sciendo
ISSN 1335-8871

MEASUREMENT SCIENCE REVIEW

Journal homepage: https://content.sciendo.com

Modeling Extreme Values with Alpha Power Inverse Pareto Distribution

Shumaila Ihtisham[1]*, Sadaf Manzoor[2], Alamgir Khalil[3], Sareer Badshah[4], Muhammad Ijaz[5], Hadia Atta[6]

[1]*Department of Statistics, Islamia College Peshawar, Pakistan,* shumaila.ehtisham@icp.edu.pk
[2]*Department of Statistics, Islamia College Peshawar, Pakistan,* sadaf@icp.edu.pk
[3] *Department of Statistics, University of Peshawar, Pakistan,* profalamgir@yahoo.com
[4]*Department of Statistics, Islamia College Peshawar, Pakistan* sareerbadshah@gmail.com
[5]*Department of Mathematics & Statistics, University of Haripur, KP, Pakistan* ijaz.statistics@gmail.com
[6]*Department of Mathematics, Islamia College Peshawar, Pakistan* hadiaatta@gmail.com

Abstract: The study focuses on the development of a new probability distribution with applications to extreme values. The distribution is proposed by incorporating an additional parameter into the inverse Pareto distribution using the α-Power Transformation. Various properties of the new distribution are derived. The paper also explores the estimation of the parameters by the Maximum Likelihood Estimation (MLE) technique. Simulations are performed to evaluate the performance of the MLEs. In addition, two real data sets with extreme values are used to evaluate the efficacy of the proposed model. It is concluded that the proposed model performs well in the case of extreme values compared to the existing distributions.

Keywords: α-Power transformation, Entropy, Extreme Values, Inverse Pareto distribution, Stochastic Ordering, Stress-Strength Parameter.

1. INTRODUCTION

The problem of extreme values in statistical theory is very common. The behavior of extreme values is studied even if it has a low likelihood of occurring but can have a major impact on the observed events [1]. Extensive literature is available on the applications of extreme values in different fields. For example, in meteorological phenomena, extreme values are found in temperatures, precipitation, wind speeds, etc. Extreme values are essential in predicting flood levels for the construction of bridges, dams, and hydroelectric power plants. The study of extreme values is also significant in droughts to model the problems that arise from river pollution. Furthermore, their applications are found in mechanical, industrial, ocean, earthquake engineering, etc. Since the extreme values can be found either in the upper or lower tails, fitting heavy-tailed distributions are necessary for the extreme value theory. Some of the heavy-tailed distributions that provide a good fit to the data with extreme values include the Pareto and inverse Pareto distributions. These distributions provide a good fit to the data sets with monotone hazard rate functions (hrf), but they may not be suitable for data with non-monotone hrf. Although there are numerous modifications of existing distributions, these do not provide adequate fit to the data sets with extreme values. Some special statistical distributions are necessary to describe the data sets with extreme values that have monotone and non-monotone hrf in order to achieve a good match.

In the recent development of statistical theory, many new probability distributions are proposed to model various data sets better than the existing models. For example, [2] introduced the idea of skew-symmetric distributions by introducing an additional skewness parameter into a normal distribution. [3] proposed the Marshal-Olkin family of distributions for modifying the existing distributions. Following a similar concept [4], [5] and [6] offered the idea of beta, Kumaraswamy and Gamma generated distributions, respectively. [7] introduced the idea of the T-X family of continuous distributions and obtained the gamma-X family, Weibull-X family and beta-exponential-X family of distributions. The details on different generating techniques are provided by [7]. [8] defined a new approach to introduce skewness into the existing distribution(s) called Alpha Power Transformation (APT). The purpose of such modifications is to increase the model's flexibility and to improve its adequacy.

In this study, the Probability Density Function (PDF) of the inverse Pareto distribution is used as a baseline distribution in the APT family to derive a new probability distribution. The Cumulative Distribution Function (CDF) and the PDF of the APT family introduced by [8] are:

DOI: 10.2478/msr-2023-0007

*Corresponding author: shumaila.ehtisham@icp.edu.pk (S. Ihtisham)

$$F_{APT}(x) = \frac{\alpha^{F(x)}-1}{\alpha-1}, \qquad \alpha > 0,\ \alpha \neq 1 \tag{1}$$

$$f_{APT}(x) = \frac{\log\alpha}{\alpha-1}\alpha^{F(x)} f(x), \quad \alpha > 0,\ \alpha \neq 1 \tag{2}$$

First, the exponential distribution is transformed using the APT family of distributions to obtain a two-parameter alpha power exponential distribution. The above generator has been studied by many researchers to introduce the APT Weibull distribution [9], APT Lindley distribution [10], APT inverse Lindley distribution [11], APT Pareto distribution [12], APT exponentiated inverse Rayleigh distribution [13], APT extended generalized exponential distribution [14], etc.

