RT&A, No 1 (67)
Volume 17, March 2022

The New Mixed Erlang Distribution: A Flexible Distribution for Modeling Lifetime Data

Therrar Kadri[1], Souad Kadri[2], Seifedine Kadry[3*], and Khaled Smaili[4]

•

Department of Education, Lebanese University, Beirut, Lebanon[1]
Department of Mathematics and Physics, Lebanese International University, Khyara, Lebanon[1,2]
Department of Applied Data Science, Noroff University College, Norway[3]
Department of Applied Mathematics, Faculty of Sciences, Lebanese University, Zahle, Lebanon[4]
[*]Corresponding author

Abstract

We introduce a new mixed distribution of the Erlang distribution that is generated from the convolution of the Extension Exponential distribution denoted by the Mixed Erlang distribution (ME). We derive an exact closed expression of the probability density function which is used to obtain closed expressions of the cumulative function, reliability function, hazard function, moment generating function and kth moment. The method of maximum likelihood and method of moments is used for estimating the model parameters. Two applications to real data sets are given to illustrate the potentiality of this distribution.

Keywords: Erlang Distribution, Extension Exponential Distribution, Probability Density Function, Maximum likelihood estimation, Moments, Akaike Information Criterion

1. Introduction

Numerous classical distributions have been extensively used over the past decades for modeling data in many applied areas such as lifetime analysis, finance and insurance, as the Exponential distribution and its alternatives the Erlang and Gamma distribution, see [1] and [2]. There is a clear need for extended forms of these distributions. In recent statistical literature modified extensions of the Exponential distributions have been proposed to give more flexibility to model real data. For example, Gupta and Kundu [6] introduced an extension of the Exponential distribution typically called the generalized exponential (GE) distribution and Mudholkar et al. [9] introduced the exponentiated Weibull (EW) distribution as another extension. Gómez et al. [5] introduced a new extension of the Exponential distribution denoted as the Extended Exponential (EE) distribution of two positive parameters.

On the other hand, the sum of independent random variables, the convolution of random variables, also plays a significant role in modeling many events in most domains of science, as communications, computer science, and teletraffic engineering (Trivedi [12]; Jasiulewicz and Kordecki [7]), Markov process, reliability and performance evaluation (Kadri et al. [8]; Smaili et al. [10]). A comprehensive study of these distributions is needed for modeling and the importance of providing closed and exact forms of probability density function (PDF), cumulative distribution function (CDF), reliability and hazard functions, moment generating function (MGF), and k^{th} moments...

The main aim of this paper is to study a new distribution generated from the convolution of the Extension Exponential distribution which is observed to have the form of a mixed distribution of the Erlang distribution. We denote this distribution by the Mixed Erlang Distribution (ME). We provide a comprehensive account of the mathematical properties of this new distribution by deriving closed and exact forms of PDF, CDF, reliability and hazard functions, MGF, and k^{th} moments. Moreover, we propose that our distribution is quite flexible, evidence by its closed and simple expressions. Also, this distribution can be used quite effectively in analyzing positive data in place of proposed distribution in the literature, which indicates that the ME distribution is a serious competitor to the others. Thus, we preform a parameter estimation of the model by the method of maximum likelihood and the method of moment. Next, we fit the new distribution to two real data sets to examine the performance of the new model and compare it to lately new distributions proposed in the statistical literature.

2. Some Preliminaries

2.1. Erlang Distribution

Erlang distribution is a two-parameter continuous probability distribution with shape integer parameter n and scale parameter $\alpha > 0$. It is considered as the sum of n independent identical Exponential distributions of parameter α, so for $n = 1$ the Erlang distribution is simplified to the Exponential distribution. Erlang distribution like Exponential is widely used in life time analysis, see [4].

Let $Y \sim Erl(n, \alpha)$. The PDF of Y is given as:

$$f_Y(t) = \frac{(\alpha t)^{n-1} \alpha e^{-\alpha t}}{(n-1)!}, t > 0 \tag{1}$$

and CDF of Y is

$$F_Y(t) = 1 - \frac{\Gamma(n, \alpha t)}{(n-1)!} \tag{2}$$

where $\Gamma(\cdot, \cdot)$ is the incomplete Gamma function. Also the MGF of Y is

$$\phi_Y(t) = \left(\frac{\alpha}{\alpha - t}\right)^n, t < \alpha \tag{3}$$

and the moment of order k of Y is

$$E[Y^k] = \frac{\Gamma(n+k)}{\alpha^k \Gamma(n)} \tag{4}$$

2.2. Extension Exponential Distribution

Gómez et al. [5], introduced a new extension of the Exponential distribution denoted as the Extended Exponential (EE) distribution of two positive parameters, denoted by $EE(\alpha, \beta)$. They characterized this distribution having $X \sim EE(\alpha, \beta)$ with PDF

$$f_X(t) = \frac{\alpha^2(1+\beta t)e^{-\alpha t}}{\alpha + \beta} \quad \alpha, \beta, t > 0$$

where α is a scale parameter and β is a shape parameter.

