Pak. J. Statist.
2020 Vol. 36(4), 287-304

WEIGHTED GAMMA-PARETO DISTRIBUTION AND ITS APPLICATION

Aijaz Ahmad Dar[1§], Aquil Ahmed[2] and **Javaid Ahmad Reshi[3]**

[1] Department of Mathematics and Actuarial Science
BSAR Crescent Institute of Science and Technology
Vandalur, Chennai, Tamil Nadu, India.

[2] Department of Statistics and Operations Research
Aligarh Muslim University, Aligarh, U.P., India.

[3] Govt. Degree College Anantnag, J&K, India.

[§] Corresponding author Email: aijazamu9@gmail.com

ABSTRACT

The current article explores a probability distribution termed as weighted gamma-Pareto distribution (WGPD) by employing the notion of weighted probability and the concept of T-X family. It is assumed that the probability of an observation to get ascertained/ recorded is proportional to distribution function of Pareto. Therefore the weight function considered in this article is distribution function of Pareto. One of the members of T-X family known as gamma-Pareto distribution with transformed random variable following gamma and transformer random variable following Pareto is considered. The explored distribution acts as a generalization of some of the well-known probability distribution viz., gamma-Pareto distribution, weighted exponential-Pareto distribution, Pareto distribution, weighted gamma distribution, gamma distribution, weighted exponential distribution, generalized exponential distribution and exponential distribution. Different properties of WGPD have been studied. Finally the application of the introduced distribution is illustrated through its fitting to a real life and a simulated data set.

KEYWORDS

Weighted probability distribution, T-X family, M.L.E., Von-Neumann sampling, Akaike information criterion (AIC).

1. INTRODUCTION

There are various techniques by virtue of which one can introduce a new extension of an already existing probability model and can result in a vast variety of probability distributions. Some of these techniques include the concept of generalized distributions (e.g., Kumaraswamy G-distribution by Cordeiro and Castro (2011), Marshell-Olkin G-distribution due to Marshell and Olkin (1997), beta-extended family of distribution by Cordeiro et al. (2012) etc.), compound distribution, slash distribution, convolution, T-X family due to Alzaatreh et al. (2013), truncation, transmutation and many more. Different techniques result in the formation of different models which can be used in analyzing a vast variety of random phenomenon, e.g., two different approaches of introducing the weighted version of an exponential distribution result in the formation of two different models, viz., gamma and generalized exponential distribution. Alzaatreh et al. (2013) introduced a new approach of deriving a vast family of probability distributions known as T-X family after considering two random variables one known as *transformer* and another the *transformed*. The methodology behind the derivation of T-X family is given in Section 2. Since the concept of T-X family came into existence, it is widely used to derive various probability distributions on considering different combinations of *transformer* and *transformed* random variables see e.g., Weibull-Pareto distribution by Alzaatreh et al. (2013), gamma-normal distribution by Alzaatreh et al. (2014), logestic-x family of distribution due to Tahir et al. (2016), Korakmaz (2018) etc.

This article is devoted to the study of weighted version of gamma-Pareto distribution (GPD). To discuss the derivation of weighted gamma-Pareto distribution it is imperative to have a brief outline about the importance and genesis of weighted probability. Every statistical investigation begins with the collection of data by making using of various sampling techniques. In other words, one can say that sampling plays a fundamental role in every sort of statistical investigation. Quite often people don't look thoroughly into the sampling techniques used in drawing a sample from the underlined population as a result of which some of the rare events which exist in the universe get unnoticed, unobserved or simply unrecognized. The best example of such a rare event is the family with an albino child. Due to the un-ascertainment of such a rare event, the results drawn can't be considered valid in the real life scenario because the inference is to be drawn about the whole population which contains the rare cases also. Therefore, to make a valid inference about the whole population, it is necessary to look into the methods of ascertainment of rare events. Initially, it was Fisher (1934), who studied how the methods of ascertainment can affect the estimation of frequencies and laid down the concept of weighted probability distribution. The concept of Fisher (1934) was later formulated in general terms by Rao (1965). It is obvious, that in a population consisting of unequally likely observations, different events will have different probability of getting recorded. To deal with such a population a function known as weight function is searched for expressing the proportionality in which the events are observed/ ascertained. Consider a random variable X with $f(x)$ as its density function and let us assume the probability of observing $X = x$ is proportional to a weight function $w(x) \geq 0$. Therefore by the definition of weighted probability the density function of observed X is given by.

$$f_w(x) = \frac{w(x) f(x)}{\int w(x) f(x) dx} = \frac{w(x) f(x)}{E[w(x)]} \tag{1}$$

The density function given by eqn. (1) is referred to as the weighted density and the function $w(x) \geq 0$ as weight function. Different situations give rise to different forms of weight functions and some of weight functions suggested by various authors can be seen in Patil, Rao and Ratnaparkhi (1986).

In this article, gamma-Pareto distribution (GPD) introduced by Alzaatreh et al. (2012) is treated as baseline/original distribution and its weighted version is suggested on assuming the probability of ascertainment is proportional to distribution function of Pareto random variable. The derivation of the weighted version of gamma-Pareto distribution is given in Section 2.

2. DERIVATION OF WEIGHTED GAMMA-PARETO DISTRIBUTION

In order to generate the beta-generated family of distributions or Kumaraswamy G-family of distributions, one is confined to use the generator with [0,1] as the support set. This confinement led Alzaatreh et al. (2013) to think about the use of generator with support other than [0,1] and explored the concept of transformed-transformer family of distributions also known as T-X family the definition of which is given as follows.

Let $F(x)$ and $v(t)$ be the respective c.d.f. and p.d.f. of a continuous random X and T such that support set of T is $[a,b]$, $-\infty \leq a < b \leq \infty$. Let $\exists$ a function say $W(.)$ satisfying the following conditions.

i. $W(F(x)) \in [a,b]$.
ii. $W(F(x))$ is differentiable and monotonically non-decreasing function.
iii. $W(F(x)) \to a$ as $x \to -\infty$ and $W(F(x)) \to b$ as $x \to \infty$.

Then, according to Alzaatreh et al. (2013) the distribution and density function of T-X family are respectively given by eqn. (2) and eqn. (3).

$$G(x) = \int_a^{W(F(x))} v\,(t)dt = \Pr[T \leq W(F(x))] = V(W(F(x))). \tag{2}$$

$$g(x) = \left\{\frac{d}{dx} W(F(x))\right\} v\{W(F(x))\}, \tag{3}$$

where $V(t) = V(T \leq t)$ is the c.d.f. of T. From eqn. (2), it can be seen that the distribution function of T-X family is a composition of V, W and F, i.e., $G(x) = (V\text{o}(W\text{o}(F)))(x)$.

