Journal of King Saud University – Science 31 (2019) 362–371

Contents lists available at ScienceDirect

Journal of King Saud University – Science

journal homepage: www.sciencedirect.com

Original article

Exponentiated generalized exponential Dagum distribution

Suleman Nasiru [a,*], Peter N. Mwita [b], Oscar Ngesa [c]

[a] Pan African University, Institute for Basic Sciences, Technology and Innovation, P. O. Box 62000-00200, Nairobi, Kenya
[b] Machakos University, Department of Mathematics, P. O. Box 136-90100, Machakos, Kenya
[c] Taita Taveta University, Mathematics and Informatics Department, P. O. Box 635-80300, Voi, Kenya

ARTICLE INFO

Article history:
Received 27 June 2017
Accepted 11 September 2017
Available online 14 September 2017

Keywords:
Dagum
Quantile
Moment
Entropy
Reliability measure
Order statistics

ABSTRACT

In this study, the exponentiated generalized exponential Dagum distribution has been proposed and studied. This family of distribution consists of a number of sub-models such as the exponentiated generalized Dagum distribution, Dagum distribution, Fisk distribution, Burr III distribution and exponentiated generalized exponential Burr III distribution among others. Statistical properties of the new family were also derived. Maximum likelihood estimators of the parameters of the distribution were developed and simulation studies performed to assess the properties of the estimators. Applications of the model was demonstrated to show its usefulness.

1. Introduction

Identifying an appropriate distribution for modeling data sets is very important in statistical analysis. Knowing the appropriate distribution a particular data sets follow helps in making sound inference about the data. Because of this, barrage of techniques have been developed for modifying existing statistical distributions to make them more flexible in modeling data sets that arise in different fields of study. The Dagum distribution (Dagum, 1977) just like other existing statistical distributions has received much attention recently due to its usefulness in modeling of size distribution of personal income and reliability analysis among others. For an extensive review on the genesis and on empirical applications of the Dagum see (Kleiber and Kotz, 2003; Kleiber, 2008).

With the goal of increasing the flexibility of the Dagum distribution in modeling lifetime data, different modifications of the distribution have been proposed in literature recently and includes: Dagum-Poisson distribution (Oluyede et al., 2016), Mc-Dagum distribution (Oluyede and Rajasooriya, 2013), gamma-Dagum distribution (Oluyede et al., 2014), transmuted Dagum distribution (Elbatal and Aryal, 2015), exponentiated Kumaraswamy-Dagum distribution (Huang and Oluyede, 2014), extended Dagum distribution (Silva et al., 2015), beta-Dagum distribution (Domma and Condino, 2013), weighted Dagum distribution (Oluyede and Ye, 2014) and log-Dagum distribution (Domma and Perri, 2009).

In addition, other authors have studied the properties and methods of estimation of the parameters of the Dagum distribution. Shahzad and Asghar (2013) employed the TL-moments to estimate the parameters of the Dagum distribution. Dey et al. (2017) studied the properties and different methods of estimating the parameters of the Dagum distribution. Domma et al. (2011) estimated the Dagum distribution with censored sample using maximum likelihood estimation. In another study, Al-Zahrani (2016) proposed a reliability test plan to determine the termination time of the experiment for a given sample size, producer risk and termination number when the quantity of interest follows the Dagum distribution.

Thus, in this study a new extension of the Dagum distribution called the exponentiated generalized exponential Dagum distribution with tractable cumulative distribution function is proposed with the basic motivation of modeling lifetime data with both monotonic and non-monotonic failure rates, control skewness, kurtosis and tail variations. The rest of the paper is organized as follows: in Section 2, the cumulative distribution function, probability density function, survival function and hazard function of the new distribution were defined. In Section 3, some sub-models of the new distribution were presented. In Section 4,

* Corresponding author.
E-mail addresses: sulemanstat@gmail.com, snasiru@uds.edu.gh (S. Nasiru), petermwita@mksu.ac.ke (P.N. Mwita), oscanges@ttu.ac.ke (O. Ngesa).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jksus.2017.09.009
1018-3647/

statistical properties of the new distribution were discussed. In Section 5, the parameters of the new distribution were estimated using maximum likelihood estimation and Monte Carlo simulation performed to assess the stability of the parameters. In Section 6, the applications of the new model was demonstrated using two data sets. Finally, the concluding remarks of the study was given in Section 7.

2. New model

Let T be a random variable with probability density function (PDF) $\lambda e^{-\lambda t}, t > 0, \lambda > 0$ and let X be a continuous random variable with cumulative distribution function (CDF) $F(x)$. Then the CDF of the exponentiated generalized exponential (EGE)-X family of distribution is defined as

$$G(x) = \int_0^{-\log\left[1-\left(1-\bar{F}^d(x)\right)^c\right]} \lambda e^{-\lambda t} dt = 1 - \left\{1 - \left[1 - (1 - F(x))^d\right]^c\right\}^\lambda, \tag{1}$$

where $\bar{F}(x) = 1 - F(x)$.

For positive integers λ and c, a physical interpretation of the EGE-X family of distribution CDF is given as follows. Eq. (1) represents the CDF of the lifetime of a series-parallel system consisting of independent components with the CDF $1 - (1 - F(x))^d$ corresponding to the Lehman type II distribution. Given that a system is formed by λ independent component series subsystems and that each of the subsystems is made up of c independent parallel components. Suppose $X_{ij} \sim 1 - (1 - F(x))^d$, for $1 \leqslant i \leqslant c$ and $1 \leqslant j \leqslant \lambda$, represents the lifetime of the i^{th} component in the j^{th} subsystem and X is the lifetime of the entire system. Then, we have

$$\begin{aligned}\mathbb{P}(X \leqslant x) &= 1 - [1 - \mathbb{P}(X_{11} \leqslant x, \ldots, X_{1c} \leqslant x)]^\lambda \\ &= 1 - [1 - \mathbb{P}^c(X_{11} \leqslant x)]^\lambda,\end{aligned}$$

and X has the CDF defined in Eq. (1).

Suppose $F(x) = (1 + \alpha x^{-\theta})^{-\beta}, x > 0, \alpha > 0, \beta > 0, \theta > 0$ is the CDF of type I Dagum distribution, then the CDF of the exponentiated generalized exponential-Dagum distribution (EGEDD) is given by

$$G(x) = 1 - \left\{1 - \left[1 - \left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^d\right]^c\right\}^\lambda, \; x > 0, \tag{2}$$

where the parameters $\alpha, \beta, \theta, \lambda, c$ and d are non-negative, with $\beta, \theta, \lambda, c$ and d being shape parameters and α being a scale parameter. The corresponding PDF of the EGEDD is given by

$$g(x) = \frac{\alpha\beta\lambda\theta c d (1 + \alpha x^{-\theta})^{-\beta-1}\left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^{d-1}\left[1 - \left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^d\right]^{c-1}}{x^{\theta+1}\left\{1 - \left[1 - \left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^d\right]^c\right\}^{1-\lambda}}, x > 0. \tag{3}$$

Lemma 1. *The PDF of the EGEDD can be expressed in terms of the density function of the Dagum distribution as*

$$g(x) = \lambda c d \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \omega_{ijk} f_D(x; \alpha, \theta, \beta_{k+1}), \lambda > 0, \alpha > 0, \theta > 0,$$
$$\beta_{k+1} > 0, c > 0, d > 0, x > 0, \tag{4}$$

where $f_D(x; \alpha, \theta, \beta_{k+1})$ *is the PDF of the Dagum distribution with parameters* α, θ *and* $\beta_{k+1} = \beta(k + 1)$ *and*

$$\omega_{ijk} = \frac{(-1)^{i+j+k}\Gamma(\lambda)\Gamma(c(i+1))\Gamma(d(j+1))}{i!j!(k+1)!\Gamma(\lambda-i)\Gamma(c(i+1)-j)\Gamma(d(j+1)-k)}, \Gamma(a+1) = a!.$$

