Journal of Data Science 377-396 , DOI: 10.6339/JDS.201804_16(2).0008

On Families of Generalized Pareto Distributions: Properties and Applications

D. Hamed[1] and F. Famoye[2] and C. Lee[2]

[1] *Department of Mathematics, Winthrop University, Rock Hill, SC, USA;* [2]*Department of Mathematics, Central Michigan University, Mount Pleasant, Michigan, USA*

In this paper, we introduce some new families of generalized Pareto distributions using the *T-R*{*Y*} framework. These families of distributions are named T-Pareto{Y} families, and they arise from the quantile functions of exponential, log-logistic, logistic, extreme value, Cauchy and Weibull distributions. The shapes of these T-Pareto families can be unimodal or bimodal, skewed to the left or skewed to the right with heavy tail. Some general properties of the T-Pareto{Y} family are investigated and these include the moments, modes, mean deviations from the mean and from the median, and Shannon entropy. Several new generalized Pareto distributions are also discussed. Four real data sets from engineering, biomedical and social science are analyzed to demonstrate the flexibility and usefulness of the T-Pareto{Y} families of distributions.

Key words: *Shannon entropy; quantile function; moment; T-X family.*

1. Introduction

The Pareto distribution is named after the well-known Italian-born Swiss sociologist and economist Vilfredo Pareto (1848-1923). Pareto [1] defined Pareto's Law, which can be stated as $N = Ax^{-a}$, where N represents the number of persons having income $\geq x$ in a population. Pareto distribution is commonly used in modelling heavy tailed distributions, including but not limited to income, insurance and city size populations.

In the literature, many applications of Pareto distribution can be found in different fields such as social studies, economics, and physics. Modelling observable environmental extreme events such as earthquakes and forest fires areas using Pareto distribution was discussed by Burroughs and Tebbens [2]. For detailed review of Pareto distribution and related topics, one may refer to Arnold [3] and the references therein.

The Pareto distribution is useful for fitting data that is skewed to the right. However, the real world data are much more complex, which may be skewed to the left or bimodal. To add more flexibility to the Pareto distribution, various generalizations were developed prior to the 1990s {e.g., Pickands [4], Johnson et al. [5] and the references therein.} During the recent decades, several new generalized Pareto distributions have been developed owing to the development of new methodologies for generating new families of distributions. Examples include the exponentiated Pareto distribution by Gupta et al. [6], the beta-Pareto distribution by Akinsete et al. [7] and the beta generalized Pareto distribution by Mahmoudi [8]. Sarabia and Prieto [9] proposed Pareto positive stable distribution to study city size data. Recently, Gómez-Déniz and Calderín-Ojeda [10, 11] developed the ArcTan Pareto distribution and successfully applied it to model insurance data and population size data.

The probability density function (PDF) of Pareto distribution is given by

$$f(x) = \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}}, x \geq \theta \tag{1}$$

where $\alpha > 0$ is a shape parameter and $\theta > 0$ is a location parameter. The cumulative distribution function (CDF) corresponding to Equation (1) is

$$F(x) = 1 - (\theta/x)^{\alpha} \tag{2}$$

Several generalizations of Equation (1) can be found in Johnson et al. [5]. Eugene et al. [12] defined the beta-generated family of distributions. The CDF of a beta-generated random variable *X* is given by

$$G(x) = \int_0^{F(x)} b(t)dt,$$

where $b(t)$ is the PDF of the beta distribution, which is used as a generator to obtain the beta-generated family of distributions and $F(x)$ is the CDF of any random variable.

Replacing beta random variable by any random variable with support (0, 1), a new family of distributions can be developed. For example, the Kum-F family proposed by Jones [13] is obtained by replacing the beta distribution with the Kumaraswamy distribution. The Kumaraswamy-Pareto distribution was studied in detail by Bourguignon et al. [14]. The use of a generator with support between 0 and 1 was extended to the use of any generator distribution with support $(-\infty, \infty)$ by Alzaatreh et al. [15], who defined the *T-X* family as follows: Let $r(t)$be the PDF of a continuous random variable *T* where$T \epsilon [a, b]$ $-\infty \leq a < b \leq \infty$ and define $W(F(x))$ to be a monotonic and absolutely continuous function of the CDF $F(x)$ of any random variable *X*.

The CDF of the *T-X* family of distributions is defined as

$$G(x) = \int_a^{W(F(x))} r(t)dt = R\{W(F(x))\}$$

where $R(t)$ is the CDF of the random variable *T*. It is easy to see that beta-generated and Kum-generated families are special cases of *T-X* family. Alzaatreh et al. [15] provided a list of $W(F(x))$ for three different supports of *T* in (0, 1), $(0, \infty)$ and $(-\infty, \infty)$.

Aljarrah et al. [16] refined the *T-X* family method by defining the $W(F(x))$ to be $Q_Y(F(x))$, the quantile function of any random variable *Y*, and defined the *T-R*$\{Y\}$ framework (see also Alzaatreh et al. [17]) as follows: Let *T*, *R* and *Y* be random variables with respective CDFs

$F_T(x) = P(T \leq x), F_R(x) = P(R \leq x)$,and$F_Y(x) = P(Y \leq x)$,. The PDFs are $f_T(x), f_R(x)$,and $f_Y(x)$, respectively. Define the quantile function as $Q_Z(p) = inf\{z: F_Z(z) \geq p\}, 0 < p < 1$ Then, the corresponding quantile functions for the random variables *T*, *R* and *Y* are $Q_T(p), Q_R(p)$and $Q_Y(p)$ The CDF and the PDF of the random variable *X* are respectively defined as

$$F_X(x) = \int_a^{Q_Y(F(x))} f_T(t)dt = F_T\left(Q_Y\left(F_R(x)\right)\right), \tag{3}$$

$$f_X(x) = f_R(x) \times \frac{f_T\left(Q_Y\left(F_R(x)\right)\right)}{f_Y\left(Q_Y\left(F_R(x)\right)\right)}, \tag{4}$$

It is interesting to note that given a random variable *R*, *T-R*$\{Y\}$ results in a generalized *R* distribution for any non-uniform *T* and *Y* random variables. Thus, one can apply the *T-R*$\{Y\}$ methodology to generate different families of generalized *R* distributions. Note that for a given random variable *T*, *T-R*$\{Y\}$ does not generate families of generalized *T* distributions using different *R* or *Y* random variables. This can be seen by the fact that the support of *R* is the same as that of *T-R*$\{Y\}$; while the support of *T* can be different from that of *T-R*$\{Y\}$.

Many new generalized families of distributions using the T-$R\{Y\}$ framework have been studied (e.g., Mansoor et al. [18], Yousof et al. [19], Aldeni et al. [20]). Some members of T-X family with $W\big(F(x)\big) = -log(1 - F(x))$ include the gamma-Pareto distribution studied by Alzaatreh et al. [21] and the Weibull-Pareto distribution studied by Alzaatreh et al. [22]. Alzaatreh et al. [17] investigated the family of generalized normal distributions and Almheidat et al. [23] investigated the family of generalized Weibull distributions. Using the quantile functions of different Y random variables, we develop several new generalizations of Pareto distribution, the T-Pareto$\{Y\}$ family. Many existing generalizations of Pareto distribution are special cases of the T-Pareto$\{Y\}$ family.

The outline of this paper is as follows: Section 2 introduces different generalizations of the Pareto distribution. Section 3 investigates some general properties of the proposed families. Section 4 defines some new members of the proposed families and some of their properties are discussed. Section 5 presents a simulation study to investigate the properties of the maximum likelihood estimators for a generalized Pareto distribution, namely, normal-Pareto {Cauchy}. Four data sets from engineering, biomedical and social science are applied in Section 6 to illustrate the flexibility and usefulness of the T-Pareto distributions. Section 7 gives a brief summary.

2. Some *T*-Pareto families of distributions

Applying different random variables T or Y, the resulting distribution of T-Pareto$\{Y\}$ family is a generalized Pareto distribution. In this section we define the following six families of generalized Pareto (GP) distributions; T-Pareto$\{Y\}$ using quantile functions of exponential, log-logistic, Weibull, logistic, Cauchy, and extreme value random variables. The corresponding quantile functions are listed in Table 1. There are other possible random variables Y with closed-form quantile functions that can be used to generate the T-Pareto$\{Y\}$ families. For practical purpose we focus on these six in Table 1 so that the resulting new families of distributions have at most five parameters.

