Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/120410

Estimating the Mixing Proportion of Mixture of two Chi-Square Distributions

Parviz Nasiri and AmirAbbas Azarian*

Department of Statistics, University of Payam Noor, 19395-4697 Tehran, I.R. Iran

Received: 9 Sep. 2017, Revised: 23 Apr. 2018, Accepted: 28 Apr. 2018
Published online: 1 Jul. 2018

Abstract: In mixture distributions, mixing parameter is an important parameter. In this paper we estimate this parameter in mixture of two chi-square distributions using method of moments, EM-algorithm and Bayesian approach. These estimators are compared empirically; and their bias is investigated using numerical techniques. At the end, we conjecture that EM-algorithm is better than method of moments, but bayesian approach is better than method of moments and EM-algorithm. Analysis of a real data set is considered for illustrative purposes.

Keywords: Method of Moments, EM-Algorithm, mixture of two independent Chi-Square distributions, Bayesian approach

1 Introduction

Mixture models in general and parametric mixture models in particular are very useful methods for modelling a population and have a lot of useful examples of applications in medicine, industry and economics. Wang, Tan and Louis [10] used this models for modelling time-to-event data to evaluate treatment effects in randomized clinical trials. Teel, Park and Sampson [11] considered the use of an EM algorithm for fitting finite mixture models when mixing component size is known. Render and Walker [1] showed that the EM algorithm is a useful method for predicting parameters of mixture distributions. Baudry and Celeux [12] showed that, although maximum likelihood through the EM algorithm is widely used to estimate the parameters in hidden structure models and has certain drawbacks, it is a good method in these problems.
Zaman et al. [6] introduced mixture of chi-square distributions using Poisson elements. Chen, Ponomareva and Tamer [8] introduced likelihood inference in some finite mixture models and discussed different situations for mixture models.
Hory and Martin [3] proposed a mixture of chi-square distributions to illustrate the time distribution of light in the imaging; also, in the same year, Martin, Hory and Huchard introduced this distribution to describe an unstructured distribution model for distributing light's distribution time. Sahuguede et al. [7] proposed this distribution for modeling the incoming light in imaging systems with multiple inputs, but did not provide a method for estimating the parameters of this distribution. Rindskopf [5] showed by examples that abnormal distributions may not necessarily be considered in the form of mixed distributions, but also showed that mixture distributions are not necessarily distributions with two or more modes, they are modeled as mixed distributions and obtained better estimates for describing the community. Jaspers, Komrek and Aerts [14] estimated mixing weights of multivariate mixture weights using Bayesian approach.
In this paper, we want to introduce a mixture chi-square distribution and estimate its mixing proportion using EM-algorithm and compare this estimator with method of moment. We also provide a real example of the use of mixture of two chi-square distributions for age of patients of a hospital. Nasiri and Azarian [13] defined Mixture of two chi-square distributions as below:

The mixture of two chi-square distributions (MTChD) has its pdf as:

$$f(x;\Theta) = \tau f_1(x;\theta_1) + (1-\tau) f_2(x;\theta_2), \qquad 0<\tau<1$$

Where $\Theta = (\tau, \theta_1, \theta_2)$ and $f_i(x;\theta_i)$, the density of the ith component, is given by

$$f_i(x;\theta_i) = \frac{x^{\frac{\theta_i}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{\theta_i}{2}} \Gamma\left(\frac{\theta_i}{2}\right)}, \qquad x \geq 0,\ \theta_i > 0,\ i = 1,2$$

* Corresponding author e-mail: pnasiri45@yahoo.com

2 Em Algorithm

Suppose Y is a p-dimensional random vector with probability density function $g(y;\Theta)$ where $\Theta=(\theta_1,\theta_2,\ldots,\theta_d)^\tau$ is the vector of unknown parameters in probability density function of Y. The likelihood function of Θ, which calculated on the observed values of y is

$$L(\Theta)=g(y;\Theta)$$

In some cases vector Y may include incomplete data due to presence of missing data or censored data; but in other situations it may be a complete data from a mixture of two or more distributions where the proportions of allocation to different distributions is unknown and we wish to find these proportions. In these cases, it is very difficult to use the method of maximum likelihood, but EM-algorithm is a useful method. EM-algorithm is an iterative method for estimating parameters using maximum likelihood method in case of incomplete data. Since $\log\ L(\Theta)$ includes incomplete data, suppose that $\log L_c(\Theta)$is the logarithm of the likelihood function of the complete data. Suppose that $\log L_c(\Theta)$ is the logarithm of the likelihood function of the complete data. We want to maximise expectation of $\log\ L_c(\Theta)$ with condition of complete data of Y. In other words, let $\Theta^{(k)}$ be the value of Θ after k^{th} iteration. At step $k+1$, the E and M steps are specified as follows.