Proof. For a real non-integer $\eta > 0$, a series expansion for $(1 - z)^{\eta-1}$, for $|z| < 1$ is

$$(1 - z)^{\eta-1} = \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma(\eta)}{i!\,\Gamma(\eta - i)} z^i \tag{5}$$

Applying the series expansion in Eq. (5) twice and the fact that $0 < (1 + \alpha x^{-\theta})^{-\beta} < 1$, implies that

$$\left[1 - \left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^d\right]^{c-1}\left\{1 - \left[1 - \left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^d\right]^c\right\}^{\lambda-1}$$
$$= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^{i+j}\Gamma(\lambda)\Gamma(c(i+1))\left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^{dj}}{i!j!\Gamma(\lambda - i)\Gamma(c(i+1) - j)}. \tag{6}$$

Substituting Eq. (6) into Eq. (3) yields

$$g(x) = \lambda c d \alpha\beta\theta x^{-\theta-1}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \times \frac{(-1)^{i+j}\Gamma(\lambda)\Gamma(c(i+1))(1+\alpha x^{-\theta})^{-\beta-1}\left(1-(1+\alpha x^{-\theta})^{-\beta}\right)^{d(j+1)-1}}{i!j!\Gamma(\lambda-i)\Gamma(c(i+1)-j)}.$$

Applying the series expansion again to $\left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^{d(j+1)-1}$ gives us the expansion of the density as

$$g(x) = \lambda c d \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \omega_{ijk} f_D(x; \alpha, \theta, \beta_{k+1}), x > 0.$$

□

Eq. (4) revealed that the PDF of the EGEDD can be written as a linear combination of the Dagum distribution with different shape parameters. The expansion of the PDF is important in providing the mathematical properties of the EGEDD. The triple infinite series in Eq. (4) is convergent for all $\lambda > 0, \alpha > 0, \theta > 0, \beta_{k+1} > 0$, $c > 0, d > 0$ and $x > 0$. This can easily be verified using symbolic computational softwares such as MATHEMATICA, MAPLE and MATLAB. Fig. 1 displays different shapes of the PDF of the EGEDD for different parameter values. The survival function of this distribution is

$$S(x) = \left\{1 - \left[1 - \left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^d\right]^c\right\}^\lambda, \tag{7}$$

and the hazard function is

$$\tau(x) = \frac{\alpha\beta\lambda\theta c d (1 + \alpha x^{-\theta})^{-\beta-1}\left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^{d-1}\left[1 - \left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^d\right]^{c-1}}{x^{\theta+1}\left\{1 - \left[1 - \left(1 - (1 + \alpha x^{-\theta})^{-\beta}\right)^d\right]^c\right\}}, \; x > 0. \tag{8}$$

The plots of the hazard function display various attractive shapes such as monotonically decreasing, monotonically increasing, upside down bathtub, bathtub and bathtub followed by upside down bathtub shapes for different combination of the values of the parameters. These features make the EGEDD suitable for modeling monotonic and non-monotonic failure rates that are more likely to be encountered in real life situation. Fig. 2 displays the various shapes of the hazard function.

3. Sub-models

The EGEDD consists of a number of important sub-models that are widely used in lifetime modeling. These include: exponentiated generalized Dagum distribution (EGDD), Dagum distribution (DD),

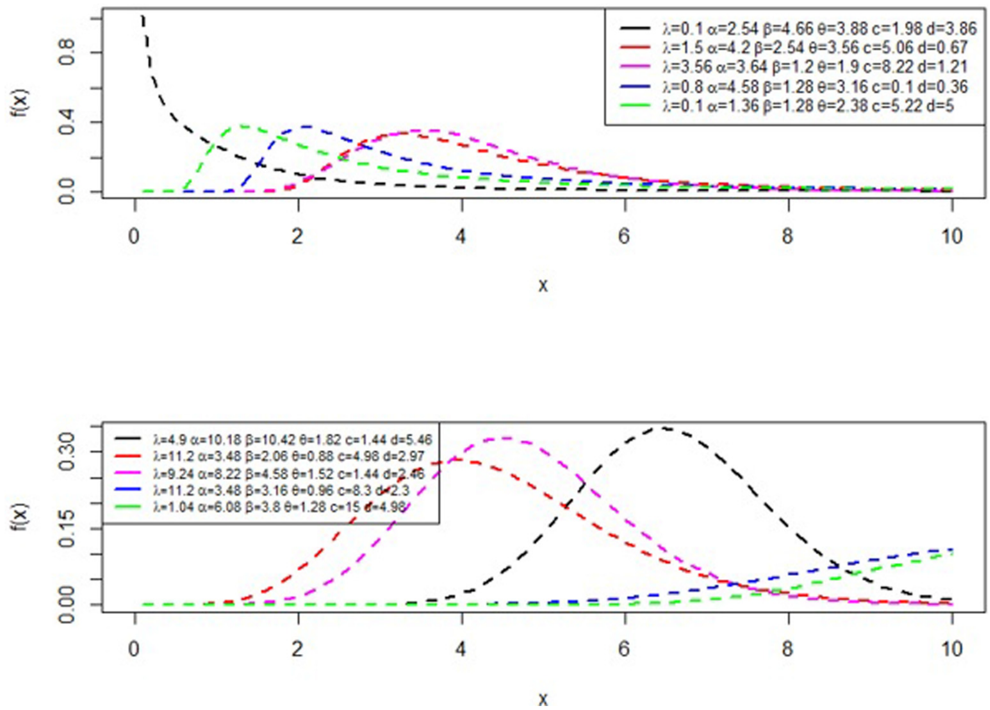

Fig. 1. EGEDD density function.

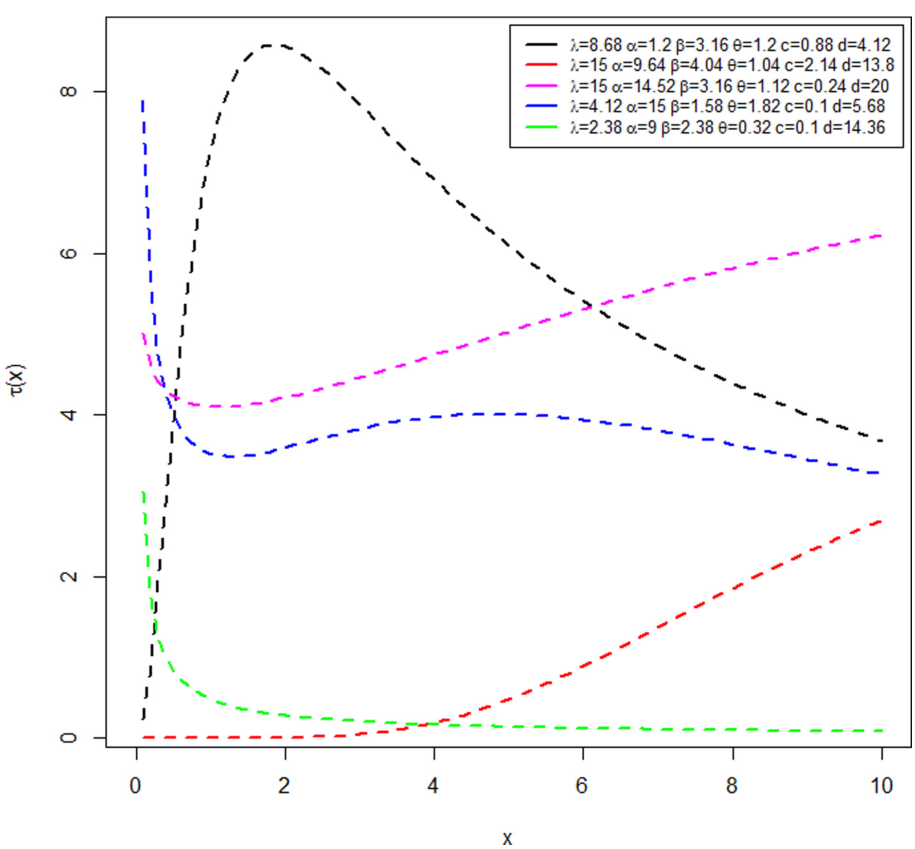

Fig. 2. Plots of the EGEDD hazard function.

exponentiated generalized exponential Burr III distribution (EGEBD), Burr III distribution, exponentiated generalized Burr III distribution (EGBD), exponentiated generalized exponential Fisk distribution (EGEFD), exponentiated generalized Fisk distribution (EGFD) and Fisk distribution (FD). Table 1 displays a list of models that can be derived from the EGEDD.

