On Using Conventional and TL moments for the Estimation of a Mixture of Exponential Distributions, A Theoretical Review

Habib Ahmed Elsayir

Department of Mathematics, Al Qarqadha University College, Umm Al Qura University (K.S.A), P.O.Box 1109-21912, Saudi Arabia

Introduction

An essential step in frequency analysis is the selection of suitable distribution, for this purpose many distributions such as exponential, Gamma, Gumbel, Kappa, Log-normal, Weibull, Pareto, Pearson type-III, Logistic, Normal, Generalized Extreme Value Type III, mixed Gamma, mixed exponential and many others are considered. Mixture models have significant role in the field of biology, geology, medicine, chemistry, actuarial science and remaining all the fields in which pattern recognition is important. Mixture distributions also have much attention in Statistics. The main goal of the mixture distributions analysis is to analyze the mixture population by decomposing it into its individual subpopulations with their associated proportion.

At the present, the mixture of exponential distribution is considered, which is the widely used and applicable. McCullagh and Peter (1994), Ghosh and Ebrahimi (2001) and Hebert and Scariano (2004) studied and defined the characteristics of the exponential mixture distributions using classical and Bayesian methods. Yoo et al. (2005) used the mixed gamma distribution to study the rainfall frequency analysis. Wilks (1998) also proved through many goodness of fit criteria that the mixed exponential distribution is preferable then Gamma distribution. Shoji and Kitaura (2006) also considered the exponential distribution including normal lognormal and Weibull to model the wet-day and full-record daily precipitation.

Linear moments (L-moments) and Trimmed Linear moments (TL-moments) used before for the mixture models by Tartaglia et al. (2006) and Hussein and Liu (2009). Generally conventional moments and EM-Algorithm are used for the mixture density parameter estimation. L-moments method usually employed in hydrology, climatology and meteorology in the research of extreme precipitation, Kysely and Picek (2007), having mostly used smaller data sets. Mendoza-Rosas and De la Cruz-Reyna (2009) concluded in their study that the exponential mixture density is flexible, more precise and much easier to apply than the Weibull distribution and they used the mixture density for the assessment of the volcanic hazard. Shabri et al. (2011) discussed the method of trimmed L-moments with one smallest value and introduced an alternative ways to reduce undesirable influence of small sample using TL-moments and L-moments method for generalized logistic distribution. The Gamma distribution frequently preferred for rainfall data. But, Oseni and Ayoola (2012) fitted many popular distributions to model the rainfall data. Finally they found that the Exponential is the best density for describing data of Ibadan metropolis.

Discussed L-Moment models graphically and mathematically, Khan (2012) in his paper presented the L-Moment and Inverse Moment estimation of the inverse Generalized Exponential distribution including the properties of the L-Moment interested mainly in the relationship between $\beta$ and various L-Moment; measure of variability for L-Moment as the numerical quantities that describe the spread of the values in a set of data Fig. 1.

Abstract: This paper focuses on derivation of a mixture of exponential distribution and Linear Moments (LM) and Trimmed Linear moments (TL) is derived for estimation of its parameters. A comparison was made between L-moments, TL-moments of two component mixture of exponential distributions and conventional moments. In intensive simulation study, empirical study results are in the favor of conventional moments, the reason is that the L and TL-moments provided sometimes nearly unbiased results; sometimes much biased. Due to this vast variation, admittedly this study is in the favor of the conventional moments even its results are not much accurate and efficient.

Keywords Distribution Function, Exponential Distribution L-moments and TL-moments, Order Statistics, Parameter Estimation, Probability Density Function
A modified alternative robust version of L-moments is introduced by Bílková (2011) and Bílková (2014), called “trimmed L-moments” and it is termed TL-moments. In a study, Mirza (2015) focused on L-moments and TL-moments of Power function distribution and obtained the coefficient of variation, skewness and kurtosis by method of moments, L-moments and TL-moments and Parameters of the density are estimated using linear moments and compared with method of moments and MLE on the basis of bias, root mean square error and coefficients where L-moments proved to be superior for the parameter estimation. According to Dutang (2017), (L-moments) and (TL-moments) are appealing alternatives to the conventional moments.

