Comparing the Estimation Methods of the Parameters for the Mixture Gamma Distribution(*)

Asst.Prof.Dr. Awat Sirdar Wahdi
Salahaddin University
College of Administration and Economics
awat.wahdi@su.edu.krd

Master student: Awaz Omer Ahmed
Salahaddin University
College of Administration and Economics
awaz.ahmed@su.edu.krd

ISSN 2709-6475 DOI: https://dx.doi.org/10.37940/BEJAR.2022.3.3.20

تأريخ استلام البحث 2022/5/23 **تأريخ قبول النشر 2022/6/22** **تأريخ النشر 2022/7/30**

Abstract

Gamma mixture models are widely used in financial and trustworthiness as well as to describe biological diversity and complex clustered or longitudinal data. Because of the intricacy of the model structure in these models, estimation is hard. In this paper we used two methods: expectation maximization (EM) and the Markov chain Monte Carlo (MCMC) Bayesian method for parameters to be estimated of the mixture gamma model by using R programming, and we compared the two methods by using simulation and by using mean square error (MSE). It shows that estimation of parameters of the mixture gamma distribution by using (EM) is better than the (MCMC) Bayesian method. The sample sizes that we used in this research are (500, 1000, and 1500).

Keywords: Mixture Gamma distribution, MCMC algorithm, EM algorithm.

مجلة اقتصاديات الأعمال
المجلد (3) العدد (3) 2022
الصفحات: 367-380

(*) Research extracted from a master's thesis

1. Introduction:

A mixed distribution, also known as a mixture density, is a distribution made up of two or more element distributions that have been weighted together. There are two types of component distributions: multivariate and statistical. Using mixed techniques, we model complicated clustered data and information with numerous sources of variation and polymorphisms. A mixed distribution, also known as a mixture density, is a distribution made up of two or more element distributions that have been weighted together (Susanto,*et.al.*,2019).

$$f(x|\theta) = \sum_{i=1}^{k} \pi_i \, GD(x|\alpha_i, \beta_i) \qquad \cdots (1)$$

The i-th component $GD(x|\alpha_i, \beta_i)$ is defined by:

$$GD(x|\alpha_i, \beta_i) = \frac{1}{\Gamma(\alpha_i)(\beta_i)^{\alpha_i}} x^{\alpha_i - 1} \exp\left\{-\frac{x}{\beta_i}\right\} \qquad \cdots (2)$$

If we have two components ($k = 2$) of a mixture gamma distribution, the model can be defined as follows: (Aljohani & Alfaer, 2021)

$$f(x|\theta) = \frac{\pi_1}{\Gamma(\alpha_1){\beta_1}^{\alpha_1}} x^{\alpha_1 - 1} exp\{-\frac{x}{\beta_1}\} + \frac{(1-\pi_1)}{\Gamma(\alpha_2){\beta_2}^{\alpha_2}} x^{\alpha_2 - 1} exp\{-\frac{x}{\beta_2}\} \qquad \cdots (3)$$

Where:

x: is the lifetime random process.

$\theta = (k, \pi, \alpha, \beta)$: parameters of mixture gamma distribution.

$\pi = (\pi_1, \cdots, \pi_k)$: is a collection of is a combination of blending

variables, $\pi_i > 0$ and $\sum_{i=1}^{k} \pi_i$, $\pi_2 = 1 - \pi_1$

$\alpha_i = (\alpha_1, \cdots, \alpha_k)$: shape parameters.

$\beta_i = (\beta_1, \cdots, \beta_k)$: scale parameters.

$x, \alpha_1, \alpha_2, \beta_1, \beta_2 > 0$

$\Gamma(\bullet)$: the Gamma function.

Figure (1) displays an example mixture gamma distribution with $\alpha = (12,1)$, $\beta = (3,2)$, and $\pi = (0.4,0.6)$.

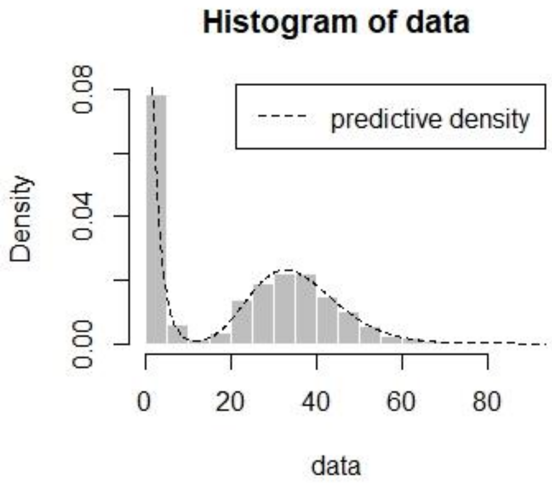

Figure (1): An instance of a two-component mixed gamma distribution.

