RESEARCH ARTICLE

Generalized X-Exponential Bathtub Shaped Failure Rate Distribution

V. M. Chacko[1] · K. S. Deepthi[1]

Abstract

Generalized Lindley, Generalized Gamma, Exponentiated Weibull and X-Exponential distributions are proposed for modelling lifetime data having bathtub shaped failure rate model. These distributions have several desirable properties. In this paper, we introduce a new distribution which generalizes the X-Exponential distribution. Several properties of the distribution are derived including the hazard rate function, moments, moment generating function etc. Moreover, we discussed the maximum likelihood estimation of this distribution. The usefulness of the new distribution is illustrated by means of two real data set. We hope that the new distribution will serve as an alternative model to bathtub models available in the literature.

Keywords Bathtub shaped failure rate · Moments · Maximum likelihood estimation · X-Exponential distribution

1 Introduction

In many applied sciences such as medicine, engineering and finance, amongst others, modeling and analyzing lifetime data are crucial. Several lifetime distributions have been used to model such kinds of data. For instance, the Exponential, Weibull, Gamma, Rayleigh distributions and their generalizations (Gupta and Kundu 1999; Nadarajah and Gupta 2007; Pal et al. 2006; Sarhana and Kundu 2009). Each distribution has its own characteristics due specifically to the shape of the failure rate function which may be only monotonically decreasing or increasing or constant in its behavior (Lai et al. 2001). The Weibull distribution is more popular than Gamma and Lognormal distribution because the survival function of the latter cannot be expressed in closed

✉ V. M. Chacko
  chackovm@gmail.com

  K. S. Deepthi
  dipthiks@gmail.com

[1] Department of Statistics, St.Thomas College (Autonomous), Thrissur, Kerala, India

forms. The Weibull distribution has closed form survival and hazard rate functions (Murthy et al. 2004). Generalized Lindley distribution has many properties, which exhibit bathtub shape for its failure rate function. Nadarajah et al. (2011) introduced Generalized Lindley distribution and discussed its various properties and applications.

Here we consider a new distribution having bathtub shaped failure rate function, which generalizes the X-Exponential distribution. The X-Exponential distribution was introduced by Chacko (2016). In order to get more flexibility in its failure rate, a small change in exponential part of X-Exponential distribution is made, the resulting distribution is named as Generalized X-Exponential. The distribution can be considered as distribution of a series system having distribution $F(x) = 1 - (1 + \lambda x^2)e^{-\lambda(x^2+x)}$, $x > 0$, $\lambda > 0$, for its components in which, a parabolic function $ax^2 + bx + c$ is used as exponent with $a = b = \lambda$, $c = 0$ whereas in X-Exponential distribution, the parameter values were $a = \lambda$, $b = 0$, $c = 0$.

In this paper, we introduce a new distribution which generalizes the X-Exponential distribution. Section 2 discussed the definition of the Generalization of X-Exponential distribution. Section 3 discussed the statistical behaviors of the distribution. Section 4 discussed the distribution of maximum and minimum. The maximum likelihood estimation of the parameters is determined in Sect. 5. Real data sets are analyzed in Sect. 6 and the results are compared with existing distributions. Conclusions are given in Sect. 7.

2 Generalized X-Exponential Distribution

Consider a life time random variable X has a cumulative distribution function (cdf) with parameter α and λ,

$$F(x) = \left(1 - \left(1 + \lambda x^2\right)e^{-\lambda(x^2+x)}\right)^{\alpha}, \quad x > 0, \ \lambda > 0, \ \alpha > 0. \tag{2.1}$$

Clearly $F(0) = 0$, $F(\infty) = 1$, F is non-decreasing and right continuous. More over F is absolutely continuous. Then the probability density function (pdf) with scale parameter λ is given by

$$f(x) = \alpha e^{-\lambda(x^2+x)}\left(\lambda\left(1 + \lambda x^2\right)(2x + 1) - 2\lambda x\right)\left(1 - \left(1 + \lambda x^2\right)e^{-\lambda(x^2+x)}\right)^{\alpha-1},$$
$$x > 0, \ \alpha > 0, \ \lambda > 0. \tag{2.2}$$

Here α and λ are shape and scale parameters. It is positively skewed distribution. The distribution with pdf of the form (2.2) is named as Generalized X-Exponential distribution with parameters α and λ and will be denoted by GXED(α, λ).

Failure rate function of GXED distribution is

$$h(x) = \frac{\alpha e^{-\lambda(x^2+x)}\left(\lambda\left(1 + \lambda x^2\right)(2x + 1) - 2\lambda x\right)\left(1 - \left(1 + \lambda x^2\right)e^{-\lambda(x^2+x)}\right)^{\alpha-1}}{1 - \left(1 - \left(1 + \lambda x^2\right)e^{-\lambda(x^2+x)}\right)^{\alpha}},$$
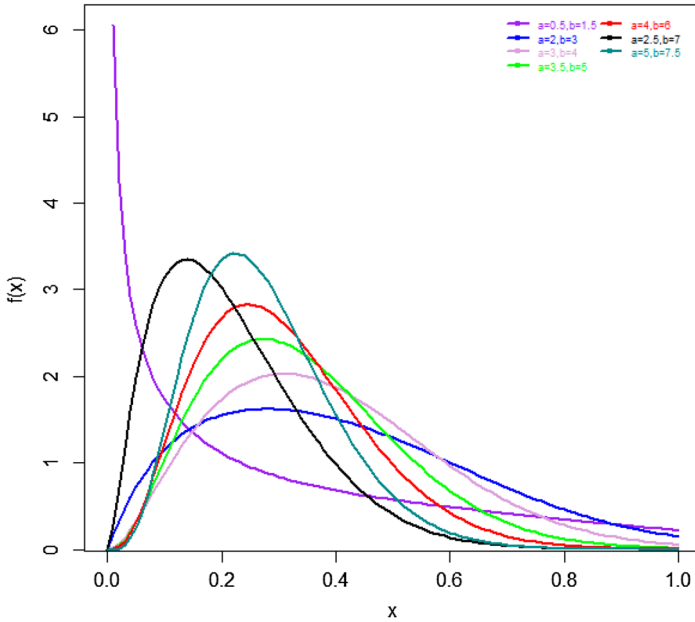$$x > 0, \ \alpha > 0, \ \lambda > 0. \tag{2.3}$$