1.Step E (Expectation): compute $Q(\Theta;\Theta^{(k)})$ where

$$Q(\Theta;\Theta^{(k)})=E_{\Theta^{(k)}}\{\log L_c(\Theta)|y\}$$

2.Step M (Maximization): choose $\Theta^{(k+1)}$ for each $\Theta\in\Omega$, (Ω is parameter space) that maximizes $Q(\Theta;\Theta^{(k)})$ with respect to Θ. In other words,

$$Q(\Theta^{(k+1)};\Theta^{(k)})\geq Q(\Theta;\Theta^{(k)})\quad\forall\,\Theta\in\Omega$$

Steps E and M are repeated until the sequence $L(\Theta^{(k)})$ converges.

In 1977 Rubin, Dumpster and Layrd showed that the likelihood function of incomplete data $(L(\Theta))$ after one iteration is non-decreasing. So the sequence $L(\Theta^{(k)})$ converges uniformly to some L^*, where L^* is a stationary value for Θ^*, which implies $\frac{\partial L(\Theta)}{\partial\Theta}=0$, or equivalently $\frac{\partial \log L(\Theta)}{\partial\Theta}=0$. In some cases, it is possible that Θ^* is a local maximum and in very rare cases it may be a saddle point and not a local maximum or minimum. Value of Θ^* depends on the initial value $\Theta^{(0)}$. It should be noted that iterations in EM-algorithm increase the likelihood, and converges to a value under some general conditions.

3 Estimating Parameters of a Mixture Chi-Square Distribution

In this section estimators of mixing parameter of mixture of two chi-square distribution will be introduced. Our approach is finding estimators using EM-algorithm, method of moments and Bayesian approach.

3.1 Estimation using EM-Algorithm

Suppose we have a population formed from a mixture of two independent chi-Square distributions, and also suppose that one part of population has chi-square distribution with parameter θ_1 and other part has chi-square distribution with parameter θ_2 independent of the first distribution. Let τ_1 be the unknown proportion of population in the first part, and $\tau_2=1-\tau_1$ be the proportion of population in the second part. In this situation, we have a mixture of two independent chi-square distributions. As we mentioned above we can use EM-algorithm for estimating the parameters of this population because τ is unknown. To formulate this population, we have:

$$\begin{aligned}L(\theta;x,z)&=\prod_{i=1}^{n}\left(\sum_{j=1}^{2}I(z_i=j)\tau_j f(x_i,\theta_j)\right)\\&=\prod_{i=1}^{n}\left(\sum_{j=1}^{2}I(z_i=j)\tau_j\left(\frac{x_i^{\frac{\theta_j}{2}-1}e^{-\frac{x_i}{2}}}{2^{\frac{\theta_j}{2}}\Gamma\left(\frac{\theta_j}{2}\right)}\right)\right)\end{aligned}$$

where $I(z_i=j)$ is an indicator function identifying the distribution of i^{th} observation and n is the number of observations. $L(\theta;x,z)$ is the likelihood function for a mixture distribution with mixing proportions τ_1 and τ_2. If we denote the logarithm of likelihood function by $l(\theta;x,z)$, then

$$\begin{aligned}l(\theta;x,z)=\sum_{j=1}^{2}\sum_{i=1}^{n}I(z_i=j)\Bigg[&\log(\tau_j)+\left(\frac{\theta_j}{2}-1\right)\log(x_j)\\&-\frac{x_i}{2}-\frac{\theta_j}{2}\log 2-\log\left(\Gamma\left(\frac{\theta_j}{2}\right)\right)\Bigg]\end{aligned}$$

In this case, we have a mixture of two independent chi-square distributions (MTChD) with parameters (τ,θ_1,θ_2).

In this case, steps E and M of EM-algorithm are as follows.