The main issue in this research is to find a strong estimation of model parameter of mixture gamma distribution $\theta = (k, \pi, \alpha, \beta)$, and for more studies on the estimation of model parameters of mixture gamma distribution can be returned to each of the sources of the following:

In 2013 (Mohammadi et al.) a technique based on birth-death statistics that is Bayesian was introduced. Using MCMC approaches, the parameters are determined using the described Bayesian methodology and a parametric model based on a blend of Beta distributions. Many numerical results based on various simulation studies are used to explain this approach (Mohammadi,*et.al.*,2013)

In 2012 (Zou & Li), a fully Bayesian approach for such a generalized gamma mixture model (GΓMM) was presented. The model can be estimated to use the (MCMC) algorithm, which employs Gibbs and Metropolis-Hastings sampling and depends on the mixture model's missing data structure. A Monte Carlo simulation analysis using synthetic and real data was conducted to demonstrate the algorithm's superior performance (Zou & Li,2012).

In 2019 (Liu,*et.al.*) a Bayesian approach is based on birth-death statistics. The proposed probabilistic approach and the parameterization model based on a blend of beta distributions are used to determine the parameters using MCMC methods. Many numerical results based on various simulation studies are used to explain this approach (Liu,*et.al.*, 2019).

The aim of this research is to estimate parameters of mixture gamma distribution by using some estimation methods and comparing estimation methods for parameters by (MSE) to find the best estimation methods. The research problem is that we can't use a closed formula to estimate parameters, so that's why we used an algorithm.

So, in this research, we suggested employing a method for the purpose of finding a strong estimate of the parameters of this model and the method estimates.

2. Maximum likelihood Estimation Method:

Consider a random sample of size n generated from a mixture of two parameter gamma distributions. Use the general formulation of such a log-likelihood function for a mixed model and get:

$$logL(x|\theta) = \sum_{j=1}^{n} log\left[\frac{\pi_1}{\Gamma(\alpha_1)\beta_1{}^{\alpha_1}} x^{\alpha_1-1}\, exp\{-\frac{x}{\beta_1}\} + \frac{(1-\pi_1)}{\Gamma(\alpha_2)\beta_2{}^{\alpha_2}} x^{\alpha_2-1}\, exp\{-\frac{x}{\beta_2}\}\right] \quad \cdots (4)$$

As illustrated in, the ensuing likelihood calculations for estimating the parameters are non-linear and do not have any closed-form solutions (4). We cannot use maximum likelihood estimation. As a result, iterative processes or gradient search algorithms must be used to estimate the parameters (Lakshmi & Vaidyanathan,2016).

3. Bayesian Inference Estimation Method:

We assume the following for determining a Bayesian approach for generic distributions $B(.)$ based on a blend of Gamma distributions:

$$B(x|\theta) = \sum_{i=1}^{k} \pi_i\, G(x|\alpha_i, \beta_i) \quad \cdots (5)$$

Where $\theta = (k, \pi, \alpha, \beta)$. First, we apply a data augmentation approach, as is common in mixture models, by creating component indicator variables, Z_j, for each real, S_j.

$$P\left(Z_j = 1\middle|k, \pi\right) = \pi_i \quad , \quad i = 1, \dots, k \quad \cdots (6)$$

Suppose that the mixture Gamma parameters have a joint prior distribution. $\theta = (k, \pi, \alpha, \beta)$ could be factored as follows: (Richardson & Green,1997)

$$f(k, \pi, \alpha, \beta, z) \propto f(k)f(\pi|k)f(z|\pi, k)f(\alpha|k)f(\beta|k) \quad \cdots (7)$$

To find the prior distributions for mixture distribution parameters, For the mixture size k, we first suppose a truncated Poisson distribution, as illustrated below:

$$P(K = k) \propto \frac{\gamma^k}{k!}\ ,\ \ k = 1, \dots, k_{max} \qquad \cdots (8)$$

For the remaining model parameters, we assume proper but relatively diffuse distributions of the form: (Wiper,*etal.*,2001)

$$\left.\begin{array}{r} \pi|k \sim D(\phi_1, \dots, \phi_k) \\ \alpha_i|k \sim G(\nu, \upsilon)\ , i = 1, \dots, k \\ \beta_i|k \sim G(\eta, \tau)\ , i = 1, \dots, k \end{array}\right\} \qquad \cdots (9)$$

where $D(\phi_1, \dots, \phi_k)$ is a Dirichlet distribution with $\phi_i > 0$ parameters. Usually, we would set $\phi_i > 0$. for all $i = 1, \dots, k$, providing the weights an uniform U(0,1) prior.