Table 1. Some quantile functions of Y and the domains of T

Random variable Y		The quantile function $Q_Y(p)$	Domain of T
i	Exponential	$-\log(1-p)$	$(0,\infty)$
ii	Log-logistic	$[p/(1-p)]^{1/\beta}, \beta > 0$	$(0,\infty)$
iii	Weibull	$\lambda(-\log(1-p))^{1/k}$, $\lambda, k > 0$	$(0,\infty)$
iv	Logistic	$\lambda \log[p/(1-p)]$, $\lambda > 0$	$(-\infty,\infty)$
v	Cauchy	$\tan(\pi(p-0.5))$	$(-\infty,\infty)$
vi	Extreme value	$\log(-\log(1-p))$	$(-\infty,\infty)$

The CDF and PDF for each of these families can be derived by using the corresponding quantile function in Equations (3) and (4) respectively. The hazard function is obtained by using the definition $h_X(x) = f_X(x)/(1 - F_X(x))$. The generalized Pareto families of distributions are now derived:

i. T-Pareto{exponential}: The CDF, PDF and hazard function of T-Pareto{exponential} are respectively given by

$$F_X(x) = F_T\{-log\big(1 - F_R(x)\big)\} = F_T\big(H_R(x)\big), \quad (5)$$

$$f_X(x) = \frac{f_R(x)}{1-F_R(x)} f_T\{-\log\big(1 - F_R(x)\big)\} = h_R(x) f_T\big(H_R(x)\big), \text{and} \quad (6)$$

$$h_X(x) = h_R(x) h_T(H_R(x))$$

where $f_R(x)$and $F_R(x)$are PDF and CDF of Pareto random variable given in Equations (1) and (2), and $h_T(x)$, $h_R(x)$, and $H_R(x)$ are the hazard function of the *T* random variable, the hazard and cumulative hazard functions for the Pareto distribution, respectively. It is noticed that the *T*-Pareto{exponential} defined above is a function of hazard and cumulative hazard functions. Thus, this family of GP can be considered as GP arising from hazard function.

ii. *T*-Pareto{log-logistic}: The CDF, PDF and hazard function of *T*-Pareto{log-logistic} are respectively given by

$$F_X(x) = F_T\{[F_R(x)/(1-F_R(x))]^{1/\beta}\}, \tag{7}$$

$$f_X(x) = \frac{f_R(x)}{\beta {F_R}^{(\beta-1)/\beta}(x)(1-F_R(x))^{(\beta+1)/\beta}} f_T\{[F_R(x)/(1-F_R(x))]^{1/\beta}\}, \text{and}$$

$$h_X(x) = \frac{h_R(x) h_T\{[F_R(x)/(1-F_R(x))]^{1/\beta}\}}{\beta {F_R}^{(\beta-1)/\beta}(x)(1-F_R(x))^{1/\beta}}$$

The Weibull-Pareto{log-logistic} distribution defined and studied by Aljarrah et al. [24] is a member of this family.

iii. *T*-Pareto{Weibull}: The CDF, PDF and hazard function of *T*-Pareto{Weibull} are respectively given by

$$F_X(x) = F_T\{\lambda[-\log(1-F_R(x))]^{1/k}\}, \tag{8}$$

$$f_X(x) = \frac{\lambda f_R(x)[-\log(1-F_R(x))]^{(1-k)/k}}{k(1-F_R(x))} f_T\{\lambda[-\log(1-F_R(x))]^{1/k}\}, \text{and}$$

$$h_X(x) = (\lambda/k) h_R(x)(H_R(x))^{(1-k)/k} h_T\{\lambda[H_R(x)]^{1/k}\}.$$

This family of GP can be considered as GP arising from hazard function. The *T*-Pareto{exponential} is a special case of this family.

iv. *T*-Pareto{logistic}: The CDF, PDF and hazard function of *T*-Pareto{logistic} are respectively given by

$$F_X(x) = F_T\{\lambda\, log[F_R(x)/(1-F_R(x))]\}, \tag{9}$$

$$f_X(x) = \frac{\lambda f_R(x)}{F_R(x)(1-F_R(x))} f_T\{\lambda\, log[F_R(x)/(1-F_R(x))]\}, \text{and}$$

$$h_X(x) = \frac{\lambda}{F_R(x)} h_R(x) h_T\{\lambda\, log[F_R(x)/(1-F_R(x))]\}.$$

v. *T*-Pareto{Cauchy}: The CDF, PDF and hazard function of *T*-Pareto{Cauchy} are respectively given by

$$F_X(x) = F_T\{tan[\pi(F_R(x)-0.5)]\}, \tag{10}$$

$$f_X(x) = \pi f_R(x) sec^2[\pi(F_R(x)-0.5)]\, f_T\{tan[\pi(F_R(x)-0.5)]\}, \text{and}$$

$$h_X(x) = \pi f_R(x) sec^2[\pi(F_R(x)-0.5)]\, h_T\{tan[\pi(F_R(x)-0.5)]\}.$$

vi. *T*-Pareto{extreme value}: The CDF, PDF and hazard function of *T*-Pareto{extreme value} are respectively given by

$$F_X(x) = F_T\{log[-log(1-F_R(x))]\}, \tag{11}$$

$$f_X(x) = \frac{f_R(x)}{-[1-F_R(x)]\log[1-F_R(x)]} f_T\{log[-log(1-F_R(x))]\}, \text{ and}$$

$$h_X(x) = \frac{h_R(x)}{H_R(x)} h_T\{log[H_R(x)]\}.$$

3. Some properties of the *T*-Pareto family of distributions

In this section, some of the general properties of the *T*-Pareto family will be discussed.

Lemma 1 (**Transformation**): Consider any random variable *T* with PDF $f_T(x)$, then the random variable

(i) $X = \theta e^{T/\alpha}$ has the *T*-Pareto{exponential} family of distribution in Equation (5).
(ii) $X = \theta(1 + T^{\beta})^{1/\alpha}$ has the *T*-Pareto{log-logistic} family of distribution in Equation (7).
(iii) $X = \theta e^{(T/\lambda)^{k}/\alpha}$ has the *T*-Pareto{Weibull} family of distribution in Equation (8).
(iv) $X = \theta(e^{T/\lambda} + 1)^{1/\alpha}$ has the distribution of *T*-Pareto{logistic} family of distribution in Equation (9).
(v) $X = \theta(0.5 - (arctanT)/\pi)^{-1/\alpha}$ has the *T*-Pareto{Cauchy} family of distribution in Equation (10).
(vi) $X = \theta e^{e^{T}/\alpha}$ has the *T*-Pareto{extreme value} family of distribution in Equation (11).

Proof: The result follows from Equations (5), (7), (8), (9), (10) and (11).

The importance of Lemma 1 is that it shows the relationship between the random variable *X* and the random variable *T*, which allows us to generate random samples from *X* by using the random variable *T*. As an example, we can generate the random variable *X* that follows the *T*-Pareto{exponential} distribution in Equation (5) by first simulating the random variable T from the PDF $f_T(x)$ and then computing $X = \theta e^{T/\alpha}$, which has the CDF $F_X(x)$.

Lemma 2 (**Quantiles**): Let $Q_X(p), 0 < p < 1$ denote a quantile function of the random variable *X*. Then the quantile functions for the (i) *T*-Pareto{exponential}, (ii) *T*-Pareto{log-logistic}, (iii) *T*-Pareto{Weibull}, (iv) *T*-Pareto{logistic}, (v) *T*-Pareto{Cauchy} and (vi) *T*-Pareto{extreme value} distributions, are respectively,

(i) $Q_X(p) = \theta e^{(Q_T(p)/\alpha)}$,
(ii) $Q_X(p) = \theta((Q_T(p))^{\beta} + 1)^{1/\alpha}$,
(iii) $Q_X(p) = \theta e^{((Q_T(p)/\lambda)^{k}/\alpha)}$,
(iv) $Q_X(p) = \theta(e^{(Q_T(p)/\lambda)} + 1)^{1/\alpha}$,
(v) $Q_X(p) = \theta(0.5 - (arctanQ_T(p))/\pi)^{-1/\alpha}$,
(vi) $Q_X(p) = \theta e^{(e^{Q_T(p)}/\alpha)}$.

Proof: The result follows by using $F_X(Q_X(p)) = p$ and then solving for $Q_X(p)$ such that $F_X(.)$ is the CDF defined in Equations (5), (7), (8), (9), (10), and (11).

Theorem 1: Let $\bar{F}_R(x) = 1 - F_R(x)$be the survival function of the Pareto distribution. The mode(s) of the (i) *T*-Pareto{exponential}, (ii) *T*-Pareto{log-logistic}, (iii) *T*-Pareto{Weibull}, (iv) *T*-Pareto{logistic}, (v) *T*-Pareto{Cauchy} and (vi) *T*-Pareto{extreme value} distributions, respectively, are the solutions of the equations

(i) $$\frac{f'_T\{-log\bar{F}_R(x)\}}{f_T\{-log\bar{F}_R(x)\}} = \frac{1}{\alpha} \quad (12)$$

(ii) $$\frac{f'_T\{F_R(x)/\bar{F}_R(x)\}^{1/\beta}}{f_T\{F_R(x)/\bar{F}_R(x)\}^{1/\beta}} = -\left\{\frac{1}{\beta} - \frac{1}{\alpha} + \left(\frac{1}{\beta} - 1\right)\frac{\bar{F}_R(x)}{F_R(x)}\right\}\frac{\beta\bar{F}_R^{\frac{1}{\beta}}}{F_R^{\frac{1-\beta}{\beta}}}, \quad (13)$$

(iii) $$\frac{f'_T\{\lambda(-log\bar{F}_R(x))^{1/k}\}}{f_T\{\lambda(-log\bar{F}_R(x))^{1/k}\}} = \frac{-[\alpha^{-1}log\bar{F}_R(x)+k^{-1}-1]}{(\lambda/k)(-log\bar{F}_R(x))^{1/k}}, \quad (14)$$