The most interesting property of T-X family is that it can be used to generate a vast variety of probability distributions on considering different combinations T, X and $W(.)$. For example if $T \in [0,1]$ and $W(F(x)) = F(x)$ or $W(x) = [F(x)]^{\alpha}$, then the corresponding T-X family will be same as the beta- generated family of distribution which was suggested by Eugene et al. (2002). Moreover, if $W(F(x)) = F(x)$ and $T \in [a,b]$, then the density given by eqn. (3) reduces to:

$$g(x) = \left\{\frac{d}{dx} F(x)\right\} v\{F(x)\} = f(x)v(F(x)),$$

which is same as the density of univariate skew distribution introduced by Ferreira and Steel (2006) with weight function $v(.)$.

Alzaatreh et al. (2012) introduced gamma-Pareto distribution on assuming $T \sim \text{gamma}(\alpha,\beta)$ and $X \sim \text{Pareto}(\theta,k)$ random variable respectively. $W(F(x)) = -\log[1 - F(x)]| = H(x)$ was used for the derivation of GPD. The derived distribution and density function of the resulting gamma-Pareto random variable are respectively given by eqn. (4) and eqn. (5).

$$G(x;\alpha,\beta,k,\theta) = \frac{\gamma\left(\alpha, k\beta^{-1}\log\left(\frac{x}{\theta}\right)\right)}{\Gamma(\alpha)}, x > \theta \geq 0, k, \alpha, \beta > 0. \tag{4}$$

$$g(x;\alpha,\beta,k,\theta) = \frac{k^{\alpha}(\theta/x)^{k\beta^{-1}}\left(\log(x/\theta)\right)^{\alpha-1}}{x\beta^{\alpha}\Gamma(\alpha)}, x > \theta \geq 0, k, \alpha, \beta > 0. \tag{5}$$

Re-parameterizing eqn. (4) and eqn. (5) by $k\beta^{-1} = \lambda$ we get

$$G(x;\alpha,\lambda,\theta) = \frac{\gamma\left(\alpha,\lambda\log(x/\theta)\right)}{\Gamma(\alpha)}, x > \theta \geq 0, \alpha, \lambda > 0. \quad (6)$$

$$g(x;\alpha,\lambda,\theta) = \frac{\lambda^{\alpha}(\theta/x)^{\lambda}\left(\log(x/\theta)\right)^{\alpha-1}}{x\Gamma(\alpha)}, x > \theta \geq 0, \alpha, \lambda > 0. \quad (7)$$

where, $\gamma(a,b) = \int_0^b x^{a-1}\, e^{-x} dx$ is known as lower incomplete gamma integral.

Theorem 1

GPD is form invariant under size biased sampling.

Proof:

The size biased version of order ω of a density $d(x)$ is given by

$$d_{SB}(x;\omega) = \frac{x^{\omega} d(x)}{\int x^{\omega} d(x) dx},\ \omega > 0.$$

Therefore, the size biased version of GPD is given as:

$$g_{SB}(x;\alpha,\lambda,\theta,\omega) = \frac{x^{\omega} g(x;\alpha,\lambda,\theta)}{\int_{\theta}^{\infty} x^{\omega} g(x;\alpha,\lambda,\theta) dx}$$

$$g_{SB}(x;\alpha,\lambda,\theta,\omega) = \frac{\frac{\lambda^{\alpha}\theta^{\lambda}x^{\omega-\lambda-1}}{\Gamma(\alpha)}\left(\log(x/\theta)\right)^{\alpha-1}}{\int_{\theta}^{\infty}\frac{\lambda^{\alpha}\theta^{\lambda}x^{\omega-\lambda-1}}{\Gamma(\alpha)}\left(\log(x/\theta)\right)^{\alpha-1} dx}$$

$$g_{SB}(x;\alpha,\lambda,\theta,\omega) = \frac{x^{\omega-\lambda-1}\left(\log(x/\theta)\right)^{\alpha-1}}{\int_{\theta}^{\infty} x^{\omega-\lambda-1}\left(\log(x/\theta)\right)^{\alpha-1} dx}.$$

Put $\log(x/\theta) = y$ in the denominator and after simplification, we get

$$g_{SB}(x;\alpha,\lambda,\theta,\omega) = \frac{(\lambda-\omega)^{\alpha}\left(\log(x/\theta)\right)^{\alpha-1}(\theta/x)^{\lambda-\omega}}{x\Gamma(\alpha)}$$

which is again the density of GPD.

2.1. Weight Function

It is shown that GPD is form invariant under size bias sampling. Thus, a new sort of weight function which is actually the c.d.f. of Pareto distribution with scale = θ and shape = $\lambda\omega$ is considered and given by:

$$w(x) = 1 - (\theta/x)^{\lambda\omega}, \theta \geq 0, \lambda, \omega > 0. \quad (8)$$

The reason for considering c.d.f. of Pareto distribution as a weight function is that it is non-negative and $w(x) \to 1$ as $\omega \to \infty$. Interestingly this property of $w(x)$ will be helpful in generalizing some of the well-known probability distributions, viz., baseline distribution, i.e., GPD by Alzaatreh et al. (2012), weighted exponential-Pareto, exponential, Pareto, gamma, weighted gamma by Jain et al. (2014), weighted exponential by Gupta and Kundu (2009), and generalized exponential distribution by Gupta and Kundu (1999).

Thus, on using the definition of weighted distribution we can write.

$$g_w(x;\omega,\alpha,\lambda,\theta)=\frac{w(x)g(x;\alpha,\lambda,\theta)}{\mathbb{E}[w(x)]}, \tag{9}$$

where $\mathbb{E}[w(x)]=1-\int_\theta^\infty\left(\frac{\theta}{x}\right)^{\lambda\omega}\frac{\lambda^\alpha\left(\frac{\theta}{x}\right)^\lambda\left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1}}{x\Gamma(\alpha)}dx.$

Put $\log(x/\theta)=y$, we get

$$\mathbb{E}[w(x)]\quad=1-\frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty e^{-\lambda(\omega+1)y}\,y^{\alpha-1}dy=1-(\omega+1)^{-\alpha}. \tag{10}$$

Therefore, on making use of eqn. (7), eqn. (8) and eqn. (10) in eqn. (9), we obtained the p.d.f. of weighted gamma-Pareto distribution (WGPD) which is given by

$$g_w(x;\omega,\alpha,\lambda,\theta)=\frac{\lambda^\alpha\theta^\lambda(1-(\theta/x)^{\lambda\omega})(\log(x/\theta))^{\alpha-1}}{(1-(\omega+1)^{-\alpha})x^{\lambda+1}\Gamma(\alpha)},x>\theta\geq 0,\omega,\alpha,\lambda>0. \tag{11}$$

Similarly, the c.d.f. and reliability function associated with WGPD are respectively given by eqn. (12) and eqn. (13).

$$G_w(x;\omega,\alpha,\lambda,\theta)=\frac{\gamma(\alpha,\lambda\log(x/\theta))-(\omega+1)^{-\alpha}\gamma(\alpha,\lambda(\omega+1)\log(x/\theta))}{\Gamma(\alpha)(1-(1+\omega)^{-\alpha})}. \tag{12}$$

$$R_w(x;\omega,\alpha,\lambda,\theta)=1-\frac{\gamma(\alpha,\lambda\log(x/\theta))-(\omega+1)^{-\alpha}\gamma(\alpha,\lambda(\omega+1)\log(x/\theta))}{\Gamma(\alpha)(1-(1+\omega)^{-\alpha})}. \tag{13}$$

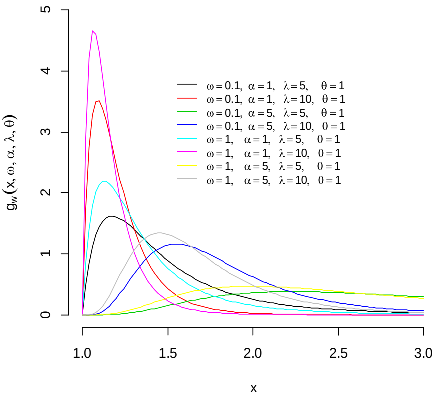

Fig. 1(a): Density.