The EE distribution is considered as a mixed distribution of the Exponential distribution, $E(\alpha)$, and Erlang distribution $Erl(2, \alpha)$, i.e.

$$f_X(t) = \frac{\alpha}{\alpha + \beta} f_{E(\alpha)}(t) + \frac{\beta}{\alpha + \beta} f_{Erl(2,\alpha)}(t) \tag{5}$$

For further use, we have the Laplace transform of $f_X(t)$ as

$$\mathcal{L}\{f_X(t)\} = \frac{\alpha^2(t + \alpha + \beta)}{(\alpha + \beta)(t + \alpha)^2}. \tag{6}$$

Other properties of EE distribution can be found in [5].

3. Mixed Erlang Distribution

Let $X_j,\ j = 1, 2, ..., n$ be n identical independent (iid) random variables that follow Extension Exponential distribution i.e. $X_j \sim EE(\alpha, \beta)$ and let $S_n = \sum_{j=1}^{n} X_j$. We denote $S_n \sim ME(\alpha, \beta, n)$ to be the Mixed Erlang distribution for $\alpha > 0, \beta > 0$ and $n \in \mathbb{N}^*$. This name is derived from the obtained expressions of this distribution in this section, which has the form of a mixed Erlang (ME) distribution. We start by deriving the PDF of this new distribution in an exact closed form. The obtained simple form will help us to derive the other mathematical functions to characterize the ME distribution.

3.1. PDF of the ME distribution

It is known that the sum of independent distributions is the convolution random variable and its PDF can be determined by the n convolution of the PDF of the summands X_j, which is an approach used. Here we take the advantage of Laplace transform over convolution to obtain our expression.

Theorem 1. Let $S_n \sim ME(\alpha, \beta, n),\ \alpha > 0, \beta > 0$ and $n \in \mathbb{N}^*$. Then the PDF of S_n is given by

$$f_{S_n}(t) = \sum_{i=0}^{n} A_i f_{Y_i}(t)$$

where

$$A_i = \frac{\binom{n}{i}\alpha^{n-i}\beta^i}{(\alpha+\beta)^n} \text{ and } Y_i \sim Erl(n+i, \alpha) \tag{7}$$

Proof. Let $X_j \sim EE(\alpha, \beta)$, $j = 1, 2, ..., n$ be n iid distributions and let $S_n = \sum_{j=1}^{n} X_j \sim ME(\alpha, \beta, n)$. We have $f_{S_n}(t)$ is the convolution of the PDF of X_j. Thus the Laplace transform of $f_{S_n}(t)$ is the product of identical distribution of EE and we get

$$\mathcal{L}\{f_{S_n}(t)\} = \left[\mathcal{L}\{f_{X_i}(t)\}\right]^n$$

From Equation (6) $\mathcal{L}\{f_{X_i}(t)\} = \frac{\alpha^2(\alpha+\beta+t)}{(\alpha+\beta)(t+\alpha)^2}$. We get

$$\mathcal{L}\{f_{S_n}(t)\} = \frac{\alpha^{2n}(\alpha+\beta+t)^n}{(\alpha+\beta)^n(t+\alpha)^{2n}}$$

and

$$f_{S_n}(t) = \mathcal{L}^{-1}\left\{\frac{\alpha^{2n}(\alpha+\beta+t)^n}{(\alpha+\beta)^n(t+\alpha)^{2n}}\right\} = \frac{\alpha^{2n}}{(\alpha+\beta)^n}\mathcal{L}^{-1}\left\{\frac{(\alpha+\beta+t)^n}{(t+\alpha)^{2n}}\right\}$$
$$= \frac{\alpha^{2n}e^{-\alpha t}}{(\alpha+\beta)^n}\mathcal{L}^{-1}\left\{\frac{(\beta+t)^n}{t^{2n}}\right\}$$

However, $(\beta+t)^n = \sum_{i=0}^{n}\binom{n}{i}\beta^i t^{n-i}$, then $\frac{(\beta+t)^n}{t^{2n}} = \sum_{i=0}^{n}\binom{n}{i}\beta^i t^{-n-i}$. Also $\mathcal{L}^{-1}\left\{t^{-n-i}\right\} = \frac{t^{(i+n-1)}}{(i+n-1)!}$. Thus we conclude that

$f_{S_n}(t) = \frac{\alpha^{2n}e^{-\alpha t}}{(\alpha+\beta)^n}\sum_{i=0}^{n}\binom{n}{i}\beta^i\frac{t^{(i+n-1)}}{(i+n-1)!}$. Next, we rearrange the sum to pull out the closest PDF which is a Erlang distribution of the form $\frac{\alpha^{n+i}t^{(n+i-1)}e^{-\alpha t}}{(n+i-1)!} = f_{Y_i}(t)$, where $Y_i \sim Erl(n+i, \alpha)$. So we can rewrite

$$f_{S_n}(t) = \sum_{i=0}^{n}\frac{\binom{n}{i}\beta^i\alpha^{n-i}}{(\alpha+\beta)^n} \times \frac{\alpha^{n+i}t^{(n+i-1)}e^{-\alpha t}}{(n+i-1)!}$$
$$= \sum_{i=0}^{n} A_i f_{Y_i}(t)$$

where $A_i = \frac{\binom{n}{i}\alpha^{n-i}\beta^i}{(\alpha+\beta)^n}$.