Inverse Pareto (IP) distribution is a heavy-tailed distribution with a monotonically increasing hrf. Since extreme values are more flexible in nature, having monotone and non-monotone hrf, the IP distribution fails to model them adequately. Therefore, there is a need to modify IP distribution to make it more suitable for modeling extreme values with monotone and non-monotone hrf. The cumulative distribution and the probability density functions of the inverse Pareto distribution are defined as

$$F(x) = \frac{x^{\beta}}{(1+x)^{\beta}} \qquad x \geq 0,\ \beta > 0 \tag{3}$$

$$f(x;\beta) = \frac{\beta x^{\beta-1}}{(1+x)^{\beta+1}} \qquad x \geq 0,\ \beta > 0 \tag{4}$$

The contemporary literature on distribution theory contains very few modifications of the IP distribution. It includes the Marshal-Olkin Extended Inverse Pareto (MOEIP) distribution [15] and the Gompertz Inverse Pareto (GoIP) distribution [16]. The MOEIP distribution is derived by using the IP distribution as the baseline distribution in the Marshal-Olkin family. The result is the MOEIP distribution with one scale and two shape parameters. The GoIP distribution, on the other hand, is obtained by using the IP distribution as the baseline distribution in the Gompertz family, resulting in a distribution with one scale and three shape parameters. Despite the fact that both modifications allow more adaptability than their baseline counterpart, they increase the number of parameters to three and four, respectively. There is undoubtedly a gap in the literature on the modification of the IP distribution with respect to various generating techniques that yield more parsimonious models.

The main objective of this study is to derive a new probability distribution that can adequately model data sets with extreme values. Second, to propose a model that captures monotone and non-monotone failure rate functions.

2. SUBJECT & METHODS

This section contains the derivation of a new probability distribution called the $\boldsymbol{\alpha}$–Power inverse Pareto distribution by substituting the CDF and the PDF of the inverse Pareto distribution into the alpha power generator.

A. *$\boldsymbol{\alpha}$–Power Inverse Pareto ($\boldsymbol{\alpha}$PIP) distribution*

A two-parameter αPIP(α,β) distribution can be obtained by substituting (3) and (4) into (2). The PDF of the (αPIP) distribution is given below:

$$f_{\alpha PIP}(x) = \frac{\beta\log\alpha}{\alpha-1}\alpha^{\left(\frac{x}{1+x}\right)^{\beta}} \frac{x^{\beta-1}}{(1+x)^{\beta+1}}, \quad \alpha \neq 1, \alpha > 0 \tag{5}$$

by substituting $z = \left(\frac{x}{1+x}\right)^{\beta}$ in (5), it can be verified that $\int_0^{\infty} f_{\alpha PIP}(x) = 1$. The corresponding CDF of the αPIP distribution is:

$$F_{\alpha PIP}(x) = \frac{\alpha^{\left(\frac{x}{1+x}\right)^{\beta}}-1}{\alpha-1} \tag{6}$$

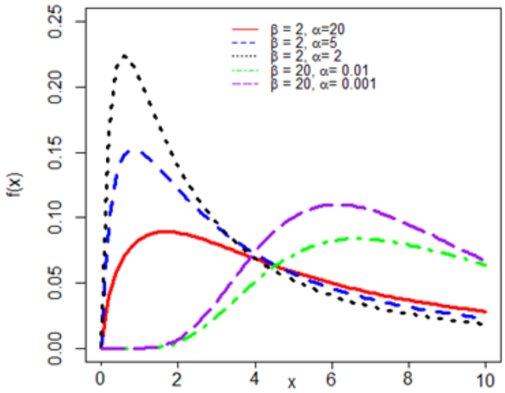

Fig. 1. The shape of the PDF of the αPIP distribution.

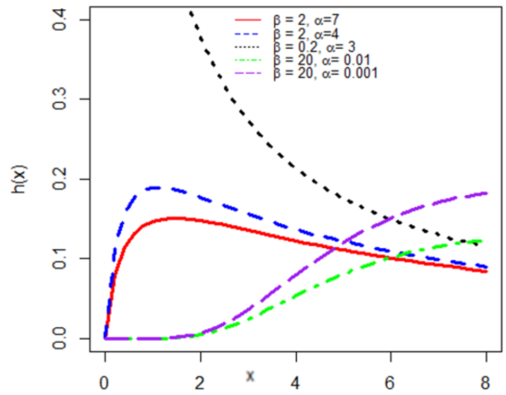

Fig. 2. The shape of the failure rate function of the αPIP distribution.

Hazard (failure) rate and survival (reliability) functions are given by:

$$h_{\alpha PIP}(x) = \frac{\beta\log\alpha}{\left[1-\alpha^{\left(\frac{x}{1+x}\right)^{\beta}-1}\right]} \frac{x^{\beta-1}}{(1+x)^{\beta+1}} \alpha^{\left(\frac{x}{1+x}\right)^{\beta}-1} \tag{7}$$

$$S_{\alpha PIP}(x) = \frac{\alpha}{\alpha-1}\left(1-\alpha^{\left(\frac{x}{1+x}\right)^{\beta}-1}\right) \tag{8}$$

Fig. 1 and Fig. 2 show various shapes of the PDF and the failure rate function of the proposed distribution for different combinations of α and β. Noticeably, the PDF of the αPIP distribution is uni-modal and positively skewed for α > 1 and fixed β, whereas, for α < 1, the shape of the αPIP distribution becomes negatively skewed. The shapes of the hazard rate function are decreasing (monotonc) and increasing-decreasing (non-monotone), depending on different values of α and β.

3. Results

This section contains statistical properties, estimates, simulations and real data applications of the proposed distribution.

