Fitting Finite Mixture Model to Exchange Rate Using Maximum Likelihood Estimation

[1]Phoong Seuk Yen and [2]Mohd Tahir Ismail

[1,2]Universiti Sains Malaysia

School of Mathematical Sciences,

11800 Minden, Penang, Malaysia

[1]yen_phoong@hotmail.com, [2]mtahir@cs.usm.my

***Abstract*- Exchange rate has great influence to the inflation and economic growth for a country. The importance of currency is that the great influence on import and export prices with the changes of exchange rate. Thus, maximum likelihood estimation (MLE) is used to fit finite mixture model. In this paper, a two-component mixture of normal distribution is used to analysis the return value of nominal monthly exchange rate for Malaysia, Thailand and Philippines by using maximum likelihood estimation. The data collected for this paper is taken from July 2005 until September 2012.**

Key words: Exchange rate, maximum likelihood estimation, finite mixture model, mixture of normal distribution

I. INTRODUCTION

Finite mixture model is increasingly getting attention over the years because it is a flexible method of modelling. Fields in which finite mixture model has widened applied include economics, engineering, meteorology, financial, biology, medicine, genetics, marketing, physical, social sciences and neural networks. In statistics, the usefulness of applied finite mixture model include modelling of over-dispersed data, fitting the zero-inflated or hurdle models, estimating heavy-tailed densities, modelling heterogeneity in a cluster analysis, calculating of switching regressions, allows measuring a two component distribution and in estimating a mixing probability in the data. In addition, finite mixture model also able to measure the observable variables, describe the complex system on the data analysis and measure the sensitivity and specificity of the diagnostic and screening procedures.

Finite mixture model made its first recorded appearance in a statistical literature paper by Newcomb [1] in modelling outliers of the data. A few years later, Pearson [2] applied method of moments to fit finite mixture model in order to estimate the parameters in the model. In Pearson study, a mixture of two univariate normal components is used to measure the ratios of forehead to body lengths for 1,000 crabs. Then, method of moments has been widely used and this can be proved by refer to the studies of Charlier and Wicksell [3] and Doetsch [4]. According to Charlier and Wicksell [3], method of moments is extended to the case of bivariate normal components. Then, Doetsch [4] further the study of Charlier and Wicksell [3] to the case of more than two univariate normal components.

With the advent of high speed computer, maximum likelihood estimation is introduced to fit finite mixture model by Fisher [5]. The main reason of applied maximum likelihood estimation instead of method of moment is that maximum likelihood estimation enable large sample sizes being analysis and the time consume is faster than method of moment. Then, Rao [6] applied maximum likelihood estimation to estimate two components of biological problem with equal standard deviations. However, Hasselblad [7] has different idea in his study. In Hasselblad [7] study, maximum likelihood estimation is used to estimate unequal variances with more than two components. Then, maximum likelihood estimation has received more attention and many statisticians start applied the maximum likelihood estimation in their study such as Peters and Coberly [8], Duda and Hart [9] and Hosmer [10] were applied maximum likelihood estimation in estimated the parameters in a mixture distribution by expressed it into an iterative form.

According Horton and Laird [11], maximum likelihood estimation is used to estimate the parameters of joint distribution and solving the missing covariates in the model. Furthermore, Yang [12] applied maximum likelihood estimation to examine biology data. A data set of HIV-1 gp120 *env* gene is analysed in order to identify a number of sites in which under the positive selection.

Moreover, maximum likelihood estimation also popular in used in meteorology field and this can refer to the studies of Castiglioni et al. [13] and Owolawi [14]. Castiglioni et al. [13] described that maximum likelihood estimation is an effective statistical method in examined the rain drop data. While in Owolawi [14] study has similar idea with Castiglioni et al. [13] that applied maximum likelihood estimation in estimated the raindrop size distribution and rain attenuation.

The main reason in which maximum likelihood estimation is vital in statistics is because it has higher probability of being close to the quantities to be estimated and has lower variance as the sample size increases in compared to other

methods. In addition, the method also performs statistically well understood and provides a consistent approach when developed for a large variety of estimation situations.

