Bayesian mixture modeling for blood sugar levels of diabetes mellitus patients (case study in RSUD Saiful Anwar Malang Indonesia)

Ani Budi Astuti1, Nur Iriawan2, Irhamah2, Heri Kuswanto2, and Laksmi Sasiarini3

1 Department of Mathematics, Statistics Study Program, Faculty of Mathematics and Natural Sciences, University of Brawijaya Malang and Department of Statistics, Faculty of Mathematics and Natural Sciences, Institut Teknologi Sepuluh Nopember Surabaya
2 Department of Statistics, Faculty of Mathematics and Natural Sciences, Institut Teknologi Sepuluh Nopember Surabaya
3 Faculty of Medicine, University of Brawijaya Malang

E-mail: ani.budi@ub.ac.id

Abstract. Bayesian statistics proposes an approach that is very flexible in the number of samples and distribution of data. Bayesian Mixture Model (BMM) is a Bayesian approach for multimodal models. Diabetes Mellitus (DM) is more commonly known in the Indonesian community as sweet pee. This disease is one type of chronic non-communicable diseases but it is very dangerous to humans because of the effects of other diseases complications caused. WHO reports in 2013 showed DM disease was ranked 6th in the world as the leading causes of human death. In Indonesia, DM disease continues to increase over time. These research would be studied patterns and would be built the BMM models of the DM data through simulation studies where the simulation data built on cases of blood sugar levels of DM patients in RSUD Saiful Anwar Malang. The results have been successfully demonstrated pattern of distribution of the DM data which has a normal mixture distribution. The BMM models have succeed to accommodate the real condition of the DM data based on the data driven concept.

1. Introduction

When a set of real data observed have a very small amount and a multimodal distribution then it is necessary need specially handling. The Bayesian statistics proposes an approach that is very flexible in the number of samples and distribution of data. These approach is based directly on the posterior distribution of the data. Therefore Bayesian approach can accommodate the real condition of the data based on the data driven concept ([1], [2], [3], [4], and [5]). Bayesian Mixture Model (BMM) is one model that uses a Bayesian approach for multimodal models and its parameter models is seen as random variables in the model parameter space [6].

Various studies have previously been carried out by BMM modeling. The modeling of BMM for the DM data in Surabaya is the research that has been done by [7] and the research of [8] has been conducted for the DM data in Malang by BMMA modeling. Some of the BMM research in the other field, among others [9] and [10]. Modeling of cases is very important in order to know the exact model and can be used for decision making in accordance with the cases observed [11].
One disease that is not contagious but it is very dangerous and deadly to humans is Diabetes Mellitus (DM). These diseases can cause complications of other very dangerous diseases so the handling costs are very expensive. According to the WHO report in 2013, DM was ranked as the world’s sixth largest as the leading causes of death [12]. 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. 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 ([7], [13], and [14]).

A study of the blood sugar levels in DM patients data modeling is very important because it can be known the exact model for policy making in the management of disease control and prevention of DM diseases. Challenge of the DM data is to have a mixture distribution. The studies of [8] and [15] indicates that the DM data have a mixture normal distribution. These paper would be studied patterns of DM data and would be built the BMM models of the DM data through simulation studies where the simulation data built on cases of blood sugar levels of DM patients in RSUD Saiful Anwar Malang.

2. Materials 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 Lowokwaru District in 2015. In this study, there is 7 villages of Lowokwaru District analyzed. To construct simulation data, we need information on real data distribution and value ranges of data from each village in the Lowokwaru District. Based on the information that has been obtained, the data generation is done as much as the number of real data and is repeated 100 times. The detection of the number of mixture components using RJMCMC algorithms and to build BMM models using MCMC with Gibbs Sampler approach by Winbugs program. Finally, to detect feasibility of this model with the Kolmogorov-Smirnov criterion. In this study hypothesized that normal mixture models (BMM-Normal) proposed as a model for the DM cases in Malang. In building a BMM-Normal model using the R software. The following are descriptions of the DM data DM in the Lowokwaru district of Malang.