Given k and x, it is simple to illustrate that perhaps the following are the subsequent conditionally distributions required for the MCMC approach:

$$P\big(Z_j = i \big| x, k, \pi, \alpha, \beta\big) \propto \pi_i G\big(x_j \big| \alpha_i, \beta_i\big),\ \ i = 1, \dots, k \qquad \cdots (10)$$

$$\pi|x, z, k \sim D(\phi_1 + n_1, \dots, \phi_k + n_k) \qquad \cdots (11)$$

$$\beta_i|x, z, k \sim G\left(\eta + n_i\alpha_i, \tau + \sum\nolimits_{j:z_j = i} x_j\right), i = 1, \dots, k \qquad \cdots (12)$$

$$f(\alpha_i|x, z, k, \beta) \propto \left(\frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)}\right)^{n_i} \left(\prod\nolimits_{j:z_j = i} x_j\right)^{\alpha_i} \alpha_i^{\nu - 1}\ e^{-\upsilon\alpha_i} \qquad \cdots (13)$$

Where $n_i = \#(z_j = i)$ for $i = 1, \dots, k$.

Permutation of the labels $i = 1, \dots, k$ has no effect on this mixed model. It is important to use distinct labeling for identifiability. Therefore, except otherwise indicated, assume that the π_i are increasing; as a result, the earlier parameter estimates, which were limited, are no longer valid for the set$\pi_1 < \cdots < \pi_k$. The product of the various gamma densities is k! times. In the case of mixture modeling with a finite number of components, there have been two primary techniques to sample the posterior distributions: one is the reversible jump MCMC methodology. The birth-death MCMC methodology is another option (Mohammadi,*et.al.*,2013).

Birth-death MCMC algorithm:

A basic mixture model for observations $X = x_1, x_2, \ldots, x_n$ a vector of latent indicator variables can be used to define$Z = z_1, z_2, \ldots, z_n$ as follows:

$$x_i|z_i = k \sim f_k(x|\theta) \qquad \cdots (14)$$

Using a BD-MCMC technique, take the sample from one of the parameter posterior distributions, $\theta = (k, \pi, \alpha, \beta)$.. This technique is based on the birth-death process. In this regard, normal mixes The parameters of the model are read as facts from a marked point process, and the mixture size, k, is varied so that the births and deaths of the components of the mixture occur in real time. The rates at which this occurs establish the stationary distribution of anything like the process (Stephens,2000).

In the birth-death process, mixed-item births occur at a constant rate, which can be made equal to the variable k from the statistical model.

The number of mixture components increases by one when a child is born. The weight of a new component is derived using a Beta distribution with parameters (1,k), with the applied to monitors selected from posterior distributions. The preceding component weights are correspondingly reduced to fit the new element, such that almost all of the values, including the new one, are equal. Add up to 1, i.e., $\pi_i = \frac{\pi_i}{1+\pi}$. Every mixed component's death rate is calculated as a likelihood ratio of the model with and without this component, as follows:

$$\Delta_j = \prod_{r=1}^{n} \left(\frac{B(x_r) - \pi_j g(x_r|\alpha_j, \beta_j)}{(1-\pi_j) B(x_r)} \right) \;,\; j = 1, \ldots, k \qquad \cdots (15)$$

At each point in time, the overall death rate of the process $\Delta = \sum_j \Delta_j$,is equal to the sum of the individual death rates. The number of mixture components is reduced by one when someone dies.

The total death rate, $\Delta = \sum_j \Delta_j$, of the process at any time is the sum of the individual death rates. A death decreases the number of mixture components by one.

The birth and death processes are Poisson processes that are independent of one another. As a result, the birth-death event time is exponentially distributed, with a mean of $\frac{1}{\Delta+\gamma}$. As a result, the probability of a birth or death is proportional to γ and Δ respectively. Define a BD-MCMC algorithm as follows:

Algorithm: (Mohammadi,*et.al.*,2013) (Shi,*et.al.*,2002) (Wiper,*et.al.*, 2001)**.**

Iterate the steps below starting with the initial values $k^{(0)}, \pi^{(0)}, \alpha^{(0)}, \beta^{(0)}$.

1. Run the birth-death process for a set period t_0 while keeping the birth rate γ constant.
   1.a. Calculate the death rates for each component, Δ_j, as well as the overall death rate, $\Delta = \sum_j \Delta_j$.
   1.b. Using an exponential distribution with a mean of $\frac{1}{\Delta+\gamma}$, time till the next jump is simulated.
   1.c. Continue if the run time is below t_0; else, progress to step 2.
   1.d. With probability, simulate the sort of jump: birth or death.

$$Pr(birth) = \frac{\gamma}{\Delta+\gamma}, \quad Pr(death) = \frac{\Delta}{\Delta+\gamma} \qquad \cdots (16)$$

   1.e. Adjust the components of the mixture.