Fig. 1 Probability density function of GXED for values of parameters $\alpha = 0.5, 2, 3, 3.5, 4, 2.5, 5$ and $\lambda = 1.5, 3, 4, 5, 6, 7, 7.5$ with color shapes purple, blue, plum, green, red, black and dark cyan, respectively

Considering the behavior near the change point x_0, $0 < x_0$ and if $\frac{d}{dx}h(x_0) = 0$

(i) If $0 < \alpha < \frac{1}{2}$, and $0 < \lambda < 1$, then $\frac{d}{dx}h(x) < 0$ when $0 < x < x_0$ and $\frac{d}{dx}h(x) > 0$ when $x > x_0$, $\frac{d^2}{dx^2}h(x) > 0$ for $x = x_0$.

(ii) If $0 < \alpha < \frac{1}{2}$, and $\lambda > 1$, then $\frac{d}{dx}h(x) < 0$ when $0 < x < x_0$, $\frac{d}{dx}h(x) > 0$ when $x > x_0$, $\frac{d^2}{dx^2}h(x) > 0$ for $x = x_0$.

(iii) If $\frac{1}{2} < \alpha < 1$, and $0 < \lambda < 1$, then $\frac{d}{dx}h(x) < 0$ when $0 < x < x_0$ and $\frac{d}{dx}h(x) > 0$ when $x > x_0$, $\frac{d^2}{dx^2}h(x) > 0$ for $x = x_0$.

(iv) If $\frac{1}{2} < \alpha < 1$, and $\lambda > 1$, then $\frac{d}{dx}h(x) < 0$ when $0 < x < x_0$ and $\frac{d}{dx}h(x) > 0$ when $x > x_0$, $\frac{d^2}{dx^2}h(x) > 0$ for $x = x_0$.

(v) If $\alpha > 1$, and $\lambda > 1$, then $\frac{d}{dx}h(x) > 0$ for $x > 0$

The shape of Eq. (2.3) appears monotonically decreasing or to initially decrease and then increase, a bathtub shape if $\alpha < 1$. The proposed distribution allows for monotonically decreasing, monotonically increasing and bathtub shapes for its hazard rate function. As α decreases from 1 to 0, the graph shift above whereas if λ increases from 1 to ∞ the shape of the graph concentrate near to 0. It is the distribution of the failure of a series system with independent components (Figs. 1, 2, 3).
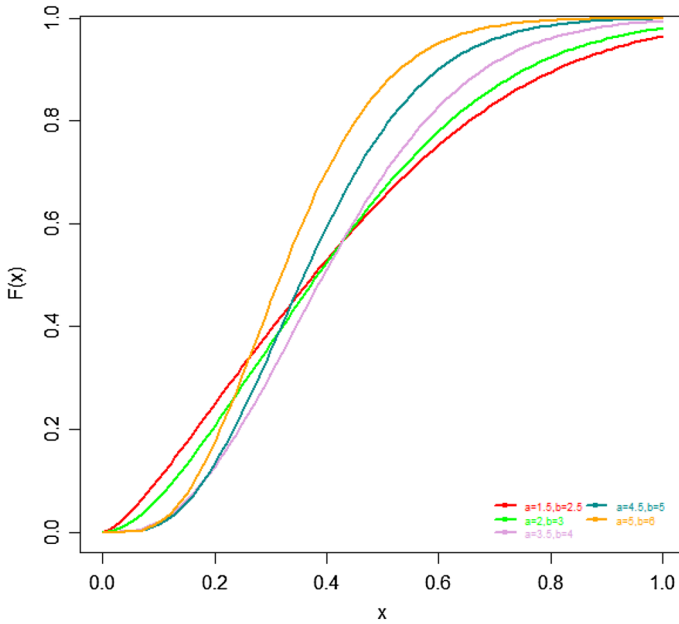
Fig. 2 Cumulative distribution function of GXED for values of parameters $\alpha = 1.5, 2, 3.5, 4.5, 5$ and $\lambda = 2.5, 3, 4, 5, 6$ with color shapes red, green, plum, dark cyan and orange respectively
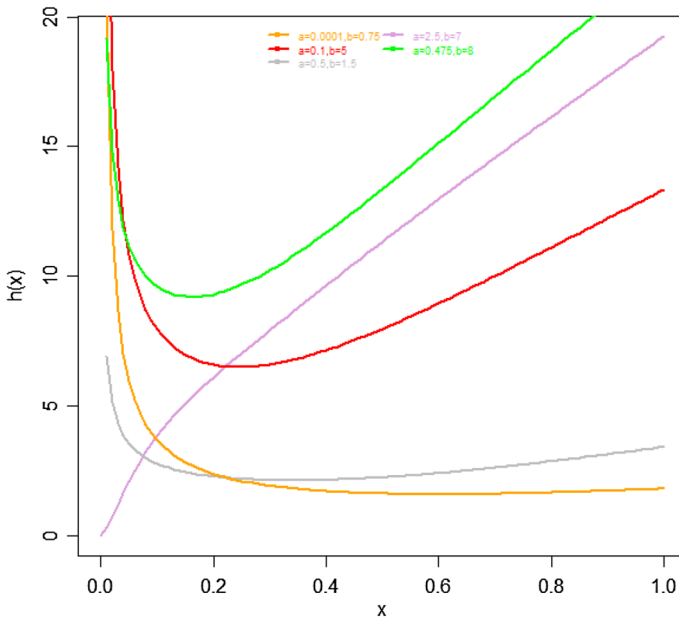

Fig. 3 Failure rate function GXED for values of parameters $\alpha = 0.0001, 0.1, 0.5, 2.5, 0.475$ and $\lambda = 0.75, 5, 1.5, 7, 8$ with color shapes orange, red, grey, plum and green, respectively

3 Moments

Calculating moments of X requires the following lemma.