A. Quantile function

The following expression is used to obtain the quantile function of the αPIP distribution:

$$F(X) = U \Rightarrow X = F^{-1}(U)$$

where U follows a uniform distribution with the range [0,1] by solving, the p^{th} quantile of the αPIP distribution is given by:

$$x_p = \frac{[log(p(\alpha-1)+1]^{1/\beta}}{(log\alpha)^{1/\beta} - [log(p(\alpha-1)+1)]^{1/\beta}} \tag{9}$$

To obtain the median of the proposed distribution, setting $p = 1/2$, in (9)

$$x_{1/2} = \frac{[log(1/2(\alpha+1))]^{1/\beta}}{(log\alpha)^{1/\beta} - \left[log\left(1/2(\alpha+1)\right)\right]^{1/\beta}} \tag{10}$$

B. Moments

The Moment Generating Function (MGF) of the $\alpha PIP(\alpha, \beta)$ can be obtained as:

$$M_x(t) = E[e^{tx}] = \int_0^\infty e^{tx}\frac{\beta log\alpha}{\alpha-1}\alpha^{\left(\frac{x}{1+x}\right)^\beta}\frac{x^{\beta-1}}{(1+x)^{\beta+1}}dx \tag{11}$$

using $u = \left(\frac{x}{1+x}\right)^\beta$ and the following representation in series form

$$e^{tx} = \sum_{j=0}^\infty \frac{t^j x^j}{j!}$$

$$\alpha^u = \sum_{k=0}^\infty \frac{(log\alpha)^k}{k!}u^k \tag{12}$$

MGF can be obtained as

$$M_x(t) = \frac{\beta}{\alpha-1}\sum_{k=0}^\infty\sum_{n=0}^\infty\sum_{j=0}^\infty \frac{(log\alpha)^{k+1}}{k!}\binom{j+n-1}{n}\frac{t^j}{j!(j+\boldsymbol{k}\beta+\beta+n)} \tag{13}$$

Hence,

$$E(X) = \frac{\beta}{\alpha-1}\sum_{k=0}^\infty\sum_{n=0}^\infty \frac{(log\alpha)^{k+1}}{k!(\boldsymbol{k}\beta+\beta+n+1)} \tag{14}$$

similarly,

$$E(X^2) = \frac{\beta}{\alpha-1}\sum_{k=0}^\infty\sum_{n=0}^\infty \frac{(log\alpha)^{k+1}(n+1)}{k!(\boldsymbol{k}\beta+\beta+n+2)} \tag{15}$$

$$E(X^3) = \frac{\beta}{\alpha-1}\sum_{k=0}^\infty\sum_{n=0}^\infty \frac{(log\alpha)^{k+1}(n+2)(n+1)}{k!(\boldsymbol{k}\beta+\beta+n+3)} \tag{16}$$

$$E(X^4) = \frac{\beta}{\alpha-1}\sum_{k=0}^\infty\sum_{n=0}^\infty \frac{(log\alpha)^{k+1}(n+3)(n+2)(n+1)}{k!(\boldsymbol{k}\beta+\beta+n+4)} \tag{17}$$

C. Skewness and kurtosis

The coefficients of Skewness and kurtosis based on quantiles are given as:

$$Sk = \frac{q_{075} - 2q_{0.50} + q_{0.25}}{q_{0.75} - q_{0.25}}$$

$$K = \frac{q_{0.875} - q_{0.625} - q_{0.375} + q_{0.125}}{q_{0.75} - q_{0.25}}$$

where $q(.)$ is a quantile function. The results of skewness and kurtosis are shown in Fig. 3 and Fig. 4. Since the values of skewness are positive, they indicate a right-skewed distribution. Also, it can be seen that the skewness decreases as α increases. Furthermore, from Fig. 4 it is visible that the kurtosis decreases as β increases.

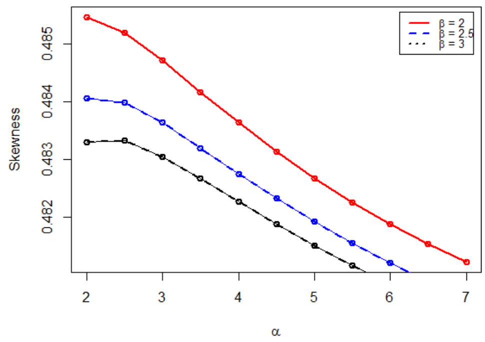

Fig. 3. Skewness for fixed β and varying α.

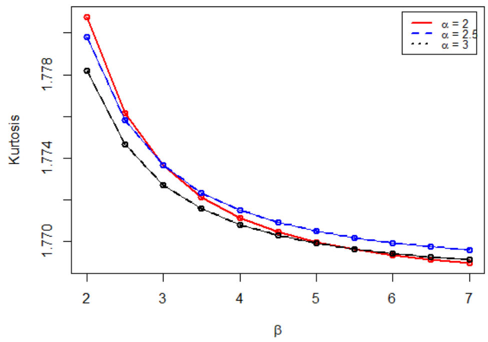

Fig. 4. Kurtosis for fixed α and varying β.

Lemma 1: Let $X_1 \sim \alpha PIP(\alpha_1, \beta)$ and $X_2 \sim \alpha PIP(\alpha_2, \beta)$ be two independently distributed random variables. If $\alpha_1 < \alpha_2$ then $X_1 \leq_{lr} X_2 \quad \forall X$

Proof: The likelihood ratio is given by:

$$\frac{f_{X_1}(x)}{f_{X_2}(x)} = \left(\frac{log\alpha_1}{log\alpha_2}\right)\left(\frac{\alpha_2 - 1}{\alpha_1 - 1}\right)\left(\frac{\alpha_1}{\alpha_2}\right)^{\left(\frac{x}{1+x}\right)^\beta}$$

$$\frac{d}{dx}\left(log\frac{f_{X_1}(x)}{f_{X_2}(x)}\right) = \beta\left(\frac{x}{1+x}\right)^{\beta-1}\frac{1}{(1+x)^2}log\left(\frac{\alpha_1}{\alpha_2}\right) < 0$$

$$if \quad \alpha_1 < \alpha_2, \forall\, x > 0$$

Hence,

$$\text{for } \alpha_1 < \alpha_2,\ X_1 \leq_{lr} X_2 \qquad \forall\, x$$

similarly:

$$X_1 \leq_{hr} X_2 \Rightarrow X_1 \leq_{st} X_2$$

Hence, it is verified that distribution with a likelihood ratio (*lr*) ordering has the same ordering in the hazard rate (*hr*) and distribution (*st*). This suggests the existence of a Uniformly Most Powerful Test UMPT [17].