Dependent on k, MCMC steps:

2. Update the variables by sampling from $z^{(i+1)} \sim z|t, k^{(i+1)}, \pi^{(i)}, \alpha^{(i)}, \beta^{(i)}$
3. By taking a sample of the weights, you can keep them up to date.
   $\pi^{(i+1)} \sim \pi|t, k^{(i+1)}, z^{(i+1)}$
4. for $r = 1, \ldots, k^{(i+1)}$
   4.a. Update β_r by sampling from $\beta_r^{(i+1)} \sim \beta_r|x, k^{(i+1)}, z^{(i+1)}$
   4.b. Update α_r
5. Set $i = i + 1$ and go to step 1.

The above-mentioned birth-death process is used in step one of the algorithms, choosing $t_0 = 1$ with a birth rate of the same magnitude as the parameter γ. Birth rates with higher values provide better mixing in practice, but they take longer to compute in the algorithm.

Steps 2, 3, and 4.a are standard Gibbs sampling moves, whereby the model parameters are updated conditional on the mixture size, k. The only complicated is step 4.2, incorporates a Metropolitan step to choose from the posterior distribution of α_i. present the parameters $G(\nu, \upsilon)$ for a Gamma distribution based on the geometry of the target distribution. The probability of a candidate being accepted α^*, using this proposal distribution, is:

$$\Pr(\alpha_r, \alpha^*) = min\left\{1, \left(\frac{\Gamma(\alpha_r)}{\Gamma(\alpha^*)}\right)^n \left(\prod_{j:z_j=r} \beta_r \, x_j\right)^{\alpha^*-\alpha_r}\right] \qquad \cdots (17)$$

The posterior distributions are used in Algorithm 3.1 to generate a sample. As a result, this BD-MCMC approach for time densities parameters may be used, $\theta = (k, \pi, \alpha, \beta)$. (Mohammadi,*et.al.*,2013)

To estimate density of time distribution use:

$$\hat{B}_r(t|\theta_r) = \sum_{i=1}^{\hat{k}_r} \hat{\pi}_{ri} G\left(t \mid \hat{\alpha}_{ri}, \hat{\beta}_{ri}\right) \quad \cdots (18)$$

4. Expectation-Maximization (EM) Algorithm Method:

The iterative approach of expectation–maximization (EM) is used to determine the maximum likelihood estimator of a parameter θ in a parametric probability distribution (Gupta & Chen,2011).

The EM algorithm is a method for computing maximum likelihood estimates from partial data without having to calculate the observed values log-likelihood explicitly (Steele,1996) (Dempster,*et.al.*,1977).

In mixture models, the maximum likelihood algorithm has been utilized as an effective method for estimating the parameters of the EM algorithm. In the EM framework, the observed data $x_1, \cdots, x_n$ is considered as incomplete data and latent class variables $z_1, \cdots z_g$ to be missing where $z_{ki} = z_k(x_i) = 1$ if observation x_i belongs to k th class and 0 otherwise for $k = 1, \dots, g$ and $i = 1, \dots, n$. By treating z as missing data, the EM algorithm is applied to mixed distributions. The E and M-steps of the EM algorithm are the two parts of the algorithm (Erişoğlu,*et.al.*, 2011).

The missing observation vector $z_i = (z_{1i}, \cdots, z_{gi})$ is examined in the EM algorithm, and its distributions function is provided as:

$$f(z_i) = \prod_{k=1}^{g} \pi_k^{z_{ki}} \quad \cdots (19)$$

Thus, the pdf of x_i, given z_i is:

$$f(x_i|z_i) = \prod_{k=1}^{g} (f_k(x_i|\theta_k))^{z_{ki}} \quad \cdots (20)$$

Gaussian random distributions in a group X and Z is defined by:

$$f(x_1, \cdots, x_n, z_1, \cdots, z_n) = \prod_{i=1}^{n} \prod_{k=1}^{g} (f_k(x_i|\theta_k))^{z_{ki}} \pi^{z_{ki}} \quad \cdots (21)$$

Thus, the log likelihood for complete data is given by

$$\ln L\,(\pi, \theta | x_1, \cdots, x_n, z_1, \cdots, z_n) = \sum_{i=1}^{n} \sum_{k=1}^{g} z_{ki} \ln\left(\pi_k f_k(x_i|\theta_k)\right) \quad \cdots (22)$$

E-step:

