Pakistan Journal of Statistics and Operation Research

Chen Pareto Distribution: Properties and Application

Phillip Oluwatobi Awodutire

*Corresponding author

1. Department of Mathematical Sciences, University of Africa, Toru Orua, Nigeria , phillip.awodutire@gmail.com

Abstract

In this work, a new distribution called the Chen Pareto distribution was derived using the Chen-G family of distributions. The mixture representation of the distribution was obtained. Furthermore, some statistical properties such as moments, moment generating functions, order statistics properties of the distribution were explored. The parameter estimation for the distribution was done using the maximum likelihood estimation method and the performance of estimators was assessed by conducting an extensive simulation study. The distribution was applied to a real data set in which it performed best when compared to some related distributions
Key Words: Chen-G, Pareto distribution; statistical properties; estimation; simulation

Mathematical Subject Classification: 60E05, 62E15.

1. Introduction

In a bid to get distributions that are more flexible and better fit to real-life data, generalization(extension) of the distributions is done by the introduction of additional parameters to the baseline distribution. Recently, researchers do such an extension by introducing the baseline distribution to a family of distribution. Several mathematical properties of the extended distributions may be easily explored using mixture forms of exponentiated-G (exp-G) of the distributions.
The Pareto distribution has been extensively studied and applied to many situations. This includes health, economics, finances, survival analysis, engineering, and Actuarial sciences. The distribution gains interest due to its ability to model heavy tailed data Lee and Kim (2018). It has many applications in actuarial science, survival analysis, economics, life testing, hydrology, finance, telecommunication, reliability analysis, physics, and engineering such as Brazauskas and Serfling (2003), Farshchian and Posner (2010) and Korkmaz *et.al* (2018).
The Pareto distribution has been studied and extended by several authors. For example, Weibull Pareto Alzaatreh *et.al* (2013), Beta Pareto Akinsete and Famoye (2008), Kumaraswamy Pareto Pereira *et.al* (2012), Exponentiated Pareto Gupta et .al (1998) Nadarajah(2005), Beta Generalized Pareto Mahmoudi (2011), Gamma Pareto Alzaatreh *et.al* (2012), Transmuted Pareto Faton and Llukan (2014) e.t.c. The c.d.f of Pareto distribution is

$$G(x) = 1 - (\frac{k}{x})^{\sigma} \quad x \geq k, k, \sigma > 0 \tag{1}$$

$$g(x) = \frac{\sigma k^{\sigma}}{x^{\sigma+1}} \quad x \geq k, k, \sigma > 0 \tag{2}$$

where k and σ is the scale and shape parameter respectively Considering a baseline cumulative distribution function (c.d.f) G(x) with a corresponding probability distribution function (p.d.f) g(x), the c.d.f of the Chen-G family of distribution is

$$F(x) = W[1 - e^{\mu(1-e^{G(x)^{\phi}})}] \quad x > 0, \phi, |\mu| > 0 \tag{3}$$

and the p.d.f as

$$f(x) = W\mu\phi g(x)G(x)^{\phi-1}e^{G(x)^{\beta}}e^{\mu(1-eG(x)^{\beta})} \quad x > 0, \phi, |\mu| > 0 \tag{4}$$

where

$$W = \frac{1}{1 - e^{\mu(1-e)}}$$

It is of importance to note that W is a normalizing constant, μ is the scale parameter, and ϕ is the shape parameter
For this research, we considered deriving a generalized form of Pareto distribution using the Chen-G family of distribution Anzagra (2020). The plan of this article is as follows. Section 2 discusses the derivation of the Chen-Pareto distribution. In section 3, we studied the mixture representation of the p.d.f and the c.d.f of the Chen-Pareto distribution. Some statistical properties of the derived distribution were considered in section 4 after which the parameter estimation of the distribution is done in section 5. Section 6 discussed the simulation study to assess the performance of the parameters of the distribution. Section 7 discussed the application to real-life data. The conclusions were done in section 8.

2. Derivation of the Chen-Pareto Distribution

In this section, we discussed the derivation of the c.d.f and p.d.f of the Chen-Pareto Distribution.
Inserting Equation (1) in (3) and Equation (2) in (4) respectively, a random variable X is said to be distributed to Chen Pareto distribution when its c.d.f F(x) is

$$F(x) = W[1 - e^{\mu(1-e^{(1-(\frac{k}{x})^{\sigma})^{\phi}})}] \quad x \geq k, \mu, \sigma, \phi, k > 0 \tag{5}$$

and the corresponding p.d.f is

$$f(x) = \frac{W\mu\phi\sigma k^{\sigma}(1-(\frac{k}{x})^{\sigma})^{\phi-1}e^{(1-(\frac{k}{x})^{\sigma})}e^{\mu(1-e^{(1-(\frac{k}{x})^{\sigma})})^{\phi}}}{x^{\sigma+1}} \quad x \geq k, \mu, \sigma, \phi, k > 0 \tag{6}$$

The Chen-Pareto distribution becomes pareto distribution when μ=1 and ϕ=1. It becomes Chen distribution when k=1 and σ=1
Figures 1 and 2 show the p.d.f and c.d.f of the Chen-Pareto distribution with assignment values for the parameters.

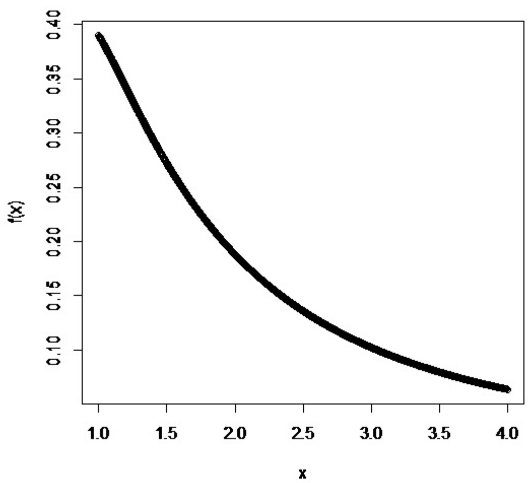
3. Mixture Representation

In this section, we derived the mixture representation of the Chen-Pareto distribution. This analytical derivation will find use subsequently for the study of the statistical properties of the Chen-Pareto Distribution.
From the p.d.f of the Chen-Pareto distribution, Equation (6) can be re-written as

