

A Comparison Between Bayesian and Maximum Likelihood Estimations in Estimating Finite Mixture Model for Financial Data

(Perbandingan antara Bayesian dan Anggaran Kebolehhajian Maksimum dalam Menganggar Model Campuran Terhingga untuk Data Kewangan)

SEUK-YEN PHOONG* & MOHD TAHIR ISMAIL

ABSTRACT

Over the years, maximum likelihood estimation and Bayesian method became popular statistical tools in which applied to fit finite mixture model. These trends begin with the advent of computer technology during the last decades. Moreover, the asymptotic properties for both statistical methods also act as one of the main reasons that boost the popularity of the methods. The difference between these two approaches is that the parameters for maximum likelihood estimation are fixed, but unknown meanwhile the parameters for Bayesian method act as random variables with known prior distributions. In the present paper, both the maximum likelihood estimation and Bayesian method are applied to investigate the relationship between exchange rate and the rubber price for Malaysia, Thailand, Philippines and Indonesia. In order to identify the most plausible method between Bayesian method and maximum likelihood estimation of time series data, Akaike Information Criterion and Bayesian Information Criterion are adopted in this paper. The result depicts that the Bayesian method performs better than maximum likelihood estimation on financial data.

Keywords: Akaike information criterion; Bayesian information criterion; Bayesian method; finite mixture model; maximum likelihood estimation

ABSTRAK

Sejak beberapa tahun, anggaran kebolehhajian maksimum dan kaedah Bayesian menjadi alat statistik popular yang sesuai digunakan untuk model campuran terhingga. Trend ini bermula dengan adanya teknologi komputer sejak sedekad yang lalu. Selain itu, sifat asimptot bagi kedua-dua kaedah statistik juga menjadi salah satu daripada faktor utama dalam meningkatkan populariti kaedah ini. Perbezaan antara kedua-dua kaedah ini adalah parameter untuk anggaran kebolehhajian maksimum adalah tetap tetapi tidak diketahui manakala parameter bagi kaedah Bayesian bertindak sebagai pemboleh ubah rawak dengan taburan yang dikenali sebelum ini. Dalam kertas ini, kedua-dua anggaran kebolehhajian maksimum dan kaedah Bayesian digunakan untuk mengkaji hubungan antara kadar pertukaran wang dan harga getah bagi Malaysia, Thailand, Filipina dan Indonesia. Untuk mengenal pasti kaedah yang paling munasabah antara kaedah Bayesian dan anggaran kebolehhajian maksimum untuk data siri masa, kriteria maklumat Akaike dan kriteria maklumat Bayesian diguna pakai dalam kertas ini. Kesimpulannya, keputusan menunjukkan bahawa kaedah Bayesian mempunyai prestasi yang lebih baik daripada anggaran kebolehhajian maksimum dalam menganalisis data kewangan.

Kata kunci: Anggaran kebolehhajian maksimum; kaedah Bayesian; kriteria maklumat Akaike; kriteria maklumat Bayesian; model campuran terhingga

INTRODUCTION

The finite mixture model was a k -component finite-dimensional mixture model which suggested by Newcomb (1886) in computing outliers in the data. Each component of the model belongs to the same family distribution, but with different parameters. The finite mixture model was then arisen widely range from financial to social science because of the flexibility properties in both practical and theoretical points of view.

In addition, Titterton et al. (1985) reviewed that finite mixture model can be interpreted in two different ways; direct and indirect interpretation. Direct interpretation means that each observation of a subpopulation from a

population has no knowledge of identity or it also means that, the finite mixture model is intended to be a direct representation of the underlying phenomenon. While for the indirect interpretation, the subpopulations have no physical interpretation so the mixture model was used to approximate an unknown distribution.

Method of moments is one of the oldest statistical methods that was adopted in modelling the parameters of distributions, including mean, variance, covariance, skewness and kurtosis. However, this method only effective in measuring simple computations and adapt to the small sample sizes. Thus, other approaches such as maximum likelihood estimation and Bayesian method

are then introduced in order to fit the finite mixture model instead of the method of moments.

Both the maximum likelihood estimation and Bayesian method have advantages in analyzing moderate or large sample sizes and widely applied along with the availability of computer technology. Furthermore, these two methods provide asymptotic properties such as asymptotically normal and asymptotically efficient for data analysis.