A collection of original research topics and survey articles on mathematical related issues and their numerous applications in diverse areas were found in Agarwal (2018). Bayes Estimators of the unknown shape parameter of the exponentiated Moment Exponential Distribution (EMED) have been derived by Fatima and Ahmad (2018), by using two informative (gamma and chi-square) priors and two non-informative (Jeffrey’s and uniform) priors under different loss functions, namely, Squared Error. Ahmad used Partial Linear Moments (PLM) for censored samples to reach to the best fit distribution selected by implying the Z-statistic (goodness of fit test) based on L-moments. Agarwal et al. (2019a) investigated the Wick-type stochastic (3+1)-dimensional modified equations and produced a new set of solutions for the (3+1)-dimensional modified equations includes solutions of exponential, hyperbolic and trigonometric types. Other related issues concerning this topic including mathematical derivations and application were found in Wazwaz, (2017), (Hyder, 2018; Agarwal, 2019b; Rekhviashvili, 2019; Sahoo and Saha Ray (2019).

This paper considered the Linear moments (L-moments) and trimmed linear moments (TL-moments) of two component mixture of Exponential distributions and compare it with the conventional moments. The purpose of the paper is to find the most appropriate method to get the true estimates of the parameters from the three considered moments. An empirical study is performed assuming simulated data of mixtures density. Paper is organized as in section 2, two-component mixtures of the exponential distribution discussed. In section 3, the estimation method conventional moments, L-moments and TL-moments are introduced, described how to find the moments and estimate the parameters by these methods. In section 4, we derive the first four moments of all considered moments and last section about the Monte Carlo simulation study.

The Exponential Distribution and Moments

The Exponential Distribution

Let us consider a random variable (X). This variable is said to have an exponential distribution ($\lambda$) if it has the following Probability Density Function (PDF):

$$f_x (\{x | \lambda\} = \{e^{-\lambda x}, \text{for} \ x > 0, \text{and} \ f_x (\{x | \lambda\} = 0 \text{for} \ x \leq 0 \}$$

where, $\lambda > 0$ is called the distribution.
As increasing numbers of complexity in this modern era, the issue of heterogeneity arises. In this way a population has some or many subpopulations in it. To model such population’s data, much interest manifesting in mixture distributions because standard duration model not provide approximate results. As the exponential distribution is extensively applicable in many practical situations, in which data have more than one mode, the mixture models are required that provide more precise results, which is defined as:

\[
f(y | \Omega) = \sum_{i=1}^{m} p_i f_i(y | \theta_i)
\]

(2)

where, \( y \) is the values, \( m \) is the total number of the components (subgroups) and the parameter space \( \Omega \) = \((p_1, p_2, \ldots, p_m, \theta_1, \theta_2, \ldots, \theta_k)\). The weighting factor \( p_i \) always greater than zero and sum of all the weighting factor is one. And \( f_i \) is the density function of a component with the parameter \( \theta_i \).

In some situations the parent population from where the observations are taken, may consist two subgroups which mixed together with an unknown but fixed proportion. In this study we considered two subpopulation and each subpopulation follows the exponential pattern but with distinct parameter value. In such circumstances only the two component mixture is recommended to model the data. So the Probability Density Function (PDF) of Two Component Mixture of Exponential (TCME) distribution is as:

\[
f(y) = p_1(\theta_1 e^{-\theta_1 y}) + p_2(\theta_2 e^{-\theta_2 y})
\]

(3)

\( 0 \leq y < \infty, 0 \leq p_1, p_2 < \infty, 0 \leq p_1, p_2, \alpha, \beta < \infty. \)

We can also write this two component PDF as:

\[
f(y) = p(\theta_1 e^{-\theta_1 y}) + (1-p)(\theta_2 e^{-\theta_2 y})
\]

(4)

The Cumulative Distribution Function (CDF) has the form:

\[
F(y) = 1 - p e^{-\theta_1 y} - (1-p) e^{-\theta_2 y}
\]

(5)

This mixed density converts to the single parameter exponential density when the scale parameters for both components have the same values. In this study we consider the generalized case in which these parameters may or may not be same. To estimate the values of unknowns, all the method of moments equate the theoretical moments to their corresponding sample moments. Each considered moments discussed in the following subsection.