Table 1
Summary of sub-models from the EGEDD.

Distribution	α	λ	β	θ	c	d
EGDD	α	1	β	θ	c	d
DD	α	1	β	θ	1	1
EGEBD	1	λ	β	θ	c	d
BD	1	1	β	θ	1	1
EGBD	1	1	β	θ	c	d
EGEFD	α	λ	1	θ	c	d
EGFD	α	1	1	θ	c	d
FD	α	1	1	θ	1	1

4. Statistical properties

In this section, various statistical properties of the EGEDD such as the quantile, moment, reliability measure, entropy and order statistics were derived.

4.1. Quantile function

The distribution of a random variable can be described using its quantile function. The quantile function is useful in computing the median, kurtosis and skewness of the distribution of a random variable.

Lemma 2. *The quantile function of the EGEDD for $p \in (0,1)$ is given by*

$$Q_X(p) = \left\{\frac{1}{\alpha}\left[\left(1-\left(1-\left(1-(1-p)^{\frac{1}{\lambda}}\right)^{\frac{1}{c}}\right)^{\frac{1}{d}}\right)^{-\frac{1}{\beta}}-1\right]\right\}^{-\frac{1}{\theta}}. \quad (9)$$

Proof. By definition, the quantile function returns the value x such that

$$G(x_p) = \mathbb{P}(X \leqslant x_p) = p.$$

Thus

$$1-\left\{1-\left[1-\left(1-\left(1+\alpha x_p^{-\theta}\right)^{-\beta}\right)^{d}\right]^{c}\right\}^{\lambda} = p. \quad (10)$$

Letting $x_p = Q_X(p)$ in Eq. (10) and solving for $Q_X(p)$ using inverse transformation yields

$$Q_X(p) = \left\{\frac{1}{\alpha}\left[\left(1-\left(1-\left(1-(1-p)^{\frac{1}{\lambda}}\right)^{\frac{1}{c}}\right)^{\frac{1}{d}}\right)^{-\frac{1}{\beta}}-1\right]\right\}^{-\frac{1}{\theta}}.$$

□

When $p = 0.25, 0.5$ and 0.75, we obtain the first quartile, the median and the third quartile of the EGEDD respectively.

4.2. Moment

It is imperative to derive the moments when a new distribution is proposed. They play a significant role in statistical analysis, particularly in applications. Moments are used in computing measures of central tendency, dispersion and shapes among others.

Proposition 1. *The r^{th} non-central moment of the EGEDD is given by*

$$\mu'_r = \lambda c d \alpha^{\frac{r}{\theta}} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \omega_{ijk}\, \beta_{k+1}\, B\left(\beta_{k+1}+\frac{r}{\theta}, 1-\frac{r}{\theta}\right), r<\theta, \quad (11)$$

where $B(\cdot,\cdot)$ is the beta function and $r = 1, 2, \ldots$.

Proof. By definition

$$\begin{aligned}\mu'_r &= \int_0^{\infty} x^r g(x)dx \\ &= \int_0^{\infty} x^r \lambda c d \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \omega_{ijk} f_D(x;\, \alpha,\, \theta,\, \beta_{k+1})dx \\ &= \lambda c d \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \omega_{ijk} \int_0^{\infty} x^r f_D(x;\, \alpha,\, \theta,\, \beta_{k+1})dx \\ &= \lambda c d \alpha^{\frac{r}{\theta}} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \omega_{ijk}\, \beta_{k+1}\, B\left(\beta_{k+1}+\frac{r}{\theta}, 1-\frac{r}{\theta}\right), r<\theta.\end{aligned}$$

□

The triple infinite series in Eq. (11) is convergent for all $\lambda > 0,\ \alpha > 0,\ \beta > 0,\ \theta > 0,\ c > 0,\ d > 0$ and $x > 0$.

4.3. Entropy

Entropy plays a vital role in science, engineering and probability theory, and has been used in various situations as a measure of variation of uncertainty of a random variable (Rényi, 1961). The Rényi entropy of a random X having the EGEDD is given by the following proposition.

Proposition 2. *If $X \sim EGEDD(\alpha, \lambda, \beta, \theta, c, d)$, then the Rényi entropy is given by*

$$I_R(\delta) = \frac{1}{1-\delta}\log\left[(\lambda\beta c d)^{\delta}\alpha^{\frac{1-\delta}{\theta}}\theta^{\delta-1}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\varpi_{ijk} B\left(\beta(\delta+k)+\frac{1-\delta}{\theta}, \delta+\frac{\delta-1}{\theta}\right)\right], \quad (12)$$

where $\delta \neq 1,\ \delta > 0,\ \beta(\delta+k)+\frac{1-\delta}{\theta} > 0,\ \delta+\frac{\delta-1}{\theta} > 0$ and

$$\varpi_{ijk} = \frac{(-1)^{i+j+k}\Gamma(\delta(\lambda-1)+1)\Gamma(c(\delta+i)-\delta+1)\Gamma(d(\delta+j)-\delta+1)}{i!j!k!\Gamma(\delta(\lambda-1)-i+1)\Gamma(c(\delta+i)-\delta-j+1)\Gamma(d(\delta+j)-\delta-k+1)}.$$

Proof. The Rényi entropy (Rényi, 1961) is defined as

$$I_R(\delta) = \frac{1}{1-\delta}\log\left[\int_0^{\infty} g^{\delta}(x)dx\right],\ \delta \neq 1,\ \delta > 0.$$

Using the same approach for expanding the density,

$$g^{\delta}(x) = (\alpha\lambda\beta\theta c d)^{\delta}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\varpi_{ijk} x^{-\delta(\theta+1)}\left(1+\alpha x^{-\theta}\right)^{-\beta(\delta+k)-\delta}.$$

Thus

$$\begin{aligned}I_R(\delta) &= \frac{1}{1-\delta}\log\left[\int_0^{\infty}(\alpha\lambda\beta\theta c d)^{\delta}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\varpi_{ijk} x^{-\delta(\theta+1)}\left(1+\alpha x^{-\theta}\right)^{-\beta(\delta+k)-\delta}\right] \\ &= \frac{1}{1-\delta}\log\left[(\alpha\lambda\beta\theta c d)^{\delta}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\varpi_{ijk}\int_0^{\infty} x^{-\delta(\theta+1)}\left(1+\alpha x^{-\theta}\right)^{-\beta(\delta+k)-\delta}\right].\end{aligned}$$

Letting $y = (1+\alpha x^{-\theta})^{-1}$, when $x = \infty, y = 1$ and when $x = 0, y = 0$. Also, $dy = \alpha\theta x^{-\theta-1}(1+\alpha x^{-\theta})^{-2}dx$ and $x = (\alpha y)^{\frac{1}{\theta}}(1-y)^{-\frac{1}{\theta}}$. Hence

$$I_R(\delta) = \frac{1}{1-\delta}\log\left[(\alpha\lambda\beta\theta cd)^{\delta}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\varpi_{ijk}\int_0^1 y^{\beta(\delta+k)+\delta-2}\right.$$
$$\left.\times\left((\alpha y)^{\frac{1}{\theta}}(1-y)^{-\frac{1}{\theta}}\right)^{-\delta(\theta+1)+\theta+1}dy\right]$$
$$=\frac{1}{1-\delta}\log\left[(\lambda\beta cd)^{\delta}\alpha^{\frac{1-\delta}{\theta}}\theta^{\delta-1}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\varpi_{ijk}B(\beta(\delta+k)\right.$$
$$\left.+\frac{1-\delta}{\theta},\delta+\frac{\delta-1}{\theta}\Big)\right],$$

where $\delta\neq 1$, $\delta>0$, $\beta(\delta+k)+\frac{1-\delta}{\theta}>0$ and $\delta+\frac{\delta-1}{\theta}>0$.

