Journal of King Saud University – Science 34 (2022) 101714

Contents lists available at ScienceDirect

Journal of King Saud University – Science

journal homepage: www.sciencedirect.com

Original article

A new flexible extension of the Lindley distribution with applications

Adel Alrasheedi [a], Abdulrahman Abouammoh [a], Mohamed Kayid [a,*]

[a] Department of Statistics and Operations Research, College of Science, King Saud University, Riyadh, Saudi Arabia

ARTICLE INFO

Article history:
Received 29 September 2021
Revised 9 November 2021
Accepted 10 November 2021
Available online 17 November 2021

Mathematics Subject Classification::
62N01
62N05

Keywords:
Mixture models
Gamma distribution
Failure rate
Mean residual life
Mean inactivity time

ABSTRACT

A new general family of distributions based on the Lindley model was introduced. Some properties of the proposed model, including moments, quantiles, and order statistics are presented. In addition, some reliability measures of this model were investigated. The parameters were estimated using the moments and the maximum likelihood methods for complete and right-censored data. Then, the behavior of the maximum likelihood estimator was investigated in a simulation study. Finally, the model was applied to the analysis of two data sets to demonstrate its usefulness.

1. Introduction

In recent years, there have been many papers dealing with the Lindley distribution and its applications; see Lindley (1958), Ghitany et al. (2008), Bakouch et al. (2012), Al-Mutairi et al. (2013) and Cakmakyapan and Kadilar (2017). Many authors have introduced some generalizations and/or extensions to the Lindley model by increasing either the number of underlying parameters or the number of mixed density functions, see for example (Ghitany et al., 2011, Al-Babtain et al., 2014, Abouammoh et al., 2015, Cordeiro et al., 2018, Abouammohm et al., 2020). The main objective of these works is to introduce more flexible probability distributions to model different types of lifetime variables in real applications. The Lindley model is defined by its probability density function (PDF) as

$$f(x)=\frac{\theta^2}{1+\theta}(1+x)e^{-\theta x},\quad \theta>0,\ x\geqslant 0,$$

which is a mixture of the gamma distribution with shape parameter 1 and scale parameter $\theta, G(1,\theta)$, and $G(2,\theta)$ with weights $p=\frac{\theta}{\theta+1}$ and $1-p$, respectively. The cumulative distribution function (CDF) of the Lindley model is:

$$F(x)=1-\frac{1+\theta+\theta x}{1+\theta}e^{-\theta x},\quad \theta>0,\ x\geqslant 0.$$

In connection with the generalization of the Lindley distribution, Shanker (2016a,b) introduced the Aradhana distribution with the following PDF and CDF, respectively:

$$f(x)=\frac{\theta^3}{\theta^2+2\theta+2}(1+x)^2e^{-\theta x},\quad \theta>0,\ x\geqslant 0,$$

and

$$F(x)=1-\frac{2+2\theta+2\theta x+\theta^2+2\theta^2x+\theta^2x^2}{\theta^2+2\theta+2}e^{-\theta x},\quad \theta>0,\ x\geqslant 0.$$

This version of the Aradhana model was used to fit the lifetime data of 20 patients receiving an analgesic reported in Gross and Clark (1976) to relief times (in minutes) and a dataset representing

* Corresponding author.
E-mail address: drkayid@ksu.edu.sa (M. Kayid).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksus.2021.101714
1018-3647/

aircraft window thickness cited in Fuller et al. (1994). In addition, Welday and Shanker (2018) have provided a generalization of this model to the two-parameter Aradhana distribution with the PDF:

$$f(x)=\frac{\theta^3}{\theta^2+2\alpha\theta+2\alpha^2}(1+\alpha x)^2e^{-\theta x},\quad \alpha\geqslant 0,\ \theta>0,\ x\geqslant 0, \tag{1}$$

and the corresponding CDF

$$F(x)=1-\frac{2\alpha^2+2\alpha\theta+2\alpha^2\theta x+\theta^2+2\alpha\theta^2x+\alpha^2\theta^2x^2}{\theta^2+2\alpha\theta+2\alpha^2}e^{-\theta x},$$
$$r\quad \alpha\geqslant 0,\ \theta>0,\ x\geqslant 0.$$

Moreover, Shanker and Shukla (2018) introduced the power Aradhana model whose PDF is given by the PDF of $X^{\frac{1}{\alpha}}$, namely

$$f(x)=\frac{\alpha\theta^3}{\theta^2+2\theta+2}x^{\alpha-1}(1+2x^\alpha+x^{2\alpha})e^{-\theta x^\alpha},\quad \alpha\geqslant 0,\ \theta>0,\ x\geqslant 0,$$

and the CDF

$$F(x)=1-\left(1+\frac{\theta x^\alpha(\theta x^\alpha+2\theta+2)}{\theta^2+2\theta+2}\right)e^{-\theta x^\alpha},\quad \alpha\geqslant 0,\ \theta>0,\ x\geqslant 0.$$

They used this model to analyze the tensile strength measured in GPA of 69 carbon fibers tested in tension at a length of 20 mm (Bader and Priest, 1982).

Most data sets are mixtures of multiple populations, and usually no information is available to determine the associated subpopulation of each data point. For example, the lifetime of a device may be recorded without regard to manufacturer or date of production, or some measurements of humans may be reported without regard to geographic location or blood type. When the measured characteristics depend on data that are not available (manufacturer, production date, geographic location, or blood type), the data are said to be mixed. It is not easy to find data sets that are not mixed in some way, since in almost every case some relative covariates are not observed. There are many applications and statistical frameworks in which mixture models occur. For detailed discussions, see Titterington et al. (1985), Lindsay (1995) and Ord (1972).

The above models are mixtures of two or three gamma distributions with different shape parameters. However, in many situations, the data may come from more than two or three subpopulations. For example, a device may be manufactured by more than three factories in a company, etc. Therefore, it is better to use a mixture model with an adjustable number of underlying submodels. The aim of this paper is to investigate such a flexible model, which also generalizes the previous models.

In this paper, we present in Section 2 a new generalized model that incorporates the above and many other models. Section 3 presents the main statistical and reliability properties of this new model. Parameter estimation for complete and right-censored data is discussed in Section 4. In Section 5, a simulation study was conducted to investigate the behavior of the MLE. Finally, the proposed model and some competing models were fitted to two datasets to show their usefulness.

2. The Abouammoh-Alrasheedi model

One can include most of the previously mentioned models in a general family from which many more special cases can be derived and studied. Statistical and reliability properties, estimation of the underlying parameters, and fitting the derived model to real data will make the models available in the literature even richer with more flexible manipulations. We now give the following definition.

Definition 1. The random variable X is said to have Abouammoh-Alrasheedi distribution with parameters (m,α,θ), denoted by $AA(m,\alpha,\theta)$ if its PDF be of the form

$$f(x)=\frac{\theta^m}{\sum_{i=1}^{m}\frac{\Gamma(m)}{\Gamma(i)}\alpha^{m-i}\theta^{i-1}}(1+\theta x)^{m-1}e^{-\theta x},\quad m>0,\ \alpha>0,$$
$$\theta>0,\ x\geqslant 0. \tag{2}$$

One can, without difficulty, verify that this is a PDF, i.e., $\int_0^\infty f(x)dx=1$. For integer m, it can be written as a mixture of gamma distributions, so by Theorem 3 of Atienza et al. (2006), it is identifiable. More precisely

$$f(x)=\sum_{i=1}^{m}w_if_i(x),\quad x\geqslant 0,$$

where

$$w_i=\frac{\frac{1}{\Gamma(i)}\left(\frac{\theta}{\alpha}\right)^i}{\sum_{j=1}^{m}\frac{1}{\Gamma(j)}\left(\frac{\theta}{\alpha}\right)^j}, \tag{3}$$

and

$$f_i(x)=\frac{\theta^{m-i+1}}{\Gamma(m-i+1)}x^{m-i}e^{-\theta x},\quad x>0,$$

is the PDF of the gamma distribution $G(m-i+1,\theta)$.

