Exponentiated Generalized Geometric Burr Iii Distribution

Suleman Nasiru[1], Peter N. Mwita[2] and Oscar Ngesa[3]

[1]*Pan African University, Institute for Basic Sciences,Technology and Innovation.*

[2]*Machakos University, Department of Mathematics*

[3]*Taita Taveta University, Mathematics and Informatics Department*

Abstract

Statistical distributions play a major role in parametric statistical modeling and inference. However, most of the existing classical distributions do not provide reasonable parametric fits to data sets. Thus, the need to develop generalized versions of these classical distributions has become an issue of interest to many researchers in the field of distribution theory. This study proposes a new generalization of the Burr III distribution called the exponentiated generalized geometric Burr III distribution. Various statistical properties of the distribution such as the quantile function, moment, moment generating function, incomplete moment, mean residual life, entropy, reliability, stochastic orders and order statistics were derived. The method of maximum likelihood estimation was employed to estimate the parameters of the distribution and simulation studies were performed to investigate the properties of the estimators for the parameters of the distribution. The simulation results revealed that the estimators for the parameters were stable as the sample size increases. Application of the distribution was demonstrated using real data set to showits usefulness.

Keywords: Burr III, geometric, quantile function, stochastic orders, order statistics, entropy.

INTRODUCTION

The modification of standard distributions through the induction of extra parameters plays a vital role in the development of new families of distributions with a range of skewness and light and heavy tails. In addition, the induction of parameters has been proved to be imperative in determining tail properties and improving the goodness-of-fit of the resulting distribution (Tahir and Nadarajah, 2015).

The Burr III distribution (Burr, 1942) which is sometimes referred to as the inverse Burr distribution (Klugman et al., 1998) in the actuarial literature and kappa distribution in the meteorological field (Mielke, 1973) has been modified in recent time by a number of researchers to improve its flexibility in modeling lifetime data. The usefulness of the distribution in finance, environmental studies, survival analysis and reliability theory cannot be ignored (see Gove et al., 2008; Lindsay et al., 1996). Some of the modified versions of the Burr III distribution includes: gamma Burr III distribution (Cordeiro et al., 2017), extended Burr III distribution (Cordeiro et al., 2014), beta Burr III distribution (Gomes et al., 2013) and Kumaraswamy Burr III distribution (Behairy et al., 2016).

In this study, a new generalization of the Burr III distribution called exponentiated generalized geometric Burr III (EGGB) distribution is proposed and studied using the exponentiated generalized geometric (EGG) family of distributions developed by Nasiru et al. (2018). The cumulative distribution function (CDF) of the EGG is defined as:

$$F(x) = 1 - \frac{(1-\lambda)\left[1-\left(1-(1-G(x))^d\right)^c\right]}{1-\lambda\left[1-(1-(1-G(x))^d)^c\right]}, c>0,\ d>0,\ 0<\lambda<1, x\in\mathbb{R}, \quad (1)$$

and the corresponding probability density function (PDF) is given by

$$f(x) = \frac{(1-\lambda)cdg(x)(1-G(x))^{d-1}(1-(1-G(x))^d)^{c-1}}{[1-\lambda[1-(1-(1-G(x))^d)^c]]^2}.$$

The rest of the paper is organized as follows: In section 2, the cumulative distribution function (CDF), probability density function (PDF), survival function and hazard function of the EGGB distribution were defined. In section 3, statistical properties of the EGGB were derived. In section 4, the parameters of the new family were estimated using maximum likelihood estimation. In section 5, simulation was performed to examine the finite sample properties of the estimators for the parameters of the EGGB distribution. In section 6, application of the model was demonstrated using real data set. Finally, the concluding remarks of the study were given in section 7.

Generalized Geometric Burr III Distribution

Suppose the random variable X follows the Burr III distribution with CDF

$$G(x) = (1+x^{-\theta})^{-\beta}, \theta>0, \beta>0, x>0. \quad (2)$$

By substituting equation (2) into (1), the CDF of the EGGB is defined as

$$F(x) = 1 - \frac{(1-\lambda)[1-(1-(1-(1+x^{-\theta})^{-\beta})^d)^c]}{1-\lambda[1-(1-(1-(1+x^{-\theta})^{-\beta})^d)^c]}, 0<\lambda<1,\ \theta>0,\ \beta>0,\ c>0,\ d>0,\ x>0.$$

(3)

By differentiating equation (3), the PDF of the EGGB distribution is

$$f(x) = \frac{(1-\lambda)\theta\beta cdx^{-\theta-1}(1+x^{-\theta})^{-\beta-1}(1-(1+x^{-\theta})^{-\beta})^{d-1}[1-(1-(1+x^{-\theta})^{-\beta})^d]^{c-1}}{\{1-\lambda[1-(1-(1-(1+x^{-\theta})^{-\beta})^d)^c]\}^2}, x>0.$$

(4)

Lemma 1. The density function of the EGGB distribution has a mixture representation of the form

$$f(x) = (1-\lambda)cd\sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i}\omega_{ijkl}\,g(x;\theta,\beta_{l+1}),\ x>0, \quad (5)$$

where $g(x;\theta,\beta_{l+1}) = \theta\beta_{l+1}x^{-\theta-1}(1+x^{-\theta})^{-\beta l+1-1}$ is the PDF of the Burr III distribution with parameters θ and $\beta_{l+1} = \beta(l+1)$, and

$$\omega_{ijkl} = \frac{(-1)^{j+k+l}\lambda^i\Gamma(i+2)\Gamma(i+1)\Gamma(c(j+1))\Gamma(d(k+1))}{i!j!K!(l+1)!\Gamma(i-j+1)\Gamma(c(j+1)-k)\Gamma(d(k+1)-l)}.$$

