FACTA UNIVERSITATIS
Series: **Economics and Organization** Vol. 12, No 4, 2015, pp. 323 - 332

EXAMINATION OF THE PERFORMANCES OF MAXIMUM LIKELIHOOD METHOD AND BAYESIAN APPROACH IN ESTIMATING SALES LEVEL

UDC 31:658.8

Nataša Papić-Blagojević[1], Vinko Lepojević[2], Sanja Lončar[3]

[1]Novi Sad Business School, Novi Sad, Serbia
[2]University of Niš, Faculty of Economics, Niš, Serbia
[3] Novi Sad Business School, Novi Sad, Serbia

Abstract. *The method of maximum likelihood and Bayesian method are widely used in data processing, not only in economics but also in other fields of research. In order to identify which approach has better performances, these methods are analyzed on the selected economic data. By comparing the estimated values obtained by applying the maximum likelihood method and Bayesian method on the data that was obtained from the company CaliVita Int., it was concluded that the Bayesian inference with informative priors gives more accurate estimates.*

Key words: *method of maximum likelihood, Bayesian method, estimation, prior, sales level*

INTRODUCTION

Historically, classical statistics have had the main role in researches compared to Bayesian methods, but in the near past, Bayesian approach has become very popular in creating statistical models for solving problems in different research fields. Bayesian inference is known as an analytical method that combines information obtained from the experiment with prior knowledge. Contrary to the maximum-likelihood approach, the Bayesian framework requires the explicit prescription of a prior probability distribution for the unknown signal parameters [15].

On the other side, the method of maximum likelihood, as a method of classical statistics, does not include any prior information that maybe exists from previous research. Very often, researchers are faced with the problem that is common for one data set and they have to estimate parameters in the moment of taking data.

Received December 10, 2015 / Accepted December 29, 2015
Corresponding author: Nataša Papić-Blagojević
Novi Sad Business School, Novi Sad, Serbia
E-mail: npapic.blagojevic@gmail.com

Many supporters of classical statistics do not accept the use of subjective prior information in objective economic science. The debate about the role of prior information in statistics is still going on. Actually, prior information is a controversial aspect of Bayesian methods, but Bayesians, as the final line of defence, have developed non-informative priors for many classes of model [9]. In order to get conclusions about the importance of prior information, the results of research with and without presence of prior information are compared in this paper.

1. LITERATURE REVIEW

Even though Bayesian methods are powerful and can be used in a variety of analytic models, the strenuous programming and computational demands have discouraged many researchers from using them [19]. Today, the use of Bayesian methods in empirical researches is rapidly growing, because there is a lot of appropriate statistical software that can be successfully applied in different situations. According to the complexity, the models are very different, so the use of some is very simple and allows the researcher to easily reach the desired estimates and test statistics, while others require researchers to possess programming skills.

The maximum likelihood method has broad and significant application in determining the statistical estimations with good characteristics. Its application, as well as application of Bayesian methods, is not only limited to economics, but it can be successfully implemented in psychology, medicine, biology, tourism, etc.

Lemmon, Brown, Stanger-Hall and Lemmon [11] study the effect of ambiguous data, or missing values for research in biology using the method of maximum likelihood and Bayesian method. Ward [18] in his paper compared Bayesian and classical methods for estimation of ecological models, where the maximum likelihood criteria consistently favoured simpler population models when compared to Bayesian criteria.

On the other hand, Flurry and Shepard [6] studied the wide application of Bayesian inference and likelihood methods in microeconomics, macroeconomics and financial econometrics. In doing so, they illustrate these methods on four problems in econometrics, producing rather generic methods. Taken together, these methods imply that if we can simulate from an economic model we can carry out likelihood based inference using its simulations [6].

Pitt, Silva, Giordani and Kohn [13] are concerned with developing a methodology for Bayesian inference for general time series state space models using Markov chain Monte Carlo (MCMC) simulation with the likelihood estimation. Fernández-Villaverde and Rubio-Ramírez [5] showed how to undertake likelihood-based inference in dynamic macroeconomic models. They also describe how to use the output to estimate the structural parameters of the model, those characterizing preferences and technology, and to compare different economies. Both tasks can be implemented from either a classical or a Bayesian perspective.

Bayesian estimation may also be used for solving some problems that are commonly encountered in traditional statistics; for example, obtaining estimations for impossible parameters, identification of the model [8], and obtaining more precise parameters estimations [3]. Also, Bayesian methods are more plausible ways to analyze small sample data compared with the maximum likelihood method [19]. In this paper, the comparison of the maximum likelihood and Bayesian method is made on the data that was obtained from the company CaliVita Int. Calculations were made for two types of prior information that can be applied in Bayesian inference, informative and non-informative, and the results are compared with maximum likelihood estimates.

