International Journal of Pure and Applied Mathematics
Volume 114 No. 1 2017, 29-40
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: 10.12732/ijpam.v114i1.3
ijpam.eu

COMPARING DIFFERENT METHODS FOR ESTIMATING PARAMETERS OF MIXTURE OF TWO CHI-SQUARE DISTRIBUTIONS

P. Nasiri[1 §], A.A. Azarian[2]
[1,2]Department of Statistics
University of Payam Noor
19395-4697, Tehran, I.R. IRAN

Abstract: At this paper, at first we introduce mixture of two chi-square distributions and then find its parameters from method of moments and maximum likelihood estimators. EM algorithm and its application for finding parameters will be discussed.

AMS Subject Classification: 43E70
Key Words: mixture, chi-square distribution, method of moments, maximum likelihood estimators, EM algorithm

1. Introduction

Mixture models in general and parametric mixture models in particular are very useful methods for modelling a population and have lots of applied useful examples in medicine, industry and economics. Render and Walker (1984)[1] introduced using EM algorithm to find maximum likelihood estimates for mixture densities. Mixture models introduced by McLachlan *et,al* (2000)[2]. Wang, Tan and Louis (2014)[3] used mixture models for modelling time-to-event data to evaluate treatment effects in randomised clinical trials. Teel, Park and Sampson (2015)[4] considered the use of an EM algorithm for fitting finite mixture models when mixture component size is known.

Received: September 22, 2016
Revised: March 9, 2017
Published: April 19, 2017

§Correspondence author

Baudry and Celeux (2015)[5] showed that, however, Maximum likelihood through the EM algorithm is wide[6] introduced additive and multiplicative mixed normal distributions. Zaman *et,al.* (2006)[7] introduced mixture of chi-square distributions using Poisson elements. Chen, Ponomareva and Tamer (2014)[8] introduced likelihood inference in some finite mixture models and discussed different situations for mixture models. Chi-square distribution is one of the most applied statistical distributions, which is used in different branches of industry and technology. For example, you can fit this distribution on the amount of extraction petroleum products from crude oil in refineries. Density function of this distribution is as below:

$$f(x;\theta) = \frac{x^{\frac{\theta}{2}-1} e^{\frac{-x}{2}}}{2^{\frac{\theta}{2}}\Gamma(\frac{\theta}{2})}. \tag{1}$$

Here $\Gamma(.)$ is gamma function and θ is the distribution parameter and we can write: $X \sim \chi^2(\theta)$ Non central moments of this distribution are given as:

$$E(X^m) = \theta(\theta+2)(\theta+4)...(\theta+2m-2) = 2^m \frac{\Gamma(m+\frac{\theta}{2})}{\Gamma(\frac{\theta}{2})}. \tag{2}$$

As you know, there are two common methods for estimating parameters, method of moments and maximum likelihood estimators. We will discuss about these methods for chi-square distribution. At method of moments, parameters will be estimated by putting sample moments equal to distribution moments. Because chi-square distribution has only one parameter, we need the first order central moment.

Also, because E(X) $=\theta$, which is the first order moment of this distribution, should be set equal to the samples first order moment, $\bar{X}$, we can write:$\hat{\theta} = \bar{X}$.

We need maximum likelihood function to estimating parameters by using maximum likelihood method for chi-square distribution. The maximum likelihood function of chi-square distribution is written as bellow:

$$L(x;\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \frac{x_i^{\frac{\theta}{2}-1} e^{\frac{-x_i}{2}}}{2^{\frac{\theta}{2}}\Gamma(\frac{\theta}{2})} \tag{3}$$

and its logarithm is:

$$l(x;\theta) = \sum_{i=1}^{n} log(f(x_i;\theta)) = (\frac{\theta}{2}-1)\sum_{i=1}^{n} log x_i - \sum_{i=1}^{n}\frac{x_i}{2} - n log(2^{\frac{\theta}{2}}\Gamma(\frac{\theta}{2})). \tag{4}$$

The maximum likelihood estimator of θ is the value which maximizes l(x;θ). You can find maximum with derivation from l(x;θ), with respect to θ and set the result equal to zero. In the other words:

$$\frac{\partial l(x;\theta)}{\partial\theta} = \frac{1}{2}\sum_{i=1}^{n} logx_i - n(\frac{\frac{1}{2}2^{\frac{\theta}{2}}log(2)\Gamma(\frac{\theta}{2}) + \frac{1}{2}\Gamma'(\frac{\theta}{2})2^{\frac{\theta}{2}}}{2^{\frac{\theta}{2}}\Gamma(\frac{\theta}{2})}) = 0. \quad (5)$$

So

$$\sum_{i=1}^{n} logx_i = n(\frac{log(2)\Gamma(\frac{\theta}{2}) + \frac{1}{2}\Gamma'(\frac{\theta}{2})}{\Gamma(\frac{\theta}{2})}), \quad (6)$$

Moreover, we estimate θ by solving

$$\frac{\Gamma'(\frac{\theta}{2})}{\Gamma(\frac{\theta}{2})} = \overline{logx_i} - log2, \quad (7)$$

which is impossible with common methods and will be solved numerically.