Lemma 1. *The constant coefficient of the PDF (2) can be written as the following form:*

$$c=\frac{\theta^m}{\sum_{i=1}^{m}\frac{\Gamma(m)}{\Gamma(i)}\alpha^{m-i}\theta^{i-1}}=\frac{e^{-\frac{\theta}{\alpha}}\theta^m}{\alpha^{m-1}\Gamma(\frac{\theta}{\alpha},m)}.$$

Proof. Since $f(x)$ is a PDF, we have

$$\int_0^\infty(1+\alpha x)^{m-1}e^{-\theta x}dx=\frac{1}{c}.$$

On the other hand, by straightforward algebra we have

$$\int_0^\infty(1+\alpha x)^{m-1}e^{-\theta x}dx=\frac{e^{\frac{\theta}{\alpha}}}{\theta^m}\alpha^{m-1}\Gamma\left(\frac{\theta}{\alpha},m\right),$$

which shows the result immediately. □

Some of special cases of the $AA(m,\alpha,\theta)$ are listed in the following.

- For $m=1$ and $\alpha=1$, it reduces to the exponential distribution.
- For $m=2$ and $\alpha=1$ it gives the Lindley distribution.
- For $m=3$ and $\alpha=1, AA(m,\alpha,\theta)$ reduces to Aradhana distribution which is a mixture of $G(1,\theta), G(2,\theta)$ and $G(3,\theta)$ with weights $\frac{\theta^2}{\theta^2+2\theta+2}, \frac{2\theta}{\theta^2+2\theta+2}$ and $\frac{2}{\theta^2+2\theta+2}$, see Shanker (2016a,b).
- For $\alpha=1$, and $m\geqslant 3, AA(m,\alpha,\theta)$ becomes a generalized Aradhana distribution and is also a generalization of the Lindley distribution.

3. Statistical and reliability properties

Now, we derive the main statistical and reliability properties of the proposed distribution. The cumulative distribution function of $AA(m,\alpha,\theta)$ is

$$F(x) = 1 - \frac{\Gamma\left(\frac{\theta}{\alpha} + \theta x, m\right)}{\Gamma\left(\frac{\theta}{\alpha}, m\right)}, \tag{4}$$

where $\Gamma(t,s) = \int_t^\infty y^{s-1}e^{-y}dy$ is the incomplete upper gamma function.

Proposition 1. *The k^{th} moments of $AA(m,\alpha,\theta)$ model is*

$$E\left(X^k\right) = \frac{1}{\Gamma\left(\frac{\theta}{\alpha}, m\right)} \sum_{j=0}^{k} \left(\frac{1}{\theta}\right)^j \left(-\frac{1}{\alpha}\right)^{k-j} \binom{k}{j} \Gamma\left(\frac{\theta}{\alpha}, m+j\right). \tag{5}$$

Proof. The k^{th} moments of a distribution with the reliability function R equals to

$$E\left(X^k\right) = \int_0^\infty k t^{k-1} R(t) dt. \tag{6}$$

Thus, for $AA(m,\alpha,\theta)$ we have

$$E\left(X^k\right) = \frac{1}{\Gamma\left(\frac{\theta}{\alpha}, m\right)} \int_0^\infty k t^{k-1} \Gamma\left(\frac{\theta}{\alpha} + \theta t, m\right) dt. \tag{7}$$

Now, we have

$$\begin{aligned} \int_0^\infty k t^{k-1}\Gamma\left(\tfrac{\theta}{\alpha} + \theta t, m\right)dt &= \int_0^\infty \int_{\frac{\theta}{\alpha}+\theta t}^\infty k t^{k-1} y^{m-1} e^{-y} dy dt \\ &= \int_{\frac{\theta}{\alpha}}^\infty \int_0^{\frac{y}{\theta}-\frac{1}{\alpha}} k t^{k-1} y^{m-1} e^{-y} dt dy \\ &= \int_{\frac{\theta}{\alpha}}^\infty y^{m-1} e^{-y} \left(\tfrac{y}{\theta} - \tfrac{1}{\alpha}\right)^k dy \\ &= \sum_{j=0}^{k} \left(\tfrac{1}{\theta}\right)^j \left(-\tfrac{1}{\alpha}\right)^{k-j} \binom{k}{j} \Gamma\left(\tfrac{\theta}{\alpha}, m+j\right), \end{aligned} \tag{8}$$

and the result follows by (7) and (8). □

By the fact that $AA(m,\alpha,\theta)$ model can be written as a mixture of the corresponding moments of the underlying gamma distributions, we have another representation of the k^{th} moments

$$E\left(X^k\right) = \sum_{i=1}^{m} w_i \frac{\Gamma(k+m-i+1)}{\theta^k \Gamma(m-i+1)}, \tag{9}$$

where w_i is defined in (3).

By straightforward algebra the moment generating function of $AA(m,\alpha,\theta)$ can be simplified as the following form.

$$M(t) = E\left(e^{tX}\right) = \frac{\Gamma\left(\frac{\theta - t}{\alpha}, m\right)}{\Gamma\left(\frac{\theta}{\alpha}, m\right)} \left(\frac{\theta}{\theta - t}\right)^m e^{-\frac{t}{\alpha}}. \tag{10}$$

The quantile function is of the form

$$q(p) = \frac{1}{\theta}\left(\Gamma^{-1}\left(\bar{p}\Gamma\left(\frac{\theta}{\alpha}, m\right), m\right) - \frac{\theta}{\alpha}\right), \quad 0 < p < 1, \tag{11}$$

where $\bar{p} = 1 - p$ and $\Gamma^{-1}(a,m)$ shows the inverse of the incomplete upper gamma function at a.

The Lorenz curve provides a graphical representation of the wealth distribution and is defined to be

$$L(p) = \frac{\int_0^{q(p)} x dF(x)}{\int_0^{q(1)} x dF(x)}.$$

It can be checked easily that for $AA(m,\alpha,\theta)$ it can be simplified as follows.

$$\begin{aligned} L(p) &= \frac{\frac{1}{\theta}\left(\Gamma\left(\frac{\theta}{\alpha}, m+1\right) - \Gamma\left(\frac{\theta}{\alpha} + \theta q(p), m+1\right)\right) - \frac{1}{\alpha}\left(\Gamma\left(\frac{\theta}{\alpha}, m\right) - \Gamma\left(\frac{\theta}{\alpha} + \theta q(p), m\right)\right)}{\frac{1}{\theta}\Gamma\left(\frac{\theta}{\alpha}, m+1\right) - \frac{1}{\alpha}\Gamma\left(\frac{\theta}{\alpha}, m\right)} \\ &= p\,\frac{\alpha\left(m\Gamma\left(\frac{\theta}{\alpha}, m\right) + \left(\frac{\theta}{\alpha}\right)^m e^{-\frac{\theta}{\alpha}} - u^m e^{-u}\right) - \theta\Gamma\left(\frac{\theta}{\alpha}, m\right)}{\alpha\left(m\Gamma\left(\frac{\theta}{\alpha}, m\right) + \left(\frac{\theta}{\alpha}\right)^m e^{-\frac{\theta}{\alpha}}\right) - \theta\Gamma\left(\frac{\theta}{\alpha}, m\right)}, \end{aligned}$$

where $u = \Gamma^{-1}\left(\bar{p}\Gamma\left(\frac{\theta}{\alpha}, m\right), m\right)$.

