Kurdish Studies
September 2024
Volume: 12, No: 5, pp 1423 - 1434
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

DOI: 10.53555/ks.v12i5.3486

Exponent Beta Pareto Distribution with Applications in Real and Simulated Dataset

Sofia[1], Muhammad Farooq[1], Muhammad Ali[2], Qasim Shah[1], Muhammad Atif[1*], Qamruz Zaman[1]

[1]Department of Statistics, University of Peshawar, Peshawar, Pakistan
[2]Department of Statistics, Government Postgraduate College Charsadda, Pakistan

***Corresponding author:** Muhammad Atif
*Email; dratif96@gmail.com

Abstract
Statistical theory frequently involves adding an extra parameter to standard distributions. This work introduces a novel distribution called the Exponent Beta Pareto distribution, which includes an additional parameter. Several aspects of the suggested distribution are determined, including the moment generating function, mode, quantiles, entropies, mean residual life function, stochastic orders, and order statistics etc. The parameters of the proposed distribution were estimated using the maximum likelihood estimation technique. The performance of the proposed distribution was evaluated on two real datasets. The proposed distribution outperforms over the various versions of Pareto distributions based on model selection criteria's.

Introduction

In statistical theory, it has become normal practice to improve on the standard distributions over the past few decades. To create new distributions, generators are usually used to add more parameters to the baseline distributions or merge existing ones [1].The goal is to (provide better fit of the complex data) improve the analysis of complicated data structures by simplifying classical distributions. [2] And [3] proposed a way for incorporating a new parameter into current distributions. [4] Introduced beta-generated distributions, where the parent distribution is beta and the baseline distribution is the cumulative distribution function (cdf) of a continuous random variable. [5] Replaced the beta distribution with the Kumaraswamy distribution, modifying the notion from [4]. Furthermore, [6] suggested the T-X family of continuous distributions. [7] Provides a detailed review of constructing univariate continuous distributions. More recently, [8] proposed a novel approach, called alpha power transformation (APT), for (the insertion of an extra parameter) including an additional parameter in a continuous distribution. The objective was to add skewness into the baseline distribution. The alpha power transformation is described as follows:

$$F_{APT(x)} = \begin{cases} \dfrac{\alpha^{F(x)} - 1}{\alpha - 1} & if\, \alpha > 0, \alpha \neq 0 \\ \mathrm{F}(x) & if\, \alpha = 1 \end{cases} \tag{1}$$

The corresponding probability density function is

$$F_{APT(x)} = \begin{cases} \dfrac{\log \alpha}{\alpha - 1} \alpha^{\mathrm{F}(x)} f(x) & if\, \alpha > 0, \alpha \neq 0 \\ f(x) & if\, \alpha = 1 \end{cases} \tag{2}$$

The generator transformed a one-parameter exponential distribution into a two-parameter alpha power distribution. The proposed distribution's features were explored, including explicit formulas for the survival and hazard functions. Quantiles, medians, moments, moments generating functions, order statistics, mean residual life function, and entropy. The shape behavior of pdf, hazard rate function, and survival function were analyzed. [9] And [1] employed the above generator to convert a two parameter Weibull distribution to a three-parameter alpha power distribution. Researchers have used the alpha power transformation to create various distributions, such as the generalized exponential distribution [10], Lindly distribution [11], extended exponential distribution [12], and inverse Lindly distribution [13] etc. The Pareto distribution is widely used to model heavy-tailed phenomena [14]. It has numerous applications, including actuarial science, survival analysis, economics, life testing, hydrology, finance, telecommunications, reliability analysis, physics, and engineering [15-17]. [18] Found that the Pareto distribution is effective for predicting losses in insurance companies, as well as assessing hospital liability. [16] Used the Pareto distribution to model sea clutter intensity returns. [19] Used the Pareto distribution to investigate wealth in society.

[20] Used a generalized Pareto distribution to describe exceedances over a margin in flood control. There are many different forms of Pareto distributions and their generalizations in the literature. According to [21], the first-kind Pareto distribution has the following CDF:

$$F(x) = 1 - \left(\frac{k}{x}\right)^{\beta} \qquad k > 0; \beta > 0; x \geq k \tag{3}$$

The model comprises two parameters: α and k, where k represents the data's lower bound. [18] normalized the data by dividing each observation by a pre-selected lower bound, yielding k = 1. Finally, the CDF and PDF of the Pareto distribution can be represented as:

$$F(x) = 1 - x^{-\beta} \qquad x \geq 1, \beta > 0 \tag{4}$$

$$f(x;\beta) = \frac{\beta}{x^{\beta+1}} \qquad x \geq 1, \beta > 0 \tag{5}$$

Where β represents the scaling parameter. The Pareto distribution's falling hazard rate function and reversed J-shaped pdf may not always provide a good match to data. Risk and loss projections can be made using numerous methods, such as machine life cycle. Human mortality exhibits more adaptable behaviour. That is why scholars offered several modifications and extensions to the Pareto distribution with varying numbers of parameters [17]. Examples include Generalised P [22], Exponentiated P [23, 24], Beta P [25], Beta Generalised P [26], Weibull P [27,28], Kumaraswamy P [29], Kumaraswamy Generalised P [30], Exponentiated Weibull P [31], The Burr X-P [17], and Exponentiated Generalised P [14]. The goal of this work is to propose a new and more flexible distribution, which we call the Exponent Beta Pareto (EBP) distribution, by adding an additional parameter to the Basic Pareto distribution in order to get an adequate match. Numerous features of the EBP distribution are investigated in the next section, along with more appealing forms for the quantile function, median, mode, moments, order statistics, mean residual life function, and stress strength parameter. Lemmas 1 and 2 contain formulae for stochastic ordering, Shannon, and Renyi entropies, respectively. In addition to simulation studies, the following section discusses a method for estimating parameters using maximum likelihood. Two real-world applications are utilised to assess the effectiveness of the suggested paradigm. Conclusions are offered in the final part.

Exponent Beta Pareto (EBP) distribution

Random variable X is said to have a EBP distribution if its pdf and cdf is of the form

$$f(x) = \frac{\alpha\beta\theta x^{-\theta-1}(1-x^{-\theta})^{\alpha-1} e^{-\beta(1-x^{-\theta})^{\alpha}}}{(e^{-\beta}-1)} \tag{6}$$

$$F(x) = \frac{e^{-\beta(1-x^{-\theta})\alpha} - 1}{(e^{-\beta}-1)} \qquad x > 1 \tag{7}$$

The following Figure 1 represents the CDF and PDF of the EBP distribution respectively.