■

In the following, we give another proof of the previous theorem by using the approach of convolution of PDF of independent random variables instead of the Laplace inverse approach.

Proof. [Alternate Proof]Let $X_j \sim EE(\alpha, \beta)$, $j = 1, 2, ..., n$ and S_n is the convolution random variable of EE. Then

$$f_{S_n}(t) = (f_{X_1} * f_{X_2} * ... * f_{X_n})(t) \tag{8}$$

However, from Equation (5), the PDF of EE can be expressed as

$$f_{X_j}(t) = \frac{\alpha}{\alpha+\beta} f_{E(\alpha)}(t) + \frac{\beta}{\alpha+\beta} f_{Erl(2,\alpha)}(t)$$

Substitute $f_{X_j}(t)$ in 8 to obtain

$$f_{S_n}(t) = \overset{n}{\circledast}\left(\frac{1}{\alpha+\beta}\left(\alpha f_{E(\alpha)} + \beta f_{Erl(2,\alpha)}\right)\right)(t)$$

where $\overset{n}{\circledast}(g)$ means that the expression is convoluted n times by itself. Furthermore, convolution is associative with scalar multiplication, thus

$$f_{S_n}(t) = \frac{1}{(\alpha+\beta)^n} \overset{n}{\circledast}\left(\alpha f_{E(\alpha)} + \beta f_{Erl(2,\alpha)}\right)(t)$$

Now, using the generalized binomial expansion over the convolution operation, we obtain

$$\begin{aligned} f_{S_n}(t) &= \frac{1}{(\alpha+\beta)^n} \sum_{i=0}^{n} \binom{n}{i} \overset{n-i}{\circledast} \alpha f_{E(\alpha)}(t) * \overset{i}{\circledast} \beta f_{Erl(2,\alpha)}(t) \\ &= \frac{1}{(\alpha+\beta)^n} \sum_{i=0}^{n} \binom{n}{i} \alpha^{n-i}\beta^i \left(\overset{n-i}{\circledast} f_{E(\alpha)} * \overset{i}{\circledast} f_{Erl(2,\alpha)}\right)(t) \end{aligned}$$

Also the convolution of $n-i$ identical Exponential distribution is the Erlang distribution $Erl(n-i, \alpha)$ or $\overset{n-i}{\circledast} f_{E(\alpha)} = f_{Erl(n-i,\alpha)}$ and the convolution of i Erlang distributions $Erl(2, \alpha)$ is the Erlang distribution $Erl(2i, \alpha)$ or $\overset{i}{\circledast} f_{Erl(2,\alpha)} = f_{Erl(2i,\alpha)}$. Thus, we get

$$f_{S_n}(t) = \frac{1}{(\alpha+\beta)^n} \sum_{i=0}^{n} \binom{n}{i} \alpha^{n-i}\beta^i \left(f_{Erl(n-i,\alpha)} * f_{Erl(2i,\alpha)}\right)(t).$$

On the other hand $Erl(n-i, \alpha) * Erl(2i, \alpha) = Erl(n+i, \alpha)$ and thus

$$\begin{aligned} f_{S_n}(t) &= \sum_{i=0}^{n} \frac{\binom{n}{i}\alpha^{n-i}\beta^i}{(\alpha+\beta)^n} f_{Erl(n+i,\alpha)}(t) \\ &= \sum_{i=0}^{n} A_i f_{Y_i}(t) \end{aligned}$$

where $A_i = \frac{\binom{n}{i}\alpha^{n-i}\beta^i}{(\alpha+\beta)^n}$ and $Y_i \sim Erl(n+i, \alpha)$.

■

In the following corollary, we give the PDF of ME in one expression, related to regularized confluent hypergeometric function.

Corollary 1. Let $S_n \sim ME(\alpha, \beta, n)$, $\alpha > 0$, $\beta > 0$ and $n \in \mathbb{N}^*$. Then the PDF of S_n is given by

$$f_{S_n}(t) = \frac{\alpha^{2n} t^{n-1} e^{-\alpha t}}{(\alpha+\beta)^n} \; {}_1\widetilde{F}_1\left(-n; n, -t\beta\right)$$

where ${}_1\widetilde{F}_1\,(a;b;x)$ is the regularized confluent hypergeometric function.

Proof. From Theorem 1 we have $f_{S_n}(t) = \sum\limits_{i=0}^{n} A_i f_{Y_i}(t)$ with $A_i = \frac{\binom{n}{i}\alpha^{n-i}\beta^i}{(\alpha+\beta)^n}$ and $Y_i \sim Erl(n+i,\alpha)$. However, the PDF of Y_i from Equation (1) is given by $f_{Y_i}(t) = \frac{(\alpha t)^{n+i-1}\alpha e^{-\alpha t}}{(n+i-1)!} I_{(0,\infty)(t)}$. Thus