3. Bayesian Analysis, BMM Analysis, and Kolmogorov-Smirnov
3.1. Bayesian Analysis
The posterior probability distribution model has become the basis of Bayesian statistical analysis method. These distribution is a blend of the prior information and the likelihood of the data. Thomas Bayes had found this concept in 1702-1761, where in model parameter, \( \theta \in \Omega \), is treated as a random variable ([16], [8], [16], [17], [18] and [19]).

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

\[
p(\theta \mid x) = \frac{f(x \mid \theta)p(\theta)}{f(x)} \quad \text{where,}
\]

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

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

where \( f(x) \) is a normalized constant [6]. So that 1 can be written as:

\[
p(\theta \mid x) \propto f(x \mid \theta)p(\theta).
\]
Derived from the Equation 2, the posterior probability that will be used for decision making is proportional to the product of the likelihood function and the prior probability of the model parameters ([1], [20], [21], [22], [23], and [24]).

### 3.2. BMM Analysis

#### 3.2.1. Mixture Models

Mixture model is a special model for the data that have multimodal distribution, where each proportion of sub-group is a constituent component of the mixture models. The mixture model called the particular model because the model is able to combine some different data but still retains the characteristics of the original data ([6], [8], [24], [25], and [26]).

Mixture probability function of an observation \( \mathbf{x} = (x_1, x_2, \ldots, x_n) \) taken from a number of \( k \)-subgroup can be expressed as in Equation 3 ([6], [8], [18], [24], and [26]).

\[
f(\mathbf{x} | \theta, \mathbf{w}) = \sum_{j=1}^{k} w_j g_j(\mathbf{x} | \theta_j),
\]

where \( f(\mathbf{x} | \theta, \mathbf{w}) \) is a function of the probability mixture, \( g_j(\mathbf{x} | \theta_j) \) is a \( j^{th} \) probability function of \( k \) number of sub-group, and \( \theta \) is a mixture model parameters containing of \( (\theta_1, \theta_2, \ldots, \theta_k) \). Parameter \( \theta_j, j = 1, 2, \ldots, k \) 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 \( (w_1, w_2, \ldots, w_k) \), where \( 0 < w_j < 1, \forall j \) and

\[
\sum_{j=1}^{k} w_j = 1 \text{ for each model parameter} \theta_j.
\]

#### 3.2.2. BMM Models

In the mixture model, each observation \( x_i \) would be classified on each unknown number of sub-group. If the allocation of each observation on each sub-group in

---

**Figure 1.** Descriptions of blood sugar levels of DM patients data in Lowokwaru District of Malang
Equation 3 is denoted by $z$, then the allocation of each observation $z_i, i = 1, 2, \ldots, n$ could be determined based on Equation 4 ([8], [24], and [27]).

$$p(z_i = j) = w_j, j = 1, 2, \ldots, k.$$  \hfill (4)

Given the value of $z_i$ then the observation data $x_i$ can be derived from the sub-group as in 5.

$$x_i \mid z_i \sim f(x \mid \theta_{z_i}), i = 1, 2, \ldots, n.$$  \hfill (5)

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

$$p(k, w, z, \theta, x) = p(k)p(w \mid k)p(z \mid w, k)p(\theta \mid z, w, k)$$

$$p(x \mid \theta, z, w, k)$$  \hfill (6)

Furthermore, to estimate each parameter in Equation 6 by employing the full conditional distribution of each parameter [27].