In this study, a two-component mixture of normal distribution is analysis by using maximum likelihood estimation in order to model the financial time series data. The data that mention is the returns value of nominal monthly exchange rate for Malaysia, Thailand and Philippines.

The structural of this paper is as follows. The section 2 presents the literature review. Section 3 presents the sample and data that collected. Section 4 presents the methodology meanwhile section 5 is the results and discussion. Lastly, section 6 presents the conclusion of this study.

II. SAMPLE AND DATA

The variable that applied in this paper is the return of nominal monthly exchange rate for Malaysia, Thailand and Philippines from July 2005 until September 2012. This paper contains 86 observations for each country. The unit of currency for Malaysia is in Ringgit Malaysia (RM) meanwhile the unit of currency for Thailand and Philippines are in Thai Baht and Peso. The data is collected from yahoo finance. While the statistical software package that use in the study is SAS version 9.3.

III. METHODOLOGY

In order to estimate the data, the formula for finite mixture model and maximum likelihood estimation is provided in this section. The general formula for mixture of normal distributions is

$$f(x_t) = \pi\varphi_1\left(\mu_1, \sigma_1^{\,2}\right) + (1-\pi)\varphi_2\left(\mu_2, \sigma_2^{\,2}\right) \quad (1)$$

where $\varphi_i\left(\mu_i, \sigma_i^{\,2}\right)$ denotes the probability density function of a normal distribution with mean μ_i and variance $\sigma_i^{\,2}$. Meanwhile, π represents the weight of the normal distribution which is the first regime. Since the total weight for a component is equal to one, therefore, the weight of the second regime is $1-\pi$.

Then, maximum likelihood estimation is applied in this paper to fit the mixture of normal distribution. The general formula for maximum likelihood estimation is as follows.

By assuming that $x_1,...,x_n$ are independent data, then the joint density function for all observations have to be specifies and the formula for joint density function is

$$f(x_1,...,x_n|\theta) = f(x_1|\theta)\times f(x_2|\theta)\times...\times f(x_n|\theta) \quad (2)$$

However, if fixed parameters are provided for this function, the function will be known as likelihood function and the formula is

$$L(\theta|x_1,...,x_n) = f(x_1,...,x_x|\theta) = \prod_{i=1}^{n} f(x_i|\theta) \quad (3)$$

In general, the likelihood function is more convenient to work with the logarithm, called the log-likelihood.

$$\mathrm{In}L(\theta|x_1,...,x_n) = \sum_{i=1}^{n} \mathrm{In}\ f(x_i|\theta) \quad (4)$$

or the average log-likelihood

$$\hat{l} = \frac{1}{n}\mathrm{In}\ L \quad (5)$$

where the hat over *l* denotes the expected log-likelihood of observation in the model.

IV. RESULTS AND DISCUSSIONS

Before the analysis being started, a descriptive statistics table is displays to describe the basic information for the return value of nominal monthly exchange rate for each country. Information such as mean, standard deviation and skewness for the return of exchange rate for Malaysia, Thailand and Philippines are listed in table 1 meanwhile figure 1 display the histogram with normal curve of return value for exchange rates for Thailand, Malaysia and Philippines.

TABLE I.
Summary Statistics

Statistics	Country		
	Malaysia	Thailand	Philippines
Mean	-0.002361	-0.003426	-0.003446
Standard Deviation	0.01434	0.01569	0.01540
Skewness	0.233	0.453	0.324

