Journal of Modern Applied Statistical Methods
May 2012, Vol. 11, No. 1,

1538 – 9472/12/$95.00

Gamma-Pareto Distribution and Its Applications

Ayman Alzaatreh
Austin Peay State University,
Clarksville, TN

Felix Famoye   Carl Lee
Central Michigan University,
Mount Pleasant, MI

A new distribution, the gamma-Pareto, is defined and studied and various properties of the distribution are obtained. Results for moments, limiting behavior and entropies are provided. The method of maximum likelihood is proposed for estimating the parameters and the distribution is applied to fit three real data sets.

Key words: Estimation, moments, T-X family, unimodality.

Introduction

The Pareto distribution was named after Swiss economist Vilfredo Pareto (1848-1923) who discovered it while studying distributions for modeling income in Switzerland. Since that time, the Pareto distribution has been widely used in modeling heavy-tailed distributions, such as income distribution. Many applications of the Pareto distribution in economics, biology and physics can be found throughout the literature. Burroughs and Tebbens (2001) discussed applications of the Pareto distribution in modeling earthquakes, forest fire areas and oil and gas field sizes, and Schroeder, et al. (2010) presented an application of the Pareto distribution in modeling disk drive sector errors. To add flexibility to the Pareto distribution, various generalizations of the distribution have been derived, including: the generalized Pareto distribution (Pickands, 1975), the beta-Pareto distribution (Akinsete, et al., 2008), and the beta generalized Pareto distribution (Mahmoudi, 2011).

Let $F(x)$ be the cumulative distribution function (CDF) of any random variable X and $r(t)$ be the probability density function (PDF) of a random variable, T, defined on $[0,\infty)$. The CDF of the T-X family of distributions defined by Alzaatreh, et al. (2012) is given by

$$G(x)=\int_0^{-\log(1-F(x))} r(t)dt. \quad (1.1)$$

Alzaatreh, et al. (2012) named this family of distributions the Transformed-Transformer family (or T-X family). When X is a continuous random variable, the probability density function of the T-X family is

$$\begin{aligned} g(x) &= \frac{f(x)}{1-F(x)} r\left(-\log\left(1-F(x)\right)\right) \\ &= h(x)\, r\left(H(x)\right). \end{aligned} \quad (1.2)$$

If a random variable T follows the gamma distribution with parameters α and β, then

$$r(t)=\left(\beta^{\alpha}\,\Gamma(\alpha)\right)^{-1} t^{\alpha-1} e^{-t/\beta}, \quad t \ge 0,$$

and the definition in (1.2) leads to the gamma-X family with PDF

Ayman Alzaatreh is an Assistant Professor in the Department of Mathematics & Statistics. Email him at: alzaatreha@apsu.edu. Felix Famoye is a Professor in the Department of Mathematics. Email him at: felix.famoye@cmich.edu. Carl Lee is a Professor in the Department of Mathematics. Email him at: carl.lee@cmich.edu.

$$g(x) = \frac{f(x)\left(-\log\left(1-F(x)\right)\right)^{\alpha-1}\left(1-F(x)\right)^{\frac{1}{\beta}-1}}{\Gamma(\alpha)\beta^{\alpha}}. \tag{1.3}$$

When $\beta=1$, the gamma-*X* family in (1.3) reduces to

$$g(x) = \frac{1}{\Gamma(\alpha)} f(x)\left(-\log\left(1-F(x)\right)\right)^{\alpha-1}. \tag{1.4}$$

Alzaatreh, et al. (2012) noted that, when $\alpha = n$ where *n* is a positive integer, the distribution in (1.4) can be written as

$$g(x) = \frac{1}{(n-1)!} f(x)\left(-\log\left(1-F(x)\right)\right)^{n-1},$$

which is the density function of the upper record values, $X_{U(n)}$, arising from a sequence $\{X_i\}_{i=1}^{n}$ of identically independent random variables with PDF $f(x)$ and CDF $F(x)$ (Johnson, et al., 1994).

The Gamma-Pareto Distribution

If *X* is a Pareto random variable with density function $f(x) = k\theta^k / x^{k+1}, \quad x > \theta$, then (1.3) results in

$$g(x) = \frac{k^{\alpha}}{x\Gamma(\alpha)\beta^{\alpha}}\left(\frac{\theta}{x}\right)^{k/\beta}\left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1}, \quad x > \theta. \tag{2.1}$$

Setting $\beta / k = c$, (2.1) reduces to

$$g(x) = \frac{1}{x\Gamma(\alpha)c^{\alpha}}\left(\frac{\theta}{x}\right)^{1/c}\left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1}, \quad \alpha, c, \theta > 0;\ x > \theta. \tag{2.2}$$

A random variable *X* with the PDF $g(x)$ in (2.2) is said to follow the gamma-Pareto distribution. When $\alpha = 1$ the gamma-Pareto distribution reduces to the Pareto distribution with parameters $1/c$ and θ, and when $\alpha = n+1$, it reduces to the upper record value distribution arising from Pareto identically independent random variables (Ahsanullah & Houchens, 1989). From (2.2), the CDF of the gamma-Pareto distribution is obtained as

$$G(x) = \gamma\left\{\alpha, c^{-1}\log\left(x/\theta\right)\right\} / \Gamma(\alpha), \tag{2.3}$$

where $\gamma(\alpha,t) = \int_0^t u^{\alpha-1}e^{-u}du$ is the incomplete gamma function.

Johnson, et al. (1994) discussed different types of Pareto distributions and their CDFs. These are

Pareto II:	$F(x) = 1 - \dfrac{\theta^{a}}{(x+\theta)^{a}}, \ x > 0.$
Pareto III:	$F(x) = 1 - \dfrac{\theta e^{-bx}}{(x+\theta)^{a}}, \ x > 0.$
Pareto IV:	$F(x) = 1 - \left[1 + \left(\dfrac{x-\mu}{\sigma}\right)^{1/\gamma}\right]^{-a}, \ x > \mu.$

Using equation (1.3), the corresponding PDF of gamma-Pareto II, gamma-Pareto III, and gamma-Pareto IV can be written as shown in Table 1.

Some relationships among these distributions are:

- If *Y* follows the gamma-Pareto distribution in (2.2), then the translation $X = Y - \theta$ follows the gamma-Pareto II distribution.

- When *a* = 1 and *b* = 0, the gamma-Pareto III distribution reduces to the gamma-Pareto II distribution.