Maximum likelihood estimation is proposed by Fisher (1922) in maximizing the likelihood function to utilize the parameter vector. This method also acts as a fundamental part of data analysis where the parameters are fixed but unknown. Moreover, maximum likelihood estimation is popular for data analysis because the estimates always show the smallest variance in comparison with other statistical tools as the sample sizes increases and this can be referred to Hosmer (1973) study. According to Hosmer (1973), the iteration of maximum likelihood estimation will not converge to parameter values if the sample size is small and the main reason that results in was the presence of outliers. The estimation of the finite mixture model using maximum likelihood method is widely applied in many fields such as financial, economics, meteorology, social science, statistics, engineering, biology and medical. Some studies that applied maximum likelihood estimation in financial analysis are Duan and Simonato (2001), Durham and Gallant (2002), Johnson et al. (2011) and Kladviko (2007). Furthermore, Lepage (2012) proved that maximum likelihood estimation is consistent and show asymptotic normality for the parameter estimates. Meanwhile Avdis and Wachter (2013) studied that maximum likelihood estimation is reliable when the sample sizes is finite.

For Bayesian method, it was proposed by Laplace (1986) in modelling the parameters of the models. Bayesian method is different from maximum likelihood estimation because the Bayesian method involves three important concepts which are prior distribution, likelihood function and posterior distribution. Prior distribution described the basic details about the model and the likelihood function demonstrated the information after the analysis. While the posterior distribution act as an average among the other two concepts and the results that derived from the posterior distribution is then used to analyse the parameters of the model.

Bayesian method often shows consistency characteristic which means that the parameter estimates are close to the predictive distributions. In addition, Bayesian method allows to test the hypotheses without predetermining the outcome of the test based on the selection of the sample size and also need not to pre-specify an arbitrary level of significance. According to Love et al. (2007), Bayesian method is efficient in data analysis, easy and well understood by users, while Monahan (1983) also proved that Bayesian method is feasible and useful in estimating the time series data. Other studies that relate to the application of Bayesian method can be referred to Newton et al. (2001) in biology, Coles et al. (2005), Martina et al. (2008) and Olson et al. (1995) in meteorology,

Cipriani et al. (2012), Feng and Xie (2012) and Jacquier and Polson (2012) in finance.

With the nonlinear properties of real financial data, both the maximum likelihood estimation and Bayesian method are adopted to examine the parameters of data. In this paper, both the statistical tools are used to explore the relationship between rubber price and exchange rate for Malaysia, Thailand, Philippines and Indonesia. Then, the model selection criterion such as Akaike information criterion (AIC) and Bayesian information criterion (BIC) are utilized in order to distinguish the most plausible method for data analysis.

The rest of this paper is structured as follows. The materials and methods adopt for the present paper are described in the next section. In the section after that, we will show and discuss the results that was obtained. Finally, the last section made conclusion on the relationship among rubber price and exchange rate based on the countries that selected.

MATERIALS AND METHODS

The model that was adopted in this paper is a two-dimensional mixture model or known as a finite mixture model with a d -dimensional. The primary reason for employing this model is because it has no limited criteria to be fulfilled. There are only three assumptions needed to be fitted for finite mixture model. The first assumption is that group the datasets given. Secondly, the observations are uncorrelated. Lastly, the parameters of the components are distinct a priori.

For this paper, the mixture of normal distributions is applied and the general formula is

$$f(x_i) = \pi\varphi_1(\mu_1, \sigma_1^2) + (1 - \pi)\varphi_2(\mu_2, \sigma_2^2), \quad (1)$$

where $\varphi_i(\mu_i, \sigma_i^2)$ denotes the probability density function of a normal distribution with mean and variance σ_i^2 . In addition, represents the weight of the normal distribution which is the first regime. Since the assumption for the entire weights of a component is equal to one, therefore, the weight of the second regime is $1 - \pi$.

