Modeling the crashes count using finite mixture models

Safaa K. Kadhem[1] and Sadeq A. Kadhim[2]

,[1]Department of Mathematics and Computer Applications, College of Science

Al-Muthanna University, Iraq

[2]Higher Education and Scientific Research Ministry, Iraq

Received 28/3/2020 , Accepted 5/5/2020 , published:10/5/2020

DOI: 10.52113/2/07.01.2020/13-27

Abstract

This paper aims at the modeling the crashes count in Al Muthanna governance using finite mixture model. We use one of the most common MCMC method which is called the Gibbs sampler to implement the Bayesian inference for estimating the model parameters. We perform a simulation study, based on synthetic data, to check the ability of the sampler to find the best estimates of the model. We use the two well-known criteria, which are the AIC and BIC, to determine the best model fitted to the data. Finally, we apply our sampler to model the crashes count in Al Muthanna governance.

Keywords: Finite Mixture Model, Model Selection, Crashes, Bayesian Inference

1.1 Introduction

Mixture models have been developed as a flexible tool to model date with an unobserved heterogeneity, for example, different types of date can form clusters or groups. A finite mixture model (FMM) is generally used when an observation belongs to one of K groups (components) that have distinct features and can be described by different probability distributions. In other words, these models are a weighted average of a finite number of distribution (mixing components). FMMs may be a finite mixture of distributions such as Gaussian or Poisson distributions. Interest in FMMs has increased over the last decades. They can be used for cluster analysis, latent class analysis, discriminant analysis, image analysis, survival analysis, disease mapping and meta-analysis. There are many textbooks which have focused in detail on finite mixture models such as [7], [2].

Bayesian methods to model these mixtures of distributions have been used widely for inference. The wide use of those distributions resulted in the large development in posterior simulation ways, for instance, the Markov chain Monte Carlo (MCMC) procedures [7]. Therefore, MCMC procedures have been used to handle the difficulties in the estimation processes of parameters of FMM, for instance, determining the rank of the mixture model [8], and the problem of label switching [10],[4]. Moreover, the Bayesian theory has been used to facilitate the modeling the complicated structure in these models via classifying them into a set of similar structures using the augmentation procedure.

This paper includes the following aspects. In Section 2, we address the literature review concern the estimation and selection processes of finite mixture models. Section 3 introduces the definition of model. The model selection method is given in Section 4, and the simulation results and two real-data examples in Section 5. We present some conclusions in Section 6.

1.2 Definition of the finite mixture model

Let $y = (y_1, y_2, \dots, y_T)$ denote a sample of observed data of size T, the probability density function (p.d.f.) of a mixture model can be defined as a combination of K component p.d.f.:

$$\Pr(y/\Theta) = \sum_{k=1}^{k} \pi_k \Pr_k(y/\theta_k), \qquad (1.1)$$

Where $\Pr_k(y/\theta_k)$ denotes the p.d.f. of the k^{th} component, π_k is the weight of the population k such that $0 \le \pi_k \le 1$, and $\sum_{k=1}^{k} \pi_k = 1$, $\Theta = (\pi; \theta) =$

$(\pi_1, \pi_2, \dots, \pi_k;\ \theta_1, \theta_2, \dots, \theta_k)$ denotes a set of all unknown weight and parameters of a mixture model. In many applications, a family of distributions having the density in Equation (1.1) can be called a k-component finite mixture model.

The main idea of mixture model is that the observations y are generated from k distinct random processes so that each process is modelled by the density $\Pr_k(y/\theta_k)$, and π_k represents the corresponding proportion of observations from this process. For example, consider a FMM where $\Pr(y/\Theta)$ is constituted from densities which are all Normal or Poisson distribution.

1.3 Bayesian estimation of mixture model

Given an independent identically distributed (iid) random sample, $y = (y_1, y_2, \dots, y_T)$, generated from a k-component mixture model defined in Equation (1.1), the likelihood function of these observations, assuming that y_t is independently distributed, can be written as

$$\mathbf{Pr}(\boldsymbol{y}/\boldsymbol{\Theta}) = \mathbf{L}(\boldsymbol{\Theta}; \boldsymbol{y}) = \prod_{t=1}^{T} \sum_{k=1}^{k} \boldsymbol{\pi_k Pr_k(y_t/\theta_k)}. \quad (\mathbf{1.2})$$

In the FMM in Equation (1.2), the unknown parameter vector $\boldsymbol{\Theta} = (\boldsymbol{\pi}; \boldsymbol{\theta})$ needs to be estimated. In order to obtain the posterior distribution of $\boldsymbol{\Theta}$, we need to combine the data-dependent likelihood function $\mathrm{L}(\boldsymbol{\Theta}; \boldsymbol{y})$ of the mixture model and the prior distribution of the unknown parameters; $\boldsymbol{\theta}$ and $\boldsymbol{\pi}$. The posterior distribution can be given as

$$\mathbf{Pr}(\boldsymbol{\Theta}/\boldsymbol{y}) \propto \mathbf{L}(\boldsymbol{\theta}, \boldsymbol{\pi}; \boldsymbol{y})\, \mathbf{Pr}(\boldsymbol{\pi})\, \mathbf{Pr}(\boldsymbol{\theta}), \quad (\mathbf{1.3})$$

where $\mathrm{L}(\theta, \pi; y) = \prod_{t=1}^{T} \Pr(y_t/\theta, \pi) = = \prod_{t=1}^{T} \{\sum_{k=1}^{K} \pi_k f(y_t/\theta_k)\}$ is the likelihood, $\Pr(\theta)$ and $\Pr(\pi)$ represent the prior distribution of θ and π respectively.