E **Step:**

Suppose that $\theta^{(t)}$ is the current parameter value. The distribution of z_i is obtained using Bayes rule and is proportional to chi-square parameters and proportions included in $\theta^{(t)}$. In other words,

$$\begin{aligned}\tau_{j,i}^{(k)}&:=P(z_i=j|X_i=x_i;\theta^{(t)})\\&=\frac{\tau_j^{(t)}f(x_i;\theta_j^{(t)})}{\tau_1^{(t)}f(x_i;\theta_1^{(t)})+\tau_2^{(t)}f(x_i;\theta_2^{(t)})}\end{aligned}$$

So the E-step results in

$$\begin{aligned} Q(\theta|\theta^{(t)}) =& E(\log L(\theta;x,z)) \\ =& \sum_{i=1}^{n}\sum_{j=1}^{2} T_{j,i}^{(t)}\Big[\log T_j + \frac{\theta_j}{2}\log x_i - \log x_i \\ & -\frac{x_i}{2} - \frac{\theta_j}{2}\log 2 - \log\left(\Gamma\left(\frac{\theta_j}{2}\right)\right)\Big] \end{aligned}$$

M **Step:**

As mentioned above, we maximize the log-likelihood obtained in E-step with respect to all the parameters. We find the maximum of $Q(\theta|\theta^{(t)})$ with respect to τ_1, v_1 and v_2. Parameters of this distribution can be estimated independently. Since $\tau_1+\tau_2=1$, we have:

$$\begin{aligned} \tau^{(t+1)} =& \arg\max_{\tau} Q(\theta|\theta^{(t)}) \\ =& \arg\max_{\tau}\Big\{\left[\sum_{i=1}^{n}\tau_{1,i}^{(t)}\right]\log\tau_1 \\ & + \left[\sum_{i=1}^{n}\tau_{2,i}^{(t)}\right]\log\tau_2\Big\} \end{aligned}$$

And hence

$$\tau_j^{(t+1)} = \frac{\sum_{i=1}^{n} T_{j,i}^{(n)}}{\sum_{i=1}^{n}(T_{1,i}^{(n)}+T_{2,i}^{(n)})} = \frac{1}{n}\sum_{i=1}^{n} T_{j,i}^{(t)}$$

So, τ_1 can be estimated from this equation. For estimating parameter v_1, we have

$$\begin{aligned} (\vartheta_1^{(t+1)}) =& \arg\max_{\vartheta_1} Q(\theta|\theta^{(t)}) \\ =& \arg\max_{\vartheta_1}\sum_{i=1}^{n} T_{1,i}^{(t)}\Big[\log T_1^{(t)} + \left(\frac{\theta_1^{(t)}}{2}-1\right)\log x_i \\ & -\frac{x_i}{2} - \frac{\theta_1^{(t)}}{2}\log 2 - \log\left(\Gamma\left(\frac{\theta_1^{(t)}}{2}\right)\right)\Big] \end{aligned}$$

If we let

$$\begin{aligned} k =& \sum_{i=1}^{n} T_{1,i}^{(t)}\log\tau_1^{(t)} \\ & + \left(\frac{\theta_1^{(t)}}{2}-1\right)\sum_{i=1}^{n} T_{1,i}^{(t)}\log x_i - \sum_{i=1}^{n} T_{1,i}^{(t)}\frac{x_i}{2} \\ & - \sum_{i=1}^{n} T_{1,i}^{(t)}\left(\frac{\theta_1^{(t)}}{2}\log 2 - \log\left(\Gamma\left(\frac{\theta_1^{(t)}}{2}\right)\right)\right) \end{aligned}$$

Differentiate k with respect to v_1 and set the derivative equal to zero to obtain

$$\Gamma(\alpha)(\Psi(\alpha))^2 = \frac{\sum_{i=1}^{n} T_{1,i}^{(t)}\log\frac{x_i}{2}}{\sum_{i=1}^{n} T_{1,i}^{(t)}}; \quad \text{where } \alpha = \frac{\theta_1^{(t)}}{2}$$

In the same as manner, we have

$$\Gamma(\beta)(\Psi(\beta))^2 = \frac{\sum_{i=1}^{n} T_{2,i}^{(t)}\log\frac{x_i}{2}}{\sum_{i=1}^{n} T_{2,i}^{(t)}}; \quad \text{where } \alpha = \frac{\theta_2^{(t)}}{2}$$

Here $\Psi(\alpha)$ is $(\Gamma(\alpha))'$. This equation should be solved numerically and there is no algebraic solution.

3.2 Estimating Using Method of Moments

As mentioned above, we can use EM-algorithm to estimate unknown parameters of a distribution when the simple methods like maximum likelihood estimators using differentiation or other routine methods do not work. However, there are other methods for estimating unknown parameters like the method of moments, which uses the sample moments and population moments to find estimates of parameters. In this method, parameters are estimated by equating sample moments with population moments. In mixture of two chi-square distributions, method of moments estimates the proportion of mixing τ_1 from the first population. Thus, we have

$$\widetilde{\tau} = \frac{\overline{x}-\theta_2}{\theta_1-\theta_2}$$

Where $\overline{x}$ is the sample mean.