$$\begin{aligned} f_{S_n}(t) &= \sum_{i=0}^{n} \frac{\binom{n}{i}\alpha^{n-i}\beta^i}{(\alpha+\beta)^n} \times \frac{(\alpha t)^{n+i-1}\alpha e^{-\alpha t}}{(n+i-1)!} \\ &= \frac{\alpha^{2n}e^{-\alpha t}}{(\alpha+\beta)^n} \sum_{i=0}^{n} \frac{\frac{n!}{(n-i)!i!}\beta^i}{(n+i-1)!} t^{n+i-1} \\ &= \frac{\alpha^{2n}t^{n-1}e^{-\alpha t}}{(\alpha+\beta)^n}\, {}_1\widetilde{F}_1\,(-n;n,-t\beta) \end{aligned}$$

where $\sum\limits_{i=0}^{n} \frac{\frac{n!}{(n-i)!i!}(t\beta)^i}{(n+i-1)!} = {}_1\widetilde{F}_1\,(-n;n,-t\beta)$ is the regularized confluent hypergeometric function which is defined as ${}_1\widetilde{F}_1\,(a;b,x) = \frac{{}_1F_1\,(a;b,x)}{\Gamma(b)}$ having ${}_1F_1\,(a;b,x)$ be the Kummer confluent hypergeometric function.

■

3.2. CDF, MGF and other functions for ME distribution

In Theorem 1, we found a closed expression of the PDF for sum of identical *EE* random variables and we gave the PDF expression as $\sum\limits_{i=0}^{n} A_i f_{Y_i}(t)$. This expression shows that our distribution is also a mixed distribution of the Erlang distribution. We take an advantage of this expression to find the other statistical characterization as CDF, MGF, moment of order k, reliability and hazard functions for *ME* distribution. Next, we derive exact closed expressions of these functions.

Theorem 2. Let $S_n \sim ME(\alpha,\beta,n)$, $\alpha > 0$, $\beta > 0$ and $n \in \mathbb{N}^*$. Then the CDF of S_n is given by

$$F_{S_n}(t) = \sum_{i=0}^{n} A_i F_{Y_i}(t)$$

where A_i is defined in Equation (7) and F_{Y_i} is the CDF of $Y_i \sim Erl(n+i,\alpha)$.

Proof. From Theorem 1, the PDF of S_n is $f_{S_n}(t) = \sum\limits_{i=0}^{n} A_i f_{Y_i}(t)$. The CDF of $S_n$is defined as

$$F_{S_n}(t) = \int_0^t f_{S_n}(x)dx = \int_0^t \sum_{i=0}^{n} A_i f_{Y_i}(x)dx = \sum_{i=0}^{n} A_i \int_0^t f_{Y_i}(x)dx = \sum_{i=0}^{n} A_i F_{Y_i}(t).$$

■

Lemma 1. $\sum\limits_{i=0}^{n} A_i = 1$.

Proof. Let $F_{Y_i}(t)$ and $F_{S_n}(t)$ be the CDF of Y_i and S_n respectively. However, the limit at infinity of any CDF is 1. Starting from the expression of the CDF in Theorem 2, $F_{S_n}(t) = \sum_{i=0}^{n} A_i F_{Y_i}(t)$, we have $\lim\limits_{t\to\infty} F_{S_n}(t) = \lim\limits_{t\to\infty} \sum_{i=0}^{n} A_i F_{Y_i}(t)$, thus $\sum_{i=0}^{n} A_i = 1$.

■

Next, we give another expression for the CDF of S_n.

Corollary 2. Let $S_n \sim ME(\alpha,\beta,n)$, $\alpha > 0$, $\beta > 0$ and $n \in \mathbb{N}^*$. Then the CDF of S_n is

$$F_{S_n}(t) = 1 - \frac{\alpha^n}{(\alpha+\beta)^n} \sum_{i=0}^{n} \frac{\binom{n}{i}\beta^i\Gamma(n+i,\alpha t)}{\alpha^i\,(n+i-1)!}$$

where $\Gamma(\cdot,\cdot)$ is the upper incomplete gamma function.

Proof. From Theorem 2, the CDF of S_n is $F_{S_n}(t) = \sum_{i=0}^{n} A_i F_{Y_i}(t)$. However, from Equation (2) we have $F_{Y_i}(t) = 1 - \frac{\Gamma(n+i,\alpha t)}{(n+i-1)!}$. Therefore,

$$F_{S_n}(t) = \sum_{i=0}^{n} \left(A_i - A_i \frac{\Gamma(n+i,\alpha t)}{(n+i-1)!} \right),$$

but from Lemma 1, $\sum_{i=0}^{n} A_i = 1$ This leads to $F_{S_n}(t) = 1 - \sum_{i=0}^{n} A_i \frac{\Gamma(n+i,\alpha t)}{(n+i-1)!}$. Moreover, from Equation (7) $A_i = \frac{\binom{n}{i}\alpha^{n-i}\beta^i}{(\alpha+\beta)^n}$and we get

$$\begin{aligned} F_{S_n}(t) &= 1 - \sum_{i=0}^{n} A_i \frac{\Gamma(n+i,\alpha t)}{(n+i-1)!} \\ &= 1 - \frac{\alpha^n}{(\alpha+\beta)^n} \sum_{i=0}^{n} \frac{\binom{n}{i}\beta^i \Gamma(n+i,\alpha t)}{\alpha^i\,(n+i-1)!} \end{aligned}$$

■

Theorem 3. Let $S_n \sim ME(\alpha,\beta,n),\ \alpha > 0, \beta > 0$ and $n \in \mathbb{N}^*$. Then the MGF of S_n is

$$\phi_{S_n}(t) = \sum_{i=0}^{n} A_i \phi_{Y_i}(t)$$

where A_i is defined in Equation (7) and ϕ_{Y_i} is the MGF of $Y_i \sim Erl(n+i,\alpha)$.