In the E-step because all z_i are treated as missing data, the expectation $E(z_{ki}|x_i)$ should be calculated to estimate the z_i .To obtain $E(z_{ki}|x_i)$ and need to calculate $f(z_{ki}|x_i)$ as follows:

$$f(z_{ki}|x_i) = \frac{(f(x_i;\theta_k))^{z_{ki}}\,\pi^{z_{ki}}}{\sum_{j=1}^{g}(f_j(x_i;\theta_j))^{z_{ji}}\,\pi^{z_{ji}}} \qquad \cdots (23)$$

$$\hat{z}_{ki} = E(z_{ki}|x_i) = \frac{\pi_k f(x_i|\theta_k)}{\sum_{j=1}^{g}\pi_j f(x_i|\theta_j)} \qquad \cdots (24)$$

The whole dataset log likelihood function's prediction is provided by:

$$E(\ln L) = \sum_{i=1}^{n}\sum_{k=1}^{g}\hat{z}_{ki}ln(\pi_k f_k(x_i;\theta_k)) \qquad \cdots (25)$$

For the mixture distribution model.

M- steps:

The M- steps in the EM algorithm, $E(\ln L)$is maximized with the restriction that $\sum_{k=1}^{g}\pi_k = 1.$ To maximize the $E(\ln L)$ with Lagrange method $\ln\hat{L}$ is obtained as:

$$\ln\hat{L} = \sum_{i=1}^{n}\sum_{k=1}^{g}\hat{z}_{ki}ln(\pi_k f_k(x_i;\theta_k)) - \lambda(\sum_{k=1}^{g}\pi_k - 1) \qquad \cdots (26)$$

To estimate the parameters which maximize equation (26) the first derivative of $\ln\hat{L}$ with respect to all parameters and set it to be zero. The estimated mixing proportion of subgroup k is defined by:

$$\hat{\pi}_k = \frac{1}{n}\sum_{i=1}^{n}\hat{z}_{ki}\ ,\ k = 1,\cdots,g \qquad \cdots (27)$$

In equation (19), if we choose $f_k(x|\theta_k)$ as the Gamma pdf of k th group for observation vector *x*, and can evaluate a mixture of Gamma distributions and mean lifetime defined by:

$$f(x|\Psi) = \sum_{k=1}^{g}\pi_k\frac{1}{\beta_k^{\alpha_k}\Gamma(\alpha_k)}x^{\alpha_k-1}\exp\left(-\frac{x}{\beta_k}\right) \qquad \cdots (28)$$

$$E(T) = \sum_{k=1}^{g}\pi_k\,\alpha_k\beta_k \qquad \cdots (29)$$

In this form, for k th subgroup in r th iteration, the maximum likelihood estimator of parameters α and β can be obtained with EM as:

$$\hat{\alpha}_{k,(r+1)} = \hat{\alpha}_{k,r} - \frac{\ln(\hat{\alpha}_{k,r}) - \Psi(\hat{\alpha}_{k,r}) - \ln\left(\frac{\sum_{i=1}^{n}\hat{z}_{ki}x_i}{\sum_{i=1}^{n}\hat{z}_{ki}}\right) + \frac{\sum_{i=1}^{n}\hat{z}_{ki}\ln(x_i)}{\sum_{i=1}^{n}\hat{z}_{ki}}}{\frac{1}{\hat{\alpha}_{k,r}} - \acute{\Psi}(\hat{\alpha}_{k,r})} \qquad \cdots (30)$$

$$\hat{\beta}_k = (\hat{\alpha}_k\sum_{i=1}^{n}\hat{z}_{ki})^{-1}\sum_{i=1}^{n}\hat{z}_{ki}x_i \qquad \cdots (31)$$

Where $\Psi(.)$ and $\acute{\Psi}$ are a digamma and trigamma functions respectively.

5. Simulations:

This section shows how to set up a simulation experiment, which is necessary for estimating the parameters of the mixture gamma distribution. The simulation method was used with different values for parameters to be estimated and in different sample sizes. The most important stage depends on it in the program of building simulation experience, the default values, namely specifying default values for the parameters $\theta = (k, \pi, \alpha, \beta)$, after that choosing the default values for the real parameters. In this research, eight models have been considered, which are arranged as follows:

Table (1): The default value of the scale parameter

Model	π_1	α_1	α_2	β_1	β_2
1	0.6	12	1	3	2
2	0.7	12	1	3	2
3	0.6	12	1	5	3
4	0.7	12	1	5	3
5	0.6	10	3	3	2
6	0.7	10	3	3	2
7	0.6	10	3	5	3
8	0.7	10	3	5	3

The sample size was chosen to determine the effect of sample size in deciding the accuracy and bitterness of the results obtained from the estimation methods used in this study. The samples have been taken at volumes characterized by (500, 1000, and 1500).

The results of the simulation and analysis to determine the best methods for estimating the parameters for the mixture gamma distribution are shown in the following tables. The results were obtained by utilizing a program written in the R-Programming language by the researcher.

