



Exponentiated Generalized Power Series Family of Distributions

Suleman Nasiru¹ · Peter N. Mwita² · Oscar Ngesa³

Received: 22 November 2017 / Revised: 5 July 2018 / Accepted: 26 July 2018 / Published online: 31 July 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

In this paper, a new family of distributions called the exponentiated generalized power series family is proposed and studied. Statistical properties such as stochastic order, quantile function, entropy, mean residual life and order statistics were derived. Bivariate and multivariate extensions of the family was proposed. The method of maximum likelihood estimation was proposed for the estimation of the parameters. Some special distributions from the family were defined and their applications were demonstrated with real data sets.

Keywords Exponentiated generalized · Power series · Bivariate · Multivariate · Stochastic ordering · Order statistics

Mathematics Subject Classification 62E15 · 60E05

1 Introduction

The development of generalized classes of distributions by researchers has become an interesting area of research in recent time. Several generalized classes of distributions have been proposed in statistical literature with the motivation of defining models with different types of failure rates, construct heavy-tailed distributions for modeling

✉ Suleman Nasiru
sulemanstat@gmail.com; snasiru@uds.edu.gh

Peter N. Mwita
petermwita@mksu.ac.ke

Oscar Ngesa
oscanges@ttu.ac.ke

¹ Institute for Basic Sciences, Technology and Innovation, Pan African University, P.O. Box 62000-00200, Nairobi, Kenya

² Department of Mathematics, Machakos University, P.O. Box 136-90100, Machakos, Kenya

³ Mathematics and Informatics Department, Taita Taveta University, P. O. Box 635-80300, Voi, Kenya

different kinds of data sets, generate distributions with left skewed, right skewed, symmetric, or reversed J shape and to consistently provide a reasonable parametric fits to data sets. This is necessary because the nature of data arising from different areas of study such as lifetime analysis, insurance, engineering, finance and biological science among others require extended forms of the existing statistical distributions before a better fit can be obtained.

Some of the generalized classes of distributions in literature includes: the transmuted geometric-G family [1]; Kumaraswamy transmuted-G family [2]; beta transmuted-H family [3]; generalized transmuted-G family [18]; transmuted exponentiated generalized-G family [22]; Kumaraswamy family [10]; Weibull-G family [9]; beta extended Weibull family [12]; exponentiated generalized class [11]; odd generalized exponential family [21]; transformed-transformer ($T-X$) method [4]; exponentiated generalized $T-X$ method [16]; exponentiated $T-X$ method [5] and exponentiated generalized exponential-X family [15].

This study proposes a new class of distributions called the exponentiated generalized power series (EGPS) family of distributions by compounding the exponentiated generalized (EG) class of distributions and the power series (PS) family of distributions. The rest of the paper is organized as follows: In Sect. 2, the cumulative distribution function (CDF), probability density function (PDF), survival function and hazard function of the EGPS family of distributions were defined. In Sect. 3, some sub-classes of the new family were discussed. In Sect. 4, statistical properties of the EGPS class were proposed using copula. In Sect. 7, special distributions from the EGPS family were discussed. In Sect. 8, simulations was performed to examine the finite sample properties of the estimators for the parameters of the special distributions. In Sect. 9, applications of the special distributions were demonstrated using real data sets. Finally, the concluding remarks of the study were given in Sect. 10.

2 Exponentiated Generalized Power Series Family

Let N represent the number of independent subsystems of a system functioning at a given time. Suppose that N has zero truncated PS distribution with probability mass function given by

$$P(N = n) = \frac{a_n \lambda^n}{C(\lambda)}, \quad n = 1, 2, \dots, \quad (1)$$

where $a_n > 0$, $C(\lambda) = \sum_{i=1}^{\infty} a_n \lambda^n$ and $\lambda \in (0, s)$ (s can be ∞) is chosen such that $C(\lambda)$ is finite and its first, second and third derivatives are defined and shown by $C'(\cdot)$, $C''(\cdot)$ and $C'''(\cdot)$. The PS family includes; binomial, Poisson, geometric and logarithmic distributions. Detailed information of the PS family can be found in [17]. Suppose the failure time of each subsystem follows the EG class of distributions with CDF given by

$$H_{c,d}(x) = \left(1 - (1 - G(x; \psi))^d\right)^c, \quad x \in \mathbb{R}, \quad (2)$$

Table 1 Useful quantities for some power series distributions

Distribution	a_n	$C(\lambda)$	$C'(\lambda)$	$C''(\lambda)$	$C'''(\lambda)$	s
Geometric	1	$\lambda(1 - \lambda)^{-1}$	$(1 - \lambda)^2$	$2(1 - \lambda)^{-3}$	$6(1 - \lambda)^{-4}$	1
Poisson	$\frac{1}{n!}$	$e^\lambda - 1$	e^λ	e^λ	e^λ	∞
Logarithmic	n^{-1}	$-\log(1 - \lambda)$	$(1 - \lambda)^{-1}$	$(1 - \lambda)^{-2}$	$2(1 - \lambda)^{-3}$	1
Binomial	$\binom{m}{n}$	$(1 + \lambda)^m - 1$	$\frac{m}{(1 + \lambda)^{1-m}}$	$\frac{m(m-1)}{(1 + \lambda)^{2-m}}$	$\frac{m(m-1)(m-2)}{(1 + \lambda)^{3-m}}$	∞

where $c, d > 0$ are extra shape parameters, $G(x; \psi)$ is the baseline CDF depending on parameter ψ and $g(x; \psi)$ is its corresponding density function. For simplicity, $G(x; \psi)$ is written as $G(x)$. If T_j is the failure time of the j th subsystem and X represents the time to failure of the first out of the N operating subsystems such that $X = \min(T_1, T_2, \dots, T_N)$. Then the conditional CDF of X given N is

$$\begin{aligned}
 F(x|N = n) &= 1 - P(X > x|N) \\
 &= 1 - P(T_1 > x, T_2 > x, \dots, T_N > x) \\
 &= 1 - (1 - P(T_1 \leq x))^n \\
 &= 1 - \left[1 - \left(1 - (1 - G(x))^d\right)^c\right]^n.
 \end{aligned} \tag{3}$$

Hence, the marginal CDF of X is given by

$$\begin{aligned}
 F(x) &= \sum_{i=1}^{\infty} \frac{a_n \lambda^n}{C(\lambda)} \left\{1 - \left[1 - \left(1 - (1 - G(x))^d\right)^c\right]^n\right\} \\
 &= 1 - \frac{C\left(\lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right)}{C(\lambda)}.
 \end{aligned} \tag{4}$$

The PDF is given by

$$f(x) = \lambda c d g(x) (1 - G(x))^{d-1} \left(1 - (1 - G(x))^d\right)^{c-1} \frac{C'\left(\lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right)}{C(\lambda)}. \tag{5}$$

Table 1 summarizes some particular cases of zero truncated PS distributions.

The survival function and the hazard rate function of the EGPS class of distributions are respectively given by

$$S(x) = \frac{C\left(\lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right)}{C(\lambda)}, \tag{6}$$

and

$$\tau(x) = \lambda c d g(x) (1 - G(x))^{d-1} \left(1 - (1 - G(x))^d\right)^{c-1} \frac{C' \left(\lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right)}{C \left(\lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right)}. \tag{7}$$

Remark 1 If $X = \max(T_1, T_2, \dots, T_N)$, then the CDF of the EGPS class is given by

$$F(x) = \frac{C \left(\lambda (1 - (1 - G(x))^d)^c\right)}{C(\lambda)}. \tag{8}$$

Remark 2 If $C(\lambda) = \lambda$, then the EG class is a special case of the EGPS class.

