Two Component Mixture of Additive Uniform Exponential Distribution

Dr. Venkata Subbarao Uppu & Mandefro Abere
Department of Statistics, College of Natural & Computational Sciences, Adigrat University
Ethiopia (*sr.uppu@gmail.com)

ABSTRACT

This paper focuses on the Two Component Mixture of Additive Uniform Exponential Distribution which is an extension of additive uniform exponential distribution proposed by Venkata Subbarao Uppu (2010). The various distributional properties like, mean, variance, moment generating function, recurrence relation of moments, skewness, kurtosis, etc., are discussed. Some inferential aspects of this distribution are presented.

Key Words: Two component mixture, probability density function, mean, variance, moment generating function, central and raw moments, hazard rate, and estimation of parameters.

1. INTRODUCTION:

Probability distributions gained lot of importance due to their ready applicability for analyzing several data sets arising at space and biological experiments, Agricultural experiments, Business analytics, Data mining, Geo-informatics, Stochastic modeling etc., The distributions are broadly classified into discrete and continuous distributions. Continuous distributions are generally amenable to more aligent mathematical treatment than are discrete distributions. This makes them especially useful as approximations to discrete distributions. The continuous distributions are used in the construction of the models and in applying statistical techniques. Continuous distributions have been used as approximations to discrete distributions (N.L. Johnson, Samuel Kotz, N.Balakrishnan (2004)).

In the last five decades generalizations of continuous distributions are developed with various considerations. The common consideration is generalization of a distribution is introducing other parameters or ascribing a probability distribution to one or more parameters. In this fashion compound distributions were developed. Away from this another direction of generalization is considering mixture of distributions with weights to component densities. These mixtures of distributions are initially considered with the same population densities. Later they extended to the mixture of heterogeneous densities. Another generalization considered in these continuous distributions are Gram-charlier series expansions of distributions using normal as parent distribution. To accommodate a wider class of distributions one area of research is considering personean system of distributions.

However in some practical situations, the variate under study may be a sum of two or more random variables. Much work has been reported in literature regarding the distribution of the sum of random variables by considering that the variates under study are from the same homogenous population or the variation may be with reference to the parameters (but the functional form remains the same). In some other practical situations the random variable under study may be a sum of two different types of populations. For example in Manpower Modeling the complete length of service of an employee in the organization is the sum of two random time periods one for probation (temporary) period another for committed (permanent) period. Less work has been reported regarding the distribution of the sum of different random variables with different populations especially with uniform and exponential populations. In some situations the variable under study may be considered to be generated from heterogeneous population. For example in manpower planning when different types of employees like, semiskilled and skilled, are employed. The complete length of service of an employee in the organization forms a heterogeneous population. Similarly, in inventory management, when the procurement is done from different resources then the life time of a commodity forms a heterogeneous population. For these sorts of situations it is needed to consider the mixture distribution to analyze the random phenomenon.
The characteristic function is

\[ M = \int_{0}^{\infty} e^{itx} f(x) \, dx \]

On simplification we get

\[ = \frac{e^{\theta x}}{a_1} (e^{-a_1 \theta x} - 1) \quad \text{for} \quad a_1 \leq x \leq \infty \]

And

\[ f_2(x) = \frac{1}{a_2} \left( 1 - e^{-\theta x} \right) \quad \text{for} \quad 0 \leq x \leq a_2 \]

\[ = \frac{e^{-\theta x} - a_2 e^{-\theta x} - 1}{a_2} \quad \text{for} \quad a_2 \leq x \leq \infty \]

Here “P” is called the mixing parameter when \( P = 0 \) equation (1) reduces to Two parameter Additive Uniform Exponential distribution.

