Journal of Statistical Theory and Applications, Vol. 15, No.1 (March 2016), 25-35

Parameter Estimation in Gamma Mixture Model using Normal-based Approximation

R. Vani Lakshmi

Department of Statistics
Pondicherry University,
Puducherry – 605014, India
rvanilakshmi@outlook.com

V. S. Vaidyanathan

Department of Statistics
Pondicherry University
Puducherry – 605014, India
vaidya.stats@gmail.com

Received 28 March 2015

Accepted 6 October 2015

Abstract

Gamma mixture models have wide applications in hydrology, finance and reliability. Parameter estimation in this class of models is a challenging task owing to the complexity associated with the model structure. In this paper, a novel approach is proposed to estimate the parameters of Gamma mixture models using Wilson-Hilferty normal-based approximation method. The proposed methodology uses a popular clustering algorithm for Gaussian mixtures namely, MCLUST and a confidence interval based search approach to obtain the estimates. The methodology is implemented on simulated as well as real-life datasets and its performance is compared with gammamixEM() function available in R.

Keywords: Gamma Mixture Model; gammamixEM(); Maximum Likelihood; MCLUST; Mean Square Error; Wilson-Hilferty Approximation.

2010 Mathematics Subject Classification: 62F10, 62H30

1. Introduction

Mixture models are primarily used in modelling population involving two or more distributions. Over the last few decades, it has found abundant applications in the domains of finance, hydrology, medical science, psychology, reliability and life testing. In particular, Gaussian mixture models have gained prominence within the model-based clustering framework. However, in recent years, there has been a marked increase in the use of non-Gaussian mixture models pertaining to skewed and asymmetric data. Examples in this direction include finite mixtures of binomial and Poisson distributions (McLachlan and Peel [19]), Weibull distribution (Jiang and Murthy [9]), inverse Weibull distribution (Sultan et al. [22]), Multivariate t-distribution (Andrews and McNicholas [1]) and Multivariate skew t-distribution (Lee and McLachlan [15]). For a general discussion on parameter estimation in finite mixture of distributions and related applications, one may refer to Everitt and Hand [5], Titterington et al. [23], McLachlan and Peel [19] and Melnykov and Maitra [20].

Published by Atlantis Press

In this paper, we consider a finite mixture of two-parameter gamma distributions with probability density function (pdf) given by

$$f(x;\alpha,\beta) = \sum_{i=1}^{g} \pi_i\, f_i(x;\alpha,\beta) \tag{1.1}$$

where $f_i(x;\alpha,\beta) = \frac{x^{(\alpha_i-1)}\exp-\left(\frac{x}{\beta_i}\right)}{\beta_i{}^{\alpha_i}\Gamma\alpha_i}$, $x > 0, \alpha_i > 0, \beta_i > 0, i = 1,2,\ldots,g.$

The corresponding cumulative distribution function (cdf) is given by

$$F(x;\alpha,\beta) = \sum_{i=1}^{g} \pi_i\, F_i(x;\alpha,\beta) \tag{1.2}$$

where $F_i(x;\alpha,\beta) = \frac{1}{\Gamma\alpha_i}\gamma\left(\alpha_i,\frac{x}{\beta_i}\right), i = 1,2,\ldots,g$ and $\gamma(.)$ represents the lower incomplete gamma function; α_i and β_i denote respectively, the shape and scale parameters of the i-th component of the mixture distribution; $\pi_1, \pi_2, \ldots, \pi_g$ denote mixture proportions or weights that satisfy the conditions (i) $0 < \pi_i < 1\ \forall\, i = 1,2,..,g$ and (ii) $\sum_{i=1}^{g}\pi_i = 1$. Here, g denotes the number of components in the mixture.

One of the earliest attempts of parameter estimation in finite Gamma Mixture Models (GMM) was by John [10] in which method of moments and Maximum Likelihood (ML) were used to estimate the parameters of a two-component mixture with common shape parameter. Kanno [12] carried out ML estimation using Newton-Raphson method in a finite mixture of two-component GMM with known location parameters. Webb [25] used Expectation-Maximization (EM) algorithm in estimating the parameters of a finite GMM with an application in target recognition. Wiper et al. [27] implemented a Bayesian density estimation method based on GMM. Bowman and Shenton [3] discuss an iterative algorithm to find solutions for ML estimators of gamma mixtures. Evin et al. [6] used two-component mixtures of gamma, Gumbel and normal distribution respectively to model flood data using Bayesian methodology. Mallya et al. [16] used GMM under a Bayesian framework in probabilistic drought classification.