Proposition 1 *The EG class is a limiting case of the EGPS class when $\lambda \rightarrow 0^+$.*
Proof

$$\begin{aligned} \lim_{\lambda \rightarrow 0^+} F(x) &= 1 - \lim_{\lambda \rightarrow 0^+} \frac{C \left(\lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right)}{C(\lambda)} \\ &= 1 - \lim_{\lambda \rightarrow 0^+} \frac{\sum_{n=1}^{\infty} a_n \lambda^n \left[1 - (1 - (1 - G(x))^d)^c\right]^n}{\sum_{n=1}^{\infty} a_n \lambda^n}. \end{aligned}$$

Applying L'Hôpital's rule

$$\begin{aligned} \lim_{\lambda \rightarrow 0^+} F(x) &= 1 - \lim_{\lambda \rightarrow 0^+} \frac{\sum_{n=1}^{\infty} n a_n \lambda^{n-1} \left[1 - (1 - (1 - G(x))^d)^c\right]^n}{\sum_{n=1}^{\infty} n a_n \lambda^{n-1}} \\ &= 1 - \lim_{\lambda \rightarrow 0^+} \frac{a_1 \left[1 - (1 - (1 - G(x))^d)^c\right] + \sum_{n=2}^{\infty} n a_n \lambda^{n-1} \left[1 - (1 - (1 - G(x))^d)^c\right]^n}{a_1 + \sum_{n=2}^{\infty} n a_n \lambda^{n-1}} \\ &= \left(1 - (1 - G(x))^d\right)^c. \end{aligned}$$

Proposition 2 *The exponentiated PS class is a limiting special case of the EGPS class when $d \rightarrow 1$.*

Proof

$$\begin{aligned} \lim_{d \rightarrow 1} F(x) &= 1 - \lim_{d \rightarrow 1} \frac{C \left(\lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right)}{C(\lambda)} \\ &= 1 - \frac{C(\lambda (1 - G(x))^c)}{C(\lambda)}. \end{aligned}$$

Lemma 1 *The EGPS class density has a linear representation of the form*

$$f(x) = c d \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[g(x) G(x)^k \right], \tag{9}$$

where $E_N(\cdot)$ is the expectation with respect to the random variable N and

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

Proof The EGPS PDF can be written as

$$f(x) = cd \sum_{n=1}^{\infty} P(N=n) n g(x) (1-G(x))^{d-1} \left(1 - (1-G(x))^{d-1}\right)^{c-1} \left[1 - (1-G(x))^d\right]^{n-1}.$$

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

$$(1-z)^{\eta-1} = \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma(\eta)}{i! \Gamma(\eta-i)} z^i. \quad (10)$$

Using the series expansion in Eq. (10) thrice and the fact that $0 \leq 1-G(x) \leq 1$, yields

$$f(x) = cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[g(x) G(x)^k \right].$$

The linear representation of the density function makes it easy to study the statistical properties of the EGPS class. Alternatively it can be written in terms of the exp-G function as

$$f(x) = cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk}^* E_N[\varphi_{k+1}], \quad (11)$$

where $\omega_{ijk}^* = \frac{\omega_{ijk}}{(k+1)}$ and $\varphi_{k+1} = (k+1)g(x)G(x)^k$ is the exp-G density with power parameter $k+1$.

3 Sub-Families

In this section, a number of sub-families of the EGPS family were discussed. These include: EG Poisson (EGP), EG binomial (EGB), EG geometric (EGG) and logarithmic (EGL) classes of distributions.

3.1 Exponentiated Generalized Poisson Class

The zero truncated Poisson distribution is a special case of PS distribution with $a_n = \frac{1}{n!}$ and $C(\lambda) = e^\lambda - 1$, ($\lambda > 0$). Using the CDF in Eq. (4), the CDF and PDF of the EGP class of distributions are respectively given by

$$F(x) = \frac{e^\lambda - e^{\lambda[1-(1-(1-G(x))^d)^c]}}{e^\lambda - 1}, \quad (12)$$

and

$$f(x) = \lambda cdg(x)(1 - G(x))^{d-1} \left(1 - (1 - G(x))^d\right)^{c-1} \frac{e^{\lambda[1-(1-(1-G(x))^d)^c]}}{e^\lambda - 1}. \quad (13)$$

3.2 Exponentiated Generalized Binomial Class

The zero truncated binomial distribution is a special case of PS distribution $a_n = \binom{m}{n}$ and $C(\lambda) = (1 + \lambda)^m - 1$, ($\lambda > 0$), where $m(n \leq m)$ is the number of replicas and is a positive integer. The CDF of the EGB class of distributions are respectively

$$F(x) = 1 - \frac{\left[1 + \lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right]^m - 1}{(1 + \lambda)^m - 1}, \quad (14)$$

and

$$f(x) = m\lambda cdg(x)(1 - G(x))^{d-1} \left(1 - (1 - G(x))^d\right)^{c-1} \frac{\left[1 + \lambda \left[1 - (1 - (1 - g(x))^d)^c\right]\right]^{m-1}}{(1 + \lambda)^m - 1}. \quad (15)$$

The EGP class is a limiting case of the EGB class if $m\lambda \rightarrow \theta > 0$, when $m \rightarrow \infty$.

3.3 Exponentiated Generalized Geometric Class

The zero truncated geometric distribution is a special case of PS distributions with $a_n = 1$ and $C(\lambda) = \frac{\lambda}{1-\lambda}$, ($0 < \lambda < 1$). The CDF and PDF of the EGG class of distributions are respectively given by

$$F(x) = 1 - \frac{(1 - \lambda) \left[1 - (1 - (1 - G(x))^d)^c\right]}{1 - \lambda \left[1 - (1 - (1 - G(x))^d)^c\right]}, \quad (16)$$

and

$$f(x) = \frac{(1 - \lambda)cdg(x)(1 - G(x))^{d-1}(1 - (1 - G(x))^d)^{c-1}}{\left[1 - \lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right]^2}. \quad (17)$$

3.4 Exponentiated Generalized Logarithmic Class

The zero truncated logarithmic distribution is another special case of the PS distribution with $a_n = \frac{1}{n}$ and $C(\lambda) = -\log(1 - \lambda)$, ($0 < \lambda < 1$). The CDF and PDF of the EGL class are respectively given by

$$F(x) = 1 - \frac{\log\left[1 - \lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right]}{\log(1 - \lambda)}, \quad (18)$$

and

$$f(x) = \frac{\lambda cdg(x)(1 - G(x))^{d-1}(1 - (1 - G(x))^d)^{c-1}}{\log(1 - \lambda) \left[\lambda \left[1 - (1 - (1 - G(x))^d)^c\right] - 1\right]}. \quad (19)$$

4 Statistical Properties

In this section, the statistical properties of the EGPS class of distributions were discussed. These include: the quantile function, moment, moment generating function, incomplete moment, mean residual life, stochastic property, reliability, Shannon entropy and order statistics.

4.1 Quantile Function

The quantile function is another way of describing the distribution of a random variable. It plays a key role when simulating random numbers from a distribution and it provides an alternative means for describing the shapes of a distribution.

Proposition 3 *The quantile function of the EGPS class is given by*

$$Q_F(u) = G^{-1} \left\{ 1 - \left[1 - \left(1 - \frac{C^{-1}((1-u)C(\lambda))}{\lambda} \right)^{\frac{1}{c}} \right]^{\frac{1}{d}} \right\}, \quad (20)$$

where $u \in [0, 1]$ and $C^{-1}(\cdot)$ is the inverse of $C(\cdot)$.

Proof By definition, the quantile function is given by $F(x_u) = P(X \leq x_u) = u$. Thus, setting $Q_F(u) = x_u$ in Eq. (4) and solving for it yields the quantile function of the EGPS class.

The median of the EGPS class is obtained by substituting $u = 0.5$ into Eq. (20).

4.2 Moment, Generating Function and Incomplete Moment

In this subsection, the moment, moment generating function (MGF) and incomplete moment were presented.

Proposition 4 *The r th non-central moment of the EGPS class is given by*

$$\mu'_r = cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\int_{-\infty}^{\infty} x^r g(x) G(x)^k dx \right], \quad r = 1, 2, \dots \quad (21)$$

Proof By definition, the r th non-central moment is given by

$$\mu'_r = \int_{-\infty}^{\infty} x^r f(x) dx.$$

Substituting the linear representation of the density function into the definition and simplifying yields the non-central moment.

Alternatively, the moment can be expressed in terms of the quantile function of the baseline. Let $G(x) = u$, then

$$\mu'_r = cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\int_0^1 Q_G(u)^r u^k du \right]. \quad (22)$$

Proposition 5 *The MGF of EGPS class of distributions is given by*

$$M_X(t) = cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\int_{-\infty}^{\infty} e^{tx} g(x) G(x)^k dx \right]. \quad (23)$$

Proof The proof of MGF directly follows from the definition $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$.