Let $X_1, X_2, \ldots, X_n$ represents a sample of $AA(m,\alpha,\theta)$. The PDF of k^{th} order statistics, $X_{(k)}$, is

$$f_{k:n}(x) = k\binom{n}{k} \frac{e^{-\frac{\theta}{\alpha}}\theta^m}{\alpha^{m-1}\Gamma\left(\frac{\theta}{\alpha}, m\right)} \left(1 - \frac{\Gamma\left(\frac{\theta}{\alpha} + \theta x, m\right)}{\Gamma\left(\frac{\theta}{\alpha}, m\right)}\right)^{k-1} \left(\frac{\Gamma\left(\frac{\theta}{\alpha} + \theta x, m\right)}{\Gamma\left(\frac{\theta}{\alpha}, m\right)}\right)^{n-k} (1+\theta x)^{m-1} e^{-\theta x}.$$

Thus the PDF of series and parallel systems with such identical components reduces to

$$f_{1:n}(x) = n \frac{e^{-\frac{\theta}{\alpha}}\theta^m}{\alpha^{m-1}\Gamma\left(\frac{\theta}{\alpha}, m\right)} \left(\frac{\Gamma\left(\frac{\theta}{\alpha} + \theta x, m\right)}{\Gamma\left(\frac{\theta}{\alpha}, m\right)}\right)^{n-1} (1+\theta x)^{m-1} e^{-\theta x},$$

and

$$f_{n:n}(x) = n \frac{e^{-\frac{\theta}{\alpha}}\theta^m}{\alpha^{m-1}\Gamma\left(\frac{\theta}{\alpha}, m\right)} \left(1 - \frac{\Gamma\left(\frac{\theta}{\alpha} + \theta x, m\right)}{\Gamma\left(\frac{\theta}{\alpha}, m\right)}\right)^{n-1} (1+\theta x)^{m-1} e^{-\theta x}.$$

respectively.

3.1. Reliability measures

The failure rate function is an important measure in reliability theory and survival analysis. Assuming that an event has not yet occurred by time x, it represents the instantaneous risk of occurrence at time x. For $AA(m,\alpha,\theta)$, the failure rate function is:

$$h(x) = \frac{f(x)}{R(x)} = \frac{e^{-\frac{\theta}{\alpha}}\theta^m (1+\alpha x)^{m-1} e^{-\theta x}}{\alpha^{m-1}\Gamma\left(\frac{\theta}{\alpha} + \theta x, m\right)}, \tag{12}$$

where $R(x) = 1 - F(x)$ shows the reliability function. For a lifetime random variable X, the mean residual life (MRL) function is defined to be

$$m(x) = E(X - x | X \geqslant x) = \frac{\int_x^\infty R(t)dt}{R(x)}, \tag{13}$$

and measures the mean of the remaining lifetime given survival up to time x.

Proposition 2. *The MRL function of $AA(m,\alpha,\theta)$ distribution is of the form*

$$m(x) = \frac{1}{\theta}\left(m + \frac{\left(\frac{\theta}{\alpha} + \theta x\right)^m e^{-\left(\frac{\theta}{\alpha} + \theta x\right)}}{\Gamma\left(\frac{\theta}{\alpha} + \theta x, m\right)}\right) - \frac{1}{\alpha} - x, \quad x \geqslant 0.$$

Proof. We can write

$$\int_x^\infty R(t)dt = \frac{1}{\Gamma\left(\frac{\theta}{\alpha}, m\right)} \int_x^\infty \Gamma\left(\frac{\theta}{\alpha} + \theta t, m\right) dt. \tag{14}$$

The integral in the right side of (14) can be simplified as the following.

$$\begin{aligned} \int_x^\infty \Gamma\left(\tfrac{\theta}{\alpha} + \theta t, m\right)dt &= \int_x^\infty \int_{\frac{\theta}{\alpha}+\theta t}^\infty y^{m-1}e^{-y} dy dt = \int_{\frac{\theta}{\alpha}+\theta x}^\infty \int_x^{\frac{y}{\theta}-\frac{1}{\alpha}} y^{m-1} e^{-y} dt dy \\ &= \int_{\frac{\theta}{\alpha}+\theta x}^\infty \left(\tfrac{y}{\theta} - \tfrac{1}{\alpha} - x\right) y^{m-1} e^{-y} dy \\ &= \tfrac{1}{\theta}\Gamma\left(\tfrac{\theta}{\alpha} + \theta x, m+1\right) - \left(\tfrac{1}{\alpha} + x\right)\Gamma\left(\tfrac{\theta}{\alpha} + \theta x, m\right). \end{aligned} \tag{15}$$

Applying ()()()(13)–(15) the result follows. □

The p-QRL function, denoted by $q_p(x)$, is the conditional pth quantile of the remaining life of an object provided that it is still alive at x, precisely

$$q_p(x) = R^{-1}(\bar{p}R(x)) - x = F^{-1}(1 - \bar{p}R(x)) - x, \quad x > 0,$$

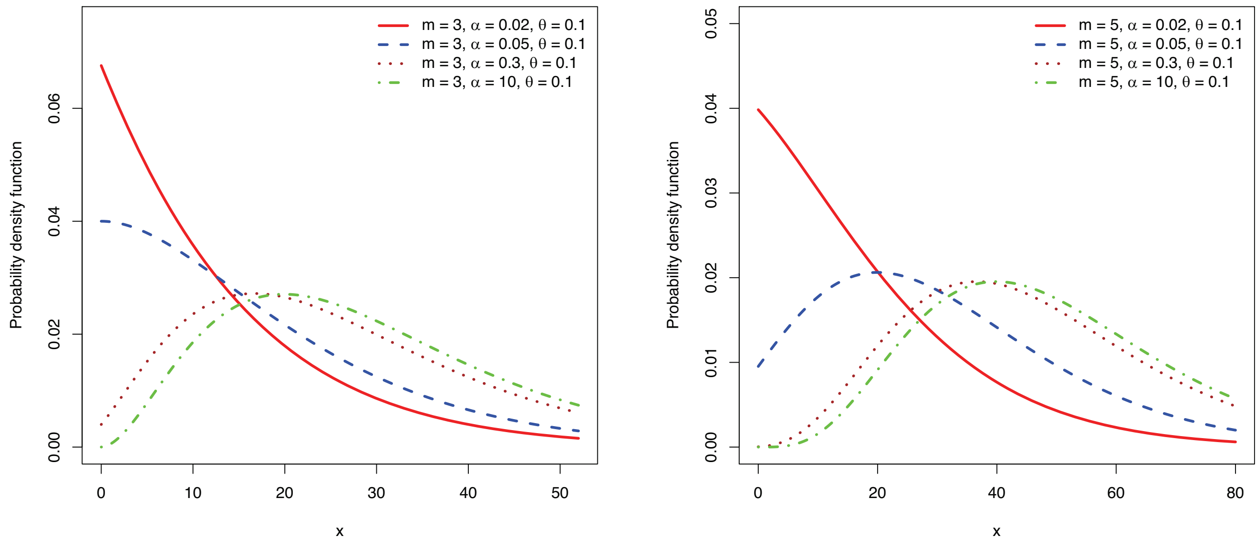

Fig. 1. The PDF of $AA(m,\alpha,\theta)$ for some values of parameters.