Proof. Referring to Theorem 1 the PDF of S_n is $f_{S_n} = \sum_{i=0}^{n} A_i f_{Y_i}(t)$.Thus,

$$\phi_{S_n}(t) = \int_{-\infty}^{+\infty} e^{tx} f_{S_n}(x)dx = \int_{-\infty}^{+\infty} e^{tx} \left(\sum_{i=0}^{n} A_i f_{Y_i}(x) \right) dx = \sum_{i=0}^{n} A_i \int_{-\infty}^{+\infty} e^{tx} f_{Y_i}(x)dx$$

$$\text{but} \int_{-\infty}^{+\infty} e^{tx} f_{Y_i}(x)dx = \phi_{Y_i}(t), \text{ thus } \phi_{S_n}(t) = \sum_{i=0}^{n} A_i \phi_{Y_i}(t).$$

■

Corollary 3. Let $S_n \sim ME(\alpha,\beta,n),\ \alpha > 0, \beta > 0$ and $n \in \mathbb{N}^*$. Then the MGF of S_n is

$$\phi_{S_n}(t) = \frac{\alpha^{2n}}{(\alpha+\beta)^n} \sum_{i=0}^{n} \frac{\binom{n}{i}\beta^i}{(\alpha-t)^{n+i}}.$$

Proof. From Theorem 3 $\phi_{S_n}(t) = \sum_{i=0}^{n} A_i \phi_{Y_i}(t)$ and the MGF of Erlang distribution Y_i is given in Equation (3) as $\phi_{Y_i}(t) = \left(\frac{\alpha}{\alpha-t}\right)^{n+i}$ which leads to $\phi_{S_n}(t) = \sum_{i=0}^{n} A_i \left(\frac{\alpha}{\alpha-t}\right)^{n+i}$. Moreover, from Equation (7) $A_i = \frac{\binom{n}{i}\alpha^{n-i}\beta^i}{(\alpha+\beta)^n}$ then

$$\begin{aligned} \phi_{S_n}(t) &= \sum_{i=0}^{n} \frac{\binom{n}{i}\alpha^{n-i}\beta^i}{(\alpha+\beta)^n} \left(\frac{\alpha}{\alpha-t}\right)^{n+i} \\ &= \frac{\alpha^{2n}}{(\alpha+\beta)^n} \sum_{i=0}^{n} \frac{\binom{n}{i}\beta^i}{(\alpha-t)^{n+i}} \end{aligned}$$

■

Theorem 4. Let $S_n \sim ME(\alpha, \beta, n)$, $\alpha > 0$, $\beta > 0$ and $n \in \mathbb{N}^*$. Then the reliability function of S_n is

$$R_{S_n}(t) = \sum_{i=0}^{n} A_i R_{Y_i}(t)$$

and the hazard function of S_n is given as

$$h_{S_n}(t) = \frac{\sum_{i=0}^{n} A_i h_i(t) R_{Y_i}(t)}{\sum_{i=0}^{n} A_i R_{Y_i}(t)}$$

where $h_{Y_i}(t)$ and $R_{Y_i}(t)$ are the hazard and reliability functions of $Y_i \sim Erl(n+i, \alpha)$ respectively, and A_i is defined in Equation (7).

Proof. $R_{S_n}(t) = 1 - F_{S_n}(t)$, and from Theorem 2 we have $F_{S_n}(t) = \sum_{i=0}^{n} A_i F_{Y_i}(t)$

$$R_{S_n}(t) = 1 - \sum_{i=0}^{n} A_i F_{Y_i}(t),$$

but $F_{Y_i}(t) = 1 - R_{Y_i}(t)$, and from Lemma 1, we have $\sum_{i=0}^{n} A_i = 1$, thus $R_{S_n}(t) = \sum_{i=0}^{n} A_i R_{Y_i}(t)$.

On the other hand, the expression of hazard function is given by

$$h_{S_n}(t) = \frac{f_{S_n}(t)}{R_{S_n}(t)} = \frac{\sum_{i=0}^{n} A_i f_{Y_i}(t)}{\sum_{i=0}^{n} A_i R_{Y_i}(t)}$$

however, $f_{Y_i}(t) = h_i(t) R_{Y_i}(t)$, then

$$h_{S_n}(t) = \frac{\sum_{i=0}^{n} A_i h_i(t) R_{Y_i}(t)}{\sum_{i=0}^{n} A_i R_{Y_i}(t)}.$$

■

Theorem 5. Let $S_n \sim ME(\alpha, \beta, n)$, $\alpha > 0$, $\beta > 0$ and $n \in \mathbb{N}^*$. Then the moment of order k of S_n is

$$E[S_n^k] = \sum_{i=0}^{n} A_i E[Y_i^k] = \frac{k! \alpha^{n-k}}{(\alpha+\beta)^n} \sum_{i=0}^{n} \frac{\beta^i}{\alpha^i} \binom{n}{i} \binom{n+i+k}{n+i}$$

where $Y_i \sim Erl(n+i, \alpha)$, and A_i is defined in Equation (7).