Suppose $p(\theta)$ is a stationary distribution of parameter $\theta$. The full conditional distribution of parameter $\theta$ is constructed by making a partition of $\theta$ as shown in Equation 7 ([8], [10], and [24]).

$$\theta = (\theta_s, \theta_{-s}),$$  \hfill (7)

where $\theta_s$ denote the-$s^{th}$ parameter to be estimated and $\theta_{-s}$ denote the complement of $\theta_s$ that is parameter $\theta$ without the-$s^{th}$ component. Therefore the full conditional distribution can be established based on the joint distribution of all the parameters as in Equation 8 ([8], [10], and [28]).

$$p(\theta_s \mid \theta_{-s}) = \int p(\theta_s, \theta_{-s})d\theta_s.$$  \hfill (8)

### 3.2.3. RJMCMC Algorithm

Reversible Jump Markov Chain Monte Carlo (RJMCMC) concept can be used to identify 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, $\theta$ updating, $z$ updating, hyperparameter $\beta$ 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 ($k + 1 \rightarrow k$) ([8], [10], and [27]). This method is more flexible to use than the Bayes Factor method.

### 3.2.4. MCMC Algorithm

The posterior distribution in the Bayesian analysis is often very complicate and requires a difficult integration process in determining the marginal posterior of a model parameter. Therefore it takes a numerical approximation with Markov Chain Monte Carlo (MCMC) algorithm ([8], [8], [23], and [29]). MCMC algorithms with Gibbs Sampler approach can be described as follows:

#### MCMC Algorithm with Gibbs Sampler Approach

### 3.3. Kolmogorov-Smirnov (KS)

The concept of Kolmogorov-Smirnov test is to compare the empirical cumulative distribution function (CDF), $F_n(x_i)$, and the hypothesis cumulative distribution function, $F(x)$. If statistic
Step 1. We will provide initial value.
\[ \theta^{(0)} = (\theta_1^{(0)}, \ldots, \theta_r^{(0)}) \]

Step 2. We will do the sampling of the parameter.
Generate the value of \( \theta_j, j = 1, \ldots, r \) from their conditional distribution as follows:

Step 2.1. Sampling \( \theta_1^{(k+1)} \) from 
\[ p \left( \theta_1 | x, \theta_2^{(k)}, \ldots, \theta_r^{(k)} \right) \]

Step 2.2. Sampling \( \theta_2^{(k+1)} \) from 
\[ p \left( \theta_2 | x, \theta_1^{(k+1)}, \theta_3^{(k)}, \ldots, \theta_r^{(k)} \right) \]

\[ \vdots \]

Step 2.r. Sampling \( \theta_r^{(k+1)} \) from 
\[ p \left( \theta_r | x, \theta_1^{(k+1)}, \theta_2^{(k+1)}, \ldots, \theta_{r-1}^{(k+1)} \right) \]

Step 3. We will do iteration.
Execute step 2 as \( K \) times with \( K \to \infty \)

order \( X_1, X_2, \ldots, X_n \) is the independent random variable with hypothesis distribution, \( \hat{F}(x) \) and \( F_n(x_i) \) is the empirical distribution as in Equation 9,
\[ F_n(x_i) = \frac{\text{the number of data } X_i \leq x_i}{n} \quad \text{for } i = 1, 2, 3, \ldots, k \text{ } n, \quad (9) \]

where \( F_n(x_i) \) is right continuous step function then the formula of test statistic \( D_n \) can be written as in Equation 10.
\[ D_n = \sup \left\{ \left| F_n(x) - \hat{F}(x) \right| \right\} \quad (10) \]

In this case, \( D_n \) is the greatest distance between \( F_n(x) - \hat{F}(x) \). If the value of \( D_n \) smaller than the tested models is better ([8], [23] and [30]).

4. Result and Discussion
4.1. Description of the Data Analysis
Description of blood sugar levels of DM patients data for Lowokwaru District in Malang 2015 can be seen in the Fig. 2.

According to Fig. 2 can be indicated that the data have a mixture distribution (multimodal distribution). These can be seen by the existence of different groupings of the data. The next step will be identified distribution data from each village. In this step, the proposed hypothesis is as follows:
\( H_0 : \text{data has a normal distribution} \)
versus
\( H_1 : \text{data don’t have a normal distribution} \)

More identification test results are shown in TABLE I.