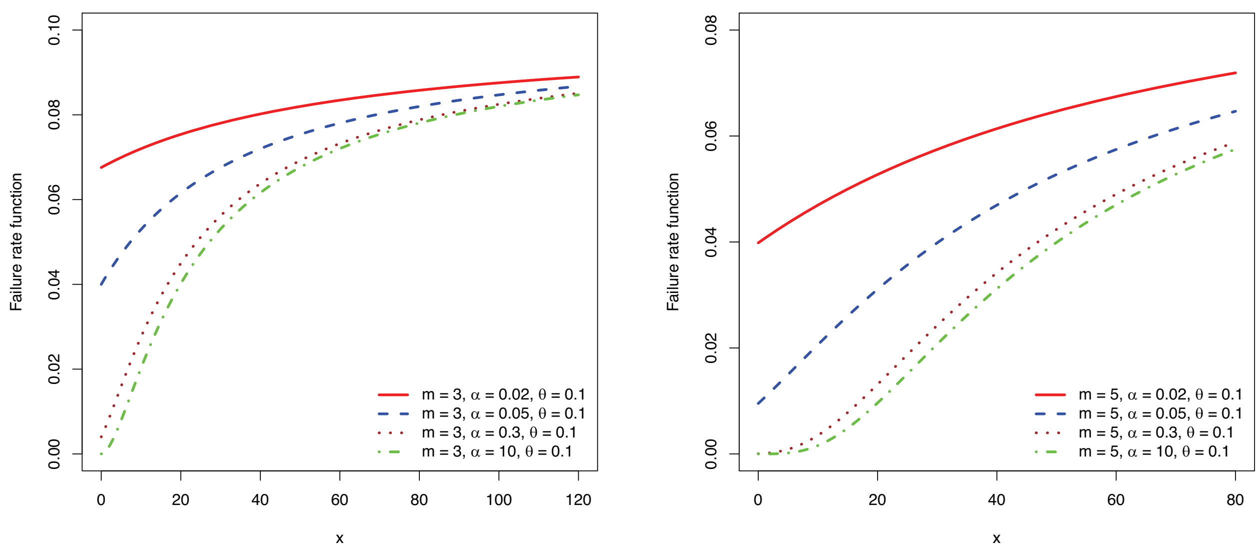

Fig. 2. The failure rate function of $AA(m,\alpha,\theta)$ for some values of parameters.

where $\bar{p} = 1 - p$. For $AA(m,\alpha,\theta)$, by (11) we have

$$q_p(x) = \frac{1}{\theta}\Gamma^{-1}\left(\bar{p}\Gamma\left(\frac{\theta}{\alpha} + \theta x, m\right), m\right) - \frac{1}{\alpha} - x, \quad x > 0.$$

Figs. 1–4 show the PDF, failure rate function, MRL function, and median residual life function, respectively. The failure rate shows an increasing shape and the MRL and median residual life functions show a decreasing shape. The MRL shows larger values than the median residual lifetime, indicating that the conditional residual lifetime is skewed to the right.

The mean inactivity time (MIT) at time x, $m^*(x)$, represents the mean of elapsed time at x given the event has been occurred before x, mathematically,

$$m^*(x) = E(x - X | X < x) = \frac{1}{F(x)}\int_0^x F(z)dz, \quad x \geqslant 0.$$

Proposition 3. *The MIT function of $AA(m,\alpha,\theta)$ distribution is of the form*

$$m^*(x) = \left(x + \frac{1}{\alpha} - \frac{1}{\theta}\right) - \frac{1}{\theta}e^{-\frac{\theta}{\alpha}}\frac{\left(\frac{\theta}{\alpha}\right)^m - \left(\frac{\theta}{\alpha} + \theta x\right)^m e^{-\theta x}}{\Gamma\left(\frac{\theta}{\alpha}, m\right) - \Gamma\left(\frac{\theta}{\alpha} + \theta x, m\right)}, \quad x \geqslant 0. \tag{16}$$

Proof. We have

$$\begin{aligned} \int_0^x F(z)dz &= \tfrac{1}{\Gamma\left(\frac{\theta}{\alpha}, m\right)}\int_0^x \Gamma\left(\tfrac{\theta}{\alpha}, m\right) - \Gamma\left(\tfrac{\theta}{\alpha} + \theta z, m\right)dz \\ &= \tfrac{1}{\Gamma\left(\frac{\theta}{\alpha}, m\right)}\left(x\Gamma\left(\tfrac{\theta}{\alpha}, m\right) - \int_0^x \Gamma\left(\tfrac{\theta}{\alpha} + \theta z, m\right)dz\right). \end{aligned} \tag{17}$$

Now, we can simplify the integral in the last expression as follows.

$$\begin{aligned} &\int_0^x \Gamma\left(\tfrac{\theta}{\alpha} + \theta z, m\right)dz = \int_0^x \int_{\frac{\theta}{\alpha}+\theta z}^{\infty} t^{m-1}e^{-t}dtdz \\ &= \int_{\frac{\theta}{\alpha}}^{\frac{\theta}{\alpha}+\theta x}\int_0^{\frac{t}{\theta}-\frac{1}{\alpha}} t^{m-1}e^{-t}dzdt + \int_{\frac{\theta}{\alpha}+\theta x}^{\infty}\int_0^x t^{m-1}e^{-t}dzdt \\ &= \tfrac{1}{\theta}\left(\Gamma\left(\tfrac{\theta}{\alpha}, m+1\right) - \Gamma\left(\tfrac{\theta}{\alpha} + \theta x, m+1\right)\right) - \tfrac{1}{\alpha}\left(\Gamma\left(\tfrac{\theta}{\alpha}, m\right) - \Gamma\left(\tfrac{\theta}{\alpha} + \theta x, m\right)\right) \\ &+ x\Gamma\left(\tfrac{\theta}{\alpha} + \theta x, m\right). \end{aligned} \tag{18}$$

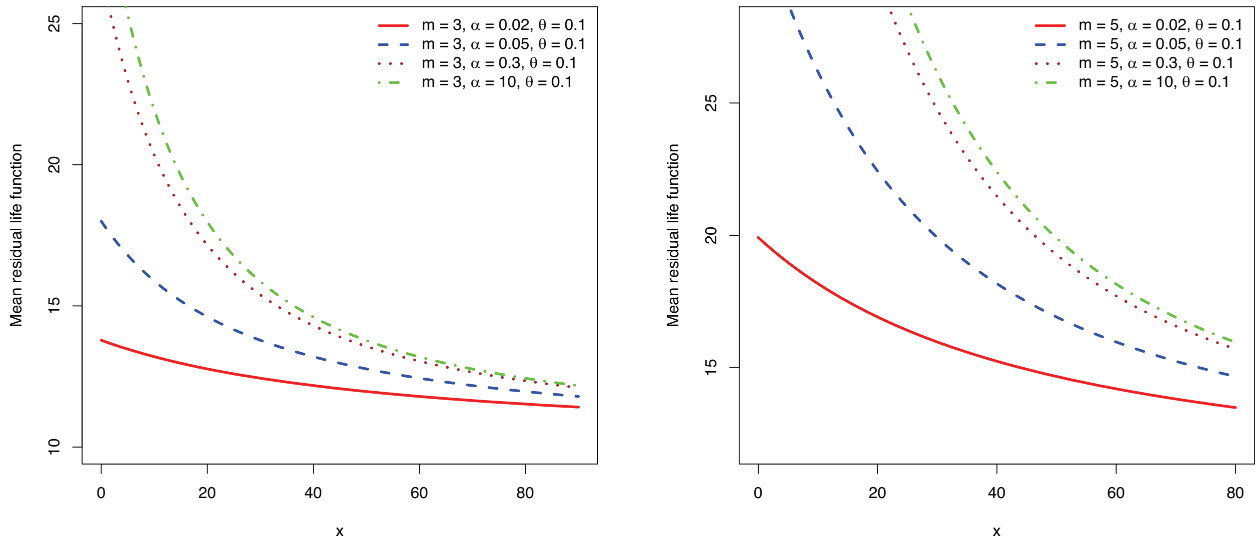

Fig. 3. The mean residual life of $AA(m,\alpha,\theta)$ for some values of parameters.