(iv) $$\frac{f'_T\{\lambda\log(F_R(x)/\bar{F}_R(x))\}}{f_T\{\lambda\log(F_R(x)/\bar{F}_R(x))\}} = \frac{(1/\alpha-1)F_R(x)+1}{\lambda} \quad (15)$$

(v) $$\frac{f'_T\{\tan(\pi(F_R(x)-0.5))\}}{f_T\{\tan(\pi(F_R(x)-0.5))\}} = \frac{-2\pi\bar{F}_R(x)\tan\big(\pi(F_R(x)-0.5)\big)+1+(1/\alpha)}{\pi\bar{F}_R(x)sec^2(\pi(F_R(x)-0.5))} \quad (16)$$

(vi) $$\frac{f'_T\{\log(-log\bar{F}_R(x))\}}{f_T\{\log(-log\bar{F}_R(x))\}} = \frac{-log\bar{F}_R(x)}{\alpha} + 1 \quad (17)$$

Proof: We will give the proof of (i). Considering the fact that $f_R{}'(x) = \frac{f_R{}^2(x)}{\bar{F}_R(x)}\left(-1 - \frac{1}{\alpha}\right)$,the derivative of Equation (6) with respect to x , can be simplified to $f_X{}'(x) = \frac{f_R{}^2(x)}{\bar{F}_R{}^2(x)}k(x)$, where $k(x) = \frac{-1}{\alpha}f_T\big(-log\bar{F}_R(x)\big) + f'_T\big(-log\bar{F}_R(x)\big)$, By setting, $f'_X(x) = 0$ .we obtain the mode of $f_X(x)$ by solving the equation $k(x) = 0$.Simplifying $k(x) = 0$ gives the result in (12). The rest of the results in Theorem 1, Equations (13)-(17), can be derived using the same technique used for Equation (12).

The results in Theorem 1 do not guarantee a unique mode for GP distributions; there could be more than one mode. For example, the normal-Pareto{Cauchy} given in Section 4 is a bimodal distribution for different values of its parameters.

Theorem 2: The Shannon entropies for the (i) *T*-Pareto{exponential}, (ii) *T*-Pareto{log-logistic}, (iii) *T*-Pareto{Weibull}, (iv) *T*-Pareto{logistic}, (v) *T*-Pareto{Cauchy} and (vi) *T*-Pareto{extreme value} distributions, respectively, are given by

(i) $$\eta_x = log\left(\frac{\theta}{\alpha}\right) + (\mu_T/\alpha) + \eta_T \quad (18)$$

(ii) $$\eta_x = log(\beta\theta/\alpha) + ((1-\alpha)/\alpha)E(log(1+T^\beta)) - (1-\beta)E(logT) + \eta_T \quad (19)$$

(iii) $$\eta_x = log(k\theta/\alpha\lambda^k) + E(T^k)/\alpha\lambda^k - (1-k)E(logT) + \eta_T \quad (20)$$

(iv) $$\eta_x = log(\theta/\alpha\lambda) + ((1-\alpha)/\alpha)E(loge^{T/\lambda} + 1) + \mu_T/\lambda + \eta_T, \quad (21)$$

(v) $$\eta_x = log(\theta/\pi\alpha) - (\alpha+1)E(\log(0.5 - (\arctan(T)/\pi)))/\alpha - E(\log(T^2+1)) + \eta_T, \quad (22)$$

(vi) $$\eta_x = log(\theta/\alpha) + \mathrm{E}(e^T)/\alpha + \mu_T + \eta_T \quad (23)$$

Here, μ_T and η_T are the mean and the Shannon entropy for the random variable *T*.

Proof: We first prove the result in Equation (18) for the *T*-Pareto{exponential} family. By the definition of the Shannon entropy,

$$\eta_x = E(-log[f_X(X)]) = -E\big(logf_R(X)\big) + E\big(log\big(1 - F_R(X)\big)\big) + E\left(-logf_T\big(log\big(1 - F_R(X)\big)\big)\right).$$

Since the random variable $T = -log\big(1 - F_R(X)\big)$ for the *T*-Pareto{exponential} family, we have

$$E\Big(-logf_T\big(-log\big(1 - F_R(X)\big)\big)\Big) = E\big(-logf_T(T)\big) = \eta_T \tag{24}$$

Now, $log(f_R(X)) = log\alpha + \alpha log\theta - (\alpha + 1)logx$, which gives

$$E\big(-log\big(f_R(X)\big)\big) = -log\alpha - \alpha log\theta + (\alpha + 1)E(log(X)).$$

From Lemma 1(i), $\mathrm{X} = \theta e^{T/\alpha}$ follows the *T*-Pareto{exponential}, hence

$$E\big(-log\big(f_R(X)\big)\big) = \log(\theta/\alpha) + ((\alpha + 1)/\alpha)E(T) \tag{25}$$

$$\text{Also, } E\big(log\big(1 - F_R(X)\big)\big) = E(\alpha\log(\theta/X)) = -E(T) \tag{26}$$

The result in Equation (18) follows from Equations (24) - (26). By the same procedure we can get the rest of the Shannon entropy formulas in Equations (19) - (23).

Moments:

Aljarrah et al. [16] proved that if $f_R(X)$ is the PDF of a non-negative random variable *R*, then the r^{th} non-central moment of the random variable *T*-*R*{*Y*} satisfies $\mathrm{E}(X^r) \leq \mathrm{E}(R^r)\mathrm{E}(\{\bar{F}_Y(T)\}^{-1})$. Thus, in order for the r^{th} non-central moment to exist, both the r^{th} non-central moment of *R* and $\mathrm{E}(\{\bar{F}_Y(T)\}^{-1})$ has to be finite. By applying this upper bound, Theorem 3 provides the r^{th} non-central moment for the (i) *T*-Pareto{exponential}, (ii) *T*-Pareto{log-logistic}, (iii) *T*-Pareto{Weibull}, (iv) *T*-Pareto{logistic}, (v) *T*-Pareto{Cauchy} and (vi) *T*-Pareto{extreme value} distributions.

Theorem 3: The r^{th} non-central moments for the (i) *T*-Pareto{exponential}, (ii) *T*-Pareto{log-logistic}, (iii) *T*-Pareto{Weibull}, (iv) *T*-Pareto{logistic}, (v) *T*-Pareto{Cauchy} and (vi) *T*-Pareto{extreme value} distributions, respectively, are given by

(i) $\mathrm{E}(X^r) = \theta^r M_T(r/\alpha)$ exists if $M_T(r/\alpha)$ exist (27)

(ii) $\mathrm{E}(X^r) = \theta^r \int_0^\infty (u^\beta + 1)^{r/\alpha} f_T(u)du$ exists if $\mathrm{E}\big(X^{r\beta/\alpha}\big)$ exists. (28)

(iii) $\mathrm{E}(X^r) = \theta^r M_{T^k}(r/\lambda^k\alpha)$ exists if $M_{T^k}(r/\lambda^k\alpha) < \infty$ (29)

(iv) $\mathrm{E}(X^r) = \theta^r \int_{-\infty}^\infty (e^{u/\lambda} + 1)^{r/\alpha} f_T(u)du$ exists if $M_T(r/\lambda\alpha) < \infty$. (30)

(v) $\mathrm{E}(X^r) = \theta^r (0.5)^{-r/\alpha} \sum_{k=0}^\infty (-1)^k \binom{r/\alpha + k - 1}{k} (-2/\pi)^k E(arctanT)^k$ (31)
exists if $E(arctanT)^k$ exists

(vi) $\mathrm{E}(X^r) = \theta^r M_{e^T}(r/\alpha)$ exists if $M_{e^T}(r/\alpha) < \infty$ (32)

Where $M_X(z) = E(e^{zX})$, $\mathrm{z} \epsilon (-\mathrm{h}, \mathrm{h}), \mathrm{h} > 0$ is the moment generating function for a random variable *X*.

Proof: We first show Equation (27). By using Lemma 1, the r^{th} non-central moment for the *T*-Pareto{exponential} distribution can be written as

$$\mathrm{E}(X^r) = \mathrm{E}(\theta e^{T/\alpha})^r = \theta^r M_T(r/\alpha).$$

The same approach is used to find the results in Equations (29) to (32). The definition of the r^{th} non-central moment or the generalized binomial expansion can be used to get the results for *T*-Pareto{log-logistic}, *T*-Pareto{logistic}, and *T*-Pareto{Cauchy} families.

The deviation from the mean and from the median are used for measuring the dispersion and the spread from the center. The mean deviation from the mean μ and the mean deviation from the median M are denoted respectively as $D(\mu)$and $D(M)$.