Table 1: Corresponding PDFs of gamma-Pareto Distributions

gamma-Pareto II:	$g(x)=\dfrac{\theta^{1/c}}{c^{\alpha}\Gamma(\alpha)}(x+\theta)^{-1-1/c}[\log(1+x/\theta)]^{\alpha-1},\ x>0,\ \text{where } c=\dfrac{\beta}{a}.$	(2.4)
gamma-Pareto III:	$g(x)=\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}\left(b+\dfrac{a}{x+\theta}\right)\left[\dfrac{\theta e^{-bx}}{(x+\theta)^{a}}\right]^{1/\beta}\left[-\log\left(\dfrac{\theta e^{-bx}}{(x+\theta)^{a}}\right)\right]^{\alpha-1},\ x>0.$	(2.5)
gamma-Pareto IV:	$g(x)=\dfrac{1}{\gamma\sigma c^{\alpha}\Gamma(\alpha)}\left(\dfrac{x-\mu}{\sigma}\right)^{-1+1/\gamma}\left(1+\left(\dfrac{x-\mu}{\sigma}\right)^{1/\gamma}\right)^{-1-1/c}\left[\log\left(1+\left(\dfrac{x-\mu}{\sigma}\right)^{1/\gamma}\right)\right]^{\alpha-1},$ $x>\mu,\ \text{where } c=\beta/a.$	(2.6)

- When $\gamma=1$ and $\mu=\sigma=\theta$, the gamma-Pareto IV distribution reduces to the gamma-Pareto distribution in (2.2) with parameters α, c and θ.

- When $\gamma=1$ and $\mu=0$, the gamma-Pareto IV distribution reduces to the gamma-Pareto II distribution.

Properties of the gamma-Pareto distribution

The following Lemma shows the relationship between the gamma-Pareto distribution and the gamma distribution.

Lemma 1

If a random variable Y follows the gamma distribution with parameters α and c, then the random variable $X=\theta e^{Y}$ follows the gamma-Pareto distribution with parameters α, c and θ.

Lemma 1 Proof

The result follows by using the transformation technique.

The hazard function associated with the gamma-Pareto distribution is

$$h(x)=\frac{g(x)}{1-G(x)}=\frac{\theta^{1/c}(\log(x/\theta))^{\alpha-1}}{x^{1+1/c}c^{\alpha}(\Gamma(\alpha)-\gamma\{\alpha,c^{-1}\log(x/\theta)\})},\quad x>\theta, \tag{3.1}$$

and the limiting behaviors of the gamma-Pareto PDF and the hazard function are given in the following theorem.

Theorem 1

The limit of the gamma-Pareto density function and the gamma-Pareto hazard function as $x \to \infty$ is 0 and the limit as $x \to \theta^+$ is given by

$$\lim_{x\to\theta^+} g(x) = \lim_{x\to\theta^+} h(x) = \begin{cases} 0, & \alpha > 1 \\ 1/(c\theta), & \alpha = 1 \\ \infty, & \alpha < 1. \end{cases} \tag{3.2}$$

Theorem 1 Proof

First it can be shown that $\lim_{x\to\infty} g(x)$ = 0. If $\alpha \le 1$, then from definition (2.2) $\lim_{x\to\infty} g(x) = 0$, and if $\alpha > 1$, then

$$\begin{aligned}\lim_{x\to\infty} g(x) &= \lim_{x\to\infty} \frac{(\theta/x)^{1/c}}{\Gamma(\alpha)c^{\alpha}} \times \lim_{x\to\infty} \left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1} \frac{1}{x} \\ &= 0 \times \lim_{x\to\infty} \left(\log\left(\frac{x}{\theta}\right) / x^{1/(\alpha-1)}\right)^{\alpha-1}.\end{aligned}$$

Using L'Hôpital's rule, it can be shown that

$$\lim_{x\to\infty} \left(\log(x/\theta)/x^{1/(\alpha-1)}\right)^{\alpha-1} = 0.$$

To show that $\lim_{x\to\infty} h(x) = 0$, we have $\lim_{x\to\infty} h(x) = \lim_{x\to\infty} g(x)/(1-G(x))$. Because $\lim_{x\to\infty} g(x) = 0$, L'Hôpital's rule can be applied and implies that

$$\begin{aligned}\lim_{x\to\infty} h(x) &= \lim_{x\to\infty} g'(x)/g(x) \\ &= \lim_{x\to\infty} \left\{ \frac{\alpha-1}{x\log(x/\theta)} - \frac{c+1}{cx} \right\} \\ &= 0.\end{aligned}$$

The result in (3.2) follows directly from the definition of (2.2) and $h(x) = g(x)/(1-G(x))$. The following theorem shows that the gamma-Pareto distribution is unimodal.

Theorem 2

The gamma-Pareto distribution has a unique mode at $x = x_0$. When $\alpha \le 1$ the mode is $x_0 = \theta$ and when $\alpha > 1$ the mode is $x_0 = \theta \mathrm{e}^{c(\alpha-1)/(c+1)}$.

Theorem 2 Proof

The derivative with respect to x of equation (2.2) is given by

$$g'(x) = \theta^{1/c} x^{2-1/c}\Gamma(\alpha) \times \left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-2} \left\{ c(\alpha-1) - (c+1)\log\left(\frac{x}{\theta}\right) \right\}. \tag{3.3}$$

From (3.3) the critical points of $g(x)$ are $x = \theta$ and $x = \theta \mathrm{e}^{c(\alpha-1)/(c+1)}$. For $\alpha \le 1$, it may be observed from (3.3) that $g'(x) < 0$, therefore $g(x)$ is strictly decreasing. Also, from Theorem 1, $\lim_{x\to\theta^+} g(x) = 1/(c\theta)$ when $\alpha = 1$ and $\lim_{x\to\theta^+} g(x) = \infty$ when $\alpha < 1$. Thus, $g(x)$ has a unique mode at $x = \theta$. Using Theorem 1, for $\alpha > 1$, $\lim_{x\to\theta^+} g(x) = 0$ implies that $x = \theta$ cannot be a modal point, hence the mode is $x_0 = \theta \mathrm{e}^{c(\alpha-1)/(c+1)}$.

Graphs of $g(x)$ and $h(x)$ are displayed in Figures 1-3. The plots show that the gamma-Pareto PDF has a very long right tail and also that when parameters c and θ increase, the peak of the distribution decreases. In addition, the graphs of the gamma-Pareto PDF indicate that $g(x)$ is a right skewed distribution. The plots in Figure 3 illustrate that the gamma-Pareto hazard function is either monotone decreasing or upside-down bathtub.

The entropy of a random variable X is a measure of variation of uncertainty (Rényi, 1961). Shannon entropy (Shannon, 1948) for a random variable X with PDF $g(x)$ is defined as $E\left\{-\log\left(g(X)\right)\right\}$. Shannon showed important applications of this entropy in communication theory and many applications have been used in