Maximum likelihood estimation is employed in this paper to fit the finite mixture of normal distribution. Hence, the general equation for maximum likelihood estimation is as follows. Assume that the observations x_1, \dots, x_n are independent and identically distributed where the likelihood function is obtained if the fixed parameters are provided as in (2) (McLachlan & Peel 2000).

$$L(\theta|x_1, \dots, x_n) = f(x_1, \dots, x_n|\theta) = \prod_{j=1}^n f(x_j|\theta), \quad (2)$$

where x_n is the n -th observations and θ represents the parameters of the model. Nonetheless, the likelihood function is more convenient to work with logarithms and known as log-likelihood function. Thus, the equation to compute the log-likelihood is

$$\ln L(\theta|x_1, \dots, x_n) = \sum_{j=1}^n \ln f(x_j|\theta). \quad (3)$$

Additionally, this paper also applied the Bayesian method to fit finite mixture model and the definition of the Bayesian method is described in the following. With k observations x_1, \dots, x_k and drawn randomly from finite mixture distributions to make inferences about the mixture model. The Bayesian inference for the standard finite mixture models are denoted as

$$p(x_j|\varphi) = \sum_{j=1}^k n_j p(x_j|\varphi_j), \quad (4)$$

where $p(x|\varphi)$ denotes the probability density with known number of components. Moreover, n_j represents the j -th sample n . Bayesian inferences is crucial to determine the parameters that characterize the data generating process given observed data.

Then, the equation for likelihood function of Bayesian method is

$$P(\theta) = \frac{1}{(\sqrt{2\pi})^n \cdot \sigma^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y - f(I, \theta))^2\right), \quad (5)$$

where σ^2 represents the variance of the residuals, θ represents as the model parameters and I denotes the observed input data.

The posterior distribution $P(\theta|x)$ is the update information of the prior distribution $P(\theta)$ with the likelihood function $P(x|\theta)$. Bayesian method can be defined as

$$P(\theta|x) = (6)$$

where $P(\theta)$ denotes the prior distribution of the possible θ values

$P(x|\theta)$ represents the likelihood function of the probabilistic data

$P(x)$ is a probability of the evidence which is constant

$P(\theta|x)$ is the posterior distribution of θ given the observed data x

In the present paper, the finite mixture model is fitted by the maximum likelihood estimation and Bayesian method in investigating the relationship among rubber price and exchange rate for Malaysia, Thailand, Philippines and Indonesia. Additionally, the monthly data obtained from DataStream and the time period that acquire is from July 2005 until September 2012. While the software that utilize is SAS version 9.3.

RESULTS AND DISCUSSION

Both the results that obtained for maximum likelihood estimation and Bayesian method are displayed in this section to describe the relationship among rubber price and exchange rate. For maximum likelihood estimation, there

are three important tables that describe the results which include model information, optimization information and parameter estimates. The findings, display that all four countries show the similar results for model information table and optimization information table.

TABLE 1. Model information table

Model information	
Response variable	Returns
Type of model	Homogeneous regression mixture
Distribution	Normal
Components	2
Link Function	Identity

Table 1 is the model information table for Malaysia, Thailand, Philippines and Indonesia. The primary role of the model information table is to display the basic information about the model. This table described that the response variable for the data is returns of the exchange rate, while the model that applied is a homogeneous regression mixture model which illustrate that all components share the same distribution family and link function. In addition, the model is a two-component normal mixture model with identity link function.

TABLE 2. Optimization information table

Optimization information	
Optimization technique	Dual Quasi-Newton
Parameters in optimization	7
Mean function parameters	4
Scale parameters	2
Mixing prob parameters	1
Number of threads	2

Based on Table 2, the optimization information is provided and the main role of this table is to demonstrate the basic information about the optimization setup to measure the maximum likelihood estimates. By referring to Table 2, a two-component normal mixture model is fitted by maximum likelihood estimation that involved seven parameters in the optimization which listed in the optimization information table. The parameters include four mean function parameters, two scale parameters and a mixing probability parameter. Moreover, Table 2 also depicts other optimization information such as the technique that applied for optimization process and number of threads that used to examine the maximum likelihood estimation. The leading role of optimization technique is used to iterate the algorithm based on the approximation of the objective function and the technique required in this paper is Dual Quasi-Newton. Dual Quasi-Newton is an algorithm that searches for the local maximum and minimum of functions. This technique also provides efficient solution for medium to moderately large sample

size. Moreover, the number of threads or number of multithreading processes that obtained for the model is equal to two.