In terms of the quantile function of the baseline, the MGF is given by

$$M_X(t) = cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\int_0^1 e^{tQ_G(u)} u^k du \right]. \quad (24)$$

Proposition 6 *The r th incomplete moment of the EGPS class of distributions is given by*

$$M_r(t) = cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\int_{-\infty}^t x^r g(x) G(x)^k dx \right], \quad r = 1, 2, \dots \quad (25)$$

Proof The proof can easily be obtained from the definition $M_r(t) = \int_{-\infty}^t x^r f(x) dx$. Letting $u = G(x)$, the incomplete moment can be expressed in terms of the baseline quantile function as

$$M_r(t) = cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\int_0^{G(t)} Q_G(u)^r u^k du \right]. \quad (26)$$

Using the power series expansion of the quantile function of the baseline as $Q_G(u) = \sum_{h=0}^{\infty} e_h u^h$, where $e_h (h = 0, 1, \dots)$ are suitably chosen real numbers that depend on the parameters of the $G(x)$ distribution.

$$(Q_G(u))^r = \left(\sum_{h=0}^{\infty} e_h u^h \right)^r = \sum_{h=0}^{\infty} e'_{r,h} u^h,$$

where $e'_{r,h} = (he_0)^{-1} \sum_{z=1}^h [z(r+1) - h] e_z e'_{r,h-z}$, $e'_{r,0} = (e_0)^h$ and $r (r \geq 1)$ is a positive integer. For more details on quantile power series expansion see [13]. The incomplete moment can now be expressed as

$$\begin{aligned} M_r(t) &= cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\sum_{h=0}^{\infty} e'_{r,h} \int_0^{G(t)} u^{k+h} du \right] \\ &= cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\sum_{h=0}^{\infty} \frac{e'_{r,h} G(t)^{k+h+1}}{k+h+1} \right]. \end{aligned}$$

4.3 Residual and Mean Residual Life

A system's residual lifetime when it is still operating at time t , is $X_t = X - t | X > t$ which has the PDF

$$\begin{aligned} f(x;t) &= \frac{f(x)}{1 - F(t)} \\ &= \lambda cd g(x) (1 - G(x))^{d-1} \left(1 - (1 - G(x))^d \right)^{c-1} \frac{C' \left(\lambda \left[1 - (1 - (1 - G(x))^d)^c \right] \right)}{C \left(\lambda \left[1 - (1 - (1 - G(x))^d)^c \right] \right)}. \end{aligned}$$

Proposition 7 *The mean residual life of X_t is given by*

$$m(t) = \frac{1}{1 - F(t)} \left[\mu - cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\sum_{h=0}^{\infty} \frac{e_h G(t)^{k+h+1}}{k+h+1} \right] \right] - t, \tag{27}$$

where $\mu = \mu'_1$ is the mean and $e_h (h = 0, 1, \dots)$ are suitably chosen real numbers that depend on the parameters of the $G(x)$ distribution.

Proof The mean residual life is defined as

$$m(t) = E[X - t | X > t] = \frac{\int_t^{\infty} (x - t) f(x) dx}{1 - F(t)} = \frac{\mu'_1 - \int_{-\infty}^t x f(x) dx}{1 - F(t)} - t. \tag{28}$$

The integral $\int_{-\infty}^t x f(x) dx$ is the first incomplete moment. Thus substituting the first incomplete moment into Eq. (28) yields the mean residual life.

4.4 Stochastic Ordering Property

Stochastic ordering is the common way of showing ordering mechanism in lifetime distribution. A random variable X_1 is said to be greater than a random variable X_2 in likelihood ratio order if $\frac{f_{X_1}(x)}{f_{X_2}(x)}$ is an increasing function of x .

Proposition 8 *Let $X_1 \sim EGPS(x; c, d, \lambda, \psi)$ and $X_2 \sim EG(x; c, d, \psi)$, then X_2 is greater than X_1 in likelihood ratio order ($X_1 \leq_{lr} X_2$) provided $\lambda > 0$.*

Proof

$$\frac{f_{X_1}(x)}{f_{X_2}(x)} = \frac{\lambda C' \left(\lambda \left[1 - (1 - (1 - G(x))^d)^c \right] \right)}{C(\lambda)},$$

which is decreasing in x provided $\lambda > 0$.

From Proposition 8, it can be considered that the hazard rate order, the usual stochastic order and the mean residual life order between X_1 and X_2 hold.

4.5 Stress-Strength Reliability

Reliability plays a useful role in the analysis of stress-strength of models. If X_1 is the strength of a component and X_2 is the stress, then the component fails when $X_1 \leq X_2$. The estimate of the reliability of the component R is $P(X_2 < X_1)$.

Proposition 9 *If $X_1 \sim EGPS(x; c, d, \lambda, \psi)$ and $X_2 \sim EGPS(x; c, d, \lambda, \psi)$, then reliability is given by*

$$R = 1 - \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\int_0^1 u^k \frac{C\left(\lambda \left[1 - (1 - (1 - G(x))^d)^c\right]\right)}{C(\lambda)} du \right]. \tag{29}$$

Proof The reliability is defined as

$$\begin{aligned} R &= \int_{-\infty}^{\infty} f(x)F(x)dx \\ &= 1 - \int_{-\infty}^{\infty} f(x)S(x)dx \\ &= 1 - \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\int_{-\infty}^{\infty} g(x)G(x)^k S(x)dx \right]. \end{aligned}$$

Letting $G(x) = u$ yields

$$R = 1 - \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\int_0^1 u^k \frac{C\left(\lambda \left[1 - (1 - (1 - u)^d)^c\right]\right)}{C(\lambda)} du \right].$$

4.6 Shannon Entropy

The entropy of a random variable is a measure of variation or uncertainty of the random variable. The Shannon entropy of a random variable X with PDF $f(x)$ is given by $\eta_X = E\{-\log f(x)\}$ [19].

Proposition 10 *The Shannon entropy of the EGPS class random variable is given by*

$$\eta_X = -\log\left(\frac{cd\lambda}{C(\lambda)}\right) - E[\log g(X)] + (1 - d)\delta_1 + (1 - c)\delta_2 - E\left[\log C'(\lambda[1 - H_{c,d}(X)])\right], \tag{30}$$

where $H_{c,d}(x)$ is the CDF of the EG class,

$$\delta_1 = -cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\sum_{q=1}^{\infty} \frac{1}{q(k+q+1)} \right],$$

and

$$\delta_2 = -cd \sum_{j,k=0}^{\infty} \sum_{i=0}^{n-1} \omega_{ijk} E_N \left[\sum_{s=0}^{\infty} \sum_{q=1}^{\infty} \frac{(-1)^s \binom{dq}{s}}{q(k+q+1)} \right].$$

Proof By definition

$$\eta_X = -\log\left(\frac{cd\lambda}{C(\lambda)}\right) - E[\log g(X)] + (1-d)E[\log(1-G(X))] + (1-c)E[\log(1-(1-G(x))^d)^c] - E[\log C'(\lambda[1-H_{c,d}(X)])]. \tag{31}$$

Let $\delta_1 = E[\log(1-G(X))]$ and $\delta_2 = E[\log(1-(1-(1-G(X))^d)^c)]$. Using the identity $\log(1-z) = -\sum_{q=1}^{\infty} \frac{z^q}{q}$, $|z| < 1$, yields

$$\log(1-G(x)) = -\sum_{q=1}^{\infty} \frac{G(x)^q}{q},$$

$$\log(1-(1-G(x))^d) = -\sum_{s=0}^{\infty} \sum_{q=1}^{\infty} \frac{(-1)^s \binom{dq}{s} G(x)^s}{q}.$$

Putting $G(x) = u$ and taking the expectation with respect to the random variable X , give the value of δ_1 and δ_2 after some algebraic manipulation.

4.7 Order Statistics

Let X_1, X_2, \dots, X_n be a random sample of size n from EGPS, then the PDF of the p th order statistic, say $X_{p:n}$ 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 EGPS class of distributions respectively, and $B(\cdot, \cdot)$ is the beta function. Thus,

$$f_{p:n}(x) = \frac{n! f(x)}{(p-1)! (n-p)!} \left[1 - \frac{C(\lambda[1-(1-(1-G(x))^d)^c])}{C(\lambda)} \right]^{p-1} \times \left[\frac{C(\lambda[1-(1-(1-G(x))^d)^c])}{C(\lambda)} \right]^{n-p}. \tag{32}$$

5 Parameter Estimation

Different methods for parameter estimation exist in literature but the maximum likelihood approach is the most commonly used. The maximum likelihood estimators have several desirable properties and can be used for constructing confidence intervals and regions. Thus, in this study, the maximum likelihood method was employed for the estimation of the parameters of the EGPS distribution. Let X_1, X_2, \dots, X_n be a random sample of size n from the EGPS distribution. Let $z_i = 1 - G(x_i; \boldsymbol{\psi})$, then the log-likelihood is given by

$$\begin{aligned} \ell = n \log(\lambda cd) + \sum_{i=1}^n \log g(x_i; \boldsymbol{\psi}) + (d-1) \sum_{i=1}^n \log(z_i) + (c-1) \sum_{i=1}^n \log(1 - z_i^d) - n \log C(\lambda) \\ + \sum_{i=1}^n \log C' \left(\lambda \left[1 - (1 - z_i^d)^c \right] \right). \end{aligned} \quad (33)$$