Table (2): Estimates of all parameters of various estimation methods, for different sample size for model (1)

N	Measure	Method	$\pi_1 = 0.6$	$\alpha_1 = 12$	$\alpha_2 = 1$	$\beta_1 = 3$	$\beta_2 = 2$
500	Mean of Parameters	EM	0.59975	12.0659	1.01094	3.00646	1.99759
		Bayes	0.59822	12.0645	1.00177	3.00970	2.05561
	MSE	EM	0.00051	0.98778	0.00852	0.06364	0.06029
		Bayes	0.00101	0.98806	0.01502	0.06552	0.12799
1000	Mean of Parameters	EM	0.59942	12.0894	1.00436	2.99086	2.00127
		Bayes	0.59963	12.0885	0.99920	2.99177	2.03487
	MSE	EM	0.00023	0.53689	0.00409	0.03242	0.03062
		Bayes	0.00044	0.53767	0.00792	0.03338	0.06564

N	Measure	Method	$\pi_1 = 0.6$	$\alpha_1 = 12$	$\alpha_2 = 1$	$\beta_1 = 3$	$\beta_2 = 2$
1500	Mean of Parameters	EM	0.60022	12.0282	1.00500	2.99968	1.99349
		Bayes	0.60044	12.0285	1.00156	2.99979	2.01523
	MSE	EM	0.00016	0.36567	0.00286	0.02264	0.01937
		Bayes	0.00032	0.36537	0.00565	0.02380	0.03982

Table (3): Estimates of all parameters of various estimation methods, for different sample size for model (2)

N	Measure	Method	$\pi_1 = 0.7$	$\alpha_1 = 12$	$\alpha_2 = 1$	$\beta_1 = 3$	$\beta_2 = 2$
500	Mean of Parameters	EM	0.70001	12.1418	1.01632	2.98151	2.00311
		Bayes	0.69871	12.1321	1.01036	2.98616	2.07278
	MSE	EM	0.00043	0.91750	0.01368	0.05491	0.09164
		Bayes	0.00080	0.92308	0.02462	0.05716	0.20420
1000	Mean of Parameters	EM	0.70014	12.0483	1.00507	2.99483	1.99613
		Bayes	0.69931	12.0447	1.00021	2.99689	2.03463
	MSE	EM	0.00021	0.42480	0.00546	0.02709	0.03946
		Bayes	0.00039	0.43248	0.01035	0.02889	0.08530
1500	Mean of Parameters	EM	0.70010	12.0565	1.00300	2.99288	2.00696
		Bayes	0.70019	12.0543	1.00188	2.99174	2.02674
	MSE	EM	0.00013	0.30933	0.00389	0.01950	0.02759
		Bayes	0.00026	0.31167	0.00746	0.01973	0.05696

Table (4): Estimates of all parameters of various estimation methods, for different sample size for model (3)

N	Measure	Method	$\pi_1 = 0.6$	$\alpha_1 = 12$	$\alpha_2 = 1$	$\beta_1 = 5$	$\beta_2 = 3$
500	Mean of Parameters	EM	0.59981	12.1323	1.00767	4.97564	2.99951
		Bayes	0.60107	12.1304	1.00407	4.98012	3.07790
	MSE	EM	0.00046	1.03032	0.00859	0.17579	0.13101
		Bayes	0.00091	1.03249	0.01633	0.18237	0.28924
1000	Mean of Parameters	EM	0.59978	12.0927	1.00515	4.98218	3.00295
		Bayes	0.59989	12.0924	0.99578	4.98261	3.05554
	MSE	EM	0.00024	0.52377	0.00417	0.09065	0.06397
		Bayes	0.00048	0.52550	0.00792	0.09285	0.13924
1500	Mean of Parameters	EM	0.60003	12.0208	1.00286	5.00206	3.00332
		Bayes	0.60010	12.0210	0.99939	5.00269	3.03673
	MSE	EM	0.00017	0.30877	0.00279	0.05589	0.04615
		Bayes	0.00033	0.30948	0.00540	0.05911	0.09339

Table (5): Estimates of all parameters of various estimation methods, for different sample size for model (4)

N	Measure	Method	$\pi_1 = 0.7$	$\alpha_1 = 12$	$\alpha_2 = 1$	$\beta_1 = 5$	$\beta_2 = 3$
500	Mean of Parameters	EM	0.69981	12.1554	1.01504	4.96378	2.98486
		Bayes	0.69871	12.1447	1.01036	4.96723	3.06921
	MSE	EM	0.00042	0.98163	0.01158	0.16609	0.17803
		Bayes	0.00086	0.99583	0.02116	0.17424	0.35573