It is pertinent to note that the approaches used for parameter estimation in the above-mentioned references involve either Bayesian methodology or iterative mathematical computations that depend on 'seed' values. As an alternative, a novel approach is proposed in this paper to estimate the parameters in finite GMM by transforming the gamma variate to normal using Wilson-Hilferty (WH) approximation. The resulting Gaussian setup is used to identify the components. Parameter estimates are then obtained by using a confidence interval based search approach. The proposed methodology is advantageous in the sense that, parameter estimation becomes less complicated because of the resulting Gaussian structure. Moreover, the proposed approach does not require an 'initial guess' of the number of clusters.

Rest of the paper is organized as follows. Section 2 discusses the difficulties in obtaining ML estimates in GMM and the limitations of gammamixEM() function (available under mixtools package in R). Subsequently, the proposed methodology is elucidated. Section 3 provides numerical illustrations of the proposed methodology and compares the performance of the resulting estimates with that of gammamixEM() based on simulated as well as real-life datasets. Section 4 concludes the paper with discussion.

2. Parameter estimation in Gamma Mixture Models

Consider a random sample of size n from a finite mixture of two-parameter gamma distributions with pdf as defined in (1.1). Using the general representation of log-likelihood function for finite mixture model given in McLachlan and Peel [19], we get

$$\log \mathrm{L}\,(\mathrm{x}; \alpha, \beta) = \sum_{\mathrm{j}=1}^{\mathrm{n}} \log\left(\sum_{\mathrm{i}=1}^{\mathrm{g}} \pi_{\mathrm{i}} \left(\frac{\mathrm{x_j}^{(\alpha_{\mathrm{i}}-1)} \exp -\left(\frac{\mathrm{x_j}}{\beta_{\mathrm{i}}}\right)}{\beta_{\mathrm{i}}{}^{\alpha_{\mathrm{i}}} \Gamma\alpha_{\mathrm{i}}} \right) \right) \tag{2.1}$$

From (2.1), it can be easily seen that the resulting likelihood equations for estimating the $(3\mathrm{g}-1)$ parameters are non-linear and have no closed form solutions. Thus, to estimate the parameters one has to resort to iterative procedures or gradient search algorithms. In the recent years, EM algorithm proposed by Dempster et al. [4] has been extensively used in the context of estimating the parameters of finite mixture models. The algorithm iteratively calculates the ML estimates and mixture proportions based on initial value specification for the parameters. However, as pointed out by McLachlan and Basford [17], the performance of the algorithm is sensitive to starting values of the model parameters. A poor choice of starting values may considerably affect the resulting estimates. Also, the estimates may not correspond to global maxima. Despite its limitations, EM algorithm is very popular in the context of mixture models. Various modifications of EM algorithm are available in the literature to address these limitations. One may refer to McLachlan and Krishnan [18] for an elaborate discussion on EM algorithm and its variants.

The gammamixEM() proposed by Benaglia et al. [2] provides estimates for the parameters of GMM through an iterative process. To begin the process, the function requires either initial values for the parameters and mixture proportions or the number of components (g) in the mixture model. In the latter case, the initial values for model parameters are estimated by method of moments after partitioning the data into g groups based on visual binning. At each successive step, the function provides estimates for the $(3\mathrm{g}-1)$ model parameters using ML equations and calculates the value of the resulting log-likelihood. The process is repeated till convergence in log-likelihood is obtained. It should be noted that knowledge of either the number of components or mixture proportions is a requirement for implementing gammamixEM(). Also, the use of method of moments for providing initial estimates for model parameters cannot be relied always. Hence, we propose a confidence interval based search approach to estimate the parameters of GMM using WH approximation and MCLUST algorithm.

WH approximation introduced by Wilson and Hilferty [26] provides normal approximation to a chi-square random variable through cube root transformation. Among other competing methods available, this transformation provides better approximation (see Johnson et al. [11]) and is sensitive to outliers. Using the fact that chi-square and gamma distributions are related, Krishnamoorthy et al. [13] proposed the following result that extends WH approximation to two-parameter gamma distribution.

"Let $\mathrm{X}_{\alpha,\beta}$ denote a two-parameter gamma random variable. Then, $X_{\alpha,\beta}^{\frac{1}{3}}$ is distributed as Gaussian with mean $\mu = \frac{\beta^{\frac{1}{3}}\,\Gamma\left(\alpha+\frac{1}{3}\right)}{\Gamma\alpha}$ and variance $\sigma^2 = \frac{\beta^{\frac{2}{3}}\,\Gamma\left(\alpha+\frac{2}{3}\right)}{\Gamma\alpha} - \mu^2$."

This result has been applied in estimation of the survival probabilities of lifetime units (Krishnamoorthy and Thomas Mathew [14]) and in estimating floods (Rao and Hamed [21]).