Conventional Moments

One of the basic logical moments, conventional moments used for the estimation of the parameter. Moreover, significant number of literature available those consider conventional moments for two-component mixtures. Conventional moments utilize the concept for parameter estimation is to equate the theoretical moments to their corresponding sample moments and simplify this relationship to obtain estimates of unknown parameters. Normally as many moments are required as many the unknown parameters.

The general formula for the moments is as:

\[
E\left( Y^r \right) = p\alpha^r + (1-p)\beta^r; r = 1, 2, 3, \ldots, n.
\]

(6)

\[ L \text{-Moments} \]

Moments have been traditionally used to characterize a probability distribution. Recently, linear moments (L-moments) and trimmed L-moments (TL-moments) are appealing alternatives to the conventional moments, Dutang (2017). Literature suggested that the use of L-Moments is advantageous for parameter estimation over the traditional method of estimation. These moments are less sensitive even in the presence of outliers (Vogel and Fennessey, 1993). According to Bilková (2014), the letter “L” in “L-moments” indicates that the \( r \)-th L-moment \( \lambda_r \) is a linear function of the expected value of a certain linear combination of order statistics. The estimate of the \( r \)-th L-moment \( \lambda_r \), based on the ample, is thus the linear combination of order data values, i.e., \( L \)-statistics. Hosking (1990) define the \( r \)th L-moments as the linear combinations of probability weighted moments of an ordered sample data \((Y_{1:k} \leq Y_{2:k} \leq \ldots \leq Y_{n:k})\):

\[
L_r = \frac{1}{r} \sum_{k=0}^{r} (-1)^k \binom{r-1}{k} E(Y_{r+k}), r = 1, 2, 3, \ldots
\]

(7)

For the real-valued random variable \( y \) with \( F(y) \) as the cumulative distribution function of ordered statistics of the sample of size \( r \), the \( E(Y_{r}) \) is define as:

\[
E(Y_{r}) = \frac{r!}{(1-r)(r-j)!} y(F)^{j-1}(1-F)^{r-j} dF
\]

(8)

The first two L-moments provides the measure of location and dispersion respectively. And the next two moments with the ratio of second L-moment provide the measure of skewness \( \tau_3 = L_3/L_2 \) and kurtosis \( \tau_4 = L_4/L_2 \) of the distribution respectively.

The four first sample L-moments are given by:

90
Moments of the TCME Distribution

Let \( Y_{1:n} \leq Y_{2:n} \leq Y_{m:n} \) be the order statistics of a random variable \( Y \). The expectation of the order statistics is as:

\[
E(Y_{r:n}) = \frac{n!}{(r-1)!(n-r)} \int x [F(y)]^{r-1} [1-F(y)]^{n-r} dF(y) \quad (12)
\]

Substituting the distribution function of the TCME, we obtain the following results respectively:

\[
E(Y_{r:n}) = \sum_{i=0}^{r-1} \sum_{j=0}^{n-r} \sum_{k=0}^{1} \binom{r-1}{i} \binom{n-r}{j} \binom{1}{k} \frac{p^i (1-p)^j}{(r-1)! (n-r)!} \times \frac{p \alpha}{(\alpha v + \beta(a+u))} \frac{1-p}{\beta (\alpha(v+1) + \beta u)}^2
\]

where, \( u = n-r-j+k \) and \( v = i+j \).