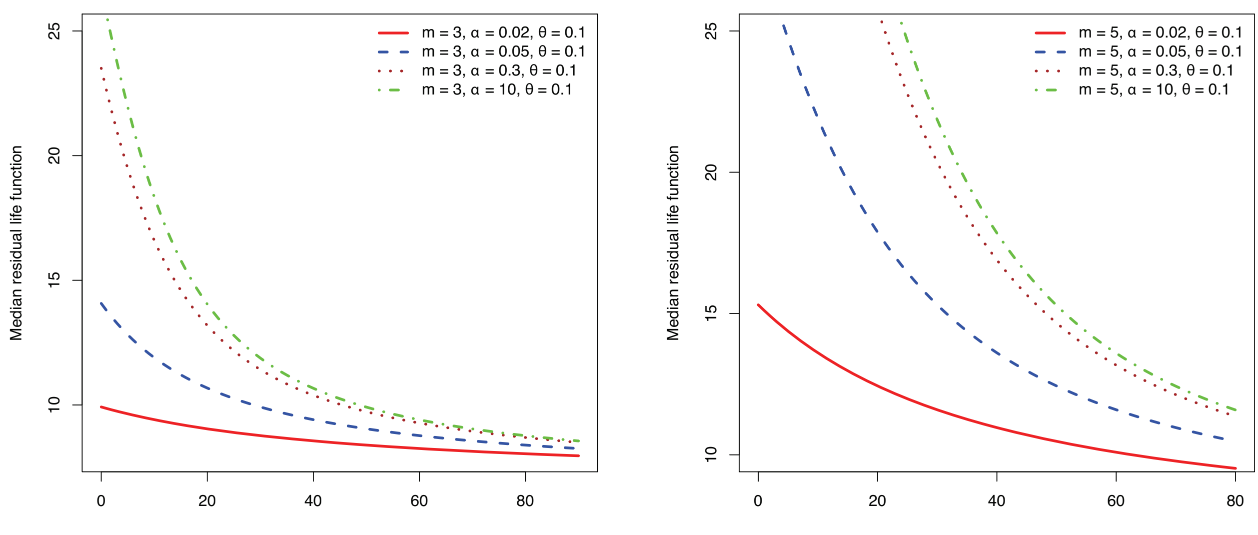

Fig. 4. The median residual life of $AA(m,\alpha,\theta)$ for some values of parameters.

On the other hand

$$\begin{aligned}\Gamma\left(\tfrac{\theta}{\alpha},m+1\right)-\Gamma\left(\tfrac{\theta}{\alpha}+\theta x,m+1\right) &= m\Gamma\left(\tfrac{\theta}{\alpha},m\right)+\left(\tfrac{\theta}{\alpha}\right)^{m}e^{-\frac{\theta}{\alpha}}\\ &\quad -m\Gamma\left(\tfrac{\theta}{\alpha}+\theta x,m\right)-\left(\tfrac{\theta}{\alpha}+\theta x\right)^{m}e^{-\left(\frac{\theta}{\alpha}+\theta x\right)}.\end{aligned} \tag{19}$$

Then, the results follows by (17), (18) and (19). □

The quantile inactivity time is an alternative for MIT and represents, at time x, the quantile of the elapsed time given the event occurred before x. For a distribution F, it is defined by:

$$q_p^*(x)=x-F^{-1}(\bar{p}F(x)),\quad x\geqslant 0. \tag{20}$$

and for $AA(m,\alpha,\theta)$ can be written as

$$q_p^*(x)=x-\frac{1}{\theta}\Gamma^{-1}\left(p\Gamma\left(\frac{\theta}{\alpha},m\right)+\bar{p}\Gamma\left(\frac{\theta}{\alpha}+\theta x,m\right),m\right)+\frac{1}{\alpha},\quad x\geqslant 0. \tag{21}$$

Fig. 5 shows the MIT and the median inactivity time ($q_{0.5}^*(x)$) of $AA(m,\alpha,\theta)$ for some parameters and shows increasing and convex forms. The greater values of the MIT indicate that the conditional distribution of the elapsed time is skewed to right.

4. Estimation of the parameters

Let m be known. By the moments method and applying (9), we can estimate (α,θ) by minimizing the following expression.

$$\Delta(\alpha,\theta)=\left(\bar{X}-E(X)\right)^2+\left(\bar{X^2}-E\left(X^2\right)\right)^2.$$

So, the moments estimator is

$$\left(\hat{\alpha},\hat{\theta}\right)=\arg\min_{\alpha,\theta}\Delta(\alpha,\theta).$$

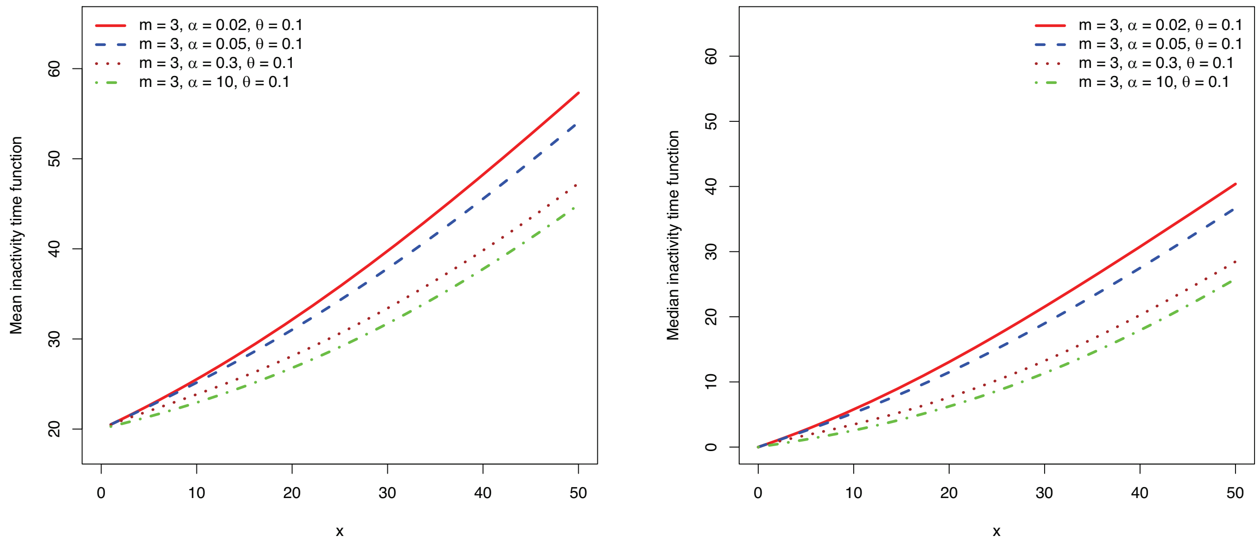

Fig. 5. Left: The MIT of $AA(m,\alpha,\theta)$ for some values of the parameters. Right: The median inactivity time of this distribution for the selected values of the parameters.

The moment estimates can be used as initial values for maximizing the log-likelihood function and finding the maximum likelihood estimator (MLE).