Lemma 2.1 *For* $\alpha,\ \lambda > 0,\ x > 0,\ \mathrm{K}(\alpha,\lambda,c) = \int_0^\infty x^c \left(1 - \left(1 + \lambda x^2\right)e^{-\lambda\left(x^2 + x\right)}\right)^{\alpha - 1} e^{-\lambda\left(x^2 + x\right)} dx$
Then,

$$\mathrm{K}(\alpha,\lambda,c) = \sum_{i=0}^{\alpha-1}\sum_{j=0}^{i}\binom{\alpha-1}{i}\binom{i}{j}(-1)^i \lambda^j \int_0^\infty x^{2j+c} e^{-\lambda\left(x^2+x\right)} dx$$

Proof We know that $(1 - z)^{\alpha-1} = \sum_{i=0}^{\alpha-1}\binom{\alpha-1}{i}(-1)^i z^i$

Therefore

$$\mathrm{K}(\alpha,\lambda,c) = \int_0^\infty x^c \left(1 - \left(1 + \lambda x^2\right)e^{-\lambda\left(x^2+x\right)}\right)^{\alpha-1} e^{-\lambda\left(x^2+x\right)} dx$$

$$= \int_0^\infty x^c \sum_{i=0}^{\alpha-1}\binom{\alpha-1}{i}(-1)^i \left[\left(1 + \lambda x^2\right)e^{-\lambda\left(x^2+x\right)}\right]^i e^{-\lambda\left(x^2+x\right)} dx$$

$$= \sum_{i=0}^{\alpha-1}\binom{\alpha-1}{i}(-1)^i \int_0^\infty x^c \left[\left(1 + \lambda x^2\right)\right]^i e^{-(i+1)\lambda\left(x^2+x\right)} dx$$

$$= \sum_{i=0}^{\alpha-1}\binom{\alpha-1}{i}(-1)^i \int_0^\infty x^c \sum_{j=0}^{i}\binom{i}{j}\left(\lambda x^2\right)^j e^{-(i+1)\lambda\left(x^2+x\right)} dx$$

$$\mathrm{K}(\alpha,\lambda,c) = \sum_{i=0}^{\alpha-1}\sum_{j=0}^{i}\binom{\alpha-1}{i}\binom{i}{j}(-1)^i \lambda^j \int_0^\infty x^{2j+c} e^{-\lambda\left(x^2+x\right)} dx$$

The result of the Lemma follows by the definition of the gamma function. The moments are

$$E(X) = \alpha\lambda\mathrm{K}(\alpha,\lambda,1) + 2\alpha\lambda^2\mathrm{K}(\alpha,\lambda,4) + \alpha\lambda^2\mathrm{K}(\alpha,\lambda,3)$$
$$E(X^n) = \alpha\lambda\mathrm{K}(\alpha,\lambda,n) + 2\alpha\lambda^2\mathrm{K}(\alpha,\lambda,n+3) + \alpha\lambda^2\mathrm{K}(\alpha,\lambda,n+2)$$

□

3.1 Moment Generating Function

Moment generating function can be obtained from following formula

$$M_X(t) = \int_0^\infty e^{tx} \alpha \left(\lambda + \lambda x^2 + 2\lambda^2 x^3\right) \left[1 - \left(1 + \lambda x^2\right) e^{-\lambda(x^2+x)}\right]^{\alpha-1} e^{-\lambda(x^2+x)} dx$$

$$M_X(t) = \int_0^\infty \alpha \left(\lambda + \lambda x^2 + 2\lambda^2 x^3\right) \left[1 - \left(1 + \lambda x^2\right) e^{-\lambda(x^2+x)}\right]^{\alpha-1} e^{-\lambda(x^2+x)+tx} dx$$

3.2 Characteristic Function

Characteristic function can be obtained from following formula

$$\varphi_X(t) = \int_0^\infty e^{itx} \alpha \left(\lambda + \lambda x^2 + 2\lambda^2 x^3\right) \left[1 - \left(1 + \lambda x^2\right) e^{-\lambda(x^2+x)}\right]^{\alpha-1} e^{-\lambda(x^2+x)} dx$$

$$\varphi_X(t) = \int_0^\infty \alpha \left(\lambda + \lambda x^2 + 2\lambda^2 x^3\right) \left[1 - \left(1 + \lambda x^2\right) e^{-\lambda(x^2+x)}\right]^{\alpha-1} e^{-\lambda(x^2+x)+itx} dx$$

3.3 Mean Deviation About Mean

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. Mean deviation about the mean defined by

$$MD(\text{Mean}) = 2\mu F(\mu) - 2\mu + 2 \int_\mu^\infty x f(x) dx$$

$$MD(Mean) = 2\,\mu F(\mu) - 2\mu$$
$$+ 2 \left(2\alpha\lambda^2 L(\alpha,\,\lambda,\,4,\,\mu) + \alpha\lambda^2 L(\alpha,\,\lambda,\,3,\,\mu) + \alpha\lambda L(\alpha,\,\lambda,\,2,\,\mu)\right)$$

where

$$L(\alpha,\,\lambda,\,c,\,\mu) = \int_\mu^\infty x^c \left[1 - \left(1 + \lambda x^2\right) e^{-\lambda(x^2+x)}\right]^{\alpha-1} e^{-\lambda(x^2+x)} dx$$

$$= \sum_{i=0}^{\alpha-1} C_i^{\alpha-1} (-1)^i \lambda^j \left(\int_\mu^\infty x^{2j+c+1} e^{-(j+1)\lambda(x^2+x)} dx\right).$$