D. Stress-strength parameter

Assuming $X_1 \sim \alpha PIP(\alpha_1, \beta)$ and $X_2 \sim \alpha PIP(\alpha_2, \beta)$ are two independently distributed random variables, then the stress strength parameter, say *R*, is given by:

$$R = \int_{-\infty}^{\infty} f_1(x) F_2(x) dx$$

using (5) and (6), stress strength parameter *R* can be obtained as:

$$R = \frac{\beta log\alpha_1}{\alpha_1 - 1} \int_0^{\infty} \frac{\alpha_1^{\left(\frac{x}{1+x}\right)^{\beta}}}{x(x-1)} \left(\frac{x}{1+x}\right)^{\beta} \left[\frac{\alpha_2^{\left(\frac{x}{1+x}\right)^{\beta}} - 1}{\alpha_2 - 1}\right] dx \tag{18}$$

$$R = \frac{\log\alpha_1}{(\alpha_1 - 1)(\alpha_2 - 1)} \left[\frac{\alpha_1\alpha_2 - 1}{\log(\alpha_1\alpha_2)} - \frac{\alpha_1 - 1}{\log(\alpha_1)}\right] \tag{19}$$

E. Mean residual life function

The Mean Residual Life (MRL) function is defined as the expected remaining lifetime of a certain object after a fixed time point. The MRL function, say $\mu(t)$, is given by:

$$\mu(t) = \frac{1}{P(X > t)} \int_t^{\infty} P(X > x) dx, \qquad t \geq 0$$

$$\mu(t) = \frac{1}{S(t)} \left(E(t) - \int_0^t x f(x) dx\right) - t, \quad t \geq 0 \tag{20}$$

where,

$$\int_0^t x f(x) dx = \frac{\beta}{\alpha - 1} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \frac{(log\alpha)^{k+1}}{k!(k\beta + \beta + n + 1)} \left(\frac{t}{1+t}\right)^{k\beta + \beta + n + 1} \tag{21}$$

substituting (8), (14) and (21) into (20), $\mu(t)$ can be written as:

$$\mu(t) = \frac{\beta}{\alpha\left(1 - \alpha^{\left(\frac{t}{1+t}\right)^{\beta}} - 1\right)} \left[\sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \frac{(log\alpha)^{k+1}}{k!(\boldsymbol{k\beta + \beta + n + 1})} \left(1 - \left(\frac{t}{1+t}\right)^{k\beta + \beta + n + 1}\right)\right] - t \tag{22}$$

F. Order statistic

The PDF of the r^{th} order statistic, $Y_{r:n}$ can be obtained as:

$$f_{r:n}(\mathrm{y}) = \frac{n!}{(r-1)!(n-r)!} fx(y) \, [Fx(y)]^{r-1} \, [1 - Fx(y)]^{n-r} \tag{23}$$

putting (5) and (6) into (23), the PDF of the r^{th}order statistic is obtained as:

$$f_{r:n}(y) = \frac{n! \, \beta log\alpha}{(r-1)!(n-r)!(\alpha-1)^n} \alpha^{\left(\frac{y}{1+y}\right)^{\beta}} \frac{y^{\beta-1}}{(1+y)^{\beta+1}} \left(\alpha^{\left(\frac{y}{1+y}\right)^{\beta}} - 1\right)^{r-1} \left(\alpha - \alpha^{\left(\frac{y}{1+y}\right)^{\beta}}\right)^{n-r} \tag{24}$$

by setting $r = 1$, the distribution of y_1 is found as:

$$f(y_1) = \frac{n\beta log\alpha}{(\alpha-1)^n} \alpha^{\left(\frac{y}{1+y}\right)^{\beta}} \frac{y^{\beta-1}}{(1+y)^{\beta+1}} \left(\alpha - \alpha^{\left(\frac{y}{1+y}\right)^{\beta}}\right)^{n-1} \tag{25}$$

by setting $r = n$, the distribution of y_{n} is found as:

$$f(y_{\mathrm{n}}) = \frac{n\beta log\alpha}{(\alpha-1)^n} \alpha^{\left(\frac{y}{1+y}\right)^{\beta}} \frac{y^{\beta-1}}{(1+y)^{\beta+1}} \left(\alpha^{\left(\frac{y}{1+y}\right)^{\beta}} - 1\right)^{n-1} \tag{26}$$