The Rényi entropy tends to Shannon entropy as $\delta\to 1$. It can easily be verified from standard calculus that the triple infinite series in Eq. (12) is convergent for all $\lambda>0$, $\alpha>0$, $\beta>0$, $\theta>0$, $c>0$, $d>0$ and $x>0$.

4.4. Reliability

The estimation of reliability is vital in stress-strength models. If X_1 is the strength of a component and X_2 is the stress, the component fails when $X_1\leqslant X_2$. Then the estimate of the reliability of the component R is $\mathbb{P}(X_2<X_1)$.

Proposition 3. *If $X_1\sim EGEDD(\alpha,\lambda,\beta,\theta,c,d)$ and $X_2\sim EGEDD(\alpha,\lambda,\beta,\theta,c,d)$, then the estimation of reliability R is given by*

$$R = 1-\lambda cd\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{v_{ijk}}{(k+1)},\tag{13}$$

where

$$v_{ijk}=\frac{(-1)^{i+j+k}\Gamma(2\lambda)\Gamma(c(i+1))\Gamma(d(j+1))}{i!j!k!\Gamma(2\lambda-i)\Gamma(c(i+1)-j)\Gamma(d(j+1)-k)}.$$

Proof. By definition

$$R=\int_0^{\infty}g(x)G(x)dx$$
$$=1-\int_0^{\infty}g(x)S(x)dx$$
$$=1-\int_0^{\infty}\alpha\lambda\beta\theta cd\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}v_{ijk}x^{-\theta-1}\left(1+\alpha x^{-\theta}\right)^{-\beta(k+1)-1}dx$$
$$=1-\alpha\lambda\beta\theta cd\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}v_{ijk}\int_0^{\infty}x^{-\theta-1}\left(1+\alpha x^{-\theta}\right)^{-\beta(k+1)-1}dx$$
$$=1-\lambda cd\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{v_{ijk}}{(k+1)}.$$

□

The triple infinite series in Eq. (13) is convergent for all $\lambda>0$, $\alpha>0$, $\beta>0$, $\theta>0$, $c>0$, $d>0$ and $x>0$.

4.5. Order statistics

Let $X_1,X_2,\ldots,X_n$ be a random sample from the EGEDD and $X_{1:n}<X_{2:n}<\ldots<X_{n:n}$ are order statistics obtained from the sample. Then the PDF, $g_{p:n}(x)$, of the p^{th} order statistic $X_{p:n}$ is given by

$$g_{p:n}(x)=\frac{1}{B(p,n-p+1)}[G(x)]^{p-1}[1-G(x)]^{n-p}g(x),$$

where $G(x)$ and $g(x)$ are the CDF and PDF of the EGEDD respectively, and $B(\cdot,\cdot)$ is the beta function. Since $0<G(x)<1$ for $x>0$, using the binomial series expansion of $[1-G(x)]^{n-p}$, which is given by

$$[1-G(x)]^{n-p}=\sum_{l=0}^{n-p}(-1)^l\binom{n-p}{l}[G(x)]^l,$$

we have

$$g_{p:n}(x)=\frac{1}{B(p,n-p+1)}\sum_{l=0}^{n-p}(-1)^l\binom{n-p}{l}[G(x)]^{p+l-1}g(x).\tag{14}$$

Therefore, substituting the CDF and PDF of the EGEDD into Eq. (14) yields

$$g_{p:n}(x)=\sum_{l=0}^{n-p}\sum_{m=0}^{p+l-1}\frac{(-1)^{l+m}n!(p+l-1)!}{l!(m+1)!(p-1)!(n-p-l)!(p+l-m-1)!}$$
$$g(x;\alpha,\lambda_{m+1},\beta,\theta,c,d),\tag{15}$$

where $g(x;\alpha,\lambda_{m+1},\beta,\theta,c,d)$ is the PDF of the EGEDD with parameters α, β, θ, c, d and $\lambda_{m+1}=\lambda(m+1)$. It is obvious that the density of the p^{th} order statistic given in Eq. (15) is a weighted function of the EGEDD with different shape parameters. The double finite series in Eq. (15) is convergent for all $\lambda>0$, $\alpha>0$, $\beta>0$, $\theta>0$, $c>0$, $d>0$ and $x>0$.

5. Parameter estimation

In this section, the maximum likelihood estimators of the unknown parameters of the EGEDD are derived and their finite sample properties assessed. Let $X_1,X_2,\ldots,X_n$ be a random sample of size n from the EGEDD. Let $z_i=(1+\alpha x_i^{-\theta})$, then the log-likelihood function is given by

$$\ell=n\log(\alpha\lambda\beta\theta cd)-(\theta+1)\sum_{i=1}^{n}\log(x_i)-(\beta+1)\sum_{i=1}^{n}\log(z_i)$$
$$+(d-1)\sum_{i=0}^{n}\log\left(1-z_i^{-\beta}\right)+(c-1)\sum_{i=1}^{n}\log\left[1-\left(1-z_i^{-\beta}\right)^d\right]$$
$$+(\lambda-1)\sum_{i=1}^{n}\log\left\{1-\left[1-\left(1-z_i^{-\beta}\right)^d\right]^c\right\}.\tag{16}$$

Taking the first partial derivatives of the log-likelihood function in Eq. (16) with respect to the parameters α, λ, β, θ, c and d, we obtain the score functions as

$$\frac{\partial\ell}{\partial\lambda}=\frac{n}{\lambda}+\sum_{i=1}^{n}\log\left\{1-[1-(1-z_i^{-\beta})^d]^c\right\},\tag{17}$$

$$\frac{\partial\ell}{\partial c}=\frac{n}{c}+\sum_{i=1}^{n}\log[1-(1-z_i^{-\beta})^d]-(\lambda$$
$$-1)\sum_{i=1}^{n}\frac{[1-(1-z_i^{-\beta})^d]^c\log[1-(1-z_i^{-\beta})^d]}{1-[1-(1-z_i^{-\beta})^d]^c},\tag{18}$$

$$\frac{\partial\ell}{\partial d}=\frac{n}{d}+\sum_{i=1}^{n}\log(1-z_i^{-\beta})+(\lambda-1)\sum_{i=1}^{n}\frac{c(1-z_i^{-\beta})^d\left[1-(1-z_i^{-\beta})^d\right]^{c-1}\log(1-z_i^{-\beta})}{1-\left[1-(1-z_i^{-\beta})^d\right]^c}-$$
$$(c-1)\sum_{i=1}^{n}\frac{(1-z_i^{-\beta})^d\log(1-z_i^{-\beta})}{1-(1-z_i^{-\beta})^d},\tag{19}$$

