

Blood Sugar Levels of Diabetes Mellitus Patients Modeling with Bayesian Mixture Model Averaging

Ani Budi Astuti¹

*Department of Mathematics,
Statistics Study Program,
Faculty of Mathematics and Natural Sciences,
University of Brawijaya, Malang, Indonesia.
Department of Statistics,
Faculty of Mathematics and Natural Sciences,
Institut Teknologi Sepuluh Nopember,
Surabaya, Indonesia.*

Nur Iriawan, Irhamah, Heri Kuswanto

*Department of Statistics,
Faculty of Mathematics and Natural Sciences,
Institut Teknologi Sepuluh Nopember,
Surabaya, Indonesia.*

Laksmi Sasiarini

*Faculty of Medicine,
University of Brawijaya,
Malang, Indonesia.*

Abstract

Diabetes Mellitus (DM) is one of the most dangerous diseases to humans because of the effects of complications caused. According to WHO, in 2013, the total number of DM patients in Indonesia was ranked 7th in the world and this disease was ranked 6th in the world as the leading causes of human death. Bayesian Mixture Model Averaging (BMMA) is a Bayesian approach for multiple mixture models with the model parameter estimation using the averaging rule. The purpose of this

¹Corresponding Author: Ani Budi Astuti, Department of Mathematics, Statistics Study Program, Faculty of Mathematics and Natural Sciences, University of Brawijaya, Malang 65145, Indonesia and PhD Student at Department of Statistics, Faculty of Mathematics and Natural Sciences, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia. E-mail: ani_budi@ub.ac.id

study would be built the BMMA models to blood sugar levels of Diabetes Mellitus (DM) patients through simulation studies where the simulation data built on cases of blood sugar levels of DM patients in RSUD Saiful Anwar Malang. The results of this study have succeeded to built the BMMA normal models with 2 components mixture that could accommodate the real condition of the DM data with driven data concept.

AMS subject classification: Primary 68U20, Secondary 62P10.

Keywords: Mixture Model, Model Averaging, Bayesian Approach, Diabetes Mellitus.

1. Introduction

Diabetes Mellitus (DM) is one of the most dangerous diseases that are not contagious but it can cause complications of other very dangerous diseases. This disease is caused by a disruption in the body's metabolic system so that the increase in blood sugar levels (hyperglycemia). A person is said to have the disease if the blood sugar levels had reached more than 200 mg/dL for the condition is not fasting, more than 126 mg/dL for fasting conditions and the average blood sugar levels exceeding 150 mg/dL ([16], [27], [28] and [29]). According to the WHO report, DM was ranked as the world's sixth largest as the leading causes of death and in 2013, Indonesia ranked seventh in the number of patients as many as 8.5 million people [17]. The impact of the DM disease is very dangerous because it can lead to death of the patient and the number of patients continues to increase sharply over time. DM disease requires huge maintenance costs because it can cause complications of other diseases.

A study of DM data is very important, in particular on the modeling of blood sugar levels in DM patient data so that it can be known the exact model for policy making in the management of disease control and prevention of DM diseases. Challenges of the data DM is to have a mixture distribution and can be approximated by some distributions. Research of [28] indicates that the DM data has a mixture distribution and research [7] indicates that the DM data can be approached by some distributions.

Bayesian Mixture Model Averaging (BMMA) is a Bayesian solution to form the best single model from all the possible mixture models. BMMA is a combination of the two methods, the Bayesian Mixture Model (BMM) and Bayesian Model Averaging (BMA) [8]. One study on the modeling of BMM for DM data is the research that has been done by [28] and some of the BMM research in the field of microarray data, among others [10]. Whilst research on the modeling of the BMA for DM data has been carried out by [7] and the modeling of the BMA for microarray data by ([3], [4], [6], and [26]). The research on BMMA has been done by [8]. The BMMA models built on data driven concept ([5], [6], [7], [9], and [22]). In this paper would be built a model of the BMMA to blood sugar levels of Diabetes Mellitus (DM) patients through simulation studies where the simulation data built on the cases of blood sugar levels of DM patients in RSUD Saiful Anwar Malang.

2. Data and Methods

The data used in this study is simulation data which is constructed based on the cases of blood sugar levels of DM patients in RSUD Saiful Anwar Malang for Kesatrian and Blimbing villages in 2015. The number of mixture components model specified by the algorithm RJMCMC and the number of models generated by the MCMC algorithm are 1, 000 models. The best model to be established in the BMMA modeling must be appropriate with the results of Occam’s Window selection model methods. The number of simulations are 10 times for each village.

2.1. Bayesian, BMM, BMA, and BMMA Analysis

2.1.1 Bayesian Analysis

Bayesian analysis is a statistical analysis method based on the posterior probability distribution model. The posterior distribution is a blend of two information i.e., the prior information and the likelihood of the data. These concept using Bayes theorem was invented by Thomas Bayes in 1702-1761, where in model parameter, $\theta \in \Omega$, is treated as a random variable ([1], [9], [12], [20] and [21]).

If a likelihood function of observational data x is $f(x | \theta)$ and the prior θ is $p(\theta)$ then the posterior probability distribution of θ is $p(\theta | x)$, determined by the rules of probability according to Bayes’ theorem [14] as in Eq. (1)

$$p(\theta | x) = \frac{f(x | \theta)p(\theta)}{f(x)} \quad \text{where,} \quad (2.1)$$

$$f(x) = E[f(x | \theta)] = \int_{\theta \in R} f(x | \theta)f(\theta)d\theta \quad \text{if } \theta \text{ continuous and}$$

$$f(x) = E[f(x | \theta)] = \sum_{\theta \in B} f(x | \theta)p(\theta) \quad \text{if } \theta \text{ discrete,}$$

where $f(x)$ is a normalized constant [12]. So that Eq. (1) can be written as:

$$p(\theta | x) \propto f(x | \theta)p(\theta). \quad (2.2)$$

Based on the Eq. (2), it can be shown the posterior probability will be used for decision making is proportional to the product of the likelihood function and the prior probability of the model parameters ([3], [4], [6], [7], [8], and [14]).