3.3 Bayesian Approach

For estimating mixing parameter τ in MTChD, we can assume Beta distribution as prior distribution of τ due to the fact that $\tau \in (0,1)$. Suppose that τ has prior distribution $Beta(\alpha,1)$.

$$\pi(\tau) = \alpha\tau^{\alpha-1}$$

If we observe only one sample of MTChD, we have:

$$f(x;\Theta)\pi(\tau) = \alpha\tau^{\alpha-1}\left[\tau\frac{x^{\frac{\theta_1}{2}-1}e^{\frac{x}{2}}}{2^{\frac{\theta_1}{2}}\Gamma\left(\frac{\theta_1}{2}\right)} + (1-\tau)\frac{x^{\frac{\theta_2}{2}-1}e^{\frac{x}{2}}}{2^{\frac{\theta_2}{2}}\Gamma\left(\frac{\theta_2}{2}\right)}\right]$$

And marginal distribution of X is:

$$m(x)=\int_0^1 \alpha\tau^{\alpha-1}\left[\tau\frac{x^{\frac{\theta_1}{2}-1}e^{\frac{x}{2}}}{2^{\frac{\theta_1}{2}}\Gamma\left(\frac{\theta_1}{2}\right)}+(1-\tau)\frac{x^{\frac{\theta_2}{2}-1}e^{\frac{x}{2}}}{2^{\frac{\theta_2}{2}}\Gamma\left(\frac{\theta_2}{2}\right)}\right]d\tau$$

$$=\frac{x^{\frac{\theta_1}{2}-1}e^{\frac{x}{2}}}{2^{\frac{\theta_1}{2}}\Gamma\left(\frac{\theta_1}{2}\right)}\int_0^1 \alpha\tau^{\alpha}d\tau+\frac{x^{\frac{\theta_2}{2}-1}e^{\frac{x}{2}}}{2^{\frac{\theta_2}{2}}\Gamma\left(\frac{\theta_2}{2}\right)}\int_0^1 \alpha\tau^{\alpha-1}(1-\tau)d\tau$$

$$=f_1(x,\theta_1)\int_0^1 \alpha\tau^{\alpha}d\tau+f_2(x,\theta_2)\int_0^1 \alpha\tau^{\alpha-1}(1-\tau)d\tau$$

$$=f_1(x,\theta_1)\frac{\alpha}{\alpha+1}+f_2(x,\theta_2)\left[1-\frac{\alpha}{\alpha+1}\right]$$

$$=f_1(x,\theta_1)\frac{\alpha}{\alpha+1}+f_2(x,\theta_2)\frac{1}{\alpha+1}$$

And posterior distribution of τ is:

$$\pi(\tau|x)=\frac{\alpha\tau^{\alpha-1}\left[\tau\frac{x^{\frac{\theta_1}{2}-1}e^{\frac{x}{2}}}{2^{\frac{\theta_1}{2}}\Gamma\left(\frac{\theta_1}{2}\right)}+(1-\tau)\frac{x^{\frac{\theta_2}{2}-1}e^{\frac{x}{2}}}{2^{\frac{\theta_2}{2}}\Gamma\left(\frac{\theta_2}{2}\right)}\right]}{\frac{x^{\frac{\theta_1}{2}-1}e^{\frac{x}{2}}}{2^{\frac{\theta_1}{2}}\Gamma\left(\frac{\theta_1}{2}\right)}\frac{\alpha}{\alpha+1}+\frac{x^{\frac{\theta_2}{2}-1}e^{\frac{x}{2}}}{2^{\frac{\theta_2}{2}}\Gamma\left(\frac{\theta_2}{2}\right)}\frac{1}{\alpha+1}}$$

$$=\frac{\alpha\tau^{\alpha-1}[\tau f_1(x,\theta_1)+(1-\tau)f_2(x,\theta_2)]}{f_1(x,\theta_1)\frac{\alpha}{\alpha+1}+f_2(x,\theta_2)\frac{1}{\alpha+1}}$$

$$=\frac{\alpha\tau^{\alpha}f_1(x,\theta_1)}{\alpha f_1(x,\theta_1)+f_2(x,\theta_2)}+\frac{\alpha\tau^{\alpha}(1+\alpha)(1-\tau)f_2(x,\theta_2)}{\alpha f_1(x,\theta_1)+f_2(x,\theta_2)}$$

With use of square error loss function, mean of this distribution is the Bayes estimator for mixing parameter of MTChD.

