 DOI: https://doi.org/10.18187/pjsor.v21i3.4950

Pakistan Journal of Statistics and Operation Research

A New Odd-Burr Pareto Distribution: Statistical Properties, Estimation, and Applications

Ibrahim Arik[1], Yeliz Mert Kantar[2,*]

* Corresponding Author

[1]Kütahya Dumlupınar University, Faculty of Economics and Administrative Sciences / Business Administration, Kütahya, Türkiye, ibrahim.arik@dpu.edu.tr
[2]Eskisehir Technical University, Faculty of Science, Department of Statistics, Eskisehir, Türkiye, ymert@eskisehir.edu.tr

Abstract

This study introduces the Odd-Burr Pareto (*OBu-P*) distribution as a novel and flexible model, which is developed by combining the Burr and Pareto distributions using the T-X generator approach (Alizadeh et al. 2017). The *OBu-P* distribution can be used to model different data types characterized by heavy tails.

Statistical properties, including moments, incomplete moments, quantile functions, and limiting behavior, generating functions and order statistics of the *OBu-P* distribution are also presented. The parameters of the *OBu-P* distribution are estimated efficiently using the maximum likelihood method. The flexibility of the distribution is shown in a real-life example compared to its alternatives.

Key Words: Burr distribution; Pareto distribution; Odd Burr-G family; Characterizations; Simulation study.

Mathematical Subject Classification: 62E15

1. Introduction

Statistical distributions play a vital role in modelling events observed in nature (Johnson et al. 1994). Among these, the Pareto distribution is particularly useful for its ability to capture heavy-tailed data, like rainfall, wind speed, and financial losses (Arnold 2008). However, the limited flexibility of the Pareto distribution has led to various generalizations, such as the exponentiated Pareto (Nadarajah 2005), beta-Pareto (Akinsete et al. 2008), Kumaraswamy-Pareto (Bourguignon et al. 2013) and the Odd-Burr Pareto distribution (Arik 2018). For example, While Nadarajah (2005) develops the exponentiated Pareto distribution and analyses its properties, Akinsete et al. (2008) propose the beta-Pareto distribution as another generalization. Other developments include the gamma-Pareto distribution derived from the T-X family of distributions (Alzaatreh et al. 2012, Alzaatreh et al. 2013). On the other hand, Mahmoudi (2011), Zea et al. (2012), Bourguignon et al. (2013), Aljarrah et al. (2015) and Tahir et al. (2016) introduce beta-generalized Pareto distribution, beta-exponentiated Pareto distribution, the Kumaraswamy Pareto distribution, Weibull-Pareto distribution and different Weibull-Pareto distribution (approach of Bourguignon et al. 2014). Furthermore, the Pareto-Rayleigh distribution, a member of the T-T, is described by Godase et al. (2017). Considering the latest studies on generalizations of the Pareto distribution, several models have been recently presented, such as the Alpha-Power Pareto distribution (Ihtisham et al. 2019) and a new three-parameter lifetime distribution that is an extension of the classical Pareto distribution (Aniyan and George, 2023) are presented. More recently, the class of the Weibull-Pareto distribution (Rashid 2024), the Alpha Power Exponentiated Pareto (Pimsap et al. 2024), and the odd-inverse Pareto-Burr XII (Olmos et al. 2024) are introduced for modelling positive data. These generalizations and developments enhance the Pareto family's capacity for modelling complex phenomena, including income inequality, catastrophic losses, and extreme weather events (Arik, 2018).

Pareto random variable with cumulative distribution function (cdf) and probability density function (pdf), are provided as:

$$G(x) = 1 - \left(\frac{x}{\theta}\right)^{-\alpha}, x \geq \theta > 0, \alpha > 0 \quad (1)$$

$$g(x) = \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}}, \quad (2)$$

where θ and α are the scale and shape parameters, respectively.

In this study, we propose a new and highly flexible distribution, the Odd Burr-Pareto (*OBu-P*) distribution (Arik 2018, Arik and Kantar 2019), derived from the T-X generator approach introduced by Alizadeh et al. (2017). This new family of continuous distributions includes two additional shape parameters that increase flexibility.

The *OBu-P* distribution can be extensively used in extreme value analysis, reliability studies, survival analysis, and engineering. Its key statistical characteristics are thoroughly explored with examinations of skewness-kurtosis relationships, density and hazard rate variations under different shape parameters (Arik, 2018).