2.1.2 BMM Analysis

Mixture Models Mixture model is a special model for the data that have sub-sub-populations or groups, where each proportion of sub-population or group is a constituent component of the mixture models. These model called the particular model because this model is able to combine some different data but still retains the characteristics of the original data ([2], [8], [12], and [24]).

According to ([8], [12], [20], and [24]), mixture probability function of an observation $\mathbf{x} = (x_1, x_2, \dots, x_n)$ taken from a number of k -subpopulation can be expressed as in Eq. (3).

$$f(\mathbf{x} | \boldsymbol{\theta}, \mathbf{w}) = \sum_{j=1}^k w_j g_j(\mathbf{x} | \theta_j), \quad (2.3)$$

where $f(\mathbf{x} | \boldsymbol{\theta}, \mathbf{w})$ is a function of the probability mixture, $g_j(\mathbf{x} | \theta_j)$ is a j^{th} probability function of k number of sub-population that make up a model, and $\boldsymbol{\theta}$ is a mixture model parameters containing of $(\theta_1, \theta_2, \dots, \theta_k)$. Parameter θ_j , $j = 1, 2, \dots, k$ represents the characteristic distribution of g_j on each component in mixture models. While \mathbf{w} is the parameter vector of proportions (weighted) mixture model containing of (w_1, w_2, \dots, w_k) , where $0 < w_j < 1, \forall j$ and $\sum_{j=1}^k w_j = 1$ for each model parameter θ_j .

BMM Models According to ([8] and [25]), to model such data into a mixture model, each observation x_i would be classified on each unknown number of sub-population. If the allocation of each observation on each sub-population in Eq. (3) is denoted by z , then the allocation of each observation $z_i, i = 1, 2, \dots, n$ could be determined based on Eq. (4).

$$p(z_i = j) = w_j, j = 1, 2, \dots, k. \quad (2.4)$$

Given the value of z_i then the observation data x_i can be derived from the sub-populations as in Eq. (5).

$$x_i | z_i \sim f(x | \theta_{z_i}), i = 1, 2, \dots, n. \quad (2.5)$$

Thus the resulting joint posterior distribution of all parameter in the mixture model can be expressed as in Eq. (6).

$$p(k, w, z, \theta, x) = p(k)p(w | k)p(z | w, k)p(\theta | z, w, k)p(x | \theta, z, w, k). \quad (2.6)$$

The next process is to estimate each parameter in Eq. (6) by employing the full conditional distribution of each parameter [25].

Suppose there is a parameter θ which has a stationary distribution of $p(\theta)$. The full conditional distribution of parameter θ is constructed by making a partition of θ as shown in Eq. (7) ([8] and [30]).

$$\theta = \theta_s, \theta_{-s}, \quad (2.7)$$

where θ_s denote the- s^{th} parameter to be estimated and θ_{-s} denote the complement of θ_s that is parameter θ without the- s^{th} component.

According to ([13] and [30]), the full conditional distribution can be established based on the joint distribution of all the parameters as in Eq. (8).

$$p(\theta_s | \theta_{-s}) = \frac{p(\theta_s, \theta_{-s})}{\int p(\theta_s, \theta_{-s}) d\theta_s} \quad (2.8)$$

RJMCMC Algorithm According to ([25] and [30]), Reversible Jump Markov Chain Monte Carlo (RJMCMC) concept can be used to determine the number of mixture components model in which the number of mixture components is unknown. In its application, RJMCMC algorithm uses the concept of birth/death and split/merge with 6 types of movement, that is w updating, θ updating, z updating, hyperparameter β updating, split/merge components, and birth/death of an empty component. The movement split/merge is a random choice between split ($k \rightarrow k + 1$) or merge ($k + 1 \rightarrow k$).

2.1.3 BMA Analysis

BMA is a Bayesian solution to model uncertainty in order to establish the best single model by considering all possible models. The completion of BMA models by averaging the posterior distribution of all the best models so that the BMA model combines the best of all possible models ([3], [4], [6], [7], [8], [15], and [18]).

If M_1, M_2, \dots, M_q is the set of possible models of M and Δ is the value would be predicted, then the BMA prediction starts with determining the prior probability distribution of all the model parameters and the model M_k ([14] and [21]). Posterior distribution of $\Delta | x$ is as in Eq. (9).

$$P(\Delta | x) = \sum_{k=1}^q P(\Delta | M_k, x)P(M_k | x), \tag{2.9}$$

where q is the sum of all the models that may be formed. Posterior distribution of $\Delta | x$ is the average of the posterior distribution whilst the posterior probability of the model M_k is:

$$P(M_k | x) = \frac{P(Y | M_k)P(M_k)}{\sum_{l=1}^q P(Y | M_l)P(M_l)} \tag{2.10}$$

where

$$P(x | M_k) = \int P(x | \theta_k, M_k)P(\theta_k | M_k)d\theta_k. \tag{2.11}$$

Eq. (11) is the marginal likelihood of the model M_k . The prior probability of $\theta_k | M_k$ is $p(\theta_k | M_k)$ and $p(x | \theta_k, M_k)$ is likelihood function and $p(M_k)$ is prior probability of the model M_k . Implicitly, all probabilities depend on the model M so the expectation value of the coefficient Δ obtained by averaging the model M as in Eq. (12).

$$E(\Delta | x) = \sum_{k=1}^q P(M_k | x)E(\Delta | M_k, x). \tag{2.12}$$

Value of $E(\Delta | x)$ in Eq. (12) shows the weighted expected value of Δ in every possible combination models whilst the variance of $(\Delta | x)$ is as in Eq. (13).

$$Var(\Delta | x) = \sum_{k=1}^q (var(\Delta | x, M_k) + (E[\Delta | M_k, x]^2)P(M_k | x) - E(\Delta | x)^2) \tag{2.13}$$

2.1.4 BMMA Analysis

BMMA is proposed method which is build by combining the two methods, BMM and BMA, wherein the model parameter estimators are obtained by averaging all of the possible mixture models. Therefore, the form of BMMA can be determined by applying Eq. (12) and Eq. (13) to the some mixture models. This finding BMMA, therefore, have the value of $E(\Delta | x)$ as the weighted value of the parameters in any combination of mixture models and $P(M_k | x)$ is the posterior distribution of the k^{th} mixture model [8]. Selection of models included in the modeling of the BMA as in Eq.(12), use Occam's Window selection methods ([3], [4], [6], [7], [15], and [23]).