Then, the parameter estimates that acquire for all four selected countries are collected and listed in Table 3. The parameter estimates for the two-component normal mixture models with seven parameters are shown in Table 3.

Table 3 shows that the fluctuation of rubber price affects the exchange rate. This is reasonable because rubber is one of the top ten export products in Malaysia, Thailand, Philippines and Indonesia. Moreover, Thailand also plays a leading role as world's largest rubber producer and exporter, followed by Indonesia and Malaysia in year 2010. In addition, the Philippines also act as one of the top ten world's largest rubber producer and exporter. Thus, the movements of rubber price influence the exchange rate of these countries.

Based on Table 3, it can be concluded that the rubber price and exchange rate show the negative effect for all selected countries. This indicates that the rubber price depreciates with the increase of exchange rate or vice versa. The result has similarity with Burger et al. (2002) findings which noted that the exchange rate represents the

development of a country. Therefore, if the exchange rate depreciates, the demand for rubber will decrease with the increase of rubber price.

The variance for both component 1 and component 2 for all four countries showed that all the variance is small and this indicates that the value is closed to the mean. Thus, the results obtained can be concluded as valid, reliable and significant. In addition, Table 3 also describes the mixing probability or weight for each country. The total weight of all components is one. Therefore, the findings described that the probability density function for Malaysia is $0.9698f_1 + 0.0302f_2$, Thailand is $0.4158f_1 + 0.5842f_2$, Philippines is $0.9304f_1 + 0.0696f_2$ and Indonesia is $0.8329f_1 + 0.1671f_2$ where f_1 denotes the probability density function for component 1 and f_2 represents a probability density function for component 2.

Moreover, the results for Bayesian method in exploring the relationship among financial time series data also exhibit in this section. There are three important tables, including Bayes information, prior distribution and posterior distribution table. The Bayes information table for Malaysia, Thailand, Philippines and Indonesia are found similar and the can be seen in Table 4.

TABLE 3. Parameter estimates

Number of component	Effect	Estimate				Probability graph
		Malaysia	Thailand	Philippines	Indonesia	
1	Exchange Rate	-0.00145	-0.00603	-0.0019	-0.00092	
1	Rubber	-0.06867	-0.03211	-0.03738	-0.04138	
2	Exchange Rate	-0.01004	-0.00124	-0.02008	0.001366	
2	Rubber	0.02728	-0.01691	0.008118	-0.1483	
1	Variance	0.000163	0.000432	0.000217	0.000176	
2	Variance	1E-8	0.000088	4.97E-8	0.003038	
Mixing probabilities		0.9698	0.4158	0.9304	0.8329	

TABLE 4. Bayes information table

Bayes information	
Sampling algorithm	Conjugate
Burn-in size	2000
MC sample size	10000
Parameters in sampling	7
Mean function parameters	4
Scale parameters	2
Mixing prob parameters	1
Number of threads	2

By referring to Table 4, the FMM procedure applied conjugate sampling algorithm with 2000 burn-in sizes followed by 10000 Monte Carlo sample sizes. The conjugate sampling algorithm is adopted because the model is a homogenous mixture. Additionally, the Bayes information table also depicts that there are seven parameters in sampling which include four mean function parameters, two scale parameters and a mixing probability parameter. Meanwhile the number of threads that used to perform multithreaded analysis in this case is equal to two.

Table 5 describes the information about the prior distribution for the sampled parameters along with the distribution family and initial values for all the selected countries. The component column in the table identified the mixture of component to which a parameter belongs. This finding indicates that the parameters for prior distribution which associated with the exchange rate and rubber price are a normal distribution. Meanwhile the prior distribution of variance parameters is an inverse gamma distribution. Lastly, the prior distribution for a mixing probability is a Dirichlet distribution.

Then, parameter estimates that was acquired from posterior distribution table for Malaysia, Thailand,

Philippines and Indonesia are collected and demonstrated in Table 6.

By referring to Table 6, it can be concluded that the fluctuation of rubber price influence the exchange rate and this can be referred to the probability graph that was placed in the last column. From the results, there is a negative effect between rubber price and exchange rate in Malaysia, Thailand, Philippines and Indonesia which means that the rubber price is increasing in value with the depreciation of the exchange rate or vice versa. In addition, the variance for both components 1 and 2 in Malaysia, Thailand, Philippines and Indonesia are small. This illustrates that the finding is valid, reliable and significant.