Figure 1: The gamma-Pareto PDF for Various Values of α, c and θ

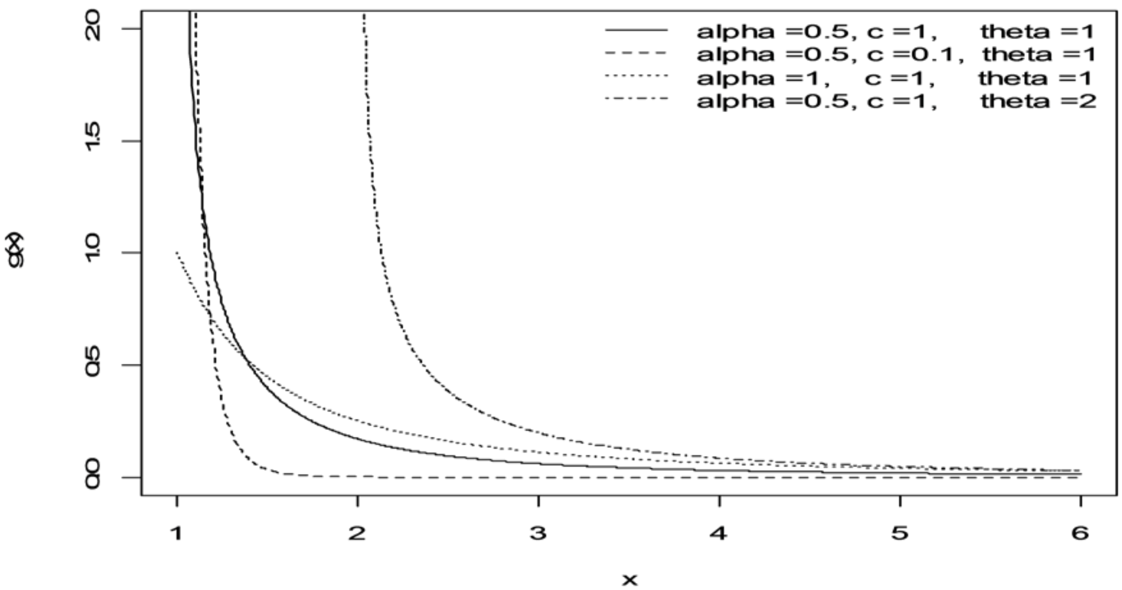


Figure 2: The gamma-Pareto PDF for Various Values of α, c and θ

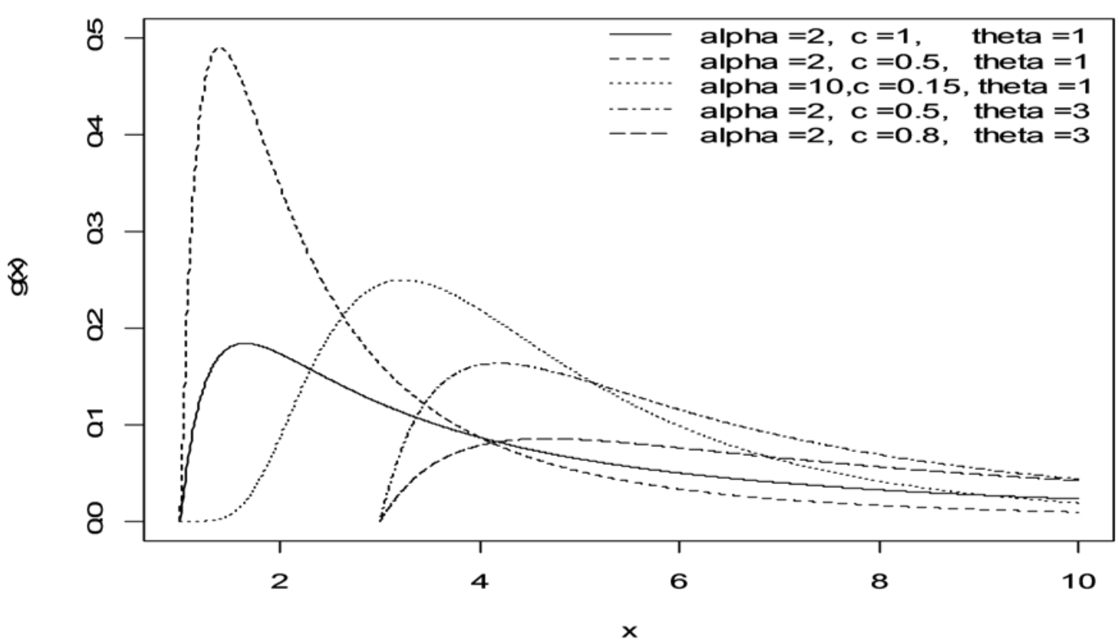
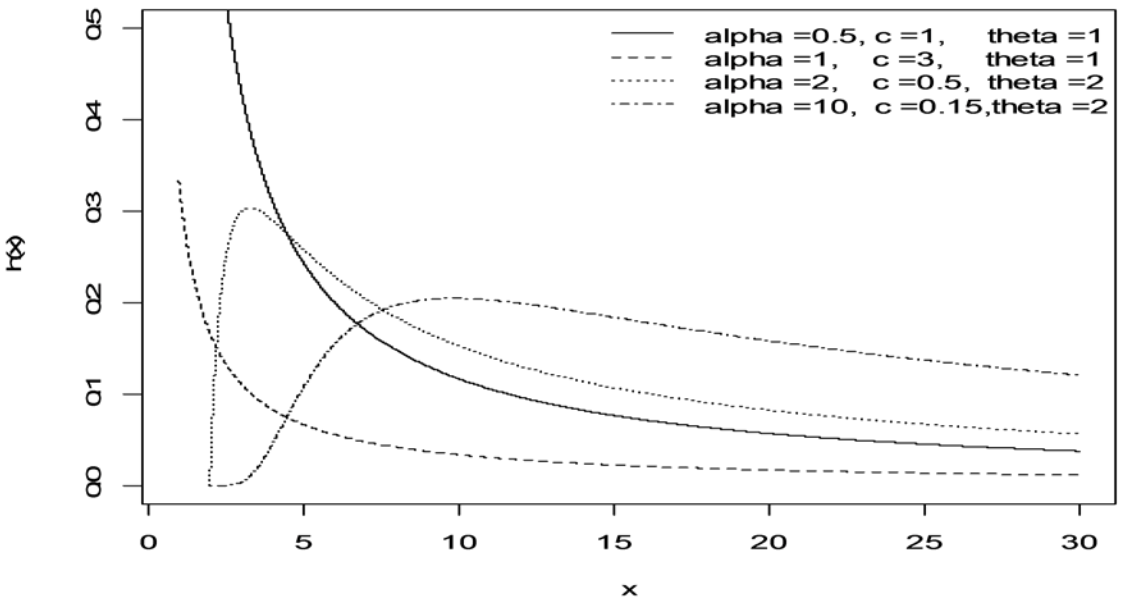


Figure 3: The gamma-Pareto Hazard Function for Various Values of α, c and θ



different areas such as engineering, physics, biology and economics.

Lemma 2

The Shannon entropy of the gamma-X family of distributions is given by

$$\eta_X = -E\left\{\log f\left(F^{-1}\left(1-e^{-T}\right)\right)\right\} + \alpha(1-\beta) + \log\beta + \log\Gamma(\alpha) + (1-\alpha)\psi(\alpha),$$

where ψ is the digamma function and T is the gamma random variable with parameters α and β.

Lemma 2 Proof

See Alzaatreh, et al. (2012) for proof details.