Theorem 4: The $D(\mu)$ and $D(M)$ for the (i) T-Pareto{exponential}, (ii) T-Pareto{log-logistic}, (iii) T-Pareto{Weibull}, (iv) T-Pareto{logistic}, (v) T-Pareto{Cauchy} and (vi) T-Pareto{extreme value} distributions, respectively, are

(i)
$$D(\mu) = 2\mu F_X(\mu) - 2\theta \sum_{i=0}^{\infty} \frac{1}{\alpha^i i!} S_u(\mu, 0, i), \quad (33)$$

$$D(M) = \mu - 2\theta \sum_{i=0}^{\infty} \frac{1}{\alpha^i i!} S_u(M, 0, i), \quad (34)$$

where $S_\xi(c, a, b) = \int_a^{Q_Y(F_R(c))} \xi^b f_T(u)du$, and $Q_Y(F_R(c)) = -log(1 - F_R(c))$

(ii)
$$D(\mu) = 2\mu F_X(\mu) - 2\theta \sum_{i=0}^{\infty} \binom{\frac{1}{\alpha}}{i} S_{u^\beta}(\mu, 0, b),$$

$$D(M) = \mu - 2\theta \sum_{i=0}^{\infty} \binom{1/\alpha}{i} S_{u^\beta}(M, 0, b),$$

with $b = i$,when $|u^\beta| < 1$ and $b = (1/\alpha) - i$ when $|u^\beta| > 1$

where$Q_Y(F_R(c)) = F_R(c)/(1 - F_R(c))^{1/\beta}$

(iii)
$$D(\mu) = 2\mu F_X(\mu) - 2\theta \sum_{i=0}^{\infty} \frac{1}{\lambda^{ki} \alpha^i i!} S_{u^k}(M, 0, i),$$

$$D(M) = \mu - 2\theta \sum_{i=0}^{\infty} \frac{1}{\lambda^{ki} \alpha^i i!} S_{u^k}(M, 0, i),$$

where $Q_Y(F_R(c)) = \lambda(-\log(1 - F_R(c)))^{1/k}$.

(iv)
$$D(\mu) = 2\mu F_X(\mu) - 2\theta \sum_{i=0}^{\infty} \binom{\frac{1}{\alpha}}{i} S_{e^{\frac{u}{\lambda}}}(\mu, -\infty, b),$$

$$D(M) = \mu - 2\theta \sum_{i=0}^{\infty} \binom{1/\alpha}{i} S_{e^{\frac{u}{\lambda}}}(\mu, -\infty, b),$$

with $b = i$ when $\left|e^{\frac{u}{\lambda}}\right| < 1$ and $b = (1/\alpha) - i$ when $\left|e^{\frac{u}{\lambda}}\right| > 1$

where $Q_Y(F_R(c)) = \lambda log(F_R(c)/(1 - F_R(c)))$.

(v)
$$D(\mu) = 2\mu F_X(\mu) - 2\theta (0.5)^{-\frac{1}{\alpha}} \sum_{i=0}^{\infty} \binom{-\frac{1}{\alpha}}{i} \left(\frac{-2}{\pi}\right)^i S_{\arctan(u)}(\mu, -\infty, i),$$

$$D(M) = \mu - 2\theta (0.5)^{-\frac{1}{\alpha}} \sum_{i=0}^{\infty} \binom{-\frac{1}{\alpha}}{i} \left(\frac{-2}{\pi}\right)^i S_{\arctan(u)}(M, -\infty, i),$$

where $Q_Y(F_R(c)) = tan(\pi(F_R(c) - 0.5))$.

(vi)
$$D(\mu) = 2\mu F_X(\mu) - 2\theta \sum_{i=0}^{\infty} \frac{1}{\alpha^i i!} S_{e^u}(\mu, -\infty, i),$$

$$D(M) = \mu - 2\theta \sum_{i=0}^{\infty} \frac{1}{\alpha^i i!} S_{e^u}(M, -\infty, i),$$

where $Q_Y(F_R(c)) = log(-log(1 - F_R(c)))$.

Proof: By the definitions of $D(\mu)$ and $D(M)$

$$D(\mu) = E(|X-\mu|) = 2\int_{-\infty}^{\mu}(\mu - x)f_X(x)dx = 2\mu F_X(\mu) - 2\int_{-\infty}^{\mu} x f_X(x)dx \quad (35)$$

$$D(M) = E(|X-M|) = 2\int_{-\infty}^{M}(M - x)f_X(x)dx + E(X) - M = \mu - 2\int_{-\infty}^{M} x f_X(x)dx \quad (36)$$

First, we will prove the results in Equations (33) and (34) for the *T*-Pareto{exponential} distribution.

Take $I_c = \int_{-\infty}^{c} x f_X(x)dx$.Using Equation (6) we can re-write I_c as

$$I_c = \int_{-\infty}^{c} x \frac{f_R(x)}{1-F_R(x)} f_T\{-\log(1 - f_R(x))\}dx \quad (37)$$

Let $u = -log(1 - F_R(x))$then Equation (37) can be written as

$$I_c = \int_0^{-log(1-F_R(c))} \theta e^{u/\alpha} f_T(\mathrm{u})du$$

Using the series representation for the exponential function $e^z = \sum_{k=0}^{\infty}\frac{z^k}{k!}$,

I_c can be simplified to $I_c = \theta\sum_{i=0}^{\infty}\frac{1}{\alpha^i i!}\int_0^{Q_Y(F_R(c))} u^i f_T(u)du = \theta\sum_{i=0}^{\infty}\frac{1}{\alpha^i i!}S_u(c,0,i)$ (38)

The expressions of $D(\mu)$ and $D(M)$for *T*-Pareto{exponential} follow from using Equation (38) in Equations (35) and (36). Applying the same technique of showing Equations (33) and (34), one can show the results of $D(\mu)$ and $D(M)$ for (ii), (iii), (iv), (v), and (vi).

4. Some new generalized Pareto distributions

In this section, we will present four new GP distributions in the *T*-Pareto{*Y*} families. The four distributions are exponentiated exponential-Pareto{exponential}, Cauchy-Pareto{logistic}, normal-Pareto{Cauchy}, and finally, log-logistic-Pareto{Weibull}. The additional parameters from the distributions of *T* and *Y* are often added to make the flexibility of characterizing the distribution shapes and tails in practical applications.

4.1 The exponentiated exponential-Pareto{exponential} distribution

Let a random variable T follow the exponentiated-exponential distribution with parameters λ and a. The PDF of T is $f_T(x) = a\lambda(1 - e^{-\lambda x})^{a-1}e^{-\lambda x}, x > 0$.Then the PDF of the exponentiated exponential-Pareto{exponential}(*EE-P*{*E*}) distribution is given by

$$f_X(x) = \frac{\alpha\lambda a}{x}(1 - (\theta/x)^{\alpha\lambda})^{a-1}(\theta/x)^{\alpha\lambda}, x \geq \theta.$$

On setting $k = \alpha\lambda$, we get

$$f_X(x) = \frac{ka}{x}(1 - (\theta/x)^{k})^{a-1}(\theta/x)^{k}, x \geq \theta\ , k, \theta, a > 0$$

Plots of exponentiated exponential-Pareto{exponential} distribution with the location parameter $\theta = 10$ and for different values of the shape paramters k and a are given in Figure 1. The graphs indicate that the distribution is either monotonically decreasing or right skewed. The paramter k is from the Pareto distribution, while the additional parameter a plays the role of charactering the shape to be reversed-J or monotonically decreasing as well as the heaviness of the tail.

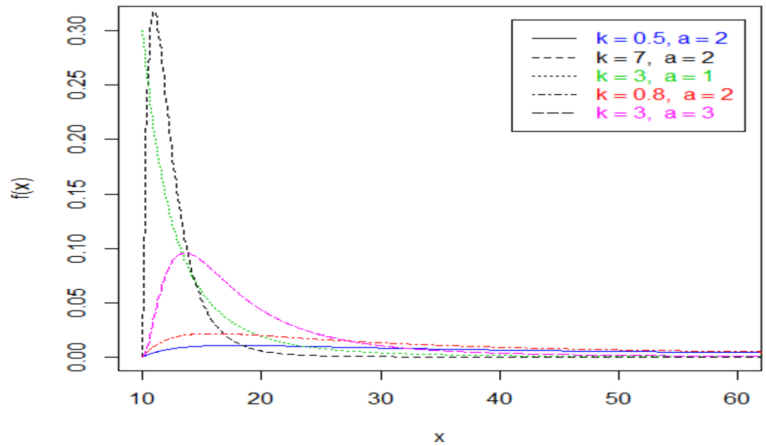

Figure 1. The PDFs of $EE\text{-}P\{E\}$ for $\theta = 10$ and various values of k and a

4.2 The Cauchy-Pareto{logistic} distribution

Let a random variable T follow the Cauchy distribution with parameters γ and μ. The PDF of T is $f_T(x) = \frac{1}{\pi}(\gamma/((x-\mu)^2+\gamma^2))$. Then the PDF of the Cauchy-Pareto{logistic}($C\text{-}P\{L\}$) distribution is given by

$$f_X(x) = \frac{(\lambda\alpha/\theta)(x/\theta)^{\alpha-1}}{\pi((x/\theta)^{\alpha}-1)}(\gamma/((\lambda log[(x/\theta)^{\alpha}-1]-\mu)^2+\gamma^2)), \qquad x \geq \theta\,, \gamma, \alpha, \theta > 0$$

Take $\gamma^* = \gamma/\lambda$ and $\mu^* = \mu/\lambda$. Then, ,

$$f_X(x) = \frac{(\alpha/\theta)(x/\theta)^{\alpha-1}}{\pi((x/\theta)^{\alpha}-1)}(\gamma^*/((\lambda log[(x/\theta)^{\alpha}-1]-\mu^*)^2+\gamma^{*2})), x \geq \theta, \qquad \mu^* \in R, \gamma^*, \alpha, \theta > 0.$$

In Figure 2, various graphs of the $C\text{-}P\{L\}$ when $\theta = 10, \mu^* = 0$ and various values of α and γ^* are provided. These plots indicate that the $C\text{-}P\{L\}$ can be monotonically decreasing (reversed J-shape) or skewed to the right and it can be either unimodal or bimodal.