First four L-Moments of the considered density are derived by using (1) in close form and the final results are as:

\[
\lambda_1 = E(Y) = p\alpha + q\beta \quad (14)
\]

\[
\lambda_2 = 0.5E(Y_{2:n} - Y_{1:n}) = 0.5 (2p\alpha - p^2\alpha - 4pq\theta_1 - q^2\beta + 2q\beta) \quad (15)
\]

\[
\lambda_3 = \frac{\left[p\alpha(6 - 9p + 4p^2) - 6\right]}{6} + \frac{\left(pq\theta_1 + pq^2\theta_2 + pq^4\theta_3\right) + q\beta(6 - 9q + 4q^2)}{6} \quad (16)
\]

\[
\lambda_4 = \frac{1}{4} E(Y_{4:n} - 3Y_{3:n} + 3Y_{2:n} - Y_{1:n}) = \frac{\left[p\alpha(1 - 3p + \frac{10p^2}{3} - \frac{5p^3}{4}) - 12pq\theta_1 + 30p^2\theta_2\right]}{4} - \frac{\left[-20p^2\theta_1 + 30pq\theta_2 - 30p^2q^2\theta_3 - 20pq\theta_3\right]}{4} + \frac{q\beta(1 - 3q + \frac{10q^2}{3} - \frac{5q^3}{4})}{4} \quad (17)
\]

where, \( \theta_j = \left(\frac{-\alpha}{\alpha v + \beta u}\right)^j \) and \( q = 1-p \).
Estimates of the parameters can be obtained by solving the system of first three L-moments, after replacing the $\lambda_i$ by their sample moments.

**TL-Moments of TCME Distribution**

The TL-moments are of two types: Either symmetric $\lambda \ (t \ m)$ or asymmetric $\lambda \ (s, \ t \ m)$, see Dutang (2017). To derive the TL-Moments TCME distribution, the expression given in the (10) is used:

$$
\lambda_{l1}^{(l)} = E (Y_{l,1}) = \frac{p^2 \alpha (9 - 4p) + 6}{6} - \frac{pq \theta_{1} - p^2 q \theta_{2} - pq \theta_{3}}{6} + \frac{q^2 \beta (9 - 4q)}{6} 
$$

$$
\lambda_{l2}^{(l)} = \frac{1}{2} E (Y_{l,4} - Y_{l,2}) 
$$

$$
\lambda_{l3}^{(l)} = \frac{1}{3} E (Y_{l,6} - 2Y_{l,5} + Y_{l,5}) 
$$

$$
\lambda_{l4}^{(l)} = \frac{20}{3} E (Y_{l,8} - 3Y_{l,6} + 3Y_{l,6} - Y_{l,6}) 
$$

Simulation Study

To understand the behavior of statistical methods, a simulation study might be adopted, because some ‘truth’ (usually some parameter/s of interest) is known from the process of generating the data. It is known that simulation studies use computer intensive procedures to assess the performance of a variety of statistical methods in relation to a known truth. Therefore, the empirical analysis is presented by carry out the Monte Carlo simulation study to evaluate the performance of all the moments for estimation performance of the parameter. Experiments are repeated in simulation study to avoid the non-sampling bias using Minitab, and sample moments estimates were calculated by using Excel, Mathematica and Matlab.

Discussion

Sample L-moments are used in a similar way as sample conventional L-moments, summarizing the basic properties of the sample distribution, which are the location. L-moments is preferred to conventional moments, since sample L-moments are less sensitive to sample variability or measurement errors in extreme observations than conventional moments. L-moments therefore lead to more accurate and robust estimates of characteristics or parameters of the basic probability distribution. L-moments has their contribution to the topic of theory of description and summarization of theoretical probability distributions aimed at obtained sample data sets, parametric estimation and hypothesis testing for theoretical probability distributions. The method of TL-moments is not intended to stand alternative to the existing robust methods but rather supplement them, particularly in situations where outliers in the data are found. The expected values of order statistics of a random sample in the definition of L-moments of probability distributions are replaced with those of a larger random sample. Thus, TL-moments have some advantages over those of conventional, L-moments and central moments. TL-moment of the probability distribution may exist despite the non-existence of the corresponding L-moment or central moment of this probability distribution. In addition to that, TL-moments obtained from the sample are expected to be more resistant to outliers in the data.