The mixing probability that displays at Table 6 denotes the weight for component 1 meanwhile the weight for component 2 is 0.0115 because the assumption for total weight of all components is one. Therefore, the probability density function for all four countries is where and denotes the probability density function for components 1 and 2.

Moreover, the findings from both maximum likelihood estimation and Bayesian method is similar in which it provides negative effect between rubber price and exchange rate for Malaysia, Thailand, Philippines and

TABLE 5. Prior distribution

Component	Effect	Distribution	Initial value			
			Malaysia	Thailand	Philippines	Indonesia
1	Exchange rate	Normal	-0.0020	-0.0033	-0.0033	0.0003
1	Rubber	Normal	-0.0551	-0.0233	-0.0281	-0.0854
2	Exchange rate	Normal	-0.0020	-0.0033	-0.0033	0.0003
2	Rubber	Normal	-0.0551	-0.0233	-0.0281	-0.0854
1	Variance	Inverse Gamma	0.0002	0.0002	0.0002	0.0007
2	Variance	Inverse Gamma	0.0002	0.0002	0.0002	0.0007
1	Mixing parameter	Dirichlet	0.6180	0.6180	0.6180	0.6180

TABLE 6. Posterior distribution





Component	Effect	Estimate				Probability graph
		Malaysia	Thailand	Philippines	Indonesia	
1	Exchange rate	-0.00187	-0.00316	-0.00314	0.000412	
1	Rubber	-0.0525	-0.0210	-0.0256	-0.0828	
2	Exchange rate	0.00323	0.0163	0.00934	0.00324	
2	Rubber	0.1049	0.0915	0.1082	0.1097	
1	Variance	0.0235	0.0236	0.0235	0.0240	
2	Variance	1.0455	1.0100	1.0470	1.0455	
	Mixing probability	0.9885	0.9885	0.9885	0.9885	

TABLE 7. Model selection criterion

Country	AIC		BIC	
	Maximum likelihood estimation	Bayesian method	Maximum likelihood estimation	Bayesian method
Malaysia	-500.7	-629.6	-486.0	-612.4196
Thailand	-466.2	-627.6	-449.0	-610.4196
Philippines	-480.5	-628.0	-463.4	-610.8196
Indonesia	-411.8	-614.4	-394.6	-597.2196

Indonesia. In order to define the most plausible method among maximum likelihood estimation and Bayesian method, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were applied. Table 7 shows the value for both AIC and BIC for each country.

Table 7 displays the AIC and BIC for both maximum likelihood estimation and Bayesian method based on different countries. The findings indicate that Bayesian method is more plausible than maximum likelihood estimation because both the AIC and BIC for Bayesian method demonstrated lowest value which denotes the better fits. This conclusion is similar with Safaa Nasir and Nashaat Jaisam (2012) study in which the Bayesian method is superior compared to maximum likelihood estimation and median estimation in modelling life time event data. Furthermore, Pandey et al. (2011) also concluded that the Bayesian method performs better than maximum likelihood estimation when the sample size is small for life time data. However, both the statistical tools are performing equally efficient for large sample sizes.

CONCLUSION

The findings of this paper describe that both maximum likelihood estimation and Bayesian method shows that there is a negative effect between rubber price and exchange rate in Malaysia, Thailand, Philippines and Indonesia. In addition, both the statistical method, maximum likelihood estimation and Bayesian method are popular in fitting finite mixture model to model the nonlinear financial time series data. Nonetheless, based on the findings of this simulation study depicts that Bayesian method is more efficient than maximum likelihood estimation.

ACKNOWLEDGEMENTS

The authors wish to thanks their respective university for the financial support and computing facilities that make the works possible to be completed.