MCLUST is a model-based clustering strategy for multivariate Gaussian mixture models. It incorporates the principles of hierarchical clustering, EM algorithm and Bayesian Information Criterion (BIC) to estimate the parameters of Gaussian mixture models. Primarily, it involves the following steps.

1. Implement agglomerative hierarchical clustering for the Gaussian mixture model to obtain classifications for g mixture components. By default, it can produce up to nine components.
2. Run EM algorithm corresponding to each initial classification obtained in Step 1.
3. Compute BIC for each resulting mixture model and identify the one for which BIC is maximum.

It is less dependent on starting values and has proved to be an efficient clustering mechanism even in the presence of noise, outliers and missing data. For more details on MCLUST, one may refer to Fraley and Raftery [7]. An implementation of MCLUST in R can be found in Fraley et al. [8].

The motivation for using WH approximation in the present work is to transform the underlying gamma components to Gaussian setup. This would facilitate the application of MCLUST to the resulting Gaussian mixture. Parameter estimation is then carried out through a confidence interval based search approach on each of the resulting components.

2.1 Proposed methodology

Consider a random sample of size n from a finite mixture of two-parameter gamma distributions with g components. First, WH approximation is applied to transform the samples to Gaussian setup. MCLUST (available under mclust package in R) is then applied to estimate the mixture proportions and also to identify component memberships. MCLUST also produces estimates for μ_i and σ_i^2, say $\hat{\mu}_i$ and $\hat{\sigma}_i^2$, $i = 1,2, \dots, g$, the mean and variance of the resulting g Gaussian components. Using these estimates, one can solve the following non-linear equations obtained through WH approximation to determine the estimates for $(\alpha_i, \beta_i), i = 1,2, \dots, g$.

$$\frac{\beta_i^{\frac{1}{3}}\Gamma\left(\alpha_i+\frac{1}{3}\right)}{\Gamma\alpha_i} = \hat{\mu}_i \tag{2.2}$$

$$\frac{\beta_i^{\frac{2}{3}}\Gamma\left(\alpha_i+\frac{2}{3}\right)}{\Gamma\alpha_i} = \hat{\sigma}_i^2 + \hat{\mu}_i^2 \tag{2.3}$$

From (2.2), we get $\beta_i = \left[\frac{\hat{\mu}_i(\Gamma\alpha_i)}{\Gamma\left(\alpha_i+\frac{1}{3}\right)}\right]^3$. Substituting for β_i in (2.3) yields

$$\frac{\Gamma\alpha_i\Gamma\left(\alpha_i+\frac{2}{3}\right)}{\left[\Gamma\left(\alpha_i+\frac{1}{3}\right)\right]^2} = \frac{\hat{\sigma}_i^2}{\hat{\mu}_i^2} + 1 \tag{2.4}$$

The above non-linear equation does not have a closed form solution for α_i and hence one has to resort to numerical methods. However, in-built functions to solve non-linear equation(s) available in R (uniroot.all() in rootSolve package, nleqslv() in nleqslv package), MATLAB (fzero(), fsolve()) and Mathematica (Solve()) either fail to provide solution for (2.4) or result in poor estimates. This may be attributed to the presence of gamma function in (2.4). As an alternative, one can directly maximize the log-likelihood L^* corresponding to every component under the Gaussian setup with $\mu_i = \frac{\beta_i^{\frac{1}{3}}\Gamma\left(\alpha_i+\frac{1}{3}\right)}{\Gamma\alpha_i}$ and $\sigma_i^2 = \frac{\beta_i^{\frac{2}{3}}\Gamma\left(\alpha_i+\frac{2}{3}\right)}{\Gamma\alpha_i} - \mu_i^2$ to obtain the estimates for (α_i, β_i). Towards this, we make use of $\hat{\mu}_i$ to construct a $100(1 - \delta)\%$ Confidence Interval (CI) for μ_i where $\delta \in (0,1)$ is a fixed value that determines the coverage probability. For small values of δ, this CI will contain μ_i with high probability. A subset S containing all values of (α_i, β_i) which lies within this CI is obtained via an extensive search in the two-dimensional space using the relation, $\mu_i = \frac{\beta_i^{\frac{1}{3}}\Gamma\left(\alpha_i+\frac{1}{3}\right)}{\Gamma\alpha_i}$. Finally, the estimates for (α_i, β_i) are identified by maximizing the log-likelihood function L^* within the respective subset S. One may also use the estimates of $\sigma_i^2, i = 1,2, \dots, g$ to construct CI in a similar manner and thereby obtain estimates for the parameters. However, this may be computationally intensive. The algorithm for implementing the proposed methodology is presented below.