Conclusion

Including they are more resistant and less prone to estimation bias, consisting the ability to characterize a wider range of the distribution, L-moments have several advantages over conventional moments. Close form of L-moments and TL-moments are delivered and then compared these moments with conventional moments. It
was very interesting situation observed in empirical study that L-moments and TL-moments yield much variation in the parameter estimates. Sometime estimated results are very accurate and sometime too much worst. It was due to the minor change in value of sample L-moments and sample TL-moments. On the other hand the results of conventional moments were not accurate but consistent. Eventuall, this paper focused on derivation of a mixture of exponential distribution and Linear Moments (LM) and Trimmed Linear moments (TL) for estimation of its parameters. In intensive simulation study, empirical study results were in the favor of conventional moments, the reason is that the L and TL-moments provided sometimes nearly unbiased results; sometimes much biased. Due to this vast variation, admittedly this study was in the favor of the conventional moments even its results were not much accurate and efficient.

Acknowledgement

I am deeply indebted to the editorial board and the reviewers of the American Journal of Applied Sciences for their valuable comments and efforts in order to publish this paper.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Adamowski, K., 2000. Regional analysis of annual maximum and partial duration flood data by nonparametric and l-moment methods. J. Hydrol., 229: 219-231
Agarwal, P., S.S. Dragomir, M. Jleli and B. Samet, 2018. Advances in mathematical inequalities and applications.
Agarwal, P., A. Hyder and M. Zakarya, 2019a. Well-posedness of stochastic modified kawahara equation. Adv. Differ. Equ.,
Agarwal, P., A.A. Hyder, M. Zakarya, G. AlNemer and C. Cesarano, 2019b. Exact solutions for a class of wick-type stochastic (3+1)-dimensional modified benjamin–bona–mahoney equations. Axioms, 8: 134–134. DOI: 10.3390/axioms8040134
Ahmad, I., M. Waqas, I.M. Almanjahie, A. Saghir and E.U. Haq, 2019. Regional flood frequency analysis using linear moments and partial linear moments: A case study. Applied Ecol. Environ. Res., 17: 3819-3836. DOI: 10.15666/aeer/1702_38193836
Bilkova, D., 2011. L-moments and their use in modeling the distribution of income and wage. Proceedings of the 58th World Statistical Congress, Aug. 21-26, Dublin, Flash Disk, pp: 1-6.
Bilková, D., 2014. Alternative tools of statistical analysis: L-moments and TL-moments of probability distributions. Pure Applied Math. J., 3: 14-25. DOI: 10.11648/j.pamj.20140302.11
Bilkova, D., 2014. L-moments and TL-moments as an alternative tool of statistical data analysis. J. Applied Math. Phys., 2: 919-929. DOI: 10.4236/jamp.2014.210104
Dutang, C., 2017. Theoretical L-moments and TL-moments using combinatorial identities and finite operators. Commun. Stat. Theory Meth. Taylor Francis, 46: 3801-3828
Elamir, E.A. and A.H. Seheult, 2003. Trimmed L-moments. Comput. Stat. Data Anal., 43: 299-314. DOI: 10.1016/S0167-9473(02)00250-5
Fatima, K. and S.P. Ahmad, 2018. Bayesian approach in estimation of shape parameter of the exponentiated moment exponential distribution. J. Stat. Theory Appl., 17: 359-374. DOI: 10.2991/jsta.2018.17.2.13
Ghosh, S.K. and N. Ebrahimi, 2001. Bayesian analysis of the mixing function in a mixture of two exponential distributions. Institut. Stat. Mimeo Series.
Hebert, J.L. and S.M. Scariano, 2004. Comparing location estimators for exponential mixtures under pitman’s measure of closeness. Common. Stat. Theory M, 33: 29-46. DOI: 10.1081/STA-120026575
Hosking, J.R.M., 1990. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. J. Royal Stat. Society (Series B), 52: 105-124. DOI: 10.1111/j.2517-6161.1990.tb01775.x
Hussein, A. and J. Liu 2009. Parametric estimation of mixtures of two uniform distributions. J. Stat. Comput. Simulat., 4: 395-410. DOI: 10.1080/00049650701810406
Hyder, A., 2018. White noise analysis combined with hypercomplex systems for solving stochastic modified KdV equations with non-Gaussian parameters. Pioneer J. Adv. Applied Math., 24: 39-61.
Khan, M.S., 2012. L-moment and inverse moment estimation of the inverse generalized exponential distribution. Int. J. Inform. Electron. Eng.
Kysely, J. and J. Ficek, 2007. Regional growth curves and improved design value estimates of extreme precipitation events in the Czech republic. Clim. Res., 33: 243-255. DOI: 10.3354/cr033243
McCullagh, P., 1994. Exponential mixtures and quadratic exponential families. Biometrika, 81: 721-729. DOI: 10.1093/biomet/81.4.721
Mendoza-Rosas, A.T. and S. De la Cruz-Reyna, 2009. A mixture of exponentials distribution for a simple and precise assessment of the volcanic hazard. Nat. Hazards Earth Syst. Sci., 9: 425-431. DOI: 10.5194/nhess-9-425-2009, 2009.