2. Mixture Distributions, see [9]

Mixture models in general and parametric mixture models in particular are very useful methods for modelling a population and have a lot of applied useful examples in medicine, industry and economics. Mixture models are developed for studying populations, which are made from some sub-populations. In more general terms, if your population made from k sub-populations and each sub-population has density $f_j(\theta)$, $j = 1, 2...k$, you can say that your population has mixture distribution f(Θ) where:

$$f(\Theta) = \sum_{j=1}^{k} \tau_j f_j(x;\theta_j). \quad (8)$$

In this distribution, mixing parameters are τ_j, j=1, 2 k and $\sum_{j=1}^{k}\tau_j = 1$. If each sub-population is independent from others, using equations (1) and (5), we can write:

$$E(X) = \sum_{j=1}^{k} \tau_j E_j(X) \quad (9)$$

And

$$Var(X) = \sum_{j=1}^{k} {\tau_j}^2 Var_j(X) \quad (10)$$

Where

$$E_j(X) = \int X dF_j \qquad Var_j(X) = \int X^2 dF_j - (\int X dF_j)^2 \tag{11}$$

So we can write the moments as:

$$E_j(X^r) = \sum_{j=1}^{k} \tau_j E_j(X^r) \tag{12}$$

And maximum likelihood function as:

$$L(\mathbf{x};\Theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i-1}^{n} (\sum_{j=1}^{k} \tau_j f_j(x_i;\theta_j)) \tag{13}$$

Suppose that random variable X is a mixture of two other random variables, X_1 and X_2 with mixing parameter τ from the first sub-population. Also you can suppose that $X_i \sim \chi^2(\theta_i)$, *i=1,2*. Using (8)to have a mixture of two chi-square distribution, we can write:

$$f(x;\Theta) = \tau f_1(x;\theta_1) + (1-\tau) f_2(x;\theta_2) \tag{14}$$

Where Θ=(τ,θ_1,θ_2)and f(x;θ_i) is the density of i*th* population. And

$$f(x;\theta_i) = \frac{x^{\frac{\theta_i}{2}-1} e^{\frac{-x}{2}}}{2^{\frac{\theta_i}{2}} \Gamma(\frac{\theta_i}{2})} \qquad ; x \geq 0; \theta_i > 0; \quad i = 1,2 \tag{15}$$

In the other words, X has Mixture of Two Chi-square Distributions (MTChD) or

$$X \sim MTChD(\tau, \theta_1, \theta_2)$$

The most important problem is estimating the parameters of MTChD;(τ,θ_1,θ_2). We continue with two common methods: Method of Moments and Maximum likelihood.

2.1. Method of Moments for Estimating Parameters of MTChD

As you know, for this method, we need moments up to order of the unknown parameters. Because this distribution has 3 unknown parameters, we have:

$$\begin{aligned} E(X) &= E(\tau X_1 + (1-\tau X_2)) \\ &= \tau E X_1 + (1-\tau) E X_2 \\ &= \tau\theta_1 + (1-\tau)\theta 2 \\ &= \theta_2 + \tau(\theta_1 - \theta_2) \end{aligned} \tag{16}$$

$$\begin{aligned} E(X^2) &= E(\tau X_1 + (1-\tau X_2))^2 \\ &= \tau^2 (EX_1)^2 + (1-\tau)^2 (EX_2)^2 + 2\tau(1-\tau)E(X_1)E(X_2) \\ &= \tau^2(\theta_1^2 + 2\theta_1) + (1-\tau)^2(\theta_2^2 + \theta_2) + 2\tau(1-\tau)\theta_1\theta_2 \end{aligned} \tag{17}$$

$$\begin{aligned} E(X^3) &= E(\tau X_1 + (1-\tau X_2))^3 \\ &= \tau^3 + (\theta_1^4 + \theta_2^4 + 9(\theta_1^3 - \theta_2^3) + 26(\theta_1^2 - \theta_2^2) + 24(\theta_1 - \theta_2) \\ &- 3(\theta_1^2\theta_2 - \theta_1\theta_2^2) + 72\theta_2) \\ &+ \tau^2(81\theta_2^4 + 27\theta_2^3 + 3\theta_1^2\theta_2 - 6\theta_1\theta_2^2 - 12\theta_1\theta_2) \\ &+ \tau(-3\theta_2^4 - 27\theta_2^3 - 78\theta_2^2 + 6\theta_1^2\theta_2 + 3\theta_1\theta_2^2 + 6\theta_1\theta_2) \\ &+ \theta_2^4 + 9\theta_2^3 + 26\theta_2^2 + 24\theta_2 \end{aligned} \tag{18}$$

Now we can estimate the parameters with solving the equations:

$$E(X) = \frac{1}{n}\sum x_i; \qquad E(X^2) = \frac{1}{n}\sum x_i^2; \qquad E(X^3) = \frac{1}{n}\sum x_i^3 \tag{19}$$