Taking the partial derivatives of the log-likelihood function with respect to the parameters yields the following score functions:

$$\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} - \frac{nC'(\lambda)}{C(\lambda)} + \sum_{i=1}^n \frac{[1 - (1 - z_i^d)^c]C''(\lambda[1 - (1 - z_i^d)^c])}{C'(\lambda[1 - (1 - z_i^d)^c])}, \quad (34)$$

$$\frac{\partial \ell}{\partial c} = \frac{n}{c} + \sum_{i=1}^n \log(1 - z_i^d) - \sum_{i=1}^n \frac{\lambda(1 - z_i^d)^c \log(1 - z_i^d)C''(\lambda[1 - (1 - z_i^d)^c])}{C'(\lambda[1 - (1 - z_i^d)^c])}, \quad (35)$$

$$\frac{\partial \ell}{\partial d} = \frac{n}{d} + \sum_{i=1}^n \log(z_i) - (c-1) \sum_{i=1}^n \frac{z_i^d \log(z_i)}{1 - z_i^d} + \sum_{i=1}^n \frac{\lambda c(1 - z_i^d)^{c-1} z_i^d \log(z_i)C''(\lambda[1 - (1 - z_i^d)^c])}{C'(\lambda[1 - (1 - z_i^d)^c])}, \quad (36)$$

$$\begin{aligned} \frac{\partial \ell}{\partial \boldsymbol{\psi}} = \sum_{i=1}^n \frac{g'(x_i; \boldsymbol{\psi})}{g(x_i; \boldsymbol{\psi})} - (d-1) \sum_{i=1}^n \frac{G'(x_i; \boldsymbol{\psi})}{z_i} + (c-1) \sum_{i=1}^n \frac{dz_i^{d-1} G'(x_i; \boldsymbol{\psi})}{1 - z_i^d} \\ - \sum_{i=1}^n \frac{\lambda cd z_i^{d-1} (1 - z_i^d)^{c-1} G'(x_i; \boldsymbol{\psi}) C''(\lambda[1 - (1 - z_i^d)^c])}{C'(\lambda[1 - (1 - z_i^d)^c])}, \end{aligned} \quad (37)$$

where $g'(x_i; \boldsymbol{\psi}) = \frac{\partial g(x_i; \boldsymbol{\psi})}{\partial \boldsymbol{\psi}}$ and $G'(x_i; \boldsymbol{\psi}) = \frac{\partial G(x_i; \boldsymbol{\psi})}{\partial \boldsymbol{\psi}}$. The score functions do not have close form, thus it is more convenient to solve them using nonlinear optimization techniques such as the quasi-Newton algorithm. For the purpose of interval estimation of the parameters, a $p \times p$ observed information matrix can be obtained as $J(\boldsymbol{\vartheta}) = \frac{\partial^2 \ell}{\partial q \partial r}$ (for $q, r = \lambda, c, d, \boldsymbol{\psi}$), whose elements can be computed numerically. Under the usual regularity conditions as $n \rightarrow \infty$, the distribution of $\widehat{\boldsymbol{\vartheta}} = (\widehat{\lambda}, \widehat{c}, \widehat{d}, \widehat{\boldsymbol{\psi}}^T)^T$ approximately converges to a multivariate normal $N_p(\mathbf{0}, J(\widehat{\boldsymbol{\vartheta}})^{-1})$ distribution. $J(\widehat{\boldsymbol{\vartheta}})$ is the observed information matrix evaluated at $\widehat{\boldsymbol{\vartheta}}$. The asymptotic normal distribution is useful for constructing approximate $100(1-\alpha)\%$ confidence intervals, where α is the significance level.

6 Extension via Copula

In this section, bivariate and multivariate extensions of the EGPS class of distributions were proposed using Clayton copula. Consider a random pair (X_1, X_2) , a copula C^* associated with the pair is simply a joint distribution of the random vector $(F_{X_1}(x_1), F_{X_2}(x_2))$. Suppose that $F_{X_1}(x_1)$ and $F_{X_2}(x_2)$ are the marginal CDFs of the random variables X_1 and X_2 respectively and C^* is the copula associated to (X_1, X_2) . Sklar [20] established that the joint CDF $F_{X_1 X_2}(x_1, x_2)$ of the pair (X_1, X_2) is given by

$$F_{X_1 X_2}(x_1, x_2) = C^*(F_{X_1}(x_1), F_{X_2}(x_2)).$$

Suppose (X_1, X_2) follows bivariate EGPS random variables with marginal distributions $F_{X_1}(x_1)$ and $F_{X_2}(x_2)$. Let the copula associated to (X_1, X_2) belong to Clayton copula family given by

$$C^*(z_1, z_2) = [z_1^{-\theta} + z_2^{-\theta} - 1]^{-\frac{1}{\theta}}, \theta \geq 0.$$

The joint CDF $F_{X_1 X_2}(x_1, x_2)$, of the bivariate EGPS class is given by

$$\begin{aligned} & F_{X_1 X_2}(x_1, x_2) \\ &= \left\{ \left[1 - \frac{C(\lambda_1 [1 - (1 - (1 - G(x_1; \psi_1))^{d_1})^{c_1}])}{C(\lambda_1)} \right]^{-\theta} \right. \\ & \quad \left. + \left[1 - \frac{C(\lambda_2 [1 - (1 - (1 - G(x_2; \psi_2))^{d_2})^{c_2}])}{C(\lambda_2)} \right]^{-\theta} - 1 \right\}^{-\frac{1}{\theta}}, \quad (38) \end{aligned}$$

where $\lambda_1, \lambda_2, c_1, c_2, d_1, d_2, \psi_1$ and ψ_2 describe the marginal parameters while θ is the Clayton copula parameter. A p -dimensional multivariate extension from the above is given by

$$F_{X_1 X_2 \dots X_p}(x_1, x_2, \dots, x_p) = \left\{ \sum_{i=1}^p \left[1 - \frac{C(\lambda_i [1 - (1 - (1 - G(x_i; \psi_i))^{d_i})^{c_i}])}{C(\lambda_i)} \right]^{-\theta} \right\}^{-\frac{1}{\theta}}. \quad (39)$$

7 Special Distribution

In this section, four special distributions were presented. These include EGP inverse exponential (EGPIE) distribution, EGB inverse exponential (EGBIE) distribution, EGG inverse exponential (EGGIE) distribution and EGL inverse exponential (EGLIE) distribution. Suppose the baseline distribution follows an inverse exponential distribution with CDF $G(x) = e^{-\gamma x^{-1}}$, $\gamma > 0, x > 0$. The densities, hazard rate functions

and quantiles of the EGPIE, EGBIE, EGGIE and EGLIE distributions are defined as follows:

7.1 EGPIE Distribution

The density function of the EGPIE distribution is obtained by substituting the baseline CDF and its corresponding PDF into Eq. (13). Thus, the PDF of EGPIE distribution is given by

$$f(x) = \lambda\gamma cd x^{-2} e^{-\gamma x^{-1}} \left(1 - e^{-\gamma x^{-1}}\right)^{d-1} \left(1 - \left(1 - e^{-\gamma x^{-1}}\right)^d\right)^{c-1} \frac{e^{\lambda[1-(1-(1-e^{-\gamma x^{-1}})^d)^c]}}{e^\lambda - 1},$$

$\lambda, \gamma, c, d > 0, x > 0.$ (40)

The corresponding hazard rate function is given by

$$\tau(x) = \lambda\gamma cd x^{-2} e^{-\gamma x^{-1}} \left(1 - e^{-\gamma x^{-1}}\right)^{d-1} \left(1 - \left(1 - e^{-\gamma x^{-1}}\right)^d\right)^{c-1} \frac{e^{\lambda[1-(1-(1-e^{-\gamma x^{-1}})^d)^c]}}{e^{\lambda[1-(1-(1-e^{-\gamma x^{-1}})^d)^c]} - 1},$$

$\lambda, \gamma, c, d > 0, x > 0.$ (41)

Figure 1 displays the plots of the density and hazard rate function of the EGPIE distribution. From the figure, the density exhibit right skewed shape with varied degrees of kurtosis and an approximately symmetric shape. The hazard rate function shows an upside down bathtub shapes.