Mean deviation about the Median defined by

$$MD(\text{Median}) = -M + 2 \int_M^\infty x f(x) dx$$

$$MD(Mean) = -M + 2\Big(2\alpha\lambda^2 L(\alpha,\ \lambda,\ 4,\ M) + \alpha\lambda^2 L(\alpha,\ \lambda,\ 3,\ M) + \alpha\lambda L(\alpha,\ \lambda,\ 2,\ M)\Big)$$

4 Distribution of Maximum and Minimum

Series, Parallel, Series–Parallel and Parallel–Series systems are general system structure of many engineering systems. The theory of order statistics provides a use-full tool for analysing life time data of such systems. Let $X_1, X_2, \ldots, X_n$ be a simple random sample from Generalized X-Exponential with cdf and pdf as in (2.1) and (2.2), respectively. Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denote the order statistics obtained from this sample. The pdf of rth order statistics $X_{(r)}$ is given by,

$$f_{r:n}(x) = \frac{1}{B(r,\ n-r+1)}[F(x;\alpha,\ \lambda)]^{r-1}[1 - F(x;\alpha,\ \lambda)]^{n-r} f(x;\alpha,\ \lambda)$$

where $F(x;\alpha,\ \lambda)$ and $f(x;\alpha,\ \lambda)$ are the cdf and pdf given by (2.1) and (2.2), respectively.

$$f_{r:n}(x) = \frac{1}{B(r,n-r+1)}\Big[\Big(1 - \big(1+\lambda x^2\big)e^{-\lambda(x^2+x)}\Big)^\alpha\Big]^{r-1}\Big[1 - \Big(1 - \big(1+\lambda x^2\big)e^{-\lambda(x^2+x)}\Big)^\alpha\Big]^{n-r}$$
$$\alpha e^{-\lambda(x^2+x)}\big(\lambda\big(1+\lambda x^2\big)(2x+1) - 2\lambda x\big)\Big(1 - \big(1+\lambda x^2\big)e^{-\lambda(x^2+x)}\Big)^{\alpha-1} \tag{4.1}$$

Then the pdf of the smallest and largest order statistics $X_{(1)}$ and $X_{(n)}$ are respectively given by:

$$f_1(x) = \frac{1}{B(1,\ n)}\Big[1 - \Big(1 - \big(1+\lambda x^2\big)e^{-\lambda(x^2+x)}\Big)^\alpha\Big]^{n-1}$$
$$\alpha e^{-\lambda(x^2+x)}\Big(\lambda\big(1+\lambda x^2\big)(2x+1) - 2\lambda x\Big)\Big(1 - \big(1+\lambda x^2\big)e^{-\lambda(x^2+x)}\Big)^{\alpha-1}$$
$$f_n(x) = \frac{1}{B(n,\ 1)}\Big[\Big(1 - \big(1+\lambda x^2\big)e^{-\lambda(x^2+x)}\Big)^\alpha\Big]^{n-1}$$
$$\alpha e^{-\lambda(x^2+x)}\Big(\lambda\big(1+\lambda x^2\big)(2x+1) - 2\lambda x\Big)\Big(1 - \big(1+\lambda x^2\big)e^{-\lambda(x^2+x)}\Big)^{\alpha-1}$$

The cdf of $X_{(r)}$ is given by,

$$F_{r:n}(x) = \sum_{j=r}^{n}\binom{n}{j}F^j(x)[1-F(x)]^{n-j}$$
$$F_{r:n}(x) = \sum_{j=r}^{n}\binom{n}{j}\Big[\Big(1 - \big(1+\lambda x^2\big)e^{-\lambda(x^2+x)}\Big)^\alpha\Big]^j\Big[1 - \Big(1 - \big(1+\lambda x^2\big)e^{-\lambda(x^2+x)}\Big)^\alpha\Big]^{n-j}$$
$$\tag{4.2}$$

Then the cdf of the smallest and largest order statistics $X_{(1)}$ and $X_{(n)}$ is respectively given by:

$$F_1(x) = 1 - \Big[1 - \Big(1 - \big(1+\lambda x^2\big)e^{-\lambda(x^2+x)}\Big)^\alpha\Big]^n$$

$$F_n(x) = \left[\left(1 - \left(1 + \lambda x^2\right)e^{-\lambda\left(x^2+x\right)}\right)^{\alpha}\right]^n.$$

Reliability of a series system having n components with independent and identically distributed (iid) GXED distribution is

$$R(x) = \left(\left(1 - \left(1 + \lambda x^2\right)e^{-\lambda\left(x^2+x\right)}\right)^{\alpha}\right)^n$$

Reliability of a parallel system having n components with iid GXED distribution is

$$R(x) = 1 - \left(\left(1 - \left(1 + \lambda x^2\right)e^{-\lambda\left(x^2+x\right)}\right)^{\alpha}\right)^n$$

Both the reliability functions can be used in various reliability calculations.

5 Estimation

In this section, point estimation of the unknown parameters of the GXED by using the method of moments and method of maximum likelihood based on a complete sample data, is explained.