Derived from TABLE I can be known that the data have a normal distribution for all 7 villages in Lowokwaru District Malang because P-value for each village is larger than 0.05, so that the null hypothesis is accepted. Therefore based on Fig. 2 and TABLE I can be indicated that the data have a normal mixture distribution.
4.2. Identification for the Number of Components Mixture by RJMCMC Algorithm.

The identification results for the number of components mixture in the normal mixture distribution by RJMCMC algorithm can be seen that in the TABLE I and TABLE II.

Based on Table I and Table II can be seen that the mixture normal distribution with 2 components for all villages in Lowokwaru District is the best model because P-value of k2 (with k-Maks= 2) from RJMCMC algorithm is the biggest. This means that the exact model to be built for the DM cases in the Lowokwaru District of Malang is a normal mixture model with 2 components. There are two groups of DM patients in the Lowokwaru district of Malang that each group has a normal distribution.
Table 1. Normality test for DM data in the Lowokwaru District

| No | Villages   | Data        | P-value | Result                 |
|----|------------|-------------|---------|------------------------|
| 1. | Tulusrejo  | Original    | 0.9222  | Normal Distribution    |
|    |            | 100 times   |         | Normal Distribution    |
|    |            | Simulation  |         |                        |
| 2. | Tlogomas   | Original    | 0.9139  | Normal Distribution    |
|    |            | 100 times   |         | Normal Distribution    |
|    |            | Simulation  |         |                        |
| 3. | Sumbersari | Original    | 0.9704  | Normal Distribution    |
|    |            | 100 times   |         | Normal Distribution    |
|    |            | Simulation  |         |                        |
| 4. | Mojolangu  | Original    | 0.7532  | Normal Distribution    |
|    |            | 100 times   |         | Normal Distribution    |
|    |            | Simulation  |         |                        |
| 5. | Lowokwaru  | Original    | 0.9551  | Normal Distribution    |
|    |            | 100 times   |         | Normal Distribution    |
|    |            | Simulation  |         |                        |
| 6. | Jatimulyo  | Original    | 0.6418  | Normal Distribution    |
|    |            | 100 times   |         | Normal Distribution    |
|    |            | Simulation  |         |                        |
| 7. | Dinoyo     | Original    | 0.7304  | Normal Distribution    |
|    |            | 100 times   |         | Normal Distribution    |
|    |            | Simulation  |         |                        |

4.3. Weighting of the BMM Models

In TABLE III are presented the results of the identification for weighted (w) values in each of mixture components on the BMM modeling. The weighting values of the first mixture component are w1 and the weighting values of the second mixture component are w2. This weighting is obtained by the use of MCMC algorithms with Gibbs Sampler approach for normal mixture distribution through Winbugs program with 3 steps execution.

Based on the weighting values shown in TABLE III it can be seen that the villages of Sumbersari, Lowokwaru, Tulusrejo, and Mojolangu have the second weighting value of mixture components (w2) is greater than the first weighting value of mixture components. Its means that in the four villages have had the proportion of DM patients with higher blood sugar levels more than the number of DM patients with lower blood sugar levels. As for the three other villages namely Tlogomas, Jatimulyo, and Dinoyo have w2 lower than w1. However, the average weighted values, Lowokwaru District has a weighted value of the second mixture components (w2) higher than a weighted value of the first mixture components (w1).

4.4. BMM Models and Goodness of Fits BMM Models by KS

BMM models and Goodness of Fits BMM models by KS for all villages in Lowokwaru District Malang can be shown in the Fig. 3 and TABLE IV.