4 Numerical Study

In last sections, we described the E-step and M-step of EM-algorithm, method of moments and Bayesian method for estimating the mixing parameter of mixture of two independent chi-square distributions with mixing proportion τ from the first population. In this section, we want to find the estimates, by simulating. For this purpose, In table 1, we simulated a mixture of two independent chi-square distributions with parameters $\theta_1=2$ and $\theta_2=8$ with mixing parameters τ=0.15, 0.35, 0.55, 0.75 and sample sizes $N=20,40,60,80,100$. Then mixing parameter τ was estimated useig EM-algorithm, method of moments and Bayesian method (with prior α= 0.2 for Bayesian method). We check these estimates by means of bias and MSE. EM-algorithm represents better estimates than method of moments in almost all cases and its bias decreased when mixing proportion and sample size increased. Also MSE of EM-algorithm is less than both method of moments and Bayes method in all cases. On the other hand, Bayes method represents better estimates than EM-algorithm for mixing proportions τ=0.15, 0.35 and 0.55, but for mixing proportion $\tau=0.75$, EM-algorithm has zero bias and smaller MSE than Bayes method.

In tables 2 and 3, similar process like table 1 were conducted but there were only parameters of chi-square distributions and prior α had changed. In table 2, we proposed $\theta_1=3$ and $\theta_2=9$ and prior α= 0.4 for Bayesian method. And in table 3, we proposed $\theta_1=4$ and $\theta_2=10$ and prior α= 0.6 for Bayesian method. In all cases similar results like table 1 were achieved.

So we can conclude that EM-algorithm has comparative advantage than method of moments, but for small mixing proportions, Bayes method represents better estimates.

5 Real Data

In the last section, we investigate statistical distribution of age of pediatrics patients referred to a hospital in Mashhad, North-east of Iran. Histogram of age of a random sample of 110 patient (74 boy and 36 girl) of this hospital represented in figure 1. As it is shown, our sample is unimodal and has positive skewness. So we can fit a chi-square distribution for the society. In figure 2, we check goodness of fit of a chi-square distribution with 4 degrees of freedom for our sample using Q-Q-plot and our idea can be accepted. But we know that patients of this hospital are both boys and girls. So we can divide population in to 2 parts and modeling the population as a mixture of two distributions. In figure 3, histogram of age of boys and girls are represented. So we can propose a chi-square distribution for age of boys and another chi-square distribution for age of girls of this hospital. Q-Q-plot in figure 4, confirms the chi-square distribution with 5 degrees of freedom for age of boys, and figure 5, confirms a chi-square with 4 degrees of freedom for girls. So we can use MTChD for modelling age of these patients.

$$f(x)=0.33\chi^2_{(4)}+0.67\chi^2_{(5)}$$

Where mixing proportion $\tau=0.33$ is the real proportion of girls from all patients.

We calculated mixing proportion using EM-algorithm, method of moments.

Table 4 represents the results. As we expected from the previous section, EM-algorithm represents better estimates than method of moments in point of view of bias and MSE, but its bias is more than Bayes estimate

Table 1: MSEs of the EM-algorithm, method of moments and Bayesian estimators with $\theta_1 = 2$, $\theta_2 = 8$

τ	N	EM-Algorithm		Method of Moments		Bayse Method ($prior\ \alpha = 0.2$)	
		BIAS	MSE	BIAS	MSE	BIAS	MSE
0.15	20	0.64733	0.00214	0.69188	0.00353	0.46255	0.05613
	40	0.60864	0.00185	0.78855	0.00438	0.45393	0.05799
	60	0.67825	0.00231	0.56448	0.00254	0.45185	0.05809
	80	0.59258	0.00175	0.66880	0.00334	0.44447	0.05954
	100	0.58738	0.00171	0.65537	0.00323	0.44179	0.05995
0.35	20	0.37389	0.00060	0.24319	0.00175	0.07043	0.07384
	40	0.37441	0.00059	0.26303	0.00187	0.07451	0.07493
	60	0.29410	0.00043	0.38848	0.00271	0.07342	0.07481
	80	0.26988	0.00031	0.38119	0.00266	0.07531	0.07668
	100	0.32187	0.00049	0.35285	0.00246	0.07724	0.07674
0.55	20	0.00000	0.00000	0.19650	0.00062	0.01168	0.13117
	40	0.04902	0.00001	0.06905	0.00115	0.00797	0.13120
	60	0.06314	0.00002	0.13772	0.00085	0.01668	0.13283
	80	0.02235	0.00000	0.08305	0.00109	0.01468	0.13324
	100	0.04081	0.00001	0.18640	0.00066	0.02271	0.13289
0.75	20	0.00000	0.00109	0.52098	0.00026	0.07945	0.16826
	40	0.00000	0.00050	0.58071	0.00014	0.07563	0.16787
	60	0.00000	0.00045	0.49063	0.00034	0.07302	0.16579
	80	0.00000	0.00036	0.55279	0.00019	0.07557	0.16761
	100	0.00000	0.00026	0.54077	0.00022	0.06898	0.16735