Proof. For a real non-integer $\eta>0$, the following identities hold:

$$(1-z)^{\eta-1} = \sum_{i=0}^{\infty}\frac{(-1)^i\,\Gamma(\eta)}{i!\,\Gamma(\eta-i)}z^i,\ |z|<1, \quad (6)$$

and

$$(1+z)^{-\eta} = \sum_{k=0}^{\infty}\frac{(-1)^k\,\Gamma(\eta+k)}{k!\,\Gamma(\eta)}z^k,\ |z|<1. \quad (7)$$

Using equations (6) and (7), and the fact that $0<(1+x^{-\theta})^{-\beta}<1$, the PDF of the EGGB distribution can be written as

$$f(x) = (1-\lambda)cd\sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i}\omega_{ijkl}\,g(x;\theta,\beta_{l+1}).$$

The linear representation of the EGGB density given in lemma 1 revealed that the EGGB distribution is a linear combination of Burr III distribution with different shape parameters. The expansion of the density is vital in deriving the mathematical properties of the EGGB distribution. The density function plot of the EGGB distribution is given in Figure 1. The density exhibits right skewed shape with varied degree of skewness and kurtosis.

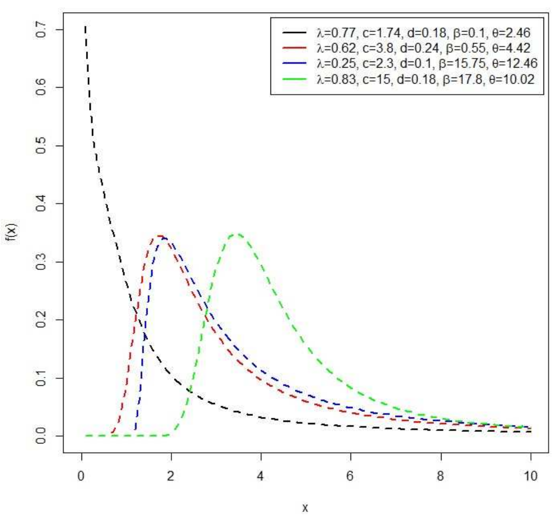

Figure 1: Plot of EGGB density function

The survival and hazard functions of the EGGB distribution are given by

$$S(x) = \frac{(1-\lambda)[1-(1-(1-(1+x^{-\theta})^{-\beta})^{d})^{c}]}{1-\lambda[1-(1-(1-(1+x^{-\theta})^{-\beta})^{d})^{c}]},\ x>0. \quad (8)$$

and

$$h(x) = \frac{(1-\lambda)\theta\beta c d x^{-\theta-1}(1+x^{-\theta})^{-\beta-1}(1-(1+x^{-\theta})^{-\beta})^{d-1}[1-(1-(1+x^{-\theta})^{-\beta})^{d}]^{c-1}}{\{1-[1-(1-(1+x^{-\theta})^{-\beta})^{d}]^{c}\}\{1-\lambda[1-(1-(1-(1+x^{-\theta})^{-\beta})^{d})^{c}]\}},\ x>0. \quad (9)$$

respectively. Figure 2 shows different plots of the hazard function of the EGGB distribution. From the figure, the hazard function exhibits unimodal shape and bathtub followed by upside down bathtub shape.

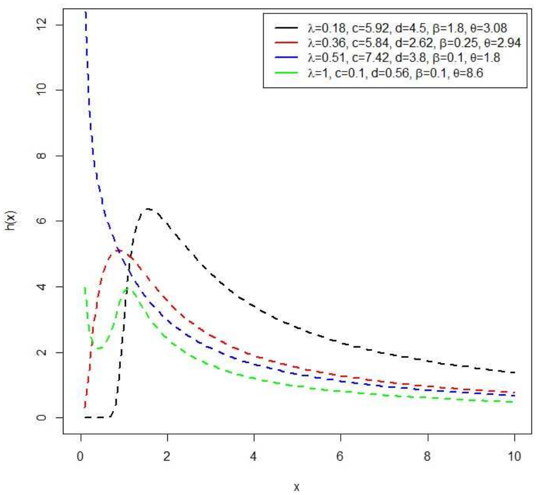

Figure 2: Plot of EGGB hazard function

Statistical Properties

Statistical properties of the EGGB distribution were derived in this section.

Quantile Function

The quantile function provides an alternative means for describing the shapes of a distribution and is vital when generating random numbers.

Lemma 2. For $u \in [0,\ 1]$, the quantile function of the EGGB distribution is given by

$$Q_X(u) = \left\{ \left[1 - \left(1 - \left(\frac{u(1-\lambda)}{1-u\lambda} \right)^{\frac{1}{c}} \right)^{\frac{1}{d}} \right]^{-\frac{1}{\beta}} - 1 \right\}^{-\frac{1}{\theta}} . \quad (10)$$

Proof. By definition, the quantile function is given by $F(x_u) = \mathrm{P}(X \leq x_u) = u, u \in [0,\ 1]$. Thus,

$$1 - \frac{(1-\lambda)[1-(1-(1-(1+x_u^{-\theta})^{-\beta})^d)^c]}{1-\lambda[1-(1-(1-(1+x_u^{-\theta})^{-\beta})^d)^c]} = u. \quad (11)$$

Replacing x_u with $Q_X(u)$ in equation (11) and solving for $Q_X(u)$ yields the quantile function. The first quartile, median and upper quartile of the EGGB random variable can easily be obtained by substituting $u = 0.25,\ 0.5,\ 0.75$ respectively into the quantile function.