Let m be known and $x_1, x_2, \ldots, x_n$ represents a realization from $AA(m,\alpha,\theta)$, the log-likelihood function is

$$l(\alpha,\theta;\mathbf{x}) = -n\frac{\theta}{\alpha} + nm\ln\theta - n(m-1)\ln\alpha - n\ln\Gamma\left(\frac{\theta}{\alpha},m\right) + \sum_{i=1}^{n}(m-1)\ln(1+\alpha x_i) - \theta\sum_{i=1}^{n}x_i.$$

Also, the likelihood equations are as follows:

$$\frac{\partial l}{\partial\alpha} = n\frac{\theta}{\alpha^2} - n\frac{m-1}{\alpha} - n\frac{\left(\frac{\theta}{\alpha}\right)^m e^{-\frac{\theta}{\alpha}}}{\alpha\Gamma\left(\frac{\theta}{\alpha},m\right)} + (m-1)\sum_{i=1}^{n}\frac{x_i}{1+\alpha x_i} = 0,$$

and

$$\frac{\partial l}{\partial\theta} = -\frac{n}{\alpha} + \frac{nm}{\theta} + n\frac{\left(\frac{\theta}{\alpha}\right)^{m-1} e^{-\frac{\theta}{\alpha}}}{\alpha\Gamma\left(\frac{\theta}{\alpha},m\right)} - \sum_{i=1}^{n}x_i = 0.$$

The MLE can be calculated by maximizing the log-likelihood function or by solving the likelihood equations. If m is not known, which is usually the case, we can estimate the parameters for a range of m values and then choose the best model based on the Kolmogorov–Smirnov (K-S) statistic or other criteria.

The Fisher information matrix for the $AA(m,\alpha,\theta)$ is of the form

$$K = \begin{bmatrix} E\left(-\frac{\partial^2 l}{\partial\alpha^2}\right) & E\left(-\frac{\partial^2 l}{\partial\alpha\partial\theta}\right) \\ E\left(-\frac{\partial^2 l}{\partial\theta\partial\alpha}\right) & E\left(-\frac{\partial^2 l}{\partial\theta^2}\right) \end{bmatrix}.$$

where Let $l = \ln f(X)$. Let $X_i, i = 1,2,\ldots,n$ stands for a sample from $AA(m,\alpha_0,\theta_0)$. Then, the MLE, $\left(\hat{\alpha},\hat{\theta}\right)$, converges in distribution to bivariate normal $N\left((\alpha_0,\theta_0), n^{-1}K^{-1}\right)$ in which K^{-1} is the inverse of the information matrix.

Suppose that events are subject to random right censorship. A random event X_i is said to be right-censored if the only information about the event is that it is greater than the random censoring variable C_i, i.e. $X_i > C_i$. So, the observations of a right censored random sample consist of $T_i = \min(X_i, C_i)$ and d_i, where $d_i = 1$ if the event is not censored, $X_i \leqslant C_i$, and $d_i = 0$ if the event is censored, $X_i > C_i$. Let we have one right censored sample $(t_i, d_i), i = 1,2,\ldots,n$. Then, the log-likelihood function is

$$l(\alpha,\theta;\mathbf{t},\mathbf{d}) = \sum_{i=1}^{n}d_i\ln f(t_i) + \sum_{i=1}^{n}(1-d_i)\ln R(t_i).$$

in which f and R are the PDF and the reliability functions of the $AA(m,\alpha,\theta)$ respectively. It is easy to check that the log-likelihood function simplifies to

$$\begin{aligned} l(\alpha,\theta;\mathbf{t},\mathbf{d}) &= n\ln\Gamma\left(\tfrac{\theta}{\alpha},m\right) + \sum_{i=1}^{n}d_i\left(-\tfrac{\theta}{\alpha} + m\ln\theta - (m-1)\ln\alpha\right. \\ &\qquad \left. + (m-1)\ln(1+\alpha t_i) - \theta t_i\right) \\ &\quad + \sum_{i=1}^{n}(1-d_i)\ln\Gamma\left(\tfrac{\theta}{\alpha} + \theta t_i, m\right). \end{aligned}$$

5. Simulation

Given that $AA(m,\alpha,\theta)$ is a mixture of gamma distributions, we can extract a random sample of size n from this model as described in the following steps:

1. Simulate one random instance of multinomial distribution with parameters $n, w_1, w_2, \ldots, w_m$, where w_i is defined by (3). Let the simulated instance be $(n_1, n_2, \ldots, n_m)$ which will satisfy $n_1 + n_2 + \ldots + n_m = n$.
2. For every n_i, simulate one random sample of the gamma distribution $G(m-i+1,\theta)$ of size n_i. Then combine these samples to provide one sample of size n from $AA(m,\alpha,\theta)$.
3. Here, the degenerate random variable with mean t_* has been used as the random censorship variable C_i. Given the censorship percentage p, we have $t_* = q(\bar{p})$ where $q()$ is defined in (11).

The simulation results have been abstracted in Table 1. We have considered two values $p = 0$ and 0.2 for censorship rate. Every cell of the table shows the results for one run. In every run, we provide $r = 1000$ replicates of samples of size $n = 80, 100$. For every sample the MLE, $\left(\hat{\alpha},\hat{\theta}\right)$ has been computed. Then the bias (B), absolute bias (AB) and the mean squared error (MSE) for both α and θ have been computed. These measures are defined in the following relations for α. They are defined similarly for θ.

$$B_\alpha = \frac{1}{r}\sum_{i=1}^{r}(\hat{\alpha}_i - \alpha),$$

Table 1
Simulation results for MLE of the parameters of $AA(m,\alpha,\theta)$. Every cell consists of triples $B_\alpha, AB_\alpha, MSE_\alpha$ in the first line and $B_\theta, AB_\theta, MSE_\theta$ in the second line.

			$p=0$			$p=0.2$		
n	m	α,θ	B	AB	MSE	B	AB	MSE
80	3	0.02, 0.05	0.004988	0.013574	0.000373	0.006054	0.018217	0.000646
			0.000600	0.008016	0.000104	-0.000659	0.011561	0.000194
		0.3, 5	0.983335	1.212043	3.712165	1.175853	1.455283	5.641408
			0.848339	1.278958	2.889727	0.980640	1.536448	4.373064
		0.05, 2	0.461190	0.499459	0.682545	0.616871	0.658195	1.280510
			0.470890	0.579849	0.623839	0.561123	0.705742	0.986987
	5	0.02, 0.05	0.001707	0.006268	0.000066	0.001591	0.007357	0.000098
			0.000680	0.005271	0.000043	-0.000078	0.006786	0.000075
		0.3, 5	0.209712	0.543812	0.572916	0.493983	0.762481	1.184417
			0.337317	1.417153	3.323895	0.958836	1.892733	5.972325
		0.05, 2	0.206730	0.251102	0.156499	0.320155	0.356562	0.356936
			0.501932	0.695099	0.956195	0.737981	0.904887	1.761379
100	3	0.02, 0.05	0.004989	0.012495	0.000293	0.004451	0.016137	0.000483
			0.000897	0.007160	0.000081	-0.001198	0.010670	0.000173
		0.3, 5	0.867955	1.100889	2.932451	1.059814	1.356333	4.983397
			0.770588	1.216026	2.658104	0.877139	1.441000	3.933093
		0.05, 2	0.431200	0.466580	0.531628	0.539306	0.582099	1.067997
			0.456392	0.558029	0.568953	0.523228	0.652936	0.872355
	5	0.02, 0.05	0.001778	0.005664	0.000055	0.001718	0.006716	0.000082
			0.000839	0.004859	0.000038	0.000443	0.006135	0.000060
		0.3, 5	0.218949	0.535202	0.552218	0.420245	0.700898	1.018933
			0.349058	1.399999	3.103568	0.795703	1.808621	5.484757
		0.05, 2	0.168868	0.215696	0.126635	0.265858	0.302138	0.246075
			0.411721	0.617124	0.793489	0.645761	0.806164	1.394646

$$AB_\alpha = \frac{1}{r}\sum_{i=1}^{r}|\hat{\alpha}_i - \alpha|,$$

and

$$MSE_\alpha = \frac{1}{r}\sum_{i=1}^{r}(\hat{\alpha}_i - \alpha)^2.$$

Some of the main observations from simulation results are listed in the following.