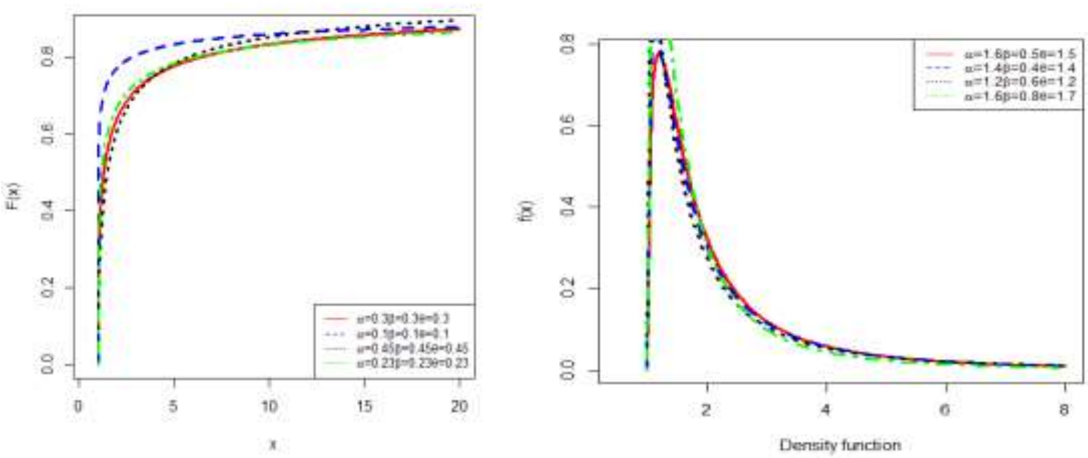

Figure 1: CDF and PDF of the EBP distribution

The survival (reliability) function and hazard rate function are obtained, respectively, as follows:

$$S_{EMAPP}(x) = \frac{e^{-\beta} - e^{-\beta\left(1-x^{-\theta}\right)^{\alpha}}}{e^{-\beta} - 1} \tag{8}$$

$$h_{EMAPP}(x) = \frac{\alpha\beta\theta x^{-\theta-1}(1-x^{-\theta})^{\alpha-1} e^{-\beta(1-x^{-\theta})^{\alpha}}}{(e^{-\beta(1-x^{-\theta})^{\alpha}} e^{-\beta})} \tag{9}$$

The following Figure 2 reprenents the hazard function of EBP

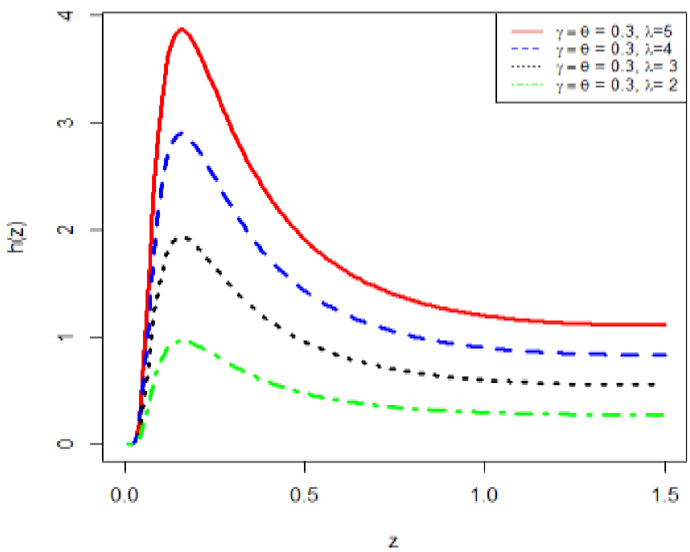

Figure 2: hazard function of EMAPP

Henceforth, a random variable X that follows the distribution in (6) is symbolized by $X\sim$ EBP (a, β, θ).

Quantile function

Quantile function is defined as an inverse of the distribution function. Consider the identity

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

where U follows standard Uniform distribution.

$$X = \left[1 - \left\{-\frac{1}{\beta}\log\right\}\left\{U\left(e^{-\beta}-1\right)+1\right\}^{\frac{1}{\alpha}}\right]^{\frac{-1}{\theta}} \tag{10}$$

Median of EBP distribution can be obtained by putting $U = 1/2$, that is,

$$m = \left[1 - \left\{-\frac{1}{\beta}\log\right\}\left\{\frac{1}{2}\left(e^{-\beta}-1\right)+1\right\}^{\frac{1}{\alpha}}\right]^{\frac{-1}{\theta}} \tag{11}$$

Which is required Median.

Mode

The mode of the distribution can be found by solving the following equation

$$\frac{d}{dx} f_{EMAPP}(X) = 0$$

By taking the derivative of Eq (6) and equating it to zero and solving for x, mode becomes

$$\frac{d}{dx}f(x)=0\Rightarrow\frac{d}{dx}[\frac{\alpha\beta\theta x^{-\theta-1}(1-x^{\theta})^{\alpha-1}e^{-\beta(1-x^{-\theta})^{\alpha}}}{(1-e^{-\beta})}]=0$$
$$-(\theta+1)x^{2\theta}+\frac{\theta(\alpha-1)}{\left(1-x^{-\theta}\right)^{\alpha}}-\alpha\beta\theta\left(1-x^{-\theta}\right)^{\alpha}=0 \tag{12}$$

Mode of the distribution satisfy the above equation.