Moments

This subsection presents the moment of the EGGB random variable.

Proposition 1. Suppose the random variable X follows the EGGB distribution. Then the r^{th} non-central moment is given by

$$\mu_r^{'} = \sum_{i,\,k,\,l=0}^{\infty} \sum_{j=0}^{i} \omega_{ijkl} \beta_{l+1} B\left(\beta_{l+1} + \frac{r}{\theta},\, 1 - \frac{r}{\theta} \right),\ r < \theta, \quad (12)$$

where $B(a\,,\,b) = \int_0^1 y^{a-1}(1-y)^{b-1}dy$ is the beta function and $r = 1,\ 2,\ \ldots$.

Proof. By definition

$$\mu_r' = \int_0^\infty x^r f(x)dx$$
$$= \int_0^\infty x^r (1-\lambda)cd \sum_{i,k,l=0}^{\infty} \sum_{j=0}^{i} \omega_{ijkl} \mathrm{g}(x;\theta,\beta_{l+1})dx$$
$$= (1-\lambda)cd \sum_{i,k,l=0}^{\infty} \sum_{j=0}^{i} \omega_{ijkl} \int_0^\infty x^r \mathrm{g}(x;\theta,\beta_{l+1})dx$$
$$= \sum_{i,\,k,\,l=0}^{\infty} \sum_{j=0}^{i} \omega_{ijkl} \beta_{l+1} B\left(\beta_{l+1}+\frac{r}{\theta},\, 1-\frac{r}{\theta}\right),\ r<\theta.$$

Moment Generating Function

Proposition 2. The moment generating function of the EGGB random variable is given by

$$M_X(z) = (1-\lambda)cd \sum_{i,k,l=0}^{\infty} \sum_{r=0}^{\infty} \sum_{j=0}^{i} \frac{\omega_{ijkl}\beta_{l+1}z^r}{r!} B\left(\beta_{l+1}+\frac{r}{\theta},\, 1-\frac{r}{\theta}\right),\ r<\theta. \tag{13}$$

Proof. By definition

$$M_X(z) = E(e^{zX}) = \int_0^\infty e^{zx} f(x)dx.$$

Using Taylor series expansion,

$$M_X(z) = \sum_{r=0}^{\infty} \frac{z^r}{r!} \int_0^\infty x^r f(x)dx$$
$$= (1-\lambda)cd \sum_{i,k,l=0}^{\infty} \sum_{r=0}^{\infty} \sum_{j=0}^{i} \frac{\omega_{ijkl}\beta_{l+1}z^r}{r!} B\left(\beta_{l+1}+\frac{r}{\theta},\, 1-\frac{r}{\theta}\right),\ r<\theta.$$

Incomplete Moment

The incomplete moment is useful when computing mean deviation, median deviation, mean residual life and measures of income inequalities. In this subsection, the incomplete moment of the EGGB random variable is derived. Proposition 3. The incomplete moment of the EGGB random variable is

$$M_r(x) = (1-\lambda)cd \sum_{i,k,l=0}^{\infty} \sum_{j=0}^{i} \omega_{ijkl} \beta_{l+1} B\left((1+x^{-\theta})^{-1}; \beta_{l+1}+\frac{r}{\theta}, 1-\frac{r}{\theta}\right),\ r<\theta, \tag{14}$$

where $B(q;a,b) = \int_0^q y^{a-1}(1-y)^{b-1}dy$ is the incomplete beta function and $r = 1, 2, \ldots$.

Proof. Using the identity

$$B(q;a,b) = \int_0^q y^{a-1}(1-y)^{b-1}dy,$$

and the technique for proving the moment, the incomplete moment of the EGGB distribution is

$$\begin{aligned} M_r(x) &= E(X^r|X \le x) \\ &= \int_0^x u^r g(u)du \\ &= (1-\lambda)cd \sum_{i,k,l=0}^{\infty} \sum_{j=0}^{i} \omega_{ijkl}\beta_{l+1} \int_0^{(1+x^{-\theta})^{-1}} y^{\beta_{l+1}+\frac{r}{\theta}-1}(1-y)^{(1-\frac{r}{\theta})-1}dy \\ &= (1-\lambda)cd \sum_{i,k,l=0}^{\infty} \sum_{j=0}^{i} \omega_{ijkl}\beta_{l+1} B\left((1+x^{-\theta})^{-1}; \beta_{l+1}+\frac{r}{\theta}, 1-\frac{r}{\theta}\right),\ r<\theta. \end{aligned}$$

Mean Residual Lifetime

The residual lifetime of a system when it is still operating at time *t*, is $X_t = X - t|X > t$ with PDF $f(x;t) = \frac{f(x)}{1-F(t)}$.

