quad (7)$$

Now, it is obvious that:

$$Sp = \frac{TN - FP + FP}{TN + FP} = 1 - \frac{FP}{TN + FP} = 1 - FPR. \quad (8)$$

The model that exhibits high sensitivity provides for a true result when number of positive cases is also high (i.e., it reveals well positive cases). And the model that exhibits high specificity usually provides for better (true) results when the number of negative cases is high (i.e., the model discovers better negative cases). The ROC curve (or Lorentz curve) graph uses the axes, Y, for sensitivity, *Se*, and axis, X, for a part of false positive cases, *FPR*, or (1-*Sp*).

The graph of ROC-curve for an ideal classifier tends to the left upper corner where part of true positive cases tends to 1 (the case of ideal sensitivity), and the part of false positive samples tends to zero. Thus, the closer approaches ROC-curve to the left upper corner the better is the model regarding prediction of a true value. As a consequence, straight diagonal line corresponds to the classifier that is unable to distinguish between the two classes. As far as visual comparison of different ROC-curves not always allows selection of better model, there is a numerical criterion in the form of area under curve (AUC) that makes the model selection easier. It is computed using the trapeze method as follows:

$$AUC = \int f(x)dx = \sum_i \left[\frac{X_{i+1} + X_i}{2} \right] \cdot (Y_{i+1} - Y_i). \quad (9)$$

As alternative model quality measure is used Gini index that is computed as an area between diagonal and Lorentz curve divided by the whole area under the diagonal. The index is widely used for resolution analysis of a system developed for credit risk estimation. In this case the model is used for dividing the clients into two groups: those who are inclined to default, and those who are not. Usually, probability is hired as a measure of inclination to default.

Application of Bayesian approach to estimation

In most cases of solving the problems of process modeling, forecasting and decision support based on statistical data we meet uncertainties of various types. As an example can serve structural, parametric and statistical uncertainties available in development and practical application of identification procedures for analysis of processes of various nature. Structural uncertainties are linked to the uncertainties of developed model structure, and parametric are referred to the model parameter estimates. Statistical uncertainties are mostly related to observations, for example to the difficulties of determining true data distribution, influence of external random factors corrupting the measurements, missing observations, extreme values etc. Theoretical studies of such problems are mostly related to analysis of reasons for emergence, classification and influence estimation of the uncertainties as well as level and probabilities of respective risks.

Very often we lack statistical data to solve the problems of risk analysis and decision making. Such problems cannot be solved with traditional statistical frequency approach because the data available cannot provide necessary information. Moreover, the situations related to decision making may change substantially and lack the results of preliminary analysis. Such particular features lead to complications with decision making and may influence negatively final results. That is why the Bayesian approach becomes in such cases useful and highly effective instrument of modeling, forecasting and decision support.

A distinctive feature of Bayesian approach to data and expert estimates processing is that researcher considers the level of his belief to possible models and forecasts before receiving data and represents his view in the form of a probability. As soon as the necessary data is received the Bayes theorem provides a possibility for computing another set of probabilities representing with them refined beliefs to the candidate models. Fig. 1 shows the process of data analysis and decision making that is at the core of Bayesian approach.