3. DISTRIBUTIONAL PROPERTIES:

The distribution function of this distribution is given by

\[ F_X(x) = P(X \leq x) \]

\[ = \frac{e^{\theta x}}{a_1} (1 - e^{-\theta x}) dt + \frac{e^{\theta x}}{a_1} \left( \frac{e^{-a_1 \theta x}}{a_1} - 1 \right) e^{-\theta x} dt \]

\[ + (1 - P) \left[ \frac{e^{\theta x}}{a_2} \left( \frac{e^{-a_2 \theta x}}{a_2} - 1 \right) e^{-\theta x} dt \right] \]

On simplification we get

\[ F_X(x) = P \left[ \frac{x}{a_1} + \frac{2e^{-\theta x} - e^{-a_1 \theta x}}{a_1} (1 + e^{-\theta x}) \right] + (1 - P) \left[ \frac{x}{a_2} + \frac{2e^{-\theta x} - e^{-a_2 \theta x}}{a_2} (1 + e^{-\theta x}) \right] \]

(2)

Mean of the distribution is given by

\[ E(X) = \frac{e^{\theta x}}{a_1} \left( 1 - e^{-\theta x} \right) dx + \frac{e^{\theta x}}{a_1} \left( \frac{e^{-a_1 \theta x}}{a_1} - 1 \right) e^{-\theta x} dx \]

\[ + (1 - P) \left[ \frac{e^{\theta x}}{a_2} \left( 1 - e^{-\theta x} \right) dx + \frac{e^{\theta x}}{a_2} \left( \frac{e^{-a_2 \theta x}}{a_2} - 1 \right) e^{-\theta x} dx \right] \]

On simplification we get

\[ E(X) = \frac{1}{\theta_1} + (1 - P) \left[ \frac{1}{\theta_2} + \frac{1}{\theta_2} \right] \]

(3)

Moment generating function is given by

\[ M_X(t) = E(e^{tx}) \]

\[ = \frac{e^{\theta x}}{a_1} (1 - e^{-\theta x}) dx + \frac{e^{\theta x}}{a_1} \left( \frac{e^{-a_1 \theta x}}{a_1} - 1 \right) e^{-\theta x} dx \]

\[ + (1 - P) \left[ \frac{e^{\theta x}}{a_2} (1 - e^{-\theta x}) dx + \frac{e^{\theta x}}{a_2} \left( \frac{e^{-a_2 \theta x}}{a_2} - 1 \right) e^{-\theta x} dx \right] \]

On simplification

\[ M_X(t) = \frac{\theta_1 \left( 1 - e^{-a_1 \theta t} \right)}{a_1 \theta t (1 - \theta_1)} + (1 - P) \frac{\theta_2 \left( 1 - e^{-a_2 \theta t} \right)}{a_2 \theta t (1 - \theta_2)} \]

(4)

Its characteristic function is

\[ \Phi_X(t) = \frac{\theta_1 \left( 1 - e^{-a_1 \theta t} \right)}{a_1 \theta t (1 - \theta_1)} + (1 - P) \frac{\theta_2 \left( 1 - e^{-a_2 \theta t} \right)}{a_2 \theta t (1 - \theta_2)} \]

(5)

The \( r \)th moment about origin is defined as

\[ \mu_r' = E(X^r) \]

\[ = \frac{e^{\theta x}}{a_1} (1 - e^{-\theta x}) dx + \frac{e^{\theta x}}{a_1} \left( \frac{e^{-a_1 \theta x}}{a_1} - 1 \right) x^r e^{-\theta x} dx \]

\[ + (1 - P) \left[ \frac{e^{\theta x}}{a_2} (1 - e^{-\theta x}) dx + \frac{e^{\theta x}}{a_2} \left( \frac{e^{-a_2 \theta x}}{a_2} - 1 \right) x^r e^{-\theta x} dx \right] \]

(6)
Taking \( r = 1,2,3,4 \) in equation (6) we get the first four raw moments of the distribution.

\[
\mu'_1 = P \left[ \frac{a_1}{\theta_1} + \frac{1}{\theta_1} \right] + (1 - P) \left[ \frac{\theta_1}{2} + \frac{1}{\theta_1} \right] 
\]

\[
\mu'_2 = P \left[ \frac{a_1}{\theta_1} + \frac{2}{\theta_1^2} + \frac{a_1^2}{3} \right] + (1 - P) \left[ \frac{a_2}{\theta_2} + \frac{2}{\theta_2} + \frac{a_2^2}{3} \right] 
\]