Proposition 4. The mean residual lifetime of EGGB random variable is given by

$$m(t) = \frac{\mu - (1-\lambda)cd\sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i}\omega_{ijkl}\beta_{l+1}B\left((1+t^{-\theta})^{-1};\beta_{l+1}+\frac{1}{\theta},1-\frac{1}{\theta}\right)}{1-F(t)} - t,\ \mu=\mu_1^{'},\ \theta<1 \quad (15)$$

Proof. By definition

$$\begin{aligned} m(t) &= E(X-t|X>t) \\ &= \frac{\int_t^{\infty}(x-t)f(x)dx}{1-F(t)} \\ &= \frac{\mu_1^{'} - \int_0^t f(x)dx}{1-F(t)} - t \\ &= \frac{\mu - (1-\lambda)cd\sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i}\omega_{ijkl}\beta_{l+1}B\left((1+t^{-\theta})^{-1};\beta_{l+1}+\frac{1}{\theta},1-\frac{1}{\theta}\right)}{1-F(t)} - t,\ \theta<1. \end{aligned}$$

The $\int_0^t xf(x)dx$ is the first incomplete moment.

Entropy

The entropy of a random variable is simply a measure of variation. It has been used extensively in the science, engineering and probability theory (R´enyi, 1961).

Proposition 5. The R´enyi entropy of the EGGB random variable is given by

$$I_R(\delta) = \frac{1}{1-\delta}\log\left[((1-\lambda)\beta cd)^{\delta}\theta^{\delta-1}\sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i}\varpi_{ijkl}B\left(\beta(\delta+l)+\frac{1-\delta}{\theta},\delta+\frac{\delta-1}{\theta}\right)\right] \quad (16),$$

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

$$\varpi_{ijkl} = \frac{(-1)^{j+k+l}\lambda^i\Gamma(2\delta+i)\Gamma(i+1)\Gamma(c(\delta+j)-\delta+1)\Gamma(d(\delta+k)-\delta+1)}{i!j!k!l!\Gamma(2\delta)\Gamma(i-j+1)\Gamma(c(\delta+j)-\delta-k+1)\Gamma(d(\delta+k)-\delta-l+1)}.$$

Proof. By definition

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

Using the concepts for expanding the density
∞ i
fδ(x) = ((1 − λ)θβcd)δx−δ(θ+1) XX$ijkl(1 + x−θ)−β(δ+l)−δ.
i,k,l=0 j=0
Hence

$$I_R(\delta) = \frac{1}{1-\delta}\log\left[(1-\lambda)\theta\beta cd)^\delta \sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i}\varpi_{ijkl}\int_0^\infty x^{-\delta(\theta+1)}(1+x^{-\theta})^{-\beta(\delta+l)-\delta}dx\right]$$

$$= \frac{1}{1-\delta}\log\left[((1-\lambda)\beta cd)^\delta\theta^{\delta-1}\sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i}\varpi_{ijkl}B\left(\beta(\delta+l)+\frac{1-\delta}{\theta}, \delta+\frac{\delta-1}{\theta}\right)\right],$$

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

Reliability

Suppose X_1 is the strength of a system and X_2 is the stress, then the component fails when $X_1 \leq X_2$. The estimate of stress-strength reliability of the system R is $P(X_2 < X_1)$.

Proposition 6. If $X_1 \sim$ EGGB$(\lambda,\theta,\beta,c,d)$ and $X_2 \sim$ EGGB$(\lambda,\theta,\beta,c,d)$, then the stress-strength reliability estimate is given by

$$R = 1-(1-\lambda)^2 cd\sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i+1}\frac{\nu_{ijkl}}{(l+1)}, \quad (17)$$

where

$$\nu_{ijkl} = \frac{(-1)^{j+k+l}\lambda^i\Gamma(i+3)\Gamma(i+2)\Gamma(c(j+1))\Gamma(d(k+1))}{2i!j!k!l!\Gamma(i-j+2)\Gamma(c(j+1)-k)\Gamma(d(k+1)-l)}.$$

Proof. By definition

$$R = \int_0^\infty f(x)F(x)dx$$

$$= 1-\int_0^\infty f(x)S(x)dx$$

$$= 1-(1-\lambda)^2\theta\beta cd\sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i}\nu_{ijkl}\int_0^\infty x^{-\theta-1}(1+x^{-\theta})^{-\beta(l+1)-1}dx$$

$$= 1-(1-\lambda)^2 cd\sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i+1}\frac{\nu_{ijkl}}{(l+1)}.$$

Stochastic Ordering Property

The simplest way of showing ordering mechanism in lifetime distribution is through stochastic ordering.

Proposition 7. Suppose X_1 follows the EGGB distribution and X_2 follows the exponentiated generalized Burr III (EGB) distribution, that is $X_1 \sim EGGB(\lambda,\theta,\beta,c,d)$ and$X_2 \sim EGB(\theta,\beta,c,d)$. Then X_1 is smaller than X_2 in likelihood ratio order.

Proof.

$$f_{X_1}(x) = \frac{(1-\lambda)\theta\beta c d x^{-\theta-1}(1+x^{-\theta})^{-\beta-1}(1-(1+x^{-\theta})^{-\beta})^{d-1}[1-(1-(1+x^{-\theta})^{-\beta})^{d}]^{c-1}}{\{1-\lambda[1-(1-(1-(1+x^{-\theta})^{-\beta})^{d})^{c}]\}^{2}},\ x>0.$$

and

$$f_{X_2}(x) = \theta\beta c d x^{-\theta-1}(1+x^{-\theta})^{-\beta-1}(1-(1+x^{-\theta})^{-\beta})^{d-1}[1-(1-(1+x^{-\theta})^{-\beta})^{d}]^{c-1},\ x>0.$$

Thus

$$\frac{f_{X_1}(x)}{f_{X_2}(x)} = \frac{(1-\lambda)}{\{1-\lambda[1-(1-(1-(1+x^{-\theta})^{-\beta})^{d})^{c}]\}^{2}},\ x>0.$$

