"Mircea cel Batran" Naval Academy Scientific Bulletin, Volume XIII – 2010
Published by "Mircea cel Batran" Naval Academy Press, Constanta, Romania

MIXTURE OF TWO EXPONENTIAL-POISSON DISTRIBUTION

Elena Carmen LUPU[1],
Sergiu LUPU [2]
[1]Ovidius University of Constanta, , Faculty of Pharmacy
[2] Naval Academy „Mircea cel Bătrân", Navy Faculty

Abstract: *The mixture model of two Exponential-Poisson distribution is investigated. First, some proprieties of the model are discussed. In order to estimate the vector of the unknown parameters the EM algorithm is proposed. Further we carry out some simulated illustration using Monte Carlo method*
Key words: *Exponential-Poisson distribution*

1. INTRODUCTION

Mixture models play a vital role in many practical applications. For example, direct applications of finite mixture models are in fisheries research, economics, medicine, psychology, palaeoanthropology, botany, agriculture, zoology, life testing and reliability, among others. Indirect applications include outliers, Gaussian sums, cluster analysis, latent structure models, modeling prior densities, empirical Bayes method and nonparametric (kernel) density estimation.

In many applications, the available data can be considered as data coming from a mixture population of two or more distributions. This idea enables us to mix statistical distributions to get a new distribution carrying the properties of its components.

Kuş (2007) [1] introduced and studied a new two - parameter distribution with decreasing failure rate, known as the Exponential–Poisson distribution. (EP). This distribution is obtained by mixing exponential and zero truncated Poisson distribution.

The mixture of two Exponential-Poisson distribution (MTEP) has its pdf as:

$$f(x,\theta) = p_1 f_1(x,\theta) + p_2 f_2(x,\theta), \quad p_2 = 1 - p_1 \qquad (1)$$

where $\theta = (p_1, \lambda_1, \lambda_2, \beta_1, \beta_2)$; $\theta_i = (\lambda_i, \beta_i)$, $i = 1,2$ and $f_i(x,\theta)$ the density function of *i* th component is given by [1]

$$f_i(x_i,\theta_i) = \frac{\lambda_i \beta_i}{1 - e^{-\lambda_i}} e^{-\lambda_i - \beta_i x + \lambda_i \exp(-\beta_i x)}, \quad i = 1,2 \quad x, \lambda_i, \beta_i \in \mathbb{R}_+ \qquad (2)$$

The cdf of the MTEP is :

$$F(x,\theta) = p_1 F_1(x,\theta_1) + p_2 F_2(x,\theta_2)$$

where $F_i(x,\theta_i)$ is the cdf of the *i* th component and has the form [1]

$$F_i(x,\theta_i) = \frac{e^{\lambda_i \exp(-\beta_i x)} - e^{\lambda_i}}{\left(1 - e^{\lambda}\right)} \qquad i = 1,2 \qquad (3)$$

It can be seen that, for all value of parameters, the MTEP density function is strictly deacrising x and tending to 0 as $x \to \infty$