An efficient method for simplifying the sampling from the posterior distribution is the date augmentation method proposed by [11]. This method is based on sampling from the complete date posterior distribution $\Pr(\Theta, z/y)$ rather than $\Pr(\Theta/y)$ by proposing auxiliary variables, called z, also referred as latent indicator variables. If we know y and z, then the analysis will be more straightforward.

We assume that there are discrete latent indicators, $z = \{z_{kt}\}$, associated with each observation of the vector $y = (y_1, y_2, \dots, y_T)$. Sine these indicators in real life are unknown parameters, the inference about a mixture model requires estimating two unknown quantities : the component indicators, z, and the component parameters, $\Theta = (\pi, \theta)$. in the Bayesian perspective, in order to obtain those quantities, these can be sampled from the fallowing complete date posterior:

$$\mathbf{Pr}(\mathbf{z}, \boldsymbol{\pi}, \boldsymbol{\theta}/\mathbf{y}) \propto \mathbf{L_c}(\boldsymbol{\theta}, \boldsymbol{\pi}; \mathbf{y}, \mathbf{z})\, \mathbf{Pr}(\boldsymbol{\pi})\, \mathbf{Pr}(\boldsymbol{\theta}), \quad \mathbf{(1.4)}$$

where $\mathrm{L_c}(\theta, \pi; y, z)$ is the complete data likelihood of a finite mixture model, $\Pr(\theta)$ and $\Pr(\pi)$ are independent prior distribution of the parameter θ and of the components weights π respectively. The complete-data likelihood can be written as

$$L_c(\theta,\pi;y,z) = \prod_{t=1}^{T} \pi_{z_t} \Pr(y_t/\theta_{z_t})$$

$$= \prod_{k=1}^{K} \prod_{t:z_1=k} \pi_k \Pr(y_t/\theta_k)$$

$$= \prod_{k=1}^{K} \pi_k^{\sum_{t=1}^{T} I(z_t=k)} \prod_{t:z_t=k} \Pr(y_t/\theta_k).$$

(1.5)

To complete the Bayesian specification of the model, we need to specify priors for the unknown parameters of the model: π and θ. The prior on the component weights is represented by a Dirichlet distribution as

$$\mathbf{Pr}(\boldsymbol{\pi}) = \prod_{k=1}^{K} \boldsymbol{\pi}_{\mathbf{k}} \propto \prod_{k=1}^{K} \mathbf{k}_{\mathbf{k}}^{\delta_k - 1} = \mathbf{Dirichlet}\,(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \dots, \boldsymbol{\delta}_k), \quad (\mathbf{1.6})$$

Where δ_k, $k = 1,2,\dots,k$ are the positive ($\delta_k > 0$) hyper-parameters of the Dirichlet distribution. The prior on the component-specific parameter, θ, based on the form of the parametric distribution assumed for observations, y. As a general case for representing the prior on the component-specific parameter, θ, we can write the following expression

$$\boldsymbol{\theta} \sim \Pr(\boldsymbol{\theta}/\boldsymbol{\varphi}), \qquad (\mathbf{1.7})$$

Where φ is referred to a collection of the hyper-parameters governing the shape of the prior distribution of θ. Common MCMC approaches can be employed. We use the Gibbs sampler [6] to simulate from the full conditional posterior distributions of the FMM.

1.3.1 Estimation using the Gibbs sampler

The posterior distribution in Equation (1.6) involves three full conditional distributions which can be written as

$$\boldsymbol{z} \sim \mathbf{Pr}(\boldsymbol{z}/\boldsymbol{y}\,,\boldsymbol{\pi},\boldsymbol{\theta}),$$

$$\boldsymbol{\pi} \sim \mathbf{Pr}(\boldsymbol{\pi}/\boldsymbol{y}\,,\boldsymbol{z}), \qquad (\mathbf{1.8})$$

$$\boldsymbol{\theta} \sim \mathbf{Pr}(\boldsymbol{\theta}/\boldsymbol{y}\,,\boldsymbol{z}).$$

It is easy to implement the Gibbs sampler [3] to sample those distributions. In Bayesian inference for FMMs, the mixture proportion $\{\pi_1, \pi_2, \dots, \pi_k\}$ can be viewed as the prior distribution that one observation belongs to sub-population k. Given the observations, y_t, the full conditional posterior distribution of z_t can be obtained as

$$\Pr(z_t = k/y_t,\pi,\theta) \propto \pi_k \Pr(y_t/\theta_k)$$

$$= \frac{\pi_k \Pr(y_t/\theta_k)}{\sum_{l=1}^{K} \pi_l \Pr(y_t/\theta_l)}. \qquad (\mathbf{1.9})$$

From Equation (1.9), the marginal distribution of the z_t is a multinomial distribution

$$z_t \sim \text{multinomial}\,\{\Pr(z_t = 1)\,,\Pr(z_t = 2)\,,\dots,\Pr(z_t = K)\}. \qquad (\mathbf{1.10})$$