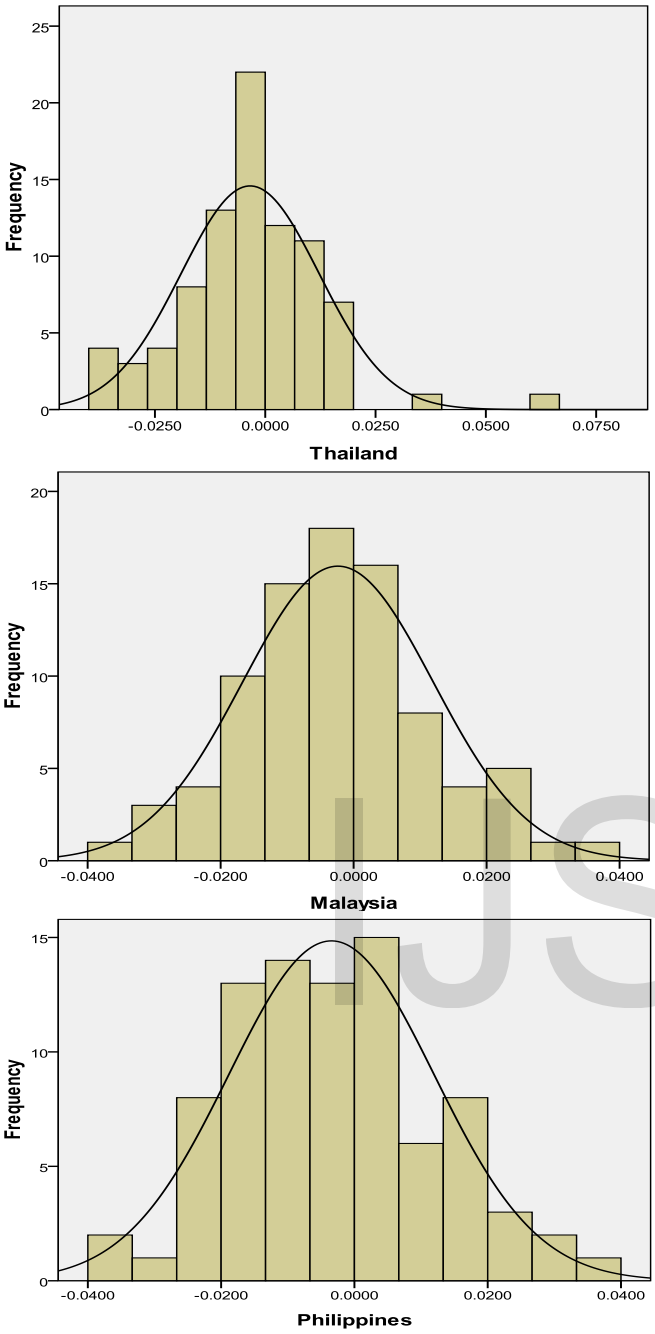


Fig. 1. Empirical Distributions of exchange return rates

By referring to the table 1 and figure 1, it can be concluded that the mean value for Malaysia is the largest in compared to Thailand and Philippines. The reason that Malaysia has highest mean value is because the value of currency in Malaysia is more valuable than others countries. However, the value of currency for Philippines is the lowest among these three countries; therefore, it has the smallest mean value.

Furthermore, table 1 is listed out all the standard deviation for the selected South Asia countries. Thailand has the largest standard deviation than others countries which indicates that it has a largest dispersion or more volatile. While Malaysia has the smallest standard deviation which illustrate that Malaysia has the smallest dispersion on the exchange return rate or less volatile than other countries.

Table 1 also listed the skewness of the distribution. A positive value is found on the skewness for all three countries which indicates that the exchange rate in these three countries can be altering by the changing of the United States currency. Moreover, all three countries are skewed to right which means that the distribution for Malaysia, Thailand and Philippines are positively skewed and this can be refer to the figure 2.

In this paper, a two-component mixture normal distribution is estimated by using maximum likelihood estimation and the normal distribution that mention in table 2 refers to recession state and growth state. The recession state in this study also known as first normal meanwhile the growth state represents the second normal. Two tables are provided which include a parameter estimates for normal model table and a mixing probabilities table with five effective parameters and the result in these tables are summarize in the table 2. Table 2 described the summary result that analysis by using maximum likelihood estimation. Meanwhile figure 2 denotes the mixture distribution graph for return of exchange rate to Malaysia, Thailand and Philippines.