Moments

R^{th} **Raw moment**

$$\mu_r^{'}=Ex^r=\int_{-\infty}^{+\infty}x^r f(x)dx$$
$$\mu_r^{'}=\frac{\alpha\beta}{(1-e^{-\beta})}\infty\sum_{k=0}\frac{(-\beta)^k}{k!}\beta\left[\alpha(1+k),\left(-\frac{r}{\theta}+1\right)\right]$$

Which is required result (13)

The moment generating function of EBP P distribution is given by

$$M_x(t)=Ee^{tx}=\int_1^{\infty}e^{tx}f(x)dx$$
$$M_x(t)=\int_1^{\infty}\frac{e^{tx}\alpha\beta\theta x^{-\theta-1}(1-x^{-\theta})^{\alpha-1}e^{-\beta(1-x^{-\theta})^{\alpha}}}{(1-e^{-\beta})}dx$$

Using R^{th} moment result we have

$$\frac{\alpha\beta}{(1-e^{-\beta})}\sum_{r=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-\beta)^k t^r}{k!r!}\beta\left[r(1+k),\left(\frac{-r}{\theta}+1\right)\right] \tag{14}$$

Lemma 1: Let $X1$~*EMAPP* ($a1,\beta,\theta$) and $X2$~ EBP ($a2,\beta,\theta$,) be two independent random variables. If $a1 < a2$ then

$$f(x)=\frac{\alpha\beta\theta x^{-\theta-1}(1-x^{\theta})^{\alpha-1}e^{-\beta(1-x^{-\theta})^{\alpha}}}{(1-e^{-\beta})}$$
$$\frac{d}{dx}\log f(x)=-\log(\alpha\beta\theta)-(\theta+1)\log x+(\alpha-1)\log(1-x^{-\theta})-\beta(1-x^{-\theta})^{\alpha}-\log(1-e^{-\beta})$$
$$\frac{d}{dx}\log f(x)=\frac{-(\theta+1)}{x}+x^{-1-\theta}\left[\theta\left(\frac{\alpha-1}{1-x^{-\theta}}\right)-\alpha\beta\theta\left(1-x^{-\theta}\right)^{\alpha-1}\right] \tag{15}$$

Leema 2

Differentiate (15) w.r.t x we have

$$\frac{d^2}{dx^2}\log f(x)=\frac{(\theta+1)}{x^2}-\frac{(\theta+1)}{x^{\theta+2}(1-x^{-\theta})}\left[\theta(\alpha-1)-\alpha\beta\theta\left(1-x^{-\theta}\right)^{\alpha}\right]-\frac{-\theta^2(\alpha-1)}{x^{2\theta+2}(1-x^{-\theta})^2}\left[1+\alpha\beta(1-x^{-\theta})^{\alpha}\right] \tag{16}$$

Order statistics

The i^{th} order statistic is given by

$$f_{i:n}(x)=\frac{n!}{(i-1)!\,(n-i)!}(F(x))^{i-1}f(x)[1-F(x)]^{n-i} \tag{16}$$

Here $f(x)=\dfrac{\alpha\beta\,\theta\,e^{-\theta}\left(1-x^{-\theta}\right)^{\alpha-1}e^{-\beta\left(1-x^{-\theta}\right)^{\alpha}}}{(1-e^{-\beta}).}$

$$F(x)=\frac{e^{-\beta\left(1-x^{-\theta}\right)^{\alpha}}-1}{(e^{-\beta}-1)}$$

$$= \frac{n!}{(i-1)!\,(n-i)!}\frac{\left[e^{-\beta\left(1-x^{-\theta}\right)^{\alpha}-1}\right]^{i-1}}{(e^{-\beta}-1)^{n-1}} \cdot \frac{\alpha\beta\,\theta\,x^{-\theta-1}\left(1-x^{-\theta-1}\right)^{\alpha-1}e^{-\beta\left(1-e^{-\theta}\right)^{\alpha}}}{(1-e^{-\beta})} \cdot \left(\left(e^{-\beta}-e^{-\beta\left(1-x^{-\theta}\right)^{\alpha}}\right)^{n-i}\right) \quad (17)$$

Which is i^{th} order statistic expression. Put $i = 1$ in (16), to have smallest (first) order statistic

$$f_{i:n}(x) = \frac{n!}{(n-1)!}\frac{\alpha\beta\theta x^{-\theta-1}\left(1-x^{-\theta}\right)^{\alpha-1}e^{-\beta\left(1-x^{-\theta}\right)^{\alpha}}}{\left(1-e^{-\beta}\right)^{n}}\left(e^{-\beta}-e^{-\beta\left(1-x^{-\theta}\right)^{\alpha}}\right)^{n-1}$$

Put $i = n$ in eq (17), to get largest order statistic.

$$f_{i:n}(x) = \frac{n!}{(n-1)!}\frac{\left[e^{-\beta\left(1-x^{-\theta}\right)^{\alpha}}-1\right]^{n-1}}{\left(e^{-\beta}-1\right)^{n-1}}\frac{\alpha\beta\theta x^{-\theta-1}\left(1-x^{-\theta}\right)^{\alpha-1}e^{-\beta\left(1-x^{-\theta}\right)^{\alpha}}}{\left(1-e^{-\beta}\right)^{n}} \quad (18)$$

Which is largest order statistic.

Stress-strength parameter

Suppose *X1* and *X2* be two continuous and independent random variables, where *X*1~*EBP(a*1, *B*, θ*)* and *X*2~*EBP (a*2, *β*, θ*)* then the stress strength parameter, say *R*, is defined as

Let $x, \sim$ EBP$(\alpha_1, \beta, \theta)$

$X_2 \sim$ EBP$\left(\alpha_2, \beta, \theta\right)$

By definition,

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

$$R = \int_0^{\infty}\frac{\alpha_1\beta\theta x^{-\theta-1}\left(1-x^{-\theta}\right)^{\alpha_1-1}e^{-\beta\left(1-x^{-\theta}\right)^{\alpha 1}}}{(1-e^{-\beta})} \cdot \frac{\{e^{-\beta}(1-x^{-})^{\alpha_2}-1\}}{(e^{-\beta}-1)}dx$$

$$R = \frac{\alpha_1\beta}{\left(1-e^{-\beta}\right)\left(e^{-\beta}-1\right)}\int_0^1 y^{\alpha_1-1}y^{-\beta y^{\alpha_1}} \cdot e^{-\beta y^{\alpha_2}}dy - \frac{1}{\left(e^{-\beta}-1\right)} \quad (19)$$

Using series

$$e^{-\beta y^{\alpha_1}} = \sum_{l=0}^{\infty}\frac{\left(-\beta y^{\alpha_1}\right)^{l}}{l!} \quad e^{-\beta y^{\alpha_2}} = \sum_{m=0}^{\infty}\frac{\left(-\beta y^{\alpha_2}\right)^{m}}{m!}$$