\[
\mu'_3 = P \left[ \frac{a_1^2}{\theta_1^2} + \frac{3a_1}{\theta_1^3} + \frac{a_1^3}{4} + \frac{6}{\theta_1^2} \right] + (1 - P) \left[ \frac{a_2^2}{\theta_2^2} + \frac{3a_2}{\theta_2^3} + \frac{a_2^3}{4} + \frac{6}{\theta_2^2} \right] 
\]

\[
\mu'_4 = P \left[ \frac{a_1^4}{5} + \frac{4a_1^2}{\theta_1^2} + \frac{a_1^3}{\theta_1^3} + \frac{12a_1}{\theta_1^4} + \frac{24}{\theta_1^5} \right] + (1 - P) \left[ \frac{a_2^4}{5} + \frac{4a_2^2}{\theta_2^2} + \frac{a_2^3}{\theta_2^3} + \frac{12a_2}{\theta_2^4} + \frac{24}{\theta_2^5} \right] 
\]

Using these raw we calculate the central moments they are as follows

\[
\mu_1 = 0 
\]

\[
\mu_2 = P \left[ \frac{1}{\theta_1} + \frac{a_1^2}{12} \right] + (1 - P) \left[ \frac{1}{\theta_2^2} + \frac{a_2^2}{12} \right] 
\]

This is called the variance of the distribution

\[
\mu_3 = P \left[ \frac{2}{\theta_1^2} \right] + (1 - P) \left[ \frac{2}{\theta_2^2} \right] 
\]

\[
\mu_4 = P \left[ \frac{a_1^2}{80} + \frac{2a_1^2}{\theta_1^2} + \frac{a_1^4}{80} \right] + (1 - P) \left[ \frac{a_2^4}{80} + \frac{2a_2^4}{\theta_2^2} + \frac{a_2^4}{80} \right] 
\]

Survival function is given by

\[
P(X) = 1 - F(X) = 1 - P \left[ \frac{x}{a_1} + \frac{2e^{-a_1 x}}{a_1 \theta_1} \right] \left[1 + e^{-\theta_1 x} \right] - (1 - P) \left[ \frac{x}{a_2} + \frac{2e^{-a_2 x}}{a_2 \theta_2} \right] \left[1 + e^{-\theta_2 x} \right] 
\]

The Hazard rate of the distribution is

\[
h(X) = \frac{f(X)}{1-F(X)} = \frac{P(f_1(x)+(1-P)f_2(x))}{(1-P) \left[ \frac{x}{a_1} + \frac{2e^{-a_1 x}}{a_1 \theta_1} \right] \left[1 + e^{-\theta_1 x} \right] - (1 - P) \left[ \frac{x}{a_2} + \frac{2e^{-a_2 x}}{a_2 \theta_2} \right] \left[1 + e^{-\theta_2 x} \right]} 
\]

Coefficient of skewness is

\[
\beta_1 = \frac{\mu_3}{\mu_2^2} = \left[ \frac{1}{\theta_1^2} + \frac{1}{\theta_2^2} \right] \left[\left(1 - P\right) \frac{1}{\theta_1^2} + \frac{1}{\theta_2^2} \right] 
\]

Coefficient of kurtosis is

\[
\beta_2 = \frac{\mu_4}{\mu_2^2} = \left[ \frac{a_1^4}{80} + \frac{a_1^2}{\theta_1^2} + \frac{a_1^3}{\theta_1^3} + \frac{12a_1}{\theta_1^4} + \frac{24}{\theta_1^5} \right] + (1 - P) \left[ \frac{a_2^4}{80} + \frac{a_2^2}{\theta_2^2} + \frac{a_2^3}{\theta_2^3} + \frac{12a_2}{\theta_2^4} + \frac{24}{\theta_2^5} \right] 
\]

This distribution does not hold additive property.