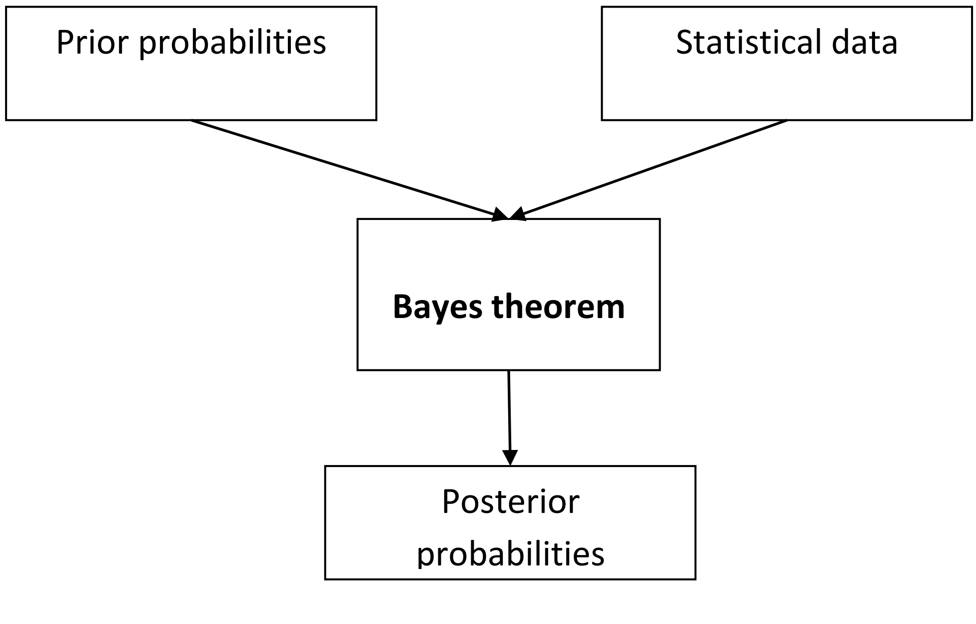

Fig. 1 The basic Bayesian paradigm

The key advantage of Bayesian approach to data expert estimate analysis is in the possibility of using any prior information related to system under study state and its model (for example, model structure and its parameters). Such information is presented in the form of prior probabilities and/or density distribution. The prior probabilities are recalculated (improved) further on using the data that are used to determine posterior distribution of parameter estimates or respective output variables. Consider as example application of conjugate distributions what means that prior and posterior distributions are of the same type.

Let experimental data, $X_1,\ldots,X_n$, is normal random sample with mean value, μ, and variance, σ^2. Suggest that prior distribution is also normal with mean, μ_0, and variance, $\sigma_0{}^2$. Then posterior distribution for, μ, with known sample, $X_1,\ldots,X_n$, and known prior distribution will also be normal with mean, μ_*, and variance, $\sigma_*{}^2$, that are computed as follows:

$$\mu_* = \frac{\sigma^2\mu_0 + n\sigma_0{}^2\,\overline{X}}{\sigma^2 + n\sigma_0{}^2}, \quad \sigma_*{}^2 = \frac{\sigma^2\,\sigma_0{}^2}{\sigma^2 + n\sigma_0{}^2},$$

where, $\overline{X} = \sum_{i-1}^{n} X_i / n$, is sample mean. Bayesian analysis often uses parameter characterizing quality of results, and determined as inverse to variance: $\eta = 1/\sigma^2$. Thus, we have for the prior distribution, $\eta_0 = 1/\sigma_0{}^2$, and for posterior distribution, $\eta_* = 1/\sigma_*{}^2$. Now the expressions for posterior parameter will take the form:

$$\eta_* = \eta_0 + n\eta\ , \ \ \mu_* = \frac{\eta_0}{\eta_*}\mu_0 + \frac{n\eta}{\eta_*}\overline{X}.$$

Thus, for the case of normal sample information about μ is contained in the sample mean, $\overline{X}$, that is complete statistic for, μ. The quality of distribution identification is determined by the relation: $X :\ \eta/\sigma^2 = n\eta$. Quality of posterior distribution description is determined by the quality of two following components: prior distribution representation, and statistical characteristics of sample data. The posterior mean is weighted prior mean and sample mean with weighting coefficient that is proportional to the quality parameter. It means that influence of prior distribution is decreasing with growing size of data sample, *n*.

Forecasting with the use of Bayesian computations

Consider the problem of forecasting random value, *x*, on the basis of historical observations, y, using Bayesian approach. That is, it is necessary to determine type of distribution for the future values of, *x*, given values of, y. From the probabilistic point of view the problem is in determining forecast density, $\pi(x \mid y)$, that describes possible changes with known values of, y.