2. Research Methodology and Results

In order to make a comparison of the methods of classical and Bayesian statistics, the data were obtained from the company CaliVita Int. representative for the Republic of Serbia, Fitco LLC, Novi Sad. The selected data related to 252 products [12]. For research purposes, the products are sorted by: product name, product type, manufacturer, price and realized sales in the period from January to June 2014 and processed in statistical packages IBM SPSS and IBM SPSS Amos *Version 21*. IBM SPSS Amos applies a general approach to data analysis known as Structural Equation Modeling –SEM. It is also known as the analysis of covariance structures or causal modeling.

A sample of 252 products is considered as sufficient for successful research tests because Bayesian statistics is not based on large samples [17]. Many articles also show the advantages of Bayesian statistics in terms of a small data set [19]; [10].

To perform the analysis, four variables were selected: *Product type*, *Manufacturer*, *Sales level in May* and *Sales level in June*. The variables *Sales level in May* and *Sales level in June* are defined as observed, endogenous variables. They are conditioned with two observed, exogenous variables: the *Product type* and the *Manufacturer*. Figure 1 shows the model for selected data.

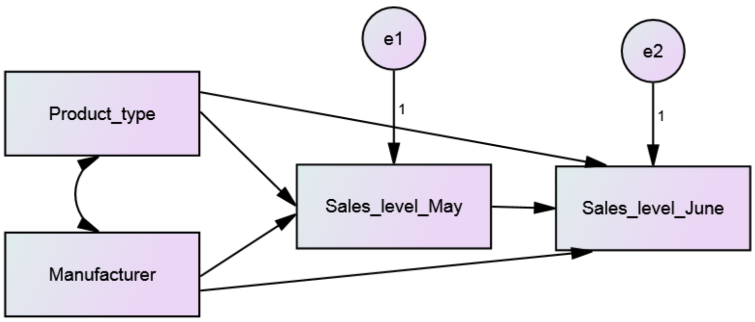

Fig. 1 Structural model

Thus defined, the model indicates the need to test the impact of product types and manufacturers to sales level in May and June. At the same time, there is a need for examining the mutual influence and a correlation between the variables *Product type* and the *Manufacturer*.

2.1. Research results obtained by method of maximum likelihood

The application of proper statistical package has enabled the estimation of the collected data using method of maximum likelihood and Bayesian method, and subsequently, drawing conclusions by comparing the results of the research. Table 1 represents the results obtained by applying maximum likelihood method on the selected data.

Table 1 Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
SL_May	<--	Manuf	-3.730	0.792	-4.709	***	
SL_May	<--	PrType	-0.941	1.502	-0.626	0.531	
SL_June	<--	Manuf	-1.090	.465	-2.346	0.019	
SL_June	<--	SL_May	0.730	.035	20.581	***	
SL_June	<--	PrType	-0.033	.845	-0.039	0.969	

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
SL_May	<---	Manuf	-0.308
SL_May	<---	PrType	-0.041
SL_June	<---	Manuf	-0.097
SL_June	<---	SL_May	0.784
SL_June	<---	PrType	-0.002

Means: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
PrType	6.996	0.142	49.276	***	
Manuf	6.246	0.269	23.200	***	

Intercepts: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
SL_May	69.800	10.095	6.914	***	
SL_June	11.827	6.191	1.910	0.056	

Covariances: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
Manuf	<->	PrType	3.926	0.654	6.000	***	

Correlations: (Group number 1 - Default model)

			Estimate
Manuf	<->	PrType	0.409

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
Manuf	18.193	1.624	11.203	***	
PrType	5.060	0.452	11.203	***	
e1	2386.501	213.029	11.203	***	
e2	754.074	67.312	11.203	***	

Source: own calculations

All values that are explained in Table 1 also appear on graphs; first, after calculating standardized estimates (Fig. 2) and second, after calculating unstandardized estimates (Fig. 3).