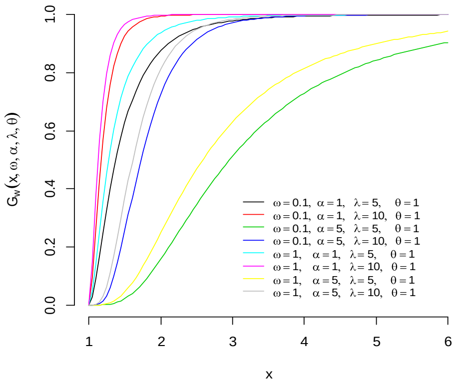

Fig. 1(b): c.d.f.

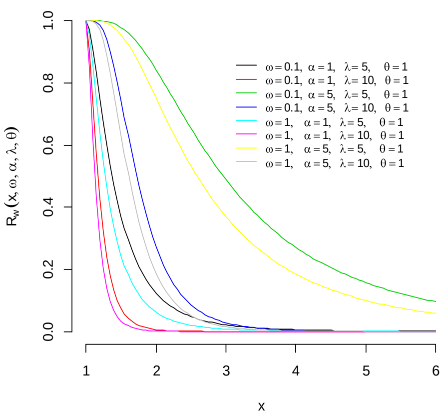

Fig. 1(c): Reliability.

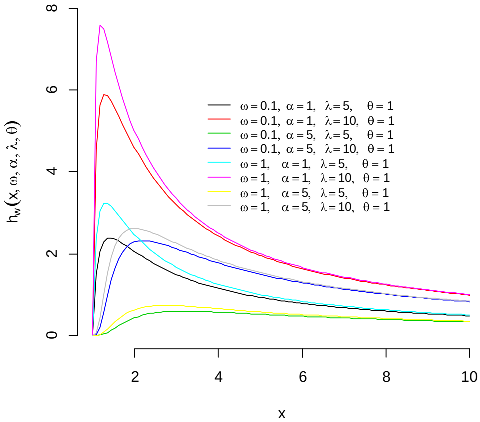

Fig. 1(d): Hazard.

Figure 1: P.d.f., c.d.f., Reliability and Hazard Rate Curves at Different Values of $\boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\lambda}$ and $\boldsymbol{\theta}$.

Notation $X \sim \text{WGPD}(\omega, \alpha, \lambda, \theta)$ wherever used in this article means a random variable X following WGPD.

3. PROPERTIES

Theorem 2

If $X \sim \text{WGPD}(\omega, \alpha, \lambda, \theta)$, then the transformation $Y = \log(X/\theta)$ follows weighted gamma distribution (WGD) introduced by Jain et al. (2014) with scale parameter λ, shape parameters α and ω.

Proof:

Using the Jacobian method, we can obtain the p.d.f. of Y as follows

$$g_w(y) = g_w(x = \theta e^y, \omega, \alpha, \lambda, \theta)\, J \tag{14}$$

where, $J = \frac{dx}{dy} = \theta e^y$ is the Jacobian of transformation $y = \log(x/\theta)$.

Therefore, on using eqn. (11) in eqn. (14) we get

$$g_w(y) = \frac{\lambda^\alpha (1 - e^{-\lambda\omega y}) y^{\alpha-1} e^{-\lambda y}}{(1 - (\omega+1)^{-\alpha}) \Gamma(\alpha)}, y > 0, \omega, \alpha, \lambda > 0,$$

which is same as the density of weighted gamma distribution introduced by Jain et al. (2014).

Corollary 1

If $X \sim \text{WGPD}(\omega, \alpha = 1, \lambda, \theta)$, then the transformation $Y = \log(X/\theta)$ follows:

i. weighted exponential distribution (WED) introduced by Gupta and Kundu (2009) with scale parameter λ and shape parameter ω.

ii. Jone's model $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} c e^{-acy} (1 - e^{-cx})^{b-1}$ with $a = \frac{1}{\omega}$, $b = 2$ and $c = \lambda\omega$, see Jones (2004) and Nadarajah and Kotez (2005).

Theorem 3

Density of WGPD can be expressed as a linear combination of densities of two gamma-Pareto distributions with respective scale parameters λ, $\lambda(\omega+1)$ and with same shape parameter α.

Proof:

Eqn. (11) can be rewritten as

$$g_w(x;\omega,\alpha,\lambda,\theta) = \frac{1}{(1-(\omega+1)^{-\alpha})}\left[\frac{\lambda^{\alpha}\theta^{\lambda}\left(\log(x/\theta)\right)^{\alpha-1}}{x^{\lambda+1}\Gamma(\alpha)} - \frac{\lambda^{\alpha}\theta^{\lambda(\omega+1)}\left(\log(x/\theta)\right)^{\alpha-1}}{x^{\lambda(\omega+1)+1}\Gamma(\alpha)}\right]$$

$$g_w(x;\omega,\alpha,\lambda,\theta) = \frac{1}{(1-(\omega+1)^{-\alpha})}\left[\frac{\lambda^{\alpha}\theta^{\lambda}\left(\log(x/\theta)\right)^{\alpha-1}}{x^{\lambda+1}\Gamma(\alpha)} - \frac{\lambda^{\alpha}\,(\omega+1)^{\alpha}\theta^{\lambda(\omega+1)}\left(\log(x/\theta)\right)^{\alpha-1}}{(\omega+1)^{\alpha}x^{\lambda(\omega+1)+1}\Gamma(\alpha)}\right]$$

$$g_w(x;\omega,\alpha,\lambda,\theta) = m\, g(x;\alpha,\lambda,\theta) - n\, g(x;\alpha,\lambda(\omega+1),\theta),$$

where, $m = (1-(\omega+1)^{-\alpha})^{-1}$ and $n = (1-(\omega+1)^{\alpha})^{-1}$.

Theorem 4

If $X \sim$ GPD(α,λ,θ) and $X_w \sim$ WGPD$(\omega,\alpha,\lambda,\theta)$ then X is smaller than X_w in stochastic, failure rate, likelihood ratio and mean residual life ordering.