Alizadeh et al. (2017) present the Odd Burr-G family, developed using the T-X generator approach, which incorporates two extra shape parameters to extend the range of continuous distributions. The T–X generator is a general framework for constructing new distribution families from a baseline cdf $G(x)$. This approach adds shape parameters that can change the skewness and tails of the distribution, making the model more flexible. cdf of the Odd Burr-G family is formulated as follows:

$$F(x) = \int_0^{\frac{G(x)}{1-G(x)}} \frac{abt^{a-1}}{(1+t^a)^{b+1}} dt = 1 - \frac{[1-G(x)]^{ab}}{\{G(x)^a + [1-G(x)]^a\}^b}, x > 0. \quad (3)$$

Taking the derivative of equation (3) yields the corresponding pdf in the following form:

$$f(x) = \frac{abg(x)G(x)^{a-1}[1-G(x)]^{ab-1}}{\{G(x)^a + [1-G(x)]^a\}^{b+1}}, x > 0, \quad (4)$$

which $a, b > 0$, additional shape parameters.

Proof (obtaining the pdf):

To obtain the pdf, i.e., the derivative of the cdf, we perform the following transformations:

$$u = G(x), M = u^a, N = (1-u)^a,$$

$$Q(u) = \frac{(1-u)^{ab}}{(M+N)^b} = (1-u)^{ab}(M+N)^{-b}$$

Hence, the cdf can be rewritten as:

$$F(x) = 1 - Q(u)$$

The derivative can be obtained as follows:

$$\frac{dF}{dx} = -\frac{dQ}{dx} = -\frac{dQ}{du}\frac{du}{dx}, \qquad where\ \frac{du}{dx} = G'(x) = g(x)$$

Using the product rule for differentiation:

$$\frac{dQ}{du} = ab(1-u)^{ab-1}(-1)(M+N)^{-b} + (1-u)^{ab}(-b)(M+N)^{-b-1}(M'+N')$$

$$\frac{dQ}{du} = -b(1-u)^{ab-1}(M+N)^{-b-1}[a(M+N) + (1-u)(M'+N')]$$

$M' = au^{a-1}$ and $N' = -a(1-u)^{a-1}$ so,

$$(1-u)(M'+N') = a(1-u)u^{a-1} - a(1-u)^a$$

$$[a(M+N) + (1-u)(M'+N')] = a[u^a + (1-u)^a] + a(1-u)u^{a-1} - a(1-u)^a = au^{a-1}$$

As a result, the derivative is obtained as follows,

$$\frac{dQ}{du} = -b(1-u)^{ab-1}(M+N)^{-b-1}au^{a-1} = -\frac{abu^{a-1}(1-u)^{ab-1}}{(M+N)^{b+1}}$$

Thus, by the chain rule

$$\frac{dF}{dx} = -\frac{dQ}{dx} = -\frac{dQ}{du}\frac{du}{dx} = \frac{abu^{a-1}(1-u)^{ab-1}}{(M+N)^{b+1}}g(x)$$

Finally, applying the reverse transformation, the pdf is obtained as:

$$F'(x) = f(x) = \frac{abg(x)G(x)^{a-1}[1-G(x)]^{ab-1}}{\{G(x)^a + [1-G(x)]^a\}^{b+1}}.$$

The remainder of the paper is organized as follows: Section 2 introduces *OBu-P*, and some properties; pdf, cdf, hazard rate function, limiting behaviours, moments, incomplete moments, generating functions, and order statistics. Section 3 provides the maximum likelihood estimation (MLE) method for *OBu-P*. In Section 4, a simulation study is conducted to evaluate the performance of the MLE. Section 5 applies the *OBu-P* distribution to a real-world dataset, demonstrating its flexibility and superior fit. Finally, Section 6 concludes the study with a summary of findings and recommendations for future research.

2. Odd-Burr Pareto Distribution (*OBu-P*)

Substituting equation (1) into equation (3), cdf of the *OBu-P* is derived as follows (Arik, 2018; Arik and Kantar 2019):

$$F(x) = 1 - \frac{\left(\frac{x}{\theta}\right)^{-\alpha ab}}{\left\{\left[1-\left(\frac{x}{\theta}\right)^{-\alpha}\right]^a + \left(\frac{x}{\theta}\right)^{-\alpha a}\right\}^b}, a, b, \alpha > 0, x \geq \theta > 0 \tag{5}$$

The pdf corresponding to (5) is,

$$f(x) = \frac{ab\alpha\theta^{\alpha ab}x^{-(\alpha ab+1)}\left[1-\left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a-1}}{\left\{\left[1-\left(\frac{x}{\theta}\right)^{-\alpha}\right]^a + \left(\frac{x}{\theta}\right)^{-\alpha a}\right\}^{b+1}} \tag{6}$$

Here, θ represents the scale parameter, while a, b, and α denote the shape parameters of *OBu-P*. A random variable X following the pdf in equation (6) is expressed as $X \sim OBu - P(a, b, \alpha, \theta)$. Figures 1-3 illustrate the pdf of the *OBu-P* for various parameter combinations. These graphs reveal the flexibility of the *OBu-P* distribution, showcasing diverse shapes, distinct tail behaviours, and varying degrees of skewness and kurtosis depending on the shape parameters.