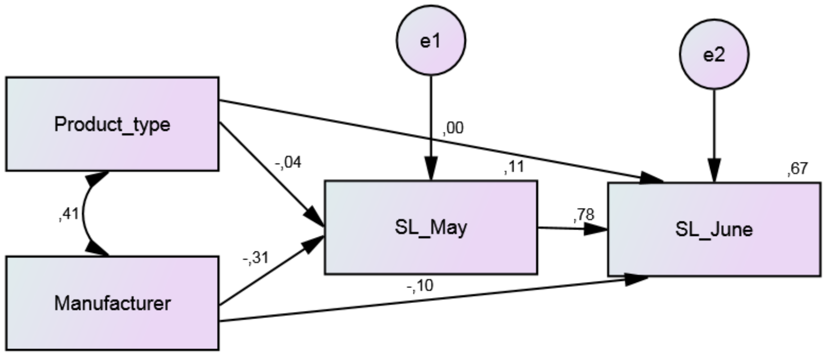

Fig. 2 Structural model with standardized estimates

The value 0.41 is correlation value between *Product type* and *Manufacturer*. Values -0.04; -0.31; -0.10; 0.00 (according to Table 1 this value is -0.002) and 0.78 are *Standardized Regression Weights*.

If we calculate unstandardized estimation, the results will be as follows:

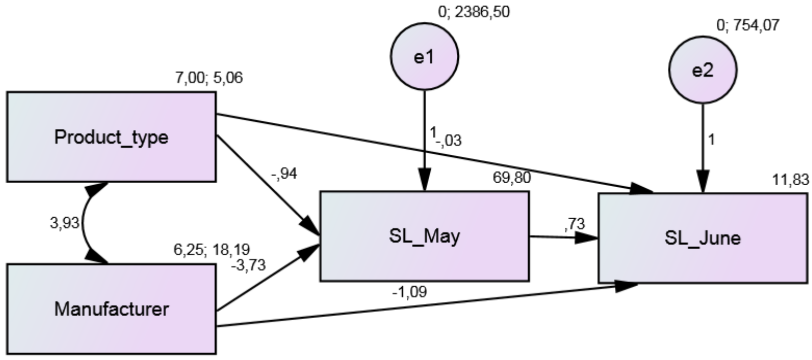

Fig. 3 Structural model with unstandardized estimates

The values from Figure 3 presented in Table 1 are sorted according to their meaning. It can be concluded, that the covariance between *Product type* and *Manufacturer* is estimated to be 3.93. Right next to the covariance in the *S.E.*column is presented the estimated standard error of about 0.654. The estimate 3.93 indicates that the observation of approximately normally distributed random variables is centred around population covariance with a standard deviation of 0.654.

The values of critical ratios are represented in the C.R. column. Dividing the covariance estimate by the estimate of its standard error gives $z = 3.926/0.654 = 6.000$. In other words, the covariance estimate is 6 standard errors above zero. The probability of getting a critical ratio as large as 6 in absolute value is less than 0.001, so the covariance between *Product type* and *Manufacturer* is significantly different from zero at the 0.001 level (two-tailed).

The estimated covariance between *Product type* and *Manufacturer* (3.93) is the value that will be compared with the Bayesian estimates. In our example, we used non-informative and informative priors in order to get more precise estimates.

2.2. Research results obtained by Bayesian method

The Bayesian paradigm is characterized by several advantages relative to the classical one, like the coherence of the whole paradigm, which is derived from the systematical applying of the Bayes law, the concept of subjective probability, the general character of the Bayesian methods which do not ask for special regularity conditions, the sounder definition of the concepts of confidence interval as well as testing [1]; [14].

For carrying out the research using Bayesian estimation it is necessary to select the appropriate prior distribution. In many cases, chosen prior will contain very little information, so the conclusions will be based only on data. Such information is called non-informative prior [12].

However, there is no prior distribution that is completely non-informative, even uniform distribution that IBM SPSS Amos used as default for each parameter, because each prior distribution carries some information. The results obtained from the Bayesian analysis will change if prior distribution changes. In an analysis conducted by the author, it will be seen that changing prior distribution affects the results of research in terms of improving their accuracy.

2.2.1. Application of non-informative uniform prior

In many cases, our prior beliefs are vague and thus difficult to translate into an informative prior. We therefore want to reflect our uncertainty about the model parameter(s) without substantially influencing the posterior parameter inference. The so-called *non-informative priors*, also called *vague* or *diffuse* priors, are employed to that end [16].

A non-informative prior might be used in the genuine absence of prior information, or if there is disagreement about the likely values of hypotheses or parameters. It may also be used in comparison with more informative priors as one aspect of a sensitivity analysis regarding posterior inferences according to the prior [2]. In literature, non-informative priors are also called objective priors and they are part of objective Bayesian analysis [7].

In our example, initially selected prior distribution is uniform prior distribution and it has the character of non-informative distribution. The results for estimated covariance between *Product type* and *Manufacturer* with the applied uniform distribution are shown in Table 2.