Proof:

To prove the statement, it is enough to show that X is smaller than X_w in likelihood ratio ordering i.e., ($X \leq_{LR} X_w$) and due to the chain of implications given in eqn. (15) other three partial orderings will follow accordingly see Lai and Xie (2006), Shaked and Shanthikumar (2007) and Jain et al. (2014).

$$\begin{array}{ccccc} X \underset{LR}{\leq} Y \Rightarrow & X & \underset{FR}{\leq} Y \Rightarrow X \underset{ST}{\leq} Y \\ & & \Downarrow & \\ & X & \underset{MR}{\leq} Y \end{array} \qquad (15)$$

Dividing eqn. (7) by eqn. (11) we get the following ratio:

$$\frac{g(x;\alpha,\lambda,\theta)}{g_w(x;\omega,\alpha,\lambda,\theta)} = \frac{1-(\omega+1)^{-\alpha}}{1-\left(\frac{\theta}{x}\right)^{\lambda\omega}}. \qquad (16)$$

Differentiating eqn. (16) w.r.t. x we get

$$\frac{\partial}{\partial x}\frac{g(x;\alpha,\lambda,\theta)}{g_w(x;\omega,\alpha,\lambda,\theta)} = -\frac{\lambda\omega[1-(\omega+1)^{-\alpha}]\theta^{\lambda\omega}}{x^{\lambda\omega+1}[1-(\theta/x)^{\lambda\omega}]^2} < 0\ \forall x,\omega,\alpha,\lambda,\theta.$$

Since, the gradient of ratio of two densities w.r.t. x comes out to be negative implying that it is decreasing in x. Hence, it can be concluded that $X \leq_{LR} X_w$ and due to the chain of implications given in eqn. (15) the rest of three partial orderings follow according, i.e., $X \leq_{FR} X_w$, $X \leq_{ST} X_w$ and $X \leq_{MR} X_w$. The statement can be also justified by using the fact that a weighted version used to be greater than the baseline distribution in likelihood ratio if the considered weighted function is monotonic increasing see Patil et al. (1986).

Since the weight function considered here is the c.d.f. of Pareto distribution which is monotonic increasing hence $X \leq_{LR} X_w$.

3.1. Descriptive Measures

The r^{th} moment about origin of WGPD is given by

$$\mu_r' = \frac{\lambda^\alpha \theta^r ((\lambda - r)^{-\alpha} - (\lambda\omega + \lambda - r)^{-\alpha})}{(1 - (\omega + 1)^{-\alpha})}, \lambda > r, r = 1,2,3,\ldots \quad (17)$$

Table 1
Characteristics of WGPD at Different Values of ω, α, λ and θ

θ	λ	α	ω	Mean μ	Variance σ^2	CV cv	Skewness γ_1	Kurtosis γ_2
5	5	5	0.5	16.022	95.678	0.611	5.452	183.227
			1.0	15.478	89.848	0.612	5.515	187.587
		10	0.5	47.019	1986.930	0.948	9.355	1522.912
			1.0	46.598	1967.018	0.952	9.366	1528.816
	15	5	0.5	7.179	1.315	0.160	1.539	4.573
			1.0	7.096	1.283	0.160	1.564	4.679
		10	0.5	10.005	5.204	0.228	1.503	4.520
			1.0	9.971	5.211	0.229	1.499	4.496
10	5	5	0.5	32.044	382.710	0.611	5.452	183.227
			1.0	30.956	359.391	0.612	5.515	187.587
		10	0.5	94.038	7947.719	0.948	9.355	1522.912
			1.0	93.195	7868.071	0.952	9.366	1528.816
	15	5	0.5	14.357	5.259	0.160	1.539	4.573
			1.0	14.193	5.132	0.160	1.564	4.679
		10	0.5	20.009	20.816	0.228	1.503	4.520
			1.0	19.942	20.842	0.229	1.499	4.496

Variance is given by

$$\sigma^2 = \frac{\theta^2 \lambda^\alpha}{((\omega+1)^{-\alpha} - 1)^2} [\{1 - (\omega + 1)^{-\alpha}\}\{(\lambda - 2)^{-\alpha} - (\lambda\omega + \lambda - 2)^{-\alpha}\} \\ -\lambda^\alpha \{(\lambda - 1)^{-\alpha} - (\lambda\omega + \lambda - 1)^{-\alpha}\}^2], \lambda > 2. \quad (18)$$

Mathematical expression for the coefficient of variation, skewness and kurtosis are not given in closed form but their estimates are obtained numerically at different values of parameters and are reported in Table 1. From Table 1, it is can be observed that on increasing θ, mean and variance increases whereas coefficient of variation, skewness and kurtosis remains unaltered. Increase in λ leads to decrease in all of the five measures, i.e., mean, variance, coefficient of variation, skewness and kurtosis. With the increase in α, mean, variance and coefficient of variation increases whereas skewness and kurtosis shows unevenness, i.e., initially increases and then starts decreasing with the increase in λ. Increase in ω leads to decline in mean and variance whereas the other three measures, viz., coefficient of variation, skewness and kurtosis increase.

3.2. Entropy Measure

Theorem 5

Renyi entropy associated with WGPD is given by

$$H_R(\delta) = \frac{1}{1-\delta}\log\left[\frac{\lambda^{\delta\alpha}\Gamma(\delta(\alpha-1)+1)}{(1-(\omega+1)^{-\alpha})^{\delta}(\Gamma(\alpha))^{\delta}\theta^{\delta-1}} \times \sum_{k=0}^{\infty}\frac{(-1)^{\delta+k}\binom{\delta}{k}}{(\lambda\omega\delta+\lambda\delta+\delta-\lambda\omega k-1)^{\delta(\alpha-1)+1}}\right], \delta \geq 0, \delta \neq 1. \quad (19)$$

Proof:

From the definition of Renyi entropy, we can write

$$H_R(\delta) = \frac{1}{1-\delta}\log\left[\int_{\theta}^{\infty}[g_w(x;\omega,\alpha,\lambda,\theta)]^{\delta}dx\right]$$

$$H_R(\delta) = \frac{1}{1-\delta}\log\left[\int_{\theta}^{\infty}\frac{\lambda^{\delta\alpha}\theta^{\delta\lambda}[1-\left(\frac{\theta}{x}\right)^{\lambda\omega}]^{\delta}[\log\left(\frac{x}{\theta}\right)]^{\delta(\alpha-1)}}{[1-(\omega+1)^{-\alpha}]^{\delta}x^{\delta(\lambda+1)}(\Gamma(\alpha))^{\delta}}dx\right]$$

Substitute $\log(x/\theta) = y$ we get

$$H_R(\delta) = \frac{1}{1-\delta}\log\left[\frac{\lambda^{\delta\alpha}\theta^{\delta\lambda}[1-(\omega+1)^{-\alpha}]^{-\delta}}{(-1)^{-\delta}[\Gamma(\alpha)]^{\delta}\theta^{\delta(\lambda+1)-1}} \times \int_0^{\infty}(e^{-\lambda\omega y}-1)^{\delta}e^{-(\delta(\lambda+1)-1)y}y^{\delta(\alpha-1)}dy\right]$$