G. Shannon and Renyi entropies

The Shannon and Renyi entropies are defined by:

$$E[-log(f(x)] = E\left[-log\left(\frac{\beta log\alpha}{\alpha - 1} \alpha^{\left(\frac{x}{1+x}\right)^{\beta}} \frac{x^{\beta-1}}{(1+x)^{\beta+1}}\right)\right]$$

$$\frac{1}{1-\rho} \log \int_{-\infty}^{\infty} f(x)^{\rho} \, dx = \frac{1}{1-\rho} \int_0^{\infty} \left(\frac{\beta log\alpha}{\alpha - 1} \alpha^{\left(\frac{x}{1+x}\right)^{\beta}} \frac{x^{\beta-1}}{(1+x)^{\beta+1}}\right)^{\rho} dx$$

using (6), the Shannon and Renyi entropies for the αPIP distribution are derived as follows:

$$SE_x = log \frac{\alpha-1}{\beta \, log\alpha} + \sum_{k=0}^{\infty} \frac{(-log\alpha)^{k+1}}{k!} \left[\frac{log\alpha}{k+2} + \beta(\beta-1) \left[\sum_{n=0}^{\infty} \frac{1}{n(\beta k + \beta + n)} - \frac{1}{(k\beta + \beta)^2}\right] - \beta(\beta+1) \sum_{n=0}^{\infty} \frac{1}{n(\beta k + \beta + n)}\right] \tag{27}$$

$$RE_x = \frac{\rho}{1-\rho} log\left[\frac{\beta log\alpha}{\alpha - 1}\right] + \frac{1}{1-\rho} log \sum_{k=0}^{\infty} \frac{(log\alpha)^k \rho^k B(k\beta + \rho\beta - \beta - \rho + 2, \; 2\rho - 1)}{\beta k!} \tag{28}$$

H. Maximum likelihood estimation

The estimates of parameters α and β are obtained through the Maximum Likelihood Estimation (MLE) method. The ML technique is used because it provides consistent, efficient and asymptotically unbiased estimates of unknown parameters. The Likelihood function of the αPIP distribution is given as follows:

$$l(\alpha,\beta) = \beta^n \left(\frac{log\alpha}{\alpha-1}\right)^n \alpha^{\sum\left(\frac{x_i}{1+x_i}\right)^\beta} \prod_{i=1}^{n} \frac{x_i{}^{\beta-1}}{(1+x_i)^{\beta+1}} \tag{29}$$

by taking the logarithm, (29) becomes:

$$logl(\alpha,\beta) = nlog\beta + nlog(log\alpha) - nlog(\alpha-1) + \sum\left(\frac{x_i}{1+x_i}\right)^\beta log\alpha + (\beta-1)\sum logx_i - (\beta+1)\sum log(1+x_i) \tag{30}$$

The partial derivatives of (30) with respect to α, β and then equating to zero, give the following results

$$\frac{\partial logl(\alpha,\beta)}{\partial\alpha} = \frac{n}{log\alpha} - \frac{n\alpha}{(\alpha-1)} + \sum\left(\frac{x_i}{1+x_i}\right)^\beta = 0 \tag{31}$$

$$\frac{\partial logl(\alpha,\beta)}{\partial\beta} = \frac{n}{\beta} + \sum\left(\frac{x_i}{1+x_i}\right)^\beta log\left(\frac{x_i}{1+x_i}\right) log\alpha + \sum logx_i - \sum log(1+x_i) = 0 \tag{32}$$

by solving the system of nonlinear equations in (31) and (32) simultaneously, the MLE of α and β can be achieved. The solution of these equations is not possible analytically, therefore, Conjugate Gradient (CG) or the Newton Raphson (N) routine can be used to obtain the solution. It is known that as $n \to \infty$, the distribution of MLEs approximates a bivariate normal distribution with mean zero and variance-covariance matrix Σ that can be derived by inverting the expected Fisher information matrix I as follows:

$$I = \begin{bmatrix} \frac{\partial^2 logl}{\partial\alpha^2} & \frac{\partial^2 logl}{\partial\alpha\partial\beta} \\ \frac{\partial^2 logl}{\partial\alpha\partial\beta} & \frac{\partial^2 logl}{\partial\beta^2} \end{bmatrix}$$

The second time derivative of (31) and (32) with respect to α and β is

$$\frac{\partial^2 logl}{\partial\alpha^2} = -\frac{n}{\alpha(log\alpha)^2} + \frac{n\alpha}{(\alpha-1)^2} \tag{33}$$

$$\frac{\partial^2 logl}{\partial\alpha\partial\beta} = \frac{1}{\alpha}\sum\left(\frac{x_i}{1+x_i}\right)^\beta log\left(\frac{x_i}{1+x_i}\right) \tag{34}$$

$$\frac{\partial^2 logl}{\partial\beta^2} = -\frac{n}{\beta^2} + log\alpha\sum\left(\frac{x_i}{1+x_i}\right)^\beta \left(log\left(\frac{x_i}{1+x_i}\right)\right)^2 \tag{35}$$

A $(1-\tau)100\%$ asymptotic confidence interval for parameter α and β is:

$$\hat{\alpha} \pm Z_{\tau/2}\sqrt{\Sigma_{11}}$$

$$\hat{\beta} \pm Z_{\tau/2}\sqrt{\Sigma_{22}}$$

I. Simulation study

Simulations are performed in R software to evaluate the average behavior of the MLEs with respect to different sample sizes. Using (9), one thousand random numbers of sample sizes n = 30, 50, 80, 100, and 120 are generated from the αPIP distribution and the average values of the MLEs, Mean Square Error (MSE), and bias are obtained. The bias and MSE are calculated using the following mathematical forms

$$Bias = \frac{1}{w}\sum_{1=1}^{w}\left(\hat{b_\iota} - b\right)$$

$$MSE = \frac{1}{w}\sum_{w}^{w}(\hat{b_\iota} - b)^2$$

where $b = (\alpha, \beta)$. Table 1 shows the average values of MLE, MSE, and Bias. From the results, it is evident that as the sample size increases, the MLEs reach their true values. Moreover, MSE and bias decrease as the sample size n increases. Hence, MLE produces consistent estimators of the parameters of the proposed distribution.