Given component indicators z, the full conditional posterior of the component weights, π, can be sampled as follows

$$\mathbf{Pr}(\boldsymbol{\pi}/\boldsymbol{y},\boldsymbol{z},\boldsymbol{\delta}) \propto \boldsymbol{L_c}(\boldsymbol{\theta},\boldsymbol{\pi};\boldsymbol{y},\boldsymbol{z})\mathbf{Pr}(\boldsymbol{\pi}/\boldsymbol{\delta})$$

$$\propto \prod_{k=1}^{K} \pi_k^{\sum_{t=1}^{T} I(z_t=k)} \prod_{t:z_t=k} \Pr(y_t/\theta_k) \prod_{k=1}^{K} \pi_k^{\delta_k-1}$$

$$\propto \prod_{k=1}^{K} \pi_k^{\sum_{t=1}^{T} I(z_t=k)+\delta_k-1}$$

$$\boldsymbol{\pi} \sim \text{Dirichlet}\,(n_1+\delta_1, n_2+\delta_2, \dots, n_k + \delta_k), \qquad (\mathbf{1.11})$$

where $n_k = \sum_{l=1}^{T} I_{z_l=k}$, $k = 1,2,\dots,K$, denote the allocation sizes. Given component indicators z and observation y, the posterior of θ is

$$\mathbf{Pr}(\boldsymbol{\theta}/\boldsymbol{y},\boldsymbol{z}) \propto \boldsymbol{L_c}(\boldsymbol{\theta},\boldsymbol{\pi};\boldsymbol{y},\boldsymbol{z})\mathbf{Pr}(\boldsymbol{\theta})$$

$$\sim \mathbf{Pr}(\boldsymbol{\theta}) \prod_{t:z_t=k} \Pr(y_t/\theta_k). \qquad (1.12)$$

Algorithm(3) provided by [6] describes the steps of sampling from the full conditional posterior distributions of a mixture model.

Algorithm 3 : Gibbs Sampler for a k-component finite mixture model

Initialization: Choose $\boldsymbol{\pi}^{(0)}$ and $\boldsymbol{\theta}^{(0)}$ arbitrarily

Iteration m $(m \geq 1)$:

1- Generate $z_t^{(m)}$ $(t = 1, 2, \dots, T)$ from

$\Pr(z_t^{(m)} = k/\pi_k^{(m-1)}, \theta_k^{(m-1)}, y_t) \propto \pi_k^{(m-1)} f(y_t/\theta_k^{(m-1)})\,; k = 1,2,\dots,K.$

2- Generate $\pi^{(m)}$ from $\Pr(\pi/z^{(m)})$,

3- Generate $\theta^{(m)}$ from $\Pr(\theta/z^{(m)}, y_t)$.

1.4 Selection model

We use two popular criteria for model selection are: the Akaike's information criterion (AIC)[1] and the Bayesian information criterion (BIC) [9]. Both AIC and BIC are based on the log-likelihood evaluated and penalized for the number of parameters in the model.

$$\text{AIC} = -2\text{LogPr}(y_t/\widehat{\boldsymbol{\theta}}) + 2K \qquad \textbf{(1.13)}$$

$$\text{BIC} = -2\text{LogPr}(y_t/\widehat{\boldsymbol{\theta}}) + h\,log(n) \qquad \textbf{(1.14)}$$

Where h is the number of parameters, n is sample size and $\text{LogPr}(y_t/\hat{\theta})$ is log likelihood function which typically is being estimated under the classic framework. We adopt modified versions of these criteria in which the log likelihood is evaluated using the Bayesian method [5].

2.1 Study design

In this section we perform a simulation study to estimate the model mixture parameters using the Gibbs sampler. In addition, under the same simulation study, we try to select the best mixture model among several competing mixture models using criteria: AIC and BIC.

2.2 Simulation study

In this section we want to implement a simulation study. This study includes generating several data set from different models

2.2.1: Generating synthetic data sets

We generated three data sets with size n=600 for each from Normal mixture models with K_0=2,3 and 4 respectively, where K_0 denotes the model order:

First model: $0.3\Pr(y_t/2,1) + 0.7\Pr(y_t/10,1)$

Second model: $0.3\Pr(y_t/2,1) + 0.5\Pr(y_t/8,1) + 0.2\Pr(y_t/12,1)$

Third model: $0.25\Pr(y_t/2,1) + 0.25\Pr(y_t/8,1) + 0.25\Pr(y_t/12,1) + 0.25\Pr(y_t/20,1)$

Figures (1), (2) and (3) show the histograms of the data sets simulated from the three models.

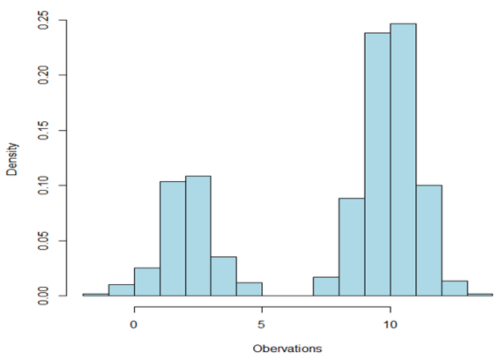

Figure (1): The probability density function of data set generated from the first model (2-components Normal mixture model).

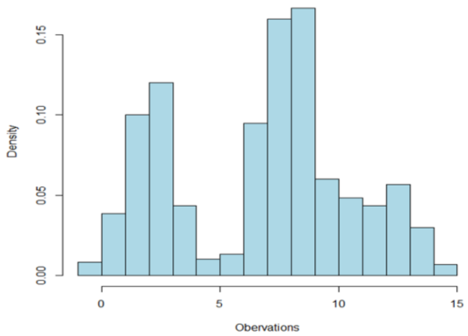

Figure (2): The probability density function of data set generated from the second model (3-components Normal mixture model).