N	Measure	Method	$\pi_1 = 0.7$	$\alpha_1 = 12$	$\alpha_2 = 1$	$\beta_1 = 5$	$\beta_2 = 3$
1000	Mean of Parameters	EM	0.70012	12.0675	1.00873	4.98817	2.98736
		Bayes	0.69989	12.0650	1.00397	4.99378	3.04684
	MSE	EM	0.00022	0.43282	0.00498	0.07706	0.08443
		Bayes	0.00022	0.43282	0.00498	0.07706	0.08443
1500	Mean of Parameters	EM	0.69971	12.0195	1.00497	4.99962	3.00517
		Bayes	0.69930	12.0183	1.00259	5.00243	3.03676
	MSE	EM	0.00014	0.26774	0.00386	0.04803	0.06117
		Bayes	0.00028	0.26903	0.00741	0.05077	0.12599

Table (6): Estimates of all parameters of various estimation methods, for different sample size for model (5)

N	Measure	Method	$\pi_1 = 0.6$	$\alpha_1 = 10$	$\alpha_2 = 3$	$\beta_1 = 3$	$\beta_2 = 2$
500	Mean of Parameters	EM	0.59871	10.2104	3.04666	2.97835	2.00839
		Bayes	0.59723	10.1832	2.97601	2.99221	2.12006
	MSE	EM	0.00059	1.40377	0.14722	0.10835	0.10161
		Bayes	0.00117	1.56436	0.37709	0.11806	0.22089
1000	Mean of Parameters	EM	0.59829	10.1424	3.00872	2.97630	2.02109
		Bayes	0.59721	10.1342	2.97430	2.98235	2.07176
	MSE	EM	0.00031	0.62524	0.06602	0.04762	0.05435
		Bayes	0.00058	0.63732	0.12388	0.05066	0.10643
1500	Mean of Parameters	EM	0.59997	10.0590	3.00954	2.99687	2.00666
		Bayes	0.59919	10.05216	2.97832	3.000836	2.046706
	MSE	EM	0.00020	0.46252	0.04370	0.03636	0.03411
		Bayes	0.00039	0.474531	0.079905	0.038513	0.06953

Table (7): Estimates of all parameters of various estimation methods, for different sample size for model (6)

N	Measure	Method	$\pi_1 = 0.7$	$\alpha_1 = 10$	$\alpha_2 = 3$	$\beta_1 = 3$	$\beta_2 = 2$
500	Mean of Parameters	EM	0.69975	10.1397	3.07168	2.99069	2.01270
		Bayes	0.69504	10.0959	2.97773	3.01253	2.17922
	MSE	EM	0.00057	1.16153	0.18805	0.08886	0.15729
		Bayes	0.00116	1.25908	0.34703	0.10703	0.45177
1000	Mean of Parameters	EM	0.63623	10.1055	3.04412	2.98569	1.99268
		Bayes	0.63515	10.0990	3.00223	2.99071	2.05126
	MSE	EM	0.00437	0.61417	0.07420	0.04848	0.05563
		Bayes	0.00480	0.62863	0.14101	0.05112	0.11616
1500	Mean of Parameters	EM	0.69978	10.0448	3.03113	2.99701	1.99744
		Bayes	0.69888	10.0386	3.00450	3.00270	2.04635
	MSE	EM	0.00020	0.36030	0.06788	0.02930	0.05093
		Bayes	0.00036	0.37360	0.12430	0.03140	0.10341

Table (8): Estimates of all parameters of various estimation methods, for different sample size for model (7)

N	Measure	Method	$\pi_1 = 0.6$	$\alpha_1 = 10$	$\alpha_2 = 3$	$\beta_1 = 5$	$\beta_2 = 3$
500	Mean of Parameters	EM	0.60046	10.1727	3.06132	4.97254	2.9951
		Bayes	0.59826	10.1420	2.97427	4.99312	3.16748
	MSE	EM	0.00057	1.27713	0.13932	0.27931	0.20752
		Bayes	0.00111	1.39066	0.30669	0.31334	0.43552
1000	Mean of Parameters	EM	0.60058	10.0683	3.03019	4.99865	2.99985
		Bayes	0.60022	10.0589	3.01335	5.00881	3.04733
	MSE	EM	0.00029	0.60425	0.06091	0.13172	0.09356
		Bayes	0.00054	0.61147	0.11007	0.13791	0.18058
1500	Mean of Parameters	EM	0.59996	10.0773	3.01939	4.97982	3.00425
		Bayes	0.59923	10.0711	2.99742	4.98272	3.04487
	MSE	EM	0.00019	0.41795	0.04166	0.09412	0.06738
		Bayes	0.00035	0.42675	0.07656	0.10060	0.12602

Table (9): Estimates of all parameters of various estimation methods, for different sample size for model (8)