Table 2: MSEs of the EM-algorithm, method of moments and Bayesian estimators with $\theta_1 = 3$, $\theta_2 = 9$

τ	N	EM-Algorithm		Method of Moments		Bayse Method ($prior\ \alpha = 0.4$)	
		BIAS	MSE	BIAS	MSE	BIAS	MSE
0.15	20	0.68839	0.00246	0.78743	0.00437	0.40132	0.01548
	40	0.52653	0.00139	0.59526	0.00276	0.39953	0.01564
	60	0.59899	0.00180	0.69525	0.00355	0.39741	0.01609
	80	0.61900	0.00192	0.58893	0.00272	0.39669	0.01610
	100	0.58041	0.00168	0.67781	0.00341	0.39588	0.01595
0.35	20	0.40916	0.00083	0.37548	0.00262	0.22026	0.04380
	40	0.17158	0.00014	0.14139	0.00120	0.22388	0.04439
	60	0.32974	0.00048	0.15685	0.00128	0.22174	0.04413
	80	0.31141	0.00046	0.29002	0.00204	0.22460	0.04420
	100	0.27668	0.00038	0.40973	0.00287	0.22229	0.04519
0.55	20	0.09947	0.00004	0.08016	0.00198	0.16818	0.08659
	40	0.07216	0.00002	0.18321	0.00067	0.17068	0.08631
	60	0.00000	0.00000	0.14257	0.00083	0.16761	0.08673
	80	0.02657	0.00000	0.00875	0.00146	0.17178	0.08634
	100	0.00000	0.00000	0.00636	0.00154	0.17370	0.08732
0.75	20	0.00000	0.00028	0.38733	0.00065	0.11387	0.13505
	40	0.00000	0.00050	0.46493	0.00040	0.11700	0.13441
	60	0.00000	0.00036	0.53295	0.00023	0.11703	0.13558
	80	0.00000	0.00042	0.47731	0.00037	0.11613	0.13456
	100	0.00000	0.00054	0.55422	0.00019	0.11833	0.13407

which represents in table 5.
In table 5 we used 3 different priors $\alpha = 0.3, 0.4, 0.5$ for estimating mixing proportion τ.
Bayesian approach provides better estimates than EM-algorithm and method of moments with all priors, but the best estimate achieved with prior $\alpha = 0.3$ which is closest to the real mixing proportion $\tau = 0.33$.
According to the results of the previous section, this result was predictable. Because real mixing proportion τ is less than 0.55 and at this conditions, Bayesian method represents the best estimate.

Table 3: MSEs of the EM-algorithm, method of moments and Bayesian estimators with $\theta_1 = 4$, $\theta_2 = 10$

τ	N	EM-Algorithm		Method of Moments		Bayse Method ($prior\ \alpha = 0.6$)	
		BIAS	MSE	BIAS	MSE	BIAS	MSE
0.15	20	0.54750	0.00149	0.66930	0.00334	0.31016	0.00387
	40	0.52270	0.00139	0.68110	0.00344	0.30880	0.00390
	60	0.59040	0.00176	0.65890	0.00326	0.30811	0.00397
	80	0.57780	0.00168	0.50434	0.00213	0.30711	0.00400
	100	0.63400	0.00203	0.64624	0.00315	0.30549	0.00409
0.35	20	0.38490	0.00072	0.41602	0.00292	0.28279	0.02185
	40	0.25790	0.00031	0.36775	0.00256	0.28277	0.02168
	60	0.27000	0.00036	0.13972	0.00119	0.28306	0.02188
	80	0.24410	0.00028	0.41081	0.00288	0.28483	0.02144
	100	0.25500	0.00030	0.27487	0.00194	0.28792	0.02162
0.55	20	0.00000	0.00000	0.22564	0.00052	0.28257	0.05416
	40	0.00460	0.00000	0.11642	0.00094	0.28112	0.05444
	60	0.00000	0.00002	0.22873	0.00051	0.28166	0.05418
	80	0.04480	0.00001	0.05546	0.00122	0.28424	0.05367
	100	0.03070	0.00000	0.21956	0.00054	0.28392	0.05492
0.75	20	0.00000	0.00016	0.63928	0.00006	0.28824	0.09731
	40	0.00000	0.00036	0.52148	0.00026	0.29120	0.09734
	60	0.00000	0.00027	0.39237	0.00064	0.29449	0.09605
	80	0.00000	0.00055	0.44978	0.00045	0.29030	0.09649
	100	0.00000	0.00040	0.51619	0.00027	0.30145	0.09453