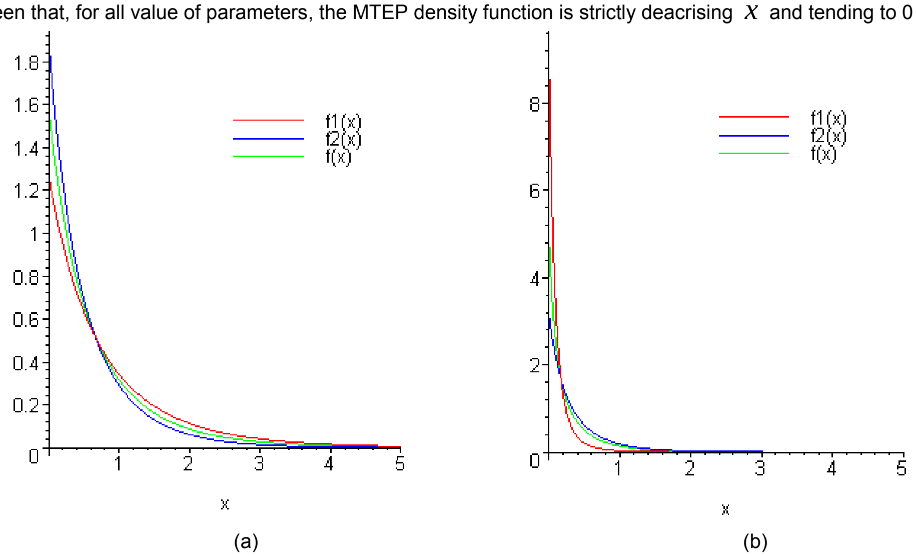


Fig. 1: Density function: components and their mixture with parameters (a) (0.5, 0.5, 0.5, 1, 1.5), (b) (0.3, 3, 1, 3, 2)

2. PROPRIETIES

1. Mean and variance

The mean of the MTEP distribution is:

$$E(X) = p_1 \frac{\lambda_1}{(e^{\lambda_1}-1)\beta_1} F_{22}([1,1],[2,2],\lambda_1) + p_2 \frac{\lambda_2}{(e^{\lambda_2}-1)\beta_2} F_{22}([1,1],[2,2],\lambda_2)$$

where $F_{p,q}(n,d,\lambda)$ is the. The definition of $F_{p,q}(n,d,\lambda)$ is:

$$F_{p,q}(n,d,\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^k \prod_{i=1}^{p} \Gamma(n_i+k)\Gamma^{-1}(n_i)}{\Gamma(k+1)\prod_{i=1}^{q}\Gamma(d_i+k)\Gamma^{-1}(d_i)}$$

where $n = \lfloor n_1,\ldots,n_p \rfloor$, p is the number of operands of n, $d = \lfloor d_1,\ldots,d_q \rfloor$ and q the number of operands of d. Generalized hypergeometric function can be quickly evaluated and readily avaible in standard softwares, such as Maple, Mathematica.

Hence, the variance is given by:

$$\begin{aligned} Var(x) = {} & \frac{p_1\lambda_1}{(e^{\lambda_1}-1)\beta_1^{2}}\left[2F_{33}([1,1,1],[2,2,2],\lambda_1) - \frac{p_1\lambda_1}{(e^{\lambda_1}-1)}F_{22}^{2}([1,1],[2,2],\lambda_1)\right] + \\ & + \frac{p_2\lambda_2}{(e^{\lambda_2}-1)\beta_2^{2}}\left[2F_{33}([1,1,1],[2,2,2],\lambda_1) - \frac{p_2\lambda_2}{(e^{\lambda_1}-1)}F_{22}^{2}([1,1],[2,2],\lambda_2)\right] - \\ & - \frac{2p_1p_2\lambda_1\lambda_2}{\beta_1\beta_2(e^{\lambda_1}-1)(e^{\lambda_2}-1)} F_{22}([1,1],[2,2],\lambda_1)\cdot F_{22}([1,1],[2,2],\lambda_2) \end{aligned}$$