Using the following series in eq (19), we have

$$R = \frac{\alpha_1\beta}{\left(1-e^{-\beta}\right)\left(e^{-\beta}-1\right)}\int_0^1 y^{\alpha_1-1}\cdot\sum_{l=0}^{\infty}\frac{(-\beta)^{l}\cdot y^{\alpha_1 l}}{l!}\cdot\sum_{m=0}^{\infty}\frac{(-\beta)^{m}y^{\alpha_2 m}}{m!}dy - \left(1-e^{-\beta}\right)$$

$$R = \frac{\alpha_1\beta}{\left(1-e^{-\beta}\right)\left(e^{-\beta}-1\right)}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-\beta)^{l+m}}{l!m!}\cdot\left.\frac{y^{\alpha_1+\alpha_1 l+\alpha_2 m_1}}{\{\alpha_1+\alpha_1 l+\alpha_2 m\}}\right\} - \frac{1}{\left(e^{-\beta}-1\right)}$$

$$R = \frac{\alpha_1\beta}{\left(1-e^{-\beta}\right)\left(e^{-\beta}-1\right)}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-\beta)^{l+m}}{l!m!}\frac{1}{\alpha_1+\alpha_1 l+\alpha_2 m} - \frac{1}{\left(e^{-\beta}-1\right)} \quad (20)$$

Parameters estimation

Maximum likelihood estimation

Let *X1*, *X2*, *X3*, . . ., *Xn* be a random sample from EBP (*a1*, *β*, θ) then the likelihood function is given by

$$f(x) = \frac{\alpha\beta\theta x^{-\theta-1}(1-x^{\theta})^{\alpha-1}e^{-\beta(1-x^{-\theta})^{\alpha}}}{(1-e^{-\beta})}$$

Taking likelihood on both sides we have,

$$\log L\left(f\left(x\right)\right)=n\log\left(\alpha\beta\theta\right)-n\log\left(1-e^{-\beta}\right)-(1+\theta)\sum_{i=1}^{n}\log x_i+\left(\alpha-1\right)\left(\sum_{i=1}^{n}\log(1-x_i^{-\theta})^{\alpha}-\beta\left[\sum_{i=1}^{n}(1-x_i^{-\theta})^{\alpha}\right]\right.$$

(21)

Differentiate equation (21) w.r.t " α " we have

$$\frac{d}{d\alpha}\log L(fx)=\frac{n}{\alpha}+\sum_{i=1}^{n}\log(1-x_i^{-\theta})+\left(\alpha-1\right)\sum_{i=1}^{n}\log(1-x_i^{-\theta})-\beta\sum_{i=1}^{n}\left(1-x_i\right)^{\alpha}\log\left(1-x_i^{-\theta}\right) \quad (22)$$

Differentiate equation (21) w.r.t " β " we have

$$\frac{d}{d\beta}\log Lf(x)=\frac{n}{\beta}-\frac{ne^{-\beta}}{\left(1-e^{-\beta}\right)}-\sum_{i=1}^{n}\left(1-x_i^{-\theta}\right)^{\alpha} \quad (23)$$

Differentiate equation (21) w.r.t " θ " we have

$$\frac{d}{d\theta}\log L(fx)=\frac{n}{\theta}+\sum_{i=1}^{n}\log x_i-\alpha(\alpha-1)\sum_{i=1}^{n}.\frac{x_i^{-\theta}\log x_i}{(1-x_i^{-\theta})}+\alpha\beta\sum_{i=1}^{n}x_i^{-\theta}\log x_i\left(1-x_i^{-\theta}\right)^{\alpha-1} \quad (24)$$

Again Differentiate equation (22) w.r.t " α " we have

$$\frac{d^2}{d\alpha^2}\log Lf(x)=-\frac{n}{\theta^2}-\alpha(\alpha-1)\sum_{i=1}^{n}x_i^{-\theta}\left(\frac{\log x_i}{1-x_i^{-\theta}}\right)^2+\alpha\beta\sum_{i=1}^{n}x_i^{-\theta}\left(\log x_i\right)^2(1-x_i^{-\theta})^{\alpha}(1-\alpha x_i^{-\theta}) \quad (25)$$

Again Differentiate equation (23) w.r.t " β " we have

$$\frac{d^2}{d\beta^2}\log Lf(x)=-\frac{n}{\beta^2}+\frac{ne^{-\beta}}{\left(1-e^{-\beta}\right)} \quad (26)$$

Again Differentiate equation (24) w.r.t " θ " we have

$$\frac{d^2}{d\theta^2}\log Lf(x)=\frac{n}{\theta}-\sum_{i=1}^{n}\log x_i-\alpha(\alpha-1)\sum_{i=1}^{n}x_i^{-\theta}\frac{\log x_i}{(1-x_i^{-\theta})}+\alpha\beta\sum_{i=1}^{n}x_i^{-\theta}\log x_i(1-x_i^{-\theta})^{\alpha-1} \quad (27)$$

Again Differentiate equation (22) w.r.t " β "we have

$$\frac{d^2}{d\alpha\beta}\log Lf(x)=-\sum_{i=1}^{n}(1-x_i^{-\theta})^{\alpha}\log(1-x_i^{-\theta}) \quad (28)$$

Again Differentiate equation (22) w.r.t " $\alpha\theta$ " we have

$$\frac{d}{d\alpha d\theta}\log Lf(x)=-\sum_{i=1}^{n}\frac{(\alpha x_i^{-\theta}\log x_i)}{(1-x_i^{-\theta})}-(\alpha-1)\sum_{i=1}^{n}x_i^{-\theta}\frac{\log x_i}{(1-x_i^{-\theta})}+\beta\sum_{i=1}^{n}x_i^{-\theta}(1-x_i^{-\theta})^{\alpha-1}\log x_i\left\{1+\alpha\log\left(1-x_i^{-\theta}\right)\right\}$$

(29)

Again Differentiate equation (23) w.r.t " θ " we have

$$\frac{d}{d\beta d\theta}\log Lf(x)=\alpha\sum_{i=1}^{n}x_i^{-\theta}(1-x_i^{-\theta})^{\alpha-1}\log x_i \quad (30)$$