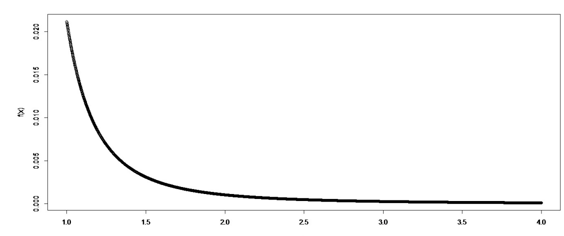
$$f(x) = W\mu\phi e^{\mu}\frac{\sigma k^{\sigma}}{x^{\sigma+1}}(1-(\frac{k}{x})^{\sigma})^{\phi-1}e^{(1-(\frac{k}{x})^{\sigma})}e^{-\mu e^{(1-(\frac{k}{x})^{\sigma})^{\phi}}} \tag{7}$$

Using Taylor series expansion

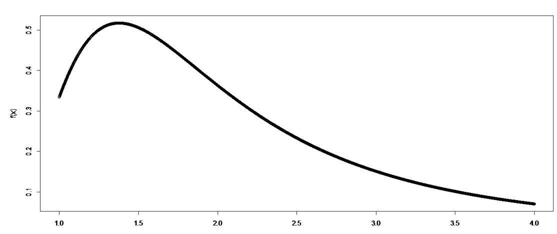
$$e^{x} = \sum_{n=0}^{\infty}\frac{x^{n}}{n!}$$



(a) k=0.5,σ=0.5,μ=2,ϕ=2

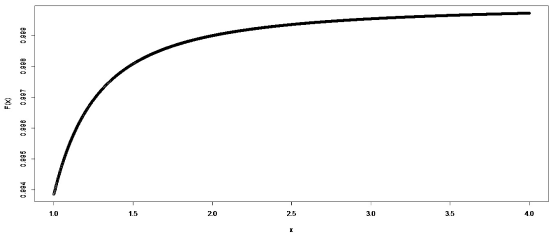


(b) k=0.5,σ=2,μ=0.5,ϕ=10

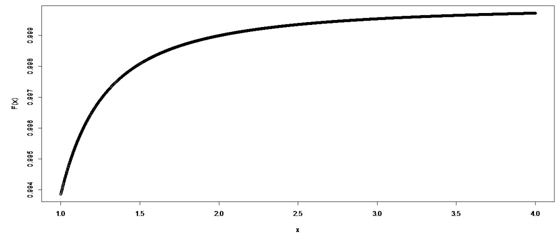


(c) k=0.5,σ=1.5,μ=4, ϕ=0.5

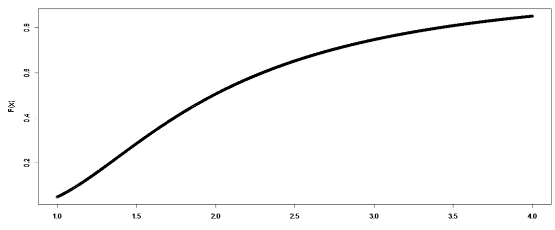
Figure 1: Graph of the p.d.f of Chen Pareto Distribution with various parameter values



(a) k=0.5,σ=0.5,μ=2,ϕ=2



(b) k=0.5,σ=2,μ=0.5,ϕ=10



(c) k=0.5,σ=1.5,μ=4, ϕ=0.5

Figure 2: Graph of the c.d.f of Chen-Pareto Distribution with various parameter values

Therefore,

$$e^{(1-(\frac{k}{x})^{\sigma})^{\phi}} = \sum_{l=0}^{\infty} \frac{(1-(\frac{k}{x})^{\sigma})^{l\phi}}{l!} \tag{8}$$

Let

$$q = e^{-\mu e^{(1-(\frac{k}{x})^{\sigma})^{\phi}}}$$

Let

$$a = e^{(1-(\frac{k}{x})^{\sigma})^{\phi}}$$

then

$$q = e^{-\mu a} \tag{9}$$

Using the Taylor expression in (9)

$$q = \sum_{m=0}^{\infty} \frac{(-\mu a)^m}{m!}$$

$$q = \sum_{m=0}^{\infty} \frac{(-1)^m \mu^m e^{(1-(\frac{k}{x})^{\sigma})^{\phi}}}{m!} \tag{10}$$

$$q = \sum_{m=0}^{\infty} \sum_{r=0}^{\infty} \frac{(-1)^m \mu^m m^r (1-(\frac{k}{x})^{\sigma})^{r\phi}}{m!r!} \tag{11}$$

Inserting (11) and (8) in (7), we have the mixture representation for the p.d.f as

$$f(x) = W\mu\phi e^{\mu} \frac{\sigma k^{\sigma}}{x^{\sigma+1}} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{r=0}^{\infty} \frac{(-1)^m \mu^m m^r (1-(\frac{k}{x})^{\sigma})^{\phi(r+l+1)-1}}{m!l!r!}$$

$$f(x) = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{r=0}^{\infty} W\mu\phi e^{\mu} \frac{(-1)^m \mu^m m^r}{m!l!r!} \frac{\sigma k^{\sigma}}{x^{\sigma+1}} (1-(\frac{k}{x})^{\sigma})^{\phi(r+l+1)-1}$$

$$f(x) = \sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty} e_k \frac{\sigma k^{\sigma}}{x^{\sigma+1}}(1-(\frac{k}{x})^{\sigma})^{\zeta-1} \tag{12}$$

where $\zeta = \phi(r+l+1)$ and

$$e_k = W\mu\phi e^{\mu}\frac{(-1)^m \mu^m m^r}{m!l!r!}$$

Equation (12) can be written in terms of Exp-G p.d.f as

$$f(x) = \sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty} j_k \Pi_{\zeta} \tag{13}$$

where

$$j_k = W\mu\phi e^{\mu}\frac{(-1)^m \mu^m m^r}{m!l!r!\zeta}$$

and

$$\Pi_{\zeta} = \zeta\frac{\sigma k^{\sigma}}{x^{\sigma+1}}(1-(\frac{k}{x})^{\sigma})^{\phi(r+l+1)-1}$$

Π_{ζ} is the p.d.f of the Exponentiated Pareto Distribution with $\phi(r+l+1)$ power parameter. Therefore the corresponding c.d.f for the mixture representation of the Chen-Pareto is derived by Integrating (13) to have

$$f(x) = \sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty} j_k \Theta_{\zeta}$$

where Θ_{ζ} is the c.d.f of the Exponentiated-Pareto Distribution with power parameter $\phi(r+l+1)$

4. Statistical Properties

In this section, we study some statistical properties of the Chen-Pareto distribution. This includes survival function, hazard function, moments, moment generating function, order statistics, and entropies.