$$\frac{\partial\ell}{\partial\beta}=\frac{n}{\beta}-\sum_{i=1}^{n}\log(z_i)+(d-1)\sum_{i=1}^{n}\frac{z_i^{-\beta}\log(z_i)}{1-z_i^{-\beta}}-(c-1)\sum_{i=1}^{n}\frac{dz_i^{-\beta}(1-z_i^{-\beta})^{d-1}\log(z_i)}{1-(1-z_i^{-\beta})^d}$$
$$+(\lambda-1)\sum_{i=1}^{n}\frac{cdz_i^{-\beta}(1-z_i^{-\beta})^{d-1}[1-(1-z_i^{-\beta})^d]^{c-1}\log(z_i)}{1-[1-(1-z_i^{-\beta})^d]^c},\tag{20}$$

$$\frac{\partial \ell}{\partial \theta}=\frac{n}{\theta}-\sum_{i=1}^{n}\log(x_i)+(\beta+1)\sum_{i=1}^{n}\frac{\alpha x_i^{-\theta}\log(x_i)}{z_i}-(d-1)\sum_{i=1}^{n}\frac{\alpha\beta x_i^{-\theta}z_i^{-\beta-1}\log(x_i)}{1-z_i^{-\beta}}$$
$$-(\lambda-1)\sum_{i=1}^{n}\frac{\alpha\beta c d x_i^{-\theta}z_i^{-\beta-1}(1-z_i^{-\beta})^{d-1}[1-(1-z_i^{-\beta})^{d}]^{c-1}\log(x_i)}{1-[1-(1-z_i^{-\beta})^{d}]^{c}}$$
$$+(c-1)\sum_{i=1}^{n}\frac{\alpha\beta d x_i^{-\theta}z_i^{-\beta-1}(1-z_i^{-\beta})^{d-1}\log(x_i)}{1-(1-z_i^{-\beta})^{d}}, \quad (21)$$

$$\frac{\partial \ell}{\partial \alpha}=\frac{n}{\alpha}-(\beta+1)\sum_{i=1}^{n}\frac{x_i^{-\theta}}{z_i}+(d-1)\sum_{i=1}^{n}\frac{\beta x_i^{-\theta}z_i^{-\beta-1}}{1-z_i^{-\beta}}-(c-1)\sum_{i=1}^{n}\frac{\beta d x_i^{-\theta}z_i^{-\beta-1}(1-z_i^{-\beta})^{d-1}}{1-(1-z_i^{-\beta})^{d}}$$
$$+(\lambda-1)\sum_{i=1}^{n}\frac{\beta c d x_i^{-\theta}z_i^{-\beta-1}(1-z_i^{-\beta})^{d-1}[1-(1-z_i^{-\beta})^{d}]^{c-1}}{1-[1-(1-z_i^{-\beta})^{d}]^{c}}. \quad (22)$$

The estimates for the parameters α, λ, β, θ, c and d are obtained by equating the score functions to zero and solving the system of non-linear equations numerically. In order to construct confidence intervals for the parameters, the observed information matrix $J(\boldsymbol{\vartheta})$ is used since the expected information matrix is complicated. The observed information matrix is given by

$$J(\boldsymbol{\vartheta})=-\begin{bmatrix}\frac{\partial^2\ell}{\partial^2\lambda} & \frac{\partial^2\ell}{\partial\lambda\partial c} & \frac{\partial^2\ell}{\partial\lambda\partial d} & \frac{\partial^2\ell}{\partial\lambda\partial\beta} & \frac{\partial^2\ell}{\partial\lambda\partial\theta} & \frac{\partial^2\ell}{\partial\lambda\partial\alpha}\\ & \frac{\partial^2\ell}{\partial^2 c} & \frac{\partial^2\ell}{\partial c\partial d} & \frac{\partial^2\ell}{\partial c\partial\beta} & \frac{\partial^2\ell}{\partial c\partial\theta} & \frac{\partial^2\ell}{\partial c\partial\alpha}\\ & & \frac{\partial^2\ell}{\partial^2 d} & \frac{\partial^2\ell}{\partial d\partial\beta} & \frac{\partial^2\ell}{\partial d\partial\theta} & \frac{\partial^2\ell}{\partial d\partial\alpha}\\ & & & \frac{\partial^2\ell}{\partial^2\beta} & \frac{\partial^2\ell}{\partial\beta\partial\theta} & \frac{\partial^2\ell}{\partial\beta\partial\alpha}\\ & & & & \frac{\partial^2\ell}{\partial^2\theta} & \frac{\partial^2\ell}{\partial\theta\partial\alpha}\\ & & & & & \frac{\partial^2\ell}{\partial^2\alpha}\end{bmatrix},$$

where $\boldsymbol{\vartheta}=(\alpha,\lambda,\beta,\theta,c,d)'$. The explicit expression for the elements of the observed information matrix are available upon request. When the usual regularity conditions are fulfilled and that the parameters are within the interior of the parameter space, but not on the boundary, $\sqrt{n}(\hat{\boldsymbol{\vartheta}}-\boldsymbol{\vartheta})$ converges in distribution to $N_6(\mathbf{0},I^{-1}(\boldsymbol{\vartheta}))$, where $I(\boldsymbol{\vartheta})$ is the expected information matrix. The asymptotic behavior is still valid when $I(\boldsymbol{\vartheta})$ is replaced by the observed information matrix evaluated at $J(\hat{\boldsymbol{\vartheta}})$. The asymptotic multivariate normal distribution $N_6(\mathbf{0},J^{-1}(\hat{\boldsymbol{\vartheta}}))$ can be used to construct an approximate $100(1-\eta)\%$ two-sided confidence intervals for the model parameters, where η is the significance level.

5.1. Monte Carlo simulation

In this sub-section, a simulation study is carried out to examine the average bias (AB) and root mean square error (RMSE) of the maximum likelihood estimators of the parameters of the EGEDD. The experiment was conduct through various simulations for different sample sizes and different parameter values. The quantile function given in Eq. (9) was used to generate random samples from the EGEDD. The simulation experiment was repeated for $N=1000$ times each with sample sizes $n=25,50,75,100,200$ and parameter values $I:\alpha=2.5,\lambda=1.5,\beta=0.4,\theta=0.5,c=1.0,d=0.2$ and $II:\alpha=0.3,\lambda=0.5,\beta=0.8,\theta=0.2,c=0.7,d=1.5$. The AB and the RMSE of the parameters were computed using the following relations:

$$\text{AB}=\frac{1}{N}\sum_{i=1}^{n}\left(\hat{\vartheta}-\vartheta\right),$$

and

$$\text{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{n}\left(\hat{\vartheta}-\vartheta\right)^2},$$

Table 2
Monte Carlo simulation results: AB and RMSE.

		I		II	
Parameter	n	AB	RMSE	AB	RMSE
λ	25	13.724	58.702	17.105	98.897
	50	0.681	12.444	2.634	32.134
	75	0.268	0.980	1.124	27.832
	100	0.204	0.891	0.286	1.249
	200	0.105	0.365	0.187	0.507
α	25	105.848	532.196	40.717	211.657
	50	3.892	59.077	7.484	96.125
	75	0.806	7.613	1.728	45.595
	100	0.195	2.625	0.332	3.204
	200	-0.031	1.268	0.097	0.354
β	25	0.763	2.226	0.030	1.703
	50	1.039	2.960	0.198	1.989
	75	0.891	2.571	0.258	2.138
	100	0.759	2.205	0.259	1.727
	200	0.382	1.089	0.031	1.175
θ	25	-0.041	0.263	0.133	0.247
	50	-0.090	0.221	0.059	0.158
	75	-0.107	0.209	0.033	0.122
	100	-0.109	0.197	0.017	0.110
	200	-0.095	0.158	0.008	0.082
c	25	9.384	41.310	6.311	32.040
	50	0.499	5.113	1.481	21.073
	75	0.254	0.904	0.270	2.625
	100	0.207	0.716	0.222	0.658
	200	0.106	0.323	0.143	0.299
d	25	3.668	0.676	1.950	0.376
	50	0.233	0.062	0.518	0.084
	75	0.155	0.204	0.395	0.064
	100	0.114	0.011	0.471	0.053
	200	0.074	0.008	0.299	0.044

where $\vartheta=\alpha,\lambda,\beta,\theta,c,d$. Table 2 presents the AB and RMSE values of the parameters λ, α, β, θ, c and d for different sample sizes. From the results, it can be seen that as the sample size increases, the RMSE decay towards zero. In addition, the AB decreases as the sample size increases. Hence, the maximum likelihood estimates and their asymptotic properties can be used for constructing confidence intervals even for reasonably small sample size.