Mirza, N. S., 2015. Parameter estimation of power function distribution with TL-moments. Rev. Columbiana de Estadística, 38: 321-334. DOI: 10.15446/RCE.V38N2.51663

Oseni, B.A. and F.J. Ayoola, 2012. Fitting the statistical distribution for daily rainfall in Ibadan, based on chi-square and Kolmogorov-Smirnov goodness-of-fit tests. Eur. J. Bus. Manage., 7: 93-100.

Rekhviashvili, S.S.H., 2019. Application of the fractional oscillator model to describe damped vibrations. Turkish J. Phys., 43: 236-242. DOI: 10.3906/fiz-1811-16

Sahoo, S. and S.A. Saha Ray, 2019. Novel approach for stochastic solutions of wick-type stochastic time-fractional Benjamin–Bona–Mahony equation for modeling long surface gravity waves of small amplitude. Stoch. Anal. Applied.

Shabri, A., U. Ahmad and Z.A. Zakaria, 2011. TL-moments and L-moments estimation of the generalized logistic distribution. J. Math. Res. DOI: 10.5539/jmr.v3n1p97

Shoji, T. and H. Kitaura, 2006. Statistical and geo-statistical analysis of rainfall in central Japan. Comput. Geosci., 32: 1007-1024. DOI: 10.1016/j.cageo.2004.12.012

Tartaglia, V., E. Caporali, E. Cavigli and A. Moro, 2006. L-moments based assessment of a mixture model for frequency analysis of rainfall extremes, Adv. Geosci., 2: 331-334. DOI: 10.5194/adgeo-2-331-2006

Vogel, R.M. and N.M. Fennessey, 1993. L-moment diagrams should replace product moment diagrams. Water Resour. Res., 29: 1745-1752. DOI: 10.1029/93WR00341

Wazwaz, A.M., 2017. Exact solution and kink solutions for new (3+1)-dimensional nonlinear modified equations of wave propagation. Open Eng., 7: 169-174. DOI: 10.1515/eng-2017-0023

Wilks, D. S., 1998. Multi-site generalization of a daily stochastic precipitation model. J. Hydrol., 210: 178-191. DOI: 10.1016/S0022-1694(98)00186-3

Yoo, C., K.S. Jung and W.K. Tae, 2005. Rainfall frequency analysis using a mixed Gamma distribution: Evaluation of the global warming effect of daily rainfall. Hydrol. Proc., 19: 3851-3861. DOI: 10.1002/hyp.5985.