4.1. Reliability Function

The survival function s(x) of a Chen Pareto distribution is given as

$$s(x) = 1 - W[1 - e^{\mu(1-e^{(1-(\frac{k}{x})^{\sigma})^{\phi}})}] \tag{14}$$

It has the hazard function h(x) as equation (15) and the reverse hazard $\bar{h}(x)$ as equation (16). Figure 3 shows the plots of the survival function and the hazard function with assigned parameter values.

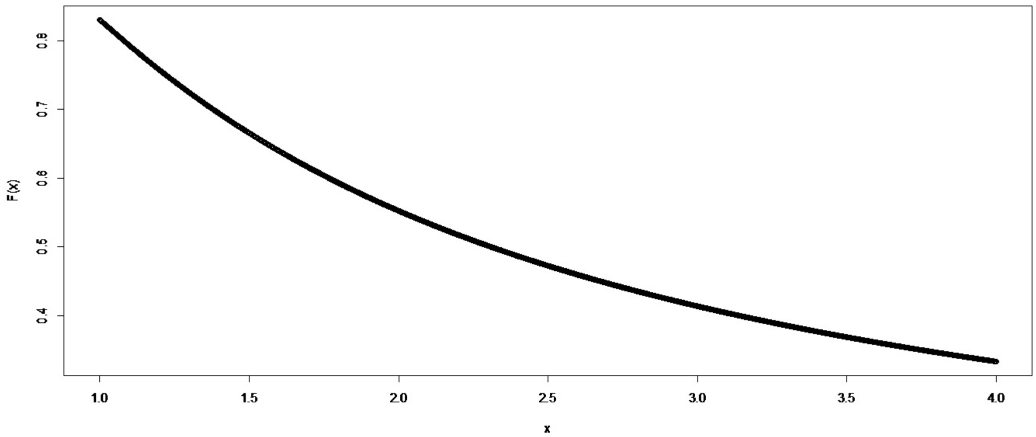
$$h(x) = \frac{W\mu\phi\sigma k^{\sigma}(1-(\frac{k}{x})^{\sigma})^{\phi-1}e^{(1-(\frac{k}{x})^{\sigma})}e^{\mu(1-e^{(1-(\frac{k}{x})^{\sigma})})^{\phi}}}{x^{\sigma+1}(1-W[1-e^{\mu(1-e^{(1-(\frac{k}{x})^{\sigma})^{\phi}})}])} \tag{15}$$

$$\bar{h}(z) = \frac{W\mu\phi\sigma k^{\sigma}(1-(\frac{k}{x})^{\sigma})^{\phi-1}e^{(1-(\frac{k}{x})^{\sigma})^{\phi}}e^{\mu(1-e^{(1-(\frac{k}{x})^{\sigma})})^{\phi}}}{x^{\sigma+1}W[1-e^{\mu(1-e^{(1-(\frac{k}{x})^{\sigma})^{\phi}})}} \tag{16}$$

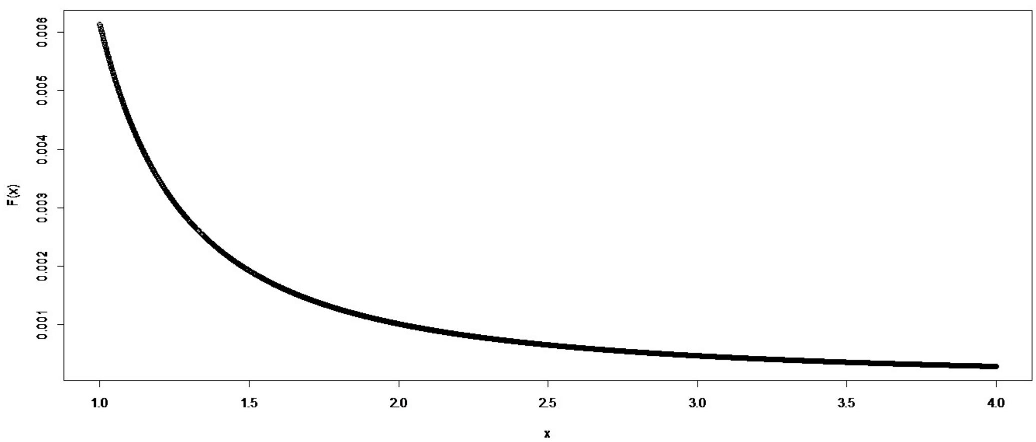
4.2. Quantile Function

The quantile function (x_u) of the Chen-Pareto distribution is derived by solving for x in equation 6 to have

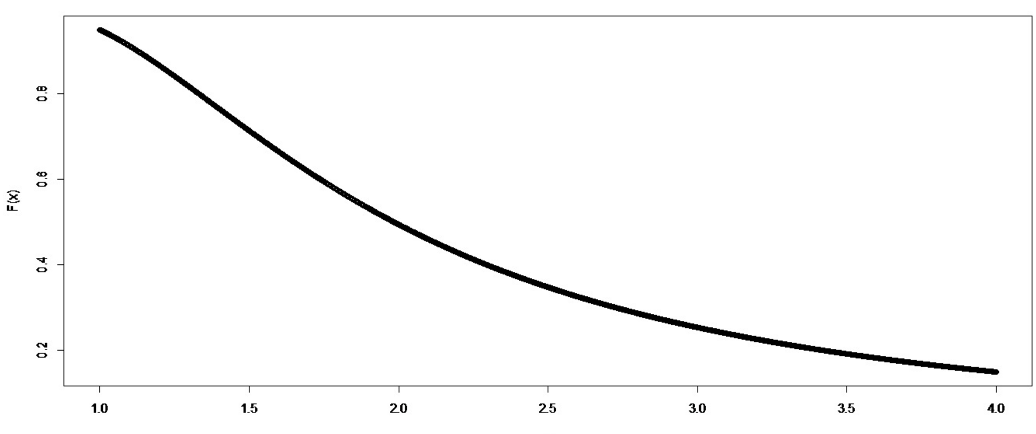
$$u = W[1 - e^{\mu(1-e^{(1-\frac{k}{x})^{\sigma}})^{\phi}}]$$