Notations:

X : Random sample of size n from a finite mixture of g gamma distributions.
Y : Cube root of X.
π_i: Mixture proportions.
C : $100(1 - \delta)\%$ CI for μ_i.
I : Factor of incrementation.
L^* : Log-likelihood function corresponding to every component under Gaussian setup.

```
Step 1 : Define $Y = \sqrt[3]{X}$
Step 2 : Run MCLUST on Y
Step 3 : Obtain the estimates for $\pi_i, \mu_i, \sigma_i^2, i = 1,2,3,\ldots,g$
Begin loop
Step 4 : For each component i
        4.1 Evaluate C
        4.2 Identify all $(\alpha_i, \beta_i)$ satisfying $\mu_i = \frac{\beta_i^{\frac{1}{3}}\Gamma\left(\alpha_i+\frac{1}{3}\right)}{\Gamma\alpha_i}$ within C
        4.3 Define $S = \left(Range(\alpha_i), Range(\beta_i)\right)$
Step 5: For X in i
        5.1 Specify I
        5.2 Evaluate $L^*$within S by I
        5.3 Determine $(\hat{\alpha}, \hat{\beta})$ that maximizes $L^*$
End loop
```

The factor of incrementation I defines the 'jumps' in S at which L^* is evaluated. Smaller values of I result in an exhaustive search for parameter values within S. An implementation of the above algorithm will provide estimates for the $(3g - 1)$ parameters of GMM.

3. Numerical Illustrations

In this section, the proposed methodology is implemented on simulated and real-life datasets. A simulation study is carried out for a two-component mixture of gamma distributions with equal mixture proportions. The parameter choices considered for this study are presented in Table 1. Mixture data of sizes $n = 50, 100$ and 500 respectively are generated and the proposed methodology is implemented for 5000 Monte Carlo (MC) runs using a program developed in R version 3.1.0. One more example is also considered that uses a simulated dataset of sample size $n = 300$ having three components with equal mixture proportions as presented in Table 2. Through this simulation study, we attempt to assess the performance of the proposed algorithm vis-à-vis gammamixEM() in terms of Bias, Standard Deviation (SD) and Mean Square Error (MSE) of the estimates.

Table 1: Parameter choices for two-component mixture

Case	Component 1		Component 2	
	α_1	β_1	α_2	β_2
1	9.4	3.2	2.1	1.3
2	8.9	3.6	7.5	3.1
3	5.2	3.6	6.5	2.9

Table 2: Parameter choice for three-component mixture

Component 1		Component 2		Component 3	
α_1	β_1	α_2	β_2	α_3	β_3
2.7	1.2	5.6	3.7	9.2	1.3

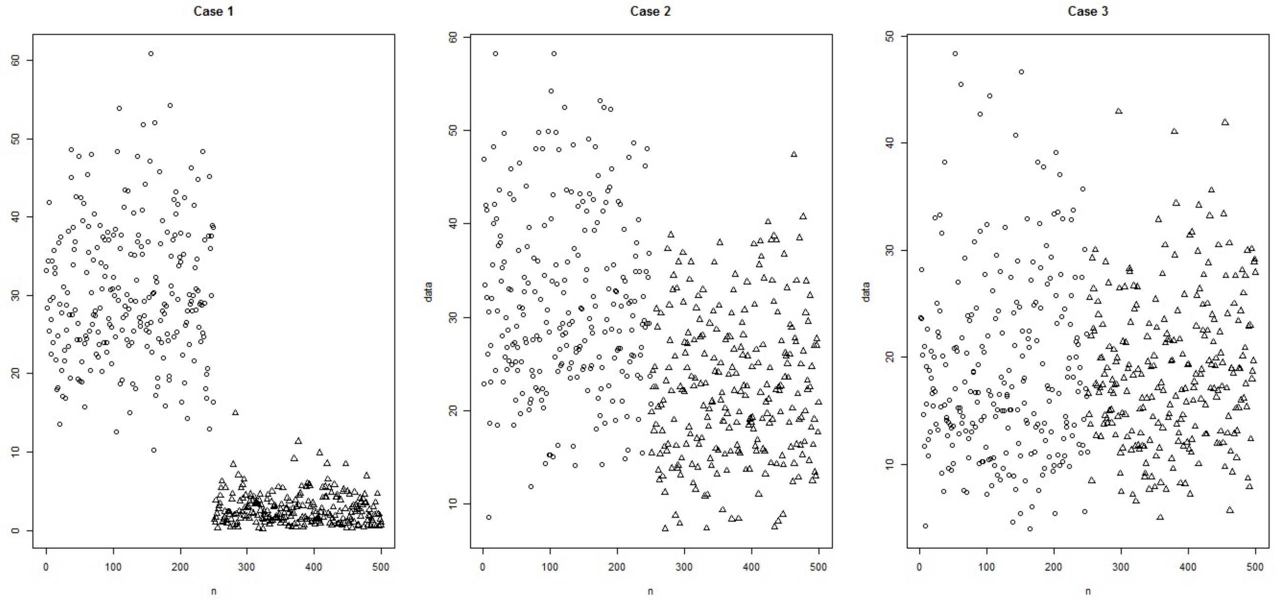

Figure 1: Plots of mixture data with 500 observations corresponding to various choices of parameter values