The derivative of the ratio of the densities yields

$$\frac{d}{dx}\frac{f_{X_1}(x)}{f_{X_2}(x)} = \frac{-2\lambda(1-\lambda)\theta\beta c d x^{-\theta-1}(1+x^{-\theta})^{-\beta-1}(1-(1+x^{-\theta})^{-\beta})^{d-1}[1-(1-(1+x^{-\theta})^{-\beta})^{d}]^{c-1}}{\{1-\lambda[1-(1-(1-(1+x^{-\theta})^{-\beta})^{d})^{c}]\}^{3}}.$$

Since $\frac{d}{dx}\frac{f_{X_1}(x)}{f_{X_2}(x)}$ is a decreasing function for $x>0$, $X_1 \leq_{lr} X_2$. From proposition 7, the hazard rate order, the stochastic order and the mean residual life order between X_1 and X_2 hold.

Order Statistics

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

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

where $F(x)$ and $f(x)$ are the CDF and PDF of the EGGB distribution respectively and $B(\cdot,\cdot)$ is the beta function. Using the binomial series expansion and the fact that $0 < F(x) < 1$ for $x>0$, yields

$$f_{p:n}(x) = \frac{n!}{(p-1)!(n-p)!}\sum_{i=0}^{n-p}(-1)^{i}\binom{n-p}{i}[F(x)]^{p+i-1}f(x). \qquad (18)$$

Substituting equation (3) into (18), the PDF of the p^{th} order statistic $X_{p:n}$ of the EGGB distribution is defined as

$$f_{p:n}(x) = \sum_{i=0}^{n-p}\frac{n!(-1)^{i}}{(p-1)!(n-p)!}\binom{n-p}{i}f(x)\left\{1-\frac{(1-\lambda)[1-(1-(1-(1+x^{-\theta})^{-\beta})^{d})^{c}]}{1-\lambda[1-(1-(1-(1+x^{-\theta})^{-\beta})^{d})^{c}]}\right\}^{p+i-1} \qquad (19)$$

Parameter Estimation

In this section, the maximum likelihood estimators for the parameters of the EGGB model were determined. Let $X_1, X_2, \ldots, X_n$ be random sample of size n from the EGGB distribution. Let $z_i = (1+x_i^{-\theta})^{-\beta}$ and $\bar{z}_i = 1-(1+x_i^{-\theta})^{-\beta}$, then the total log-likelihood for the complete sample is given by

$$\ell = n\log((1-\lambda)\theta\beta c d) - (\theta+1)\sum_{i=1}^{n}\log(x_i) - (\beta+1)\sum_{i=1}^{n}\log(1+x_i^{-\theta}) + (d-1)\sum_{i=1}^{n}\log(\bar{z}_i) + (c-1)\sum_{i=1}^{n}\log(1-\bar{z}_i^{d}) - 2\sum_{i=1}^{n}\log[1-\lambda(1-(1-\bar{z}_i^{d})^{c})]. \qquad (20)$$

Finding the partial derivatives of the log-likelihood function with respect to the parameters gives the components of the score function as:

$$\frac{\partial \ell}{\partial \lambda} = -\frac{n}{1-\lambda} - 2\sum_{i=1}^{n} \frac{(1-\bar{z}_i^d)^c - 1}{1-\lambda(1-(1-\bar{z}_i^d)^c)}, \quad (21)$$

$$\frac{\partial \ell}{\partial c} = \frac{n}{c} + \sum_{i=1}^{n} \log(1-\bar{z}_i^d) - 2\sum_{i=1}^{n} \frac{\lambda(1-\bar{z}_i^d)^c \log(1-\bar{z}_i^d)}{1-\lambda(1-(1-\bar{z}_i^d)^c)}, \quad (22)$$

$$\frac{\partial \ell}{\partial d} = \frac{n}{d} + \sum_{i=1}^{n} \log(\bar{z}_i) - (c-1)\sum_{i=1}^{n} \frac{\bar{z}_i^d \log(\bar{z}_i)}{1-\bar{z}_i^d} + 2\sum_{i=1}^{n} \frac{\lambda c \bar{z}_i^d (1-\bar{z}_i^d)^{c-1} \log(\bar{z}_i)}{1-\lambda(1-(1-\bar{z}_i^d)^c)}, \quad (23)$$

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} - \sum_{i=1}^{n} \log(1+x_i^{-\theta}) + (d-1)\sum_{i=1}^{n} \frac{z_i \log(1+x_i^{-\theta})}{\bar{z}_i} - (c-1)\sum_{i=1}^{n} \frac{d z_i \bar{z}_i^{d-1} \log(1+x_i^{-\theta})}{1-\bar{z}_i^d} +$$
$$2\sum_{i=1}^{n} \frac{\lambda c d z_i \bar{z}_i^{d-1} (1-\bar{z}_i^d)^{c-1} \log(1+x_i^{-\theta})}{1-\lambda(1-(1-\bar{z}_i^d)^c)}, \quad (24)$$