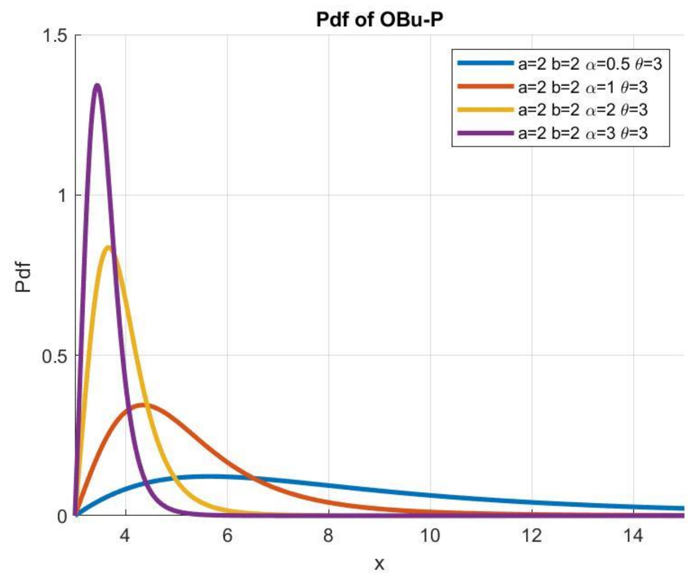

Figure 1. Plots of the *OBu-P* pdf for fixed values of a, b *and* θ, with varying values of α

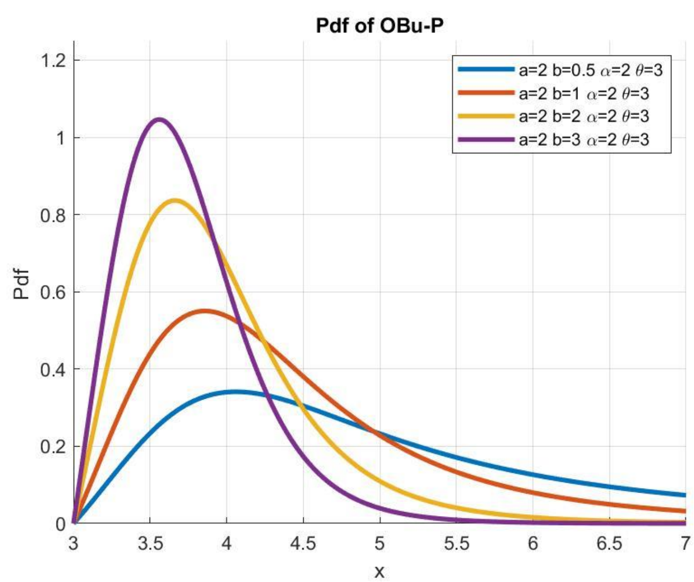

Figure 2. Plots of the *OBu-P* pdf for fixed values of a, α and θ, with varying values of b

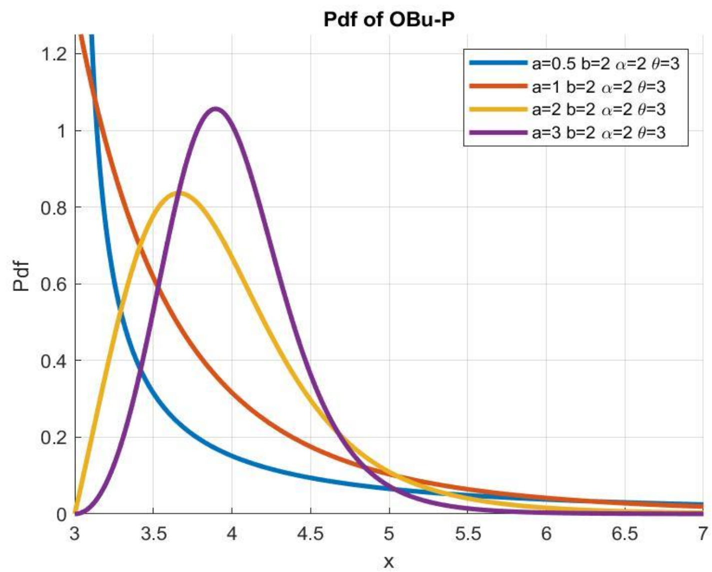

Figure 3. Plots of the *OBu-P* pdf for fixed values of b, α and θ, with varying values of a

Special Cases

a) For a=b=1 the *OBu-P* distribution reduces to the Pareto distribution.

b) When a=1, $O\ Bu-P(a,b,\alpha,\theta)$ reduces to the Pareto distribution with parameters $k=b\alpha$ and θ and its pdf is,

$$g(x)=\frac{k\theta^k}{x^{k+1}}, k>0, x\geq\theta>0.$$

c) If a random variable $Y \sim BurrXII(a,b)$, then the random variable,

$$X=\theta(Y+1)^{\frac{1}{\alpha}}$$

follows the $OBu-P(a,b,\alpha,\theta)$ distribution.