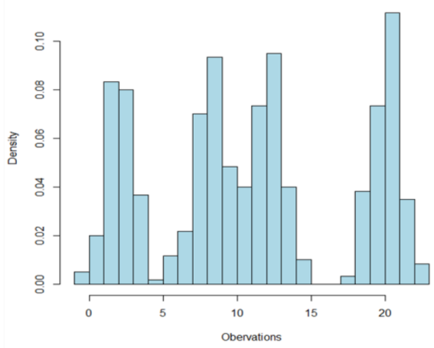

Figure (3): The probability density function of data set generated from the third model (4-components Normal mixture model).

2.2.2: Estimation of the model parameters

In this section, we estimate the model parameters using the Gibbs sampler [3] of each model of the three models that generated the data sets. The parameters of the model is explained as following:

$$\sigma_j^2 \sim InvGamma(a_j, b_j)$$

$$\mu_j/\sigma_j^2 \sim N(\eta_j, \sigma_j^2/\xi_j)$$

$$\pi_j \sim Dir(\delta_j)$$

Where $\eta_j, \zeta_j, a_j, b_j$ and δ_j are known hyper-parameters, $j = 1,2$. The hyper-parameters need to be specified [6]. We specify non-informative priors: $a = 0.001, b = 0.001, \eta = 0$ and $\zeta = 0.001$. The weigh parameter, π, is given a Dirichlet prior with non-informative value, $\delta_k = 1, k = 1,2, \ldots, K$. Given the above parameterization on the hyper-parameters of the priors distributions, we follow algorithm (2), given by [6], to implement the sampling process.

2.3: Results

2.3.1: Results of the model estimation

We run separately the Gibbs sampler for 10000 iterations for each model. We adopted the last 5000 iterations for inference and discarded 5000 iterations as a burn-in period. Tables (1), (2) and(3) show estimation result of each model receptively. Figures (4), (5) and(6) show the MCMC result of each model receptively. As we see that the sampler performs well for estimating the true parameters of all three models.

3.1: Modeling data count using Poisson Mixture model

Let $y = (y_1, y_2, \ldots, y_T)$ denote count data length T , the Poisson mixture model of K component can be given by

$$Pr(y/\lambda) = \sum_{k=1}^{k} \pi_k \, Poi_k(y/\lambda_k), \quad (1.1)$$

Where $Pr_k(y/\lambda)$ denotes the probability mass function (p.m.f.) of the Poisson mixture model , π_k is the weight of the population k such that $0 \le \pi_k \le 1$, and $\sum_{k=1}^{k} \pi_k = 1$.

By following the same procedure with respect to the normal mixture model mentioned in the section 2, with replace the p.d.f. of normal distribution by the p.m.f. of Poisson distribution, we apply the Poisson mixture model to application including monthly accidents count data as described in the next section.

3.1.1- Real application data

In this section we consider an application including a series of monthly accidents count data occur in period (2014-2017) in the motorway 8 which links the Nasiriya city by the hila city that pass through the Al Muthanna city. Figure (7) shows the map of these accident count data.

Algorithm 2 : Gibbs Sampler for a K-component Normal mixture model with conjugate priors

1. Initialization: Choose $\pi_k^{(0)}$ and $\theta_k^{(0)}$,$k = 1,1, \ldots, K$

2. Iteration: for $m = 1,2, \dots, M$

(a) Generate $z_t^{(m)}; t = 1, \dots., T$ from $(k = 1,2, \dots K)$

$$\Pr\left(z_t^{(m)} = 1\right) = 1 - \Pr\left(z_t^{(m)} = 2\right) \propto \frac{\pi_k^{(m-1)}}{\sigma_k^{(m-1)}} \exp\left(-\frac{y_t - \mu_k^{(m-1)^2}}{2(\sigma_k^2)^{(m-1)}}\right).$$

Compute: $n_k^{(m)} = \sum_{l=1}^{n} I_{z_l^{(m)}=k}$ and $s_k^{y(m)} = \sum_{l=1}^{n} I_{z_l^{(m)}=k}\, y_l$.

(b) Update $\pi_k^{(m)}$ from $Dir(\delta_1 + n_1^{(m)}, \delta_2 + n_2^{(m)}, \dots., \delta_k + n_k^{(m)})$

(c) Generate $\mu_k^{(m)}; k = 1,2, \dots, K$ from

$$N\left(\frac{\eta_k \zeta_k + s_k^{y(m)}}{\zeta_k + (n_k)^{(m)}}, \frac{\sigma_k^{2(m-1)}}{\zeta_k + (n_k)^{(m)}}\right).$$

Compute: $s_k^{v(m)} = \sum_{l=1}^{n} I_{z_l^{(m)}=k} (y_t - \mu_k^{(m)})^2.$

(d) Generate $\sigma_k^{2(m)}; k = 1,2, \dots, K$ from

InvGamma$(a_k + 0.5\left(n_k^{(m)} + 1\right), b_k + 0.5\zeta_k(\mu_k^{(m)} - \eta_k)^2 + 0.5(s_k^{v(m)}))$.