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} - \sum_{i=1}^{n} \log(x_i) + (\beta+1)\sum_{i=1}^{n} \frac{x_i^{-\theta} \log(x_i)}{1+x_i^{-\theta}} - (d-1)\sum_{i=1}^{n} \frac{\beta x_i^{-\theta}(1+x_i^{-\theta})^{-\beta-1} \log(x_i)}{\bar{z}_i} +$$
$$(c-1)\sum_{i=1}^{n} \frac{\beta d x_i^{-\theta} \bar{z}_i^{d-1}(1+x_i^{-\theta})^{-\beta-1} \log(x_i)}{1-\bar{z}_i^d} - 2\sum_{i=1}^{n} \frac{\lambda \beta c d x_i^{-\theta} \bar{z}_i^{d-1}(1-\bar{z}_i^d)^{c-1}(1+x_i^{-\theta})^{-\beta-1} \log(x_i)}{1-\lambda(1-(1-\bar{z}_i^d)^c)}, \quad (25)$$

Setting equations (21) to (25) to zero and solving them simultaneously yield the maximum likelihood estimates for the model parameters. The equations do not have a closed form and have to be solved using numerical techniques such as the quasiNewton algorithms. In order to construct confidence intervals for the parameters, a 5 × 5 observed information matrix can be obtained as $J = -\left\{\frac{\partial^2 \ell}{\partial q \partial r}\right\}$(for $q,\ r = \lambda, c, d, \beta, \theta$), whose element can estimated numerically.

Monte Carlo Simulation

To investigate the finite sample properties of the maximum likelihood estimators for the parameters of the EGGB distribution, simulation studies were performed. The results of the simulation were obtained from 1,000 Monte Carlo repetitions. In each repetition, a random sample of size $n = 25, 50, 75$ and 100 were generated from the EGGB distribution. The simulation results revealed that the average estimates (AE) were quite close to the actual values, the average bias (AB) and root mean square error (RMSE) for the parameters were small and decay towards zero on average as the sample size increases. Hence, it can be concluded from the results that the estimates of the parameters are stable and their asymptotic properties can be used for constructing confidence intervals and regions even for a reasonably small sample size.

λ	c	d	β	θ	n	Parameters	AE	Bias	RMSE
						λ	0.084	-0.016	0.095
						c	0.606	0.106	0.136
						d	0.752	-0.048	0.187
						β	0.311	-0.089	0.152
						θ	0.343	0.043	0.003
					50	λ	0.092	-0.008	0.094
						c	0.599	0.099	0.137
						d	0.751	-0.049	0.171
						β	0.314	-0.086	0.146
						θ	0.344	0.044	0.003
					75	λ	0.100	0.001	0.095
						c	0.587	0.087	0.136
						d	0.757	-0.043	0.167
						β	0.320	-0.080	0.141
						θ	0.343	0.043	0.002
					100	λ	0.103	0.003	0.094
						c	0.584	0.084	0.134
						d	0.755	-0.045	0.157
						β	0.321	-0.079	0.136
						θ	0.342	0.042	0.002
						λ	0.734	-0.066	0.191
						c	0.552	-0.148	0.183
						d	0.312	0.012	0.109
						β	2.796	0.796	1.060
						θ	0.507	0.007	0.003
					50	λ	0.726	-0.074	0.177
						c	0.558	-0.142	0.184
						d	0.312	0.012	0.102
						β	2.691	0.691	0.987
						θ	0.512	0.012	0.003
					75	λ	0.732	-0.068	0.166
						c	0.562	-0.138	0.182
						d	0.319	0.019	0.093
						β	2.694	0.694	0.980
						θ	0.503	0.003	0.003
					100	λ	0.741	-0.059	0.153
						c	0.570	-0.130	0.174
						d	0.321	0.021	0.088
						β	2.684	0.684	0.972
						θ	0.498	-0.002	0.003
						λ	0.369	-0.131	0.309
						c	1.531	0.031	0.396
						d	2.846	0.346	0.484
						β	1.299	0.099	0.244
						θ	3.464	-0.036	0.018
					50	λ	0.384	-0.116	0.296
						c	1.493	-0.007	0.397
						d	2.795	0.295	0.477
						β	1.312	0.112	0.245
						θ	3.476	-0.025	0.016
					75	λ	0.415	-0.085	0.267
						c	1.501	0.001	0.381
						d	2.784	0.284	0.475
						β	1.311	0.111	0.238
						θ	3.442	-0.058	0.016
					100	λ	0.439	-0.061	0.248
						c	1.484	-0.016	0.378
						d	2.740	0.240	0.477
						β	1.319	0.119	0.240

Table 1: Simulation results: AE, AB and RMSE

0.1	0.5	0.8	0.4	0.3	25
0.8	0.7	0.3	2.0	0.5	25
0.5	1.5	2.5	1.2	3.5	25

Application

In this section, the usefulness of the EGGB distribution was demonstrated empirically by means of a real data set. The performance of the EGGB model with regards to providing an appropriate parametric fit to the dataset was compared to that of the extended Burr III (EBIII) distribution (Cordeiro et al., 2014) and beta Burr III (BBIII) distribution (Gomes et al., 2013) using the Akaike information criterion (AIC), corrected Akaike information criterion (AICc) and Bayesian information criterion (BIC). The maximum likelihood estimates for the parameters of the fitted models were obtained by maximizing the log-likelihood function via the subroutine *mle2* uisng the *bbmle* package in R (Bolker, 2014). The PDFs of the EBIII and BBIII distributions are:

$$f(x) = \frac{\alpha\beta ab}{s(x/s)^{\alpha+1}}\left(\frac{(x/s)^{\alpha}}{1+(x/s)^{\alpha}}\right)^{\beta+1}\frac{\left[1-(1+(x/s)^{-\alpha})^{-\beta}\right]^{a-1}}{\{1-[1-(1+(x/s)^{-\alpha})^{-\beta}]^{a}\}^{1-b}}$$