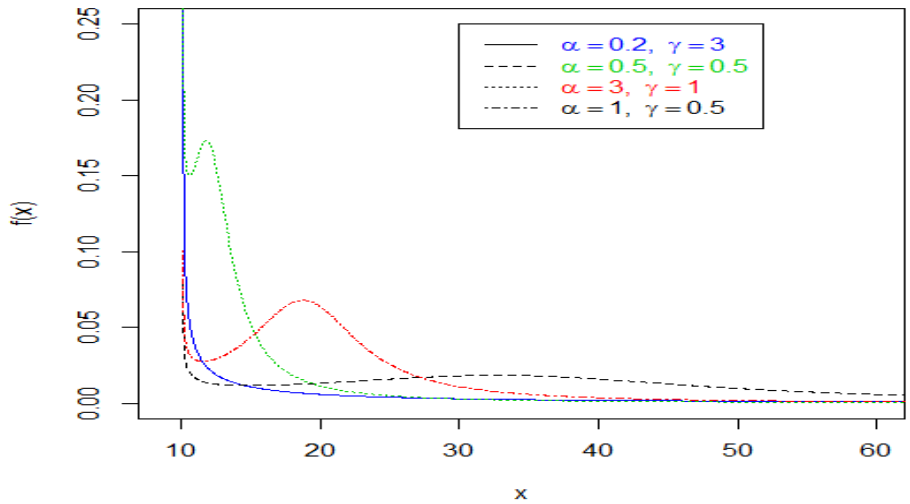

Figure 2. PDFs of *C-P*$\{L\}$ for various values of α *and* γ^*

4.3 The normal-Pareto{Cauchy} distribution

Let a random variable T follow the normal distribution with parameters μ and σ^2. The PDF of T is $f_T(x) = \frac{1}{\sqrt{2\pi\sigma}} exp(-(x-\mu)^2/2\sigma^2)$ .Then the PDF of the normal-Pareto{Cauchy}(*N-P*$\{C\}$) distribution is given by

$$f_X(x) = \frac{\sqrt{\pi}(\theta/x)^{\alpha} sec^2(\pi[0.5-(\theta/x)^{\alpha}])}{\sqrt{2}\sigma x} \exp(-(\tan(\pi[0.5-(\theta/x)^{\alpha}])-\mu)^2/2\sigma^2, x \geq \theta$$

where $\sigma^2, \alpha, \theta > 0$. In Figure 3, various graphs of $f_X(x)$ when $\theta = 10$ and for various values of α, σ and μ are provided. The figure shows that *N-P*$\{C\}$ PDF can be right skewed, left skewed, unimodal and bimodal. For fixed σ and μ the peak increases as α increases. When $\alpha>1$ and σ are both fixed, the shapes shift from right skewed, to bimodal, then, to left skewed, as μ increases.

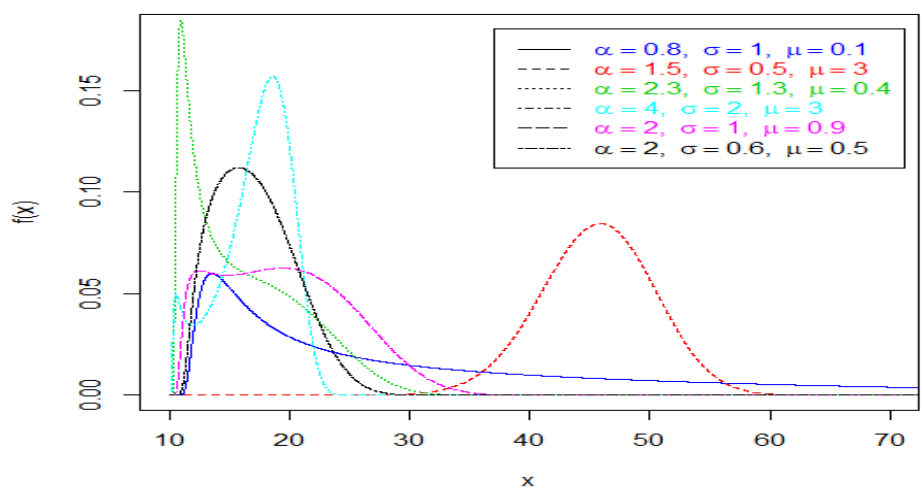

Figure 3. PDFs of *N-P*$\{C\}$ for various values of α, σ and μ

4.4 Log-logistic-Pareto{Weibull} distribution

Let a random variable T follow the log-logistic distribution with parameter β The PDF of T is $f_T(x) = \beta x^{\beta-1}/(1+x^{\beta})^2, x > 0$ .The quantile function of Weibull is given in Table 1 with parameters *(λ, k)*. The PDF of the log-logistic-Pareto{Weibull}(*LL-P*{*W*}) distribution, when setting λ=1, is given by

$$f_X(x) = \frac{\alpha\beta}{xk}\frac{(-\alpha\log(\theta/x))^{1/k-1}}{(1+(-\alpha\log(\theta/x))^{\beta/k})^2}, x \geq \theta, \beta, \alpha, \theta, k > 0$$

Note that the parameter λ is set to 1 since the resulting distribution has five parameters and λ is a scale parameter.

In Figure 4, various graphs of the *LL-P*{*W*} PDF when $\theta = 10$ and various values of α, k and β are provided. These plots indicate that the *LL-P*{*W*} can be monotonically decreasing (reversed J-shape) or skewed to the right. Moreover, the peak increases as α or β increases with the other parameter values fixed.

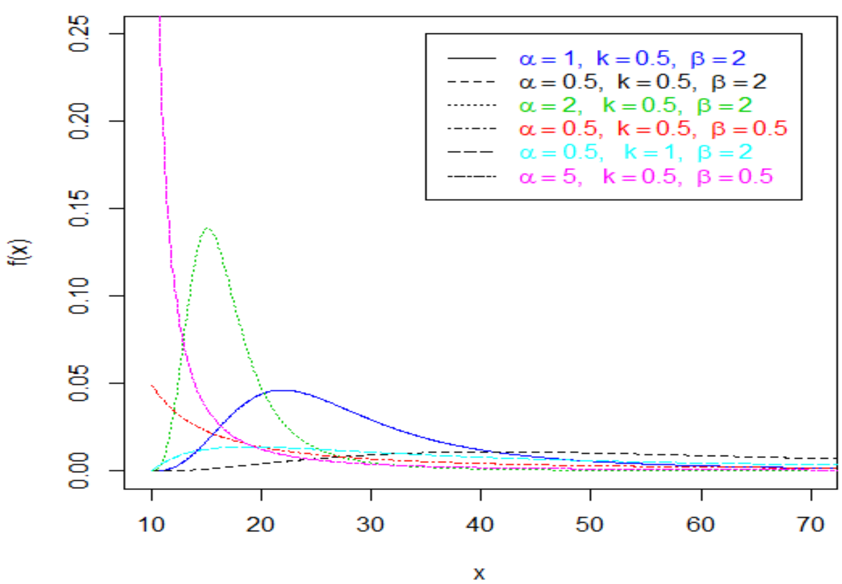

Figure 4. PDFs of *LL-P*{*W*} for various values of α, k and β

5. A simulation study of the properties of maximum likelihood estimators for *N-P*{*C*} distribution.

Suppose $X_1, X_2, \ldots, X_n$ constitute a random sample from a normal-Pareto{Cauchy} distribution as defined in Sub-section 4.3, the likelihood function, L, for the normal-Pareto{Cauchy} distribution has the following form

$$L = L(\alpha,\theta,\mu,\sigma) = \prod_{i=1}^{n}\{\frac{\sqrt{\pi}\alpha(\theta/x_i)^{\alpha}sec^2(\pi[0.5-(\theta/x_i)^{\alpha}])}{\sqrt{2}\sigma x_i}exp(-\{tan(\pi[0.5-(\theta/x_i)^{\alpha}])-\mu\}^2/2\sigma^2)\},$$

for $x_i \geq \theta$.