Proof of c: Y can be obtained as follows:

$$Y=\left(\frac{X}{\theta}\right)^{\alpha}-1$$

Using transformation method (Bhatt, 2023), we have

$$F_Y(y)=P(Y\leq y)=P\left(\left(\frac{X}{\theta}\right)^{\alpha}-1\leq y\right)=P\left(X\leq\theta(y+1)^{\frac{1}{\alpha}}\right)=F_X\left(\theta(y+1)^{\frac{1}{\alpha}}\right)$$

$$F_Y(y)=F_X\left(\theta(y+1)^{\frac{1}{\alpha}}\right)=1-\frac{\left(\frac{\theta(y+1)^{\frac{1}{\alpha}}}{\theta}\right)^{-\alpha ab}}{\left\{\left[1-\left(\frac{\theta(y+1)^{\frac{1}{\alpha}}}{\theta}\right)^{-\alpha}\right]^{a}+\left(\frac{\theta(y+1)^{\frac{1}{\alpha}}}{\theta}\right)^{-\alpha a}\right\}^{b}}$$

$$F_Y(y)=F_X\left(\theta(y+1)^{\frac{1}{\alpha}}\right)=1-\frac{(y+1)^{-ab}}{\{[1-(y+1)^{-1}]^{a}+(y+1)^{-a}\}^{b}}$$

$$F_Y(y) = F_X\left(\theta(y+1)^{\frac{1}{\alpha}}\right) = 1 - \left(\frac{(y+1)^{-a}}{\frac{y^a+1}{(y+1)^a}}\right)^b = 1-(1+y^a)^{-b}$$

$$F_Y(y) = 1-(1+y^a)^{-b}$$

The latest equation obtained is the cdf of Burr XII distribution.

d) If a random variable Y follows the Lomax distribution with the shape parameter *b* and scale parameter 1, then the random variable,

$$X = \theta\left(Y^{\frac{1}{a}}+1\right)^{\frac{1}{\alpha}},$$

follows the $OBu-P(a,b,\alpha,\theta)$ distribution.

Proof of d: Y can be obtained as follows:

$$Y = \left[\left(\frac{X}{\theta}\right)^{\alpha} - 1\right]^{a}$$

Using transformation method (Bhatt, 2023):

$$F_Y(y) = P(Y \le y) = P\left(\left[\left(\frac{X}{\theta}\right)^{\alpha} - 1\right]^{a} \le y\right) = P\left(X \le \theta\left(y^{\frac{1}{a}}+1\right)^{\frac{1}{\alpha}}\right) = F_X\left(\theta\left(y^{\frac{1}{a}}+1\right)^{\frac{1}{\alpha}}\right)$$

$$F_Y(y) = F_X\left(\theta(Y+1)^{\frac{1}{\alpha}}\right) = 1 - \frac{\left(\frac{\theta\left(y^{\frac{1}{a}}+1\right)^{\frac{1}{\alpha}}}{\theta}\right)^{-\alpha ab}}{\left\{\left[1-\left(\frac{\theta\left(y^{\frac{1}{a}}+1\right)^{\frac{1}{\alpha}}}{\theta}\right)^{-\alpha}\right]^{a} + \left(\frac{\theta\left(y^{\frac{1}{a}}+1\right)^{\frac{1}{\alpha}}}{\theta}\right)^{-\alpha a}\right\}^{b}}$$

$$F_Y(y) = F_X\left(\theta(y+1)^{\frac{1}{\alpha}}\right) = 1 - \frac{\left(y^{\frac{1}{a}}+1\right)^{-ab}}{\left\{\left[1-\left(y^{\frac{1}{a}}+1\right)^{-1}\right]^{a} + \left(y^{\frac{1}{a}}+1\right)^{-a}\right\}^{b}}$$

$$F_Y(y) = F_X\left(\theta(y+1)^{\frac{1}{\alpha}}\right) = 1 - \frac{\left(y^{\frac{1}{a}}+1\right)^{-ab}}{\frac{(y+1)^b}{(y^{\frac{1}{a}}+1)^{ab}}} = 1 - \frac{1}{(y+1)^b}$$

$$F_Y(y) = 1-(y+1)^{-b}$$

The final equation obtained is the cdf of the Lomax distribution with the scale parameter 1.