(a) k=0.5,σ=0.5,μ=2,ϕ=2

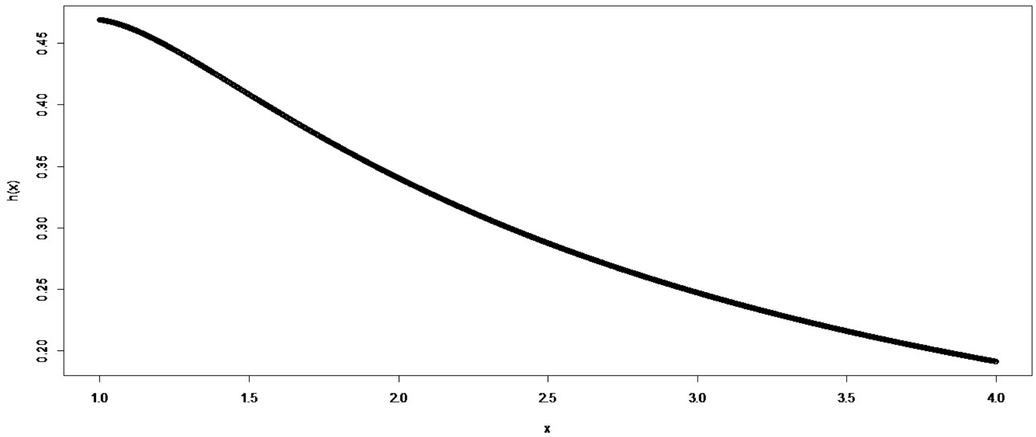


(b) k=0.5,σ=2,μ=0.5,ϕ=10

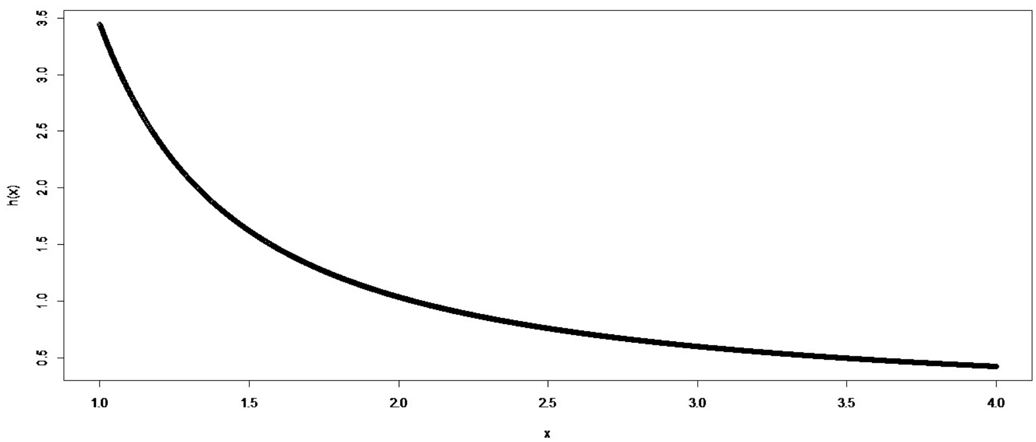


(c) k=0.5,σ=1.5,μ=4, ϕ=0.5

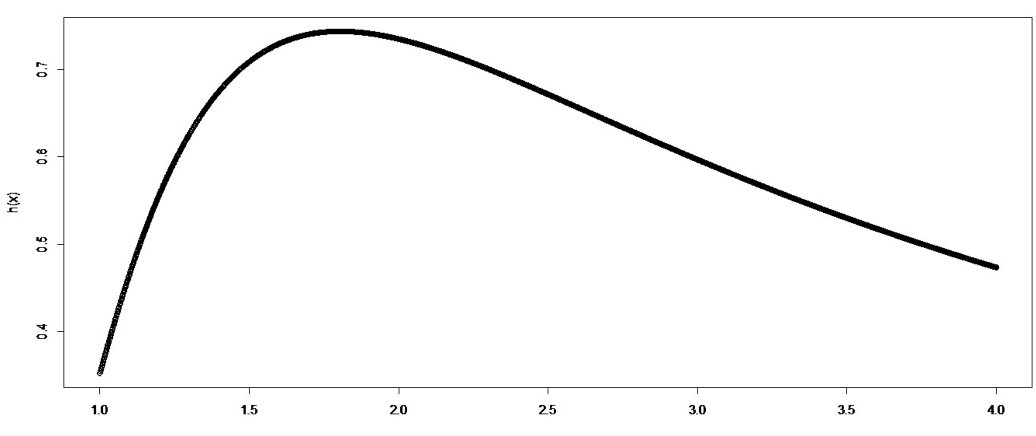
Figure 3: Graph of the survival function of Chen-Pareto Distribution with various parameter values



(a) k=0.5,σ=0.5,μ=2,ϕ=2



(b) k=0.5,σ=2,μ=0.5,ϕ=10



(c) k=0.5,σ=1.5,μ=4, ϕ=0.5

Figure 4: Graph of the hazard function of Chen Pareto Distribution with various parameter values

where u∈ Uniform (0,1). Making x the subject of the formula, we have

$$x_u = \frac{k}{(1-(\ln(1-\frac{1}{\mu}\ln(1-\frac{u}{W})))^{\frac{1}{\phi}})^{\frac{1}{\sigma}}} \tag{17}$$

To obtain the three quartiles, Q_1, Q_2, Q_3 we assign u= 0.25, 0.5 and 0.75 respectively. Equation (17) will also be useful for simulating data that is distributed to the Chen-Pareto distribution.