2.2. Maximum Likelihood Estimation of Parameters of MTChD

To use maximum likelihood method, we should find maximum likelihood function. We use from indicator variable I(z_i=j)to define likelihood function, where

$$I(z_i = j) = \begin{cases} 1, & \text{if } i^{th} \text{ sample belongs to } j^{th} \text{ sub-population.} \\ 0, & \text{Other Wise.} \end{cases} ,$$

$$j = 1,2 \quad i = 1, ..., n \tag{20}$$

And likelihood function is:

$$\begin{aligned} L(\Theta; x, z) &= \prod_{i=1}^{n}(\sum_{j=1}^{2} I(z_i = j)\tau_j f(x_i, \theta_j)) \\ &= \prod_{i=1}^{n}(\sum_{j=1}^{2} I(z_i = j)\tau_j (\frac{x_i^{\frac{\theta_j}{2}-1} e^{\frac{-x_i}{2}}}{2^{\frac{\theta_j}{2}}\Gamma(\frac{\theta_j}{2})})) \end{aligned} \tag{21}$$

Where τ_2=1-τ_1. Logarithm of this function is

$$l(\Theta;x,z)=\sum_{i=1}^{n}(\sum_{j=1}^{2}I(z_i=j)log(\tau_j)+(\frac{\theta_j}{2}-1)log(x_i)-\frac{x_i}{2}-\frac{\theta_j}{2}log2-log(\Gamma(\frac{\theta_j}{2}))) \quad (22)$$

As mentioned above, you can find estimators, by derivation from $l(\Theta$;x,z) with respect to each unknown parameter and set the result equal to zero and solving the equation. But this method is very difficult for estimating parameters of MTChD. So we use from an alternative method, EM algorithm, which is an iterative method for finding the parameters in Maximum Likelihood method.

3. Estimating Parameters of MTChD Using EM-Algorithm

Suppose Y is a p-dimensional random vector with probability density function g(y; Θ) where $\Theta = (\theta_1, \theta_2, ..., \theta_d)^T$ 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 be included incomplete data which made from missing data or censored data; but in other situations it may be a complete data from a mixture of two or more distributions which the proportions of allocation in different distributions is unknown and we should find these proportions. In these cases using the method of Maximum Likelihood is very difficult, but EM-algorithm is a useful method. EM-algorithm is an iterative method for estimating parameters in case of incomplete data using maximum likelihood method.Since log L(Θ) included incomplete 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 the other words, suppose that Θ(k) is the value of after k^{th} iteration. At step k+1, steps E and M are as below:

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

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

2. Step M (Maximization): choosing $\Theta^(k+1)$ for each $\Theta \in \Omega$ which maximizing $Q(\Theta;\Theta^{(k)})$ with respect to Θ. In the other words

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

Step M is repeated until the sequence L $(\Theta^{(k)})$ converges.

In 1977 Rubin, Dumpster and Layrd showed that the likelihood function of incomplete data L(Θ) after one iteration is not decreasing. So the sequence $L(\Theta^{(k)})$ converges uniformly to some L*, which L* is a stationary value for Θ*, which is true in $\frac{\partial L(\Theta)}{\partial \Theta} = 0$ or equivalently $\frac{\partial logL(\Theta)}{\partial \Theta} = 0$. 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 locally maximum or minimum. Value of Θ* depends on the initial value of Θ(0). It should be noted that iteration in EM-algorithm will be increasing the likelihood, and will converge to a value in some generally good conditions.
Now suppose that we have a population which is 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 which is independent from the last distribution; and τ_1 is the proportion of population which comes from the first part; and τ_2=1- τ_1 is the proportion of population which comes from the second part which is unknown. At this situation we have a mixture of two independent chi-square distributions. As we mentioned above you can use EM-algorithm for finding the parameters of this population because τ is unknown. To formulate this population, we have:

$$\begin{aligned} L(\Theta; x, z) &= \prod_{i=1}^{n}(\sum_{j=1}^{2} I(z_i = j)\tau_j f(x_i, \theta_j)) \\ &= \prod_{i=1}^{n}(\sum_{j=1}^{2} I(z_i = j)\tau_j (\frac{x_i^{\frac{\theta_j}{2}-1} e^{\frac{-x_i}{2}}}{2^{\frac{\theta_j}{2}}\Gamma(\frac{\theta_j}{2})})) \end{aligned} \tag{23}$$

Where z_i is an indicator variable for identifying the distribution of x_i and n is the number of observations. L is the likelihood function for a mixture distribution with proportion of mixing τ_1. If we show the logarithm of likelihood function with l(Θ; x, z), then:

$$l(\Theta; x, z) = \sum_{i=1}^{n}(\sum_{j=1}^{2} I(z_i = j)log(\tau_j) + (\frac{\theta_j}{2} - 1)log(x_i) - \frac{x_i}{2} - \frac{\theta_j}{2}log2 - log(\Gamma(\frac{\theta_j}{2}))) \tag{24}$$

In this case we have a mixture of two independent chi-square distributions (MTChD) with parameters$(\tau, \theta_1, \theta_2)$. Steps E and M of EM-algorithm are as:

E step:
Suppose that $\theta^{(t)}$ is the current parameter value which is known. The distribution of z_i will be achieved using Bayes rule which is proportional to chi-square parameters and proportions which are introduced in $\theta^{(t)}$. In the other words:

$$\begin{aligned}\tau_{j,i} &= P(Z_i = j|X_i = x_i; \theta^{(t)}) \\ &= \frac{\tau_j^{(t)} f(x_i; \theta_j^{(t)})}{\tau_1^{(t)} f(x_i; \theta_1^{(t)}) + \tau_2^{(t)} f(x_i; \theta_2^{(t)})}\end{aligned} \tag{25}$$

So E-step is equal to:

$$\begin{aligned}Q(\theta|\theta^{(t)}) =& E(\log L(\theta; x, z)) \\ =& \sum_{i=1}^{n}(\sum_{j=1}^{2} T_{j,i}^{(t)}(I(z_i = j)log(\tau_j) + (\frac{\theta_j}{2} - 1)log(x_i) - \frac{x_i}{2} \\ & - \frac{\theta_j}{2}log2 - log(\Gamma(\frac{\theta_j}{2})))\end{aligned} \tag{26}$$

M Step: As mentioned above, we should maximize the equation achieved in E-step with respect to all of those parameters. We should find the maximum of $Q(\theta|\theta((t))$ with respect to τ_1, θ_1 and θ_2 . Parameters of this distribution can be maximized independently. Since $\tau_1 + \tau_2 = 1$, we have:

$$\begin{aligned}\tau^{(t+1)} &= \arg\max_{\tau} Q(\Theta|\Theta^t) \\ &= \arg\max_{\tau}\{log\tau_1 \sum_{i=1}^{n} \tau_{1,i}^{(t)} + log\tau_2 \sum_{i=1}^{n} \tau_{2,i}^{(t)}\}\end{aligned} \tag{27}$$

And hence:

$$\tau_j^{(t+1)} = \frac{\sum_{i=1}^{n} T_{j,i}^{(t)}}{\sum_{i=1}^{n}(T_{1,i}^{(t)} + T_{2,i}^{(t)})} = \frac{1}{n}\sum_{i=1}^{n} T_{j,i}^{(t)} \tag{28}$$

Which can be used for estimating τ_1. For estimating parameters θ_1 and θ_2 we have:

$$\begin{aligned}\theta_1^{(t+1)} &= \arg\max_{\theta_1} Q(\Theta|\Theta^t) \\ &= \arg\max_{\theta_1} \sum_{i=1}^{n} T_{1,i}^{(t)}\{log\tau_1^{(t)} + (\frac{\theta_1^{(t)}}{2} - 1)\log x_i - \frac{x_i}{2} - \frac{\theta_1^{(t)}}{2}\log 2 \\ & - \log(\Gamma(\frac{\theta_1^{(t)}}{2}))\}\end{aligned} \tag{29}$$

We can define

$$K = \sum_{i=1}^{n} T_{1,i}^{(t)} \{log\tau_1^{(t)} + (\frac{\theta_1^{(t)}}{2} - 1)\log x_i - \frac{x_i}{2} - \frac{\theta_1^{(t)}}{2}\log 2 - \log(\Gamma(\frac{\theta_1^{(t)}}{2}))\}$$

Derivative from k with respect to θ_1 and set the equation equal to zero, we have:

$$(\Gamma(\alpha))(\Psi(\alpha))^2 = \frac{\sum_{i=1}^{n} T_{1,i}^{(t)} \log \frac{x_i}{2}}{\sum_{i=1}^{n} T_{1,i}^{(t)}}; \qquad where \quad \alpha = \frac{\theta_1^{(t)}}{2} \tag{30}$$

The same as manner we have:

$$(\Gamma(\beta))(\Psi(\beta))^2 = \frac{\sum_{i=1}^{n} T_{2,i}^{(t)} \log \frac{x_i}{2}}{\sum_{i=1}^{n} T_{2,i}^{(t)}}; \qquad where \quad \beta = \frac{\theta_2^{(t)}}{2} \tag{31}$$

Where $\Psi(\alpha)$ is $(\Gamma(\alpha))^{\prime}$. This equation should be solved numerically and there is no straight solution.

4. Numerical Study

We provided the results of numerical study in two situations: At first we supposed that parameters θ_1 and θ_2 were known and the only unknown parameter of MTChD was τ. We compared parameter estimates from method of moments and EM-algorithm in figure 1. Results were achieved by using relations 19 and 28. After that, we supposed parameters θ_1 and θ_2 are unknown and the only known parameter of MTChD was τ. At this situation, we compared the parameter estimates from method of moments, maximum likelihood and EM-algorithm by relations 19, 22, 30. Results are presented in figures 2, 3 and 4. For solving non-linear equations, we used `nleqslv` package in `R` software[1]0. We used from numerical methods to find the best estimators in maximum likelihood method.