Figure 1 represents the plots of the simulated two-component mixture data with 250 observations from each component for various cases in Table 1. From the plots, it is evident that the three cases considered in Table 1 represent well-separated, mild-overlapping and high-overlapping components respectively. Thus, the parameter choices considered in this study intends to assess the performance of the proposed approach in the presence of well-separated as well as overlapping clusters.

For implementing the proposed algorithm, the values for δ and I are chosen as 0.001 and 0.1 respectively. This would ensure a large coverage probability for the CI resulting in a wider search space and also provide an exhaustive search to identify the estimates accurately. For implementing gammamixEM(), the number of components in the dataset is chosen as its initial value. The results obtained corresponding to proposed methodology and gammamixEM() under various cases are presented respectively in Tables 3 to 5. The measures suffixed with _G and _P in the tables denote those obtained under gammamixEM() and proposed approach respectively. The notation 'e' in the table values denotes the standard exponential notation used for representing extremely small/large numerical values in decimal form.

Table 3: Bias, SD and MSE of estimates for Case 1

n	Measure	Component 1			Component 2		
		$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\pi}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\pi}_2$
50	Bias_G	269.1932	0.2988	-0.0014	7.4796	0.5544	0.0014
	SD_G	9.63e+03	2.9078	0.1157	250.5569	2.8966	0.1157
	MSE_G	9.27e+07	8.5426	0.0134	6.28e+04	8.6957	0.0134
	Bias_P	-1.2622	0.4788	0.0025	-0.1336	0.1105	-0.0025
	SD_P	1.1342	0.6242	0.0315	0.3212	0.2872	0.0315
	MSE_P	2.8794	0.6188	0.0033	0.1210	0.0947	0.0033
100	Bias _G	94.6000	0.2998	0.0072	4.28e+03	0.2058	-0.0072
	SD_G	3.28e+03	2.3327	0.1053	3.00e+05	1.9185	0.1053
	MSE_G	1.08e+07	5.5304	0.0111	9.01e+10	3.7221	0.0111
	Bias_P	-0.6414	0.2450	-0.0002	-0.0209	0.0392	0.0002
	SD_P	0.8097	0.3843	0.0142	0.3282	0.2139	0.0142
	MSE_P	1.0668	0.2077	0.0012	0.1081	0.0473	0.0012
500	Bias_G	3.5329	0.4184	0.0142	0.4100	0.0841	-0.0142
	SD_G	44.8854	2.2063	0.1095	2.9906	1.3798	0.1095
	MSE_G	2.03e+03	5.0419	0.0122	9.1100	1.9105	0.0122
	Bias_P	0.0411	0.0026	-0.0008	-0.4150	0.7729	0.0008
	SD_P	0.5841	0.2273	0.0028	0.5374	0.8415	0.0028
	MSE_P	0.3428	0.0517	0.0004	0.4609	1.3054	0.0004

Table 4: Bias, SD and MSE of estimates for Case 2

n	Measure	Component 1			Component 2		
		$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\pi}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\pi}_2$
50	Bias_G	140.1000	-1.4127	0.0004	1.47e+04	-1.1854	-0.0004
	SD_G	1.96e+03	1.7224	0.3076	8.01e+05	1.6628	0.3076
	MSE_G	3.87e+06	4.9619	0.0946	6.42e+11	4.1694	0.0946
	Bias_P	0.0639	0.6729	-0.0114	0.4938	-0.4924	0.0114
	SD_P	1.2290	1.1167	0.2235	1.7469	1.2224	0.2235
	MSE_P	1.5141	1.6995	0.0501	3.2948	1.7365	0.0501
100	Bias _G	122.1000	-1.1795	-0.0057	31.1100	-0.8701	0.0057
	SD_G	3.28e+03	1.6540	0.3221	2.74e+02	1.7132	0.3221
	MSE_G	1.10e+07	4.1264	0.1038	7.59e+04	3.6917	0.1038
	Bias_P	0.1490	0.4937	-0.0174	0.6274	-0.6920	0.0174
	SD_P	1.0504	1.0208	0.2117	1.5839	0.9327	0.2117
	MSE_P	1.1252	1.2856	0.0451	2.9019	1.3486	0.0451
500	Bias _G	1.06e+05	-0.3975	0.0120	8.0861	-0.3920	-0.0120
	SD_G	4.34e+06	1.5642	0.2973	26.6376	1.3891	0.2973
	MSE_G	1.88e+13	2.6042	0.0885	7.75e+02	2.0827	0.0885
	Bias_P	-0.0961	0.8768	-0.0286	0.9535	-0.5315	0.0286
	SD_P	1.2588	1.0224	0.1241	1.6987	1.0989	0.1241
	MSE_P	1.5934	1.8138	0.0162	3.7942	1.4898	0.0162