4.3. Moments

In this section, the moments of the Chen-Pareto is derived. The moments help to describe the distribution. Let $X_1, X_2, X_3, ..., X_n$ be distributed to a particular distribution, the r^{th} is defined as

$$E[X^r] = \int_0^\infty x^r f(x)\partial x$$

Therefore, the r^{th} moment of the Chen-Pareto is given as

$$E[X^r] = \int_0^\infty x^r \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{r=0}^\infty j_k \zeta \frac{\sigma k^\sigma}{x^{\sigma+1}}(1-(\frac{k}{x})^\sigma)^{\zeta-1}\partial x$$

$$E[X^r] = \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{r=0}^\infty j_k \int_0^\infty x^r \zeta \frac{\sigma k^\sigma}{x^{\sigma+1}}(1-(\frac{k}{x})^\sigma)^{\zeta-1}\partial x$$

$$E[X^r] = \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{r=0}^\infty j_k \zeta\sigma k^\sigma \int_0^\infty x^{r-\sigma-1}(1-(\frac{k}{x})^\sigma)^{\zeta-1}\partial x$$

$$(1-(\frac{k}{x})^\sigma)^{\zeta-1} = \sum_{f=0}^\infty \binom{\zeta}{f}(k^\sigma x^{-\sigma})^f \partial x$$

$$E[X^r] = \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{r=0}^\infty \sum_{f=0}^\infty j_k \zeta\sigma k^{\sigma(f+1)}\binom{\zeta}{f}\int_0^\infty x^{r-\sigma-\sigma f-1}\partial x$$

$$E[X^r] = \zeta\sigma \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{r=0}^\infty \sum_{f=0}^\infty j_k k^{\sigma(f+1)}\binom{\zeta}{f}\frac{\Gamma(\tau)}{\tau^{r-\sigma-\sigma f}}\int_0^\infty \frac{\tau^{r-\sigma-\sigma f}}{\Gamma(\tau)}x^{r-\sigma-\sigma f-1}\partial x$$

Finally,

$$E[X^r] = \zeta\sigma \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{r=0}^\infty \sum_{f=0}^\infty j_k k^{\sigma(f+1)}\binom{\zeta}{f}\frac{\Gamma(\tau)}{\tau^{r-\sigma-\sigma f}} \tag{18}$$

From equation (18), we can derive the mean (E[X]), second moment (E[X^2]), Variance ($E[X^2]-(E[X])^2$), Skewness, and Kurtosis.

4.4. Moment Generating Function

In this section, we derived the moment generating function of the Chen-Pareto. It is also a function that can help in the derivation of the moments of the distribution. The moment generating function of Chen Pareto distribution is obtained as

$$M_X(t) = E[e^{tx}] = \int_0^\infty e^{tx} f(x)\partial x$$

$$M_X(t) = \int_0^\infty e^{tx}\partial x$$

$$= \int_0^\infty \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{r=0}^\infty j_k \zeta \frac{\sigma k^\sigma}{x^{\sigma+1}} (1 - (\frac{k}{x})^\sigma)^{\zeta - 1} \partial x$$

$$= \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{r=0}^\infty \sum_{f=0}^\infty j_k \zeta \sigma k^{\sigma(f+1)} \binom{\zeta}{f} \int_0^\infty e^{tx} x^{-\sigma - \sigma f - 1} \partial x \tag{19}$$

Expanding e^{tx}

$$e^{tx} = \sum_{d=0}^\infty \frac{(tx)^d}{d!} \tag{20}$$

Inserting equation (20) in (19), we have

$$M_X(t) = \zeta \sigma \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{r=0}^\infty \sum_{f=0}^\infty \sum_{d=0}^\infty j_k \frac{t^d}{d!} k^{\sigma(f+1)} \binom{\zeta}{f} \int_0^\infty x^{t-\sigma-\sigma f - 1} \partial x$$

which finally gives

$$M_X(t) = \zeta \sigma \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{r=0}^\infty \sum_{f=0}^\infty \sum_{d=0}^\infty j_k \frac{t^d}{d!} k^{\sigma(f+1)} \binom{\zeta}{f} \frac{\Gamma(\tau)}{\tau^{t-\sigma-\sigma f}}$$

4.5. Order Statistics

If $X_{(1)}, X_{(2)}, X_{(3)}, ..., X_{(n)}$ is an order statistics taken from a randomly sampled values $X_1, X_2, X_3, ..., X_n$ with p.d.f f(x) and c.d.f F(x), then the p.d.f of the h^{th} samples is

$$f_{h:n}(x) = n \binom{n-1}{h-1} f(x) [F(x)]^{h-1} [1 - F(x)]^{n-h} \quad h = 1, 2., ..., n \tag{21}$$

Substituting equations (3) and (4) in (21), we have the order statistics of the Chen-Pareto distribution as

$$f_{h:n}(x) = n \binom{n-1}{h-1} \frac{W\mu\phi\sigma k^\sigma (1 - (\frac{k}{x})^\sigma)^{\phi-1} e^{(1-(\frac{k}{x})^\sigma)} e^{\mu(1-e^{(1-(\frac{k}{x})^\sigma)})^\phi}}{x^{\sigma+1}} \left(A[1 - e^{\mu(1-e^{(1-(\frac{k}{x})^\sigma)^\phi}})] \right)^{h-1}$$

$$\left(1 - W[1 - e^{\mu(1-e^{(1-(\frac{k}{x})^\sigma)^\phi}})] \right)^{n-h} \tag{22}$$

Therefore, from (22), the p.d.f of the first order, $X_{(1)}$ (i.e h=1) is

$$f_{1:n}(x) = n \frac{W\mu\phi\sigma k^\sigma (1 - (\frac{k}{x})^\sigma)^{\phi-1} e^{(1-(\frac{k}{x})^\sigma)} e^{\mu(1-e^{(1-(\frac{k}{x})^\sigma)})^\phi}}{x^{\sigma+1}} \left(1 - W[1 - e^{\mu(1-e^{(1-(\frac{k}{x})^\sigma)^\phi}})] \right)^{n-1} \tag{23}$$

and the p.d.f of the last order, $X_{(n)}$ is

$$f_{h:n}(x) = n \frac{W\mu\phi\sigma k^\sigma (1 - (\frac{k}{x})^\sigma)^{\phi-1} e^{(1-(\frac{k}{x})^\sigma)} e^{\mu(1-e^{(1-(\frac{k}{x})^\sigma)})^\phi}}{x^{\sigma+1}} \left(W[1 - e^{\mu(1-e^{(1-(\frac{k}{x})^\sigma)^\phi}})] \right)^{n-1}$$