Shape Properties

In this section the shape characteristics of the *OBu-P* pdf is examined. The first derivative of (6) is,

$$\frac{\partial f(x)}{\partial x} = -ab\alpha\theta^{\alpha ab}x^{-(\alpha ab+1)-1}\left(\frac{x}{\theta}\right)^{\alpha(a+1)}\left[1-\left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a}$$
$$\left\{\left(\frac{x}{\theta}\right)^{\alpha a}\left[1-\left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a}\left[(\alpha ab+1)\left(\frac{x}{\theta}\right)^{\alpha}+\alpha-1\right]+(1-\alpha a)\left(\frac{x}{\theta}\right)^{\alpha}+\alpha-1\right\}$$
$$\left[\left(\frac{x}{\theta}\right)^{\alpha}-1\right]^{-2}\left\{\left[1-\left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a}+\left(\frac{x}{\theta}\right)^{-\alpha a}\right\}^{-b}\left\{\left(\frac{x}{\theta}\right)^{\alpha a}\left[1-\left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a}+1\right\}^{-2} \tag{7}$$

Equation (7) can yield multiple solutions. At any solution point x_0, the nature of the point—whether it is a local maximum, local minimum, or an inflection point—depends on the sign of the second derivative: if $f''(x_0) < 0$ it is a maximum, if $f''(x_0) > 0$ it is a minimum, and if $f''(x_0) = 0$ it represents an inflection point (Arik, 2018).

Lifetime Characteristics

For the *OBu-P* random variable, $S(x)$ survival function and $h(x)$ hazard rate function (*Hrf*) are defined as follows:

$$S(x) = \frac{\left(\frac{x}{\theta}\right)^{-\alpha ab}}{\left\{\left[1-\left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a}+\left(\frac{x}{\theta}\right)^{-\alpha a}\right\}^{b}} \tag{8}$$

$$h(x) = \frac{ab\alpha\left[1-\left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a-1}}{x\left\{\left[1-\left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a}+\left(\frac{x}{\theta}\right)^{-\alpha a}\right\}} \tag{9}$$

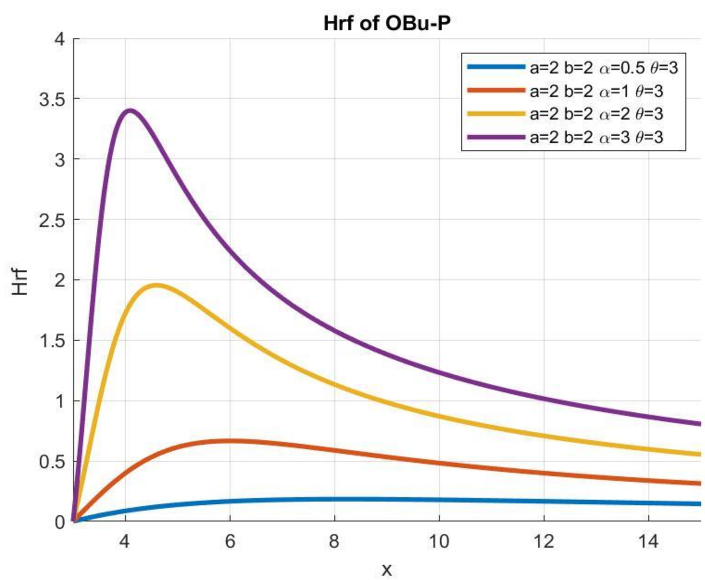

Figure 4. *OBu-P*'s *Hrf* function plots corresponding to different parameter values

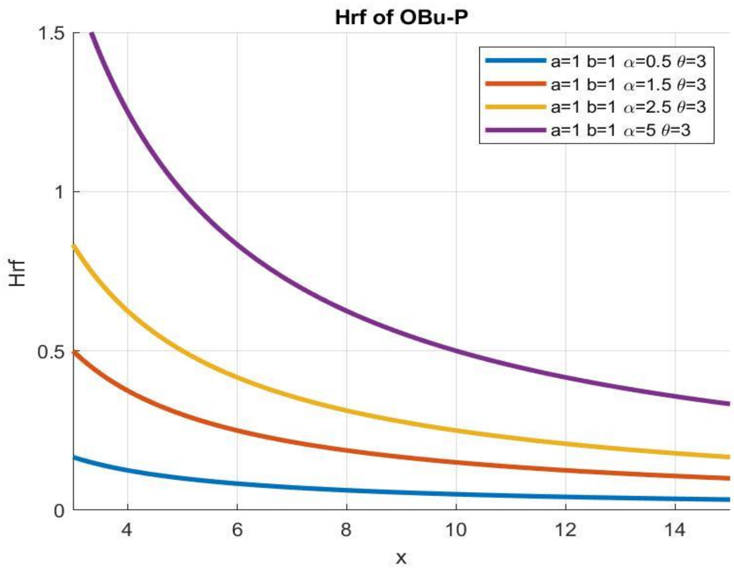

Figure 5. *OBu-P*'s *Hrf* function plots corresponding to different parameter values

Figures 4 and 5 present *Hrf* for various parameter values. These graphs reveal that *Hrf* of the *OBu-P* distribution exhibits diverse shapes, including monotonically decreasing and increasing-decreasing patterns, depending on the parameter configurations. This flexibility makes the *OBu-P* distribution particularly well-suited for analysing lifetime data (Arik, 2018).