Mean residual life function

Assuming X is a continuous random variable with a survival function as stated in Eq (8), the mean residual life function is the expected additional lifetime that a component has survived until time t. The mean residual life function, called $\mu(t)$, is given by

$$\mu\left(t\right)=\frac{1}{S(t)}\left[E(t)-\int_{0}^{t}xf(x)dx\right] \quad (31)$$

$$E(t)=\int_{-\infty}^{+\infty}tf(x)dt=\frac{\int_{1}^{0}t\alpha\beta\theta^{-\theta-1}(1-t^{-\theta})^{\alpha-1}e^{-\beta(1-t^{-\theta})\alpha}dt}{(1-e^{-\beta})} \quad (32)$$

$$E(t)=\frac{\alpha\beta}{(1-e^{-\beta})}\int_0^1(1-y)^{-\frac{1}{\theta}}y^{\alpha-1}e^{-\beta y^{\alpha}}dy \qquad (33)$$

Using series

$$e^{-\beta y^{\alpha}}=\sum_{n=0}^{\infty}\frac{(-\beta y^{\alpha})^n}{n!}$$

$$E(t)=\frac{\alpha\beta}{(1-e^{-\beta})}\sum_{n=0}^{\infty}\frac{(-\beta)^n}{n!}\int_0^1(1-y)^{-\frac{1}{\theta}+1-1}y^{\alpha(n+1)-1}dy \qquad (34)$$

Using beta function in (34)

$$\beta(m,n)=\int_0^1 x^{m-1}(1-x)^{n-1}dx$$

$$E(t)=\frac{\alpha\beta}{(1-e^{-\beta})}\sum_{n=0}^{\alpha}\frac{(-\beta)^n}{n!}\beta\left[\alpha(n+1),-\frac{1}{\theta}+1\right] \qquad (35)$$

Now

$$\int_0^t xf(x)dx=\int_1^t\frac{x\alpha\beta\theta x^{-\theta-1}(1-x^{-\theta})^{\alpha-1}e^{-\beta(1-x^{-\theta})^{\alpha}}}{(1-e^{-\beta})}dx \qquad (36)$$

$$\int_0^t f(x)dx=\frac{\alpha\beta}{(1-\beta^{-\beta})}\sum_{l=0}^{\infty}\frac{(-\beta)^{\alpha}}{l!}\int_1^{1-t}(1-y)^{\frac{-1}{\theta}}y^{\alpha(l+1)-1}dy \qquad (37)$$

Using equation (8),(35),(37) in (31) we have

$$\int_0^t f(x)dx=\frac{\alpha\beta}{e^{-\beta(1-x^{-\theta})\alpha}-e^{-\beta}}\left[\sum_{n=0}^{\infty}\frac{(-\beta)^n}{n!}\beta\left\{\alpha(n+1)-\frac{1}{\theta}+1\right\}-\sum_{l=0}^{\infty}\frac{(-\beta)^l}{l!}\int_0^{1-t}(1-y)^{-\frac{1}{\theta}}y^{\alpha(l+1)-1}dy\right]$$

(38)

Which is required solution

The Shannon and Renyi entropy for a random variable X with Efficient modified alpha power Pareto distribution is as follows.

$$S.Ex=E\left[-\log f(x)\right]=-E\left[\log f(x)\right]=-\log\left[Ef(x)\right]$$

$$R.E=\frac{1}{1-V}\log\left[\int_{-\infty}^{+\infty}f(x)dx\right]$$

Proof:

For EBP distribution, the Shannon and Renyi entropies are given respectively as

$$S.Ex=-\log\left[\frac{\alpha^2\beta^2}{(1-e^{-\beta})}\sum_{n=0}^{\infty}\frac{(-2\beta)^n}{n!}\beta\left(n\alpha+2\alpha-1,2+\frac{1}{\theta}\right)\right] \qquad (39)$$

$$R.E=\frac{1}{1-V}\log\left[\frac{\alpha^v\beta^v\theta^v}{(1-e^{-\beta})^v}\sum_{l=0}^{\infty}\frac{(-\beta V)^l}{l!}\beta\left\{\left(\alpha(\alpha+l)-V+1\right),\left(V+\frac{1}{\theta}(V-1)\right)\right\}\right] \qquad (40)$$

Simulations study

Simulation study has been performed for average MLEs, Mean Square Error (MSE) and bias.

W= 100 samples of size n = 70, 110, and 170 were produced form EMAPP distribution. Random numbers were generated by the following expression

$$X=\left[\log\left(\frac{\alpha/\left(U(\alpha-1)+1\right)}{\log\alpha}\right)\right]^{-\frac{1}{\beta}}$$

where U is uniform random numbers with parameter [0,1] Bias and MSE are calculated by

$$Bias=\frac{1}{W}\sum_{1=1}^{w}(\hat{b}_1-b)$$

$$MSE = \frac{1}{W}\sum_{1=1}^{w}(\hat{b}_1 - b)^2$$

Where b equals (α, β, θ). Simulation results were obtained for various combinations of α, θ and β. Table 2 displays the average MSE and bias values. These estimations are consistent and close to genuine parameter values based on the sample. Size rises. Increasing sample size leads to lower MSEs and biases for all parameter combinations. The MLE technique accurately estimates EBP distribution characteristics.

Table 1: MSE and bias

Parameters	N	MeanZ0	MeanZ1	MeanZ2	MSE0	MSE1	MSE2	BIAS0	BIAS1	BIAS2
A0=2	70	1.9623	2.1645	5.3930	0.18011	7.4612	7.3431	-0.0376	-0.8354	1.3930
B0=3	110	1.9798	2.8001	4.7052	0.1294	6.4172	4.4098	-0.0201	-0.1998	0.7052
C0=4 w=100	170	1.9791	2.9522	4.4471	0.0379	3.8103	2.4992	-0.0208	-0.0477	0.4471

Applications

Two data sets have been analyzed to demonstrate the performance of the proposed model. The first data set consists of 40 wind related catastrophes used by [33]. It includes claims of $2,000,000. The sorted values, observed in millions are as follows.

Table 2: Dataset

1.1	1.4	1.3	1.7	1.9	1.8	1.6	2.2	1.7	2.7
4.1	1.8	1.5	1.2	1.4	3, 1	7	2.3	1.6	2

The second data set is related with the monthly actual taxes revenue in Egypt from January 2006 to November 2010.The data has been analyzed by [17, 34]. The data values are as follows.