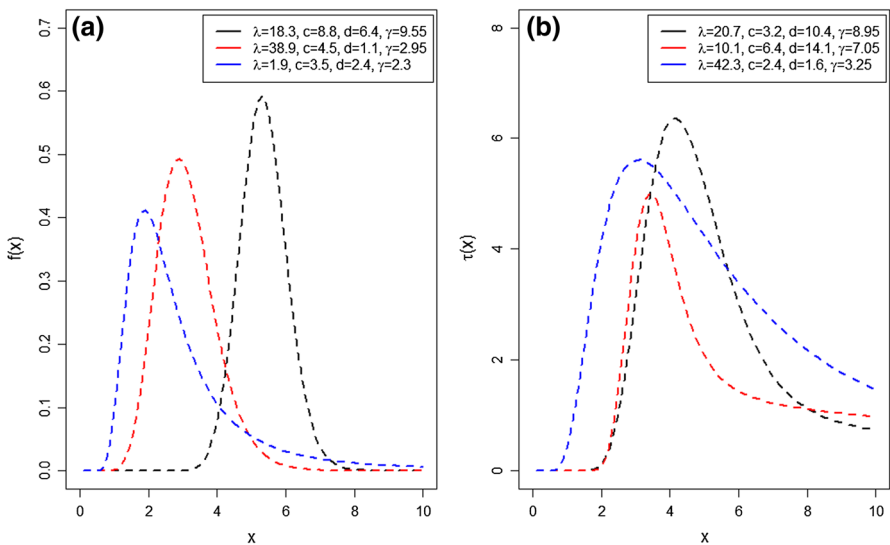


Fig. 1 Plots of EGPIE a PDF and b hazard rate function

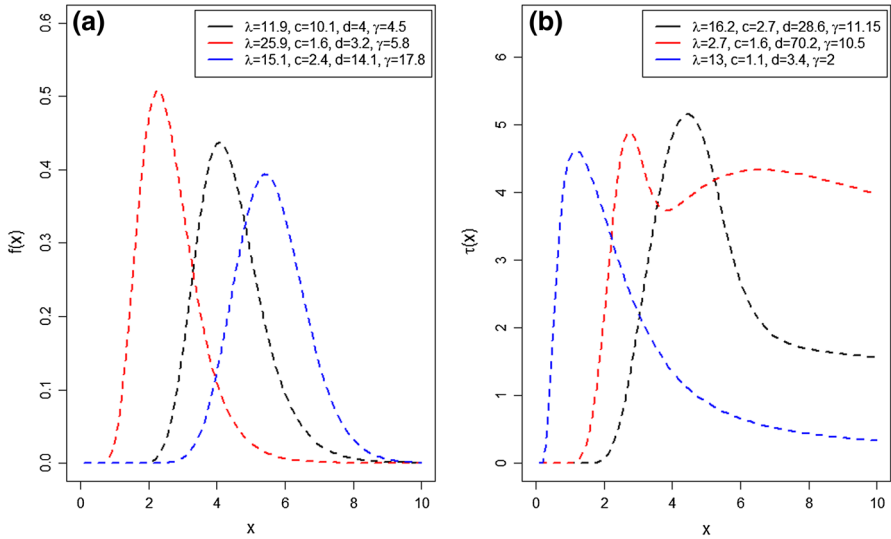


Fig. 2 Plots of EGBIE a PDF and b hazard rate function

The quantile function of the EGPIE distribution is given by

$$Q(u) = \left\{ \frac{-1}{\gamma} \log \left[1 - \left[1 - \left(1 - \frac{\log(e^\lambda - u(e^\lambda - 1))}{\lambda} \right)^{\frac{1}{c}} \right]^{\frac{1}{d}} \right] \right\}^{-1}. \tag{42}$$

7.2 EGBIE Distribution

Using Eq. (15), the PDF of EGBIE distribution is given by

$$f(x) = m\lambda\gamma cd x^{-2} e^{-\gamma x^{-1}} \left(1 - e^{-\gamma x^{-1}}\right)^{d-1} \left(1 - \left(1 - e^{-\gamma x^{-1}}\right)^d\right)^{c-1} \frac{[1 + \lambda[1 - (1 - (1 - e^{-\gamma x^{-1}})^d)^c]]^{m-1}}{(1 + \lambda)^m - 1}, \tag{43}$$

$\lambda, \gamma, c, d > 0, x > 0.$

The corresponding hazard rate function is given by

$$\tau(x) = m\lambda\gamma cd x^{-2} e^{-\gamma x^{-1}} \left(1 - e^{-\gamma x^{-1}}\right)^{d-1} \left(1 - \left(1 - e^{-\gamma x^{-1}}\right)^d\right)^{c-1} \frac{[1 + \lambda[1 - (1 - (1 - e^{-\gamma x^{-1}})^d)^c]]^{m-1}}{[1 + \lambda[1 - (1 - (1 - e^{-\gamma x^{-1}})^d)^c]]^m - 1}, \tag{44}$$

$\lambda, \gamma, c, d > 0, x > 0.$

The plots of the density and hazard rate functions of the EGNIE distribution for $m = 5$ are shown in Fig. 2. The density function exhibits right skewed and approximately symmetric shapes. The hazard rate function exhibits an upside down bathtub shapes and an upside down bathtub shape followed by a bathtub and then upside down bathtub shape.

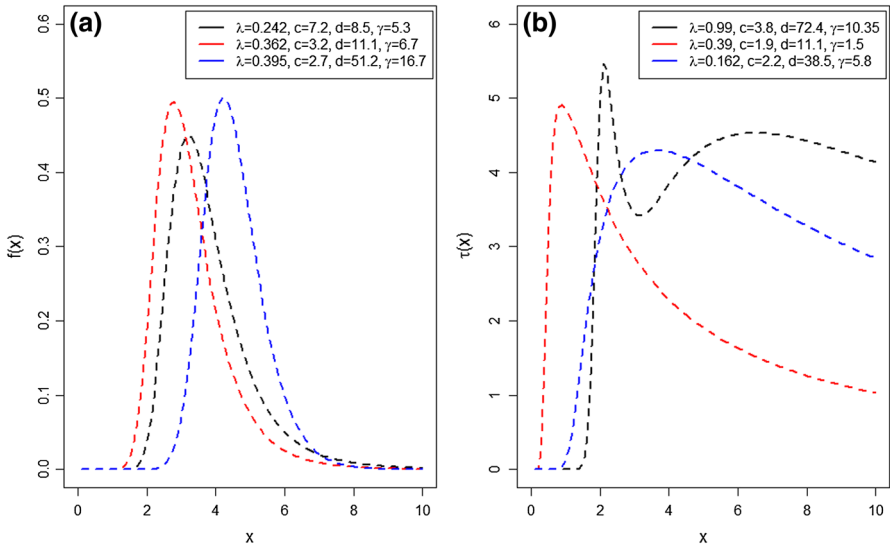


Fig. 3 Plots of EGGIE a PDF and b hazard rate function

The quantile function of the EGGIE distribution is given by

$$Q(u) = \left\{ \frac{-1}{\gamma} \log \left[1 - \left[1 - \left(1 - \frac{(((1 + \lambda)^m - 1)(1 - u) + 1)^{\frac{1}{m}} - 1}{\lambda} \right)^{\frac{1}{c}} \right]^{\frac{1}{d}} \right] \right\}^{-1} \tag{45}$$

7.3 EGGIE Distribution

From Eq. (17), the PDF of EGGIE distribution is given by

$$f(x) = \frac{(1 - \lambda)\gamma cd x^{-2} e^{-\gamma x^{-1}} (1 - e^{-\gamma x^{-1}})^{d-1} (1 - (1 - e^{-\gamma x^{-1}})^d)^{c-1}}{[1 - \lambda[1 - (1 - (1 - e^{-\gamma x^{-1}})^d)^c]]^2}, \quad 0 < \lambda < 1, \gamma, c, d > 0, x > 0. \tag{46}$$

The associated hazard rate function is given by

$$\tau(x) = \frac{\gamma cd x^{-2} e^{-\gamma x^{-1}} (1 - e^{-\gamma x^{-1}})^{d-1} (1 - (1 - e^{-\gamma x^{-1}})^d)^{c-1}}{[1 - (1 - (1 - e^{-\gamma x^{-1}})^d)^c][1 - \lambda[1 - (1 - (1 - e^{-\gamma x^{-1}})^d)^c]]}, \quad 0 < \lambda < 1, \gamma, c, d > 0, x > 0. \tag{47}$$

The density and hazard rate function plots of the EGGIE distribution are displayed in Fig. 3. The density and hazard rate function exhibits similar shapes like that of the EGGIE distribution.

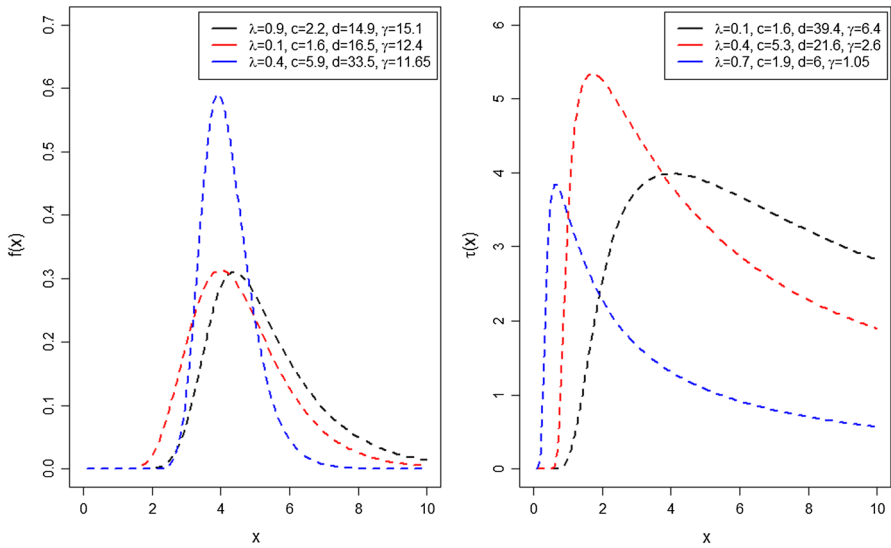


Fig. 4 Plots of EGLIE a PDF and b hazard rate function

The quantile function of the EGGIE distribution is given by

$$Q(u) = \left\{ \frac{-1}{\gamma} \log \left[1 - \left(1 - \left(\frac{u(1-\lambda)}{1-u\lambda} \right)^{\frac{1}{c}} \right)^{\frac{1}{d}} \right] \right\}^{-1}. \tag{48}$$