Most often, the model for which the forecast density is needed does not exist. But known is the probabilistic model for, *x,* that is expressed in terms of distribution, $g(x \mid \theta)$, depending on unknown parameter, θ, that is based on the model describing observations, y. If we denote posterior density of, θ, given, y, as $P(\theta \mid y)$, then forecasting density for, *x*, can be written as follows:

$$\pi(x \mid y) = \int_{\Theta} g(x \mid \theta) P(\theta \mid y)\, d\theta .$$

If a point forecast is required, it is enough to use the point estimate of, $\pi(x \mid y)$. To find an interval forecast it is enough to compute interval estimates for, $\pi(x \mid y)$.

Example: forecasting actuarial risk in car insurance using GLM

As an example of actuarial process modeled by GLM the problem of estimating (forecasting) financial loss in car insurance had been selected. The experimental data includes one dependent variable, "*Loss*", reflecting paid volume of insurance among the cars of the following brands: VAZ, Mitsubishi, and Toyota. The regions of selling (distribution) of the car policies include the cities of Kyiv, Donetsk, and Odesa. The data relates to the years starting from 2006 with the sample size of 9546 examples. The results of GLM constructing for different distribution laws are presented in Tables 4 – 6.

Table 4. Distributions and link functions used

№	Model	
	Distribution of dependent variable	Link function
1.	Gamma	LOG
2.	Normal	LOG
3.	Poisson	LOG
4.	Normal	Identity

Table 5. Numerical characteristics of the models constructed

Total loss	Mean	Std. deviance	Min	Max	Mean of standard error	Variance, %

102008320.905	**11805.690**	**15358.118**	**6273.867**	**18549.819**	**0.075**	**130.091**
18111231.380	1897.457	939.910	4010.978	634.054	0.120	49.535
17921032.574	**1877.531**	**1027.567**	**4234.951**	**558.354**	**0.176**	**54.730**
17921032.589	1877.531	999.302	3535.396	118.004	224.348	53.224

Table 6. Statistical characteristics of the models

Total loss	Log-likelihood	Actual total loss	Difference	Risk
102008320.905	**-15742.754**	17921032.581	**84087288.32**	**1.301**
18111231.380	-98700.167		190198.799	0.495
17921032.574	**-42173677.24**		**0.007**	**0.547**
17921032.589	-98700.167		0.009	0.532

Table 6 shows that the risk of financial loss for the models constructed varies approximately between 40-60% what is marginally acceptable but requires undertaking of some measures regarding its minimization.

Discussion

Comparison of the models hiring normal distribution with logarithmic and identity link functions leads to results that information Akaike criterion accepts approximately the same value of about 20.78. It means that better model selecting should be based upon the total forecast of possible loss. Thus, adequate model in this study is the one based upon Poisson distribution and exponential link function. This choice is supported by the minimum loss estimation error, maximum approach of forecast to actual data, and true estimate of possible risk of loss. Normal model with identity link function provides for a small loss error of about 1.65% but "weak" forecasts of loss and erroneous risk estimate.

An important issue of the study is development and implementation of a system that would provide a substantial help in solving the tasks of selected process model building, forecast estimation, alternatives generation, and selection of the best one of them on the purpose of its further practical implementation. Usually this is specialized intellectual decision support system (DSS) containing all necessary computational procedures and sets of statistical criteria to estimate quality of intermediate and final results of process analysis. Thus, the main ideas related to constructing and practical application of the DSS are as follows [14]: (1) to get handy and useful instrument for carrying out necessary computations and get results within short period of time; (2) to test immediately quality of computations using various sets of quality analysis criteria in automatic mode of system operation; (3) to identify possible uncertainties and take countermeasures to reduce their negative influence on the computations; (4) to calculate possible alternatives and select the best one of them for further practical implementation. Availability of intellectual features in the system simplifies the tasks to be performed by the system user and provides substantial help for him in form of advices, extra information on the methods to be used for model structure and parameter estimation, statistical quality criteria to be applied in specific cases, on the quality of intermediate and final results of computations and some other features.

In the case of analysis statistical/experimental data possible uncertainties are touching upon data itself. They can be encountered in the form of state and measurement noise, missing observations, short samples, and multicollinearity. All these problems can be solved using appropriate techniques mentioned earlier.