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

The fit of the proposed EBP distribution is compared with several other competitive models namely Pareto distribution by [35], generalized pareto Distribution, by [22] alpha power pareto Distribution by [29], Kumaraswamy Pareto distribution, by [14] ,[36] with the following pdfs

Pareto Distribution

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

Generalized Pareto Distribution

$$f(x) = \frac{1}{\delta}(1+\frac{\xi x}{\delta})^{-\frac{1}{\xi}1} \qquad \xi \ne 0, X \ge 0, \delta > 0$$

Kumaraswamy Pareto distribution

$$f(x) = \frac{abk\beta^k}{x^{k+1}}\left[1-\left(\frac{\beta}{x}\right)^k\right]^{a-1}\left[(1-\left(\frac{\beta}{x}\right)^k)^a\right]^{b-1} \qquad x \ge \beta, a, b, k > 0$$

Alpha Power Pareto Distribution

$$f_{APP}(x) = \begin{cases} \dfrac{\beta \log x}{\alpha - 1}\alpha^{1-x-\beta}\, x^{-\beta-1} & \alpha \ne 1 \\ f(x) \quad \alpha \ne 1 & \end{cases}$$

Using R for Adequacy Model package, the goodness of fit test is used to evaluate the performance of EBP and other Pareto distributions. Goodness of fit criteria include the result of Akaike's Information Criteria. There are several information criteria, including AIC, CAIC, BIC, and HQIC. Tables 3 and 4 show the Kulmogrov-Smirnov test (KS) results and p-values. A model is called good fit if it meets all of the criteria and has a higher p value.

Table 3 Goodness of fit result for data set 1.

Distribution	MLE				AIC	CAIC	BIC	HQIC	P-value
EBP	2.3789	-2.8682	4.2116		36.4760	37.976	39.463	37.059	0.995
PD	1.6970				44.4142	44.636	45.410	44.608	0.077
GPD	3.0451	-0.7223			59.6902	60.396	61.681	60.079	
APPD	22.648	3.12158			36.8918	37.597	38.883	37.280	0.660
KPD	6.2153	1.68050	2.7222	0.8728	38.7520	41.418	42.735	39.529	6.777

Table 4. Goodness of fit result for data set 2.

Distribution	MLE				AIC	CAIC	BIC	HQIC	P-value
EBP	20.6192	8.1199	0.9027		383.728	384.165	389.961	386.161	0.965
PD	0.40654				516.459	516.529	518.536	517.270	1.273
GPD	18.8328	-0.429			417.771	417.985	421.926	419.393	0.002
APPD	29.5775	0.7695			465.322	465.536	469.477	466.944	1.029e-06
KPD	7.94960	3.4110	1.0904	2.4508	384.571	385.312	392.881	387.815	2.2e-16

Tables 3 and 4 show that the EBP distribution has lower AIC, CAIC, BIC, HQIC, and -log-likelihood values than alternative fitted distributions. Figures 3 and 4 show promising results for the proposed distribution. Figures 5 and 6. QQ-plot and PP-plot are included. Although certain QQ-plot values deviate from the fitted line, this is typical of heavy-tailed distributions [37].