Derived from Fig. 2, it can be shown that the BMM models for DM cases in the Lowokwaru
Table 2. RJMCMC Algorithm for DM data in the Lowokwaru District

| No | Villages  | Data   | P-value of $k_2$ | k-Maks | Result                                      |
|----|----------|--------|-----------------|--------|---------------------------------------------|
| 1  | Tulusrejo| Original| 0.3839          | 2      | Mixture Normal Distribution with 2 Components |
|    |          |        | 100 times       | 0.381437| Mixture Normal Distribution with 2 Components |
|    |          | Simulation (average) |        |        | Mixture Normal Distribution with 2 Components |
| 2  | Tlogomas | Original| 0.4015          | 2      | Distribution with 2 Components              |
|    |          |        | 100 times       | 0.394409| Mixture Normal Distribution with 2 Components |
|    |          | Simulation (average) |        |        | Mixture Normal Distribution with 2 Components |
| 3  | Sumberari| Original| 0.3216          | 2      | Distribution with 2 Components              |
|    |          |        | 100 times       | 0.334833| Mixture Normal Distribution with 2 Components |
|    |          | Simulation (average) |        |        | Mixture Normal Distribution with 2 Components |
| 4  | Mojolangu| Original| 0.347           | 2      | Distribution with 2 Components              |
|    |          |        | 100 times       | 0.348434| Mixture Normal Distribution with 2 Components |
|    |          | Simulation (average) |        |        | Mixture Normal Distribution with 2 Components |
| 5  | Lowokwaru| Original| 0.3727          | 2      | Distribution with 2 Components              |
|    |          |        | 100 times       | 0.373755| Mixture Normal Distribution with 2 Components |
|    |          | Simulation (average) |        |        | Mixture Normal Distribution with 2 Components |
| 6  | Jatimulyo| Original| 0.3649          | 2      | Distribution with 2 Components              |
|    |          |        | 100 times       | 0.367506| Mixture Normal Distribution with 2 Components |
|    |          | Simulation (average) |        |        | Mixture Normal Distribution with 2 Components |
| 7  | Dinoyo   | Original| 0.4868          | 2      | Distribution with 2 Components              |
|    |          |        | 100 times       | 0.488151| Mixture Normal Distribution with 2 Components |
### Table 3. The Weighting Values of the BMM Models

| No | Villages | Data                   | w1-value | w2-value |
|----|----------|------------------------|----------|----------|
| 1  | Tulusrejo| Original               | 0.3384   | 0.6616   |
|    |          | 100 times simulation   | 0.3384   | 0.6616   |
|    |          | (Average)              | 0.3384   | 0.6616   |
| 2  | Tlogomas | Original               | 0.5041   | 0.4959   |
|    |          | 100 times simulation   | 0.504    | 0.496    |
|    |          | (Average)              | 0.504    | 0.496    |
| 3  | Sumbersari| Original              | 0.5041   | 0.4959   |
|    |          | 100 times simulation   | 0.2544   | 0.7456   |
|    |          | (Average)              | 0.2544   | 0.7456   |
| 4  | Mojolangu| Original               | 0.4527   | 0.5473   |
|    |          | 100 times simulation   | 0.4527   | 0.5473   |
|    |          | (Average)              | 0.4527   | 0.5473   |
| 5  | Lowokwaru| Original              | 0.2935   | 0.7065   |
|    |          | 100 times simulation   | 0.2935   | 0.7065   |
|    |          | (Average)              | 0.2935   | 0.7065   |
| 6  | Jatimulyo| Original              | 0.5687   | 0.4313   |
|    |          | 100 times simulation   | 0.5687   | 0.4313   |
|    |          | (Average)              | 0.5687   | 0.4313   |
| 7  | Dinoyo   | Original              | 0.6058   | 0.3942   |
|    |          | 100 times simulation   | 0.6058   | 0.3942   |
|    |          | (Average)              | 0.6058   | 0.3942   |

Average of w  | 0.448914286 | 0.551085714

District in Malang have a normal mixture distribution with two components, namely DM patients with average blood sugar levels less than 200 mg/dL and greater than or equal to 200 mg/dL. Each village has a different characteristic of the normal mixture distribution that vary according to the condition data. Based on the shape of the distribution and the value of the blood sugar levels of DM patients in the Lowokwaru District, the villages that need attention from the Local Government in Malang respectively are Lowokwaru, Tulusrejo, Mojolangu, Sumbersari, Dinoyo, Tlogomas, and Jatimulyo. From TABLE IV it can be seen that the average value of KS as Goodness of Fits BMM models is 0.35739667. Its means that the BMM models that have been built have high accuracy and the BMM models are able to explain the condition of the real data with data driven concept. Derived from the exact model that has been constructed for DM cases, it can be utilized for the Department of Health and related agencies in the prevention and control of the DM diseases.