J. Applications

This section contains the application of the αPIP distribution to two real data sets with extreme values. The first data set consists of fourteen observations representing the failure time of an air conditioning system of a Boeing 720 airplane. The data set is taken from [18]. This data set is also used by [19] for the application of inverse Pareto distribution.

Dataset 1:

12 21 26 27 29 48 57 59 70 74 153 326 386 502

The second data set is taken from [20], contains seventy-two observations for the years 1958-1984, showing exceedances of flood peaks of the Wheaton River, Canada. [20] applied Kumaraswamy-Pareto distribution to fit this data set.

Dataset 2:

1.70	2.20	14.4	1.10	0.40	20.6	5.30	0.70	1.90
13.0	12.0	9.30	1.40	18.7	8.50	25.5	11.6	14.1
22.1	1.10	2.50	14.4	1.70	37.6	0.60	2.20	39.0
0.30	15.0	11.0	7.30	22.9	1.70	0.10	1.10	0.60
9.00	1.70	7.00	20.1	0.40	2.80	14.1	9.90	10.4
10.7	30.0	3.60	5.60	30.8	13.3	4.20	25.5	3.40
11.9	21.5	27.6	36.4	2.70	64.0	1.50	2.50	27.4
1.00	27.1	16.8	5.30	9.70	27.5	2.50	27.0	

Table 1. The Average values of MLEs, MSE, and Bias.

Parameter	N	Mean($\hat{\alpha}$)	Mean($\hat{\beta}$)	MSE($\hat{\alpha}$)	MSE($\hat{\beta}$)	Bias($\hat{\alpha}$)	Bias($\hat{\beta}$)
$\alpha = 1.5$ $\beta = 2.5$	30	1.400115	2.631222	0.567685	0.404840	0.099884	0.124203
	50	1.445268	2.624203	0.514986	0.265051	0.054731	0.131222
	80	1.452058	2.605978	0.438589	0.206711	0.047941	0.105978
	100	1.467999	2.596364	0.398013	0.179021	0.032005	0.096363
	120	1.509865	2.570933	0.365937	0.131222	0.009864	0.070933
$\alpha = 2$ $\beta = 2$	30	1.871960	2.323901	1.574074	0.530220	0.128031	0.323902
	50	1.971150	2.211210	1.331714	0.365116	0.028844	0.211209
	80	2.070460	2.164870	1.092099	0.255814	0.070469	0.164879
	100	2.034600	2.103851	0.126548	0.103856	0.034604	0.103856
	120	2.010460	2.093921	0.085362	0.093920	0.010469	0.093920
$\alpha = 2$ $\beta = 4$	30	1.712811	3.884625	0.763510	1.323776	0.287188	0.115375
	50	1.813151	3.937342	0.673690	1.028092	0.186848	0.102658
	80	1.856336	3.912156	0.607480	0.853057	0.143663	0.097844
	100	1.874032	3.925662	0.514501	0.791503	0.125967	0.074338
	120	1.903017	3.947414	0.509920	0.731333	0.096983	0.052585

To compare the goodness of fit of the αPIP distribution with other distributions, the Pareto, the basic Pareto, the Rayleigh, the Kumaraswamy Pareto and the inverse Pareto distributions with the following densities are considered:

- Pareto Distribution (PD)

$$f(x) = \frac{\sigma\beta^{\sigma}}{(x+\beta)^{\sigma+1}} \qquad \sigma,\ \beta > 0, X \geq 0$$

- Basic Pareto Distribution (BP)

$$f(x) = \frac{\beta}{x^{\beta+1}} \qquad \beta > 0, x \geq 1$$

- Rayleigh Distribution (RD)

$$f(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \qquad x > 0, \sigma > 0$$

- Kumaraswamy Pareto Distribution (KPD)

$$f(x) = \frac{abk\beta^k}{x^{k+1}}\left[1-\left(\frac{\beta}{x}\right)^k\right]^{a-1}\left[1-\left\{1-\left(\frac{\beta}{x}\right)^k\right\}^a\right]^{b-1}$$

$$x \geq \beta, a, b, k > 0.$$

- Inverse Pareto Distribution (IPD)

$$f(x) = \frac{\beta x^{\beta-1}}{(1+x)^{\beta+1}} \qquad x > 0, \beta > 0$$

- Two Parameter Inverse Pareto Distribution (TIPD)

$$f(x) = \frac{\alpha\beta x^{\alpha-1}}{(\beta+x)^{\alpha+1}} \qquad x > 0, \alpha, \beta > 0$$

The model adequacy is assessed through Akaike's Information Criteria (AIC), -$ln(\hat{\theta})$, the Kolmogorov-Smirnov test (KS) and *p*-value. Table 2 and Table 3 illustrate the numerical values of the criteria MLE, AIC, -$ln(\hat{\theta})$, KS and *p*-value. Generally, the model is a good fit if the *p*-value is greater, and the values of AIC and -$ln(\hat{\theta})$ are smaller than others. To classify the shape of the hazard rate function, a scaled Total Time on Test (TTT) plot introduced by [21] is applied. The shape of the TTT plot is concave for increasing hrf, convex for decreasing hrf, and shifting curvature of non-monotone hrf.