7.4 EGLIE Distribution

From Eq. (19), the PDF of EGLIE distribution is given by

$$f(x) = \frac{\lambda\gamma cd x^{-2} e^{-\gamma x^{-1}} (1 - e^{-\gamma x^{-1}})^{d-1} (1 - (1 - e^{-\gamma x^{-1}})^d)^{c-1}}{\log[1 - \lambda][\lambda[1 - (1 - (1 - e^{-\gamma x^{-1}})^d)^c] - 1]}, \quad 0 < \lambda < 1, \gamma, c, d > 0, x > 0. \tag{49}$$

The corresponding hazard rate function is given by

$$\tau(x) = \frac{\lambda\gamma cd x^{-2} e^{-\gamma x^{-1}} (1 - e^{-\gamma x^{-1}})^{d-1} (1 - (1 - e^{-\gamma x^{-1}})^d)^{c-1}}{\log[1 - \lambda[1 - (1 - (1 - e^{-\gamma x^{-1}})^d)^c]][\lambda[1 - (1 - (1 - e^{-\gamma x^{-1}})^d)^c] - 1]}, \quad 0 < \lambda < 1, \gamma, c, d > 0, x > 0. \tag{50}$$

Figure 4 shows the density and hazard rate function of the EGLIE distribution. The density exhibits approximately symmetric shapes with different degrees of kurtosis. The hazard rate function shows upside down bathtub shapes.

Table 2 Monte Carlo simulation results: AE and RMSE for EGPIE and EGBIE distributions

λ	c	d	γ	n	Parameter	EGPIE		EGBIE	
						AE	RMSE	AE	RMSE
0.8	0.5	0.2	0.3	25	λ	0.692	0.313	0.638	0.325
					c	0.450	0.097	0.438	0.109
					d	0.219	0.065	0.236	0.077
					γ	0.370	0.091	0.368	0.091
				50	λ	0.654	0.337	0.635	0.311
					c	0.448	0.088	0.434	0.103
					d	0.214	0.054	0.229	0.067
					γ	0.364	0.087	0.367	0.089
				75	λ	0.628	0.354	0.638	0.300
					c	0.445	0.085	0.440	0.102
					d	0.213	0.049	0.228	0.064
					γ	0.360	0.084	0.360	0.087
100	λ	0.609	0.363	0.638	0.295				
	c	0.444	0.308	0.438	0.101				
	d	0.211	0.043	0.227	0.306				
	γ	0.357	0.081	0.360	0.086				
0.1	0.9	0.1	0.5	25	λ	0.156	0.094	0.131	0.083
					c	0.873	0.129	0.885	0.121
					d	0.121	0.032	0.120	0.033
					γ	0.587	0.095	0.585	0.095
				50	λ	0.153	0.092	0.113	0.077
					c	0.902	0.100	0.898	0.105
					d	0.119	0.027	0.118	0.026
					γ	0.582	0.089	0.580	0.089
				75	λ	0.143	0.091	0.102	0.074
					c	0.909	0.089	0.904	0.089
					d	0.117	0.022	0.117	0.023
					γ	0.579	0.086	0.576	0.084
100	λ	0.140	0.091	0.091	0.073				
	c	0.910	0.082	0.901	0.085				
	d	0.116	0.020	0.117	0.021				
	γ	0.574	0.082	0.572	0.081				

The quantile function of the EGLIE distribution is given by

$$Q(u) = \left\{ \frac{-1}{\gamma} \log \left[1 - \left(1 - \left(1 - \frac{1 - (1 - \lambda)^{1-u}}{\lambda} \right)^{\frac{1}{c}} \right)^{\frac{1}{d}} \right] \right\}^{-1}. \quad (51)$$

Table 3 Monte Carlo simulation results: AE and RMSE for EGGIE and EGLIE distributions

λ	c	d	γ	n	Parameter	EGGIE		EGLIE	
						AE	RMSE	AE	RMSE
0.8	0.5	0.2	0.3	25	λ	0.817	0.099	0.771	0.144
					c	0.556	0.083	0.508	0.080
					d	0.174	0.047	0.218	0.080
					γ	0.282	0.039	0.318	0.078
				50	λ	0.834	0.072	0.784	0.101
					c	0.558	0.078	0.505	0.073
					d	0.175	0.043	0.215	0.071
					γ	0.285	0.031	0.316	0.068
				75	λ	0.834	0.068	0.780	0.092
					c	0.554	0.075	0.502	0.073
					d	0.175	0.042	0.219	0.069
					γ	0.285	0.029	0.311	0.062
100	λ	0.835	0.065	0.779	0.081				
	c	0.552	0.072	0.503	0.073				
	d	0.175	0.040	0.220	0.065				
	γ	0.285	0.030	0.308	0.060				
0.1	0.9	0.1	0.5	25	λ	0.087	0.161	0.073	0.138
					c	0.939	0.112	0.936	0.113
					d	0.121	0.035	0.122	0.035
					γ	0.561	0.100	0.560	0.100
				50	λ	0.066	0.139	0.066	0.133
					c	0.933	0.100	0.933	0.097
					d	0.119	0.028	0.119	0.027
					γ	0.559	0.090	0.562	0.088
				75	λ	0.053	0.123	0.055	0.121
					c	0.926	0.090	0.926	0.090
					d	0.118	0.024	0.118	0.024
					γ	0.555	0.081	0.555	0.082
100	λ	0.043	0.118	0.045	0.114				
	c	0.923	0.085	0.924	0.085				
	d	0.117	0.022	0.117	0.022				
	γ	0.553	0.076	0.554	0.076				

8 Monte Carlo Simulation

In this section, Monte Carlo simulations were performed to examine the finite sample properties of the maximum likelihood estimators for the parameters of the EGPIE, EGBIE, EGGIE and EGLIE distributions. For the case of the EGBIE distribution, $m = 5$ was used during the simulation. The simulation steps are as follows

Table 4 Failure times of Kevlar 49/epoxy strands with pressure at 90%

0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.04	0.05	0.06	0.07	0.07
0.08	0.09	0.09	0.10	0.10	0.11	0.11	0.12	0.13	0.18	0.19	0.20
0.23	0.24	0.24	0.29	0.34	0.35	0.36	0.38	0.40	0.42	0.43	0.52
0.54	0.56	0.60	0.60	0.63	0.65	0.67	0.68	0.72	0.72	0.72	0.73
0.79	0.79	0.80	0.80	0.83	0.85	0.90	0.92	0.95	0.99	1.00	1.01
1.02	1.03	1.05	1.10	1.10	1.11	1.15	1.18	1.20	1.29	1.31	1.33
1.34	1.40	1.43	1.45	1.50	1.51	1.52	1.53	1.54	1.54	1.55	1.58
1.60	1.63	1.64	1.80	1.80	1.81	2.02	2.05	2.14	2.17	2.33	3.03
3.03	3.34	4.20	4.69	7.89							

Table 5 Maximum likelihood estimates of parameters and standard errors

Model	$\hat{\lambda}$	\hat{c}	\hat{d}	$\hat{\gamma}$
EGPIE	26.062 (0.009)	7.320 (1.770)	0.175 (0.019)	0.002 (0.002)
EGBIE	11.644 (1.925×10^{-5})	8.862 (1.136×10^{-4})	0.313 (2.141×10^{-2})	0.003 (8.755×10^{-5})
EGGIE	0.664 (2.412×10^{-1})	20.525 (4.798×10^{-3})	0.498 (1.360×10^{-1})	0.002 (9.954×10^{-4})
EGLIE	0.018 (5.004×10^{-1})	19.277 (3.845×10^{-3})	0.616 (6.431×10^{-2})	0.002 (6.301×10^{-4})

Table 6 Log-likelihood, goodness-of-fit statistics and information criteria

Model	ℓ	AIC	AIC _c	BIC	K–S	W*
EGPIE	− 116.660	241.314	241.946	251.774	0.182	0.738
EGBIE	− 122.930	253.868	254.500	264.328	0.195	0.926
EGGIE	− 140.090	288.170	288.802	298.631	0.237	1.386
EGLIE	− 134.010	276.025	276.657	286.486	0.203	1.211

The bold values means best based on the model selection technique

1. Specify the values of the parameters λ, c, d, γ and the sample size n .
2. Generate random samples of size $n = 25, 50, 75, 100$ from EGPIE, EGBIE, EGGIE and EGLIE distributions using their respective quantiles.
3. Find the maximum likelihood estimates of the parameters.
4. Repeat steps 2–3 for $N = 1500$ times.
5. Calculate the average estimates (AE) and root mean square error (RMSE) for the estimators of the parameters of the distributions.