- The MLE of the parameters are consistent, i.e., AB and MSE decrease with n.
- The AB and MSE have larger values for censored data ($p = 0.2$) rather than uncensored data.

6. Applications

Shanker (2016a,b) analyzed a data set consisting of the number of cycles to failure for 25 yarn samples. The data are: 15, 20, 38, 42, 61, 76, 86, 98, 121, 146, 149, 157, 175, 176, 180, 180, 198, 220, 224, 251, 264, 282, 321, 325, 653. The MLE of the parameters of $AA(m,\alpha,\theta)$ was calculated from $m = 1, 2, \ldots, 20$. For $m = 10$, we have the smallest value of the K-S statistic. Thus, the estimated model is $AA(10, 0.003368, 0.022551)$. The K-S statistic and corresponding p-value are 0.105787 and 0.942375, respectively. The AIC value, which corresponds to the MLE, is 306.2316. Fig. 6, left panel, shows the empirical CDF and the CDF of the fitted model, and graphically confirms that the model describes the data well. Also, Fig. 8, left panel, draws the histogram of data and the estimated PDFs.

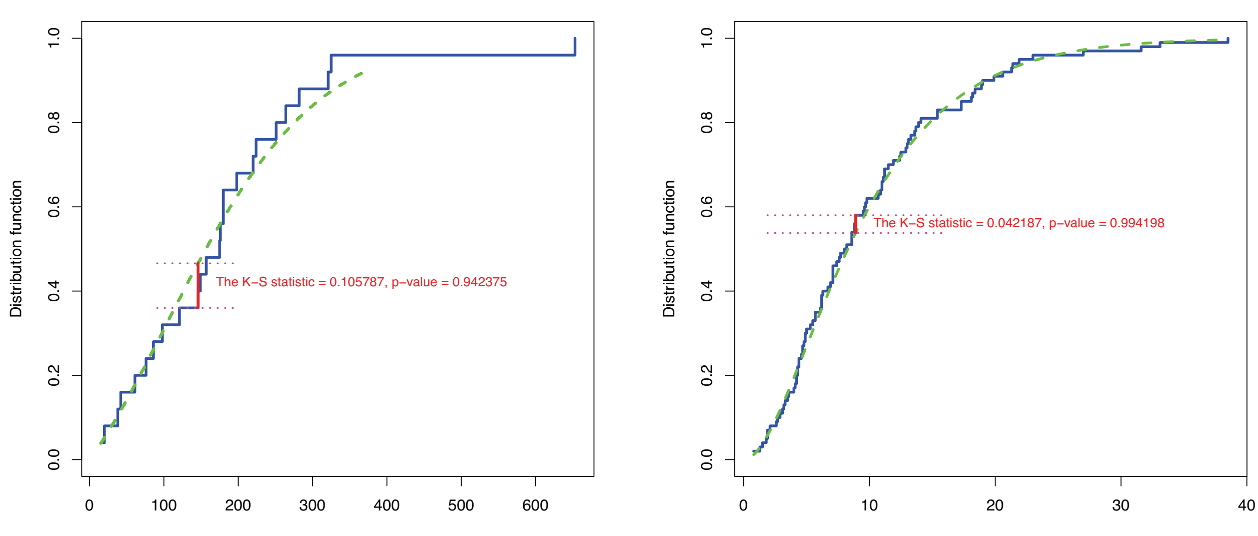

Fig. 6. The empirical distribution function along with fitted $AA(m,\alpha,\theta)$.

Table 2
The results of fitting the Lindley distribution, the Aradhana distribution, the generalized Aradhana distribution and the $AA(m, \alpha, \theta)$ to data sets.

Data	Model	$\hat{\theta}$	$\hat{\alpha}$	K-S statistics	K-S p-value	AIC
The first data set	Lindley	0.01115		0.127756	0.762682	307.019
	Aradhana	0.016728		0.12346	0.7967	311.0772
	Generalized Aradhana	0.013388	0.019097	0.11755	0.8407	308.864
	$AA(10, \alpha, \theta)$	0.022551	0.003368	0.105787	0.942375	306.2316
The second data set	Lindley	0.1866		0.067677	0.7495	640.0784
	Aradhana	0.27655		0.080136	0.5419	642.343
	Generalized Aradhana	0.2597	0.5556	0.060349	0.8596	641.5076
	$AA(2, \alpha, \theta)$	0.202477	1545.077	0.042187	0.991498	638.6034

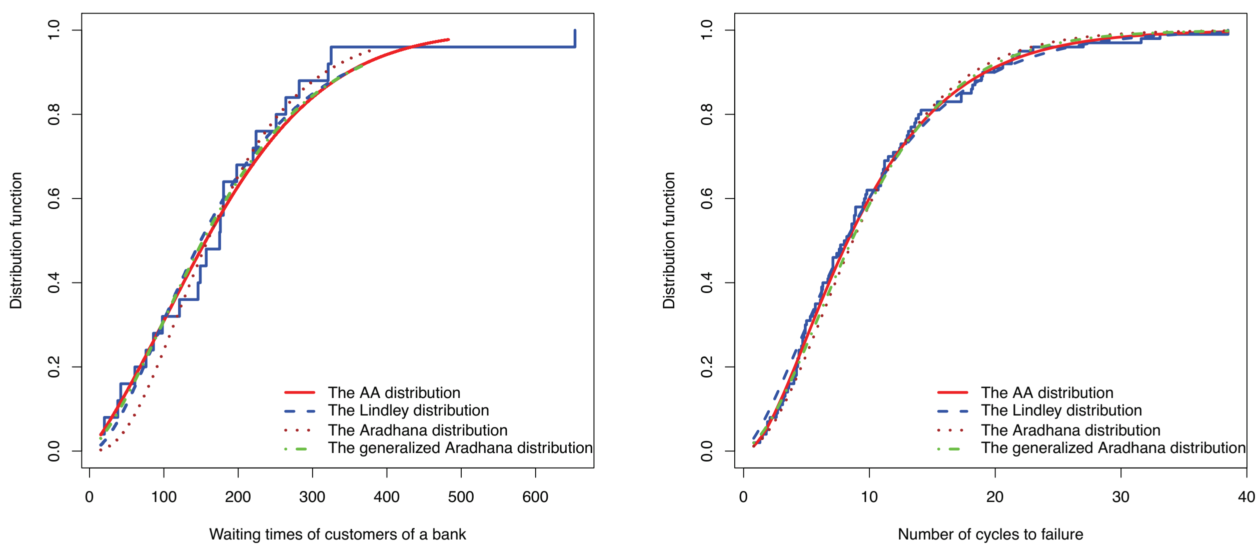

Fig. 7. The empirical distribution function along with some fitted models for the first data set (left) and the second data set (right).