As you see, in figure 1, estimates from EM-algorithm have smaller bias and MSE than Method of Moments estimates. In figures 2, 3 and 4, we estimated parameters θ_1 and θ_2 when τ is known. Estimates are achieved from three methods: Maximum likelihood, Method of Moments and EM algorithm. Estimates from method of moments are out of parameter space, when τ is smaller than 0.5. This is one of the disadvantages of method of moments. But if τ is

		θ_1=2, θ_2=8						θ_1=3, θ_2=9					
		EM-ALGORITHM			MME			EM-ALGORITHM			MME		
τ	n	EST.	MSE	BIAS	EST.	MSE	BIAS	EST.	MSE	BIAS	EST.	MSE	BIAS
0.35	20	0.72	0.0006	0.37	0.59	0.0017	0.24	0.76	0.0008	0.41	0.73	0.0026	0.37
	40	0.72	0.0006	0.37	0.61	0.0018	0.26	0.52	0.0001	0.17	0.49	0.0012	0.14
	60	0.64	0.0004	0.30	0.74	0.0027	0.39	0.68	0.0005	0.33	0.50	0.0012	0.16
	80	0.62	0.0003	0.27	0.73	0.0026	0.38	0.66	0.0005	0.31	0.64	0.0020	0.29
	100	0.67	0.0004	0.327	0.70	0.0024	0.35	0.63	0.0004	0.28	0.76	0.0029	0.41
0.55	20	0.55	2.6E-6	0	0.35	0.0006	0.20	0.65	4.4E-5	0.09	0.63	0.0020	0.08
	40	0.60	1.1E-5	0.05	0.48	0.0011	0.07	0.62	2.4E-5	0.07	0.37	0.0007	0.18
	60	0.61	1.7E-5	0.06	0.41	0.0008	0.14	0.55	3.2E-6	0	0.41	0.0008	0.14
	80	0.57	2.1E-6	0.02	0.47	0.0010	0.08	0.58	3.5E-6	0.02	0.54	0.0014	0.01
	100	0.59	7.1E-6	0.04	0.36	0.0006	0.19	0.55	9.8E-12	0	0.56	0.0015	0.01
0.75	20	0.75	0.0010	0	0.23	0.0002	0.52	0.75	0.0002	0	0.36	0.0007	0.39
	40	0.75	0.0004	0	0.17	0.0001	0.58	0.75	0.0005	0	0.29	0.0004	0.47
	60	0.75	0.0004	0	0.26	0.0003	0.49	0.75	0.0003	0	0.22	0.0002	0.53
	80	0.75	0.0003	0	0.20	0.0001	0.55	0.75	0.0004	0	0.27	0.0003	0.48
	100	0.75	0.0002	0	0.21	0.0002	0.54	0.75	0.0005	0	0.20	0.0001	0.55

Figure 1: Comparing EM-Algorithm and Method of Moments estimators of MTChD, τ is unknown,θ_1 and θ_2 are known