Let $X_1, X_2, \ldots, X_n$ are random sample taken from GXED. Let $m_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$, $m_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$. Equating sample moments to population moments we get moment estimators for parameters.

$$m_1 = 2\alpha\lambda^2 K(\alpha, \lambda, 4) + \alpha\lambda^2 K(\alpha, \lambda, 3) + \alpha\lambda K(\alpha, \lambda, 1)$$
$$m_2 = 2\alpha\lambda^2 K(\alpha, \lambda, 5) + \alpha\lambda^2 K(\alpha, \lambda, 4) + \alpha\lambda K(\alpha, \lambda, 2)$$

The solution of these equations are moment estimators.
To find maximum likelihood estimator, consider likelihood function as,

$$L(\alpha, \lambda) = \prod_{i=1}^{n} f(x_i)$$

$$L(\alpha, \lambda) = \prod_{i=1}^{n} \alpha e^{-\lambda\left(x_i^2+x_i\right)}\left(\lambda\left(1 + \lambda x_i^2\right)(2x_i + 1) - 2\lambda x_i\right)\left(1 - \left(1 + \lambda x_i^2\right)e^{\left(-\lambda\left(x_i^2+x_i\right)\right)}\right)^{\alpha-1}$$

$$L(\alpha, \lambda) = (\alpha)^n e^{-\lambda\sum_{i=1}^{n}\left(x_i^2+x_i\right)} \prod_{i=1}^{n}\left(\lambda\left(1 + \lambda x_i^2\right)(2x_i + 1) - 2\lambda x_i\right)\prod_{i=1}^{n}\left(1 - \left(1 + \lambda x_i^2\right)e^{-\lambda\left(x_i^2+x_i\right)}\right)^{\alpha-1}$$

$$\log L(\alpha, \lambda) = n\log\alpha - \lambda\sum_{i=1}^{n}\left(x_i^2 + x_i\right) + \sum_{i=1}^{n}\log\left(\lambda\left(1 + \lambda x_i^2\right)(2x_i + 1) - 2\lambda x_i\right)$$

$$+ (\alpha - 1)\sum_{i=1}^{n}\log\left(1 - \left(1 + \lambda x_i^2\right)e^{-\lambda\left(x_i^2+x_i\right)}\right)$$

The first partial derivatives of the log-likelihood function with respect to the two-parameters are

$$\frac{\partial}{\partial \alpha} \log L(\alpha, \lambda) = \frac{n}{\alpha} + \sum_{i=1}^{n} \log\left(1 - \left(1 + \lambda x_i^2\right)e^{-\lambda\left(x_i^2 + x_i\right)}\right)$$

$$\frac{\partial}{\partial \alpha} \log L(\alpha, \lambda) = 0 \text{ implies } \hat{\alpha} = -\frac{1}{n}\sum_{i=1}^{n} \log\left(1 - \left(1 + \lambda x_i^2\right)e^{-\lambda\left(x_i^2 + x_i\right)}\right) \quad (5.1)$$

and

$$\frac{\partial}{\partial \lambda} \log L(\alpha, \lambda) = -\sum_{i=1}^{n}\left(x_i^2 + x_i\right) + \sum_{i=1}^{n} \frac{\left((2x_i + 1)\left(1 + 2\lambda x_i^2\right) - 2x_i\right)}{\left(\lambda\left(1 + \lambda x_i^2\right)(2x_i + 1) - 2\lambda x_i\right)}$$

$$+ (\alpha - 1)\sum_{i=1}^{n} \frac{\left(1 + \lambda x_i^2\right)e^{-\lambda\left(x_i^2 + x_i\right)}\left(x_i^2 + x_i\right) - e^{-\lambda\left(x_i^2 + x_i\right)}x_i^2}{\left(1 - \left(1 + \lambda x_i^2\right)e^{-\lambda\left(x_i^2 + x_i\right)}\right)}$$

$$\frac{\partial}{\partial \lambda} \log L(\alpha, \lambda) = 0 \text{ implies}$$

$$\sum_{i=1}^{n}\left(x_i^2 + x_i\right) = \sum_{i=1}^{n} \frac{\left((2x_i + 1)\left(1 + 2\lambda x_i^2\right) - 2x_i\right)}{\left(\lambda\left(1 + \lambda x_i^2\right)(2x_i + 1) - 2\lambda x_i\right)}$$

$$+ (\alpha - 1)\sum_{i=1}^{n} \frac{\left(1 + \lambda x_i^2\right)e^{-\lambda\left(x_i^2 + x_i\right)}\left(x_i^2 + x_i\right) - e^{-\lambda\left(x_i^2 + x_i\right)}x_i^2}{\left(1 - \left(1 + \lambda x_i^2\right)e^{-\lambda\left(x_i^2 + x_i\right)}\right)} \quad (5.2)$$

Solving this system in α and λ gives the maximum likelihood estimates (MLE) of α and λ. It can be obtain estimates using R software by numerical methods. The initial values have been chosen arbitrarily. Parameter estimation is done using the non-linear method in R software.