The corresponding log-likelihood function is

$$l(\alpha,\theta,\mu,\sigma) = nlog\frac{\sqrt{\pi}\alpha\theta^{\alpha}}{\sqrt{2}\sigma} - (\alpha+1)\sum_{i=1}^{n} logx_i + \sum_{i=1}^{n}\log[sec^2(\pi[0.5-(\theta/x_i)^{\alpha}])]$$
$$-\sum_{i=1}^{n}\{tan\pi[0.5-(\theta/x_i)^{\alpha}]-\mu\}^2/2\sigma^2 \quad (39)$$

The maximum likelihood estimates for α,θ,μ and σ are the values of α,θ,μ and σ that make the log-likelihood as large as possible. Since $x_i \geq \theta$ the maximum likelihood estimator for the parameter θ is the sample minimum given by $\hat{\theta} = \min(x_i)$. On taking partial derivatives of the log-likelihood in (39) with respect to α,μ and σ, and equating the derivatives to zero we get the likelihood equations of the normal-Pareto{Cauchy} distribution as follows:

$$0 = \frac{\partial l}{\partial\alpha} = \frac{n}{\alpha} + nlog\theta - \sum_{i=1}^{n} logx_i - \sum_{i=1}^{n} 2\pi(\theta/x_i)^{\alpha}\ln(\theta/x_i)\tan(\pi[0.5-(\theta/x_i)^{\alpha}])$$
$$+\sum_{i=1}^{n}\frac{\pi}{\sigma^2}(\theta/x_i)^{\alpha}\ln(\theta/x_i)(\tan(\pi[0.5-(\theta/x_i)^{\alpha}])-\mu)sec^2(\pi[0.5-(\theta/x_i)^{\alpha}]), \quad (40)$$

$$0 = \frac{\partial l}{\partial\alpha} = \sum_{i=1}^{n}\frac{1}{\sigma^2}\left(\tan\left(\pi\left[0.5-\left(\frac{\theta}{x_i}\right)^{\alpha}\right]\right)-\mu\right) \quad (41)$$

$$0 = \frac{\partial l}{\partial\sigma} = \frac{-n}{\sigma} + \sum_{i=1}^{n}\frac{1}{\sigma^3}(\tan(\pi[0.5-(\theta/x_i)^{\alpha}])-\mu)^2 \quad (42)$$

By using $\hat{\theta} = x_{(1)}$ and using the moment estimates of α,μ and σ as the initial estimates, Equations (40)-(42) are solved iteratively, we obtain $\hat{\alpha},\hat{\mu}$ and $\hat{\sigma}$, the maximum likelihood estimates for α,μ and σ respectively. To evaluate the performance of the MLE, a simulation study is conducted for different sample sizes (n = 100, 200, 500 and 1000) and different combinations of the parameters (α = 2, 4, μ = 0, 0.8, 2, 3, σ = 1, 2 and θ =10). For each different parameter combination, 1000 simulations are conducted. The averages and standard deviations of the MLEs are presented in Table 2.

Table 2: Average maximum likelihood estimates in N-P{C} distribution and their Monte Carlo standard deviations

N	Actual Values			Average (Standard Deviation) of MLE			
	α	μ	σ	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\mu}$	$\hat{\sigma}$
100	2	0.0	1.0[r]	10.2075 (0.2198)	2.1980 (0.3028)	0.0439 (0.1818)	1.1444 (0.2158)
	2	2.0	1.0[l]	12.0280 (1.3837)	2.3646 (0.4302)	1.8543 (0.5991)	1.1529 (0.4089)
	4	0.8	1.0[b]	10.2128 (0.2083)	4.4048 (0.6172)	0.8799 (0.2705)	1.1695 (0.2826)
	4	3.0	2.0[b]	10.3407 (0.3407)	4.2297 (0.5111)	3.1138 (0.9394)	2.1770 (0.7321)
200	2	0.0	1.0	10.1309 (0.1515)	2.1165 (0.2054)	0.0182 (0.1281)	1.0786 (0.1395)
	2	2.0	1.0	11.2779 (1.0180)	2.2108 (0.3015)	1.8756 (0.4114)	1.0662 (0.2709)
	4	0.8	1.0	10.1321 (0.1381)	4.2375 (0.4161)	0.8384 (0.1836)	1.0903 (0.1807)
	4	3.0	2.0	10.1744 (0.1788)	4.1194 (0.3389)	3.0515 (0.5940)	2.0782 (0.4585)
500	2	0.0	1.0	10.0650 (0.0958)	2.0588 (0.1126)	0.0115 (0.0766)	1.0379 (0.0826)
	2	2.0	1.0	10.6309 (0.6274)	2.1049 (0.1781)	1.9445 (0.2457)	1.0342 (0.1592)
	4	0.8	1.0	10.0667 (0.0855)	4.1241 (0.2546)	0.8230 (0.1094)	1.0460 (0.1056)
	4	3.0	2.0	10.0806 (0.0876)	4.0694 (0.2010)	3.0509 (0.3523)	2.0546 (0.2705)
1000	2	0.0	1.0	10.0382 (0.0660)	2.0326 (0.0880)	0.0052 (0.0548)	1.0233 (0.0576)
	2	2.0	1.0	10.3892 (0.3825)	2.0651 (0.1204)	1.9682 (0.1807)	1.0243 (0.1136)
	4	0.8	1.0	10.0406 (0.0577)	4.0745 (0.1747)	0.8135 (0.0799)	1.0295 (0.0732)
	4	3.0	2.0	10.0487 (0.0562)	4.0466 (0.1405)	3.0379 (0.2484)	2.0437 (0.1914)

[r]Unimodal and skewed to the right, [l]Unimodal and skewed to the left, [b]Distribution is bimodal

In Table 2, it is noticed that the bias of the MLE of θ is relatively large when the distribution is skewed to the left and the sample size is small. However, as n increases the bias reduces. As discussed in Alzaatreh [22], the reason that the overestimate of θ is mainly due to the fact that the minimum observation in a sample is larger than the population minimum, especially when sample size is small. The results from the simulation indicate that the MLE is appropriate for estimating the parameters of the N-P{C} distribution. Simulations of different generalized Pareto distributions are also conducted. The results are similar. It is anticipated that MLE method is appropriate for estimating the parameters of T-Pareto{Y} families of distributions.

6. Some applications of *T*-Pareto{*Y*} family of distributions

As demonstrated in section 4, one can derive many different GP distributions, which can capture a wide range of distribution shapes. This section presents some applications of the normal-Pareto{Cauchy} distribution using three real data sets and also presents an application of the Cauchy-Pareto{logistic} distribution using the Australian city size data. The maximum likelihood estimation method is used to estimate the parameters of the fitted distributions (with the corresponding standard errors in parentheses). The likelihood equations are given in the Appendix. The maximized log-likelihood value, the Bayesian Information Criterion (BIC) value, and the Kolmogorov-Smirnov (K-S) test statistic for the fitted distributions are reported in Tables 3, 4, 5 and 7 in order to compare the *T*-Pareto{*Y*} distributions with other distributions.

6.1 Strengths of 1.5 cm glass fibers data

This data set consists of the breaking strength of 63 glass fibers of length 1.5 cm, originally obtained by workers at the UK National Physical Laboratory [25]. The distribution of the data is skewed to the left (skewness = –0.922 and kurtosis = 1.103). Barreto-Souza et al. [26] applied the beta generalized exponential distribution (BGED) to fit the data and Alzaghal et al. [27] fitted the data using the exponentiated Weibull-exponential distribution (EWED). More recently, Almheidat et al. [23] fitted the data to the Lomax-Weibull{log-logistic} distribution (LWD). The results of *N*-*P*{*C*} in fitting this data set compared to the other distributions are presented in Table 3.

Table 3. Estimates of the model parameters for the strength of 1.5cm glass fibers data

Distribution	BGED	EWED	LWD	*N*-*P*{*C*}
Parameter Estimates	$\hat{a}$ = 0.412 (0.302) $\hat{b}$ = 93.466 (120.085) $\hat{\alpha}$ = 22.612 (21.925) $\hat{\lambda}$ = 0.923 (0.501)	$\hat{\alpha}$ = 23.614 (3.954) $\hat{\gamma}$ = 7.249 (0.994) $\hat{c}$ = 0.003 (0.003)	$\hat{\alpha}$ = 1.191 (0.723) $\hat{\theta}$=21.964 (9.417) $\hat{k}$= 2.984 (1.233) $\hat{\lambda}$ = 1.089 (0.311)	$\hat{\alpha}$ = 2.359 (0.332) $\hat{m}$= 3.605 (1.390) $\hat{\sigma}$= 1.612 (0.725) $\hat{\theta}$ = 0.55
K-S (p-value)	0.167 (0.059)	0.137 (0.195)	0.101 (0.537)	0.092 (0.675)
Log Likelihood	-15.599	-14.330	-11.991	-9.7
BIC	47.8	41.1	35.2	31.8

Table 3 shows that the N-$P\{C\}$ provides the best fit to this left skewed data set based on the different criteria presented. Figure 5 contains the histogram of the data and the PDFs of the fitted distributions.