2.1.5 MCMC Algorithm

In Bayesian Analysis, to obtain the posterior distribution is often very complicate and requires a difficult integration process in determining the marginal posterior of a model parameter. So that it takes a numerical approximation with Markov Chain Monte Carlo (MCMC) algorithm ([7], [11], and [19]). MCMC algorithms can be described as follows:

MCMC Algorithm with Gibbs Sampler Approach

Step 1. In this step we will provide initial value.

$$\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_r^{(0)})$$

Step 2. In this step we will do the sampling of the parameter.

Generate the value of θ_j , $j = 1, \dots, r$ from their conditional distribution as follows:

Step 2.1. Sampling $\theta_1^{(k+1)}$ from $p(\theta_1 | \mathbf{x}, \theta_2^{(k)}, \dots, \theta_r^{(k)})$

Step 2.2. Sampling $\theta_2^{(k+1)}$ from $p(\theta_2 | \mathbf{x}, \theta_1^{(k+1)}, \theta_3^{(k)}, \dots, \theta_r^{(k)})$

\vdots

Step 2.r. Sampling $\theta_r^{(k+1)}$ from $p(\theta_r | \mathbf{x}, \theta_1^{(k+1)}, \theta_2^{(k+1)}, \dots, \theta_{r-1}^{(k+1)})$

Step 3. In this step we will do iteration.

Execute step 2 as K times with $K \rightarrow \infty$.

2.1.6 Goodness of Fit Model with Kolmogorov-Smirnov (KS)

According to ([7] and [18]), the theoretical concept of Kolmogorov-Smirnov test is to compare the empirical cumulative distribution function (CDF), $F_n(x_i)$, and the hypothesis cumulative distribution function, \hat{F} . If $X_{(1)}, X_{(2)}, \dots, X_{(n)}$, is statistic order to the independent random variable with hypothesis distribution, \hat{F} and the empirical distribution is as in Eq. (14).

$$F_n(x_i) = \frac{\text{the number of data } X_i \leq x_i}{n} \quad \text{for } i = 1, 2, 3, \dots, k < n, \quad (2.14)$$

where $F_n(x_i)$ is right continuous step function.

The formula of test statistic D_n can be written as in Eq. (15).

$$D_n = \sup \left\{ |F_n(x) - \hat{F}| \right\} \quad (2.15)$$

If the value of D_n smaller then the tested models is better.

3. Results and Discussions

3.1. Description and Distribution of the Data from the Kesatrian and Blimbing Villages

The description and distribution results for the blood sugar levels of DM patients data in Saiful Anwar Malang Hospital for Kesatrian and Blimbing villages in 2015 can be seen that in the Figure 1 and Figure 2.

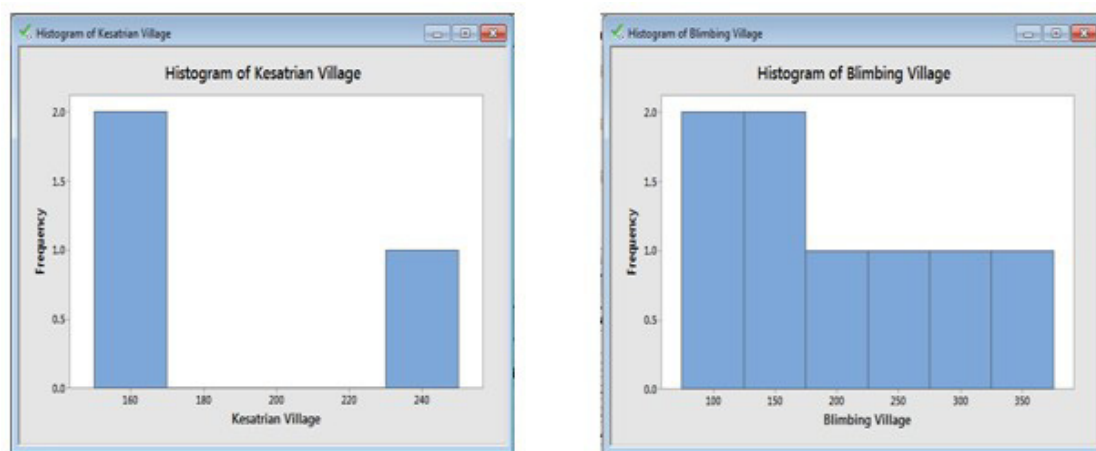


Figure 1: Histogram for the blood sugar levels of DM patients data for Kesatrian and Blimbing villages.

Derived from Figure 1 can be known that data are indicated to have a mixture distribution with 2 components, both for the Kesatrian and Blimbing villages. Meanwhile, based on Figure 2 can be seen that the data have a normal distribution for the Kesatrian and Blimbing villages. It can be seen by the p-value of the normality tests for data greater than 0.05, which means accept the null hypothesis that the data has a normal distribution.

3.2. Identification for the Number of Components Mixture by RJMCMC Algorithm with Mixture Normal Distribution

The identification results for the number of components mixture in the normal mixture distribution by RJMCMC algorithm can be seen that in the Table 1 and Table 2.