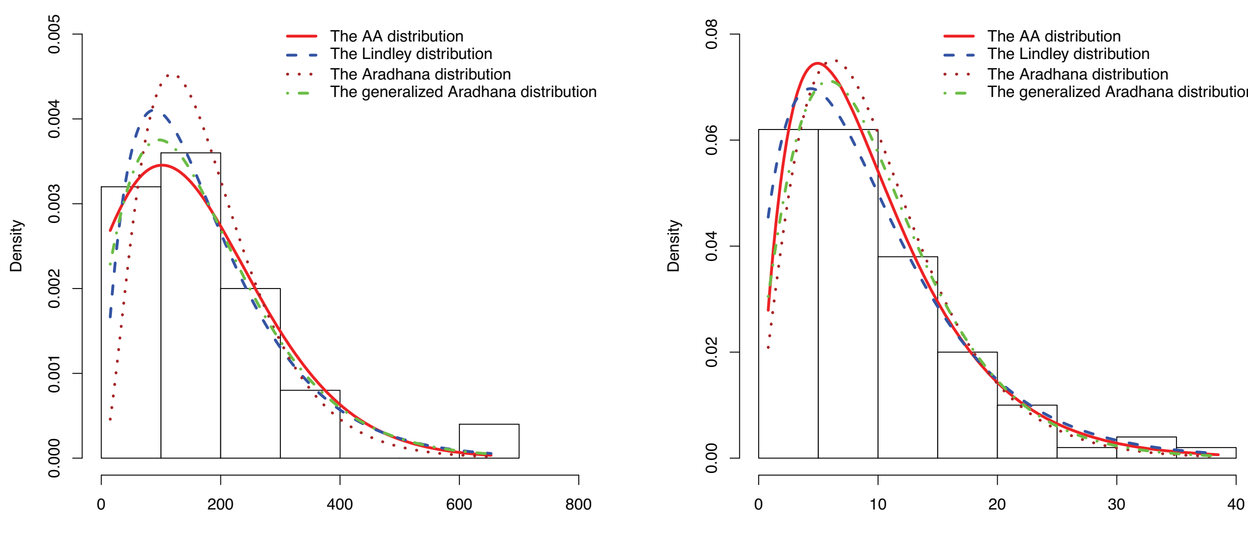

Fig. 8. The histogram along with some fitted models for the first data set (left) and the second data set (right).

The second data set consist of 100 waiting times (in minutes) of customers to be served in a bank: 0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4.0, 4.1, 4.2, 4.2, 4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5.0, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1, 7.1, 7.1, 7.4, 7.6, 7.7, 8.0, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11.0, 11.0, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13.0, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19.0, 19.9, 20.6, 21.3, 21.4, 21.9, 23.0, 27.0, 31.6, 33.1, 38.5, see Ghitany et al. (2008) and Shanker (2015).

For $m = 1, 2, \ldots, 20$, the MLEs of the parameters of $AA(m, \alpha, \theta)$ were calculated. Based on the K-S statistics, the best model among them is $AA(2, 1545.077, 0.202477)$. The corresponding AIC value is 638.6034, and the K-S statistics and p-value are 0.042187 and

0.994198, respectively, indicating a good fit. Fig. 6, right panel, shows the empirical and fitted CDFs and confirms a good fit.

In a comparative analysis, we fitted the Lindley distribution, the Aradhana distribution, and the generalized Aradhana distribution defined by (1) to these data sets. The MLEs of the parameters of the models were estimated and the results, summarized in Table 2, show that the proposed model $AA(m, \alpha, \theta)$ performs better than the others in both examples. Figs. 7 and 8 show the fitted distributions and PDFs respectively.

7. Conclusion

Because of its applicability and usefulness, the Lindley model and its generalizations have been considered by many authors. Here, a new generalization of the Lindley distribution has been introduced to extend this collection. Some properties of this distribution have been studied. Parameter estimation was discussed using moments and maximum likelihood methods. Simulation studies show that the MLE is efficient and consistent for both complete and right-censored data. The results of fitting the presented model to two real data sets show that it is useful in analyzing data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the two anonymous reviewers for their suggestions and very constructive comments that improved the presentation and readability of the paper. This work is supported by Researchers Supporting Project number (RSP-2021/323), King Saud University, Riyadh, Saudi Arabia.

References

Abouammoh, A.M., Alshangiti, A.M., Ragab, I.E., 2015. A new generalized Lindley distribution. J. Stat. Comput. Simul. 85, 3662–3678.

Abouammohm, A.M., Kayid, M., 2020. A New Flexible Generalized Lindley Model: Properties, Estimation and Applications. Symmetry 12, 1678. URL: https://www.mdpi.com/2073-8994/12/10/1678.

Al-Babtain A.A., Eid, H.A., A-Hadi N.A., Merovci, F., 2014. The five parameter Lindley distribution. Pakistan J. Stat. 31, 363–384..

Al-Mutairi, D.K., Ghitany, M.E., Kundu, D., 2013. Inferences on stress-strength reliability from Lindley distribution. Commun. Stat. Theory Methods 42, 1443–1463.

Atienza, N., Garcia-Heras, J., Muñoz-Pichardo, J.M., 2006. A new condition for identifiability of finite mixture distributions. Metrika 63, 215–221. https://doi.org/10.1007/s00184-005-0013-z.

Bader, M.G., Priest, A.M., 1982. Statistical aspects of fiber and bundle strength in hybrid composites. In hayashi, T., Kawata, K. Umekawa, S. (Eds), Progress in Science in engineering Composites, ICCM-IV, Tokyo. 1129–1136..

Bakouch, H.S., Al-Zahrani, B.M., Al-Shaomrani, A.A., Marchi, V.A.A., Louzada, F., 2012. An extended Lindley distribution. J. Korean Stat. Soc. 41, 75–85.

Cakmakyapan, S., Kadilar, O.G., 2017. The Lindley family of distributions: properties an applications. Hacettepe J. Math. Stat. 46, 1113–1137.

Cordeiro, G.M., Afify, A.Z., Yousof, H.M., Cakmakyapan, S., Ozel, G., 2018. The Lindley Weibull Distribution: properties and applications. Acad. Bras. Cienc. 90 (3), 2579–2598. https://doi.org/10.1590/0001-3765201820170635. PMID:30304208.

Fuller, E.R.J., Frieman, E.W., Quinn, B., Quinn, G.D., Craig, C., 1994. Fracture mechanics approach to the design of glass aircraft windows: A case study. SPIE Proc. 2286, 419–430.

Ghitany, M.E., Atieh, B., Nadarajah, S., 2008. Lidley distribution and its application. Math. Comput. Simul. 78, 493–506.

Ghitany, M.E., Alqallaf, F., Al-Mutairi, D.K., Husain, H.A., 2011. A two-parameter Lindley distribution and its applications to survival data. Math. Comput. Simul. 81, 1190–1201.

Gross, A.J., Clark, V.A., 1976. Survival Distributions: Reliability Applications in the Biometrical Sciences. John Wiley, New York.

Lindley, D.V., 1958. Fiducial distributions and Bayes' theorem. J. R. Stat. Soc. B 20, 102–107.

Lindsay, B.G., 1995. Mixture models: theory, geometry and applications. In: NSF-CBMS Regional Conference Series in Probability and Statistics, p. 5.

Ord, J.K., 1972. Families of frequency distributions. Charles Griffin, London, U.K..

Shanker, R., 2015. Shanker distribution and its applications. Int. J. Stat. Appl. 5, 338–348.

Shanker, R., 2016a. Aradhana distribution and its applications. Int. J. Stat. Appl. 6, 23–34.

Shanker, R., 2016b. On Generalized Lindley Distribution and Its Applications to Model Lifetime Data from Biomedical Science and Engineering. Insights Biomed. 1 (2).

Shanker, R., Shukla, K.K., 2018. A two-parameter power Aradhana distribution with properties and application. Indian J. Ind. Appl. Math. 9, 210–220.

Titterington, D.M., Smith, A.F.M., Makov, U.E., 1985. Statistical analysis of finite mixture distributions. John Wiley and Sons, Chichester, U.K..

Welday, D., Shanker, R., 2018. A generalized Aradhana distribution with properties and applications. Biometrics Biostat. Int. J. 7 (4), 374–385.