Table 4: EM-Algorithm and method of moments, estimates for real data

Method	Mixing Proportion Estimate	BIAS	MSE
EM-algorithm	0.56373	0.26373	0.00045
Method of moments	0.62399	0.32399	0.00194

Table 5: Bayesian approach for estimating mixing Parameter for real data

teta1	teta2	prior	Bayes Estimate	Bias	MSE-Posterior
4	5	0.3	0.43459	0.08459	0.00453
		0.4	0.50493	0.15493	0.00422
		0.5	0.55342	0.20342	0.00351

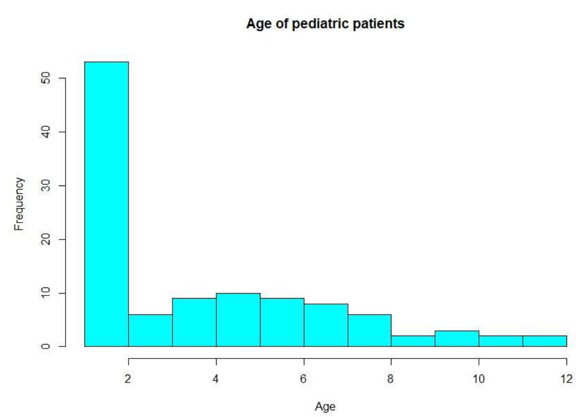


Fig. 1: Histogram of age of all patients

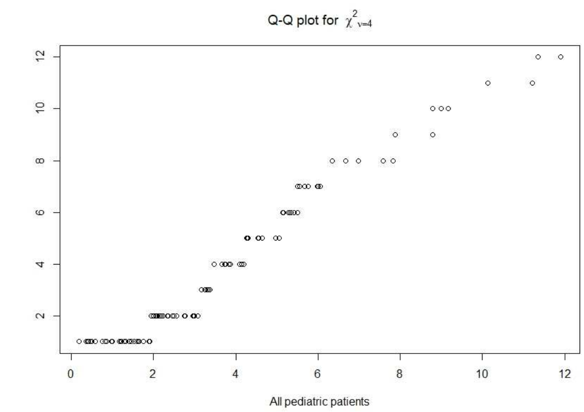


Fig. 2: *Q-Q*-plot for all patients

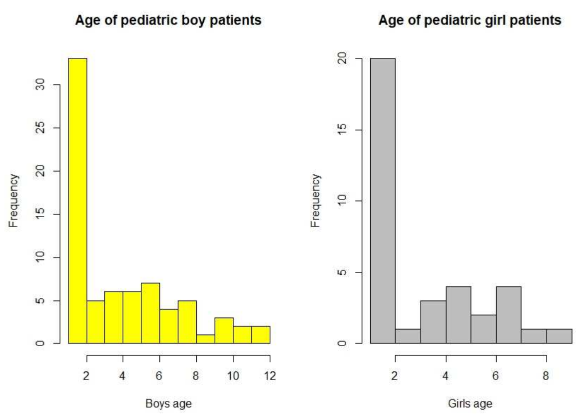


Fig. 3: Histogram of age by gender

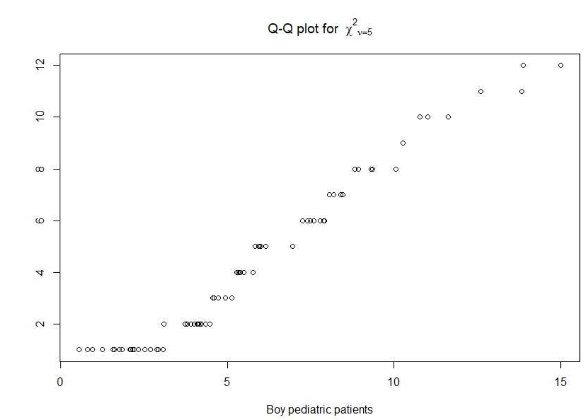


Fig. 4: *Q-Q*-plot for boys age distribution

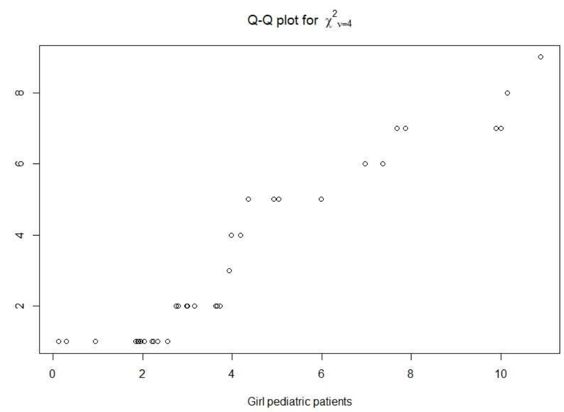


Fig. 5: *Q-Q*-plot for girls age distribution

Acknowledgement

The authors are thankful to the referees for their valuable comments and suggestions. And also the authors are thankful to Professor Shred Gore for his kind guides and editing the paper.

References

[1] R. Render and H. Walker, Mixture Densities, Maximum Likelihood and the EM Algorithm. SIAM review, Vol. 26, pp. 195-239 (1984).

[2] G.J McLachlan and D. Peel, Finite Mixture Models, Wiley, New York (2000).

[3] N. Martin, C. Hory and M. Huchard, Chi-Square Law Mixture for an Unsupervised Time-Frequency Pattern Extraction, Proceedings of the Tenth International Congress on Sound and Vibration, Stockholm, Sweden, 7-10 July, Code 62825, (2003).

[4] C. Hory and N. Martin, Time-Frequency Modelization as a Mixture of Chi-Square Distributions, Statistical signal processing, Conference paper. St. Louis, MO, USA, 28 Sept -1 Oct, pp. 246-249, (2003).

[5] D. Rindskopf, Mixture or Homogeneous? Comment on Bauer and Curran (2003), Psychological Methods, Vol. 8, No. 3, pp. 364368, (2003).

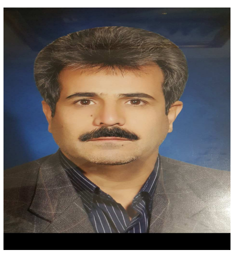
[6] M.R. Zaman, D. Paul, N. Akhter, M.H. Howlader and M.S. Kabir, Chi-square Mixture of Chi-square Distributions, Journal of Applied Sciences, Vol. 6, No. 2, pp. 243-246 (2006).