6. Applications

In this section, the application of the EGEDD is provided by fitting the distribution to two real data sets. The goodness-of-fit of the EGEDD is compared with that of its sub-models, the exponentiated Kumaraswamy Dagum (EKD) distribution and the McDagum (McD) distribution using Kolmogorov-Smirnov (K-S) statistic and Cramér-von (W*) misses distance values, as well as Akaike information criterion (AIC), corrected Akaike information criterion (AICc) and Bayesian information criterion (BIC). The maximum likelihood estimates of the fitted model parameters were computed by maximizing the log-likelihood function via the subroutine mle2 using the bbmle package in R (Bolker, 2014). This was done using a wide range of initial values. The process often leads to more than one maximum, thus in such situation, the maximum likelihood estimates corresponding to the largest maxima is chosen. In few cases were no maximum is identified for the selected initial values, new sets of initial values are employed in order to get a maximum. The PDF of EKD distribution is given by

$$g(x)=\alpha\lambda\delta\phi\theta x^{-\delta-1}\left(1+\lambda x^{-\delta}\right)^{-\alpha-1}\left[1-\left(1+\lambda x^{-\delta}\right)^{-\alpha}\right]^{\phi-1}$$
$$\times\left\{1-\left[1-\left(1+\lambda x^{-\delta}\right)^{-\alpha}\right]^{\phi}\right\}^{\theta-1}, \quad (23)$$

Table 3
Failure time data on 100 cm yarn subjected to 2.3% strain level.

86	146	251	653	98	249	400	292	131	169
175	176	76	264	15	364	195	262	88	264
157	220	42	321	180	198	38	20	61	121
282	224	149	180	325	250	196	90	229	166
38	337	65	151	341	40	40	135	597	246
211	180	93	315	353	571	124	279	81	186
497	182	423	185	229	400	338	290	398	71
246	185	188	568	55	55	61	244	20	289
393	396	203	829	239	236	286	194	277	143
198	264	105	203	124	137	135	350	193	188

Table 4
Maximum likelihood estimates of parameters and standard errors for yarn data.

Model	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{c}$	$\hat{d}$
EGEDD	0.026	75.310	0.017	3.513	45.692	0.090
	(0.007)	(0.007)	(0.005)	(0.631)	(0.036)	(0.011)
EGDD	1.992		10.480	4.733	75.487	0.223
	(0.251)		(13.022)	(0.587)	(27.669)	(0.032)
DD	19.749		11.599	1.126		
	(10.814)		(5.008)	(0.069)		
EGEBD		35.463	35.965	4.859	15.667	0.070
		(0.271)	(0.120)	(0.666)	(2.714)	(0.011)
EGBD			24.801	4.196	73.9120	0.258
			(15.068)	(1.808)	(22.832)	(0.112)
EGEFD	20.662	34.477		5.217	16.438	0.65
	(2.365)	(0.278)		(0.578)	(2.708)	(0.009)
EGFD	10.537			5.239	21.341	0.140
	(1.115)			(0.429)	(4.089)	(0.015)
	$\hat{\lambda}$	$\hat{\delta}$	$\hat{\beta}$	$\hat{a}$	$\hat{b}$	$\hat{c}$
McD	0.027	0.600	98.780	0.333	25.042	46.276
	(1.848×10^{-2})	(9.647×10^{-2})	(2.180×10^{-5})	(1.504×10^{-1})	(4.507×10^{-4})	(4.654×10^{-5})
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\delta}$	$\hat{\phi}$	$\hat{\theta}$	
EKD	46.109	39.413	5.188	0.203	31.169	
	(1.295)	(5.006)	(0.961)	(0.040)	(11.023)	

Table 5
Log-likelihood, goodness-of-fit statistics and information criteria for yarn data.

Model	ℓ	AIC	AICc	BIC	K-S	W*
EGEDD	**−628.170**	**1268.336**	**1269.553**	**1283.967**	**0.124**	**0.249**
EGDD	−653.070	1316.137	1317.040	1329.163	0.172	0.948
DD	−649.260	1304.517	1304.938	1312.333	0.164	0.821
EGEBD	−630.870	1271.745	1272.648	1284.771	0.136	0.340
EGBD	−653.030	1314.056	1314.694	1324.447	0.174	0.969
EGEFD	−630.760	1271.523	1272.426	1284.549	0.139	0.339
EGFD	−666.880	1341.757	1342.395	1352.177	0.236	0.760
McD	−628.200	1268.399	1269.616	1284.030	0.128	0.285
EKD	−653.960	1317.913	1318.816	1330.938	0.178	0.985

for $\alpha > 0, \lambda > 0, \delta > 0, \phi > 0, \theta > 0, x > 0$, and that of McD distribution is

$$g(x) = \frac{c\beta\lambda\delta x^{-\delta-1}}{B(a,b)}(1+\lambda x^{-\delta})^{-\beta ac-1}\left[1-(1+\lambda x^{-\delta})^{-c\beta}\right]^{b-1}, \quad (24)$$

for $a > 0, b > 0, c > 0, \lambda > 0, \beta > 0, \delta > 0, x > 0$.

6.1. Yarn data

The data in Table 3 represents the time to failure of a 100 cm polyster/viscose yarn subjected to 2.3% strain level in textile experiment in order to assess the tensile fatigue characteristics of the yarn. The data set can be found in Quesenberry and Kent (1982) and Pal and Tiensuwan (2014).

The maximum likelihood estimates of the parameters of the fitted models with their corresponding standard errors in brackets are given in Table 4. All the parameters of the EGEDD are

Table 6
Likelihood ratio test statistic for yarn data.

Model	Hypotheses	LRT	*P*-values
EGDD	$H_0: \lambda = 1$ vs $H_1: H_0$ is false	49.801	< 0.001
DD	$H_0: \lambda = c = d = 1$ vs $H_1: H_0$ is false	42.181	< 0.001
EGEBD	$H_0: \alpha = 1$ vs $H_1: H_0$ is false	5.409	0.020
EGBD	$H_0: \lambda = \alpha = 1$ vs $H_1: H_0$ is false	49.721	< 0.001
EGEFD	$H_0: \beta = 1$ vs $H_1: H_0$ is false	5.187	0.023
EGFD	$H_0: \lambda = \beta = 1$ vs $H_1: H_0$ is false	77.421	< 0.001

Table 7
Failure Times for 36 appliances subjected to an automatic life test.

11	35	49	170	329	381	708	958	1062	1167	1594	1925
1990	2223	2327	2400	2451	2471	2551	2565	2568	2694	2702	2761
2831	3034	3059	3112	3214	3478	3504	4329	6367	6976	7846	13403

Table 8
Maximum likelihood estimates of parameters and standard errors for appliances data.