$$\alpha>0,\ \beta>0,\ a>0,\ b>0,\ s>0,\ x>0,$$

,

and

$$f(x) = \frac{\alpha\beta}{s(x/s)^{\alpha+1}B(c,\,d)}\left(\frac{(x/s)^{\alpha}}{1+(x/s)^{\alpha}}\right)^{\beta c+1}\left[1-\left(\frac{(x/s)^{\alpha}}{1+(x/s)^{\alpha}}\right)^{\beta}\right]^{d-1}$$

$$\alpha>0,\ \beta>0,\ c>0,\ d>0,\ s>0,\ x>0,$$

,

respectively. The data set comprises the survival times in weeks, of 33 patients suffering from acute Myelogeneous Leukaemia. The data set was previously analyzed by Feigl and Zelen (1965) and are: 65, 156, 100, 134, 16, 108, 121, 4, 39, 143, 56, 26, 22, 1, 1, 5, 65, 56, 65, 17, 7, 16, 22, 3, 4, 2, 3, 8, 4, 3, 30, 4, 43. Table 2 shows the descriptive statistics of the data. The minimum and maximum survival times were 1.000 and 156.000 weeks respectively. The average survival time was 40.879 weeks with a standard deviation of 46.703 weeks. The survival time was right skewed with a coefficient of skewness of 1.165 week. The distribution of the survival time is fat-tailed with coefficient of kurtosis of 3.122 weeks.

Table 2: Descriptive Statistics

Statistic	Value
Mean	40.879
Median	22.000
Minimum	1.000

Maximum	156.000
Standard deviation	46.703
Skewness	1.165
Kurtosis	3.122

Further exploratory analysis of the data using the total time on test (TTT) transform plot revealed that the data exhibit a bathtub failure rate since the TTT curve is first convex below the 45 degrees line and the followed by a concave shape above it as shown in Figure 3.

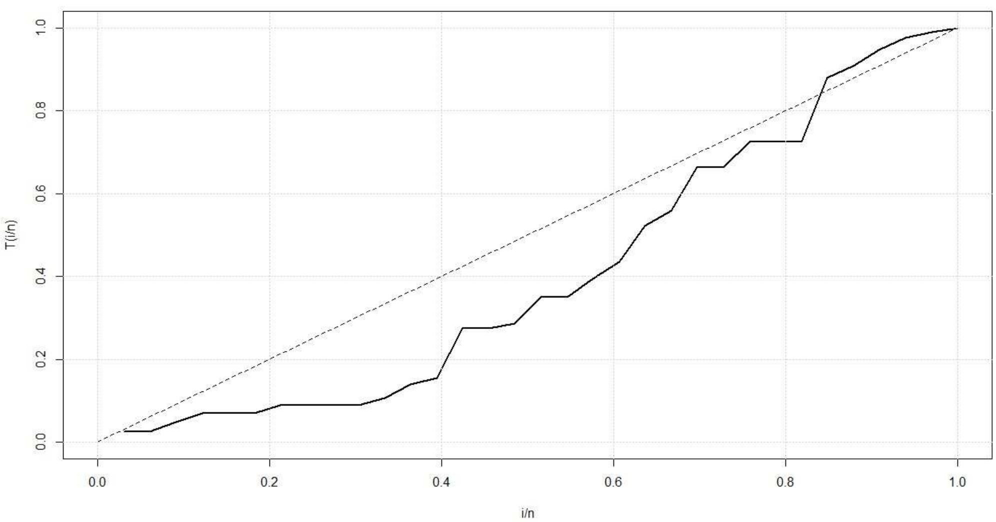

Figure 3: TTT plot of Myelogeneous Leukaemia data

Table 3 displays the maximum likelihood estimates for the parameters of the fitted models with their corresponding standard errors in parentheses.

Table 3: Maximum likelihood estimates and corresponding standard errors in parentheses

Distribution	Parameter estimates				
EGGB(λ, c, d, β, θ)	0.119	3.415	0.094	0.088	6.014
	(0.671)	(1.011)	(0.031)	(0.013)	(0.913)
EBIII(α, β, a, b, s)	5.208	400.523	0.101	2.144	0.280
	(2.130×10^{-1})	(7.105×10^{-4})	(1.848×10^{-2})	(5.418×10^{-1})	(1.253×10^{-2})
BBIII(α, β, c, d, s)	1.652	86.587	0.002	20.737	498.784
	(7.308×10^{-1})	(4.647×10^{-3})	1.057×10^{-3}	(8.598×10^{-4})	(1.093×10^{-3})

The EGGB distribution provides a reasonable parametric fit to the survival data than the EBIII and BBIII distributions as shown in Table 4. From Table 4, the EGGB model has the highest log-likelihood and the smallest AIC, AICc and BIC values compared to the other candidate models.

Table 4: Goodness-of-fit statistics

Distribution	log-likelihood	AIC	AICc	BIC
EGGB	-157.480	324.960	327.160	332.442
EBIII	-158.070	326.140	328.340	333.623
BBIII	-167.360	344.720	346.920	352.203

The asymptotic variance-covariance matrix of the maximum likelihood estimates for the parameters of the EGGB distribution is given by

$$J^{-1} = \begin{pmatrix} 0.450 & 0.145 & -0.014 & -0.001 & -0.047 \\ 0.145 & 1.021 & 0.012 & -0.004 & -0.316 \\ -0.014 & 0.012 & 0.001 & -1.355\times10^{-4} & -0.013 \\ -0.001 & -0.004 & -1.355\times10^{-4} & 1.776\times10^{-4} & 0.009 \\ -0.047 & -0.316 & -0.013 & 0.009 & 0.833 \end{pmatrix}.$$

Figure 4 displays the plot of the empirical density and the fitted densities of the distributions to the data.