Proof. From Theorem 3, we have $\phi_{S_n}(t) = \sum_{i=0}^{n} A_i \phi_{Y_i}(t)$. Now the moment of S_n of order k is given by $E[S_n^k] = \frac{d^k \phi_{S_n}(t)}{dt^k}\Big|_{t=0} = \sum_{i=0}^{n} A_i \frac{d^k \phi_{Y_i}(t)}{dt^k}\Big|_{t=0} = \sum_{i=0}^{n} A_i E[Y_i^k]$. Moreover, $E[Y_i^k] = \frac{\Gamma(n+i+k)}{\alpha^k \Gamma(n+i)}$ and from Equation (7) $A_i = \frac{\binom{n}{i} \alpha^{n-i} \beta^i}{(\alpha+\beta)^n}$ then

$$\begin{aligned} E[S_n^k] &= \sum_{i=0}^{n} \frac{\binom{n}{i} \alpha^{n-i} \beta^i}{(\alpha+\beta)^n} \frac{\Gamma(n+i+k)}{\alpha^k \Gamma(n+i)} \\ &= \frac{\alpha^{n-k}}{(\alpha+\beta)^n} \sum_{i=0}^{n} \frac{\beta^i}{\alpha^i} \frac{\binom{n}{i} \Gamma(n+i+k)}{\Gamma(n+i)} \\ &= \frac{k! \alpha^{n-k}}{(\alpha+\beta)^n} \sum_{i=0}^{n} \frac{\beta^i}{\alpha^i} \binom{n}{i} \binom{n+i+k}{n+i} \end{aligned}$$

as

$$\frac{\binom{n}{i} \Gamma(n+i+k)}{\Gamma(n+i)} = \frac{\frac{n!}{(n-i)! i!} k! \, (n+i+k-1)!}{(n+i-1)! k!} = k! \binom{n}{i} \binom{n+i+k}{n+i}$$

■

We end this part to point out the importance of writing the PDF of our ME distribution as linear combination PDF of the known Erlang distribution. This expression facilitates in determining the other statistical expressions as CDF, reliability and hazard functions, moment generating function (MGF), and k^{th} moments. This procedure was adopted by Smaili et al. [10] and [11]. Later, these expressions are used to give an estimated model for a real-life data.

4. Real Life Data

To illustrate the new results presented in this paper, we fit the ME distribution to two examples of real data. The MLE and MME approaches are employed to estimate the parameters of the real-life data and MATHEMATICA software is used. We analyze real data sets to show that the ME distribution can be a better model than other existing distributions. We consider the distributions in recent papers that proposed their distribution to fit the model data. For each data set, we compare the fitted distributions using the three criteria: AIC (Akaike Information Criterion), $AICC$ (Akaike Information Criterion Corrected) and BIC (Bayesian Information Criterion). Let us be precise that $\log(L)$ is the log-likelihood taking with the estimate values, $AIC = 2k - 2\log(L)$, $AICC = AIC + \frac{2k(k+1)}{n-k-1}$ and $BIC = -2\log(L) + k\log(n)$, where k denotes the number of estimated parameters and n denotes the sample size. The best fitted distribution corresponds to lower AIC, $AICC$ and BIC. Also the histogram and the estimated PDFs and CDFs for the best fitted models to the two data are displayed in Figures 1 and 2, respectively.

Data set 1: The data set contains n = 63 measures related to the strength of 1.5cm glass bers. It is reported in Smith and Naylor (1987): 0.55, 0.93, 1.25, 1.36, 1.49, 1.52, 1.58, 1.61, 1.64, 1.68, 1.73, 1.81, 2, 0.74, 1.04, 1.27, 1.39, 1.49, 1.53, 1.59, 1.61, 1.66, 1.68, 1.76, 1.82, 2.01, 0.77, 1.11, 1.28, 1.42, 1.5, 1.54, 1.6, 1.62, 1.66, 1.69, 1.76, 1.84, 2.24, 0.81, 1.13, 1.29, 1.48, 1.5, 1.55, 1.61, 1.62, 1.66, 1.7, 1.77, 1.84, 0.84, 1.24, 1.3, 1.48, 1.51, 1.55, 1.61, 1.63, 1.67, 1.7, 1.78, 1.89.