		MAXIMUM LIKELIHOOD						METHOD OF MOMENTS						EM-ALGORITHM					
		θ_1			θ_2			θ_1			θ_2			θ_1			θ_2		
τ	n	Est.	Bias	MSE	Est.	Bias	MSE	Est.	Bias	MSE	Est.	Bias	MSE	Est.	MSE	MSE	Est.	Bias	MSE
0.35	20	4.210	2.210	1.506	7.065	0.065	0.663	-1.870	3.870	3.332	18.49	11.49	18.128	5.629	3.629	0.265	4.992	2.008	0.332
	40	3.960	1.960	0.908	7.170	0.170	0.456	-2.025	4.025	1.903	18.80	11.80	11.018	5.646	3.646	0.146	5.002	1.998	0.185
	60	3.965	1.965	1.186	7.195	0.195	0.399	-2.022	4.022	1.170	18.75	11.75	6.395	5.636	3.636	0.095	4.994	2.006	0.119
	80	4.000	2.000	0.955	7.075	0.075	0.397	-2.027	4.027	0.869	18.75	11.75	4.805	5.614	3.614	0.069	4.970	2.030	0.087
	100	3.995	1.995	0.922	7.145	0.145	0.401	-2.037	4.037	0.741	18.76	11.76	4.045	5.635	3.635	0.056	4.991	2.009	0.073
0.45	20	3.955	1.955	0.990	7.045	0.045	0.465	-0.379	2.379	2.469	10.96	3.964	5.987	5.214	3.214	0.220	4.654	2.346	0.275
	40	3.725	1.725	0.592	7.180	0.180	0.368	-0.583	2.583	1.519	11.27	4.271	3.556	5.202	3.202	0.128	4.646	2.354	0.157
	60	3.785	1.785	0.676	7.080	0.080	0.469	-0.615	2.615	0.884	11.32	4.325	2.109	5.193	3.193	0.077	4.644	2.356	0.097
	80	3.585	1.585	0.440	7.195	0.195	0.319	-0.654	2.654	0.700	11.35	4.359	1.708	5.189	3.189	0.058	4.639	2.361	0.073
	100	3.605	1.605	0.371	7.255	0.255	0.287	-0.669	2.669	0.563	11.37	4.373	1.340	5.207	3.207	0.047	4.638	2.362	0.061
0.55	20	3.805	1.805	0.809	7.130	0.130	0.603	0.907	1.093	2.415	6.976	0.024	2.794	4.767	2.767	0.182	4.251	2.749	0.227
	40	3.610	1.610	0.428	7.185	0.185	0.378	0.786	1.214	1.382	7.104	0.104	1.541	4.759	2.759	0.093	4.250	2.750	0.118
	60	3.455	1.455	0.285	7.375	0.375	0.267	0.796	1.204	0.918	7.052	0.052	1.026	4.752	2.752	0.063	4.246	2.754	0.079
	80	3.580	1.580	0.399	7.255	0.255	0.337	0.755	1.245	0.648	7.107	0.107	0.767	4.763	2.763	0.045	4.257	2.743	0.057
	100	3.325	1.325	0.082	7.415	0.415	0.060	0.755	1.245	0.469	7.100	0.100	0.537	4.746	2.746	0.038	4.240	2.760	0.047
0.65	20	3.665	1.665	0.805	7.330	0.330	1.066	2.305	0.305	3.174	4.532	2.468	1.635	4.281	2.281	0.147	3.806	3.194	0.183
	40	3.395	1.395	0.276	7.535	0.535	0.516	2.291	0.291	1.534	4.536	2.464	0.755	4.277	2.277	0.075	3.802	3.198	0.094
	60	3.190	1.190	0.064	7.725	0.725	0.082	2.170	0.170	1.199	4.597	2.403	0.552	4.269	2.269	0.048	3.792	3.208	0.058
	80	3.195	1.195	0.059	7.655	0.655	0.063	2.118	0.118	0.873	4.637	2.363	0.433	4.271	2.271	0.037	3.793	3.207	0.047
	100	3.170	1.170	0.061	7.705	0.705	0.065	2.156	0.156	0.669	4.601	2.399	0.339	4.263	2.263	0.030	3.784	3.216	0.038
0.75	20	3.190	1.190	0.309	8.215	1.215	0.656	3.853	1.853	6.664	3.078	3.922	1.245	3.769	1.769	0.119	3.295	3.705	0.132
	40	3.090	1.090	0.037	8.245	1.245	0.117	3.680	1.680	3.508	3.097	3.903	0.667	3.775	1.775	0.069	3.297	3.703	0.074
	60	3.045	1.045	0.020	8.290	1.290	0.106	3.587	1.587	2.249	3.134	3.866	0.417	3.766	1.766	0.043	3.279	3.721	0.045
	80	3.045	1.045	0.025	8.310	1.310	0.079	3.555	1.555	1.781	3.158	3.842	0.339	3.775	1.775	0.032	3.286	3.714	0.036
	100	3.010	1.010	0.005	8.245	1.245	0.072	3.587	1.587	1.271	3.141	3.859	0.247	3.759	1.759	0.023	3.271	3.729	0.028

Figure 2: Comparing EM-Algorithm and Method of Moments estimators of MTChD, τ is known, θ_1=2 and θ_2=7 are unknown