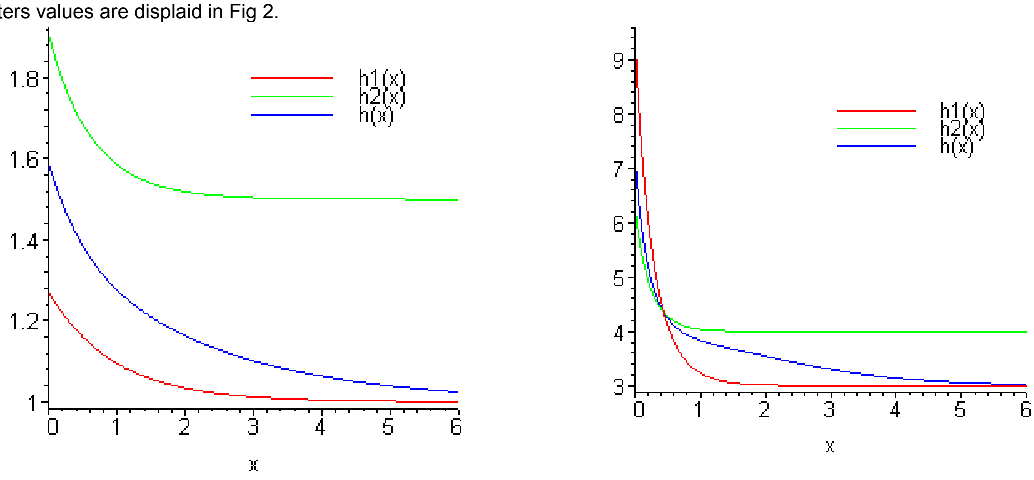
2. The survival and hazard functions

Using (1) and (2), survival function (reability function) and hazard function (failure rate function) of MTEP are given, respectively, by

$$S(x,\theta) = p_1\left(1 - \frac{e^{\lambda_1\exp(-\beta_1 x)} - e^{\lambda_1}}{1-e^{\lambda_1}}\right) + p_2\left(1 - \frac{e^{\lambda_2\exp(-\beta_2 x)} - e^{\lambda_2}}{1-e^{\lambda_2}}\right)$$

$$h(x,\theta) = \frac{f(x,\theta)}{S(x,\theta)} = \frac{p_1 \frac{\lambda_1\beta_1}{(1-e^{-\lambda_1})} e^{-\lambda_1-\beta_1x+\lambda_1\exp(-\beta_1x)} + p_2 \frac{\lambda_2\beta_2}{(1-e^{-\lambda_2})} e^{-\lambda_2-\beta_2x+\lambda_2\exp(-\beta_2x)}}{p_1\left(1 - \frac{e^{\lambda_1\exp(-\beta_1 x)} - e^{\lambda_1}}{1-e^{\lambda_1}}\right) + p_2\left(1 - \frac{e^{\lambda_2\exp(-\beta_2 x)} - e^{\lambda_2}}{1-e^{\lambda_2}}\right)}$$

The MTEP hazard function is deacrising and both the initial and the long-term hazard are finite[1]. The MTEP hazard function for selected parameters values are displaid in Fig 2.



(a) (b)

Fig. 2: Hazard function: components and their mixture with parameters (a) (0.5, 0.5, 0.5, 1, 1.5), (b) (0.3, 3, 1, 3, 4)

3. PARAMETER ESTIMATION. EM ALGORITHM

In this section, the EM algorithm is used to estimate the parameters of the pdf of the MTEP given in (1) and (2). The EM algorithm was intoduced by Dempster et al. (1977) [5] and provides a simple computational method for fitting mixture models.

Let $x=(x_1,\dots,x_n)$ be a sample of independent observations from a mixture of two EP distributions of dimension *d*, and let $z=(z_1,\dots,z_n)$ be the latent variables that determine the component from which the observation originates.

$X_i/(Z_i=1)\sim EP(\lambda_1,\beta_1)$ and $X_i/(Z_i=2)\sim EP(\lambda_2,\beta_2)$

where $P(Z_1=1)=p_1$ and $P(Z_1=2)=p_2=1-p_1$

Let $\theta=\left(p_1,p_2,\lambda_1,\lambda_2,\beta_1,\beta_2\right)$ the current estimate of mixture parameters. Then the likelihood function is

$$L(\theta,x,z)=\prod_{j=1}^{n}\sum_{i=1}^{2}I\left(z_j=i\right)p_i f\left(x_j;\lambda_i,\beta_i\right)$$

Given our current estimate of the parameters $\theta^{(k)}$, the conditional distribution of the Z_i is determined by Bayes theorem:

$$t_{i,j}^{(k)}=P\left(Z_j=i\mid X_j=x_j;\theta^{(k)}\right)=\frac{p_i^{(k)}f(x_j,\theta_i)}{\sum_{i=1}^{2}p_i^{(k)}f(x_j,\theta_i)},\ i=1,2,\ j=1..n \qquad (4)$$