Put $\lambda\omega y = t$ we get

$$H_R(\delta) = \frac{1}{1-\delta}\log\left[\frac{\lambda^{\delta\alpha}\theta^{\delta\lambda}[1-(\omega+1)^{-\alpha}]^{-\delta}}{(-1)^{-\delta}[\Gamma(\alpha)]^{\delta}\theta^{\delta(\lambda+1)-1}}\int_0^{\infty}(e^{-t}-1)^{\delta}\exp\left(-\frac{(\delta(\lambda+1)-1)}{\lambda\omega}t\right)\frac{t^{\delta(\alpha-1)}}{(\lambda\omega)^{\delta(\alpha-1)+1}}dt\right]. \quad (20)$$

Consider the following integral:

$$\int_0^{\infty}(e^{-t}-1)^{n}e^{-mt}t^{\nu-1}dt = \Gamma(\nu)\sum_{k=0}^{\infty}\frac{(-1)^{k}\binom{n}{k}}{(n+m-k)^{\nu}}. \quad (21)$$

Therefore, on making use of eqn. (21) with $n=\delta, m=\frac{(\delta(\lambda+1)-1)}{\lambda\omega}, \nu=\delta(\alpha-1)+1$ in eqn. (20) we get

$$H_R(\delta) = \frac{1}{1-\delta}\log\left[\frac{\lambda^{\delta\alpha}\Gamma(\delta(\alpha-1)+1)}{(1-(\omega+1)^{-\alpha})^{\delta}(\Gamma(\alpha))^{\delta}\theta^{\delta-1}} \times \sum_{k=0}^{\infty}\frac{(-1)^{\delta+k}\binom{\delta}{k}}{(\lambda\omega\delta+\lambda\delta+\delta-\lambda\omega k-1)^{\delta(\alpha-1)+1}}\right].$$

4. RANDOM NUMBER GENERATION

Random numbers from WGPD are generated by employing *Rejection* method also known as Von-Neumann sampling, on treating the Pareto with shape = k and scale = θ_1 as proposal density. Thus, the ratio of target to proposal density is given by

$$\rho(y) = \frac{g_w(y;\omega,\alpha,\lambda,\theta)}{k\theta_1 y^{-(k+1)}}$$

$$\rho(y) = \frac{\lambda^\alpha \theta^\lambda \left(1-(\theta/y)^{\lambda\omega}\right)\left(\log(y/\theta)\right)^{\alpha-1} y^{(k-\lambda)}}{(1-(\omega+1)^{-\alpha})\Gamma(\alpha)k\theta_1} \tag{22}$$

Differentiating eqn. (22) w.r.t. y we get:

$$\frac{\partial\rho(y)}{\partial y} = \frac{1}{k\theta_1((\omega+1)^\alpha-1)\Gamma(\alpha)}\left[\frac{y^{k-1}}{(\lambda(\omega+1))^{-\alpha}}\left(\frac{\theta}{y}\right)^{\lambda(1+\omega)}\left(\log\left(\frac{y}{\theta}\right)\right)^{\alpha-2} \times\left\{\log\left(\frac{y}{\theta}\right)\left((k-\lambda)\left(\frac{y}{\theta}\right)^{\lambda\omega}+\lambda\omega-k+\lambda\right)+(\alpha-1)\left(\left(\frac{y}{\theta}\right)^{\lambda\omega}-1\right)\right\}\right]. \tag{23}$$

To determine the constant m such that $\rho(y) \le m \; \forall \; y$, we solve eqn. (24) for y at preassigned parameter values of target and proposal density:

$$\frac{\partial\rho(y)}{\partial y} = 0 \tag{24}$$

After obtaining the critical value m on solving eqn. (24) such that $\frac{\partial^2\rho(y)}{\partial y^2}\Big|_{y=m} < 0$, the following two steps are followed:

i. Simulate a random number $Y = y$ and $U = u$ respectively from the Pareto(θ_1, k) and $U(0,1)$ distribution.
ii. Retain $Y = y$ and consider it a random number from WGPD$(\omega,\alpha,\lambda,\theta)$ if $u \le \frac{\rho(y)}{m}$ otherwise return to step (i).

The above discussed procedure/algorithm is written in R programming and is given in Appendix-I.

5. MAXIMUM LIKELIHOOD ESTIMATION

Let $\theta < x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$ be an ordered sample of size n from WGPD. Therefore its log likelihood function is given by

$$\log[l(\Theta\,|\,x)] = n\alpha\log\lambda - \sum_{i=1}^{n}\log x_i - n\log(\Gamma(\alpha)) - n\log[1-(\omega+1)^{-\alpha}] + n\lambda\log\theta + (\alpha-1)\sum_{i=1}^{n}\log\left(\log\left(\frac{x_i}{\theta}\right)\right) - \lambda\sum_{i=1}^{n}\log x_i + \sum_{i=1}^{n}\log\left(1-\left(\frac{\theta}{x_i}\right)^{\lambda\omega}\right) \tag{25}$$

Before moving forward to find the estimates of parameters, it is worth to discuss the procedure of estimating an unknown parameter, support of the distribution depends on. Smith (1985) discussed the maximum likelihood estimation in a class of non-regular cases wherein the probability density is zero for x less than an unknown parameter θ. Distribution of such a nature considered by Smith (1985) is three-parameter versions of Weibull, gamma, beta and log gamma. The density of all these distributions is zero for x less than an unknown parameter θ in other words one can say that support of the distribution depends on θ (threshold parameter). Smith (1985) suggested estimation of the unknown parameter θ by sample minimum $x_{(1)}$ and the remaining parameters by maximum likelihood estimation after excluding $x_{(1)}$ from the sample. In parameter estimation of WGPD the case is same as the density is zero for $x < \theta$ moreover the likelihood function is increasing w.r.t. θ. Therefore the maximum likelihood estimator (M.L.E.) of θ under the restriction $x < \theta$ is $x_{(1)}$, i.e., $\hat{\theta}_{mle} = x_{(1)}$ The M.L.E.'s of remaining three parameters, viz., α, ω and λ are obtained by using maximum likelihood estimation procedure after excluding $x_{(1)}$ from the sample.

Differentiating eqn. (25) w.r.t. α, ω, λ after replacing θ by its M.L.E. $x_{(1)}$ and excluding $x_{(1)}$ form the sample we get the respective gradients. Equating the derived gradients to zero, we obtained the following system of three nonlinear equations:

$$n\log\lambda - n\Psi(\alpha) + \frac{n\log(1+\omega)}{1-(1+\omega)^{\alpha}} + \sum_{x_i \neq x_{(1)}} \log\left(\log\left(\frac{x_i}{x_{(1)}}\right)\right) = 0 \qquad (26)$$

$$\sum_{x_i \neq x_{(1)}} \frac{\lambda\log\left(x_{(1)}/x_i\right)\left(x_{(1)}/x_i\right)^{\lambda\omega}}{\left(x_{(1)}/x_i\right)^{\lambda\omega} - 1} - \frac{n\alpha}{(1+\omega)((1+\omega)^{\alpha}-1)} = 0. \qquad (27)$$

$$\frac{n\alpha}{\lambda} + n\log x_{(1)} - \sum_{x_i \neq x_{(1)}} \frac{\omega(x_{(1)}/x_i)^{\lambda\omega}\log\left(\frac{x_{(1)}}{x_i}\right)}{1-(x_{(1)}/x_i)^{\lambda\omega}} - \sum_{x_i \neq x_{(1)}} \log x_i = 0. \qquad (28)$$

where $\Psi(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$. It is very hectic to solve eqn. (26), eqn. (27) and eqn. (28) manually because of being non-linear in nature. Therefore the derived system of nonlinear equations is defined in R programming as given in Appendix-II and solved by using the function "nleqslv".