Theorem 3

The Shannon entropy for the gamma-Pareto distribution is given by

$$\eta_X = \log c + \log\theta + \log\Gamma(\alpha) + \alpha(c+1) + (1-\alpha)\psi(\alpha). \tag{3.4}$$

Theorem 3 Proof

First it is necessary to find $-E\left\{\log f\left(F^{-1}\left(1-e^{-T}\right)\right)\right\}$, where $f(x) = k\theta^k / x^{k+1}$. It can be shown that $F^{-1}(x) = \theta(1-x)^{-1/k}$, thus

$$-E\left\{\log f\left(F^{-1}\left(1-e^{-T}\right)\right)\right\} = \log\theta - \log k + (1+1/k)E(T)$$

The result follows from Lemma 2 by noting that $E(T) = \alpha\beta$ and $c = \beta / k$ (see equation 2.1).

The Rényi (1961) entropy for the random variable X with PDF $g(x)$ is defined as

$$I_R(s) = \frac{1}{1-s}\log\left\{\int g^s(x)dx\right\},\ \ s>0,\ s\neq 1. \tag{3.5}$$

By using the gamma-Pareto PDF in (2.2), we have

$$\int_\theta^\infty g^s(x)dx = c^{-s\alpha}\Gamma^{-s}(\alpha) \times \int_\theta^\infty x^{-s}\left(\frac{\theta}{x}\right)^{s/c}\left[\log\left(\frac{x}{\theta}\right)\right]^{s(\alpha-1)} dx. \tag{3.6}$$

Substituting $u = \log(x/\theta)$, (3.6) can be re-written as

$$\int_\theta^\infty g^s(x)dx = \frac{\theta^{1-s}\Gamma(s(\alpha-1)+1)}{c^{s\alpha}\Gamma^s(\alpha)[s(1+1/c)-1]^{s(\alpha-1)+1}}. \tag{3.7}$$

Using equation (3.7), the Rényi entropy in (3.5) can be written as

$$I_R(s) = \log\theta - \tfrac{1}{1-s}\left\{s\alpha\log c + s\log\Gamma(\alpha) - \log\Gamma(\xi) + \xi\log(s(1+c^{-1})-1)\right\}, \tag{3.8}$$

where $\xi = s(\alpha-1)+1$. Shannon entropy is a special case of Rényi entropy obtained by taking the limit of Rényi entropy as $s\to 1$. The result in (3.4) follows by using the L'Hôpital's rule for evaluating the limit of equation (3.8) as $s\to 1$.

Moments and Mean Deviations

The non-central moments for the gamma-Pareto distribution in (2.2) can be written as

$$E(X^r) = c^{-\alpha}\Gamma^{-1}(\alpha) \times \int_\theta^\infty x^{r-1}\left(\frac{\theta}{x}\right)^{1/c}\left[\log\left(\frac{x}{\theta}\right)\right]^{\alpha-1} dx. \tag{4.1}$$

Substituting $u = \log(x/\theta)$ reduces (4.1) to

$$E(X^r) = \theta^r(1-rc)^{-\alpha},\ \ c<1/r. \tag{4.2}$$

Hence, the mean for the gamma-Pareto distribution is

$$\mu = \theta(1-c)^{-\alpha},\ \ c<1.$$

Note that when $\alpha = 1$ in equation (4.2), $E(X^r) = \theta^r(1-rc)^{-1}$ which represents the non-central moments for the Pareto distribution with parameters $1/c$ and θ.

Using the binomial expansion for $(X-\mu)^r$, the central moments $E(X-\mu)^r$ for any random variable X can be written as

$$E(X-\mu)^r = \sum_{k=0}^{r}\binom{r}{k}(-1)^{r-k}\mu^{r-k}E(X^k). \tag{4.3}$$

Using equations (4.2) and (4.3), the central moments for the gamma-Pareto random variable X can be simplified to

$$E(X-\mu)^r = \theta^r\sum_{k=0}^{r}\binom{r}{k}(-1)^{r-k}(1-c)^{(k-r)\alpha}(1-kc)^{-\alpha}. \tag{4.4}$$

Note that equation (4.4) indicates that the central moments of the gamma-Pareto distribution is an increasing function of θ. Using (4.4), the variance, the skewness and the kurtosis for the gamma-Pareto distribution are respectively expressed as

$$\sigma^2 = \theta^2[(1-2c)^{-\alpha} - (1-c)^{-2\alpha}],\ \ c<0.5 \tag{4.5}$$

$$\gamma_1 = \frac{(1-3c)^{-\alpha} + 2(1-c)^{-3\alpha} - 3(1-2c)^{-\alpha}(1-c)^{-\alpha}}{((1-2c)^{-\alpha} - (1-c)^{-2\alpha})^{3/2}} \tag{4.6}$$

$$\gamma_2 = [(1-2c)^{-\alpha} - (1-c)^{-2\alpha}]^{-2}\left\{(1-4c)^{-\alpha} - 3(1-c)^{-4\alpha} + 6(1-2c)^{-\alpha}(1-c)^{-2\alpha} - 4(1-3c)^{-\alpha}(1-c)^{-\alpha}\right\}. \tag{4.7}$$

Equations (4.6) and (4.7) show that the skewness and the kurtosis are free of θ. Theorem 4 shows that when $\alpha \geq 1$ (or $\alpha < 1$), the non-central moments of gamma-Pareto distribution is bounded below (or above) by the non-central moments of the Pareto distribution.

Theorem 4

Let X be a random variable that follows the gamma-Pareto distribution. If $\alpha \geq 1$, then $E(X^s) \geq \theta^r / (1-rc)$ and if $\alpha < 1$, then $E(X^s) < \theta^r / (1-rc)$.

Theorem 4 Proof

Because $0 < rc < 1$ or $0 < 1-rc < 1$ this implies that for $\alpha \geq 1$, $(1-rc)^{\alpha} \leq 1-rc$ and for $\alpha < 1$, $(1-rc)^{\alpha} > 1-rc$. Thus, if $\alpha \geq 1$, then $\theta^r (1-rc)^{-\alpha} \geq \theta^r (1-rc)^{-1}$ and if $\alpha < 1$, then $\theta^r (1-rc)^{-\alpha} < \theta^r (1-rc)^{-1}$.

Table 2 provides the mode, mean, variance, skewness and kurtosis of the gamma-Pareto distribution for various values of α and c when θ = 1. For fixed α and θ, the mean, variance, skewness and kurtosis are increasing functions of c. For fixed c and θ, the mean, median and variance are increasing functions of α. When α > 1, the following trends for the mode are observed: (1) it increases as c increases with fixed α and θ, and (2) it increases as α increases with fixed c and θ. Table 2 also shows that the skewness is always positive and for fixed α and it increases rapidly as c increases.