Model	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{c}$	$\hat{d}$
EGEDD	0.001 (1.000×10^{-4})	27.198 (0.001)	4.560 (0.847)	2.838 (0.123)	20.866 (0.010)	0.070 (0.003)
EGDD	7.977 (0.651)		0.404 (0.044)	3.570 (0.391)	15.862 (5.196)	0.130 (0.021)
DD	0.018 (0.0062)		1495.519 (1.058×10^{-7})	0.509 (0.056)		
EGEBD		25.705 (0.514)	14.152 (0.110)	3.412 (0.247)	8.332 (1.934)	0.047 (0.009)
EGBD			9.504 (3.205)	3.392 (0.388)	11.226 (3.440)	0.129 (0.022)
EGEFD	13.048 (1.817)	27.555 (0.071)		3.561 (0.392)	9.084 (2.186)	0.047 (0.009)
EGFD	8.4843 (1.550)			3.429 (0.711)	16.533 (5.833)	0.143 (0.034)
	$\hat{\lambda}$	$\hat{\delta}$	$\hat{\beta}$	$\hat{a}$	$\hat{b}$	$\hat{c}$
McD	1.427 (0.092)	3.455 (0.212)	1.275 (6.875)	10.505 (56.906)	0.064 (0.012)	500.556 (6.796)
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\delta}$	$\hat{\phi}$	$\hat{\theta}$	
EKD	5.562 (1.517)	12.683 (2.158)	3.716 (0.755)	0.128 (0.029)	11.609 (3.922)	

significant at the 5% significance level. The EGEDD provides a better fit to the yarn data than its sub-models, the McD distribution and the EKD distribution. From Table 5, the EGEDD has the highest log-likelihood and the smallest K-S, W*, AIC, AICc, and BIC values compared to the other models. Although the EGEDD provides the best fit to the data, the McD distribution, EGEBD and EGEFD are alternatively good models for the data since their measures of fit values are close to that of the EGEDD.

In order to make a complete statistical inference about a model, it is imperative to reduce the number of parameters of the model and examine how that affects the ability of the reduce model to fit the data. The likelihood ratio test (LRT) is therefore performed to compare the EGEDD with its sub-models. The LRT statistic and their corresponding *P*-values in Table 6 revealed that the EGEDD provides a good fit than its sub-models.

The asymptotic variance-covariance matrix for the estimated parameters of the EGEDD for the yarn data is given by

$$J^{-1} = \begin{bmatrix} 5.0338 \times 10^{-5} & 2.1232 \times 10^{-5} & 1.3887 \times 10^{-5} & 0.0045 & 2.5246 \times 10^{-4} & -7.5812 \times 10^{-5} \\ 2.1232 \times 10^{-5} & 5.1601 \times 10^{-5} & 1.0316 \times 10^{-5} & 0.0019 & 1.0648 \times 10^{-4} & -2.5628 \times 10^{-5} \\ 1.3887 \times 10^{-5} & 1.0316 \times 10^{-5} & 2.2466 \times 10^{-5} & 0.0012 & 6.9642 \times 10^{-5} & -1.6719 \times 10^{-5} \\ 4.4786 \times 10^{-3} & 1.8887 \times 10^{-3} & 1.2354 \times 10^{-3} & 0.3985 & 2.2462 \times 10^{-2} & -6.7451 \times 10^{-3} \\ 2.5246 \times 10^{-4} & 1.0648 \times 10^{-4} & 6.9642 \times 10^{-5} & 0.0225 & 1.2662 \times 10^{-3} & -3.8023 \times 10^{-4} \\ -7.5812 \times 10^{-5} & -2.5628 \times 10^{-5} & -1.6719 \times 10^{-5} & -0.0067 & -3.8023 \times 10^{-4} & 1.1654 \times 10^{-4} \end{bmatrix}.$$

Thus, the approximate 95% confidence interval for the parameters λ, α, β, θ, c and d of the EGEDD are $[75.296, 75.324]$, $[0.012, 0.040]$, $[0.007, 0.027]$, $[2.276, 4.750]$, $[45.621, 45.763]$ and $[0.068, 0.111]$ respectively.

6.2. Appliances data

The appliances data was obtained from (Lawless, 1982). The data set consists of failure times for 36 appliances subjected to an automatic life test. The data set are given in Table 7.

Table 8 provides the maximum likelihood estimates of the parameters with their corresponding standard errors in brackets for the models fitted to the appliances data. From Table 8, all the parameters of the EGED are significant at the 5% significance level.

From Table 9, it is clear that the EGEDD provides a better fit to the appliances data than the other models. It has the highest log-likelihood and the smallest K-S, W*, AIC, AICc and BIC values. Alternatively, the EGEBD and EGEFD are good models since their goodness-of-fit measures are close to that of the EGEDD.

The LRT was performed in order to compare the EGEDD with its sub-models. From Table 10, the LRT revealed that the EGEDD provides a better fit to the appliances data than its sub-models. Although the LRT favored the EGEFD at the 5% level of significance, the EGEDD was better than it at the 10% level of significance.

Table 9
Log-likelihood, goodness-of-fit statistics and information criteria for appliances data.

Model	ℓ	AIC	AICc	BIC	K-S	W*
EGEDD	**−328.870**	**669.740**	**670.957**	**679.241**	**0.253**	**0.569**
EGDD	−340.910	691.818	692.721	699.736	0.264	0.882
DD	−339.610	685.225	685.646	689.976	0.257	0.858
EGEBD	−330.910	671.823	672.726	679.741	0.272	0.634
EGBD	−341.520	691.037	691.675	697.371	0.268	0.881
EGEFD	−330.730	671.460	672.363	679.377	0.269	0.625
EGFD	−341.030	690.054	690.692	696.388	0.269	0.907
McD	−356.480	724.955	728.950	734.456	0.347	0.986
EKD	−341.650	693.295	694.198	701.213	0.269	0.925

Table 10
Likelihood ratio test statistic for appliances data.

Model	Hypotheses	LRT	P-values
EGDD	$H_0 : \lambda = 1$ vs $H_1 : H_0$ is false	24.078	< 0.001
DD	$H_0 : \lambda = c = d = 1$ vs $H_1 : H_0$ is false	21.486	< 0.001
EGEBD	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	4.084	0.043
EGBD	$H_0 : \lambda = \alpha = 1$ vs $H_1 : H_0$ is false	25.297	< 0.001
EGEFD	$H_0 : \beta = 1$ vs $H_1 : H_0$ is false	3.720	0.054
EGFD	$H_0 : \lambda = \beta = 1$ vs $H_1 : H_0$ is false	24.315	< 0.001

The asymptotic variance-covariance matrix for the estimated parameters of the EGEDD for the appliances data is given by

$$J^{-1} = \begin{bmatrix} 1.7033\times10^{-6} & 1.5346\times10^{-8} & 1.1045\times10^{-3} & 3.7492\times10^{-5} & 1.2695\times10^{5} & -6.6696\times10^{-8} \\ 1.5346\times10^{-8} & 1.4494\times10^{-8} & 8.8310\times10^{-6} & 5.7406\times10^{-6} & 1.1008\times10^{-7} & -8.3473\times10^{-8} \\ 1.1045\times10^{-3} & 8.8310\times10^{-6} & 7.1688\times10^{-1} & 2.1348\times10^{-2} & 8.2348\times10^{-3} & 1.3547\times10^{-5} \\ 3.7492\times10^{-5} & 5.7406\times10^{-6} & 2.1348\times10^{-2} & 1.5185\times10^{-2} & 2.6827\times10^{-4} & -2.8002\times10^{-4} \\ 1.2695\times10^{-5} & 1.1008\times10^{-7} & 8.2348\times10^{-3} & 2.6827\times10^{-4} & 9.4629\times10^{-5} & -2.9359\times10^{-7} \\ -6.6696\times10^{-8} & -8.3473\times10^{-8} & 1.3547\times10^{-5} & -2.8002\times10^{-4} & -2.9359\times10^{-7} & 8.4565\times10^{-6} \end{bmatrix}.$$

Thus, the approximate 95% confidence interval for the parameters λ, α, β, θ, c and d of the EGEDD are [27.1955, 27.2005], [0.0008, 0.0012], [2.9005, 6.2195], [2.5965, 3.0795], [20.8470, 20.8850] and [0.0643, 0.0757] respectively.