The above defined system of equations can be solved by using function "nleqslv" with the following general syntax in R programming:

```
>nleqslv(xstart, S, data,...)
xstar= vector of guess values.
S= name of the system to get solved.
data= numeric vector to be supplied.
...= additional arguments.
```

6. APPLICATION

For illustration purpose two types of data sets are considered and WGPD along with its special cases is fitted to them. Performance of WGPD and its special cases is studied in terms of fitting and the model of best fit for the considered data sets is found by using Akaike information criterion (AIC). Two types of data sets considered include a real life and a simulated one, and are given as follows:

6.1. Simulated Data

A random sample of size 100 is simulated from WGPD with $\omega = 0.5, \alpha = 5, \lambda = 5$, $\theta = 10$ by employing the algorithm discussed in Section 4. The scale and shape of the *proposal* density is taken to be 10 and 1 respectively. After fixing the values of the parameters, following R command is executed to generate a random sample.

```
>Simulated_data<-VNsample(100,500,20,0.5,5,5,10,10,1)
>Simulated_data
$Data
[1]   22.85 142.12 18.95 21.17 25.27 30.64 33.98 24.13 24.20
[10]  23.22  36.09 36.14 73.39 47.21 26.88 84.48 27.67 38.23
[19]  37.15  33.69 17.77 28.97 18.52 24.50 20.10 49.70 32.45
[28]  18.72  15.36 19.35 15.20 64.05 33.93 14.56 24.95 25.20
[37]  37.70  55.20 28.85 46.61 15.25 45.58 25.38 27.31 48.15
[46]  22.67  60.37 28.69 22.18 16.65 99.43 17.21 59.21 19.28
[55]  54.98  24.19 14.07 22.82 21.01 26.34 22.52 36.16 16.99
[64]  27.26  19.63 37.50 29.20 35.24 17.69 65.48 27.19 45.12
[73]  15.41  16.79 19.06 22.63 24.88 27.36 17.46 62.47 20.03
[82]  72.01  20.81 29.99 18.89 15.84 37.78 46.55 73.07 31.18
[91]  21.82  45.09 21.72 30.30 28.53 47.77 34.53 33.35 24.11
[100] 45.24
$Point_Max.
[1] 28.609356 2.535282
```

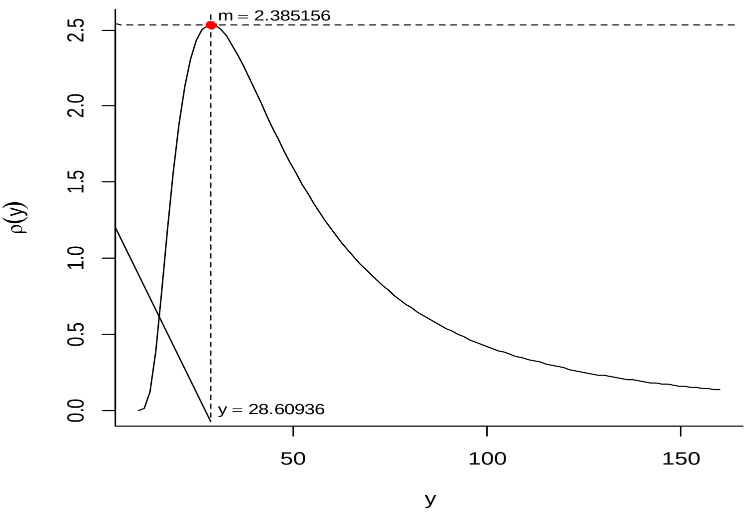

Figure 2: Ratio of Target to Proposal Density at
$\boldsymbol{\omega = 0.5, \alpha = 5, \lambda = 5, \theta = 10, \theta_1 = 10, k = 1}$

From Figure 2 it is clear that the ratio of target to proposal density ($\rho(y)$) at assigned values of parameters is increasing function and the value of constant m comes out to be 2.535282.

6.2. Real Life Data

The real life data set considered is related to flood discharge which is reported in Alzaatreh et al. (2012) and is given as follows:

> 1460, 4050, 3570, 2060, 1300, 1390, 1720, 6280, 1360, 7440, 5320, 1400, 3240, 2710, 4520, 4840, 8320, 13900, 71500, 6250, 2260, 318, 1330, 970, 1920, 15100, 2870, 20600, 3410, 726, 7500, 7170, 2000, 829, 17300, 4740, 13400, 2940, 5660.

Table 2
MLE's, -log Likelihood and AIC

Data	Distribution	M.L.E's				$-\log$ $(likelihood)$	AIC
		$\hat{\theta}_{mle}$	$\hat{\alpha}_{mle}$	$\hat{\lambda}_{mle}$	$\hat{\omega}_{mle}$		
Simulated data	WGPD	14.05	1.8981	3.0544	4.7707	290.5741	589.1482
	WEPD	14.05	...	0.1078	119.0706	559.5579	1125.116
	GPD	14.05	2.1436	3.3464	...	366.4171	738.8342
	WGD	...	8.2555	0.2914	0.756900	370.4862	746.9724
	WED	...	...	0.0703	0.000002	403.3208	810.6416
	GED	...	...	0.0514	...	405.9222	813.8445
	GD	...	7.1697	0.2520	...	373.4016	750.8031
	ED	...	...	0.0351	...	434.8147	871.6294
	PD	14.05	...	1.5769	...	382.1350	768.2700
Flood discharge data	WGPD	318	6.063	2.441327	1.6153	365.4407	738.8813
	WEPD	318	...	0.042914	40.506	437.8323	881.6646
	GPD	318	6.1351	2.4657	...	365.4521	736.9042
	WGD	...	1.7995	0.000304	2.6476	385.9541	777.9082
	WED	...	...	0.0002	17.44	383.4422	770.8843
	GED	...	...	0.000234	...	388.9989	779.9978
	GD	...	0.9196	0.000136	...	382.9048	769.8097
	ED	...	...	0.000148	...	382.9964	767.9929
	PD	318	...	0.412714	...	392.8100	789.6200

On fitting of WGPD and its special cases to the considered data sets, M.L.E.'s of parameters and AIC is estimated, and is given in Table 2. From Table 2 it can be noticed that for simulated data set WGPD possesses least AIC followed by GPD, WGD, GD, PD, WED, GED, ED and WEPD respectively. Whereas, for the real life data set it is GPD which proves to be the distribution of best fit followed in order by WGPD, ED, GD, WED, WGD, GED, PD and GEPD.