6 Asymptotic Confidence bounds

In this section, we derive the asymptotic confidence intervals of these parameters when $\alpha > 0$ and $\lambda > 0$, since the MLEs of the unknown parameters $\alpha > 0$ and $\lambda > 0$ cannot be obtained in closed forms, by using variance covariance matrix I^{-1}, where I^{-1} is the inverse of the observed information matrix which is defined as follows

$$I^{-1} = \begin{pmatrix} -\frac{\partial^2 L}{\partial \alpha^2} & -\frac{\partial^2 L}{\partial \alpha \partial \lambda} \\ -\frac{\partial^2 L}{\partial \lambda \partial \alpha} & -\frac{\partial^2 L}{\partial \lambda^2} \end{pmatrix}$$

$$= \begin{pmatrix} var(\hat{\alpha}) & cov(\hat{\alpha}, \hat{\lambda}) \\ cov(\hat{\lambda}, \hat{\alpha}) & var(\hat{\lambda}) \end{pmatrix}$$

The second partial derivatives are as follows

$$\frac{\partial^2 L}{\partial \alpha^2} = -\frac{n}{\alpha^2}$$

$$\frac{\partial^2 L}{\partial \lambda^2} = \sum_{i=1}^{n} \frac{\left(\lambda(1 + \lambda x_i^2)(2x_i + 1) - 2\lambda x_i\right)\left[2x_i^2(2x_i + 1)\right] - \left((2x_i + 1)(1 + 2\lambda x_i^2) - 2x_i\right)^2}{\left(\lambda(1 + \lambda x_i^2)(2x_i + 1) - 2\lambda x_i\right)^2}$$

$$+ \sum_{i=1}^{n} \frac{\left(1 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)}\right)\left[\left(e^{-\lambda(x_i^2 + x_i)}x_i^2 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)}(x_i^2 + x_i)\right) + e^{-\lambda(x_i^2 + x_i)}(x_i^2 + x_i)\right]}{\left(1 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)}\right)^2}$$

$$- \sum_{i=1}^{n} \frac{\left((1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)}(x_i^2 + x_i) - e^{-\lambda(x_i^2 + x_i)}x_i^2\right)^2}{\left(1 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)}\right)^2}$$

$$\frac{\partial^2 L}{\partial \alpha \partial \lambda} = \sum_{i=1}^{n} \frac{(1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)}(x_i^2 + x_i) - e^{-\lambda(x_i^2 + x_i)}x_i^2}{\left(1 - (1 + \lambda x_i^2)e^{-\lambda(x_i^2 + x_i)}\right)}$$

We can derive the $(1 - \delta)100\%$ confidence intervals of the parameters α and λ by using variance matrix as in the following forms

$$\hat{\alpha} \pm Z_{\frac{\delta}{2}}\sqrt{var(\hat{\alpha})} \text{ and } \hat{\lambda} \pm Z_{\frac{\delta}{2}}\sqrt{var(\hat{\lambda})}$$

where $Z_{\frac{\delta}{2}}$ is the upper $\left(\frac{\delta}{2}\right)$th percentile of the standard normal distribution.

7 Simulation

We assess the performance of the maximum likelihood estimators given by (5.1) and (5.2) with respect to sample size n. The assessment was based on a simulation study:

(i)   Generate hundred samples of size n from (2.2). The inversion method was used to generate samples, i.e., values of the GXED distribution random variable are generated using

$$\left(1 + \lambda x^2\right)e^{-\lambda(x^2 + x)} = 1 - U^{\frac{1}{\alpha}}$$

where $U \sim U(0, 1)$ is a uniform variate on the unit interval.

(ii)   Compute the maximum likelihood estimates for the hundred samples, say (α_i, λ_i) for $i = 1, 2, \ldots, 100$.

(iii)   Compute the biases and mean squared errors using $bias_h(n) = \frac{1}{100}\sum_{i=1}^{100}\left(\hat{h}_i - h\right)$ and $MSE_h(n) = \frac{1}{100}\sum_{i=1}^{100}\left(\hat{h}_i - h\right)^2$ for $h = (\alpha, \lambda)$. We repeated these steps for n = 10, 20, …, 100 with different values of parameters, for computing $bias_h(n)$ and $MSE_h(n)$ for n = 10, 20, …, 100.

	α		λ	
	Bias	MSE	Bias	MSE
$(\alpha = 0.5, \lambda = 0.001)$N				
10	0.022574	0.005095843	0.0001423755	2.027078×10^{-07}
50	0.0028631	0.0004098582	3.40494×10^{-06}	5.796808×10^{-10}
100	0.0017267	0.0002981372	1.29409×10^{-06}	1.674669×10^{-10}
$(\alpha = 1, \lambda = 0.5)$ N				
10	-0.00351507	0.0001235572	-0.00721473	0.0005205233
50	0.001504456	0.0001131694	0.001179188	6.952422×10^{-05}
100	0.001201623	0.0001443898	0.000886022	7.85035×10^{-05}
$(\alpha = 1.5, \lambda = 1)$ N				
10	0.03376	0.011398	-0.00575	0.000331
50	0.002923	0.000427	-0.000568	1.61136×10^{-05}
100	0.0005311	2.820577×10^{-05}	0.0001264	1.59848×10^{-06}

8 Application

In this section, we present the analysis of a real data set using the GXED(α, λ) model and compare it with the other fitted model Generalized Lindley distribution (GLD) (Nadarajah et al. 2011) using Akaike information criterion (AIC), Bayesian information criterion (BIC) and Kolmogorov–Smirnov (K–S) statistic. We considered the Survival data for psychiatric inpatients (Klein and Moesch Berger 1997) to estimate the parameter values. The data are given below.

1 1 2 22 30 28 32 11 14 36 31 33 33 37 35 25 31 22 26 24 35 34 30 35 40 39

Table 1 provides the parameter estimates, standard errors obtained by inverting the observed information matrix and log-likelihood values. Table 2 provides values of AIC, BIC, and p values based on the KS statistic. The corresponding probability plots are shown in Fig. 4. We can see that the GXED distribution gives the smallest AIC

Table 1 MLEs of parameters, SE and Log-likelihood	Model	ML estimates	Standard error	Log L
	GXED	$\hat{\alpha} = 0.6967$	0.03312696	-99.36048
		$\hat{\lambda} = 0.00184$	7.852131×10^{-5}	
	GLD	$\hat{\alpha} = 1.0691$	0.05636717	-107.6572
		$\hat{\lambda} = 0.07547$	0.002741764	