Table 2 Bayesian estimates with non-informative prior

	Mean	S.E.	S.D.	C.S.
Covariances				
PrType<->Manuf	4.036	0.005	0.683	1.000

Source: own calculations

If the prior distribution is uniform then the posterior mean will be close to the estimate obtained by method of maximum likelihood. It is confirmed in our example where the estimated posterior mean for covariance between *Product type* and *Manufacturer* is 4.036 and the estimated covariance obtained by the method of maximum likelihood is 3.926. In Fig. 4 we can see that the posterior distribution is centered close to 4 that corresponds to the posterior mean (4.036).

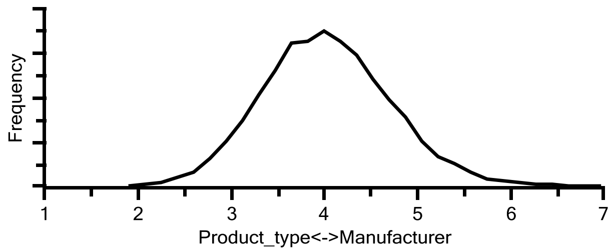

Fig. 4 Posterior distribution with non-informative prior

In this way, it was confirmed that in the case where the prior information is diffuse or non-informative, the results of classical and Bayesian statistics differ very little. In order to find more precise estimates in the next step we choose informative prior.

2.2.2. Application of informative normal prior

The prior that contains the most amount of certainty about the population parameter is an informative prior. Informative priors contain strict numerical information that is crucial to the estimation of the model and can have a large impact on final estimates [17]. The problem with using an informative prior is that people might use different background information (or interpret it differently). Thus, informative priors often seem subjective [4].

In order to prove previous claims, instead of initially selected uniform distribution, we now choose normal distribution as prior distribution. Normal distribution has the character of informative prior distribution. The results for estimated covariance between *Product type* and *Manufacturer* with the applied normal distribution are shown in Table 3.

Table 3 Bayesian estimates with informative prior

	Mean	S.E.	S.D.	C.S.
Covariances				
PrType<->Manuf	1.887	0	0.279	1.000

Source: own calculations

In Fig. 5 we can see that the posterior distribution is now centered close to 1.9 that correspond to the posterior mean (1,887).

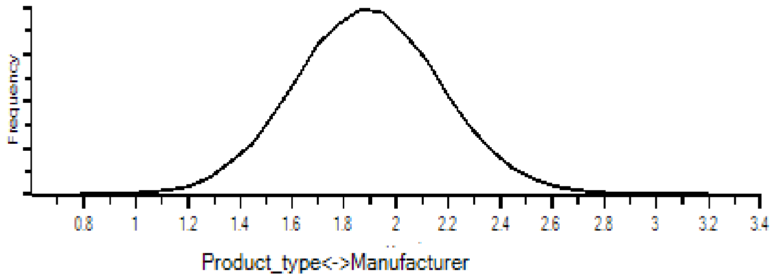

Fig. 5 Posterior distribution with informative prior

If we want to make final conclusions, we should compare results of applied methods. First, it is necessary to compare the posterior standard deviation of Bayesian statistics, (denoted S.D.) with a standard error of classical statistics (denoted S.E.) which is a useful measure of uncertainty. The value of S.D. is 0.279 and of S.E. is 0.654. Lower value of S.D. indicates that the results of Bayesian methods are more precise.

Second, we have to compare estimated covariance between *Product type* and *Manufacturer*. Unlike the case where the uniform distribution was a prior and where the results were not much different compared to classical estimates, in the case where the normal distribution is chosen for a prior, the situation is changing. Now, estimated posterior mean for covariance between *Product type* and *Manufacturer* is 1.887 and the estimated covariance obtained by the method of maximum likelihood is 3.926. So, we can conclude that we can get more accurate estimates if we use normal informative distribution as a prior distribution.

CONCLUSION

For many years, classical statistics had objective advantage compared to Bayesian approach. Supporters of Bayesian statistics did not have an opportunity to emphasize the possibility of applying Bayesian methods in data processing, because there was a real inability to perform complex methods to handle large amounts of data. The emergence of adequate software solutions has enabled intensive application of Bayesian methods in different research areas.

Comparative analysis of maximum likelihood method and Bayesian method was obtained on 252 products and their sales level from January to June 2014. By using the statistical package IBM SPSS Amos and constructing the structural model, the base for further analysis was made. The example has shown the importance of prior information in the Bayesian estimation, and thereby confirmed that the Bayesian approach is more complex because there is an obvious need for estimating a prior probability and examining its sensitivity. Including normal prior distribution, as informative prior, it was concluded that the estimates obtained from Bayesian approach represent an improvement of estimates obtained by classical method in terms of accuracy.