Table 2: Mode, Mean, Variance, Skewness and Kurtosis for Some Values of α and c with $\theta = 1$ (*: Undefined)

α	c	Mode	Mean	Variance	Skewness	Kurtosis
0.5	0.1	1	1.0541	0.0069	3.6850	27.6334
	0.2	1	1.1180	0.0410	5.5537	95.1825
	0.3	1	1.1952	0.1526	15.2326	*
	0.4	1	1.2910	0.5694	*	*
1	0.1	1	1.1111	0.0154	2.8111	17.8286
	0.2	1	1.2500	0.1042	4.6476	73.8000
	0.3	1	1.4286	0.4592	16.4438	*
	0.4	1	1.6667	2.2222	*	*
2	0.1	1.0952	1.2346	0.0383	2.2819	13.2512
	0.2	1.1814	1.5625	0.3364	4.4009	77.3004
	0.3	1.2596	2.0408	2.0851	26.1507	*
	0.4	1.3307	2.7778	17.2840	*	*
3	0.1	1.1994	1.3717	0.0714	2.1075	12.0304
	0.2	1.3956	1.9531	0.8149	4.6209	98.2327
	0.3	1.5865	2.9155	7.1251	47.9991	*
	0.4	1.7708	4.6296	103.5670	*	*

The deviation from the mean and the median are used to measure the dispersion and spread in a population from the center. If the median is denoted by M, then the mean deviation from the mean , $D(\mu)$, and the mean deviation from the median, $D(M)$, can be written as

$$D(\mu) = \int_\theta^\mu (\mu - x) g(x) dx + \int_\mu^\infty (x-\mu) g(x) dx$$
$$= 2\int_\theta^\mu (\mu - x) g(x) dx$$
$$= 2\mu G(\mu) - 2\int_\theta^\mu x g(x) dx. \tag{4.6}$$

$$D(M) = \int_\theta^M (M - x) g(x) dx + \int_M^\infty (x - M) g(x) dx$$
$$= 2\int_\theta^M (M - x) g(x) dx + E(X) - M$$
$$= 2MG(M) + \mu - M - 2\int_\theta^M x g(x) dx$$
$$= \mu - 2\int_\theta^M x g(x) dx. \tag{4.7}$$

Consider the integral

$$\int_\theta^m x g(x) dx = \frac{1}{c^\alpha \Gamma(\alpha)} \int_\theta^m (\theta / x)^{1/c} (\log(x/\theta))^{\alpha - 1} dx. \tag{4.8}$$

Using the substitution $u = \log(x/\theta)$, the equation (4.8) can be written as

$$\int_\theta^m x g(x) dx = \frac{\mu}{\Gamma(\alpha)} \gamma[\alpha, (c^{-1} - 1)\log(m/\theta)], \tag{4.9}$$

and by using equations (2.3) and (4.9), the mean deviation from the mean is

$$D(\mu) = 2\mu \left\{ \gamma[\alpha, c^{-1}\log(\mu/\theta)] - \gamma[\alpha, (c^{-1} - 1)\log(\mu/\theta)] \right\} / \Gamma(\alpha),\ c < 1,$$

and the mean deviation from the median is

$$D(M) = \mu \left\{ 1 - \frac{2}{\Gamma(\alpha)} \gamma \left[\alpha, \left(\frac{1}{c} - 1 \right) \log\left(\frac{M}{\theta} \right) \right] \right\},\ c < 1.$$

Parameter Estimation

When $\alpha < 1$, the likelihood function for the gamma-Pareto distribution goes to infinity as θ approaches the sample minimum $x_{(1)}$; thus, when $\alpha < 1$ and θ is estimated by $x_{(1)}$, no MLE for α and c exists. A similar problem was studied by Smith (1985) who proposed an alternative approach for estimating the parameters as follows: If sample data $x_1, x_2, ..., x_n$ are observed, estimate the parameter θ by the sample minimum $x_{(1)}$ and then use the MLE method to estimate α and c by excluding the sample minimum.

Applying Smith's method to obtain the MLE for the gamma-Pareto parameters, the log-likelihood function for the gamma-Pareto distribution is given by

$$\log L(\alpha, c) = \sum_{x_i \neq x_{(1)}} \log g(x_i; x_{(1)}, \alpha, c)$$
$$= \sum_{x_i \neq x_{(1)}} \left\{ -\alpha \log c - \log\Gamma(\alpha) - \log x_{(1)} - (1 + c^{-1}) \times \log(x_i / x_{(1)}) + (\alpha - 1)\log\left(\log(x_i / x_{(1)}) \right) \right\}. \tag{5.1}$$

The derivatives of (5.1) with respect to α and c are given by

$$\frac{\partial \log L}{\partial \alpha} = \sum_{x_i \neq x_{(1)}} \left\{ -\log c - \psi(\alpha) - \log\left(\log\left(\frac{x_i}{x_{(1)}} \right) \right) \right\} \tag{5.2}$$

$$\frac{\partial \log L}{\partial c} = \sum_{x_i \neq x_{(1)}} \left\{ \frac{-\alpha}{c} + \frac{1}{c^2} \log\left(\frac{x_i}{x_{(1)}} \right) \right\}. \tag{5.3}$$

Setting (5.3) to zero and simplifying results in

$$c = \frac{1}{(n-n')\alpha} \sum_{x_i \neq x_{(1)}} \log(x_i / x_{(1)}), \tag{5.4}$$

where n' is the frequency of $x_{(1)}$. Equation (5.4) can be written as

$$c = \frac{1}{\alpha} m_1^*, \tag{5.5}$$

where $m_1^* = \sum_{x_i \neq x_{(1)}} \log(x_i / x_{(1)}) / (n-n')$ is the sample mean for $\log(x_i / x_{(1)})$ after excluding $x_{(1)}$.

Setting (5.2) to zero and using equation (5.5) results in

$$\psi(\alpha) - \log(\alpha) + \log(m_1^*) - m_2^* = 0, \tag{5.6}$$

where $m_2^* = \sum_{x_i \neq x_{(1)}} \log(\log(x_i / x_{(1)})) / (n-n')$ is the sample mean for $\log(\log(x_i / x_{(1)}))$ after excluding $x_{(1)}$. The MLE $\hat{\alpha}$ of α is the solution of equation (5.6) and the MLE $\hat{c}$ of c can be determined by substituting the estimate $\hat{\alpha}$ in equation (5.5).

The initial values for the parameters α and c can be obtained by assuming the random sample $Y_i = \log(X_i / x_{(1)}),\ i = 1, \ldots, n$ are taken from the gamma distribution with parameters α and c. By equating the population mean and the population variance of gamma distribution (with parameters α and c) to the corresponding sample mean and sample variance of y_i, $i = 1, \ldots, n$ and then solving for α and c, the initial values are $c_0 = s_y^2 / \bar{y}$ and $\alpha_0 = \bar{y}^2 / s_y^2$, where s_y^2 and $\bar{y}$ are the sample variance and the sample mean for y_1, y_2, …, y_n.