We chose the analysis done by Chesneau in [3] for this data. Chesneau compared the Lindley, Exponential, Exponentiated Exponential (EExp), and Exponential Hypoexponential distribution (EHypo). The corresponding PDF of EExp and EHypo are given by

$$f_{EExp}(x) = \lambda\alpha e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^{\alpha-1}$$
$$f_{EHyp}(x) = \lambda\alpha\left(1 + 10\alpha\right)e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^{\alpha-1}\left(1 - \left(1 - e^{-\lambda x}\right)^{0.1}\right)$$

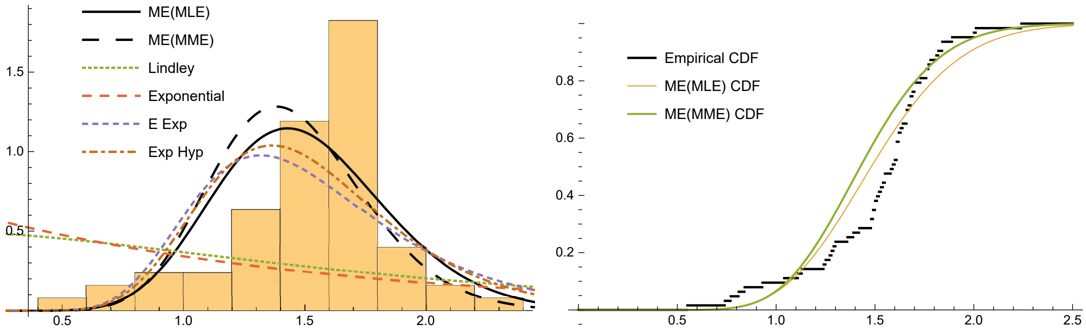
respectively. We derive two estimated distributions of the ME distribution using MLE and MME and used as a competitive distribution of the previous ones. See Table 1. This table shows that the ME model gives a better fit to this data than the other distributions. The plots in Figures 1 also indicate the same thing. So, the ME model could be chosen as the best model.

Table 1: *MLE and MME of ME Distribution with MLEs of competitor distributions and AIC, AICC and BIC of data set 1*

Model	Estimated Parameters	AIC	AICC	BIC
ME(MLE)	$\widehat{n} = 13$, $\widehat{\alpha} = 13.771948428082293$, $\widehat{\beta} = 20.34250599538599$	50.547	50.747	63.1195
ME(MME)	$\widehat{n} = 15$, $\widehat{\alpha} = 16.653713280301165$, $\widehat{\beta} = 25.644927330965096$	53.678	53.8784	66.251
Lindley	$\widehat{\theta} = 0.996116$	164.56	164.62	166.70
Exponential	$\widehat{\lambda} = 0.663647$	179.66	179.73	181.80
EExp	$\widehat{\alpha} = 31.3489$, $\widehat{\lambda} = 2.61157$	66.76	66.96	71.05
EHypo	$\widehat{\alpha} = 24.0816$, $\widehat{\lambda} = 1.83894$	59.67	59.87	63.96

Data set 2: The data consist of 101 observations. The data was presented in Birnbaum & Saunders (1969) and correspond to the fatigue time of 101 6061-T6 aluminum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second (cps). The data are: 70, 90, 96, 97, 99, 100, 103, 104, 104, 105, 107, 108, 108, 108, 109, 109, 112, 112, 113, 114, 114, 114, 116, 119, 120, 120, 120, 121, 121, 123, 124, 124, 124, 124, 124, 128, 128, 129, 129, 130, 130, 130, 131, 131, 131, 131, 131, 132, 132, 132, 133, 134, 134, 134, 134, 134, 136, 136, 137, 138, 138, 138, 139, 139, 141, 141, 142, 142, 142, 142, 142, 142, 144, 144, 145, 146, 148, 148, 149, 151, 151, 152, 155, 156, 157, 157, 157, 157, 158, 159, 162, 163, 163, 164, 166, 166, 168, 170, 174, 196, 212.

We chose the analysis done by Yousof et al. in [13] for this data. They compared the Weibull, Wei-Weibull (WW) and their Weibull-Weibull logarithmic (WWL) distribution to fit this data. The



(a) *Fitted pdfs for data 1* **(b)** *Fitted cdfs for data 1*

Figure 1: *The two figures show a best fitting for the EE distribution*

CDF of WW and WWL distribution are given by

$$F_{WW}(x;\alpha,\beta,\lambda,\gamma) = 1 - e^{-\alpha\left(e^{\lambda x^{\gamma}}-1\right)^{\beta}}$$

$$F_{WWL}\left(x;\alpha,\beta,\lambda,\gamma,p\right) = \frac{p\alpha\beta\gamma\lambda x^{\gamma-1}\left(e^{\lambda x^{\gamma}}-1\right)^{\beta-1}\left(e^{\left(\lambda x^{\gamma-1}-\alpha\left(e^{\lambda x^{\gamma}}-1\right)^{\beta}\right)}\right)}{\left(p\left(1-e^{-\alpha\left(e^{\lambda x^{\gamma}}-1\right)^{\beta}}\right)-1\right)\ln\left(1-p\right)}$$

respectively for $\alpha, \beta, \lambda, \gamma > 0$ and $0 < p < 1$. We derive two estimated distributions of the ME distribution using MLE and MME and used as a competitive distribution of the previous ones, see Table 2. This table shows that the ME model gives a better fit to this data than the other distributions. The plots in Figure 2 also indicate the same thing. So, the ME model for the second time could be chosen as the best model.