Table 5: Bias, SD and MSE of estimates for Case 3

n	Measure	Component 1			Component 2		
		$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\pi}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\pi}_2$
50	Bias_G	185.8000	-1.8638	-0.0006	31.4700	-1.2805	0.0006
	SD_G	3.27e+03	1.3691	0.3081	169.9681	1.4484	0.3081
	MSE_G	1.07e+07	5.3476	0.0949	2.99e+04	3.7371	0.0949
	Bias_P	3.2358	-0.3848	-0.0051	0.4283	-0.8872	0.0051
	SD_P	1.4821	1.2346	0.2305	1.8978	0.9949	0.2305
	MSE_P	12.6663	1.6719	0.0531	3.7845	1.7768	0.0531
100	Bias _G	43.7600	-1.5619	0.0132	42.540	-1.0322	-0.0132
	SD_G	167.1973	1.3762	0.3241	692.6855	1.4303	0.3241
	MSE_G	2.99e+04	4.3332	0.1052	4.82e+05	3.1108	0.1052
	Bias_P	3.3516	-0.4623	-0.0068	0.5817	-0.9928	0.0068
	SD_P	1.2925	0.9843	0.2215	1.6568	0.7687	0.2215
	MSE_P	12.9037	1.1823	0.0491	3.0829	1.5765	0.0491
500	Bias _G	61.0100	-0.9441	0.0200	10.3147	-0.5788	-0.0200
	SD_G	1.90e+03	1.3193	0.3023	106.1100	1.3425	0.3023
	MSE_G	3.60e+06	2.6315	0.0918	1.14e+04	2.1369	0.0918
	Bias_P	2.9917	-0.0473	-0.0074	0.1618	-0.4478	0.0074
	SD_P	1.5957	1.2196	0.1491	1.9482	1.5322	0.1491
	MSE_P	11.4960	1.4893	0.0223	3.8210	2.5478	0.0223

The results corresponding to the three-component simulated dataset is presented below.

Table 6: Bias, SD and MSE of estimates for three-component mixture

n	Measure	Component 1			Component 2			Component 3		
		gammamixEM()								
		$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\pi}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\pi}_2$	$\hat{\alpha}_3$	$\hat{\beta}_3$	$\hat{\pi}_3$
300	Bias_G	3.1042	-0.1033	-0.0267	426.5088	-2.2011	0.0033	185.3341	1.9024	0.0334
	SD_G	56.0000	0.7235	0.1308	3.93e+03	1.5918	0.2221	5.00e+03	1.8250	0.2323
	MSE_G	3.14e+03	0.5340	0.0180	1.56e+07	7.3780	0.0493	2.50e+07	6.9488	0.0549
		Proposed Approach								
300	Bias_P	0.1575	-0.1470	-0.0098	0.9523	-1.6896	0.1827	-0.0822	1.9459	-0.1629
	SD_P	0.2829	0.2458	0.0210	0.7915	0.3252	0.0538	0.8881	0.5498	0.0599
	MSE_P	0.1048	0.0820	0.0006	1.5334	2.9604	0.0351	0.7954	4.0887	0.0312

It should be noted that for each MC run of the proposed algorithm, MCLUST produces estimates for the means of the components under Gaussian setup. Based on this, CI is constructed and the search space for estimating the scale and shape parameters of the gamma mixture is identified. For purpose of brevity, these results are not presented. Based on simulation results, it can be observed that

1. the estimates produced by the proposed methodology are ‘closer’ to the true parameter values when compared with that of gammamixEM(). This is evident from the small values of Bias corresponding to the proposed approach in Tables 3 to 6.
2. the values of SD and MSE obtained by the proposed approach are smaller than the corresponding values of gammamixEM() for all cases. This indicates that the proposed approach produces reliable estimates.
3. even in the presence of high-overlapping components, the proposed methodology performs better than gammamixEM() as seen from Table 5.

The under-performance of gammamixEM() may be attributed to the fact that it uses visual binning to identify initial classification thereby resulting in poor estimates. This can be seen in the extremely large values of Bias, SD and MSE for gammamixEM() in Tables 3 to 6. However, initial classification in the proposed approach is done through MCLUST that makes use of hierarchical clustering.