		MAXIMUM LIKELIHOOD						METHOD OF MOMENTS						EM-ALGORITHM					
		θ_1			θ_2			θ_1			θ_2			θ_1			θ_2		
τ	n	Est.	Bias	MSE	Est.	Bias	MSE	Est.	Bias	MSE	Est.	Bias	MSE	Est.	Bias	MSE	Est.	Bias	MSE
0.35	20	4.700	1.700	0.915	7.630	0.370	0.313	-2.101	5.101	4.531	21.695	13.695	23.483	6.660	3.660	0.345	6.059	1.941	0.446
	40	4.475	1.475	0.792	7.640	0.360	0.175	-2.280	5.280	2.351	22.096	14.096	12.356	6.642	3.642	0.169	6.026	1.974	0.222
	60	4.400	1.400	0.500	7.645	0.355	0.131	-2.303	5.303	1.632	22.124	14.124	8.844	6.649	3.649	0.108	6.025	1.975	0.137
	80	4.440	1.440	0.721	7.645	0.355	0.151	-2.392	5.392	1.149	22.332	14.332	5.977	6.629	3.629	0.082	6.002	1.998	0.106
	100	4.385	1.385	0.694	7.645	0.355	0.121	-2.366	5.366	0.959	22.294	14.294	4.871	6.653	3.653	0.072	6.023	1.977	0.088
0.45	20	4.490	1.490	0.985	7.400	0.600	0.275	-0.459	3.459	3.357	13.263	5.263	7.836	6.226	3.226	0.276	5.688	2.312	0.334
	40	4.400	1.400	0.680	7.420	0.580	0.214	-0.739	3.739	2.045	13.671	5.671	4.763	6.242	3.242	0.150	5.703	2.297	0.176
	60	4.365	1.365	0.439	7.385	0.615	0.114	-0.692	3.692	1.148	13.622	5.622	2.693	6.217	3.217	0.091	5.673	2.327	0.112
	80	4.200	1.200	0.330	7.410	0.590	0.092	-0.779	3.779	1.046	13.723	5.723	2.502	6.210	3.210	0.071	5.658	2.342	0.088
	100	4.260	1.260	0.332	7.420	0.580	0.129	-0.739	3.739	0.772	13.671	5.671	1.819	6.206	3.206	0.058	5.657	2.343	0.068
0.55	20	4.210	1.210	0.411	7.230	0.770	0.187	1.035	1.965	3.743	8.739	0.739	3.839	5.779	2.779	0.228	5.290	2.710	0.286
	40	4.155	1.155	0.328	7.290	0.710	0.131	0.958	2.042	2.024	8.775	0.775	2.113	5.734	2.734	0.119	5.227	2.773	0.140
	60	4.150	1.150	0.282	7.255	0.745	0.147	0.910	2.090	1.332	8.776	0.776	1.388	5.757	2.757	0.077	5.245	2.755	0.091
	80	4.080	1.080	0.269	7.285	0.715	0.116	0.860	2.140	1.011	8.813	0.813	1.053	5.752	2.752	0.062	5.244	2.756	0.075
	100	4.070	1.070	0.220	7.300	0.700	0.090	0.886	2.114	0.819	8.814	0.814	0.848	5.747	2.747	0.045	5.244	2.756	0.053
0.65	20	4.105	1.105	0.431	7.160	0.840	0.309	2.724	0.276	5.038	5.876	2.124	2.368	5.285	2.285	0.210	4.799	3.201	0.260
	40	3.905	0.905	0.188	7.325	0.675	0.122	2.672	0.328	2.567	5.849	2.151	1.131	5.269	2.269	0.100	4.776	3.224	0.122
	60	3.835	0.835	0.150	7.400	0.600	0.100	2.464	0.536	1.868	5.953	2.047	0.841	5.261	2.261	0.066	4.765	3.235	0.078
	80	3.875	0.875	0.157	7.325	0.675	0.122	2.495	0.505	1.302	5.972	2.028	0.607	5.258	2.258	0.051	4.756	3.244	0.059
	100	3.845	0.845	0.113	7.370	0.630	0.093	2.433	0.567	1.074	5.996	2.004	0.510	5.248	2.248	0.039	4.750	3.250	0.047
0.75	20	3.765	0.765	0.297	7.615	0.385	0.304	4.498	1.498	10.273	4.149	3.851	1.961	4.760	1.760	0.195	4.240	3.760	0.228
	40	3.645	0.645	0.126	7.650	0.350	0.202	4.210	1.210	6.425	4.278	3.722	1.107	4.730	1.730	0.090	4.206	3.794	0.103
	60	3.590	0.590	0.072	7.630	0.370	0.078	4.142	1.142	4.456	4.295	3.705	0.791	4.743	1.743	0.057	4.208	3.792	0.064
	80	3.585	0.585	0.040	7.675	0.325	0.062	4.110	1.110	2.838	4.296	3.704	0.519	4.743	1.743	0.043	4.207	3.793	0.051
	100	3.585	0.585	0.040	7.630	0.370	0.058	4.096	1.096	2.321	4.310	3.690	0.418	4.742	1.742	0.037	4.207	3.793	0.042

Figure 3: Comparing EM-Algorithm and Method of Moments estimators of MTChD, τ is known, θ_1=3 and θ_2=8 are unknown