Based on Table 1, for simulation of Kesatrian village, it can be seen that the best model is the mixture normal distributions with 2 components mixture because it has the

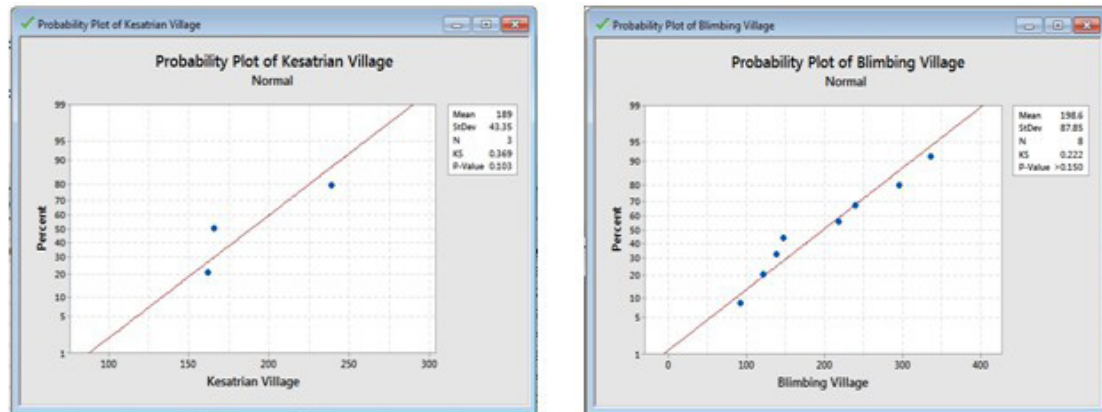


Figure 2: Normal Distribution Test for the blood sugar levels of DM patients data for Kesatrian and Blimbing villages.

Table 1: Results of RJMCMC Algorithm for Kesatrian and Simulation of Kesatrian Villages.

No.	Villages	Result		
		1	2	3
1	Kesatrian	0.2798	0.4909	0.2292
2	Kesatrian Simulation 1	0.2649	0.4962	0.2389
3	Kesatrian Simulation 2	0.2581	0.5004	0.2416
4	Kesatrian Simulation 3	0.2323	0.5125	0.2552
5	Kesatrian Simulation 4	0.1828	0.522	0.2952
6	Kesatrian Simulation 5	0.1942	0.5211	0.2847
7	Kesatrian Simulation 6	0.1597	0.5282	0.312
8	Kesatrian Simulation 7	0.2501	0.5012	0.2487
9	Kesatrian Simulation 8	0.2612	0.4975	0.2413
10	Kesatrian Simulation 9	0.1672	0.5285	0.3043
11	Kesatrian Simulation 10	0.27	0.4946	0.2354
Average of Kesatrian Simulation		0.22405	0.51022	0.26573

Table 2: Results of RJMCMC Algorithm for Blimbing and Simulation of Blimbing Villages.

No.	Villages	Result of RJMCMC Test							
		1	2	3	4	5	6	7	8
1	Blimbing	0.1692	0.3984	0.2154	0.103	0.0551	0.032	0.0201	0.0067
2	Blimbing Simulation 1	0.1479	0.3982	0.2219	0.1092	0.0588	0.0347	0.0221	0.0072
3	Blimbing Simulation 2	0.1472	0.4002	0.2204	0.1094	0.0595	0.0345	0.0217	0.0071
4	Blimbing Simulation 3	0.1557	0.4028	0.2188	0.1075	0.0567	0.0322	0.0199	0.0063
5	Blimbing Simulation 4	0.1169	0.3708	0.2264	0.1264	0.0729	0.0456	0.0305	0.0104
6	Blimbing Simulation 5	0.1689	0.393	0.213	0.1033	0.0564	0.0347	0.0228	0.0079
7	Blimbing Simulation 6	0.1701	0.4005	0.212	0.1018	0.0542	0.0319	0.022	0.0074
8	Blimbing Simulation 7	0.1365	0.3759	0.2177	0.1146	0.0679	0.0441	0.317	0.0115
9	Blimbing Simulation 8	0.1247	0.3644	0.2181	0.1188	0.0721	0.0489	0.0384	0.0147
10	Blimbing Simulation 9	0.154	0.3903	0.216	0.108	0.059	0.0379	0.0258	0.009
11	Blimbing Simulation 10	0.1265	0.3736	0.2206	0.1216	0.0709	0.0448	0.031	0.0112
Average of Blimbing Simulation		0.14484	0.387	0.21849	0.1121	0.0628	0.0389	0.0266	0.0093

average value probability of RJMCMC test the biggest, that is 0.51022. Meanwhile, according to Table 2, for Blimbing village, it can be seen also that the mixture normal distributions with 2 components mixture is the best model because it has the average value probability of RJMCMC test the biggest, that is 0.387. Therefore, in modeling BMMA will be established mixture normal model that contains 2 components mixture.

3.3. The BMMA Normal Models with 2 Components Mixture and the Goodness of Fit Models by KS

The BMMA normal models with 2 components mixture and the goodness of fit models by KS for simulation of Kesatrian and Blimbing villages can be seen that in the Table 3 and Table 4.

Based on Table 3 and Table 4 as a result of the simulation data, it can be seen that the average for the number of component models in the BMMA normal models with 2 components mixture is 13 models and 10 models from the 1,000 models have been generated. Based on this model can be shown that there are two groups of DM patients in Kesatrian and Blimbing villages. Furthermore, the goodness of fit from the BMMA normal models with 2 components is measured based on the Kolmogorov-Smirnov (KS) value. The average KS value is 0.605446154 for simulation of Kesatrian village (Table 3) and the average KS value is 0.43951321 for simulation of Blimbing village (Table 4). Based on both the value of small KS can be seen that the BMMA model with 2 components mixture is very good models because if the value of KS small then there is a suitability between a model built with a model of the real data. The BMMA normal models with 2 components mixture and the number of components model as the result of Occam’s Window selection for each simulation data can be described as in Equation (A1) to Equation (A11) for Kesatrian and simulation of Kesatrian villages and then

Table 3: The BMMA Normal Models with 2 Components Mixture and the Goodness of Fit Models by KS.