The misclassification rates corresponding to 5000 MC runs of Cases 1, 2 and 3 of both gammamixEM() and the proposed approach are presented in Table 7 below.

Table 7: Misclassification rates of estimates for two-component mixture

Case	n=50		n=100		n=500	
	gamma mixEM()	Proposed Approach	gamma mixEM()	Proposed Approach	gamma mixEM()	Proposed Approach
1	0.036	0.005	0.029	0.004	0.029	0.004
2	0.400	0.365	0.415	0.361	0.456	0.410
3	0.505	0.493	0.505	0.507	0.504	0.492

It can be observed that the proposed approach has a lower misclassification rate than gammamixEM() specifically for Case 1 (well-separated components) and Case 2 (mild-overlapping components).

3.1 Real-life illustration

Here, we consider a real-life illustration discussed in Türkan and Çaliş [24] where gamma mixture models with two components have been used to model actual time-of-death (in days) for 60 irradiated mice. The proposed methodology as well as gammamixEM() is implemented on the dataset and the results are presented below.

Table 8: Parameter Estimates for Irradiated Mice Data

Methodology	Estimates of Component 1			Estimates of Component 2		
	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\pi}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\pi}_2$
gammamixEM()	10.391	28.709	0.433	93.098	6.816	0.567
Proposed Approach	12.200	23.700	0.434	89.000	7.100	0.566

A comparison of the estimates reveals that both the methods produce ‘closer’ results. In order to get better insights in terms of bias and standard errors, bootstrap samples of size 100 are generated from the dataset. The average value, bias, standard errors and 95% confidence interval of the resulting bootstrap estimates under the proposed methodology and gammamixEM() are reported in Table 9.

From Table 9, it is observed that the proposed approach results in smaller bias and standard errors in comparison to gammamixEM(). Also, the width of the 95% bootstrap CI is smaller for the proposed approach. This indicates that the proposed methodology produces reliable estimates when compared with gammamixEM().

Table 9: Bootstrap Results for Irradiated Mice Data

Measure	Component 1			Component 2		
	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\pi}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\pi}_2$
gammamixEM()						
Average	48.3599	31.9493	0.4729	401.5431	8.5686	0.5271
Bias	37.9689	3.2403	0.0399	308.4451	1.7526	-0.0399
Standard Error	254.1724	17.9057	0.1723	1511.8318	9.9516	0.1723
95% CI	(5.67, 46.35)	(4.07,75.49)	(0.0780, 0.8457)	(8.10,1409.40)	(0.435,61.389)	(0.1543, 0.9220)
Proposed Approach						
Average	14.7820	22.4960	0.4260	76.3820	9.0880	0.5740
Bias	2.5820	-1.2040	-0.0080	-12.6180	1.9880	0.0080
Standard Error	6.1183	7.8748	0.0622	19.4198	3.3288	0.0622
95% CI	(8.40, 27.33)	(6.95, 36.00)	(0.2670,0.5174)	(42.75,94.00)	(6.700,12.551)	(0.4826,0.7330)

4. Discussion

This paper introduces a heuristic methodology to estimate the parameters of a finite mixture of gamma distributions. The approach assumes that all the $(3g - 1)$ parameters of the g-component mixture model are unknown. It places no restriction on the parameters and does not involve the use of calculus, complex iterative optimization methods and re-parameterization procedures. It encompasses the concepts of WH approximation, MCLUST algorithm and ML method to arrive at estimates in an efficient manner using a confidence interval based search approach. WH approximation paves way for transforming data from gamma to Gaussian setup. MCLUST produces estimates for the mixture proportions and identifies the components under Gaussian setup. The use of BIC for model accuracy in the MCLUST paradigm ensures that optimum number of components is identified.

From the numerical illustrations, it is clear that the performance of the proposed method is better than gammamixEM() in terms of Bias, SD and MSE across a wide range of parameter choices. The choice of parameters used in the simulation study ensures low, moderate and high overlapping between components. Unlike gammamixEM() that uses visual binning to obtain initial classification, the proposed approach identifies the components through MCLUST using hierarchical clustering which leads to reliable estimates for the parameters. As indicated in Table 7, the misclassification rate of the proposed approach is lower than that of gammamixEM(). However, computational time associated with the proposed method is relatively high owing to the fact that it searches a large two-dimensional space.

To conclude, the proposed approach is novel and it produces estimates that are reliable. Further research is being pursued to implement the proposed methodology in a finite mixture of three-parameter gamma distributions involving location parameters.

Acknowledgment

The authors express their sincere gratitude to the referees for their valuable comments and suggestions which helped in improving the article.

References

[1] J. L. Andrews and P.D. McNicholas, Model-based clustering, classification, and discriminant analysis via mixtures of multivariate t-distributions, *Stat. Comput.* **22**(5) (2012) 1021-1029.