N	Measure	Method	$\pi_1 = 0.7$	$\alpha_1 = 10$	$\alpha_2 = 3$	$\beta_1 = 5$	$\beta_2 = 3$
500	Mean of Parameters	EM	0.69917	10.1290	3.08718	4.97661	2.99388
		Bayes	0.69609	10.0872	3.02103	5.00843	3.15639
	MSE	EM	0.00052	0.96838	0.19845	0.21589	0.28791
		Bayes	0.00103	1.03911	0.30834	0.24273	0.64654
1000	Mean of Parameters	EM	0.69919	10.0416	3.04102	5.00222	2.99848
		Bayes	0.69800	10.0151	3.00406	5.01997	3.08215
	MSE	EM	0.00024	0.46544	0.09408	0.10773	0.15028
		Bayes	0.00050	0.48962	0.16520	0.12058	0.28890
1500	Mean of Parameters	EM	0.70048	10.0367	3.02445	4.99709	2.99592
		Bayes	0.69921	10.0368	2.99948	5.00158	3.04399
	MSE	EM	0.00017	0.32614	0.05380	0.07647	0.08577
		Bayes	0.00034	0.33661	0.09876	0.08229	0.16771

Tables (2, 3,..., 9) show that the estimated values for all parameters using all studied estimation methods revealed that the mean is close to the default values for this parameter in all models and sample sizes.

For all models, we noticed that most of the averages for estimated parameters in all estimation methods are close to default values for all parameters, and the MSE declines with increasing sample sizes.

6. Conclusion:

In this paper, we proposed the mixture models and estimated the parameters of the mixture gamma distribution. Maximum likelihood estimations of mixture model parameters are nonlinear and have no conservation equation approaches, so we can't use them to estimate parameters. Instead, we should utilize iterative methods or algorithms to estimate mixture gamma parameter estimates.

The simulation study compares the EM algorithm and the MCMC Bayesian approach for estimating mixed gamma model coefficients. The EM algorithm better illustrates the MCMC Bayesian method by using mean square error (MSE). For all models, we notice that most of the averages of estimated parameters in all estimation methods are very close to the default values for this parameter. When sample size increases, the MSE also decreases.

7. Reference:

1. ALJOHANI, H. M. & ALFAER, N. M., 2021. Hybrid Censoring Schemes with Mixture Gamma Failure Distribution. Journal of Mathematics.
2. DEMPSTER, A. P., LAIRD, N. M. & RUBIN, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Methodological), 39, 1-22.
3. ERIŞOĞLU, Ü., ERIŞOĞLU, M. & EROL, H., 2011. A mixture model of two different distributions approach to the analysis of heterogeneous survival data. International Journal of Computational and Mathematical Sciences, 5, 75-79.
4. GUPTA, M. R. & CHEN, Y., 2011. Theory and use of the EM algorithm, Now Publishers Inc.
5. LAKSHMI, R. V. & VAIDYANATHAN, V., 2016. Parameter Estimation in Gamma Mixture Model using Normal-based Approximation. J. Stat. Theory Appl., 15, 25-35.
6. LIU, C., LI, H.-C., FU, K., ZHANG, F., DATCU, M. & EMERY, W. J., 2019. Bayesian estimation of generalized gamma mixture model based on variational em algorithm. Pattern Recognition, 87, 269-284.
7. MOHAMMADI, A., SALEHI-RAD, M.-R. & WIT, E. C., 2013. Using mixture of Gamma distributions for Bayesian analysis in an M/G/1 queue with optional second service. Computational Statistics, 28, 683-700.
8. RICHARDSON, S. & GREEN, P. J., 1997. On Bayesian analysis of mixtures with an unknown number of components (with discussion). Journal of the Royal Statistical Society: series B (statistical methodology), 59, 731-792.
9. SHI, J., MURRAY-SMITH, R. & TITTERINGTON, D., 2002. Birth-death MCMC methods for mixtures with an unknown number of components. Citeseer.
10. STEELE, B. M., 1996. A modified EM algorithm for estimation in generalized mixed models. Biometrics, 1295-1310.
11. STEPHENS, M., 2000. Bayesian analysis of mixture models with an unknown number of components-an alternative to reversible jump methods. Annals of statistics, 40-74.
12. SUSANTO, I., IRIAWAN, N. & KUSWANTO, H., 2019. Bayesian inference for the finite gamma mixture model of income distribution. Journal of Physics: Conference Series. IOP Publishing, 012077.
13. WIPER, M., INSUA, D. R. & RUGGERI, F., 2001. Mixtures of gamma distributions with applications. Journal of computational and graphical statistics, 10, 440-454.
14. ZOU, Y.-H. & LI, H. C., 2012. MCMC estimation of finite generalized gamma mixture model. IEEE International Geoscience and Remote Sensing Symposium. IEEE, 5935-5938.