Table 2 shows an efficient performance of the αPIP distribution in comparison with other fitted distributions because the value of AIC and -$ln(\hat{\theta})$ of the αPIP is lower and the *p*-value is higher among all other distributions. Although the IP distribution fits the first data set well, its AIC and -$ln(\hat{\theta})$ values are comparatively higher than the αPIP. Similarly, for the second data set, Table 3 shows the improved performance of the proposed distribution in terms of AIC, -$ln(\hat{\theta})$ and *p*-value. For this data set the IP distribution does not provide a good fit. Despite the fact that the TIP and the αPIP distributions provide close fits for both data sets, the suggested model outperforms in terms of performance measures. Hence, the proposed model is a good choice among other probability models in the presence of extreme values.

Table 2. The goodness of fit measures for data set 1.

Dist 1	MLE	AIC	-ln($\hat{\theta}$)	KS	*p*-Value
RD	184.5	132.91	65.45	0.59	0.1014
BP	0.539	106.15	52.07	0.26	0.2266
PD	18.51 2.420	107.56	51.78	0.17	0.7426
KPD	15.78 1.910 0.120 0.980	106.56	49.80	0.13	0.9608
IPD	39.38	131.83	79.91	0.68	0.0918
TIP	0.250 16.22	103.76	49.76	0.15	0.8385
αPIP	**7.296 0.079**	**103.52**	**49.76**	**0.15**	**0.8713**

Table 3. The goodness of fit measures for data set 2.

Dist 2	MLE	AIC	$-\ln(\hat{\theta})$	KS	p-Value
RD	0.512	605.97	575.04	0.59	0.0014
BP	0.556	589.51	263.71	0.61	0.0000
PD	21.05 2.420	532.47	264.24	0.11	0.1694
KPD	2.850 85.80 0.050 0.100	548.40	262.40	0.17	0.1734
IPD	3.370	531.51	264.75	0.22	0.0011
TIP	1.853 12.21	524.12	260.06	0.13	0.1322
αPIP	**4.784** **1.254**	**522.58**	**259.29**	**0.12**	**0.2401**

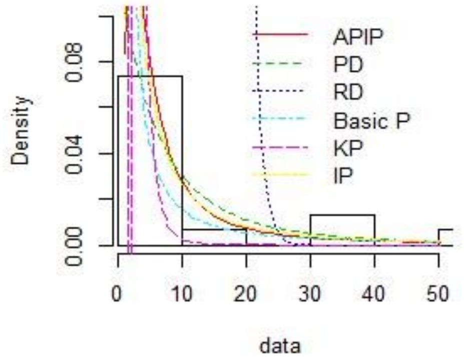

Fig. 5. Comparison between fitted distributions for data set 1.

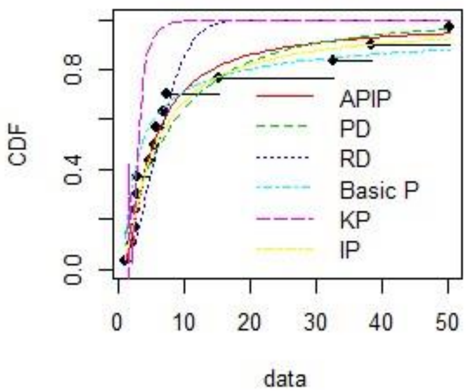

Fig. 6. Comparison between fitted CDF for data set 1.

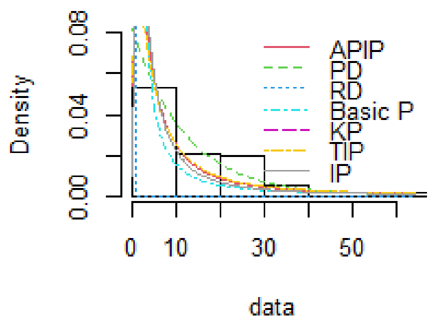

Fig. 7. Comparison between fitted distributions for data set 2.

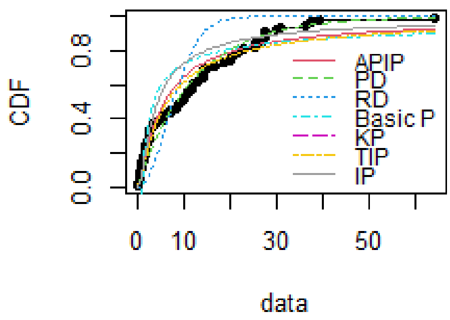

Fig.8. Comparison between fitted distributions for data set 2.

Fig. 5 - Fig. 8 display the empirical densities of the αPIP and other fitted distributions. From these figures, the efficient performance of the αPIP distribution is confirmed. Fig. 9 and Fig. 10 portray the scaled TTT plot of the first data set and the second data set, respectively. The first data set has monotone (decreasing) hrf and the second one has non-monotone hrf.

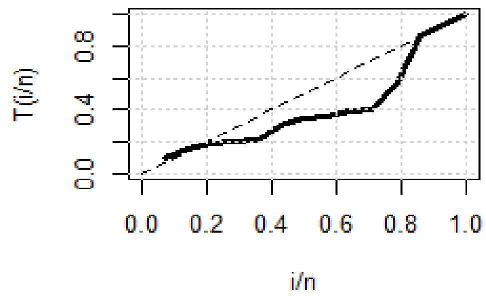

Fig. 9 Scaled TTT plot for data set 1.

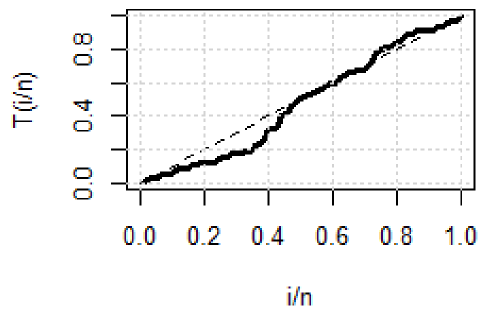

Fig. 10. Scaled TTT plot for data set 2.