No.	Villages	The Number of Generation Models in the BMMA	The Number of Component Models in the BMMA	The BMMA Models	The Fitted Models by KS
1	Kesatrian	1,000	14	Equation (A1)	0.614563927
2	Kesatrian Simulation 1	1,000	13	Equation (A2)	0.613438034
3	Kesatrian Simulation 2	1,000	13	Equation (A3)	0.613438034
4	Kesatrian Simulation 3	1,000	14	Equation (A4)	0.592135119
5	Kesatrian Simulation 4	1,000	13	Equation (A5)	0.613737704
6	Kesatrian Simulation 5	1,000	13	Equation (A6)	0.613615432
7	Kesatrian Simulation 6	1,000	13	Equation (A7)	0.614415823
8	Kesatrian Simulation 7	1,000	14	Equation (A8)	0.560690951
9	Kesatrian Simulation 8	1,000	13	Equation (A9)	0.613234424
10	Kesatrian Simulation 9	1,000	14	Equation (A10)	0.610930429
11	Kesatrian Simulation 10	1,000	13	Equation (A11)	0.608825589
Average of Kesatrian Simulation		1,000	13.3		0.605446154

Table 4: The BMMA Normal Models with 2 Components Mixture and the Goodness of Fit Models by KS.

No.	Villages	The Number of Generation Models in the BMMA	The Number of Component Models in the BMMA	The BMMA Models	The Fitted Models by KS
1	Blimbing	1,000	18	Equation (B1)	0.30890336
2	Blimbing Simulation 1	1,000	9	Equation (B2)	0.47746443
3	Blimbing Simulation 2	1,000	6	Equation (B3)	0.49178101
4	Blimbing Simulation 3	1,000	16	Equation (B4)	0.56590676
5	Blimbing Simulation 4	1,000	7	Equation (B5)	0.57912561
6	Blimbing Simulation 5	1,000	7	Equation (B6)	0.19888206
7	Blimbing Simulation 6	1,000	7	Equation (B7)	0.34575018
8	Blimbing Simulation 7	1,000	9	Equation (B8)	0.4126112
9	Blimbing Simulation 8	1,000	8	Equation (B9)	0.51248193
10	Blimbing Simulation 9	1,000	17	Equation (B10)	0.45335265
11	Blimbing Simulation 10	1,000	9	Equation (B11)	0.35777624
Average of Blimbing Simulation		1,000	9.5		0.43951321

Equation (B1) to Equation (B11) for Blimbing and simulation of Blimbing villages (Appendix A).

4. Conclusion

The BMMA models as a result of the simulation data for cases of the blood sugar levels of DM patients in RSUD Saiful Anwar Malang of Kesatrian and Blimbing villages have mixture normal distribution with 2 components mixture and have the average for the number of component models at 13 and 10 models from the 1,000 models have been generated, respectively. For Kesatrian and Blimbing villages, based on this model and information from Figure 1 can be seen that there are two groups of DM patients, namely DM patients with average blood sugar levels less than 200 mg/dL and greater than or equal to 200 mg/dL with the proportion of the first group is greater than the second group, respectively. The goodness of fit from the BMMA normal model with 2 components mixture by KS with the average KS =0.605446154 for Kesatrian village and the average KS=0.43951321 for Blimbing village. Based on the value of small KS, the BMMA models for cases of the blood sugar levels of DM patients in RSUD Saiful Anwar Malang for Kesatrian and Blimbing villages are very good models and those models able to accommodate the real condition of the DM data with driven data concept.

Acknowledgements

This paper is part of research Grants PUPT/BOPTN University of Brawijaya in 2015 and the part of Doctoral Research at Department of Statistics, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia . We would like the first thank to Directorate General of Higher Education and University of Brawijaya which have financed this research, and the second thank to RSUD Saiful Anwar Malang for allowing use of research data and thank to anonymous reviewer to this paper.

Appendix A. The BMMA Normal Models with 2 Components Mixture and the Number of Components Model as the Result of Occam’s Window Selection for Real and Simulation Data

$$\begin{aligned}
 f_{BMMA}(A1) = & \left[w_{A1(1)(1)} N_{A1(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A1(1)(2)} N_{A1(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\
 & + \dots + \left[w_{A1(14)(1)} N_{A1(14)(1)} \left(\hat{\mu}_{(14)(1)}, \hat{\sigma}_{(14)(1)}^2 \right) + w_{A1(14)(2)} N_{A1(14)(2)} \right. \\
 & \left. \left(\hat{\mu}_{(14)(2)}, \hat{\sigma}_{(14)(2)}^2 \right) \right] \tag{A1}
 \end{aligned}$$

$$\begin{aligned}
f_{BMM A}(A2) = & \left[w_{A2(1)(1)} N_{A2(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A2(1)(2)} N_{A2(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\
& + \cdots + \left[w_{A2(13)(1)} N_{A2(13)(1)} \left(\hat{\mu}_{(13)(1)}, \hat{\sigma}_{(13)(1)}^2 \right) + w_{A2(13)(2)} N_{A2(13)(2)} \right. \\
& \left. \left(\hat{\mu}_{(13)(2)}, \hat{\sigma}_{(13)(2)}^2 \right) \right] \tag{A2}
\end{aligned}$$