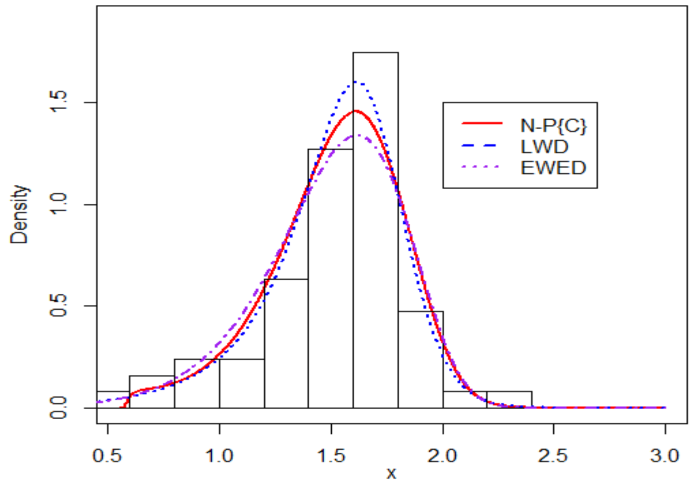

Figure 5. The fitted PDFs for the glass fibers data

6.2 The fatigue life of 6061-T6 aluminum data

The second data set was analyzed by Alzaatreh et al. [21] and Mahmoudi [8]. The data is on the fatigue life of 6061-T6 aluminum coupons cut parallel with the direction of rolling and oscillated at 18 cycles per second. The data set consists of 101 observations with maximum stress per cycle 31,000 psi.

Mahmoudi [8] fitted the data to the five-parameter beta generalized Pareto, Weibull, beta-Pareto and the three-parameter generalized Pareto distributions. Alzaatreh et al. [21] showed that the fit of a three-parameter gamma-Pareto distribution was the best among the other distributions used by Mahmoudi [8] to fit the data. The result of fitting beta-Pareto, beta-generalized Pareto, and the gamma-Pareto distributions from Mahmoudi [8] and Alzaatreh et al. [21] are reported in Table 4 along with the result of fitting the N-$P\{C\}$ distribution to the data. The results in Table 4 indicate that the beta generalized Pareto, gamma-Pareto and the N-$P\{C\}$ distributions fit the data well. Based on the K-S statistic, the two best fits are from the five-parameter beta generalized Pareto and the four-parameter N-$P\{C\}$ distributions. This suggests that the N-$P\{C\}$, with one less parameter, is a better choice than the beta generalized Pareto distribution for fitting this right skewed data with long tail.

Table 4. Estimates of the fatigue life of 6061-T6 aluminum coupons data

Distribution	*Beta-Pareto	*Beta-Generalized Pareto	*Gamma-Pareto	N-$P\{C\}$
Parameter Estimates	$\hat{\alpha}$ = 485.47 $\hat{\beta}$ = 162.06 $\hat{k}$= 0.394 $\hat{\theta}$ =3.91	$\hat{\alpha}$ = 12.112 $\hat{\beta}$ = 1.702 $\hat{m}$= 40.564 $\hat{k}$= 0.273 $\hat{\theta}$ = 54.837	$\hat{\alpha}$ = 15.021 $\hat{c}$ = 0.0429 $\hat{\theta}$ = 70	$\hat{\alpha}$ = 1.611 (0.221) $\hat{m}$= 0.482 (0.213) $\hat{\sigma}$ = 0.379 (0.0612) $\hat{\theta}$ = 70
K-S (p-value)	0.091 (0.376)	0.070 (0.700)	0.106 (0.212)	0.077 (0.599)
Log likelihood	-458.65	-457.85	-448.53	-446.55
BIC	910.6	907.0	905.9	906.8

*The standard errors were not provided by Mahmoudi [8] and Alzaatreh et al. [21]

Figure 6 displays the N-$P\{C\}$, gamma-Pareto and beta generalized Pareto fitted density functions along with the histogram for the fatigue life of 6061-T6 aluminum data. The plots in Figure 6 indicates that the N-$P\{C\}$ provides a good fit to the data which is approximately symmetric, with a long right tail.

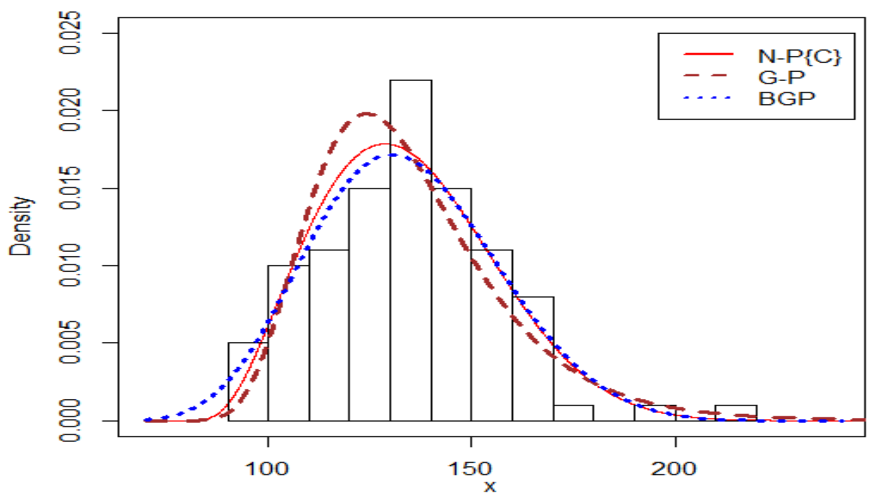

Figure 6. The fitted PDFs for the fatigue life of 6061-T6 aluminum data

6.3 The Airborne data

The airborne data represents the repair times in hours for an airborne communication transceiver. The data consist of 46 observations. The distribution of the data is highly skewed to the right (skewness = 2.99). Cordeiro et al. [28] fitted the data by using the beta generalized Raleigh (BGR), exponentiated generalized Rayleigh (EGR), and generalized Rayleigh distributions. Alzaghal et al. [27] fitted the exponentiated Weibull-exponential distribution (EWED) to the data and the distribution provided a better fit to the data than the other distributions. The results from Alzaghal et al. [27] and Cordeiro et al. [28] are provided in Table 5 in addition to the N-$P\{C\}$ distribution.

The results from Table 5 show that the N-$P\{C\}$ distribution provides the best fit to this highly right skewed data based on the different criteria presented. The plots in Figure 7 represent the fitted density functions of the N-$P\{C\}$, the EWED and the beta generalized Rayleigh distributions with the histogram of the Airborne data.

Table 5. Estimates of the model parameters for the airborne data

Distribution	Generalized Rayleigh	Exponentiated Generalized Rayleigh	Beta Generalized Rayleigh	EWED	N-$P\{C\}$
Parameter Estimates	$\hat{\alpha}$ = -0.703 (0.049) $\hat{\theta}$ = 0.008 (0.003)	$\hat{\alpha}$ = 5.712 (2.400) $\hat{\alpha}$ = -0.946 (0.024) $\hat{\theta}$ = 0.007 (0.002)	$\hat{a}$ = 10.482 (0.476) $\hat{b}$ = 20.761 (0.228) $\hat{\alpha}$ = -0.893 (0.022) $\hat{\theta}$=0.0000047 (0.00001)	$\hat{\alpha}$ = 0.498 (0.133) $\hat{\gamma}$ = 1.279 (0.702) $\hat{c}$ = 4.512 (1.992)	$\hat{\alpha}$ = 0.413 (0.051) $\hat{m}$= 0.343 (0.237) $\hat{\sigma}$ = 0.761 (0.114) $\hat{\theta}$ = 0.2
K-S (p-value)	0.176 (0.116)	0.179 (0.105)	0.122 (0.500)	0.091 (0.838)	0.072 (0.973)
Log Likelihood	-217.10	-108.15	-99.55	-100.00	-96.5
BIC	224.7	227.7	214.4	211.5	204.5

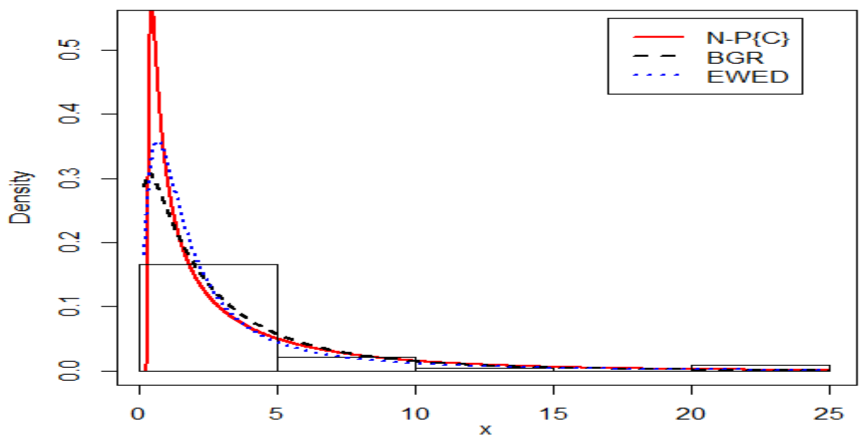

Figure 7. The fitted PDFs for the airborne data

6.4 The Australian city size data set

The Australian city size data set consists of estimates of the resident population size of significant urban areas of Australia for the following years: 1996, 2001, 2006, and 2011. The summary statistics of these population size data are given in Table 6. It is seen that the data are highly skewed to the right with skewness > 5.0.