4.6. Shanon Entropy

Entropy measures the uncertainty of a random variable X. Shannon(1948) defined the entropy of a random variable X. The entropy B for the Chen-Pareto distribution is given as

$$B = -E[log(f(x))] \tag{24}$$

$$B = -E[log(\frac{W\mu\phi\sigma k^{\sigma}(1-(\frac{k}{x})^{\sigma})^{\phi-1}e^{(1-(\frac{k}{x})^{\sigma})}e^{\mu(1-e^{(1-(\frac{k}{x})^{\sigma})})^{\phi}}}{x^{\sigma+1}})] \tag{25}$$

This can be estimated iteratively.

5. Parametric Estimation

In this section, the parametric estimation of the distribution is conducted. Let $X_1, X_2, ..., X_n$ be samples assumed to be distributed to the Chen Pareto distribution, the loglikelihood l is given as

$$l = n\ln W + n\ln\mu + n\ln\phi + n\ln\sigma + n\sigma\ln k - (\sigma+1)\sum_{i=0}^{n}\ln(x_i) + (\phi-1)\sum_{i=0}^{n}\ln(1-(\frac{k}{x_i})^{\sigma})$$

$$+\sum_{i=0}^{n}(1-(\frac{k}{x_i})^{\sigma})^{\phi} + \sum_{i=0}^{n}\mu(1-e^(1-(\frac{k}{x_i})^{\sigma})^{\phi}) \tag{26}$$

Differentiating l with respect to the distribution parameters, we have

$$\frac{\partial l}{\partial\phi} = \frac{n}{\phi} + \sum_{i=0}^{n}\ln(1-(\frac{k}{x_i})^{\sigma}) + \sum_{i=0}^{n}(1-(\frac{k}{x_i})^{\sigma})^{\phi}\ln 1-(\frac{k}{x_i})^{\sigma} + \sum_{i=1}^{n}\mu(e^{(1-(\frac{k}{x_i})^{\sigma})^{\phi}\ln 1-(\frac{k}{x_i})^{\sigma}}) \tag{27}$$

$$\frac{\partial l}{\partial\sigma} = \frac{n}{\sigma} + n\ln k - \sum_{i=1}^{n}\ln x_i - (\phi-1)\sum_{i=1}^{n}\frac{(\frac{k}{x_i})^{\sigma}\ln\frac{k}{x_i}}{1-(\frac{k}{x_i})^{\sigma}} + \phi\sum_{i=1}^{n}(1-(\frac{k}{x_i})^{\sigma})^{\phi-1}(\frac{k}{x_i})^{\sigma}\ln\frac{k}{x_i}$$

$$-\sum_{i=1}^{n}\mu\phi e^{(1-(\frac{k}{x_i})^{\sigma})^{\phi}}(1-(\frac{k}{x_i})^{\sigma})^{\phi-1}(\frac{k}{x_i})^{\sigma}\ln\frac{k}{x_i} \tag{28}$$

$$\frac{\partial l}{\partial\mu} = \frac{e^{\mu(1-e)}(1-e)}{(1-e^{\mu(1-e)})^2} + \frac{n}{\mu} + \sum_{i=1}^{n}1-e^{(1-(\frac{k}{x_i})^{\sigma})^{\phi}} \tag{29}$$

$$\frac{\partial l}{\partial k} = \frac{n\sigma}{k} - \sum_{i=1}^{n}\frac{\frac{\sigma k^{\sigma-1}}{x^{\sigma}}}{1-(\frac{k}{x_i})^{\sigma}} + \sum_{i=1}^{n}\phi(1-(\frac{k}{x_i})^{\sigma})^{\phi-1}\frac{\sigma k^{\sigma-1}}{x^{\sigma}} + \phi\sigma\mu\sum_{i=1}^{n}e^{(1-(\frac{k}{x_i})^{\sigma})^{\phi}}(1-(\frac{k}{x_i})^{\sigma})^{\phi-1}\frac{k^{\sigma-1}}{x^{\sigma}} \tag{30}$$

Setting $\frac{\partial l}{\partial\phi} = 0$, $\frac{\partial l}{\partial\mu} = 0$, $\frac{\partial l}{\partial\sigma} = 0$, $\frac{\partial l}{\partial b} = 0$ and solving the resulting equations will result in obtaining the estimates $\hat{\phi}$, $\hat{\mu}$, $\hat{\sigma}$, $\hat{k}$ respectively.
As n$|\infty$, we obtained the asymptotic distribution of the estimates obtained from equations (27)-(30) as

$$\hat{\theta} \in N(\theta, \alpha) \tag{31}$$

where $\theta= \begin{bmatrix}\phi, \sigma, \mu, k\end{bmatrix}'$ and $\alpha= \begin{bmatrix}\alpha_{\phi,\phi} & \alpha_{\phi,\sigma} & \alpha_{\phi,\mu} & \alpha_{\phi,k}\\ \alpha_{\sigma,\phi} & \alpha_{\sigma,\sigma} & \alpha_{\sigma,\mu} & \alpha_{\sigma,k}\\ \alpha_{\mu,\phi} & \alpha_{\mu,\sigma} & \alpha_{\mu,\mu} & \alpha_{\mu,k}\\ \alpha_{k,\phi} & \alpha_{k,\sigma} & \alpha_{k,\mu} & \alpha_{k,k}\end{bmatrix}$
α is the variance-covariance matrix of the estimates.An asymptotic confidence interval with significance level γ for each parameter θ_i is given by

$$ACI(\theta_i, 100(1-\gamma)) = \hat{\theta} - z_{\frac{\gamma}{2}}\sqrt{\alpha^{\hat{\theta},\hat{\theta}}}, \theta + z_{\frac{\gamma}{2}}\sqrt{\alpha^{\hat{\theta},\hat{\theta}}} \tag{32}$$

where $\alpha^{\hat{\theta},\hat{\theta}}$ is the i^th diagonal element of $K_n(\hat{\theta})^{-1}$ for $i = 1, 2, 3, 4$ and $z_{\gamma/2}$ is the quantile of the standard normal distribution. The expressions of the elements of α are in Appendix 1