Limit behaviour

The limit behaviour of *OBu-P* are provided by using Lemmas 2.1 and 2.2:

The asymptotic behaviour of the pdf and *Hrf* is described by the following lemmas

Lemma 2.1: The limit of the *OBu-P* density as $x \to \infty$ is 0 and the limit as $x \to \theta$ is given as follows:

$$\lim_{x\to\theta} f(x) = \begin{cases} \infty, when\ 0 < a < 1 \\ \dfrac{ab\alpha}{\theta}, when\ a = 1 \\ 0, when\ a > 1 \end{cases}$$

Proof: It is easy to show the above equation from the *OBu-P* density in equation (6).

Lemma 2.2: The limit of the *OBu-P* hazard function as $x \to \infty$ is 0 and the limit as $x \to \theta$ is given as follows:

$$\lim_{x\to\theta} h(x) = \begin{cases} \infty, when\ 0 < a < 1 \\ \dfrac{ab\alpha}{\theta}, when\ a = 1 \\ 0, when\ a > 1 \end{cases}$$

Proof: It is easy to obtain the results of Lemma 2.2 with taking the limit of the *OBu-P* hazard function in equation (9).

Quantile function

To obtain the quantile function, the *cdf* must be solved according to x:

$$F(x) = 1 - \frac{\left(\frac{x}{\theta}\right)^{-\alpha a b}}{\left\{\left[1 - \left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a} + \left(\frac{x}{\theta}\right)^{-\alpha a}\right\}^{b}}$$

$$[1 - F(x)]^{-\frac{1}{b}} = \frac{\left\{\left[1 - \left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a} + \left(\frac{x}{\theta}\right)^{-\alpha a}\right\}}{\left(\frac{x}{\theta}\right)^{-\alpha a}} = \frac{\left[1 - \left(\frac{x}{\theta}\right)^{-\alpha}\right]^{a}}{\left(\frac{x}{\theta}\right)^{-\alpha a}} + 1$$

$$\left([1 - F(x)]^{-\frac{1}{b}} - 1\right)^{\frac{1}{a}} = \frac{1}{\left(\frac{x}{\theta}\right)^{-\alpha}} - 1$$

$$\left([1 - F(x)]^{-\frac{1}{b}} - 1\right)^{\frac{1}{a}} + 1 = \left(\frac{x}{\theta}\right)^{\alpha}$$

$$X = \theta\left\{\left([1 - U]^{-\frac{1}{b}} - 1\right)^{\frac{1}{a}} + 1\right\}^{\frac{1}{\alpha}}, U \sim Uniform(0,1) \tag{10}$$

OBu-P data is generated by using (10).

Alternative Mathematical Representations of the *OBu-P* Distribution

This part of the study explores alternative mathematical representations of the *OBu-P* distribution using power series expansions. Beginning with the cdf given by Altun et al. (2017), these expansions enable the pdf and cdf to be reformulated in ways that simplify analysis, facilitate computation, and provide additional perspectives on how the distribution behaves under different parameter configurations.

$$F(x) = 1 - \frac{[1 - G(x)]^{ab}}{\{G(x)^a + [1 - G(x)]^a\}^b} = 1 - \left\{1 - \frac{G(x)^a}{G(x)^a + \bar{G}(x)^a}\right\}^b$$

By expanding the second term using the binomial series,

$$(1 - z)^b = \sum_{i=0}^{\infty} (-1)^i \binom{b}{i} z^i \quad | z | < 1$$

where $z = \frac{G(x)^a}{G(x)^a + \bar{G}(x)^a}$ and $\bar{G}(x) = 1 - G(x)$, thus:

$$F(x) = 1 - \sum_{i=0}^{\infty} (-1)^i \binom{b}{i} \frac{G(x)^{ai}}{[G(x)^a + \bar{G}(x)^a]^i} \tag{11}$$

Thus, the following alternative representations for the *OBu-P* can be derived through various mathematical expansions.

$$G(x)^{ai} = \{1 - [1 - G(x)]\}^{ai} = \sum_{j=0}^{\infty} (-1)^j \binom{ai}{j} [1 - G(x)]^j$$

$$[1 - G(x)]^j = \sum_{k=0}^{\infty} (-1)^k \binom{j}{k} G(x)^k$$

$$G(x)^{ai} = \sum_{k=0}^{\infty} \alpha_k G(x)^k \tag{12}$$

where $\alpha_k = \sum_{j=k}^{\infty}(-1)^{k+j}\binom{ai}{j}\binom{j}{k}$.