$$\begin{aligned}
f_{BMM A}(A3) = & \left[w_{A3(1)(1)} N_{A3(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A3(1)(2)} N_{A3(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\
& + \cdots + \left[w_{A3(13)(1)} N_{A3(13)(1)} \left(\hat{\mu}_{(13)(1)}, \hat{\sigma}_{(13)(1)}^2 \right) + w_{A3(13)(2)} N_{A3(13)(2)} \right. \\
& \left. \left(\hat{\mu}_{(13)(2)}, \hat{\sigma}_{(13)(2)}^2 \right) \right] \tag{A3}
\end{aligned}$$

$$\begin{aligned}
f_{BMM A}(A4) = & \left[w_{A4(1)(1)} N_{A4(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A4(1)(2)} N_{A4(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\
& + \cdots + \left[w_{A4(14)(1)} N_{A4(14)(1)} \left(\hat{\mu}_{(14)(1)}, \hat{\sigma}_{(14)(1)}^2 \right) + w_{A4(14)(2)} N_{A4(14)(2)} \right. \\
& \left. \left(\hat{\mu}_{(14)(2)}, \hat{\sigma}_{(14)(2)}^2 \right) \right] \tag{A4}
\end{aligned}$$

$$\begin{aligned}
f_{BMM A}(A5) = & \left[w_{A5(1)(1)} N_{A5(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A5(1)(2)} N_{A5(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\
& + \cdots + \left[w_{A5(13)(1)} N_{A5(13)(1)} \left(\hat{\mu}_{(13)(1)}, \hat{\sigma}_{(13)(1)}^2 \right) + w_{A5(13)(2)} N_{A5(13)(2)} \right. \\
& \left. \left(\hat{\mu}_{(13)(2)}, \hat{\sigma}_{(13)(2)}^2 \right) \right] \tag{A5}
\end{aligned}$$

$$\begin{aligned}
f_{BMM A}(A6) = & \left[w_{A6(1)(1)} N_{A6(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A6(1)(2)} N_{A6(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\
& + \cdots + \left[w_{A6(13)(1)} N_{A6(13)(1)} \left(\hat{\mu}_{(13)(1)}, \hat{\sigma}_{(13)(1)}^2 \right) + w_{A6(13)(2)} N_{A6(13)(2)} \right. \\
& \left. \left(\hat{\mu}_{(13)(2)}, \hat{\sigma}_{(13)(2)}^2 \right) \right] \tag{A6}
\end{aligned}$$

$$\begin{aligned}
f_{BMM A}(A7) = & \left[w_{A7(1)(1)} N_{A7(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A7(1)(2)} N_{A7(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\
& + \cdots + \left[w_{A7(13)(1)} N_{A7(13)(1)} \left(\hat{\mu}_{(13)(1)}, \hat{\sigma}_{(13)(1)}^2 \right) + w_{A7(13)(2)} N_{A7(13)(2)} \right. \\
& \left. \left(\hat{\mu}_{(13)(2)}, \hat{\sigma}_{(13)(2)}^2 \right) \right] \tag{A7}
\end{aligned}$$

$$\begin{aligned}
f_{BMM A}(A8) = & \left[w_{A8(1)(1)} N_{A8(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A8(1)(2)} N_{A8(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\
& + \cdots + \left[w_{A8(14)(1)} N_{A8(14)(1)} \left(\hat{\mu}_{(14)(1)}, \hat{\sigma}_{(14)(1)}^2 \right) + w_{A8(14)(2)} N_{A8(14)(2)} \right. \\
& \left. \left(\hat{\mu}_{(14)(2)}, \hat{\sigma}_{(14)(2)}^2 \right) \right] \tag{A8}
\end{aligned}$$

$$\begin{aligned}
f_{BMM A}(A9) = & \left[w_{A9(1)(1)} N_{A9(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A9(1)(2)} N_{A9(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\
& + \cdots + \left[w_{A9(13)(1)} N_{A9(13)(1)} \left(\hat{\mu}_{(13)(1)}, \hat{\sigma}_{(13)(1)}^2 \right) + w_{A9(13)(2)} N_{A9(13)(2)} \right. \\
& \left. \left(\hat{\mu}_{(13)(2)}, \hat{\sigma}_{(13)(2)}^2 \right) \right] \tag{A9}
\end{aligned}$$

$$f_{BMM A}(A10) = \left[w_{A10(1)(1)} N_{A10(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A10(1)(2)} N_{A10(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \quad (4.16)$$

$$+ \cdots + \left[w_{A10(14)(1)} N_{A10(14)(1)} \left(\hat{\mu}_{(14)(1)}, \hat{\sigma}_{(14)(1)}^2 \right) + w_{A10(14)(2)} N_{A10(14)(2)} \left(\hat{\mu}_{(14)(2)}, \hat{\sigma}_{(14)(2)}^2 \right) \right] \quad (A10)$$

$$f_{BMM A}(A11) = \left[w_{A11(1)(1)} N_{A11(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{A11(1)(2)} N_{A11(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{A11(13)(1)} N_{A11(13)(1)} \left(\hat{\mu}_{(13)(1)}, \hat{\sigma}_{(13)(1)}^2 \right) + w_{A11(13)(2)} N_{A11(13)(2)} \left(\hat{\mu}_{(13)(2)}, \hat{\sigma}_{(13)(2)}^2 \right) \right] \quad (A11)$$

$$f_{BMM A}(B1) = \left[w_{B1(1)(1)} N_{B1(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B1(1)(2)} N_{B1(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B1(18)(1)} N_{B1(18)(1)} \left(\hat{\mu}_{(18)(1)}, \hat{\sigma}_{(18)(1)}^2 \right) + w_{B1(18)(2)} N_{B1(18)(2)} \left(\hat{\mu}_{(18)(2)}, \hat{\sigma}_{(18)(2)}^2 \right) \right] \quad (B1)$$