TABLE II.
Maximum Likelihood Estimation for normal mixture

Country	Malaysia	
	Normal 1	Normal 2
Weight	0.9437	0.0563
Mean	-0.0041	0.0266
Variance	0.0002	0.00003

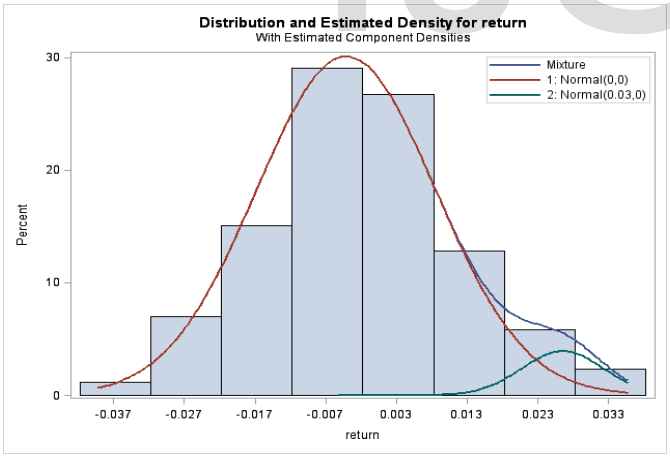
Country	Thailand	
	Normal 1	Normal 2
Weight	0.5783	0.4217
Mean	-0.00114	-0.00657
Variance	0.0001	0.0004

Country	Philippines	
	Normal 1	Normal 2
Weight	0.0737	0.9263
Mean	-0.0197	-0.0022
Variance	7.30E-7	0.0002

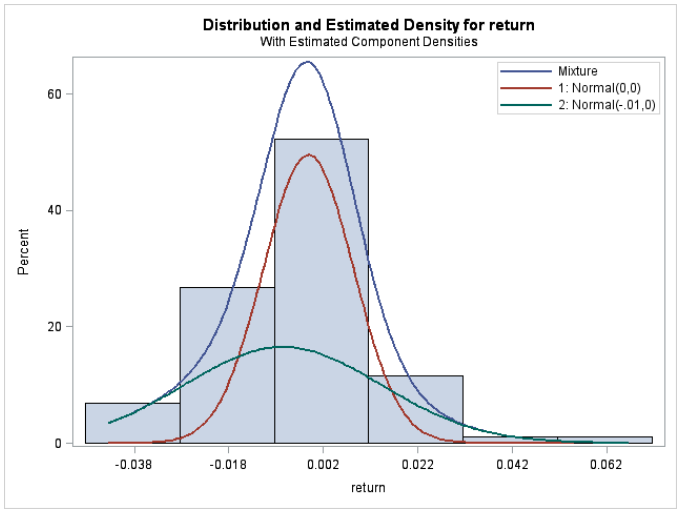
Both the table 2 and figure 2 described that the weight for Thailand and Malaysia in first normal shows that heavier than second normal and this indicates that the recession state is more dominant than growth state for Thailand and Malaysia. However, the economy in Philippines not affected by those event because the growth state in Philippines is more dominant than recession state.

Moreover, the mean for both Thailand and Philippines shows negative value for both recession state and growth state. This can be explained that the currency in Thailand is declining because the economy is not recovered from the impact of September 11 attacks. For Philippines, the negative mean value in both normal densities denotes that the currency in Philippines is declining because of the global economics crisis which beginning at 2008. However, Malaysia shows different mean value in compared to Thailand and Philippines. Malaysia has negative value in recession state and positive value in growth state. This indicates that the economy in Malaysia is recovered from the crisis.