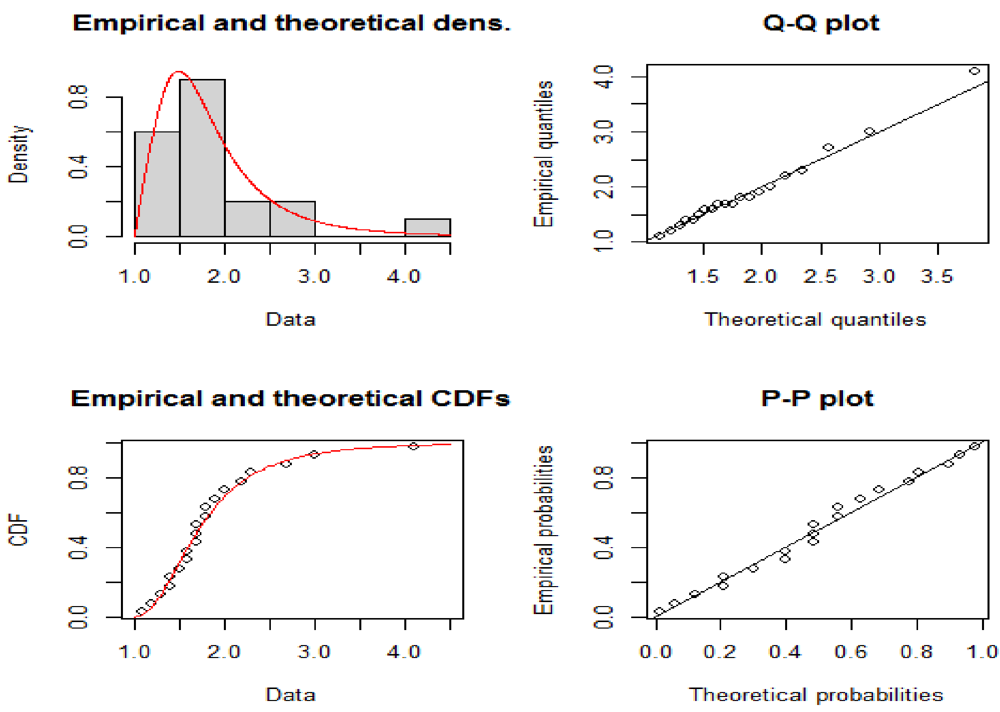

Figure 3: Plots for data set 1

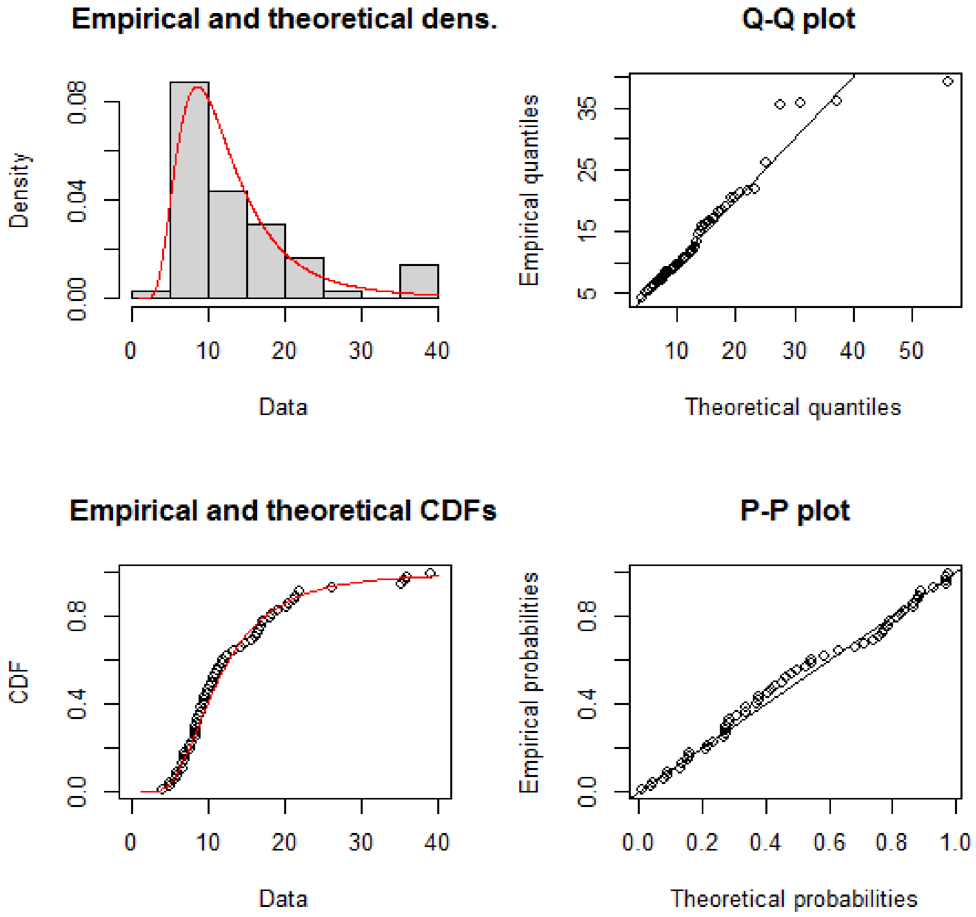

Figure 4: Plots for data set 2

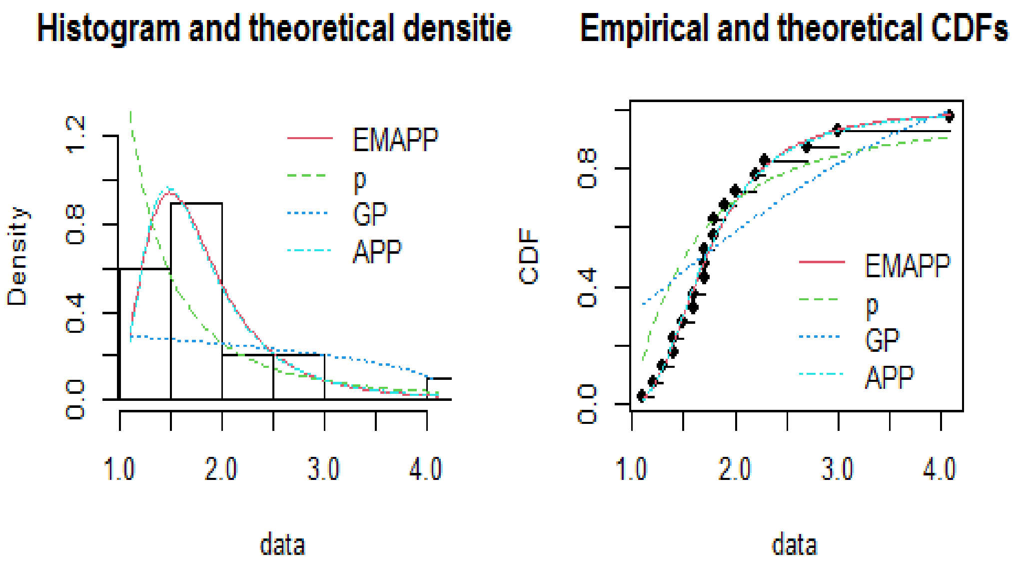

Figure 5: Comparison between fitted distributions for dataset 1.

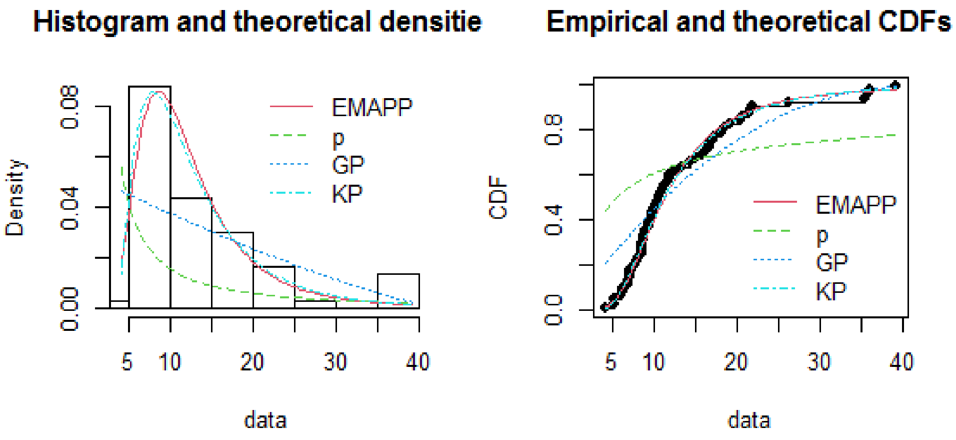

Figure 6: Comparison between fitted distributions for dataset 2

Conclusion

The new distribution, known as the EBP distribution, is introduced using alpha power Pareto distribution. The transformation adds skewness to a family of distribution functions. Various features of the distribution have been derived, including moment generation. Topics covered include function, order statistics, stress strength, mean residual life, mode, stochastic ordering, and entropy expressions. The maximum likelihood estimation approach was utilized to obtain parameter estimates for unknown parameters. The proposed distribution outperformed other Pareto distributions on two real datasets.