Real parameters					
Weight		Mean		Variance	
w_1	w_2	μ_1	μ_2	σ_1^2	σ_2^2
0.3	0.7	2	10	1	1
$\widehat{w_1}$	$\widehat{w_2}$	$\widehat{\mu_1}$	$\widehat{\mu_2}$	$\widehat{\sigma_1^2}$	$\widehat{\sigma_2^2}$
0.281	0.719	2.049	9.956	1.096	1.026

Table (1): The estimated and real values of the parameters of a Normal mixture model with K=2 using Gibbs sampler.

Real parameters								
Weight			Mean			Variance		
w_1	w_2	w_3	μ_1	μ_2	μ_3	σ_1^2	σ_2^2	σ_3^2
0.3	0.5	0.2	2	8	12	1	1	1
$\widehat{w_1}$	$\widehat{w_2}$	$\widehat{w_3}$	$\widehat{\mu_1}$	$\widehat{\mu_2}$	$\widehat{\mu_3}$	$\widehat{\sigma_1^2}$	$\widehat{\sigma_2^2}$	$\widehat{\sigma_3^2}$
0.31	0.48	0.21	2.07	7.89	12.02	0.98	1.07	1.04

Table (2): The estimated and real values of the parameters of a Normal mixture model with K=3 using Gibbs sampler.

Real parameters											
Weight				Mean				Variance			
w_1	w_2	w_3	w_4	μ_1	μ_2	μ_3	μ_4	σ_1^2	σ_2^2	σ_3^2	σ_4^2
0.25	0.25	0.25	0.25	2	8	12	20	1	1	1	1
$\widehat{w_1}$	$\widehat{w_2}$	$\widehat{w_3}$	$\widehat{w_4}$	$\widehat{\mu_1}$	$\widehat{\mu_2}$	$\widehat{\mu_3}$	$\widehat{\mu_4}$	$\widehat{\sigma_1^2}$	$\widehat{\sigma_2^2}$	$\widehat{\sigma_3^2}$	$\widehat{\sigma_4^2}$
0.26	0.27	0.24	0.22	2.09	8.34	12.19	20.07	1.33	1.04	1.13	1.07

Table (3): The estimated and real values of the parameters of a Normal mixture model with K=4 using Gibbs sampler.

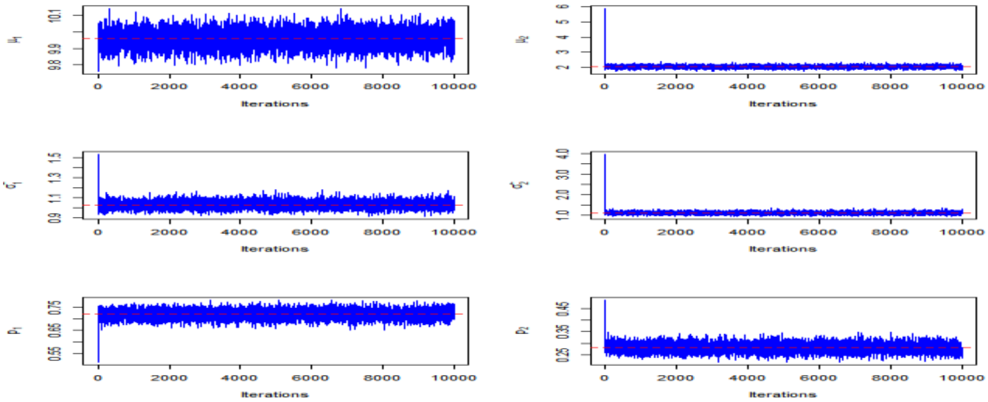

Figure (4): The posterior distributions of the model parameter with K_0=2.

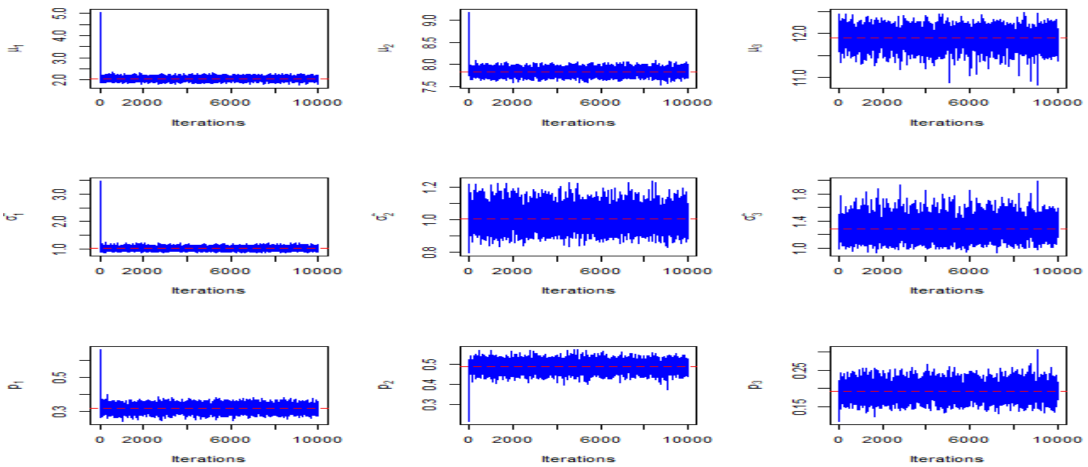

Figure (5): The posterior distributions of the model parameter with K_0=3.