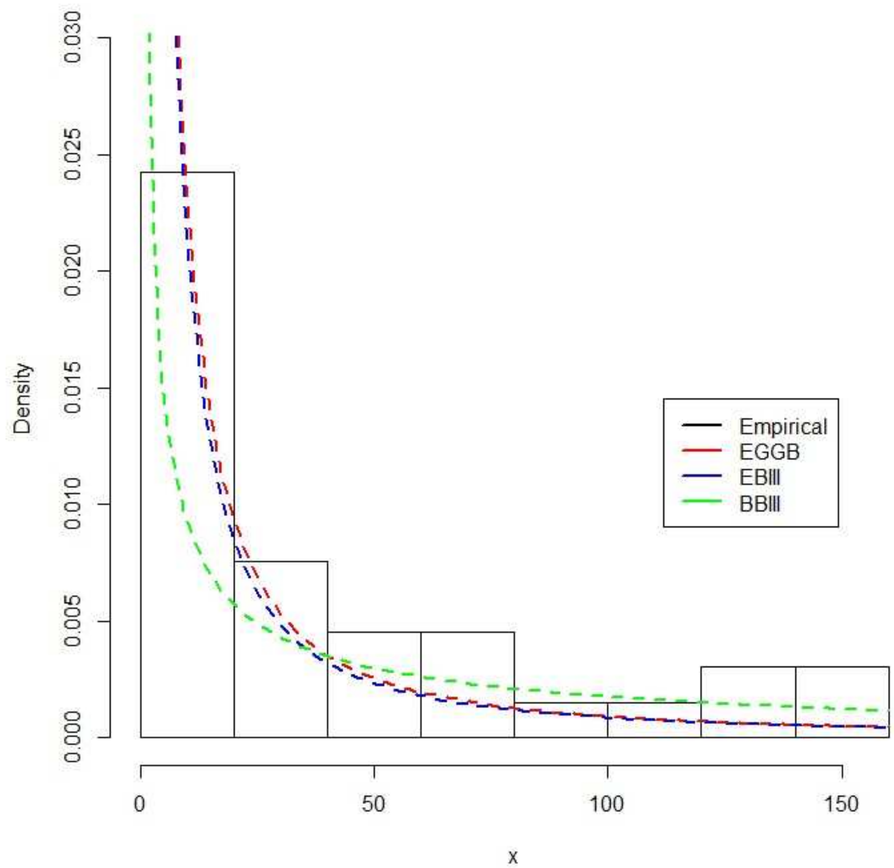

Figure 4: Empirical and fitted densities plot of Myelogeneous Leukaemia data

CONCLUSION

In this study, a new generalization of the Burr III distribution called the exponentiated generalized geometric Burr III distribution was developed. Statistical properties of the model such as the moments, moment generating function, incomplete moment, stochastic ordering, order statistics among others were derived. The maximum likelihood method was used to estimate the parameters of the model and simulation studies were performed to examine the finite sample properties of the estimators of the model parameters. Finally, the usefulness of the distribution was demonstrated empirically using a survival data.

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.
Competing interests
The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

Behairy, S. M., Al-Dayian, G. R., and El-Helbawy, A. A. (2016). The KumaraswamyBurr type III distribution: properties and estimation. *British Journal of Mathematics and Computer Science*, 14(2):1–21.

Bolker, B. (2014). Tools for general maximum likelihood estimation. R development core team.

Burr, I. W. (1942). Cumulative frequency functiions. *Annals of Mathematical Statistics*, 13(2):215–232.

Cordeiro, G. M., Gomes, A. E., and da Silva, C. Q. (2014). Another extended Burr III model: some properties and applications. *Journal of Statistical Computation and Simulation*, 84(12):2524–2544.

Cordeiro, G. M., Gomes, A. E., da Silva, C. Q., and M., O. E. M. (2017). A useful extension of the Burr III distribution. *Journal of Statistical Distributions and Applications*, 4:2–15.

Feigl, P. and Zelen, M. (1965). Estimation of exponential probabilities with concomitant information. *Biometrics*, 21:826–838.

Gomes, A. E., da Silva, C. Q., Cordeiro, G. M., and M., O. E. M. (2013). The beta Burr III model for lifetime data. *Brazilian Journal of Probability and Statistics*, 27:502–543.

Gove, J. H., Ducey, M. J., Leak, W. B., and Zhang, L. (2008). Rotated sigmoid structures in managed uneven-aged northern hardwood stands: a look at the Burr type III distribution. *Forestry*, 81(2):161–176.

Klugman, S. A., Panjer, H. H., and Willmot, G. E. (1998). *Loss models*. Wiley New York.

Lindsay, S. R., Wood, G. R., and Woollon, R. C. (1996). Modelling the diameter distribution of forest stands using the Burr distribution. *Journal of Applied Statistics*, 23:609–619.

Mielke, P. W. (1973). Another family of distributions for describing and analyzing precipitation data. *Journal of Applied Meterology*, 12:275–280.

Nasiru, S., Mwita, P. N., and Ngesa, O. (2018). Exponentiated generalized power series family of distributions. *Annals of Data Science. In review*.

R´enyi, A. (1961). On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pages 547–561. University of California Press, Berkeley, CA.