Table 2 shows the simulation results for the EGPIE and EGBIE distributions whereas Table 3 displays that of the EGGIE and EGLIE distributions. From both tables it can be seen that the AE of the parameters were quite close to the actual values of the parameters. The RMSEs of the parameters were smaller and decay towards zero as the sample size increases.

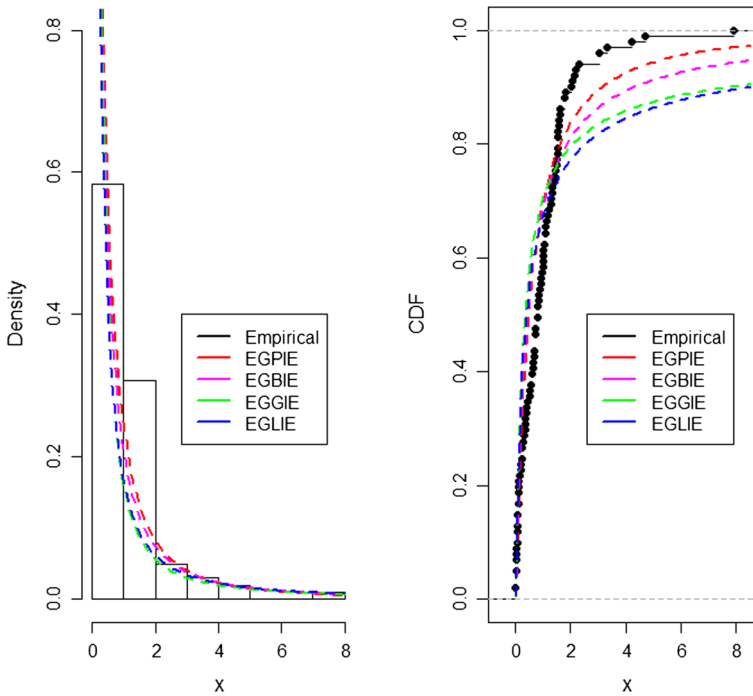


Fig. 5 Empirical and fitted density and CDF plots of Kevlar data

9 Applications

The applications of the EGPIE, EGBIE (with $m = 5$), EGGIE and EGLIE distributions were demonstrated in this section using real data sets. The performance of the distributions with regards to providing reasonable parametric fit to the data sets were compared using the Kolmogorov–Smirnov (K–S) statistic, Cramér–von (W^*) misses distance values, Akaike information criterion (AIC), corrected Akaike information criterion (AIC_c), and Bayesian information criterion (BIC). The maximum likelihood estimates for the parameters of the models were estimated by maximizing the log-likelihood functions of the models via the subroutine *mle2* using the *bbmle* package in R [8]. This was carried out using a wide range of initial values. The process often results into more than one maximum, hence in such situation, the maximum likelihood estimates corresponding to the largest maxima was chosen.

9.1 First Data Set

The data set comprises 101 observations corresponding to the failure time in hours of Kevlar 49/epoxy strands with pressure at 90%. The data set displayed in Table 4 can be found in [7, 6].

Table 5 displays the maximum likelihood estimates of the parameters of the fitted distributions with their corresponding standard errors in brackets. The parameters of

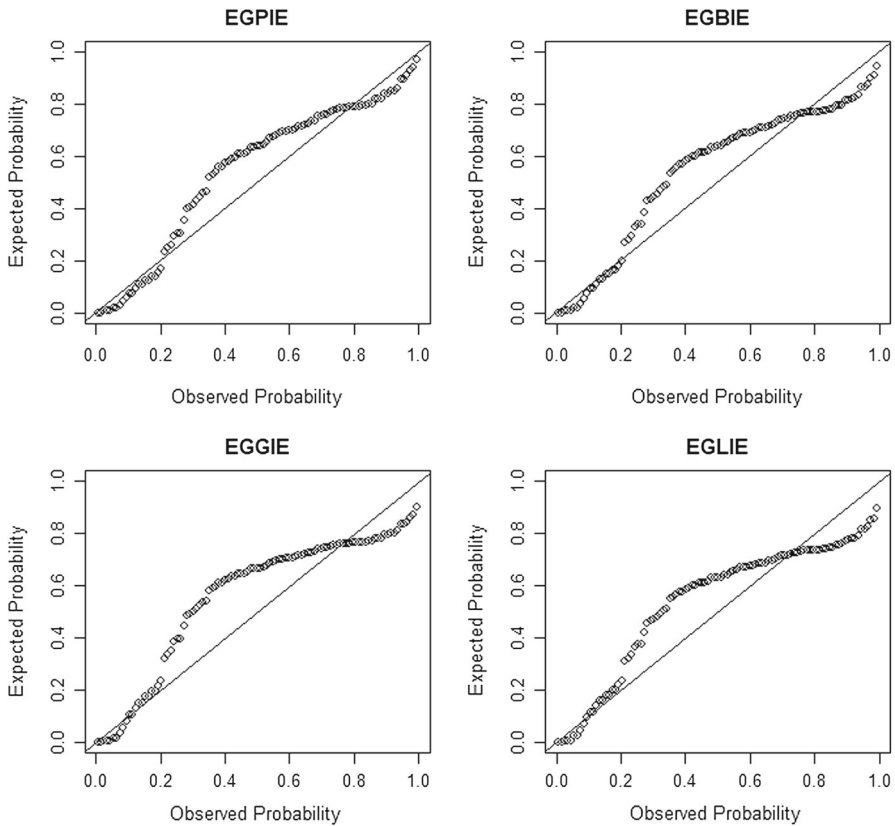


Fig. 6 P–P plots of the fitted distributions

Table 7 Egypt tax data in 1000 million Egyptian pounds

5.9	20.4	14.9	16.2	17.2	7.8	6.1	9.2	10.2	9.6	13.3	8.5
21.6	18.5	5.1	6.7	17	8.6	9.7	39.2	35.7	15.7	9.7	10
4.1	36	8.5	8	9.2	26.2	21.9	16.7	21.3	35.4	14.3	8.5
10.6	19.1	20.5	7.1	7.7	18.1	16.5	11.9	7	8.6	12.5	10.3
11.2	6.1	8.4	11	11.6	11.9	5.2	6.8	8.9	7.1	10.8	

the EGPIE and EGGIE distributions were all significant at the 5% level with the exception of the parameter γ for the two distributions. The parameters of the EGBIE distribution were all significant at the 5% level. The EGLIE distribution parameters were all significant at the 5% level with the exception of the parameter λ .

The EGPIE distribution provides a better fit to the data set compared to the other models. From Table 6, the EGPIE distribution has the highest log-likelihood and the smallest K–S, W^* , AIC, AIC_c and BIC values compared to the other fitted models.

The plots of the empirical density, the fitted densities, the empirical CDF and the CDF of the fitted distributions are shown in Fig. 5. It is obvious the EGPIE distribution provides a better fit to the data compared to the other fitted models.

Table 8 Maximum likelihood estimates of parameters and standard errors

Model	$\hat{\lambda}$	\hat{c}	\hat{d}	$\hat{\gamma}$
EGPIE	9.221 (13.530)	0.259 (0.027)	0.426 (2.377)	102.807 (0.439)
EGBIE	16.148 (75.694)	0.236 (0.027)	1.533 (2.021)	101.286 (5.312)
EGGIE	0.345 (1.493)	2.407 (9.371)	3.563 (2.911)	14.277 (42.500)
EGLIE	0.614 (0.737)	2.715 (2.461)	3.455 (1.158)	13.333 (9.358)

Table 9 Log-likelihood, goodness-of-fit statistics and information criteria

Model	ℓ	AIC	AIC _c	BIC	K–S	W*
EGPIE	− 189.020	386.035	386.776	394.345	0.065	0.062
EGBIE	− 188.950	385.893	386.634	394.203	0.082	0.075
EGGIE	− 188.360	384.720	385.461	393.030	0.211	0.047
EGLIE	− 188.330	384.658	385.398	392.968	0.070	0.039