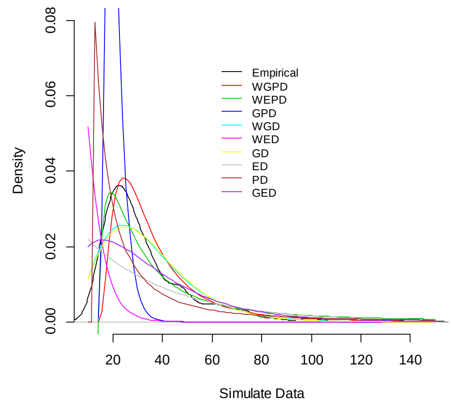

Fig. 3(a): Density.

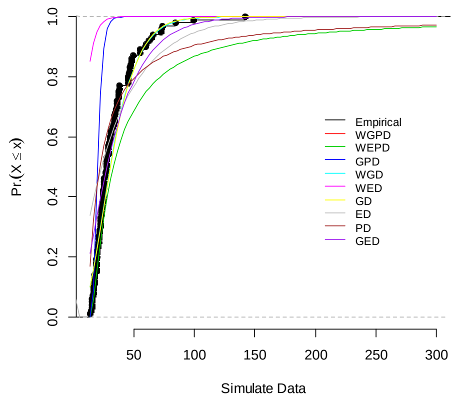

Fig. 3(b): c.d.f.

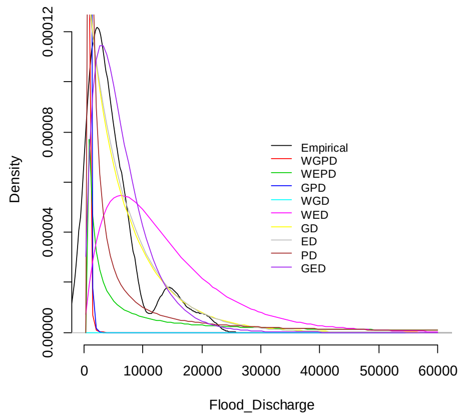

Fig. 3 (c): Density.

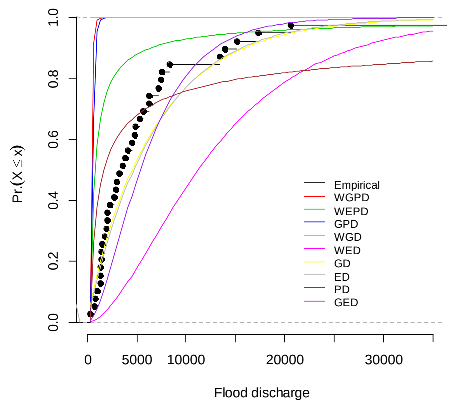

Fig. 3 (d): c.d.f.

Figure 3: Empirical Density and Distribution Curves along with the Fitted Ones

Table 3
Special Cases of Weighted Gamma-Pareto Distribution

Distribution	Transformation, Limit and Substitution	Probability Density Function
Gamma-Pareto Distribution (GPD) introduced by Alzaatreh et al. (2012)	$\omega \to \infty$(or) $\omega \to 0$	$\frac{\lambda^\alpha}{x\Gamma(\alpha)}\left[\log\left(\frac{x}{\theta}\right)\right]^{\alpha-1}\left(\frac{\theta}{x}\right)^\lambda$ (or) $\frac{\lambda^{\alpha+1}}{x\Gamma(\alpha+1)}\left[\log\left(\frac{x}{\theta}\right)\right]^{\alpha}\left(\frac{\theta}{x}\right)^\lambda, x > \theta.$
Weighted Exponential-Pareto Distribution (WEPD)	$\alpha = 1$	$[1-(\omega+1)^{-1}]^{-1}\left[1-\left(\frac{\theta}{x}\right)^{\lambda\omega}\right]\frac{\theta^\lambda}{x^{\lambda+1}}, x > \theta.$
Pareto distribution (PD)	$\alpha = 1,\ \omega \to \infty$	$\frac{\lambda\theta^\lambda}{x^{\lambda+1}}, x > \theta.$
Weighted Gamma Distribution (WGD) introduced by Jain et al. (2014)	$y = \log\left(\frac{x}{\theta}\right)$	$[1-(\omega+1)^{-\alpha}]^{-1}[1-e^{-\lambda\omega y}]\frac{\lambda^\alpha e^{-\lambda y}}{\Gamma(\alpha)}y^{\alpha-1}, y > 0.$
Gamma Distribution (GD)	$y = \log\left(\frac{x}{\theta}\right)$, $\omega \to \infty$(or) $\omega \to 0$	$\frac{\lambda^\alpha e^{-\lambda y}}{\Gamma(\alpha)}y^{\alpha-1}$ (or) $\frac{\lambda^{\alpha+1} e^{-\lambda y}}{\Gamma(\alpha+1)}y^{\alpha}, y > 0.$
Weighted Exponential Distribution (WED) introduced by Gupta and Kundu (2009).	$y = \log\left(\frac{x}{\theta}\right)$, $\alpha = 1$	$[1-(\omega+1)^{-1}]^{-1}[1-e^{-\lambda\omega y}]\lambda e^{-\lambda y}, y > 0.$
Generalized Exponential Distribution (GED) with location = 0, shape =2, Scale = $1/\lambda$) see Gupta and Kundu (1999)	$y = \log\left(\frac{x}{\theta}\right)$, $\alpha = 1,\ \omega = 1$	$2\lambda[1-e^{-\lambda x}]e^{-\lambda x}, y > 0.$
Exponential Distribution (ED)	$y = \log\left(\frac{x}{\theta}\right)$, $\omega \to \infty,\ \alpha = 1$	$\lambda e^{-\lambda y}, y > 0.$

CONCLUSION

In this article, weighted version of gamma Pareto distribution is studied. It has been shown that gamma-Pareto distribution is form invariant under size biased sampling. Therefore, a new weight function which is the c.d.f. of Pareto distribution is suggested for the first time. The weight function considered in this paper helped to generalize various known distributions, viz., gamma, exponential, Pareto, gamma-Pareto, generalized exponential, weighted gamma, weighted exponential-Pareto and weighted exponential distribution. It has been shown that WGPD can be expressed as linear combination of two gamma-Pareto distributions with different scale and same shape parameters. It is also shown that weighted gamma-Pareto distribution is greater than the base line distribution, i.e., GPD in terms of stochastic, failure rate, likelihood ratio and means residual life ordering. The random numbers from WGPD are generated by employing the rejection method after considering the Pareto as proposal density. Two types of data sets are considered including a real life and a simulated one. The simulated data set is generated at $\omega = 0.5, \alpha = 5, \lambda = 5, \theta = 10$ with shape and scale of proposal density as $\theta_1 = 10, k = 1$ respectively. Real life data set considered is related to flood discharge. WGPD along with its special cases is fitted to the considered data sets and by using AIC as a model selection tool, it has been shown WGPD proves to be the model of best fit for the simulated data set followed respectively by GPD, WGD, GD, PD, WED, GED, ED and WEPD. Whereas, for the considered real life data set it is GPD which proved to be the model of best fit followed by WGPD, ED, G, WED, WGD, GED, PD and GEPD respectively.