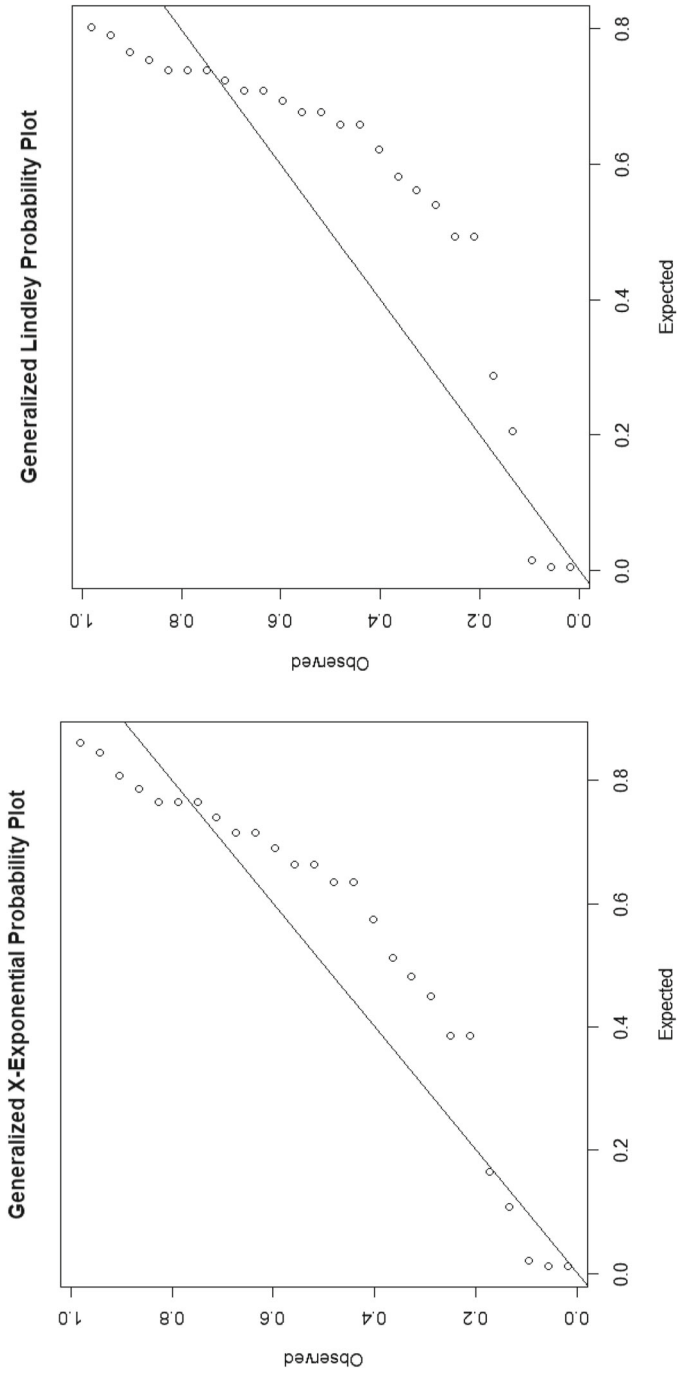| Table 2 AIC, BIC, KS Statistic and p value of the model	Model	AIC	BIC	KS Statistic	p value
	GXED	202.721	205.2371	0.21127	0.1962
	GLD	219.3145	221.8307	0.30109	0.01794

Fig. 4 Probability plot for GXED and GLD (Data set 1)

Table 3 MLEs of parameters, SE and Log-likelihood	Model	ML estimates	Standard error	Log L
	GXED	$\hat{\alpha} = 0.3181$	0.00715	− 231.6088
		$\hat{\lambda} = 0.000302$	1.045224×10^{-5}	
	GLD	$\hat{\alpha} = 0.4547$	0.01123	− 238.9909
		$\hat{\lambda} = 0.0278$	0.000691	

value, the smallest BIC value, the largest p value based on the KS statistic. Hence, the GXED distribution provides the best fit based on the AIC values, BIC values, p values based on the KS statistic. The Failure rate and probability plots again show that the GXED distribution provides the best fit.

The variance covariance matrix I^{-1} of the MLEs under the Generalized X-Exponential distribution for the data set 1 is computed as $\begin{pmatrix} 2.853228 \times 10^{-2} & 3.981029 \times 10^{-5} \\ 3.981029 \times 10^{-5} & 1.603055 \times 10^{-7} \end{pmatrix}$

Thus, the variances of the MLE of α and λ is $var(\hat{\alpha}) = 2.853228 \times 10^{-2}$ and $var(\hat{\lambda}) = 1.603055 \times 10^{-7}$. Therefore, 95% confidence intervals for α and λ are [0.64168, 0.750658] and [0.001710366, 0.001968678] respectively.

8.1 Data Set 2

The data consist of the lifetimes of 50 devices, Aarset (1987) and it is provided below.

0.1	0.2	1	1	1	1	1	2	3	6	7	11	12	18	18	18	18	18	21	32	36	40	45	46	47
50	55	60	63	63	67	67	67	67	72	75	79	82	82	83	84	84	84	85	85	85	85	85	86	86

The parameter estimates, standard errors and the various measures are given in Table 3. The corresponding probability plots are shown in Fig. 5. We can see again that the GXED distribution gives the smallest AIC value, the smallest BIC value, and largest p value based on the KS statistic, see Table 4. Hence, the GXED distribution again provides the best fit based on the AIC values, BIC values, p values based on the KS statistic. The Failure rate and probability plots again show that the GXED distribution provides the better fit.

The variance covariance matrix I^{-1} of the MLEs under the Generalized X-Exponential distribution for the data set 2 is computed as $\begin{pmatrix} 2.559555 \times 10^{-3} & 1.707983 \times 10^{-6} \\ 1.707983 \times 10^{-6} & 5.462464 \times 10^{-9} \end{pmatrix}$.