7. Conclusion

This study proposed and presented results on the statistical properties of the EGEDD. The EGEDD contains a number of submodels with potential applications to a wide area of probability and statistics. Statistical properties such as the quantile function, moment, entropy, reliability and order statistic were derived. The estimation of the parameters of the model was approached using maximum likelihood estimation and the applications of the EGEDD was also demonstrated to show its usefulness.

Addendum

During the review process, one of the reviewers referred us to a work done by Rezaei et al. (2017), we found out that our proposed CDF for the EGE-X family of distribution possess exactly analogous form with the CDF of their generalized exponentiated class of distribution. However, we conducted our research without any prior knowledge of their work. The content of that paper, is however different from ours.

Competing interests

The authors declare that there is no conflict of interest regarding the publications of this article.

Acknowledgment

The first author wishes to thank the African Union for supporting his research at the Pan African University, Institute for Basic Sciences, Technology and Innovation. The authors wish to thank the Editor-in-chief and the anonymous reviewers for their valuable comments and suggestions that have greatly improved the content of this manuscript.

Appendix A. Appendix

R Algorithm

```
### EGEDD PDF
EGEDD_PDF<-function(x,alpha,lambda,beta,theta,c,d){
A<-(1+alpha*(x^(-theta)))^(-beta-1)
B<-1-(1+alpha*(x^(-theta)))^(-beta)
fxn<-lambda*alpha*beta*theta*c*d*(x^(-theta-1))
  *A*(B^(d-1))*
((1-(B^d))^(c-1))*((1-(1-(B^d))^c)^(lambda-1))
return(fxn)
}
```

```
### EGEDD CDF
EGEDD_CDF<-function(x,alpha, lambda, beta,theta,c,d)
  {
fxn<-1-(1-(1-(1-(1+alpha*(x^(-theta)))^
  (-beta))^d)^c)^lambda
return(fxn)
}
```

```
### EGEDD survival function
EGEDD_Surv<-function(x,alpha,lambda,beta,theta,c,d)
  {
fxn<-(1-(1-(1-(1+alpha*(x^(-theta)))^
  (-beta))^d)^c)^lambda
return(fxn)
}
```

```
### EGEDD Hazard function
EGEDD_Hazard<-function(x,alpha,lambda,beta,theta,c,
  d){
PDF<-EGEDD_PDF(x,alpha,lambda,beta,theta,c,d)
Survival<-EGEDD_Surv(x,alpha,lambda,beta,theta,c,
  d)
hazard<-PDF/Survival
return(hazard)
}
```

```
### EGEDD Quantile function
Quantile<-function(alpha,lambda,beta,theta,c,d,u){
A<-(1-u)^(1/lambda)
B<-(1-A)^(1/c)
C<-(1-B)^(1/d)
D<-(1-C)^(-1/beta)
result<-((1/alpha)*(D-1))^(-1/theta)
return(result)
}
```

```
### EGEDD Moment
EGEDD_Moment<-function(alpha,lambda,beta,theta,c,d){
func<-function(x,alpha,lambda,beta,theta,c,d,r){
(x^r)*(EGEDD_PDF(x,alpha,lambda,beta,theta,c,d))}
results<-integrate(func,lower=0,
  upper=Inf,subdivisions=10000,
alpha=alpha,lambda=lambda,beta=beta,theta=theta,
  c=c,d=d,r=r)
return(results$value)
}
```

```
### Negative Log-likelihood function of EGEDD
EGEDD_LL<-function(alpha,lambda,beta,theta,c,d){
A<-(1+alpha*(x^(-theta)))^(-beta-1)
B<-1-(1+alpha*(x^(-theta)))^(-beta)
fxn<- -sum(log(lambda*alpha*beta*theta*c*d*(x^(-
  theta-1))*A*(B^(d-1))*
((1-(B^d))^(c-1))*((1-(1-(B^d))^c)^(lambda-1))))
return(fxn)
}
```

```
### Fitting EGEDD to Real Data Set
library(bbmle)
fit<-mle2(EGEDD_LL, start=list alpha=alpha,
  lambda=lambda,beta=beta,
theta=theta,c=c,d=d),method=''BFGS",data=list(x))
summary(fit)
### Computing the variance-covariance matrix
vcov(fit)
```

References

Al-Zahrani, B., 2016. Reliability test plan based on Dagum distribution. Int. J. Adv. Stat. Prob. 4 (1), 75–78.

Bolker, B., 2014. Tools for general maximum likelihood estimation. r development core team.

Dagum, C., 1977. New model of personal income distribution specification and estimation. Econ. Appl. 30 (3), 413–437.

Dey, S., Al-Zahrani, B., Basloom, S., 2017. Dagum distribution: properties and different methods of estimation. Int. J. Stat. Prob. 6 (2), 74–92.

Domma, F., Condino, F., 2013. The beta-Dagum distribution: definition and properties. Commun. Stat.-Theory Methods 42 (22), 4070–4090.

Domma, F., Giordano, S., Zenga, M.M., 2011. Maximum likelihood estimation in Dagum distribution with censored sample. J. Appl. Stat. 38 (12), 2971–2985.

Domma, F., Perri, P.F., 2009. Some developments on the log-Dagum distribution. Stat. Methods Appl. 18, 205–209.

Elbatal, I., Aryal, G., 2015. Transmuted Dagum distribution with applicatrions. Chil. J. Stat. 6 (2), 31–45.

Huang, S., Oluyede, B.O., 2014. Exponentiated Kumaraswamy-Dagum distribution with applications to income and lifetime data. J. Stat. Distrib. Appl. 1 (8), 1–20.

Kleiber, C., 2008. A guide to the Dagum distribution. In: Duangkamon, C. (Ed.), Modeling Income Distributions and Lorenz Curves Series: Economics Studies in Inequality, Social Exclusion and Well-being. Springer, New York. vol. 5.

Kleiber, C., Kotz, S., 2003. Statistical Size Distribution in Economics and Actuarial Sciences. John Wiley and Sons.

Lawless, J.F., 1982. Statistical Models and Methods for Lifetime Data. Wiley, New York.

Oluyede, B.O., Huang, S., Pararai, M., 2014. A new generalized Dagum distribution with applications to income and lifetime data. J. Stat. Econ. Methods 3 (2), 125–151.

Oluyede, B.O., Motsewabagale, G., Huang, S., Pararai, M., 2016. The Dagum-Poisson distribution: model, properties and application. Electron. J. Appl. Stat. Anal. 9 (1), 169–197.

Oluyede, B.O., Rajasooriya, S., 2013. The Mc-Dagum distribution and its statistical properties with applications. Asian J. Math. Appl. 44, 1–16.

Oluyede, B.O., Ye, Y., 2014. Weighted Dagum and related distributions. Afrika Matematika 25 (4), 1125–1141.

Pal, M., Tiensuwan, M., 2014. The beta transmuted exponentiated Weibull geometric distribution. Austrian J. Stat. 43 (2), 133–149.

Quesenberry, C.P., Kent, J., 1982. Selecting among probability distributions used in reliability. Technometrics 24 (1), 59–65.

Rényi, A., 1961. On measures of entropy and information. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley, CA, pp. 547–561.

Rezaei, S., Marvasty, A.K., Nadarajah, S., Alizadeh, M.A., 2017. A new exponentiated class of distributions: properties and applications. Commun. Stat.-Theory Methods 46, 6054–6073.

Shahzad, M.N., Asghar, Z., 2013. Comparing TL-moments, L-moments and conventional moments of Dagum distribution by simulated data. Revista Colombiana de Estadistica 36 (1), 79–93.

Silva, A.O., da Silva, L.C.M., Cordeiro, G.M., 2015. The extended Dagum distribution: properties and applications. J. Data Sci. 13, 53–72.