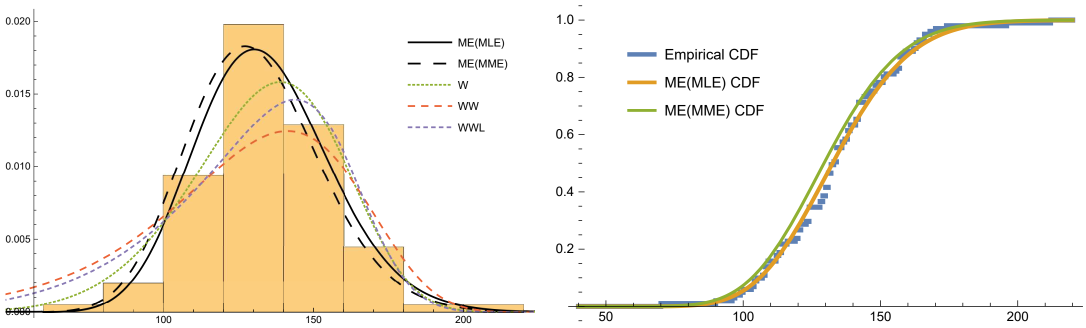
Table 2: *MLE and MME of ME Distribution with MLEs of competitor distributions and AIC, AICC and BIC of data set 2*

Model	Estimated Parameters	AIC	AICC	BIC
ME(MLE)	$\widehat{n} = 32, \widehat{\alpha} = 0.31305352844373824, \widehat{\beta} = 0.13952982315141454$	916.557	916.679	931.018
ME(MME)	$\widehat{n} = 28, \widehat{\alpha} = 0.3149943298488485, \widehat{\beta} = 0.27793752048684467$	918.586	933.047	918.709
Weibull	$\widehat{\lambda} = 0.0036, \widehat{\gamma} = 1.1516$	1167.38	1167.5	1172.61
WW	$\widehat{\lambda} = 0.0036, \widehat{\gamma} = 1.1516$	1167.38	1167.5	1172.61
WWL	$\widehat{\alpha} = 0.0060, \widehat{\beta} = 3.1600, \widehat{\lambda} = 0.0873, \widehat{\gamma} = 0.6264, \widehat{p} = 0.8732$	948.49	948.91	961.56

We see in Tables 1 and 2 that the ME distribution has the smallest AIC, $AICC$ and BIC for the two data sets, compared with lately proposed distributions, indicating that the ME distribution is a serious competitor to the other considered distributions.

5. Conclusion

A new distribution, Mixed Erlang (ME) distribution, has been proposed and its properties are studied. We derived exact closed expressions of the PDF, CDF, reliability function, hazard function, MGF, and k^{th} moments. We have studied the maximum likelihood estimators and method of moments estimators and the parameters estimation is carried out in the presence of real data. We presented two real life data sets, and our ME distribution was compared with lately proposed distributions and showed that the ME distribution is a serious competitor to the others.



(a) *Fitted pdfs for data 2* **(b)** *Fitted cdfs for data 2*

Figure 2: *The two figures show a best fitting for the EE distribution*

References

[1] Abdelkader, Y.H.: Erlang distributed activity times in stochastic activity networks. Kybernetika. 39(3), 347–358 (2003)
[2] Bocharov, P.P., D'Apice, C., Pechinkin, A.V.: Queueing theory. Walter de Gruyter. (2011)
[3] Chesneau, C.: A new family of distributions based on the hypoexponential distribution with fitting reliability data. Statistica. 78(2), 127–147 (2018)
[4] Forbes, C., Evans, M., Hastings, N., Peacock, B.: Statistical distributions. John Wiley & Sons (2011)
[5] Gómez, Y.M., Bolfarine, H., Gómez, H.W.: A new extension of the exponential distribution. Revista Colombiana de Estadística. 37(1), 25–34 (2014)
[6] Gupta, R.D., Kundu, D.: Exponentiated exponential family: An alternative to Gamma and Weibull distribution'. Biometrical Journal. 43(1), 117–130 (2001)
[7] Jasiulewicz, H., Kordecki, W.: Convolutions of Erlang and of Pascal distributions with applications to reliability. Demonstr. Math. 36(1), 231–238 (2003)
[8] Kadri, T., Smaili, K., Kadry, S. In: Markov modeling for reliability analysis using Hypoexponential distribution. In Numerical Methods for Reliability and Safety Assessment, pp. 599–620. Springer, Cham (2015)
[9] Mudholkar, G.S., Srivastava, D.K.: Exponentiated Weibull family for analysing bathtub failure data. IEEE Trans. Reliability. 42, 299–302 (1993)
[10] Smaili, K., Kadri, T., Kadry, S.: A modified-form expressions for the hypoexponential distribution. British Journal of Mathematics & Computer Science. 4(3), 322–332 (2014)
[11] Smaili, K., Kadri, T., Kadry, S.: Finding the PDF of the hypoexponential random variable using the Kad matrix similar to the general Vandermonde matrix. Comm. Statist. Theory Methods. 45(5), 1542–1549 (2016)
[12] Trivedi, K.S.: Probability & statistics with reliability, queuing and computer science applications. John Wiley & Sons (2008)
[13] Yousof, H.M., Rasekhi, M., Afify, A.Z., Alizadeh, M., Ghosh, I. and Hamedani, G.G.: The Beta Weibull-G Family of Distributions: Theory, Characterizations and Applications. Pakistan Journal of Statistics. 33, 95–116 (2017)