Table 6. Summary statistics of the Australian city size data (n = 101)

	Mean	**Standard deviation**	**Skewness**	**Kurtosis**
1996	150874	513118	5.43	31.19
2001	160552	544943	5.41	30.96
2006	171447	578239	5.33	29.97
2011	188584	636610	5.28	29.33

The dataset was downloaded from the website http://www.citypopulation.de/. Gómez-Déniz and Calderín-Ojeda [10] analyzed this data using the ArcTan Pareto (PAT) distribution and compared the results to the classical Pareto, the lognormal and Pareto Positive Stable (PPS) distributions for the years 1991, 1996, 2001, 2006 and 2011. They found that only the PAT distribution adequately fits the data for each year; while PPS and PAT performs equally well for the years 1996, 2001, 2006 and 2011. Since the data for 1991 is no longer available on the website for comparison purpose, we fit the years 1996, 2001, 2006 and 2011 and only compare with the PAT distribution. The C-$P\{L\}$ distribution is applied to fit this population size data. Results from Table 7 show that the C-$P\{L\}$ fits better than the PAT for the 1996 data using the BIC and the K-S criteria. For the year 2001, The C-$P\{L\}$ fits better than PAT using the K-S and its p-value. For the years 2006 and 2011, the C-$P\{L\}$ adequately fits the data but not better than the PAT distribution. This application shows that the C-$P\{L\}$ is a good competitor to the PAT distribution for fitting the population size data.

Table 7. Comparison between the estimates of the PAT and C-$P\{L\}$ parameters for the Australian city size data

	PAT		C-$P\{L\}$	
Year $\hat{\theta} = \min(x_i)$	Estimates (standard error)	Goodness of fit BIC K-S(p-value)	Estimates (standard error)	Goodness of fit BIC K-S(p-value)
1996 $\hat{\theta}$ = 4249	$\hat{\alpha}$ = 6.301 (1.832) $\hat{g}$= 1.116 (0.129)	2427.10 0.113 (.145)	$\hat{\alpha}$ = 0.420 (0.022) $\hat{g}$= 0.438 (0.0560)	2402.6 0.105 (.222)
2001 $\hat{\theta}$ = 10035	$\hat{\alpha}$ = 0.289 (0.936) $\hat{g}$= 0.798 (0.120)	2379.53 0.071(0.657)	$\hat{\alpha}$ = 0.792 (0.081) $\hat{g}$= 0.897 (0.125)	2381 0.060 (0.861)
2006 $\hat{\theta}$ = 10799	$\hat{\alpha}$ = 0.295 (0.887) $\hat{g}$= 0.805 (0.120)	2390.67 0.061(0.823)	$\hat{\alpha}$ = 0.848 (0.092) $\hat{g}$= 0.925 (0.125)	2396.2 0.079 (0.561)
2011 $\hat{\theta}$ = 11318	$\hat{\alpha}$ = 0.602 (0.539) $\hat{g}$= 0.826 (0.120)	2412.61 0.057(0.877)	$\hat{\alpha}$ = 0.822 (0.085) $\hat{g}$= 0.885 (0.118)	2415.1 0.67 (0.710)

7. Summary

In this article, a generalization of the two-parameter Pareto distribution to the T-Pareto$\{Y\}$ family is defined and studied using the T-$R\{Y\}$ framework presented by Aljarrah et al. [16]. Six new generalized Pareto families using the quantile functions of exponential, log-logistic, logistic, Cauchy, extreme value, and Weibull are presented. Various general properties of the new families including, moment, Shannon entropy, mean deviations from the mean and median are derived.

Four new distributions, exponentiated exponential-Pareto{exponential}, Cauchy-Pareto{logistic}, normal-Pareto{Cauchy}, and log-logistic-Pareto{Weibull} distributions are defined. Four real data sets from engineering, biomedical and social science are fitted using the normal-Pareto{Cauchy} distribution and the Cauchy-Pareto{logistic} to demonstrate the flexibility and potential applications of the proposed generalized Pareto family of distributions. The comparison with other existing generalized distributions indicates the T-Pareto family of distributions perform well for fitting real data sets that are skewed to left or skewed to right with heavy tail from different disciplines.

Acknowledgements

The second author gratefully acknowledges the financial support received from the U.S. Department of State, Bureau of Education and Cultural Affairs Fulbright Grant # PS00230565. The authors are grateful to the Editor and the referee for their comments/questions. These have improved the paper.

References

[1] Pareto, V. (1896). Cours d'Économie Politique, Reprinted as a volume of Oeuvres Complétes. Droz, Geneva, 1965.

[2] Burroughs, S.M. & Tebbens, S.F. (2001). Upper-truncated power law distributions. *Fractals*, 9(02), 209-222.

[3] Arnold, B.C. (2015). *Pareto Distribution,* 2nd edition. John Wiley & Sons, Ltd.

[4] Pickands III, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 3, 119-131.

[5] Johnson, L., Kotz, S. & Balakrishnan, N. (1994). *Continuous Univariate Distributions*. Vol. 1, 2nd edition. John Wiley and Sons, New York.

[6] Gupta, R.C., Gupta, P.L. & Gupta, R.D. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics-Theory and Methods*, 27(4), 887-904.

[7] Akinsete, A., Famoye, F. & Lee, C. (2008). The beta-Pareto distribution. *Statistics*, 42(6), 547-563.

[8] Mahmoudi, E. (2011). The beta generalized Pareto distribution with application to lifetime data. *Mathematics and Computers in Simulation*, 81, 2414-2430.

[9] Sarabia, J.M. and Prieto, F. (2009). The Pareto-positive stable distribution: A new descriptive model for city size data. *Physica A: Statistical Mechanics and its Applications.* 388, 4179-4191.

[10] Gómez-Déniz, E., & Calderín-Ojeda, E. (2015a). On the use of the Pareto ArcTan distribution for describing city size in Australia and New Zealand. *Physica A: Statistical Mechanics and its Applications*, 436, 821-832.

[11] Gómez-Déniz, E., & Calderín-Ojeda, E. (2015b). Modeling insurance data with the Pareto ArcTan distribution. *ASTIN bulletin*, 45(3), 639-660.

[12] Eugene, N., Lee, C. & Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Method*, 31, 497-512.

[13] Jones, M.C. (2009). Kumaraswamy's distribution: A beta-type distribution with tractability advantages, *Statistical Methodology*, 6, 70-81.

[14] Bourguignon, M., Silva, R.B., Zea, L.M. & Cordeiro, G.M. (2013). The Kumaraswamy Pareto distribution. *Journal of Statistical Theory and Applications*, 12(2), 129-144.

[15] Alzaatreh, A., Lee, C. & Famoye, F. (2013b). A new method for generating families of continuous distributions. *Metron*, 71(1), 63-79.

[16] Aljarrah, M.A., Lee, C. & Famoye, F. (2014). On generating *T-X* family of distributions using quantile functions. *Journal of Statistical Distributions and Applications*, 1(1), 1-17.

[17] Alzaatreh, A., Lee, C. & Famoye, F. (2014). *T*-normal family of distributions: A new approach to generalize the normal distribution. *Journal of Statistical Distributions and Applications*, 1(1), 16.

[18] Mansoor, M., Tahir, M.H., Alzaatreh, A., Cordeiro, G. M., M. Zubair, M. and Ghazali, S. S. A. (2016). An Extended Frechet Distribution: Properties and Applications. *Journal of Data Science*, 14, 167-188.

[19] Yousof, H. M., Alizadeh, M., Jahanshahi, S. M. A., Thiago G. Ramiresd, T. G., Ghosh, I. and G.G. Hamedani, G. G. (2017). The Transmuted Topp-Leone G Family of Distributions: Theory, Characterizations and Applications. *Journal of Data Science*, 15, 723-740.

[20] Aldeni, M., Lee, C. and Famoye, F. (2017). Families of distributions arising from the quantile of generalized lambda distribution. *Journal of Statistical Distributions and Applications*, 4(25).

[21] Alzaatreh, A, Famoye, F. & Lee, C. (2012). Gamma-Pareto distribution and its applications. *Journal of Modern Applied Statistical Methods*, 11(1), 78-94.

[22] Alzaatreh, A., Famoye, F. & Lee, C. (2013a). Weibull-Pareto distribution and its applications. *Communications in Statistics-Theory and Method*, 42(9), 1673-1691.

[23] Almheidat, M., Famoye, F. & Lee, C. (2015). Some Generalized Families of Weibull Distribution: Properties and Applications. *International Journal of Statistics and Probability*, 4(3), 18.

[24] Aljarrah, M.A., Famoye, F. & Lee, C. (2015). A New Weibull-Pareto Distribution. *Communications in Statistics-Theory and Methods*, 44(19), 4077-4095.

[25] Smith, R.L. and Naylor, J.C. (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Applied Statistics*, 36, 358-369.

[26] Barreto-Souza, W., Santos, A.H.S. and Cordeiro, G.M. (2010). The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation*, 80(2), 159-172.

[27] Alzaghal, A., Famoye, F. & Lee, C. (2013). Exponentiated *T-X* family of distributions with some applications. *International Journal of Statistics and Probability*, 2(3), 31-49.

[28] Cordeiro, G.M., Cristino, T., Hashimoto, E.M. & Ortega, E.M. (2013). The beta generalized Rayleigh distribution with applications to lifetime data. *Statistical Papers*, 54, 133–161.