In eq. (11) can be expressed as follows:

$$G(x)^a + \bar{G}(x)^a = \sum_{k=0}^{\infty}\eta_k G(x)^k + \sum_{k=0}^{\infty}(-1)^k\binom{a}{k}G(x)^k$$
$$= \sum_{k=0}^{\infty}\left\{\eta_k + (-1)^k\binom{a}{k}\right\}G(x)^k = \sum_{k=0}^{\infty}t_k G(x)^k,$$

where $\eta_k = \sum_{j=k}^{\infty}(-1)^{k+j}\binom{a}{j}\binom{j}{k}$ and $t_k = \eta_k + (-1)^k\binom{a}{k}$.

We apply a formula from Gradshteyn & Ryzhik (2002) for expanding a power series raised to a positive integer power, which allows us to express the series in a more manageable form for further analysis and computations,

$$\left(\sum_{k=0}^{\infty} t_k x^k\right)^i = \sum_{k=0}^{\infty} c_{i,k} x^k \tag{13}$$

where $c_{i,k} = (kt_0)^{-1}\sum_{m=1}^{k}[m(i+1)-k]t_m c_{i,k-m}$, $k \geq 1$ and $c_{i,0} = t_0^i$. Then,

$$[G(x)^a + \bar{G}(x)^a]^i = \left[\sum_{k=0}^{\infty} t_k G(x)^k\right]^i = \sum_{k=0}^{\infty} c_{i,k} G(x)^k \tag{14}$$

Thus,

$$\frac{G(x)^{ai}}{[G(x)^a + \bar{G}(x)^a]^i} = \frac{\sum_{k=0}^{\infty}\alpha_k G(x)^k}{\sum_{k=0}^{\infty}\beta_k G(x)^k} \tag{15}$$

where $\beta_k = c_{i,k}$. Based on the ratio of two power series (Gradshteyn & Ryzhik, 2002), it can be expressed as follows:

$$\frac{G(x)^{ai}}{[G(x)^a + \bar{G}(x)^a]^i} = \sum_{k=0}^{\infty}\lambda_k G(x)^k \tag{16}$$

where $\lambda_k = \beta_0^{-1}[\alpha_k - \beta_0^{-1}\sum_{r=1}^{k}\beta_r\lambda_{k-r}]$ for $k > 0$ and $\lambda_0 = \frac{\alpha_0}{\beta_0}$. Finally, the *OBu-P* cdf is presented as:

$$F(x) = 1 - \sum_{k,i=0}^{\infty}(-1)^i\binom{b}{i}\lambda_k G(x)^k$$
$$= 1 - \sum_{k=0}^{\infty} m_k G(x)^k$$
$$= \sum_{k=0}^{\infty} n_k G(x)^k = \sum_{k=0}^{\infty} n_k H_k(x) \tag{17}$$

and

$$f(x) = \sum_{k=0}^{\infty} n_{k+1} h_{k+1}(x) \tag{18}$$

where $H_k(x)$ is the exponentiated Pareto (*EP*) cdf with parameter θ, α and k, $m_k = \sum_{i=0}^{\infty}(-1)^i\binom{b}{i}\lambda_k$, $n_0 = 1 - m_0$ and $n_k = -m_k$ (Arik, 2018; Arik and Kantar 2019).

Theoretical Convergence:
The foundational expansion in Eq.(11) employs the binomial series $(1-z)^b$, which converges absolutely for $|z| < 1$. In our case $z = \frac{G(x)^a}{G(x)^a + \bar{G}(x)^a}$ Since $0 < G(x) < 1$ for any finite $x > 0$, it follows that $0 < z < 1$ for all x in the support of the distribution. This satisfies the convergence condition $|z| < 1$ universally, guaranteeing the convergence of the series in Eq. (11). All subsequent expansions (Eqs. 12, 14, 16) are power series in $G(x)$, which is bounded between 0 and 1. Therefore, these series also converge absolutely for the parameter space of interest $(a, b, \alpha, \theta > 0)$. See (Gradshteyn and Ryzhik 2007; Abramowitz and Stegun 1972).

Truncation Error Bounds:

In practice, the infinite series in Eq. (11) must be truncated after *i*=*N* terms. The resulting truncation error, $E(x)$, is the absolute difference between the true cdf $F(x)$ and its *N*-term approximation $\hat{F}(x)$:

$$E(x) = \sum_{i=N+1}^{\infty} (-1)^i \binom{b}{i} \frac{G(x)^{ai}}{[G(x)^a + \bar{G}(x)^a]^i}$$

$$|E(x)| = \left| \sum_{i=N+1}^{\infty} (-1)^i \binom{b}{i} \frac{G(x)^{ai}}{[G(x)^a + \bar{G}(x)^a]^i} \right| \leq \sum_{i=N+1}^{\infty} \left| (-1)^i \binom{b}{i} \frac{G(x)^{ai}}{[G(x)^a + \bar{G}(x)^a]^i} \right| \leq$$