4. ESTIMATION OF PARAMETERS:

In this method the theoretical moments of the population and the sample moments are equated correspondingly to deduce the estimates of the parameters. Let \( u_1, u_2, u_3 \ldots \ldots u_n \) be a random sample of size ‘n’ drawn from the population having the probability function of the form given in equation (1). This distribution is having four parameters ‘\( a_1, a_2, \theta_1, \theta_2 \)’. Hence we equate the first four moments of the population and the sample, which leads to the following equations.

\[
\frac{1}{n} \sum_{i=0}^{n} u_i = P \left[ \frac{a_1}{2} + \frac{1}{\theta_1} \right] + (1 - P) \left[ \frac{\theta_1}{2} + \frac{1}{\theta_1} \right] 
\]

\[
\frac{1}{n} \sum_{i=0}^{n} u_i^2 = P \left[ \frac{a_1}{\theta_1} + \frac{2}{\theta_1^2} + \frac{a_1^2}{3} \right] + (1 - P) \left[ \frac{a_2}{\theta_2} + \frac{2}{\theta_2} + \frac{a_2^2}{3} \right] 
\]

\[
\frac{1}{n} \sum_{i=0}^{n} u_i^3 = P \left[ \frac{a_1^2}{\theta_1^2} + \frac{3a_1}{\theta_1^3} + \frac{a_1^3}{4} + \frac{6}{\theta_1^2} \right] + (1 - P) \left[ \frac{a_2^2}{\theta_2^2} + \frac{3a_2}{\theta_2^3} + \frac{a_2^3}{4} + \frac{6}{\theta_2^2} \right] 
\]

\[
\frac{1}{n} \sum_{i=0}^{n} u_i^4 = P \left[ \frac{a_1^4}{5} + \frac{4a_1^2}{\theta_1^2} + \frac{a_1^3}{\theta_1^3} + \frac{12a_1}{\theta_1^4} + \frac{24}{\theta_1^5} \right] + (1 - P) \left[ \frac{a_2^4}{5} + \frac{4a_2^2}{\theta_2^2} + \frac{a_2^3}{\theta_2^3} + \frac{12a_2}{\theta_2^4} + \frac{24}{\theta_2^5} \right] 
\]

Solving these equations simultaneously by using numerical methods like Newton Raphson we estimate the values of the parameters.

Sample mean \( \bar{u} \) is an unbiased estimator for the population mean ‘\( \mu \)’ and variance of \( \bar{u} \) is as follows.
\[ V(\bar{u}) = V \left[ \frac{1}{n} \sum_{i=1}^{n} u_i \right] = \frac{1}{n} \left[ \frac{\alpha_1^2}{12} + \frac{1}{\beta_1^2} \right] + (1 - P) \left[ \frac{\alpha_2^2}{12} + \frac{1}{\beta_2^2} \right] \]

and

\[ V(s^2) = \frac{\mu_4 - \mu_2^2}{n} - \frac{2[\mu_4 - 2\mu_2^2]}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3} \]

where is the \( i^{th} \) central moment [Crammer (1946)] given in equation (18) and (20), we obtain the variance of \( S^2 \).

5. CONCLUSIONS:

In the concept of manpower modeling several authors have approximated the complete length of service (CLS) distribution by different distributions. Silock (1954) and Bartholomew (1974) have considered that CLS of an employee in the organization follows exponential distribution. Rangarao, V. (1994) modified the manpower models given by Silock (1954) and Bartholomew (1974) by assuming that CLS of an employee in an organization follows the right truncated exponential distribution. Prakasharao, V.V.S (1997) developed manpower models with truncated compound beta distribution. In all these papers they assumed that the employee behavior throughout the period is homogeneous. M.P Ramayya (2005) have applied the merged exponential distribution for modeling the CLS of an employee in the organization. However, the CLS of an employee varies accordingly to the behavioral factors namely; semi committed and committed states (temporary and permanent) Mc Clean(1976). Generally an employee during semi committed state will have different rate of leaving and duration of stay during this period. During this period service can be characterized by uniform distribution. During the committed state, the rate of leaving is stabilized and constant, the duration of stay during this period can be characterized by exponential distribution. Therefore the total duration that an employee stays in an organization is a sum of two random variables associated with semi committed and committed states of the employee. Hence it is possible to develop a manpower model assuming that CLS of an employee in the organization is a random variable which follows AUED. This distribution has a tremendous potential in analyzing several factors like the probability that an employee surviving in the organization, the rate of labor wastage etc.

6. ACKNOWLEDGEMENT

We sincerely thank the referees for their valuable suggestions and comments regarding this paper.

7. REFERENCES

Bartholomew, D.J., 1974. The statistical approach to manpower planning. Statistician, 20: 3-26