And the expected log-likelihood is given by:

$$Q(\theta,\theta^{(k)})=\sum_{j=1}^{n}\sum_{i=1}^{2}t_{i,j}^{(k)}\left[\log(pi)+\log(f(x_j;\lambda_i,\beta_i))\right]$$

So we need to find: $\theta^{(k+1)}=\arg\min_{\theta}Q(\theta,\theta^{(k)})$

In order to avoid solving a nonlinear system of equations, $\theta^{(k+1)}$ is determined by an iterativ system of equations. This is:

1. Set the initial values of the parameters $p_1^{(0)}$, $\lambda_1^{(0)}$, $\lambda_2^{(0)}$, $\beta_1^{(0)}$, $\beta_2^{(0)}$
2. E Step.

   Calculate: $t_{i,j}^{(k)}$ defined in (4)
3. M Step

   Calculate $p_1^{(k+1)}$, $\lambda_1^{(k+1)}$, $\lambda_2^{(k+1)}$, $\beta_1^{(k+1)}$, $\beta_2^{(k+1)}$ by maximum likelihood method. Thus

$$p_i^{(k+1)}=\frac{1}{n}\sum_{j=1}^{n}t_{i,j}^{(k)},\ i=1,2$$

Let:

$$L_1=L\left(x_1,\dots x_n,p_1^{(k)},\lambda_1^{(k)},\beta_1^{(k)},t_{1,1}^{(k)},\dots t_{1,n}^{(k)}\right)$$

$$=\sum_{j=1}^{n}\left(\ln p_1^{(k)}+\ln(\lambda_1^{(k)}\beta_1^{(k)})-\beta_1^{(k)}x_j-\lambda_1^{(k)}+\lambda_1^{(k)}e^{-\beta_1^{(k)}x_j}-\ln(1-e^{-\lambda_1^{(k)}})\cdot t_{1j}^{(k)}\right)$$

$$L_2=L\left(x_1,\dots x_n,p_2^{(k)},\lambda_2^{(k)},\beta_2^{(k)},t_{2,1}^{(k)},\dots t_{2,n}^{(k)}\right)$$

$$=\sum_{j=1}^{n}\left(\ln p_2^{(k)}+\ln(\lambda_2^{(k)}\beta_2^{(k)})-\beta_2^{(k)}x_j-\lambda_2^{(k)}+\lambda_2^{(k)}e^{-\beta_2^{(k)}x_j}-\ln(1-e^{-\lambda_2^{(k)}})\cdot t_{2,j}^{(k)}\right)$$

Setting the first derivatives of L_1 and L_2 with respect to each parameter equal to zero results the following equations for each parameter

$$\lambda_1^{(k+1)}=n\cdot\left[\sum_{j=1}^{n}\left(e^{-\beta_1^{(k)}x_j}+\frac{1}{1-e^{-\lambda_1^{(k)}}}\right)t_{1,j}^{(k)}\right]^{1}$$

$$\beta_1^{(k+)}=n\cdot\left[\sum_{j=1}^{n}\left(x_j(1+\lambda_1^{(k)}e^{-\beta_1^{(k)}x_j})\right)t_{1,j}^{(k)}\right]^{1}$$

$$\lambda_2^{(k+1)} = n \cdot \left[\sum_{j=1}^{n} \left(e^{-\beta_2^{(k)} x_j} + \frac{1}{1-e^{-\lambda_2^{(k)}}} \right) t_{2,j}^{(k)} \right]^{-1}$$

$$\beta_2^{(k+)} = n \cdot \left[\sum_{j=1}^{n} \left(x_j (1+\lambda_1^{(k)} e^{-\beta_1^{(k)} x_j}) \right) t_{2,j}^{(k)} \right]^{-1}$$