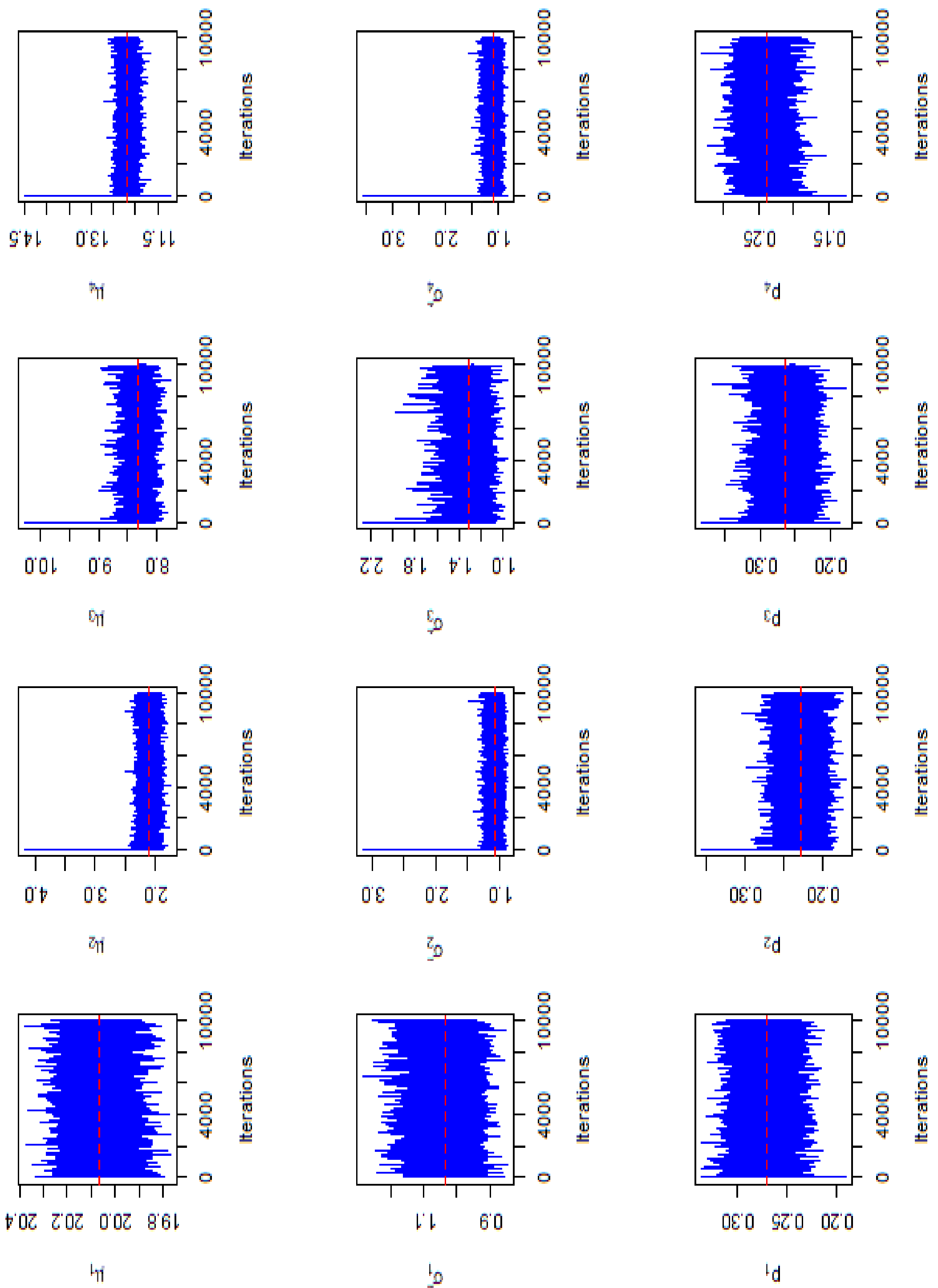

Figure (6): The posterior distributions of the model parameter with K_0=4.

Figure (7) : The map of accident count in Al Muthanna City.

Index	Mouth	Accident count	Index	Mouth	Accident count
1	Jan.2014	23	25	Jan.2016	35
2	Feb.2014	40	26	Feb.2016	30
3	Mar.2014	26	27	Mar.2016	63
4	Apr.2014	30	28	Apr.2016	49
5	May.2014	15	29	May.2016	44
6	June.2014	28	30	June.2016	52
7	July.2014	28	31	July.2016	46
8	Aug.2014	35	32	Aug.2016	56
9	Sept.2014	16	33	Sept.2016	46
10	Oct.2014	18	34	Oct.2016	37
11	Nov.2014	22	35	Nov.2016	38
12	Dec.2014	10	36	Dec.2016	25
13	Jan.2015	22	37	Jan.2017	36
14	Feb.2015	37	38	Feb.2017	42
15	Mar.2015	25	39	Mar.2017	45
16	Apr.2015	31	40	Apr.2017	29
17	May.2015	31	41	May.2017	27
18	June.2015	36	42	June.2017	37
19	July.2015	25	43	July.2017	33
20	Aug.2015	25	44	Aug.2017	51
21	Sept.2015	16	45	Sept.2017	57
22	Oct.2015	16	46	Oct.2017	43
23	Nov.2015	22	47	Nov.2017	52
24	Dec.2015	10	48	Dec.2017	50