Lemma 3

The Fisher information matrix for the gamma-Pareto distribution when θ is known is given by

$$I = n \begin{bmatrix} \psi'(\alpha) & 1/c \\ 1/c & \alpha / c^2 \end{bmatrix}. \tag{5.7}$$

Lemma 3 Proof

The Fisher information matrix is defined by $I = [I_{ij}]$ with $I_{ij} = E\left\{ \frac{-\partial^2}{\partial \tau_i \partial \tau_j} \log\left(L(x_i, \vec{\tau})\right) \right\}$, where $\tau_1 = \alpha$ and $\tau_2 = c$. To find *I*, the second derivatives of $\log L(\alpha, c) = \sum_{i=1}^{n} \log g(x_i; \theta, \alpha, c)$, are needed. These can be obtained from the derivatives of (5.2) and (5.3), where $x_{(1)}$ is replaced by θ and the sums are taken form $i = 1$ to *n*. The second derivatives of $\log L(\alpha, c)$ can be written as

$$\partial^2 \log L / \partial \alpha^2 = -n\psi'(\alpha),$$

$$\partial^2 \log L / \partial \alpha \partial c = -n / c,$$

and

$$\partial^2 \log L / \partial c^2 = n\alpha c^{-2} - 2c^{-3} \sum_{i=1}^{n} \log(x_i / \theta).$$

From Lemma 1, it may be concluded that $E(\log(X / \theta)) = \alpha c$. The results of (5.7) follow from taking the negative expected values of the second derivatives.

Theorem 5

The variance-covariance matrix for the gamma-Pareto distribution when θ is known is given by

$$\Sigma = \frac{1}{n(\alpha\psi'(\alpha) - 1)} \begin{bmatrix} \alpha & -c \\ -c & c^2\psi'(\alpha) \end{bmatrix}. \tag{5.8}$$

Theorem 5 Proof

The result follows by taking the inverse of the Fisher information matrix in (5.7). From

(5.8), the variance of the ML estimates $\hat{\alpha}$ and $\hat{c}$ are respectively given by

$$\operatorname{var}(\hat{\alpha}) = \frac{\alpha}{n(\alpha\psi'(\alpha)-1)}$$

and (5.9)

$$\operatorname{var}(\hat{c}) = \frac{c^2\psi'(\alpha)}{n(\alpha\psi'(\alpha)-1)}.$$

From (5.9), the variance of $\hat{\alpha}$ does not depend on the parameter c. Also, as c increases the variance of $\hat{c}$ increases. Using the approximation $\psi'(\alpha) \simeq \alpha^{-1} + \frac{1}{2}\alpha^{-2} + \frac{1}{6}\alpha^{-3}$ (Johnson, et al., 1994, page 357), equations (5.9) can be approximated as

$$\operatorname{var}(\hat{\alpha}) \simeq \frac{6\alpha^3}{n(3\alpha+1)}$$

and (5.10)

$$\operatorname{var}(\hat{c}) \simeq \frac{c^2(6\alpha^2+3\alpha+1)}{n\alpha(3\alpha+1)}.$$

Based on the Central Limit Theorem, $(\hat{\alpha}-\alpha)/se(\hat{\alpha}) \xrightarrow{d} N(0,\ 1)$ and $(\hat{c}-c)/se(\hat{c}) \xrightarrow{d} N(0,\ 1)$ where $se(\hat{\alpha})$ and $se(\hat{c})$ are the standard errors of $\hat{\alpha}$ and $\hat{c}$ respectively.

In the following, the uniformly minimum variance unbiased estimator (UMVUE) is derived for the parameter c assuming that the parameters α and θ are known. The following theorem by Lehmann & Scheffé (1950) is needed in order to find the UMVUE for the parameter c.

Theorem 6

Let $X_1, X_2, ..., X_n$ be a random sample from PDF $g(x,\beta),\ \beta \in \Omega$. Let T be a sufficient statistic for β and let the family $\{g(T,\beta),\ \beta \in \Omega\}$ of probability density functions be complete. If there is a function of T that is an unbiased estimator of β, then this function of T is the UMVUE of β.

Lemma 4

If α and θ are known for the gamma-Pareto distribution, then $T = \sum_{i=1}^{n} \log(x_i/\theta)$ is a sufficient statistic for the parameter c.

Lemma 4 Proof

Let $X_1, X_2, ..., X_n$ be a random sample from the gamma-Pareto distribution, the joint density function is then given by

$$g(x_1, x_2, ..., x_n \mid c) = \prod_{i=1}^{n} \frac{(\theta/x_i)^{1/c}(\log(x_i/\theta))^{\alpha-1}}{x_i c^{\alpha}\Gamma(\alpha)} = k(c,T)h(x_1, x_2, ..., x_n)$$

where $k(c,T) = \exp\left\{-n\alpha\log c + (1+c^{-1})T\right\}$,

$$h(x_1, x_2, ..., x_n) = \exp\left\{-\log(\theta\Gamma(\alpha)) + (\alpha-1)\log(\log(x_i/\theta))\right\}$$

and $T = \sum_{i=1}^{n} \log(x_i/\theta)$. Thus, by using the factorization theorem, the statistic T is sufficient for the parameter c because the parameters α and θ are known.

Theorem 7

If α and θ are known for the gamma-Pareto distribution, then $m_1 = (n\alpha)^{-1}\sum_{i=1}^{n}\log(x_i/\theta)$ is the UMVU estimator for the parameter c.

Theorem 7 Proof

By using Lemma 4, $T = \sum_{i=1}^{n}\log(x_i/\theta)$ is a sufficient statistic for c. It follows from Lemma 1 that $T = \sum_{i=1}^{n}\log(x_i/\theta)$ has a gamma distribution with parameters $n\alpha$ and c. Because gamma density belongs to the exponential family, this implies that $\{g(T,c),\ c>0\}$ is a complete family where $g(x)$ is the gamma-Pareto density. Also from Lemma 1, $E(m_1) = c$, hence m_1 is an unbiased estimator for parameter c. By applying Theorem 6, the statistic m_1 is the

UMVUE of c. From equation (5.5), it is interesting to note that when θ $(= x_{(1)})$ and α are known, the MLE of c is the UMVUE of parameter c.

Applications

The gamma-Pareto is applied to three data sets: The first data set (see Table 3) was analyzed by Akinsete, et al. (2008) and represents Floyd River flood rates for the years 1935-1973 in Iowa, USA. The second data set (see Table 5) is from Mahmoudi (2011) and it represents the fatigue life of 6061-T6 aluminum coupons cut parallel with the direction of rolling and oscillated at 18 cycles per second. The third data set (see Table 7) was analyzed by Eugene (2001) and represents the observed frequencies for Tribolium Confusum Strain #3. The maximum likelihood estimates, the log-likelihood value and the AIC (Akaike Information Criterion) values for the fitted distributions are reported in Tables 4, 6 and 8.

Akinsete (2008) fitted the data in Table 3 to the beta-Pareto distribution and compared the result with the Pareto and the generalized Pareto distribution (Pickands, 1975). Results are shown in Table 4, along with the result obtained by fitting the gamma-Pareto distribution to the data. The results show that both beta-Pareto and gamma-Pareto distributions provide adequate fit to the data. Because the gamma-Pareto distribution has only three parameters, this is an advantage for using it over the four-parameter beta-Pareto distribution. In examining the distribution of this data, observe that the data has a reversed J-shape distribution; this suggests that the gamma-Pareto distribution performs well in modeling reversed J-shape distribution. Figure 4 displays the empirical and the fitted cumulative distribution functions and supports the results shown in Table 4.