4. Verify the convergence of the algorithm

$$\max\left(\left|p_1^{(k+1)} - p_1^{(k)}\right|, \left|p_2^{(k+1)} - p_2^{(k)}\right|, \left|\lambda_1^{(k+1)} - \lambda_1^{(k)}\right|, \left|\lambda_2^{(k+1)} - \lambda_2^{(k)}\right|, \left|\beta_1^{(k+1)} - \beta_1^{(k)}\right|, \left|\beta_2^{(k+1)} - \beta_2^{(k)}\right| \right) < \varepsilon$$

For a given ε and $k+1$ number of iteration or $k+1<$ total number of iteration.

If one condition is met, then STOP, otherwise $k = k+1$, $p^{(k)} = p^{(k+1)}$, $\lambda_1^{(k)} = \lambda_1^{(k+1)}$, $\lambda_2^{(k)} = \lambda_2^{(k+1)}$, $\beta_1^{(k)} = \beta_1^{(k+1)}$, $\beta_2^{(k)} = \beta_2^{(k+1)}$ and go to step 2.

4. SIMULATION OF MTEP

In this section, we calculate the estimates of the parameters that appear in the pdf of MTEP given in (1) and (2) by using the EM Algorithm in a Monte Carlo simulatin.

Several simulation algorithms for EP distribution where introduced in [4]. One of this is the **EPExp algorithm (simulation of variable $EP(\lambda,\beta)$ by enveloping it with an Exponential density):**

repeat
Generate *U* uniform *0-1*
Generate U_1 uniform *0-1*
until $U_1 \le e^{-\lambda(1-U)}$
Take $X = -\frac{1}{\beta}\log(U)$

So, the random sample of the mixture are generated as fallow [2]:

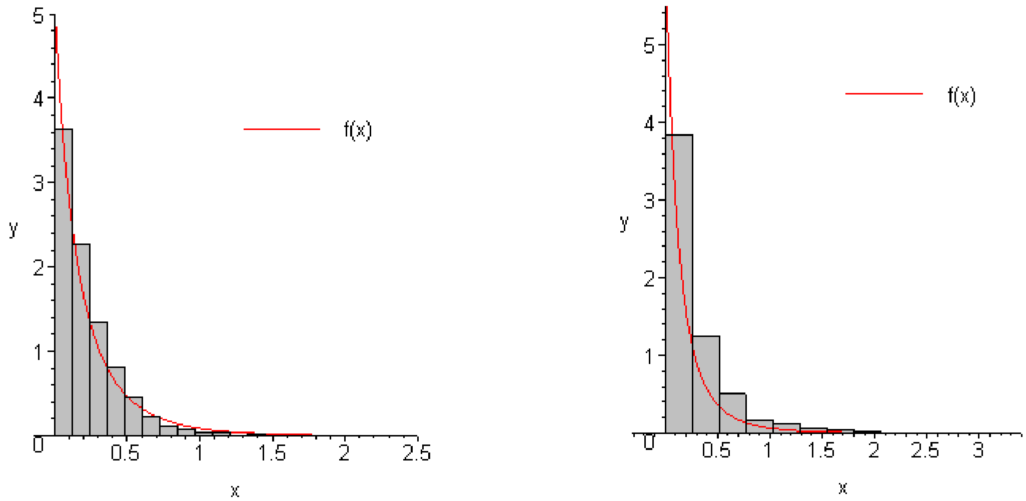
1. Generate one uniform variate U_1
2. If $U_1 < p_1$ then generate r.v *x* with pdf. $f_1(x,q_1)$
   If $U_1 > p_1$ then generate r.v *x* with pdf. $f_2(x,q_2)$

Where $f_1(x,q_1)$ and $f_2(x,q_2)$ can be simulated with EPExp algorithm.

The algorithm was implemented in the Maple; each time a number N=10000 of sampling values was generated for each choise of the vector parameters $\theta = (p_1, p_2, \lambda_1, \lambda_2, \beta_1, \beta_2)$. The result show that the mean and the variance of the simulated values are close to the theoretical value of the mean, and of the variance as well.