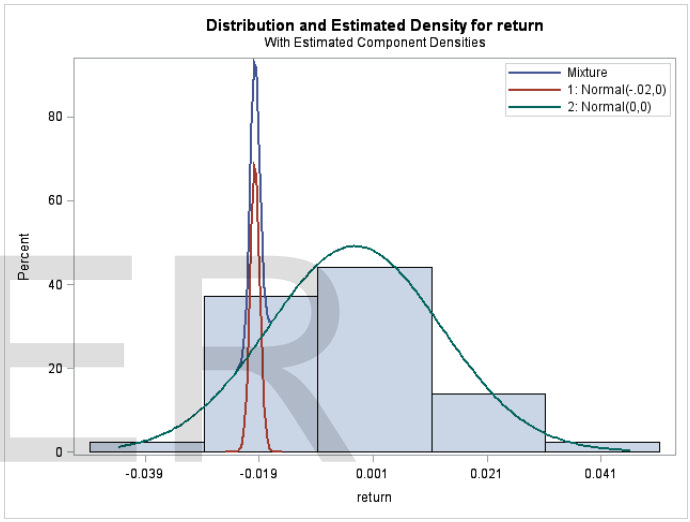
Although variance is a not necessarily tool in measure the up and down swings of the bankroll goes through but a large value of variance represents a wider swings of bankroll. Result in table 2 show that all variance for normal 1 and 2 for Thailand, Malaysia and Philippines are small thus it can be concluded that the bankroll swing in these three countries are less. This is reliable since all these countries are label as developing countries so there might have a limited or less bankroll for the country. In addition, the result is valid, reliable and significant since the variance is small.



Malaysia



Thailand



Philippines

Fig. 2. Mixture of Normal distribution

V. CONCLUSION

This paper applied maximum likelihood estimation to fit finite mixture model in order to estimate returns of nominal monthly exchange rate for Malaysia, Thailand and Philippines. A two-component mixture of normal distribution is analysing in this paper and the first normal is categorized as recession state while second normal denotes as growth state for the nominal exchange rate data. Result shows that the economy in Malaysia has recovered, but not for the economy Thailand and Philippines.

ACKNOWLEDGMENT

This paper is supported by the financial and facilities from the respective university.

REFERENCES

[1] S. Newcomb, "A generalized theory of the combination of observations so as to obtain the best result," *American Journal of Mathematics*, vol. 8, pp.343-366, 1886.

[2] K. Pearson, "Contributions to the theory of mathematics evolution," *Philosophical Transactions of the Royal Society of London*, vol. A185, pp.71-110, 1984.

[3] C. V. L. Charlier and S. D. Wicksell, "On the dissection of frequency functions," *Arkiv for Mathematics Astronomi och Fysik*, vol. 18(6), pp. 85-98, 1924.

[4] G. Doetsch, "Die elimination des dopplereffekts auf spektroskopische feinstrukturen und exakte bestimmung der komponenten," *Zeitschrift f˙r Physik*, vol. 49, pp. 705-730, 1928.

[5] R. A. Fisher, "On the Mathematical Foundations of Theoretical Statistics," *Philosophical Transactions of the Royal Society of London*, vol. 222, pp. 309-368, 1992.

[6] C. R. Rao, "The utilization of multiple measurements in problems of biological classification," *Journal of the Royal Statistical Society*, vol. B10, pp. 159-203, 1948.

[7] V. Hasselblad, "Estimation of parameters for a mixture of normal distributions," *Technometrics*, vol. 8, pp. 431-444, 1966.

[8] B. C. Peters and W. A. Coberly, "The numerical evaluation of the maximum likelihood estimate of mixture proportions," *Communications in Statistics-Theory and Methods*, vol. 5, pp. 1127-1135, 1976.

[9] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.

[10] D. W. Hosmer, "A comparison of iterative maximum likelihood estimates of the parameters of a mixture of two normal distributions under three different types of sample," *Biometrics*, vol. 29, pp. 761-770, 1973.

[11] N. J. Horton and N. M. Laird, "Maximum Likelihood Analysis of Generalized Linear Models with Missing Covariates," *Statistical Methods in Medical Research*, vol. 8, pp. 37-50, 1998.

[12] Z. H. Yang, "Maximum Likelihood Analysis Of Adaptive Evolution In Hiv-1 Gp120 *Env* GenE," *Pacific Symposium on Biocomputing*, pp. 226-237, 2001.

[13] S. Castiglioni, L. Lombardi, E. Toth, A. Castellarin, and A. Montanari, "Calibration of rainfall-runoff models in ungauged basins: A regional maximum likelihood approach," *Advances in Water Resources*, vol. 33, pp. 1235–1242, 2010.

[14] P. Owolawi, "Raindrop Size Distribution Model for the Prediction of Rain Attenuation in Durban," *Piers Online*, vol. **7**(6), pp. 516-523, 2011.