6. Simulation Studies

In this section, we carried out the simulation study to assess the performance of the maximum likelihood estimates of the Chen- Pareto distribution. The simulation study was conducted by deriving a random sample of sizes 50, 100, 200,300 from the Chen Pareto distribution for different values of the parameters. For each sample size, the maximum likelihood estimators are obtained and the procedure is repeated 10,000 times. We have computed the Average estimate, bias, and MSE of parameter estimates for these 10,000 values, and the results are presented in Tables 1, 2, 3 and 4 . The results show that the Bias is small, that is the estimated values are close to the true values. Also, the MSE reduces as the sample sizes increase for each parameter. This shows that the estimation method is adequate and consistent.

Table 1: Table displaying simulation results for CP(k=1,σ=0.5,μ=1,ϕ=3)

Sample size	Parameter	AE	Bias	M.S.E
50	σ	2.993	1.494	4.129
	μ	0.578	-0.921	2.244
	ϕ	0.559	-0.941	3.600
100	σ	2.887	1.388	3.332
	μ	0.534	-0.965	2.215
	ϕ	0.798	-0.702	2.888
200	σ	2.913	1.414	3.314
	μ	0.516	-0.982	2.126
	ϕ	0.948	-0.552	2.183
300	σ	2.901	1.436	3.306
	μ	0.504	-0.994	2.015
	ϕ	1.038	-0.462	2.034

Table 2: Table displaying simulation results for CP(k=1,σ=1.5,μ=2,ϕ=2)

Sample size	Parameter	AE	Bias	M.S.E
50	σ	2.032	0.199	0.389
	μ	2.277	0.443	0.901
	ϕ	0.915	-0.918	2.434
100	σ	1.976	0.143	0.197
	μ	2.039	0.205	0.544
	ϕ	1.195	-0.638	1.823
200	σ	1.988	0.155	0.133
	μ	1.893	0.060	0.391
	ϕ	1.451	-0.382	1.156
300	σ	2.005	0.172	0.117
	μ	1.808	-0.024	0.323
	ϕ	1.583	-0.250	0.919

Table 3: Table displaying simulation results for CP(k=1,σ=3,μ=0.5,ϕ=0.5)

Sample size	Parameter	A.E	Bias	M.S.E
50	σ	0.509	-0.825	2.114
	μ	3.197	1.864	6.182
	ϕ	0.882	-0.450	3.539
100	σ	0.500	-0.834	2.110
	μ	3.024	1.692	4.896
	ϕ	0.538	-0.793	2.580
200	σ	0.499	-0.835	2.092
	μ	3.014	1.681	4.586
	ϕ	0.529	-0.802	2.299
300	σ	0.503	-0.831	2.086
	μ	2.992	1.660	4.425
	ϕ	0.533	-0.799	2.210

7. Application to Real life Data

In this section, we applied the Chen-Pareto distribution to a real dataset. Here, the Floyd River dataset was considered. The Floyd River is located in Iowa, U.S.A. which provides the consecutive annual flood discharge rates for the year 1935–1973.In literature, this dataset has been analyzed using different distribution, for example in Akinsete et al. (2008), Rahman et. al (2018), Mudholkar and Huston (1996) e.t.c. Table 5 gives the summary statistics of the dataset. The Pareto and Cubic Transmuted Pareto were fitted using this data. The M.L.E's and their respective -loglikelihood with corresponding Akaike Information Criterion, Bayesian Information Criterion are presented in Table 6. From the results on the table, using the comparison criterion, it is sufficient to say that the Chen-Pareto is the most appropriate for this data.

Table 4: Table displaying simulation results for CP(k=1,σ=1,μ=2.5,ϕ=1)

Sample size	Parameter	AE	Bias	M.S.E
50	σ	0.972	-0.527	0.823
	μ	1.887	0.386	1.332
	ϕ	1.258	-0.240	1.803
100	σ	0.964	-0.535	0.810
	μ	1.621	0.120	1.003
	ϕ	1.554	0.055	1.623
200	σ	0.976	-0.523	0.779
	μ	1.439	-0.061	0.856
	ϕ	1.847	0.347	1.585
300	σ	0.986	-0.512	0.767
	μ	1.346	-0.154	0.755
	ϕ	1.978	0.479	1.572

Table 5: Table displaying Descriptive analysis of survival time of Floyd River Data

Minimum	First Quartile	Median	Mean	Third Quartile	Maximum
318	1590	3570	6771	6725	71500

8. Conclusion

In statistics, generalization of distributions have attracted the interest of Statisticians in a bid to get better fits for real-life data. In this work, the Chen-Pareto distribution, which is a generalized distribution of the Pareto distribution, was studied. The Chen-Pareto distribution extends the pareto distribution by adding two shape parameters. The mixture representation of the distribution which enables us to study some statistical properties such as moments, moment generating function, order statistics, entropy, reliability functions, and quantile function was provided. The estimation of the parameters of the model has also been discussed. A simulation study to assess the adequacy of the estimation method and consistency of the parameters was conducted, and It has been shown, by means of a real data sets, that the Chen-Pareto distribution can provide better fits than the Pareto distribution.