Thus, the variances of the MLE of α and λ is $var(\hat{\alpha}) = 2.559555 \times 10^{-3}$ and $var(\hat{\lambda}) = 5.462464 \times 10^{-9}$. Therefore, 95% confidence intervals for α and λ are [0.3063695, 0.3299066] and [0.0002845985, 0.0003189833] respectively.
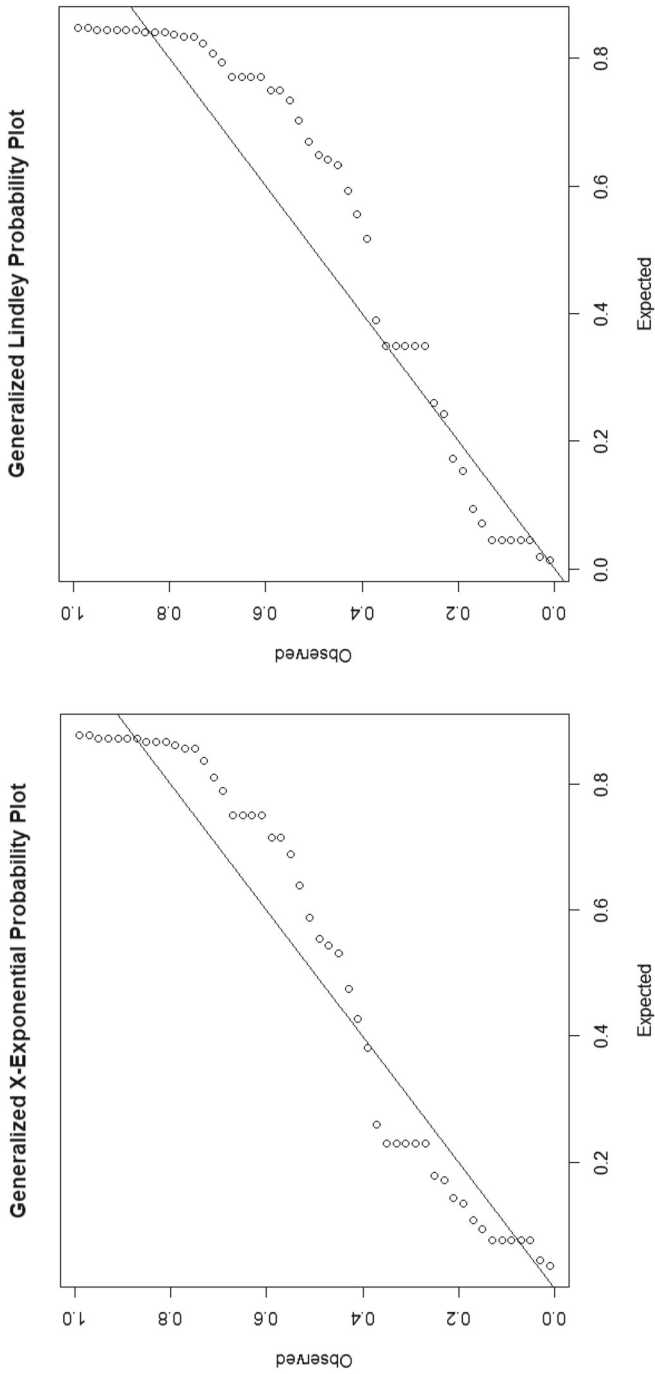
Fig. 5 Probability plot for GXED and GLD (Data set 2)

Table 4 AIC, BIC, KS Statistic and p value of the model

Model	AIC	BIC	KS statistic	p value
GXED	467.2176	471.0417	0.15546	0.1783
GLD	481.9817	485.8058	0.19356	0.04719

9 Conclusion

Here we introduced a new distribution having bathtub shaped failure rate, which generalizes the X-Exponential distribution. Several properties of the distribution are derived including the hazard rate function, moments, moment generating function etc. Also we studied the maximum likelihood estimation of this distribution. The usefulness of the new distribution is illustrated by means of two real data set. The first data set provides GXED distribution has smallest AIC value, the smallest BIC value, the largest p value based on the KS statistic, compared to GLD distribution. And also the second data set provides GXED distribution has smallest AIC value, the smallest BIC value, the largest p value based on the KS statistic, compared to GLD distribution. It shows that the proposed distribution is a better alternative among bathtub shaped failure rate models.

References

Aarset MV (1987) How to identify bathtub hazard rate. IEEE Trans Reliab R-36:106–108

Chacko VM (2016) X-Exponential bathtub failure rate model. Reliab Theor Appl 43(4):55–66

Gupta RD, Kundu D (1999) Generalized exponential distributions. Aust N Z J Stat 41:173–188

Klein JP, Moesch Berger ML (1997) Survival analysis techniques for censored and truncated data. Springer, New York

Lai CD, Xie M, Murthy DNP (2001) Bathtub shaped failure rate distributions. In: Balakrishnan N, Rao CR (eds) Handbook in reliability, vol 20. Elsevier, North-Holland, pp 69–104

Murthy DNP, Xie M, Jiang R (2004) Weibull models. Wiley, New York

Nadarajah S, Gupta AK (2007) The exponentiated gamma distribution with application to drought data. CSA Bull 59:233–234

Nadarajah S, Bakouch HS, Tahmasbi R (2011) A generalized Lindley distribution. Sankhya B 73:331–359

Pal M, Ali MM, Woo J (2006) Exponentiated Weibull distribution. Statistica, anno LXVI, n. 2

Sarhana AM, Kundu D (2009) Generalized linear failure rate distribution. Commun Stat Theor Methods 38(5):642–660