Table (5): Shows the monthly accident counts in Al Muthanna city through period (2014-2017)

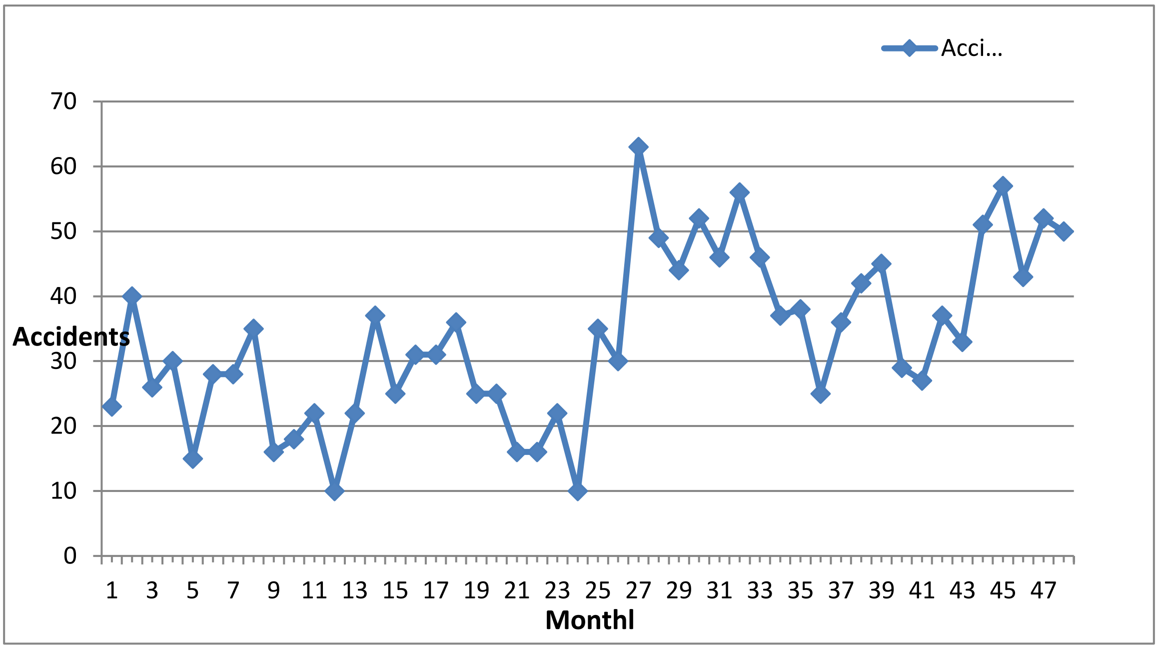

Figure (8): Shows the monthly accidents counts in Al Muthanna city

3.2: Results

Tables (6) shows the estimation and model selection results of six Poisson mixture model fitted to accident count data. The third column represent the estimated weights, while the fourth column show the estimated mean parameter of Poisson mixture model. The fifth column shows the log-likelihood of each model fitted to those data. While the sixth and seventh columns represent the AIC and BIC respectively. Note that the best model for these data is the Poisson mixture model with K=3. The mean of the first component of this model was $\hat{\lambda}_1$=16.838 with estimated weight is $\widehat{w_1}$=0.213 and the mean of the second component was $\hat{\lambda}_2$= 30.475 with estimated weight is $\widehat{w_2}$=0.452 .While the mean of the third component was $\hat{\lambda}_3$ =47.390 with estimated weight is $\widehat{w_3}$=0.333 . This result can be supported by the density fitting of our selected model as shown in figure (9).

Conclusion

In this research, we have used two modified versions, under a Bayesian principle, for AIC and BIC to select the best mixture model. Under Bayesian framework, we have used one of the well-known MCMC procedure, which called the Gibbs sampler, to estimate the model parameters. We have performed a simulation study to estimate several Normal mixture models fitted to synthetic datasets. In addition, we have checked our criteria: AIC and BIC on the same synthetic datasets. These criteria have shown a better performance with respect to select the correct model. Finally, we have applied these criteria on real

Model	components	$\hat{w}_j$	$\hat{\hat{\lambda}}_j$	Log-likelihood	AIC	BIC
2	1 2	0.485 0.514	22.913 43.531	-197.990	201.990	205.732
3	1 2 3	0.213 0.452 0.333	16.838 30.475 47.390	-191.726	**197.726**	**203.339**
4	1 2 3 4	0.130 0.264 0.321 0.284	14.810 24.331 34.537 48.836	-191.878	199.878	207.363
5	1 2 3 4 5	0.091 0.151 0.237 0.233 0.286	13.306 21.181 27.166 36.129 48.329	-192.494	202.494	211.850
6	1 2 3 4 5 6	0.070 0.119 0.145 0.233 0.182 0.249	12.134 18.441 24.375 30.259 38.136 49.270	-192.998	204.998	216.225
7	1 2 3 4 5 6 7	0.067 0.102 0.109 0.179 0.148 0.186 0.206	11.982 17.813 23.297 27.775 32.735 40.105 50.572	-193.467	207.467	220.566

application involving accidents count data occurred in the Al Muthanna city using a Bayesian Poisson mixture model.

Table (6): Shows the results of the estimation parameters and model selection (AIC, BIC) of six models fitted to the real data (accident counts).