REFERENCES

- Avdis, E. & Wachter, J.A. 2013. Maximum likelihood estimation of the equity premium. *Working paper*.
- Burger, K., Smith, H. & Vogelvang, B. 2002. Exchange rates and natural rubber prices, the effect of the Asian crisis. *Spain: 2002 International Congress Zaragoza*, no. 24958.
- Cipriani, M., Costantini, R. & Guarino, A. 2012. A Bayesian approach to experimental analysis: Trading in a laboratory financial market. *Review of Economic Design* 16(2): 175-191.
- Coles, S.G., Sisson, S.A. & Pericchi, L.R. 2005. A case for a reassessment of the risks of extreme hydrological hazards in the Caribbean. *SERRA* 20: 296-306.
- Duan, J.C. & Simonato, J.G. 2001. Maximum likelihood estimation of deposit insurance value with interest rate risk. *Journal of Empirical Finance* 9: 109-132.
- Durham, G.B. & Gallant, A.R. 2002. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *Journal of Business and Economic Statistics* 20(3): 297-316.
- Feng, X.X. & Xie, D.J. 2012. Bayesian estimation of CIR model. *Journal of Data Science* 10: 271-280.
- Fisher, R.A. 1922. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London A* 222: 309-368.
- Hosmer, D.W. 1973. A comparison of iterative maximum likelihood estimates of the parameters of a mixture of two normal distributions under three different types of sample. *Biometrics* 29: 761-770.
- Jacquier, E. & Polson, N.G. 2012. Asset allocation in finance: A Bayesian perspective. In *Hierarchical Models and MCMC: A tribute to Adrian Smith*, edited by Dellaportas, D., Polson, N. & Stephen, G. Oxford: Oxford University Press.
- Johnson, P.H. Jr., Qi, Y.X. & Chueh, Y.C. 2011. Bias-corrected maximum likelihood estimation in actuarial science. *Proceedings of 46th Actuarial Research Conference*.
- Kladivko, K. 2007. Maximum likelihood estimation of the Cox-Ingersoll-Ross process: The Matlab implementation. Technical Computing Prague.
- Laplace, P.S. 1986. Memoir on the probability of the causes of events. *Statistical Science* 1 3: 364-378.
- Lepage, G. 2012. Maximum likelihood estimation for conditionally heteroscedastic models when the innovation process is in the domain of attraction of a stable law. *Parallel Meetings*, 27-31 August 2012, Malaga, Spain.
- Love, K.R., Ye, K.Y., Smith, E.P. & Prisley, S.P. 2007. Error models in geographic information systems vector data using Bayesian methods. *International Journal of Geographical Information Science technical report No. 07-1*.
- Martina, M.L.V., Todini, E. & Libralon, A. 2008. Rainfall thresholds for flood warning systems: A Bayesian decision approach. *Water Science and Technology Library* 63(3): 203-227.
- McLachlan, G.J. & Peel, D. 2000. *Finite Mixture Models*. New York: John Wiley & Sons.
- Monahan, J.F. 1983. Fully Bayesian analysis of ARMA time series models. *Journal of Econometrics* 21: 307-331.

- Newcomb, S. 1886. A generalized theory of the combination of observations so as to obtain the best result. *American Journal of Mathematics* 8: 343-366.
- Newton, M.A., Kenzioriski, C.M., Richmond, C.S., Blattner, F.R. & Tsui, K.W. 2001. On differential variability of expression ratios: Improving statistical inference about gene expression changes from microarray data. *Journal of Computational Biology* 8: 37-52.
- Olson, D.A., Junker, N.W. & Korty, B. 1995. Evaluation of 33 years of quantitative precipitation forecasting at the NMC. *Weather Forecasting* 10: 498-511.
- Pandey, B.N., Dwivedi, N. & Bandyopadhyay, P. 2011. Comparison between Bayesian and maximum likelihood estimation of scale parameter in Weibull distribution with known shape under linex loss function. *Journal of Scientific Research* 55: 163-172.
- Safaa Nasir & Nashaat Al-Anber. 2012. A comparison of the Bayesian and other methods for estimation of reliability function for Burr-XII distribution. *Journal of Mathematics and Statistics* 8(1): 42-48.
- Titterington, D.M., Smith, A.F.M. & Markov, U.E. 1985. *Statistical Analysis of Finite Mixture Distributions*. New York: Wiley.

School of Mathematical Sciences
Universiti Sains Malaysia
11800 Penang
Malaysia

*Corresponding author; email: yen_phoong@hotmail.com

Received: 22 May 2013

Accepted: 5 February 2015