[2] T. Benaglia, D. Chauveau, R. David Hunter, Derek Young, mixtools: An R Package for Analyzing Finite Mixture Models, *Journal of Statistical Software* **32**(6) (2009) 1-29.

[3] K.O. Bowman and L.R. Shenton, Maximum Likelihood Estimators for Normal and Gamma Mixtures, *Far East Journal of Theoretical Statistics* **20**(2) (2006) 217-240.

[4] A.P. Dempster, N.M. Laird and D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Stat. Soc. B (Met.)* **39**(1) (1977) 1-38.

[5] B.S. Everitt and D.J. Hand, *Finite mixture distributions* (Chapman and Hall, London, 1981).

[6] G. Evin, J. Merleau and L. Perreault, Two-component mixtures of normal, gamma, and Gumbel distributions for hydrological applications, *Water. Resour. Res*. **47**(8) (2011),W08525, doi:10.1029/2010WR010266.

[7] C. Fraley and A.E. Raftery, How many clusters? Which clustering method? Answers via model-based cluster analysis, *Comput. J.* **41**(8) (1998) 578-588.

[8] C. Fraley, A.E. Raftery T. Brendan Murphy and Luca Scrucca, mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. (*Technical Report No. 597, Department of Statistics, University of Washington*, 2012). Available at http://www.stat.washington.edu/research/reports/2012/tr597.pdf.

[9] R. Jiang and D. Murthy, Two sectional models involving three Weibull distributions, *Qual. Reliab. Eng. Int*. **13**(2) (1997) 83-96.

[10] S. John, On identifying the population of origin of each observation in a mixture of observations from two gamma populations, *Technometrics* **12**(3) (1970) 565-568.

[11] N.L. Johnson, Samuel Kotz, N. Balakrishnan N, *Continuous Univariate Distributions (Volume-1)* (John Wiley & Sons, NY, 1994)

[12] R. Kanno, Maximum Likelihood estimation of parameters for a mixture of two gamma distribution, *Report of Statistical Application Research*, JUSE **29**(3) (1982) 14-24.

[13] K. Krishnamoorthy, T. Mathew and S. Mukherjee, Normal-based methods for a gamma distribution: Prediction and Tolerance Intervals and Stress-Strength Reliability, *Technometrics* **50**(1) (2008) 69-78.

[14] K. Krishnamoorthy and Thomas Mathew, *Statistical Tolerance Regions: Theory, Applications and Computation*, (John Wiley & Sons, USA, 2009)

[15] S. Lee and G.J. McLachlan, Finite Mixtures of Multivariate Skew t-Distributions: Some Recent and New Results, *Stat. Comput.* **24**(2) (2014) 181-202.

[16] G. Mallya, S. Tripathi and R.S. Govindaraju, Probabilistic drought classification using gamma mixture models, *J. Hydrol*. **526** (2015) 116-126.

[17] G. McLachlan and K.E. Basford, *Mixture Models: Inference and Applications to Clustering* (Marcel Dekker, NY, 1988).

[18] G. McLachlan and T. Krishnan, *The EM algorithm and extensions* (John Wiley & Sons, USA, 2007).

[19] G. McLachlan and D. Peel, *Finite mixture models* (John Wiley & Sons, USA, 2004).

[20] V. Melnykov and R. Maitra, Finite mixture models and model-based clustering, *Statistics Surveys* **4** (2010) 80-116.

[21] A.R. Rao and K. Hamed, *Flood Frequency Analysis* (CRC Press, USA, 1999).

[22] K.S. Sultan, M.A. Ismail and A.S. Al-Moisheer, Mixture of two inverse Weibull distributions: Properties and estimation, *Comput. Stat. Data. Anal.* **51**(11) (2007) 5377-5387.

[23] D.M. Titterington, A.F. Smith and U.E. Makov, *Statistical analysis of finite mixture distributions* (John Wiley & Sons, UK, 1985).

[24] A.H. Türkan and N. Çaliş, Comparison of two-component mixture distribution for heterogeneous survival datasets: A Review Study, İSTATİSTİK: *Journal of the Turkish Statistical Association* **7**(2) (2014) 33-42.

[25] A.R. Webb, Gamma mixture models for target recognition, *Pattern Recogn*. **33**(12) (2000) 2045-2054.

[26] E.B. Wilson and M.M Hilferty, The distribution of chi-square, *Proceedings of the National Academy of Sciences of the United States of America* **17** (1931) 684-688.

[27] M. Wiper, D.R. Insua and F. Ruggeri, Mixtures of gamma distributions with applications, *J. Comput. Graph. Stat.* **10**(3) (2001) 440-454.