Table 1:The theoretical and the estimated mean and variance of MTEP

Caz 1: $p_1 = 0.3$, $p_2 = 0.7$ $\lambda_1 = 1$, $\beta_1 = 2$, $\lambda_2 = 1.5$, $\beta_2 = 3$	Caz2: $p_1 = p_2 = 0.5$, $\lambda_1 = 3$, $\beta_1 = 1$, $\lambda_2 = 4$, $\beta_2 = 3$
EX =0,270434 $\overline{x}$ =0.262992 DX =0.113874 s^2 =0.1202416	EX =0.298748 $\overline{x}$ =0.0,282759 DX =0.157882 s^2 =0.159947



Caz 1 Caz 2

Fig. 3: Fit of the densities to the histograms of the estimated values

Also, simulation have been performed to investigate the convergence of the proposed EM algoritm. 10000 samples of size 100 and 500 each are randomly sampled from the MTEP for each of three value of θ are generated. The result from the simulated data are reported in Table 2, which give the average $av\left(\hat{\theta}\right)$ calculated based on 10000 Monte Carlo repetition. The convergence of the EM algorithm is assumed when the absolute differences between succesive estimates are less than 10^{-5}.

Table 2 The means of the MTEM estimators for $\theta = \left(p_1, p_2, \lambda_1, \lambda_2, \beta_1, \beta_2\right)$ from 10000 samples of size 100 an 500 generated from MTEP

n	$\theta = \left(p_1, \lambda_1, \lambda_2, \beta_1, \beta_2\right)$	$av\left(\hat{\theta}\right)$
100	(0.3, 2, 4, 1, 3) (0.5, 2, 4, 1, 3) (0.8, 2, 4, 1, 3)	(0.252, 1.842, 4.233, 0.930, 3.162) (0.483, 2.094, 4.109, 1.067, 3.098) (0.821, 2.103, 4.114, 1.131, 3.112)
500	(0.3, 2, 4, 1, 3) (0.5, 2, 4, 1, 3) (0.8, 2, 4, 1, 3)	(0.214, 2.042, 4.133, 1.090, 3.122) (0.509, 2.019, 3.903, 0.916, 3.045) (0.798, 2.032, 4.097, 0.930, 3.083)

From table 2 we see that the mean of the MTEP estimators are closer to the theoretical value as *n* increase.

5. CONCLUSION

In this paper, we discuss the properties of the MTEP. Estimation of the unknown parameters of the mixture of two Exponential-Poisson distributions denoted by model (1) is obtained using EM-Algorithm. Some Monte Carlo simulations are carried out to investigate the performance of the estimation technique. Mixture models are a natural way to build a clustering model out of an existing probabilistic model.

6. REFERENCES

1. COŞCUN KUŞ A new lifetime distribution, Computational Statistics and Data & Data Analysis 51 (2007), issue 9, p. 4497-4509;
2. VĂDUVA, I. Modele de simulare**,** Editura Universităţii din Bucureşti, 2004
3. VADUVA, I. Computer generation of gamma random variables by rejection and composition procedures, Math. Oper. Forsch. u. Statist., Ser. Statistics, Vol.8, Nr. 4, p.545-576 (Berlin);
4. LUPU, ELENA (2009). On simulation of some lifetime distributions, Analele Universitatii Bucureşti Matematică- Informatică, Anul LVIII (in print).
5. DEMPSTER, A.P., LAIRD, N.M., RUBIN, D.B., 1977. Maximum-likelihood from incomplete data via the EM algorithm. J. Roy. Statist. Soc. B 39, 1–38.
6. SULTAN K.S., ISMAIL M.A, AL-MOISHEER A.S, Mixture of two inverse Weibull distribution: Proprieties and estimation, Computational Statistics and Data & Data Analysis 51 (2007), issue 9, p.5377-5387.