[7] S. Sahuguede, A. Vergonjanne and J. Cances, OCDMA Channel Statistic as a Mixture of Chi-Square Distributions, Fifth advanced International conference on telecommunications, pp. 248-253, (2009).

[8] X. Chen, M. Ponomareva and E. Tamer. Likelihood inference in some finite mixture models. Journal of Econometrics, Vol.182, pp. 87-99 (2014).

[9] Core Team. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, (2014). http://www.R-project.org

[10] C. Wang, Z. Tan and T.A. Louis, An Exponential Tilt Mixture Model for Time-to-Event Data to Evaluate Treatment Effect Heterogeneity in Randomized Clinical Trials. Biometrics & Biostatistics International Journal, Vol. 1, Issue. 2, pp. 1-6 (2014).
[11] C. Teel, T. Park and A. R. Sampson, EM Estimation for Finite Mixture Models With Known Mixture Component Size, Communications in Statistics: Simulation and Computation, Vol. 44, Issue. 6, pp. 1545-1556 (2015).
[12] J.P. Baudry and G. Celeux, EM for mixtures, Statistics and Computing, Vol. 25, Issue. 4, pp. 713-726 (2015).
[13] P, Nasiri and A.A, Azarian.Comparing different methods for estimating parameters of mixture of two Chi-square distributions, International Journal of Pure and Applied Mathematics, Vol. 114, pp. 29-40 (2017).
[14] S. Jaspers, A. Komrek and M. Aerts, Bayesian estimation of multivariate normal mixtures with covariate-dependent mixing weights, with an application in antimicrobial resistance monitoring, Biometrical Journal, Vol. 60, pp. 7-19 (2018).



Parviz Nasiri was born in Ardabil province of Iran in 1966. In 1990, he successfully completed his bachelor's degree at Ferdowsi University of Mashhad in Iran, and in 1993 joined the University of Pune for M.Sc. and M-Phil Courses of Statistics. He successfully completed his Ph.D in 2001 at the University of Mumbai in India. At the moment, Dr. Parviz Nasiri is Associate Professor of Statistics at the Payam Noor University in Iran.

Amir Abbas Azarian is Ph.D. Candidate in statistics at Payam Noor University of Iran. At moment, He is collaborating with statistics office of Mashhad University of Medical Sciences for statistical consulting and data analyzing. In meantime, his research interests are Mixture Models and Queueing Models.