ACKNOWLEDGMENT

The authors are highly thankful to referees and the editor for their valuable suggestions which helped in improving the overall quality of the manuscript.

REFERENCES

1. Alzaatreh, A., Famoye, F and Lee, C. (2013). Weibull-pareto distribution and its applications. *Communications in Statistics - Theory and Methods*, 42(9), 1673-1691.
2. Alzaatreh, A., Famoye, F. and Lee C. (2012). Gamma-pareto distribution and its applications, *Journal of Modern Applied Statistical Methods*, 11(1), 78-94.
3. Alzaatreh, A., Lee, C. and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71(1), 63-79.
4. Alzaatreha, A., Famoye, F. and Lee, C. (2014). The gamma-normal distribution: Properties and applications. *Computational Statistics and Data Analysis*, 69, 67-80.
5. Cordeiro, G.M. and Castro M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81(7), 883-898.
6. Cordeiro, G.M., Silva, G.O. and Ortega, E.M.M. (2012). The beta extended Weibull family. *Journal of Probability and Statistical Science*, 10(1), 15-40.
7. Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods*, 31(4), 497-512.
8. Ferreira, J.T.S.A. and Steel, M.F.J. (2006). A constructive representation of univariate skewed distributions. *Journal of the American Statistical Association*, 101(474), 823-829.

9. Fisher, R.A. (1934). The effect of methods of ascertainment upon the estimation of frequencies, *Annals of Eugenics*, 6(1), 13-25.
10. Gupta, R.D. and Kundu, D. (2009). A new class of weighted exponential distributions. *Statistics*, 43(6), 621-634.
11. Gupta, R.D. and Kundu, D. (1999). Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2), 173-188.
12. Jain, K., Singla, N. and Gupta, R. (2014). A weighted version of gamma distribution. *Discussiones Mathematicae Probability and Statistics*, 34(1-2), 89-111.
13. Jones, M.C. (2004). Families of distributions arising from distributions of order statistics. *Test*, 13(1), 1-43.
14. Korakmaz, M.C. (2018). A new family of the continuous distributions: the extended weibull-g family. *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 68(1), 248-270.
15. Lai, C.D. and Xie, M. (2006). *Stochastic ageing and dependence for reliability*. Springer Science and Business Media.
16. Marshell, A.W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84(3), 641-652.
17. Nadarajah, S. and Kotz, S. (2005). The beta exponential distribution. *Reliability Engineering and System Safety*, 91(6), 689-697.
18. Patil, G.P., Rao, C.R. and Ratnaparkhi, M.V. (1986). On discrete weighted distributions and their use in model choice for observed data. *Communications in Statistics: Theory and Methods*, 15 (3) 907-918.
19. Rao, C.R. (1965). On discrete distributions arising out of methods of ascertainment. *Sankhya: Series A*, 27(2), 311-324.
20. Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*. Springer Science and Business Media.
21. Smith, R.L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1), 67-90.
22. Tahir, M.H., Cordeiro, G.M., Alzaatreh, A., Mansoor, M. and Zubair, M. (2016). The logistic-x family of distributions and its applications, *Communications in Statistics - Theory and Methods*, 45:24, 7326-7349.

APPENDICES

APPENDIX-I

```
> VNsample<-function(n,n1,xstart,w,a,L,t,scale,shape){#w=omega,
+ #a=alpha,L=lambda,t=theta,scale and shape of proposal density,
+ #n=required size, n1=size generated from proposal density
+ Ratio<-function(x)eval({w;a;L;t;scale;shape;((L^a)/(gamma(a)*x*
+ (1-(1+w)^(-a))))*(1-(t/x)^(L*w))*((t/x)^L)*((log(x/t))^(a-1))/(
+ scale*x^(-(shape+1)))})
+ D1<-function(x)eval({w;a;L;t;scale;shape;DD(expression(((L^a)/(
+ gamma(a)*x*(1-(1+w)^(-a))))*(1-(t/x)^(L*w))*((t/x)^L)*((log(x/t
+ ))^(a-1))/(scale*x^(-(shape+1)))),"x",1)})
+ D2<-function(x)eval({w;a;L;t;scale;shape;DD(expression(((L^a)/(
+ gamma(a)*x*(1-(1+w)^(-a))))*(1-(t/x)^(L*w))*((t/x)^L)*((log(x/t)
+ )^(a-1))/(scale*x^(-(shape+1)))),"x",2)})
+ library(nleqslv)
+ Criticalpoint=nleqslv(xstart, D1)#xstart=guess value
+ d2=D2(Criticalpoint$x)# second deriavtive test
+ if(d2 >= 0) stop("The point is not the point of Maximum")
+ if(Criticalpoint$x<t)stop("Abscissa is less than Min(Support)")
+ m=Ratio(Criticalpoint$x)#Ordinate of Ratio at critical point
+ curve(Ratio(x),xlim=c(t,t+150),xlab=expression(y),
+ ylab=expression(rho(y)),bty="l")
+ abline(v=Criticalpoint$x,h=Ratio(Criticalpoint$x),lty=2)
+ points(Criticalpoint$x,Ratio(Criticalpoint$x),cex=1,pch=19,col=2)
+ library(VGAM)
+ y<-rpareto(n1,scale,shape)
+ u<-runif(n1)
+ data=c()
+ for(i in 1:n1){
+ if (u[i]<=(1/m)*Ratio(y[i]))
+ data=c(data,y[i])
+ if (length(data)==n)break}
+ L=list(Data=data,Point_Max.=c(Criticalpoint$x,
+ Ratio(Criticalpoint$x)))
+ return(L)}
```

APPENDIX-II

```
> system<- function(x,D) {# x=vector of unknowns
+ t=min(D) # t=theta, D= a numeric vector, i.e., data set
+ d<-D[D!=min(D)] # exclude sample minimum
+ n<-length(d)
+ y <- numeric(3) # no. of equations
+ #Grandient equation w.r.t. omega (x[1])
+ y[1]=sum(sapply(1:n,function(i){((t/d[i])^(x[3]*x[1])*x[3]*
+ log(t/d[i]))/((t/d[i])^(x[3]*x[1])-1)-x[2]/((1+x[1])*
+ ((1+x[1])^x[2]-1))}))
+ #Grandient equation w.r.t. lambda (x[3])
+ y[2]=sum(sapply(1:n,function(i){x[2]/x[3]+log(t)-((t/d[i])^(x[3]*
+ x[1])*x[1]*log(t/d[i]))/(1-(t/d[i])^(x[3]*x[1]))-log(d[i])}))
+ #Grandient equation w.r.t. alpha (x[2])
+ y[3]=sum(sapply(1:n,function(i){log(x[3])-digamma(x[2])+(log(1+
+ x[1]))/(1-(1+x[1])^(x[2]))+log(log(d[i]/t))}))
+ y}
```