		MAXIMUM LIKELIHOOD						METHOD OF MOMENTS						EM-ALGORITHM					
		θ_1			θ_2			θ_1			θ_2			θ_1			θ_2		
τ	n	Est.	Bias	MSE	Est.	Bias	MSE	Est.	Bias	MSE	Est.	Bias	MSE	Est.	Bias	MSE	Est.	Bias	MSE
0.35	20	5.680	1.680	1.323	8.385	0.615	0.379	-2.189	6.189	4.585	24.747	15.747	23.630	7.679	3.679	0.385	7.085	1.915	0.476
	40	5.405	1.405	0.663	8.265	0.735	0.237	-2.471	6.471	2.683	25.314	16.314	14.100	7.637	3.637	0.211	7.037	1.963	0.255
	60	5.350	1.350	0.573	8.270	0.730	0.167	-2.580	6.580	2.155	25.519	16.519	10.765	7.654	3.654	0.130	7.042	1.958	0.169
	80	5.285	1.285	0.626	8.250	0.750	0.093	-2.619	6.619	1.422	25.592	16.592	7.179	7.667	3.667	0.094	7.044	1.956	0.115
	100	5.390	1.390	0.363	8.285	0.715	0.086	-2.644	6.644	1.205	25.631	16.631	6.072	7.654	3.654	0.078	7.037	1.963	0.099
0.45	20	5.645	1.645	0.806	7.925	1.075	0.242	-0.654	4.654	4.909	15.774	6.774	10.688	7.230	3.230	0.311	6.697	2.303	0.385
	40	5.390	1.390	0.508	7.965	1.035	0.126	-0.739	4.739	2.544	15.903	6.903	5.551	7.230	3.230	0.154	6.693	2.307	0.187
	60	5.255	1.255	0.407	8.020	0.980	0.120	-0.689	4.689	1.617	15.823	6.823	3.533	7.211	3.211	0.109	6.664	2.336	0.129
	80	5.110	1.110	0.303	7.955	1.045	0.125	-0.804	4.804	1.231	15.995	6.995	2.686	7.220	3.220	0.079	6.679	2.321	0.102
	100	5.095	1.095	0.238	7.980	1.020	0.055	-0.807	4.807	1.001	15.979	6.979	2.370	7.209	3.209	0.071	6.669	2.331	0.089
0.55	20	5.225	1.225	0.537	7.630	1.370	0.203	1.203	2.797	4.859	10.342	1.342	5.015	6.749	2.749	0.282	6.245	2.755	0.347
	40	5.060	1.060	0.371	7.710	1.290	0.111	1.033	2.967	2.762	10.537	1.537	2.815	6.742	2.742	0.142	6.227	2.773	0.178
	60	5.045	1.045	0.335	7.670	1.330	0.116	1.051	2.949	1.770	10.507	1.507	1.836	6.761	2.761	0.096	6.266	2.734	0.117
	80	4.880	0.880	0.261	7.735	1.265	0.097	1.081	2.919	1.235	10.481	1.481	1.233	6.746	2.746	0.073	6.246	2.754	0.087
	100	4.950	0.950	0.237	7.735	1.265	0.092	0.955	3.045	1.139	10.590	1.590	1.137	6.745	2.745	0.057	6.242	2.758	0.067
0.65	20	4.850	0.850	0.473	7.515	1.485	0.212	3.073	0.927	7.575	7.183	1.817	3.336	6.272	2.272	0.248	5.774	3.226	0.305
	40	4.720	0.720	0.287	7.540	1.460	0.118	2.907	1.093	4.154	7.287	1.713	1.858	6.262	2.262	0.132	5.761	3.239	0.150
	60	4.605	0.605	0.176	7.585	1.415	0.065	2.973	1.027	2.578	7.245	1.755	1.128	6.261	2.261	0.081	5.752	3.248	0.100
	80	4.715	0.715	0.161	7.520	1.480	0.100	2.813	1.187	2.151	7.331	1.669	0.939	6.256	2.256	0.059	5.750	3.251	0.072
	100	4.550	0.550	0.128	7.580	1.420	0.064	2.766	1.234	1.695	7.333	1.667	0.756	6.251	2.251	0.049	5.746	3.254	0.057
0.75	20	5.680	1.680	1.323	8.385	0.615	0.379	5.145	1.145	18.281	5.267	3.733	2.873	5.745	1.745	0.227	5.192	3.809	0.259
	40	5.405	1.405	0.663	8.265	0.735	0.237	4.812	0.812	8.778	5.370	3.630	1.416	5.731	1.731	0.124	5.173	3.827	0.145
	60	5.350	1.350	0.573	8.270	0.730	0.167	4.614	0.614	6.188	5.468	3.532	1.060	5.718	1.718	0.082	5.153	3.847	0.089
	80	5.285	1.285	0.626	8.250	0.750	0.093	4.539	0.539	4.593	5.498	3.502	0.739	5.726	1.726	0.061	5.162	3.838	0.068
	100	5.390	1.390	0.363	8.285	0.715	0.086	4.572	0.572	3.683	5.460	3.540	0.630	5.718	1.718	0.049	5.155	3.845	0.056

Figure 4: Comparing EM-Algorithm and Method of Moments estimators of MTChD, τ is known, θ_1=4 and θ_2=9 are unknown

greater than 0.5, method of moments has reasonable estimates. We use from numerical methods to estimating parameters in method of maximum likelihood. Estimates from maximum likelihood are more reasonable from EM algorithm and method of moments. Although, EM algorithm estimators have smaller MSEs than maximum likelihood estimators, but their biases are greater. In the other words, estimates of parameters in EM algorithm, converges to the mean of two population parameters. So the maximum likelihood method develops the best estimates of MTChD, when τ is known.

References

[1] R. Render, H. Walker, Mixture densities, maximum likelihood and the EM algorithm, *SIAM review* **26** 2 (1984), 195-239.

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

[3] C. Wang, Z. Tan, 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* **1** 2 (2014)

[4] C. Teel, T. Park , A.R. Sampson, EM estimation for finite mixture models with known mixture component size. *Communications in Statistics: Simulation and Computation* **44** 6 (2015), 1545-1556.

[5] J.P. Baudry, G. Celeux, EM for mixtures: Initialization requires special care, *Statistics and Computing* **25** 4 (2015), 713-726.

[6] S. Lipovetski, Additive and multiplicative mixed normal distributions and finding cluster centers, *Int. J. Mach. Learn. & Cyber* **4** (2013), 1-11

[7] M.R. Zaman, D.N.R. Paul, N. Akhter, M.H. Howlader, M.H. Howlader, Chi-square Mixture of Chi-square Distributions, *Journal of Applied Sciences*, **6** 2 (2006), 243-246.

[8] X. Chen, M. Ponomareva, E. Tamer, Likelihood inference in some finite mixture models, *Journal of Econometrics* **182** (2014). 8799

[9] S.M. MirHoseini, E. Pasha, EM-algorithm convergence and accelerations, *6th Iranian statistical conference,* (2002)

[10] B. Hasselman nleqslv: Solve Systems of Nonlinear Equations, *R package version 2.8.* (2015)