$$f_{BMM A}(B2) = \left[w_{B2(1)(1)} N_{B2(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B2(1)(2)} N_{B2(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B2(9)(1)} N_{B2(9)(1)} \left(\hat{\mu}_{(9)(1)}, \hat{\sigma}_{(9)(1)}^2 \right) + w_{B2(9)(2)} N_{B2(9)(2)} \left(\hat{\mu}_{(9)(2)}, \hat{\sigma}_{(9)(2)}^2 \right) \right] \quad (B2)$$

$$f_{BMM A}(B3) = \left[w_{B3(1)(1)} N_{B3(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B3(1)(2)} N_{B3(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B3(6)(1)} N_{B3(6)(1)} \left(\hat{\mu}_{(6)(1)}, \hat{\sigma}_{(6)(1)}^2 \right) + w_{B3(6)(2)} N_{B3(6)(2)} \left(\hat{\mu}_{(6)(2)}, \hat{\sigma}_{(6)(2)}^2 \right) \right] \quad (B3)$$

$$f_{BMM A}(B4) = \left[w_{B4(1)(1)} N_{B4(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B4(1)(2)} N_{B4(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B4(16)(1)} N_{B4(16)(1)} \left(\hat{\mu}_{(16)(1)}, \hat{\sigma}_{(16)(1)}^2 \right) + w_{B4(16)(2)} N_{B4(16)(2)} \left(\hat{\mu}_{(16)(2)}, \hat{\sigma}_{(16)(2)}^2 \right) \right] \quad (B4)$$

$$f_{BMM A}(B5) = \left[w_{B5(1)(1)} N_{B5(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B5(1)(2)} N_{B5(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B5(7)(1)} N_{B5(7)(1)} \left(\hat{\mu}_{(7)(1)}, \hat{\sigma}_{(7)(1)}^2 \right) + w_{B5(7)(2)} N_{B5(7)(2)} \left(\hat{\mu}_{(7)(2)}, \hat{\sigma}_{(7)(2)}^2 \right) \right] \quad (B5)$$

$$f_{BMM A}(B6) = \left[w_{B6(1)(1)} N_{B6(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B6(1)(2)} N_{B6(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B6(7)(1)} N_{B6(7)(1)} \left(\hat{\mu}_{(7)(1)}, \hat{\sigma}_{(7)(1)}^2 \right) + w_{B6(7)(2)} N_{B6(7)(2)} \left(\hat{\mu}_{(7)(2)}, \hat{\sigma}_{(7)(2)}^2 \right) \right] \quad (B6)$$

$$f_{BMM A}(B7) = \left[w_{B7(1)(1)} N_{B7(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B7(1)(2)} N_{B7(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B7(7)(1)} N_{B7(7)(1)} \left(\hat{\mu}_{(7)(1)}, \hat{\sigma}_{(7)(1)}^2 \right) + w_{B7(7)(2)} N_{B7(7)(2)} \left(\hat{\mu}_{(7)(2)}, \hat{\sigma}_{(7)(2)}^2 \right) \right] \quad (B7)$$

$$f_{BMM A}(B8) = \left[w_{B8(1)(1)} N_{B8(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B8(1)(2)} N_{B8(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B8(9)(1)} N_{B8(9)(1)} \left(\hat{\mu}_{(9)(1)}, \hat{\sigma}_{(9)(1)}^2 \right) + w_{B8(9)(2)} N_{B8(9)(2)} \right. \\ \left. \left(\hat{\mu}_{(9)(2)}, \hat{\sigma}_{(9)(2)}^2 \right) \right] \quad (B8)$$

$$f_{BMM A}(B9) = \left[w_{B9(1)(1)} N_{B9(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B9(1)(2)} N_{B9(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B9(8)(1)} N_{B9(8)(1)} \left(\hat{\mu}_{(8)(1)}, \hat{\sigma}_{(8)(1)}^2 \right) + w_{B9(8)(2)} N_{B9(8)(2)} \right. \\ \left. \left(\hat{\mu}_{(8)(2)}, \hat{\sigma}_{(8)(2)}^2 \right) \right] \quad (B9)$$

$$f_{BMM A}(B10) = \left[w_{B10(1)(1)} N_{B10(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B10(1)(2)} N_{B10(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B10(17)(1)} N_{B10(17)(1)} \left(\hat{\mu}_{(17)(1)}, \hat{\sigma}_{(17)(1)}^2 \right) + w_{B10(17)(2)} N_{B10(17)(2)} \right. \\ \left. \left(\hat{\mu}_{(17)(2)}, \hat{\sigma}_{(17)(2)}^2 \right) \right] \quad (B10)$$

$$f_{BMM A}(B11) = \left[w_{B11(1)(1)} N_{B11(1)(1)} \left(\hat{\mu}_{(1)(1)}, \hat{\sigma}_{(1)(1)}^2 \right) + w_{B11(1)(2)} N_{B11(1)(2)} \left(\hat{\mu}_{(1)(2)}, \hat{\sigma}_{(1)(2)}^2 \right) \right] \\ + \cdots + \left[w_{B11(9)(1)} N_{B11(9)(1)} \left(\hat{\mu}_{(9)(1)}, \hat{\sigma}_{(9)(1)}^2 \right) + w_{B11(9)(2)} N_{B11(9)(2)} \right. \\ \left. \left(\hat{\mu}_{(9)(2)}, \hat{\sigma}_{(9)(2)}^2 \right) \right] \quad (B11)$$