### 5. Conclusion

The BMM models for cases of the blood sugar levels of DM patients in RSUD Saiful Anwar Malang in the Lowokwaru District have normal mixture distribution with 2 components, namely DM patients with average blood sugar levels less than 200 mg/dL and greater than or equal to...
Villages that need attention from Local Government in Malang are Lowokwaru, Tulusrejo, Mojolangu, Sumbersari, Dinoyo, Tlogomas, and Jatimulyo, respectively. The BMM-Normal model for Lowokwaru District has an average KS value of 0.35739667 so that these models have high accuracy and able to accommodate the real condition of the DM data with driven data concept.

Acknowledgement
This paper is part of research Grants PUPT/BOPTN University of Brawijaya in 2015-2016 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.
### Table 4. The Goodness of Fits BMM Models by KS

| No | Villages   | Data               | KS-value       |
|----|------------|--------------------|----------------|
| 1  | Tulusrejo  | Original           | 0.226512714    |
|    |            | 100 times simulation | 0.22355145    |
|    |            | (Average)          | 0.22355145     |
| 2  | Tlogomas   | Original           | 0.300381216    |
|    |            | 100 times simulation | 0.29200553    |
|    |            | (Average)          | 0.29200553     |
| 3  | Sumbersari | Original           | 0.396896742    |
|    |            | 100 times simulation | 0.40181093    |
|    |            | (Average)          | 0.40181093     |
| 4  | Mojolangu  | Original           | 0.214107821    |
|    |            | 100 times simulation | 0.22167467    |
|    |            | (Average)          | 0.22167467     |
| 5  | Lowokwaru  | Original           | 0.518545269    |
|    |            | 100 times simulation | 0.52749470    |
|    |            | (Average)          | 0.52749470     |
| 6  | Jatimulyo  | Original           | 0.35239443     |
|    |            | 100 times simulation | 0.36528055    |
|    |            | (Average)          | 0.36528055     |
| 7  | Dinoyo     | Original           | 0.381005007    |
|    |            | 100 times simulation | 0.38113163    |
|    |            | (Average)          | 0.38113163     |

**Average KS Simulation** 0.35739667

### References

[1] Gosh J K, Delampady M and Samanta T, 2006 *An Introduction to Bayesian Analysis Theory and Method* (New York: Springer)

[2] Hoeting J A, Madigan D, Raftery A E and Volinsky C T 1999 Bayesian model averaging: tutorial (with discussion), *Stat.Sci.*, **14** pp 382–401

[3] Iriawan N 2000 *Computationally Intensive Approaches to Inference in Neo-Normal Linear Models* (Australia: CUT).

[4] Iriawan N 2012 *Modeling and Analysis of Data-Driven* (in Indonesian) Vol.1 (Surabaya: ITS)

[5] Sebastiani P, Xie H and Ramoni M F 2006 Bayesian analysis of comparative microarray experiments by model averaging, *Internat.Soc.Bayesian.Anal.* **1** pp 707–732

[6] Gelman A, Carlin J B, Stern H S and Rubin D B 1995 *Bayesian Data Analysis* (London: Chapman & Hall)

[7] Suyono S, *Textbook of Medicine* (in Indonesian) Vol.3 4ed (Jakarta: FKUI) (2006)

[8] Astuti A B, Iriawan N, Irhamah, Kuswanto H, and Sasiarini L 2016 Blood sugar levels of diabetes mellitus patients modeling with bayesian mixture model averaging *Glob. J. Pure Appl. Math.* **12** pp 3143–3158