The bold values means best based on the model selection technique

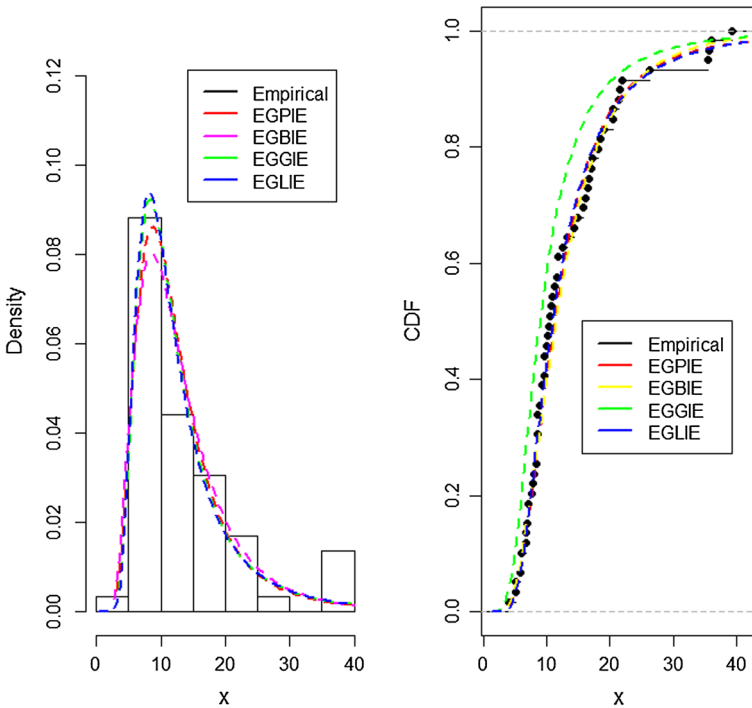


Fig. 7 Empirical and fitted density and CDF plots of tax data

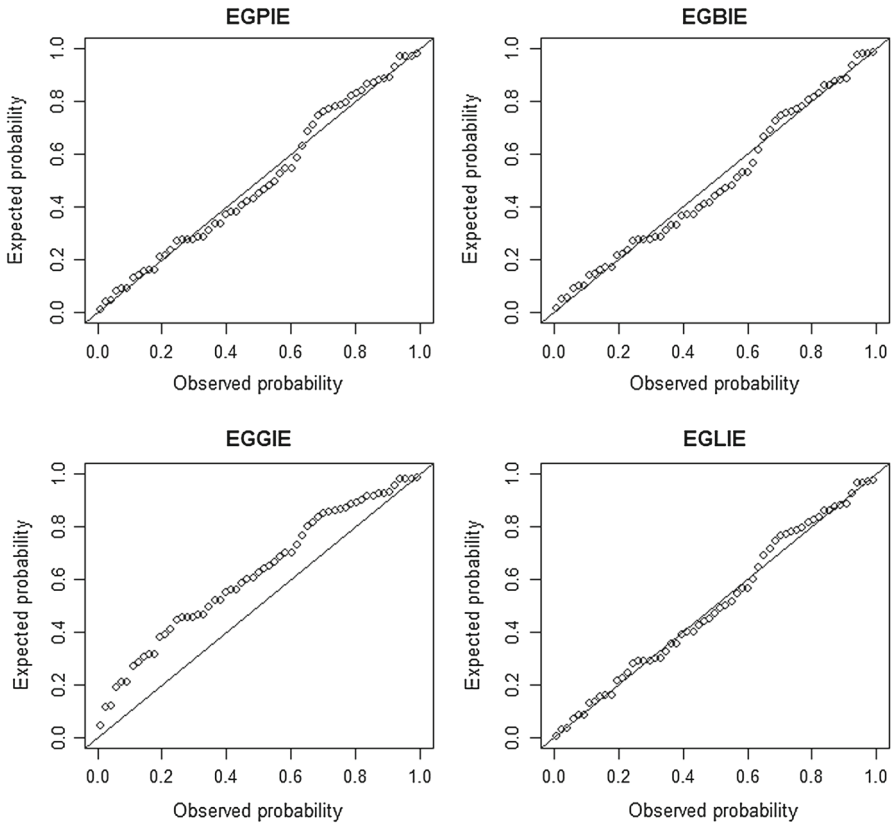


Fig. 8 P–P plots of the fitted distributions for tax data

In addition, the P–P plots in Fig. 6 shows that the EGPIE and EGBIE distributions provide a more reasonable fit to the data compared to the EGGIE and EGLIE distributions.

9.2 Second Data Set

The data set comprises monthly actual taxes revenue in Egypt from January 2006 to November 2010 in 1000 million Egyptian pounds. The data can be found in [14] and are given in Table 7.

The maximum likelihood estimates with their corresponding standard errors in brackets for the fitted distributions are given in Table 8.

The EGLIE distribution provides a better fit to the tax data as compared to the other fitted distributions. With the exception of the K–S, all other model selection technique selected the EGLIE distribution as the best distribution for the tax data as shown in Table 9.

Figure 7 displays the plots of the empirical density, the fitted distributions, the empirical CDF and the CDF of the fitted distributions for the tax data.

The P–P plots for the fitted distributions are given in Fig. 8. From Fig. 8, it can be seen that the EGLIE distribution provides a more reasonable parametric fit to the tax data.

10 Conclusion

This study proposed and studied the properties of the EGPS family of distributions. Various statistical properties such as the quantile function, moment, moment generating function, incomplete moment, reliability, residual life, mean residual life, Shannon entropy and order statistics were derived. The method of maximum likelihood estimation was proposed for the estimation of the parameters of the family. Bivariate and multivariate extensions of the family was proposed using the Clayton copula. Some special distributions were defined and Monte Carlo simulations were performed to investigate the finite sample properties of the estimators for the parameters of the special distributions. The results of the simulation revealed the parameters of the distributions were stable with regards to the estimation techniques. Finally, applications of the special distributions were illustrated using real data sets.

Acknowledgements 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.

Compliance with Ethical Standards

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

References

1. Afify AZ, Alizadeh M, Yousof HM, Aryal G, Ahmad M (2016) The transmuted geometric-G family of distributions: theory and applications. *Pak J Stat* 32:139–160
2. Afify AZ, Cordeiro GM, Yousof HM, Alzaatreh A, Nofal ZM (2016) The Kumaraswamy transmuted-G family of distributions: properties and applications. *J Data Sci* 14:245–270
3. Afify AZ, Yousof HM, Nadarajah S (2017) The beta transmuted-H family of distributions: properties and applications. *Stat Its Inference* 10:505–520
4. Alzaatreh A, Lee C, Famoye F (2013) A new method for generating families of continuous distributions. *Metron* 71(1):63–79
5. Alzaghal A, Famoye F, Lee C (2013) Exponentiated T–X family of distributions with some applications. *Int J Stat Probab* 2(3):31–49
6. Andrews DF, Herzberg AM (2012) *Data: a collection of problems from many fields for the student and research worker*. Springer, New York
7. Barlow R, Toland R, Freeman T (1984) A Bayesian analysis of stress-rupture life of Kevlar 49/epoxy spherical pressure vessels. In: *Proceedings of conference on applications of statistics*. Marcel Dekker, New York
8. Bolker B (2014) *Tools for general maximum likelihood estimation*. R development core team
9. Bourguignon M, Silva RB, Cordeiro GM (2014) The Weibull-G family of probability distributions. *J Data Sci* 12:53–68
10. Cordeiro GM, de Castro M (2011) A new family of generalized distributions. *J Stat Comput Simul* 81(7):883–898

11. Cordeiro GM, Ortega EMM, da Cunha CCD (2013) The exponentiated generalized class of distributions. *J Data Sci* 11:1–27
12. Cordeiro GM, Ortega EMM, Silva GO (2012) The beta extended Weibull family. *J Probab Stat Sci* 10(10):15–40
13. Gradshteyn IS, Ryzhik IM (2007) *Tables of integrals, series, and products*. Academic Press, New York
14. Mead M (2014) An extended Pareto distribution. *Pak J Stat Oper Res* 10(3):313–329
15. Nasiru S, Mwita PN, Ngesa O (2017) Exponentiated generalized exponential Dagum distribution. *J King Saud Univ Sci*. <https://doi.org/10.1016/j.jksus.2017.09.009>
16. Nasiru S, Mwita PN, Ngesa O (2017) Exponentiated generalized transformed-transformer family of distributions. *J Stat Econ Methods* 6(4):1–17
17. Noack A (1950) A class of random variables with discrete distributions. *Ann Math Stat* 21(1):127–132
18. Nofal ZM, Afify AZ, Yousof HM, Cordeiro GM (2017) The generalized transmuted-G family of distributions. *Commun Stat Theory Methods* 46:4119–4136
19. Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–432
20. Sklar A (1959) Fonctions de repartition à n dimension et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* 8:229–231
21. Tahir MH, Cordeiro GM, Alizadeh M, Mansoor M, Zubair M, Hamedani GG (2015) The odd generalized exponential family of distributions with applications. *J Stat Distrib Appl* 2(1):1–28
22. Yousof HM, Afify AZ, Alizadeh M, Butt NS, Hamedani GG, Ali MM (2015) The transmuted exponential generalized-G family of distributions. *Pak J Stat Oper Res* 11:441–464