CONCLUSION

The study focuses on the development of a new probability model with applications to extreme values. The proposed distribution αPIP is obtained by substituting the cumulative distribution and probability density function of the inverse Pareto distribution into the α-power transformed family of distributions. Various statistical properties of the proposed distribution have been obtained including: quantile function, median, order statistics, moment generating function, mean residual life function, stress strength parameter, and expressions for the Shannon and Renyi entropies. The parameters are estimated through the MLE approach. Two real data sets with extreme values are studied to delineate the significance of the proposed model. It is concluded that the αPIP distribution leads to better results in the presence of extreme values rather than other probability models. Using the proposed model, one can estimate the expected number of failures in the air conditioning system of an airplane. Future research may be conducted to modify the proposed distribution by using the transmutation technique, defining the shape parameter to the power of CDF, or adding a scale parameter. A conclusion section is required. It presents a critical analysis, interpretation and evaluation of the obtained results.

REFERENCES

[1] Hamdi, Y., Haigh, I. D., Parey, S., Wahl, T. (2021). Preface: Advances in extreme value analysis and application to natural hazards. *Natural Hazards and Earth System Sciences*, 21 (5), 1461-1465. https://doi.org/10.5194/nhess-21-1461-2021

[2] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12 (2), 171-178. https://www.jstor.org/stable/4615982

[3] Marshall, A. W., 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. https://www.jstor.org/stable/2337585

[4] Eugene, N, Lee, C, Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods*, 31 (4), 497-512. https://doi.org/10.1081/STA-120003130

[5] Jones, M. C. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, 6 (1), 70-81. https://doi.org/10.1016/j.stamet.2008.04.001

[6] Zografos, K., Balakrishnan, N. (2009). On families of beta- and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, 6 (4), 344-362. https://doi.org/10.1016/j.stamet.2008.12.003

[7] Alzaatreh, A., Lee, C., Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71 (1), 63-79. https://doi.org/10.1007/s40300-013-0007-y

[8] Mahdavi, A., Kundu, D. (2017). A new method for generating distributions with an application to exponential distribution. *Communications in Statistics - Theory and Methods*, 46 (13), 6543-6557. https://doi.org/10.1080/03610926.2015.1130839

[9] Nassar, M., Alzaatreh, A., Mead, M., Abo-Kasem, O. (2017). Alpha power Weibull distribution: Properties and applications. *Communications in Statistics - Theory and Methods*, 46 (20), 10236-10252. https://doi.org/10.1080/03610926.2016.1231816

[10] Dey, S., Ghosh, I., Kumar, D. (2019). Alpha-power transformed Lindley distribution: Properties and associated inference with application to earthquake data. *Annals of Data Science*, 6, 623-650. https://doi.org/10.1007/s40745-018-0163-2

[11] Dey, S., Nassar, M., Kumar, D. (2019). Alpha power transformed inverse Lindley distribution: A distribution with an upside-down bathtub-shaped hazard function. *Journal of Computational and Applied Mathematics*, 348, 130-145. https://doi.org/10.1016/j.cam.2018.03.037

[12] Ihtisham, S., Khalil, A., Manzoor, S., Khan, S. A., Ali, A. (2019). Alpha-Power Pareto distribution: Its properties and applications. *PloS ONE*, 14 (6), e0218027. https://doi.org/10.1371/journal.pone.0218027

[13] Ali, M., Khalil, A., Ijaz, M., Saeed, N. (2021). Alpha-Power Exponentiated Inverse Rayleigh distribution and its applications to real and simulated data. *PloS ONE*, 16 (1), e0245253. https://doi.org/10.1371/journal.pone.0245253

[14] Eghwerido, J.T., Nzei, L.C., Zelibe, S.C. (2022). The alpha power extended generalized exponential distribution. *Journal of Statistics and Management Systems*, 25 (1), 187-210. https://doi.org/10.1080/09720510.2021.1872692

[15] Gharib, M., Mohammed, B. I., Aghel, W. E. R. (2017). Marshll-Olkin extended inverse Pareto distribution and its application. *International Journal of Statistics and Probability*, 6 (6), 71-84. https://doi.org/10.5539/ijsp.v6n6p71

[16] Bashir, S., Naqvi, I. B. (2018). The Gompertz inverse Pareto distribution and extreme value theory. *American Review of Mathematics and Statistics*, 6 (2), 30-37. https://doi.org/10.15640/arms.v6n2a4

[17] Shaked, M, Shanthikumar, J. G. (eds.) (2007). *Stochastic Orders*. Springer. https://doi.org/10.1007/978-0-387-34675-5

[18] Srinivasa Rao, G., Kantam, R. R. L., Rosaiah, K., Pratapa Reddy, J. (2013). Estimation of stress–strength reliability from inverse Rayleigh distribution. *Journal of Industrial and Production Engineering*, 30 (4), 256-263. https://doi.org/10.1080/21681015.2013.828787

[19] Guo, L., Gui, W. (2018). Bayesian and classical estimation of the inverse Pareto distribution and its application to strength-stress models. *American Journal of Mathematical and Management Sciences*, 37 (1), 80-92. https://doi.org/10.1080/01966324.2017.1383217

[20] 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. https://doi.org/10.2991/jsta.2013.12.2.1

[21] Aarset, M. V. (1987). How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, 36 (1), 106-108. https://doi.org/10.1109/TR.1987.5222310

Received August 3, 2022
Accepted March 20, 2023