Table 6: Table displaying results of the analysis of Floyd River dataset

Model	Parameters	Estimates	-L	AIC	BIC
Chen-P	σ	1.294	376.46	760.92	767.57
	μ	-2.272			
	ϕ	87.280			
	k	20.414			
Cubic Transmuted-P	k	318	381.190	768.379	773.370
	ϕ	0.71			
	ρ	-0.90			
Pareto	ϕ	318	392.810	787.620	789.283
	k	0.412			

Reference

1. Anzagra,L. , Sarpong,S., Suleman Nasiru (2020).Chen-G class of distributions, Cogent Mathematics and Statistics, 7:1, 1721401
2. Lee, S., Kim, J. H. (2018) Exponentiated generalized Pareto distribution: Properties and applications towards extreme value theory. Communications in Statistics-Theory and Methods.1–25.
3. Brazauskas, V., Serfling, R.(2003). Favorable estimators for fitting Pareto models: A study using goodness-of-fit measures with actual data. ASTIN Bulletin: The Journal of the IAA. 33(2):365–81.
4. Farshchian, M., Posner, F. L.(2010)The Pareto distribution for low grazing angle and high resolution X-band sea clutter. Naval Research Lab Washington DC. 5.Korkmaz, M., Altun, E., Yousof, H., Afify, A., Nadarajah, S. (2018). The Burr X Pareto Distribution: Properties, Applications and VaR Estimation. Journal of Risk and Financial Management.11(1):1
6. Alzaatreh, A., Famoye, F., Lee, C. (2013) Weibull-Pareto distribution and its applications. Communications in Statistics Theory and Methods.42(9):1673–1691.
7.Akinsete A, Famoye F, Lee C. (2008) The beta-Pareto distribution. Statistics. 42(6):547–63.
8. Pereira, M.B., Silva, R.B., Zea, L.M., Cordeiro, G.M. (2012) The kumaraswamy Pareto distribution. arXiv preprint arXiv:1204.1389.
9. Gupta, R.C., Gupta, P.L., Gupta, R.D. (1998) Modeling failure time data by Lehman alternatives. Communications in Statistics-Theory and methods.27(4):887–904.
10. Nadarajah, S. (2005) Exponentiated Pareto distributions. Statistics.39(3):255–60.
11. Mahmoudi, E.(2011) The beta generalized Pareto distribution with application to lifetime data. Mathematics and computers in Simulation.81(11):2414–2430.
12. Alzaatreh, A., Famoye, F., Lee, C.(2012) Gamma Pareto Distribution and Applications, Journal of Modern Applied Statistical Methods. 11(1):7
13. Faton, M., Llukan, P.(2014) Transmuted Pareto Distribution. ProbStat Forum. 7:1-11
14. Shannon, C.E. (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423
15. Akinsete, A., Famoye, Felix., Lee, Carl.(2008) The beta-Pareto distribution,Statistics,42:6,547- 563
16.Rahman, M., Al-Zahrani, B., Shahbaz, M.Q. (2018) Cubic Transmuted Pareto Distribution, Annals of Data Science,https://doi.org/10.1007/s40745-018-0178-8
17. Mudholkar, G.S., Huston, A.D.(1996) The exponentiated Weibull family: some properties and a flood data application. Commun Stat Theory Methods 23:1149–1171

Appendix Elements of the Variance-Covariance Matrix of α

$$\frac{\partial^2 l}{\partial \mu^2} = -\frac{n}{\mu^2}$$

$$\frac{\partial^2 l}{\partial \mu \partial k} = \frac{\partial^2 l}{\partial k \partial \mu} = \sum_{i=1}^{n} \frac{\phi\sigma(1-(\frac{k}{x})^\sigma)^{\phi-1} e^{(}1-(\frac{k}{x})^\sigma)^\phi k^{\sigma-1}}{x^\sigma}$$

$$\frac{\partial^2 l}{\partial \mu \partial \sigma} = \frac{\partial^2 l}{\partial \sigma \partial \mu} = -\sum_{i=1}^{n} \phi e^{(}1-(\frac{k}{x})^\sigma)^\phi (1-(\frac{k}{x})^\sigma)^{(}\phi-1)(\frac{k}{x})^\sigma \ln(\frac{k}{x})$$

$$\frac{\partial^2 l}{\partial \mu \partial \phi} = \frac{\partial^2 l}{\partial \phi \partial \mu} = \sum_{i=1}^{n} e^{(1-(\frac{k}{x})^\sigma)^\phi \ln(1-(\frac{k}{x})^\sigma)}$$

$$\frac{\partial^2 l}{\partial \mu \partial \phi} = \frac{\partial^2 l}{\partial \phi \partial \mu} = \sum_{i=1}^{n} \frac{(\frac{k}{x})^\sigma \ln(\frac{k}{x})}{1-(\frac{k}{x})} + \sum_{i=1}^{n} (\frac{k}{x})^\sigma \ln(\frac{k}{x})(1-(\frac{k}{x})^\sigma)^{\phi-1}(1+\phi\ln(1-(\frac{k}{x})^\sigma))$$

$$+\mu \sum_{i=1}^{n} (\frac{k}{x})^\sigma \ln(\frac{k}{x})(1-(\frac{k}{x})^\sigma)^{\phi-1}(1+\phi\ln(1-(\frac{k}{x})^\sigma)) e^{(1-(\frac{k}{x})^\sigma)^\phi \ln(1-(\frac{k}{x})^\sigma)}$$

$$\frac{\partial^2 l}{\partial k \partial \phi} = \frac{\partial^2 l}{\partial \phi \partial k} = -\sum_{i=1}^{n} \frac{\sigma k^{\sigma-1}}{(1-(\frac{k}{x})^\sigma)x^\sigma} - \sum_{i=1}^{n} \frac{\sigma k^{\sigma-1}}{x^\sigma}(1+\phi\ln(1-(\frac{k}{x})^\sigma))$$

$$-\sum_{i=1}^{n}\frac{\sigma k^{\sigma-1}}{x^{\sigma}}(1-(\frac{k}{x})^{\sigma})^{\phi-1}(1+\phi\ln(1-(\frac{k}{x})^{\sigma}))e^{(1-(\frac{k}{x})^{\sigma})^{\phi}\ln(1-(\frac{k}{x})^{\sigma})}$$

$$\frac{\partial^2 l}{\partial\phi^2}=-\frac{n}{\phi^2}+\sum_{i=1}^{n}(1-(\frac{k}{x})^{\sigma})^{\phi}(\ln(1-(\frac{k}{x})^{\sigma}))^2+\sum_{i=1}^{n}\mu(1-(\frac{k}{x})^{\sigma})^{\phi}(\ln(1-(\frac{k}{x})^{\sigma}))^2$$