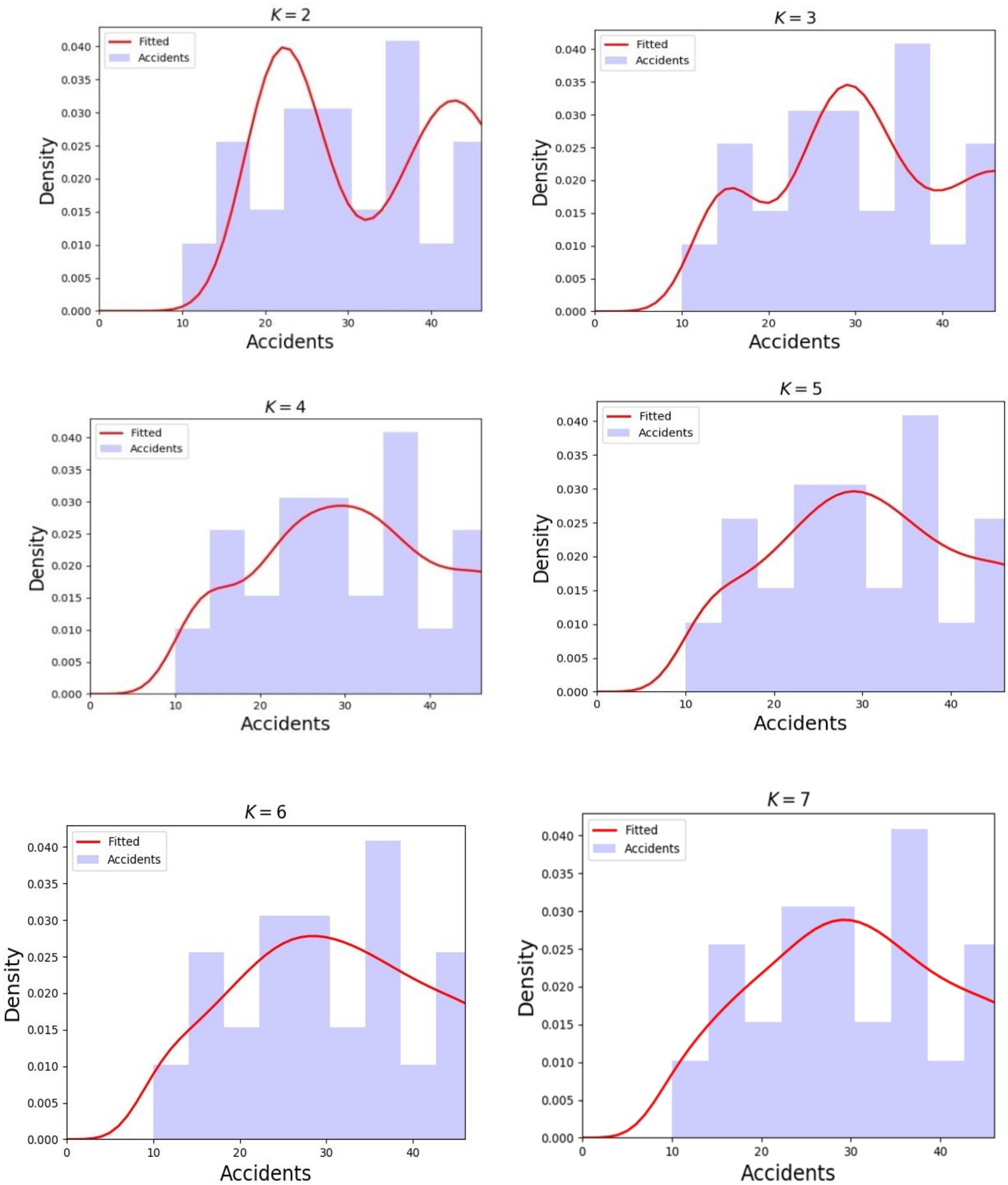

Figure (9): Shows the densities of six Poisson mixture models fitted the accidents count.

References

1- Akaike, H. (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control* 19: 716-723.

2- Fruhwirth-Schnatter, S., Celeux, G., & Robert, C. P. (Eds.). (2019). *Handbook of mixture analysis*. CRC Press.

3- Geman, S, and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *-IEEE Transaction on Pattern Analysis and Machine Intelligence 6*: 721-741.

4- Jasra, A., Holmes, C. C. and Stephens, D. A. (2005). Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling. *Statistical Science* 20(1): 50-67.

5- Kadhem, S. K., Hewson, P., & Kaimi, I. (2018). Using hidden Markov models to model spatial dependence in a network. *Australian & New Zealand Journal of Statistics*, 60(4), 423-446.

6- Marin, J. M., Mengersen, K., & Robert, C. P. (2005). Bayesian modeling and inference on mixtures of distributions. *Handbook of statistics*, 25, 459-507.

7- McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. New York: Wiley.

8- Richardson, S. and Green, P. J. (1997). On Bayesian analysis of mixtures with an unknown number of components. *Journal of the Royal Statistical Society, Series B* 59: 731-792.

9- Schwarz G. Estimating the dimension of a model. *Annals of Statistics*. (1978), 6:461–464.

10- Stephens, M. (2000). Bayesian analysis of mixture models with an unknown number of components - An alternative to reversible jump methods. *The Annals of Statistics* 28: 40-74.

11- Tanner, M. Y. and Wong, W. H. (1987). The calculation of posterior distribution by data augmentation. *Journal of the American Statistical Association* 67: 702-708.