REFERENCES

1. Caraiani, P. (2010) Forecasting Romanian GDP using a BVAR Model, Romanian Journal of Economic Forecasting, 4/2010, 76-87.
2. Congdon, P. (2006), Bayesian Statistical Modelling, Second Edition, John Wiley and Sons Ltd, Chichester, West Sussex, England.
3. Depaoli, S. (2013) Mixture class recovery in GMM under varying degrees of class separation: Frequentist versus Bayesian estimation, Psychological Methods, 18(2), 186-219.
4. Downey, A.B. (2012), Think Bayes, Bayesian Statistics Made Simple, Version 1.0., Green Tea Press, Needham, Massachusetts.
5. Fernández-Villaverde, J., Rubio-Ramírez, J.F. (2007) Estimating macroeconomic models: A likelihood approach, The Review of Economic Studies, 74(4), 1059-1087.
6. Flurry, T., Shephard, N. (2011) Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models, Econometric Theory, 27(05), 933-956.
7. Ghosh, J.K, Delampady, M., Samanta, T. (2006), An Introduction to Bayesian Analysis, Theory and Methods, Springer Science+Business Media, LLC, New York.
8. Kim, S.Y., Suh, Y., Kim, J.S., Albanese, M., Langer M.M. (2013) Single and multiple ability estimation in the SEM framework: a non-informative Bayesian estimation approach, Multivariate and Behavioral Research, 48(4), 563-591.
9. Koop, G. (2003), Bayesian Econometrics, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England.
10. Lee, S.Y., Song, X.Y. (2004) Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes, Multivariate Behavioral Research, 39, 653–686.
11. Lemmon, A.R., Brown, J. M., Stanger-Hall, K., Lemmon, E.M. (2009) The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference, Systematic Biology, 58(1), 130-145.
12. Papić-Blagojević, N. (2015), Comparative analysis of classical inference and Bayesian approach in the processing of economic data, Doctorial dissertation, Niš, Faculty of Economics.
13. Pitt, M.K., Silva, R.D.S., Giordani, P., Kohn, R. (2012) On some properties of Markov chain Monte Carlo simulation methods based on the particle filter, Journal of Econometrics, 171(2), 134-151.
14. Poirier, D. (2008), Bayesian Econometrics, In S.N. Durlauf and L.E. Blume (eds), The New Palgrave Dictionary of Economics, Palgrave Macmillan.
15. Prix, R., Krishnan, B. (2009) Targeted search for continuous gravitational waves: Bayesian versus maximum-likelihood statistics, Classical and Quantum Gravity, 26 (20), 204013.
16. Rachev, T.S., Hsu, J.S.J., Bagasheva, B.S., Fabozzi, F.J. (2008), Bayesian Methods in Finance, John Wiley & Sons, Inc., New Jersey.
17. Schoot, R.V.D., Depaoli, S. (2014) Bayesian analyses: where to start and what to report, The European Health Psychologist, Volume 16, Issue 2, 75-84.
18. Ward, E.J. (2008) A review and comparison of four commonly used Bayesian and maximum likelihood model selection tools, Ecological Modelling, 211(1), 1-10.
19. Zhang, Z., Hamagami, F., Wang, L., Grimm, K.J., Nesselroade, J.R. (2007) Bayesian analysis of longitudinal data using growth curve models, International Journal of Behavioral Development, 31 (4) , 374-383.

ISPITIVANJE PERFORMANSI METODA MAKSIMALNE VERODOSTOJNOSTI I BAJESOVOG PRISTUPA U OCENJIVANJU NIVOA PRODAJE

Metod maksimalne verodostojnosti i Bajesov metod nalaze široku primenu u obradi podataka, ne samo u oblasti ekonomije već i u drugim poljima istraživanja. Cilj rada je da se kroz komparativnu analizu ova dva metoda na odabranim ekonomskim podacima identifikuje koji od pristupa ima bolje karakteristike u datim okolnostima. Poređenjem ocenjenih vrednosti dobijenih primenom metoda maksimalne verodostojnosti i Bajesovog metoda na podacima o ostvarenoj prodaji kompanije CaliVita Int. izvedeni su zaključci o performansama oba pristupa. U radu je pokazano da odabirom informativnih apriornih informacija, Bajesov pristup daje preciznije ocene u odnosu na rezultate dobijene primenom metoda maksimalne verodostojnosti.

Ključne reči: *metod maksimalne verodostojnosti, Bajesov metod, ocenjivanje, apriorna informacija, nivo prodaje*