References

1. Dey S, Sharma VK, Mesfioui M. A new extension of Weibull distribution with application to lifetime data.Annals of Data Science. 2017; 4(1):31–61.
2. Mudholkar G S, Srivastava DK. Exponentiated Weibull family for analyzing bathtub failure-rate data.IEEE transactions on reliability. 1993; 42(2):299–302.
3. Marshall AW, Olkin I. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika. 1997; 84(3):641–52.
4. Eugene N, Lee C, Famoye F. Beta-normal distribution and its applications. Communications in Statistics-Theory and methods. 2002; 31(4):497–512.
5. Jones M. C. Kumaraswamy's distribution: A beta-type distribution with some tractability advantages.Statistical Methodology. 2009; 6(1):70–81.
6. Alzaatreh A, Lee C, Famoye F. A new method for generating families of continuous distributions.Metron. 2013; 71(1):63–79.
7. Lee C, Famoye F, Alzaatreh AY. Methods for generating families of univariate continuous distributionsin the recent decades. Wiley Interdisciplinary Reviews: Computational Statistics. 2013; 5(3):219–38.
8. Mahdavi A, Kundu D. A new method for generating distributions with an application to exponential distribution.Communications in Statistics-Theory and Methods. 2017; 46(13):6543–57.
9. Nassar M, Alzaatreh A, Mead M, Abo-Kasem O. Alpha power Weibull distribution: Properties and applications.Communications in Statistics-Theory and Methods. 2017 18; 46(20):10236–52.
10. Dey S, Alzaatreh A, Zhang C, Kumar D. A new extension of generalized exponential distribution with application to Ozone data. Ozone: Science & Engineering. 2017; 39(4):273–85.
11. Dey S, Ghosh I, Kumar D. Alpha-Power Transformed Lindley Distribution: Properties and Associated Inference with Application to Earthquake Data. Annals of Data Science. 2018:1–28.
12. Hassan AS, Mohamd RE, Elgarhy M, Fayomi A. Alpha power transformed extended exponential distribution: properties and applications. Journal of Nonlinear Sciences and Applications. 2018; 12(4), 62–67.
13. Dey S, Nassar M, Kumar D. Alpha power transformed inverse Lindley distribution: A distribution with an upside-down bathtub-shaped hazard function. Journal of Computational and Applied Mathematics.2019; 348:130–45.
14. Lee S, Kim J H. Exponentiated generalized Pareto distribution: Properties and applications towards extreme value theory. Communications in Statistics-Theory and Methods. 2018; 1–25.
15. Brazauskas V, Serfling R. Favorable estimators for fitting Pareto models: A study using goodness-of-fit measures with actual data. ASTIN Bulletin: The Journal of the IAA. 2003; 33(2):365–81.
16. Farshchian M, Posner F L. The Pareto distribution for low grazing angle and high resolution X-band sea clutter. Naval Research Lab Washington DC; 2010.
17. Korkmaz M, Altun E, Yousof H, Afify A, Nadarajah S. The Burr X Pareto Distribution: Properties, Applications and VaR Estimation. Journal of Risk and Financial Management. 2018; 11(1):1.
18. Philbrick S W. A practical guide to the single parameter Pareto distribution. PCAS LXXII. 1985; 44–85.

19. Levy M, Levy H. Investment talent and the Pareto wealth distribution: Theoretical and experimental analysis. Review of Economics and Statistics. 2003; 85(3):709–25.
20. Castillo E, Hadi AS. Fitting the generalized Pareto distribution to data. Journal of the American Statistical Association. 1997; 92(440):1609–20.
21. Johnson N. L., and Kotz S., Balakrishnan N. Continuous Univariate Distributions-I. New York: John Wiley; 1994.
22. Pickands J III. Statistical inference using extreme order statistics. The Annals of Statistics. 1975;3(1):119–31.
23. Gupta RC, Gupta PL, Gupta RD. Modeling failure time data by Lehman alternatives. Communications in Statistics-Theory and methods. 1998; 27(4):887–904.
24. Nadarajah S. Exponentiated Pareto distributions. Statistics. 2005; 39(3):255–60.
25. Akinsete A, Famoye F, Lee C. The beta-Pareto distribution. Statistics. 2008; 42(6):547–63.
26. Mahmoudi E. The beta generalized Pareto distribution with application to lifetime data. Mathematics and computers in Simulation. 2011; 81(11):2414–30.
27. Alzaatreh A, Famoye F, Lee C. Weibull-Pareto distribution and its applications. Communications in Statistics-Theory and Methods. 2013; 42(9):1673–91.
28. Tahir M H, Cordeiro GM, Alzaatreh A, Mansoor M, Zubair M. A new Weibull–Pareto distribution: properties and applications. Communications in Statistics-Simulation and Computation. 2016; 45(10):3548–67.
29. Pereira MB, Silva RB, Zea LM, Cordeiro GM. The kumaraswamy Pareto distribution. arXiv preprint arXiv:1204.1389. 2012.
30. Nadarajah S, Eljabri S. The kumaraswamy gp distribution. Journal of Data Science. 2013; 11(4):739–66.
31. Afify AZ, Yousof HM, Hamedani GG, Aryal G. The exponentiated Weibull-Pareto distribution with application. J. Stat. Theory Appl. 2016; 15:328–46.
32. Shaked M, Shanthikumar Jeyaveerasingam G. Stochastic Orders. Series: Springer Series in Statistics. New York: Springer; 2007.
33. Hogg R. and Klugman S.A. Loss Distributions. New York: Wiley; 1984.
34. Mead ME, Afify AZ, Hamedani GG, Ghosh I. The beta exponential Fre´chet distribution with applications. Austrian Journal of Statistics. 2017; 46(1):41–63.
35. Johnson N. L., and Kotz S. Continuous Univariate Distributions-2. Boston: Houghton Mifflin; 1970.
36. Guo L, Gui W. Bayesian and Classical Estimation of the Inverse Pareto Distribution and Its Application
37. to Strength-Stress Models. American Journal of Mathematical and Management Sciences. 2018; 37 (1):80–92.
38. Brodin E, Rootze´n H. Univariate and bivariate GPD methods for predicting extreme wind storm losses. Insurance: Mathematics and Economics. 2009; 44(3):345–56.