[9] Do K A, Muller P and Tang F 2005 *A Bayesian Mixture Model For Differential Gene Expression* (Texas: Research in Department of Biostatistics, The University of Texas)

[10] Wati D M 2006 *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) (Surabaya: Statistics Department, ITS)

[11] Astuti A B 2014 Cox regression model for lifetime of a newborn baby (neonates) in Dr. Saiful Anwar Malang Hospital *Glob. J. Pure Appl. Math.* **10** pp 591–600.
http://www.tempo.co/read/news/2013/11/14/060529766/ Indonesian-Log-10-Big-State-with- Diabetes-Sufferers, (in Indonesian).
http://medicastore.com/diabetes/terapidia, Treatment of Diabetes Mellitus. (in Indonesian).
Smeltzer S and Bare B 2008 Textbook of MedicalSurgical Nursing 11ed (Philadelphia: Lippincott)
Suharjo B 2003 Implementation of Bayesian Markov Chain Monte Carlo in Normal Mixture Distributions (in Indonesian) (Surabaya: Statistics Department, ITS)
Aitkin M 2001 Likelihood and bayesian analysis of mixtures Stat.Model. 1 (2001) pp 287–304
Box G E P and Tiao 1973 Bayesian Inference in Statistical Analysis (Boston MA: Addison-Wesley, Massachusetts)
Iriawan N 2001 Univariable normal mixture model estimation: a bayesian methods approach with MCMC (in Indonesian) Proc. of National Seminar and Konferda VII on Mathematics in DIY & Central Java Region pp 105–110
Iriawan N 2003 Teaching Modules of Simulation Technique (in Indonesian) (Surabaya: ITS)
Astuti A B, Iriawan N, Irhamah and Kuswanto H 2014 Model components selection in bayesian model averaging using occams window for microarray data, Proc. of 4th Annual Basic Science International Conference 2014 in conjunction with The 5th International Conference on Global Resource Conservation 2014
Astuti A B, Iriawan N, Irhamah and Kuswanto H 2014 Kolmogorov-smirnov and continuous ranked probability score validation on the bayesian model averaging for microarray data Appl. Math. Sci. 146 pp 7277–7287
Astuti A B, Iriawan N, Irhamah, and Kuswanto H 2015 Occams window selection in bayesian model averaging for gene expression data from chickpea plant Int. J.Appl. Stat. 53 pp 160–165.
Astuti A B, Iriawan N and Sasiarini L 2015 Bayesian model averaging with occams window selection components model on diabetes mellitus data (in Indonesia) SEMIRATA 2015 MIPAb KS-PTN West University of Tanjungpura, Pontianak.
Astuti A B, Iriawan N, Irhamah, and Kuswanto H 2015 An algorithm for determining the number of mixture components on the bayesian mixture model averaging for microarray data J.Mat.Stat. 11 pp 45–51
Astuti E Y 2006 Implementation of the Bayesian Markov Chain Monte Carlo in Optimal Portfolio Modeling with Mixture Model Approach of Multiple Mixture (in Indonesian) (Surabaya: Statistics Department, ITS)
McLachlan G J and Basford K E 1988 Mixture Models Inference and Applications to Clustering (New York: Marcel Dekker)
Richardson S and Green P J 1997 On Bayesian analysis of mixtures with an unknown number of components (with discussion) J. Roy. Statist. Soc 59 pp 731–792
Gilks W R 1995 Full Conditional Distribution, in Markov Chain Monte Carlo in Practise. (eds) (London: Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. Chapman & Hall) pp 75–88
Gamerman D 1997 Markov Chain Monte Carlo (London: Chapman & Hall)
Hustianda V F 2012 Comparison of Bayesian Model Averaging and Multiple Linear Regression in Predicting Factors Affecting Number of Infant Death in East Java, (in Indonesian) (Surabaya: Statistics Department, ITS)