Mahmoudi (2011) proposed a five-parameter beta generalized Pareto distribution. He fitted the data (shown in Table 5) and compared the result with beta-Pareto, three-parameter generalized Pareto, Weibull and Pareto distributions. To conserve space, only the results of fitting beta generalized Pareto and beta-Pareto from Mahmoudi (2011) are reported in Table 6 along with the result of fitting the gamma-Pareto distribution to the data. The results in Table 6 indicate that the gamma-Pareto distribution provides the best fit among the distributions. The distribution of this data indicates that the data is approximately symmetric. This example suggests that the gamma-Pareto distribution does very well in fitting the distribution of data which is approximately symmetric. Figure 5 displays the empirical and the fitted cumulative distribution functions and supports the results shown in Table 6.

Table 3: Annual Flood Discharge Rates of the Floyd River Data

Years	Flood Discharge (ft3/s)									
1935-1944	1460	4050	3570	2060	1300	1390	1720	6280	1360	7440
1945-1954	5320	1400	3240	2710	4520	4840	8320	13900	71500	6250
1955-1964	2260	318	1330	970	1920	15100	2870	20600	3810	726
1965-1973	7500	7170	2000	829	17300	4740	13400	2940	5660	

Table 4: Parameter Estimates for the Floyd River Flood Data

Distribution	Parameter Estimates	Log Likelihood	AIC
Pareto	$\hat{k}$ = 0.4125 $\hat{\theta}$ = 318	-392.81	789.62
Generalized Pareto	$\hat{k}$ = -0.3071 $\hat{\theta}$ = 4520	-379.55	763.09
Beta-Pareto	$\hat{\alpha}$ = 6.1550 $\hat{\beta}$ = 24.2434 $\hat{k}$ = 0.0926 $\hat{\theta}$ = 318	-365.45	738.9
Gamma-Pareto	$\hat{\alpha}$ = 5.1454 $\hat{c}$ =0.4712 $\hat{\theta}$ = 318	-365.81	734.9

Figure 4: CDF for Fitted Distributions for Floyd River Flood Data

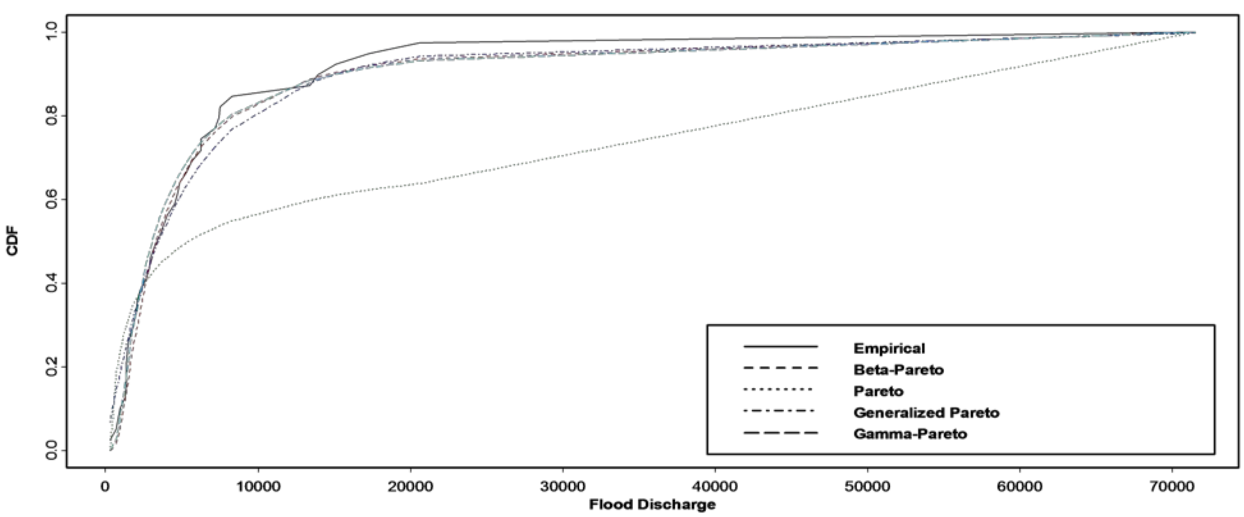


Table 5: Fatigue Life of 6061-T6 Aluminum Coupons Data

70	90	96	97	99	100	103	104	104	105
107	108	108	108	109	109	112	112	113	114
114	114	116	119	120	120	120	121	121	123
124	124	124	124	124	128	128	129	129	130
130	130	131	131	131	131	131	132	132	132
133	134	134	134	134	134	136	136	137	138
138	138	139	139	141	141	142	142	142	142
142	142	144	144	145	146	148	148	149	151
151	152	155	156	157	157	157	157	158	159
162	163	163	164	166	166	168	170	174	196
212									

Table 6: Parameter Estimates for the Fatigue Life of 6061-T6 Aluminum Coupons Data

Distribution	Parameter Estimates	Log Likelihood	AIC
Pareto	$\hat{\alpha}$ =1.579 $\hat{\theta}$ = 70	-548	1100
Beta Pareto	$\hat{\alpha}$ = 485.47 $\hat{\beta}$ = 162.06 $\hat{k}$ = 0.3943 $\hat{\theta}$ = 3.91	-458.65	925.3
Beta- Generalized Pareto	$\hat{\alpha}$ = 12.112 $\hat{\beta}$ =1.702 $\hat{\mu}$ = 40.564 $\hat{k}$ = 0.273 $\hat{\theta}$ = 54.837	-457.85	925.7
Gamma-Pareto	$\hat{\alpha}$ = 15.0209 $\hat{c}$ =0.04258 $\hat{\theta}$ = 70	-448.53	900.6

Eugene (2001) proposed the beta-normal distribution and termed it the generalized normal distribution. Eugene (2001) fitted the data in Table 7 and compared the result with gamma distribution and Lagrange-gamma distribution proposed by Famoye and Govindarajulu (1998). These results are reported in Table 8 along with the result of fitting the data to the gamma-Pareto distribution. The results from the log-likelihood and AIC values indicate that the gamma-Pareto and the generalized normal distributions fit the data best. Figure 6 displays the empirical and the fitted cumulative distribution functions. Figure 6 shows that the generalized normal distribution does not fit the left tail very well, however, the gamma-Pareto distribution does provide a good fit. The distribution shows that the data has a long right tail. This example suggests that the gamma-Pareto distribution does very well in fitting the distributions of data with a long right tail characteristic.

Conclusion

This article defined the gamma-X family and studied a special case of the gamma-X family, the gamma-Pareto distribution. Various properties of the gamma-Pareto distribution were investigated, including moments, deviations from the mean and median, hazard function, unimodality, entropies and Fisher information matrix. Results of the uniformly minimum variance unbiased estimator was obtained for one of the shape parameters of the gamma-Pareto distribution. Three real data sets were fitted to the gamma-Pareto distribution and compared with other known distributions. Results show that the gamma-Pareto distribution provides a good fit to each data set and suggests that the gamma-Pareto distribution can be a good model to fit data with a reversed J-shape, approximately symmetric and long right tail characteristics.