References

- [1] M. Aitkin, Likelihood and Bayesian Analysis of Mixtures. *Statistical Modelling*, 1 (2001): 287–304, DOI:10.1177/1471082X0100100404.
- [2] E. Y. Astuti, Implementation of the Bayesian Markov Chain Monte Carlo in Optimal Portfolio Modeling with Mixture Model Approach of Multiple Mixture. (in Indonesian). *Master Thesis*, (2006), Statistics Department, ITS, Surabaya.
- [3] A. B. Astuti, N. Iriawan, Irhamah, and H. Kuswanto, Model Components Selection in Bayesian Model Averaging using Occam's Window for Microarray Data, *Proceedings of 4th Annual Basic Science International Conference 2014 in conjunction with The 5th International Conference on Global Resource Conservation 2014*, (2014a), FMIPA, UB, Malang.
- [4] A. B. Astuti, N. Iriawan, Irhamah, and H. Kuswanto, Kolmogorov-Smirnov and Continuous Ranked Probability Score Validation on the Bayesian Model Averaging for Microarray Data, *Applied Mathematical Sciences*, 146 (8) (2014b): 7277–7287. DOI: <http://dx.doi.org/10.12988/ams.2014.49760>.
- [5] A. B. Astuti, Cox Regression Model for Lifetime of a Newborn Baby (Neonates) in Dr. Saiful Anwar Malang Hospital. *Global Journal of Pure and Applied Mathematics*, 10 (4) (2014c): 591–600. ISSN 0973–1768.
- [6] A. B. Astuti, N. Iriawan, Irhamah, and H. Kuswanto, Occam's Window Selection in Bayesian Model Averaging Modeling for Gene Expression Data from

- Chickpea Plant, *International Journal of Applied Mathematics and Statistics*, 53 (4) (2015a): 160–165. ISSN 0973–1377 (Print), ISSN 0973-7545 (Online).
- [7] A. B. Astuti, N. Iriawan, and L. Sasirini, Bayesian Model Averaging with Occam's Window Selection Components Model on Diabetes Mellitus Data, (in Indonesia), SEMIRATA2015 MIPA BKS-PTN West, (2015b), University of Tanjungpura, Pontianak.
- [8] A. B. Astuti, N. Iriawan, Irhamah, and H. Kuswanto, An Algorithm for Determining the Number of Mixture Components on the Bayesian Mixture Model Averaging for Microarray Data, *Journal of Mathematics and Statistics*, 11 (2) (2015c): 45–51. DOI: 10.3844/jmssp.2015.45.51.
- [9] G. E. P. Box, and Tiao, *Bayesian Inference in Statistical Analysis*. MA: Addison-Wesley, Massachusetts, (1973).
- [10] K. A. Do, P. Muller, and F. Tang, A Bayesian Mixture Model For Differential Gene Expression, *Research in Department of Biostatistics*, (2005), The University of Texas.
- [11] D. Gamerman, *Markov Chain Monte Carlo*. Chapman & Hall, London, (1997).
- [12] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*. Chapman & Hall, London, (1995).
- [13] W.R. Gilks, *Full Conditional Distribution*, in *Markov Chain Monte Carlo in Practice*. (eds). Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. Chapman & Hall, London, pp: 75–88, (1995).
- [14] J. K. Gosh, M. Delampady, and T. Samanta, *An Introduction to Bayesian Analysis Theory and Method*, Springer, New York, (2006).
- [15] J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky, Bayesian Model Averaging: A Tutorial (with Discussion), *Statistical Science*, **14** (1999), 382–401.
- [16] <http://medicastore.com/diabetes/terapidia>, Treatment of Diabetes Mellitus. (in Indonesian).
- [17] <http://www.tempo.co/read/news/2013/11/14/060529766/> Indonesian-Log-10-Big-State-with- Diabetes-Sufferers, (in Indonesian).
- [18] V. F. Hustianda, Comparison of Bayesian Model Averaging and Multiple Linear Regression in Predicting Factors Affecting Number of Infant Death in East Java, (in Indonesian), *Master Thesis*, (2012), Statistics Department, ITS, Surabaya.
- [19] N. Iriawan, Computationally Intensive Approaches to Inference in Neo-Normal Linear Models, *Ph.D. Thesis*, (2000), CUT-Australia.
- [20] N. Iriawan, Univariable Normal Mixture Model Estimation: A Bayesian Methods Approach with MCMC. (in Indonesian). *Proceedings of National Seminar and Konferda VII on Mathematics in DIY & Central Java Region*, (2001), Yogyakarta, pp: 105–110.

- [21] N. Iriawan, Simulation Technique, (in Indonesian), *Teaching Modules*, ITS, Surabaya, (2003).
- [22] N. Iriawan, Modeling and Analysis of Data-Driven, (in Indonesian), Volume I, ITS, Surabaya, (2012).
- [23] D. Madigan, and A. E. Raftery, Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam's Window, *Journal of the American Statistical Association*, **89** (1994), 1535–1546.
- [24] G. J. McLachlan and K. E. Basford, *Mixture Models Inference and Applications to Clustering*. Marcel Dekker, New York, (1988).
- [25] S. Richardson and P. J. Green, On Bayesian Analysis of Mixtures with an Unknown Number of Components (with discussion), *Journal of the Royal Statistical Society*, 59 (4) (1997): 731–792, DOI: 10.1111/1467-9868.00095.
- [26] P. Sebastiani, H. Xie, and M. F. Ramoni, Bayesian Analysis of Comparative Microarray Experiments by Model Averaging, *International Society for Bayesian Analysis*, **1** (2006), 707–732.
- [27] S. Smeltzer and B. Bare, *Textbook of Medical–Surgical Nursing*, Eleventh Edition, Lippincott, (2008).
- [28] B. Suharjo, Implementation of Bayesian Markov Chain Monte Carlo in Normal Mixture Distributions, (in Indonesian), *Master Thesis*, (2003), Statistics Department, ITS, Surabaya.
- [29] S. Suyono, *Textbook of Medicine*, (in Indonesian), Volume III, Edition IV, FKUI, Jakarta, (2006).
- [30] D. M. Wati, Analysis of Bayesian Markov Chain Monte Carlo in Normal Mixture Modeling with Unknown Number of Components (Case Study: Rate of Inflation in Indonesia). (in Indonesian). *Master Thesis*, (2006), Statistics Department, ITS, Surabaya.