Conclusions

Thus, a short review of modern data processing Bayesian techniques was presented. An example of modeling and forecasting of financial actuarial process is presented. It was shown that generalized linear models can serve as effective instrument of analysis financial and economic processes that helps to take into consideration actual complicated factor interactions and their influence on dependent variable. There also exists a possibility for model and forecasts quality analysis (of the results achieved) using a set of appropriate statistical criteria.

Further research in this direction should be focused on the problems touching upon the following issues: refinement of the forecasting model; more profound analysis of factors influencing dependent variable; more active use of Bayesian and neural networks and other methods of intellectual data analysis to modeling and forecasting actuarial processes; development and use of commercial intellectual decision support system. The DSS will help to construct and study more sophisticated combined model consisting of linear and nonlinear parts, to reach higher quality forecasts of dependent variable, and consequently improve estimates of possible financial loss. This approach to financial processes analysis will help to minimize financial risks in insurance as well as in many other spheres of human activities. Finally such studies will positively influence macro economy as a whole.

References

1. De Gooijer, J. G. (2017) *Elements of Nonlinear Time Series Analysis and Forecasting*. – Cham, Switzerland: Springer, 618 p.
2. Bidyuk, P., Romanenko V.D., Tymoshchuk O.L. (2012) *Time Series Analysis*. – Kyiv: Polytechnika Publisher at the National Technical University of Ukraine “Igor Sikorsky KPI”, 600 p.

3. Tsay R.S. (2010) Analysis of Financial Time Series. – Hoboken (NJ, USA): Wiley & Sons, Inc., 715 p.
4. Hansen B.E. (2017) Econometrics. – University of Wisconsin, 427 p.
5. Rossi P.E., Allenby G.M., McCulloch R. (2005) Bayesian Statistics and Marketing. – Hoboken (New Jersey): John Wiley & Sons, Ltd, 348 p.
6. Press S.J. (2003) *Subjective and Objective Bayesian Statistics*: *Principles, Models, Applications*. – Hoboken (New Jersey): John Wiley & Sons, Inc., 558 p.
7. Bidyuk P., Terentyev A., Konovalyuk M., (2010) Bayesian networks in data mining technologies // *Artificial Intelligence*, № 2, pp. 104 – 113. Available at: <http://dspace.nbuv.gov.ua/bitstream/handle/123456789/56141/13-Bidyuk.pdf?sequence=1>.
8. Zgurovsky M.Z., Bidyuk P.I., Terentiev O.M., Prosyankina-Zharova T.I. (2014) *Bayesian Networks in Decision Support Systems*. – Kyiv: NTUU KPI, 320 p.
9. Jensen F.V. (2001) *Bayesian networks and Decision Graphs*. – New York: Springer-Verlag, 268 p.
10. Lee Sik-Yum. (2007) *Structural Equation Modeling*. – Chichester (England): John Wiley & Sons Ltd, 432 p.
11. Chen Ming-Hui, Shao Qi-Man, Ibrahim J.G. (2000) *Monte Carlo Methods in Bayesian Computation*. – New York: Springer-Verlag, 386 p.
12. De Jong P., Heller G.Z. (2008) Generalized Linear Models for Insurance Data. – New York: Cambridge University Press, 197 p.
13. McCullagh P., Nelder J. A. (1989) *Generalized Linear Models*. New York: Chapman & Hall, 526 p.
14. Sugumaran V. (2002) *Intelligent Support Systems Technology: Knowledge Management*. – London: IRM Press, 318 p.
15. SAS Institute Inc. SAS/OR 14.2 User's Guide: Mathematical Programming. Cary, NC: SAS Institute Inc. SAS/OR 14.2 User's Guide: Mathematical Programming.2016. Available online: http://support.sas.com/thirdpartylicenses
16. Kuznietsova N., Bidyuk P. (2020) Intelligence information technologies for financial data processing in risk management. *Communications in Computer and Information Science* №1158. P. 539–558.
17. Trofymchuk O., et al. (2019) Decision Support Systems for Modelling, Forecasting and Risk Estimation, LAP Lambert Academic Publishing, Riga. 160 p.
18. Davidson R., MacKinnon J. G. (2004). Econometric Theory and Methods, Oxford University Press, Offord.