Figure 5: CDF for Fitted Distributions for Fatigue Life of 6061-T6 Aluminum Data

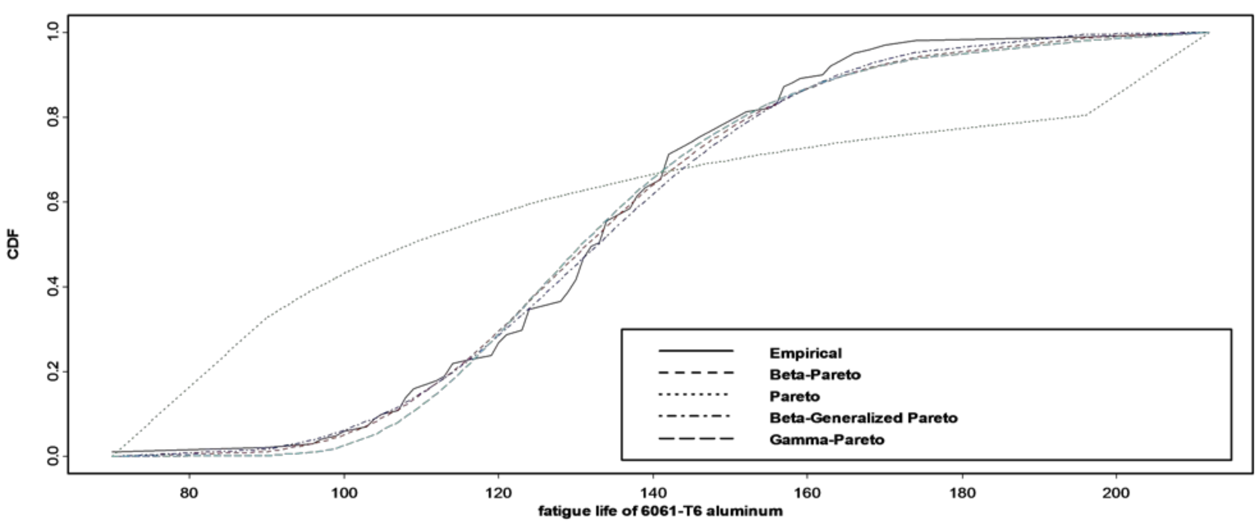


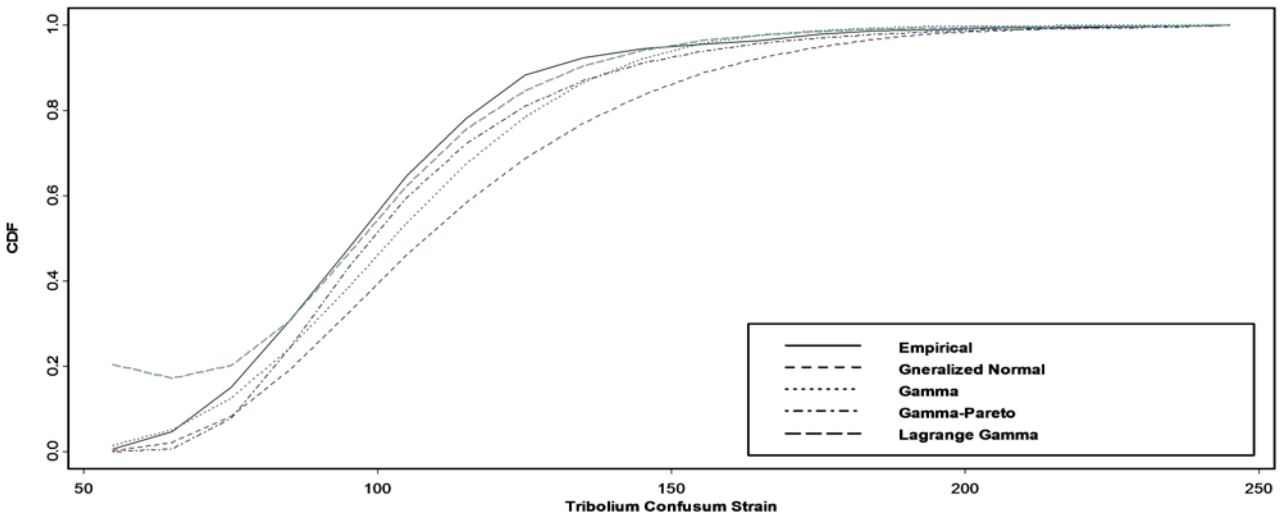
Table 7: Observed frequencies for Tribolium Confusum Strain # 3

x-Values	Frequency	x-Values	Frequency	x-Values	Frequency
55	3	125	51	195	1
65	20	135	20	205	2
75	53	145	11	215	0
85	78	155	6	225	1
95	86	165	4	235	1
105	86	175	7	245	1
115	68	185	5		

Table 8: Parameter Estimates for the Tribolium Confusum Strain # 3 Data

Distribution	Parameter Estimates	Log Likelihood	AIC
Gamma	$\hat{\alpha}$ = 15.15 $\hat{\beta}$ = 6.92	-2335.31	4674.62
Lagrange-gamma	$\hat{r}$ = 31 $\hat{\lambda}$=0.36842 $\hat{\theta}$ = 0.02913	-2314.2	4640.41
Generalized Normal	$\hat{\alpha}$ = 28.68 $\hat{\beta}$ = 0.20 $\hat{\mu}$ = 30.65 $\hat{\sigma}$ = 22.04	-2290.85	4597.71
Gamma-Pareto	$\hat{\alpha}$ = 6.3513 $\hat{c}$ =0.09743 $\hat{\theta}$ = 55	-2297.7	4599.4

Figure 6: CDF for Fitted Distributions for Tribolium Confusum Strain # 3 Data



References

Ahsanullah, M., & Houchens, R. L. (1989). A note on record values from a Pareto distribution. *Pakistan Journal of Statistics*, *5*(*1*), 51-57.

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

Alzaatreh, A., Lee, C., & Famoye, F. (2012). A new method for generating families of continuous distributions. Sumbitted for publication.

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

Eugene, N. (2001). *A Class of Generalized Normal Distribution: Properties, Estimation and Applications* (Doctoral Dissertation). Central Michigan University, Mount Pleasant, MI.

Famoye, F., & Govindarajulu, Z. (1998). On the Lagrange gamma distribution. *Computational Statistics and Data Analysis*, *27*, 421-431.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994) *Continuous Univariate Distributions, Vol. 1, 2nd Ed.* Hoboken, NJ: John Wiley and Sons, Inc.

Lehmann, E. L., & Scheffé, H. (1950). Completeness, similar regions, and unbiased estimation. *Sankhyā*, *10*, 305-340.

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

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

Rényi, A. (1961). On measures of entropy and information. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, *I*, University of California Press, Berkeley, 547-561.

Schroeder, B., Damouras, S., & Gill, P. (2010). Understanding latent sector error and how to protect against them. *ACM Transactions on Storage (TOS)*, *6*(*3*), Article 8.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379-432.

Smith, L. R. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, *72*(*1*), 67-90.