$$\sum_{i=N+1}^{\infty} \left| (-1)^i \binom{b}{i} \right| z^i \leq \left| (-1)^{N+1} \binom{b}{N+1} \right| z^{N+1}$$

where $z = \frac{G(x)^a}{[G(x)^a + \bar{G}(x)^a]}$. This result guarantees that for any desired tolerance $\varepsilon > 0$, one can always find a sufficiently large *N* such that $E(x) < \varepsilon$ for all x in the support of the distribution. This uniform bound ensures the computational feasibility and numerical stability of the series representation for the *OBu-P* distribution. On the other hand, Table 1 presents the minimum number of terms (*N*) required in the series expansion to calculate the *OBu-P* distribution's properties at a given point x within a specified error tolerance. It demonstrates that a higher number of terms is needed both for larger values of x (further into the tail of the distribution) and for stricter (smaller) error tolerances.

Table 1. For $a = 2, b = 1.5, \theta = 1.0, \alpha = 2.0$, error values.

x	$Errors$		
	10^{-3}	10^{-6}	10^{-9}
0.5	8	12	16
1.0	10	15	20
2.0	12	17	22
5.0	15	20	25

Based on equation (17), cdf of *OBu-P* is expressed as a weighted sum of power series expansions of the baseline distribution function, $G(x)$ (Alizadeh et al., 2017; Altun et al., 2017). This formulation is particularly valuable for deriving key statistical properties of the *OBu-P* distribution, such as its moments, incomplete moments, and order statistics. The ability to represent the cdf in this manner, simplifies the calculation and analysis of statistical properties of the *OBu-P* variable (Arik, 2018).

Moments

By using (18), the r*th* moment corresponding to the *OBu-P* can be derived. To compute this, the r*th* moment of the EP distribution is expressed as follows:

$$\mu'_{EP} = E_{EP}(Z^r) = \alpha\beta\theta^\alpha \int_\theta^\infty z^{r-\alpha-1} \left(1 - \left(\frac{z}{\theta}\right)^{-\alpha}\right)^{\beta-1} dz \tag{19}$$

Here, α and β represent the shape parameters, while θ is the scale parameter of the EP distribution. Through some algebraic manipulation, the rth moment of the EP distribution is given as follows (Stoppa, 1990):

$$\mu'_{EP} = E_{EP}(Z^r) = \beta\theta^r B\left(1 - \frac{r}{\alpha}, \beta\right),\ r \leq \alpha \tag{20}$$

where $B(x, y) = \int_0^1 t^{(x-1)}(1-t)^{(y-1)} dt$ is the beta function. Moments of *OBu-P* random variable is:

$$\mu'_{OBu-P} = E_{OBu-P}(X^r) = \sum_{k=0}^{\infty} n_{k+1} \int_\theta^\infty x^r h_{(k+1)}(x) dx = \theta^r \sum_{k=0}^{\infty} n_{k+1} (k+1) B\left(1 - \frac{r}{\alpha}, (k+1)\right), r \leq \alpha. \tag{21}$$

In particular setting *r*=1, the mean of the *OBu-P* distribution is,

$$\mu_1' = \theta \sum_{k=0}^{\infty} n_{k+1}(k+1)B\left(1-\frac{1}{\alpha},(k+1)\right)$$

The *j*th central moment of *X* can be easily obtained from the equation (21) as follows:

$$\mu_j = E[(X-\mu)^j] = \sum_{k=0}^{j}\binom{j}{k}\mu_j'(-\mu_1')^{j-k}.$$

The interchange of summation and integration in Equation (21) is justified by the absolute and uniform convergence of the series representation of the *OBu-P* density for all $x > \theta$ and across the specified parameter ranges, which follows from the properties of power series expansions and the ratio of convergent series under the given conditions (see, Cordeiro and Castro 2011 for analogous justification in generalized distribution families).

Tables 2 and 3 report the mean, variance, skewness, and kurtosis of the OBu-P distribution (Caudill, 2012).

$$K \geq S^2 + 1. \tag{22}$$

From Table 3, all skewness–kurtosis pairs of the *OBu-P* distribution comply with Equation (22).

Table 2. Mean and variance values for the *OBu-P* distribution

a	α	b=0.75 (μ, σ²)	b=2 (μ, σ²)	b=5 (μ, σ²)
0.75	2.0	(3.03, 3.71)	(2.24,0.12)	(2.06,0.01)
2.0	2.0	(2.43, 0.12)	(2.23,0.02)	(2.14,0.01)
5.0	2.0	(2.33, 0.01)	(2.26,0.00)	(2.22,0.00)

Table 3. Skewness (*S)* and kurtosis (*K*) values for the *OBu-P* distribution

a	α	b=0.75 (*S*, *K*)	b=2.0 (*S*, *K*)	b=5.0 (*S*, *K*)
0.75	5.0	(12.66, 478.0)	(3.75,30.4)	(3.48,23.2)
2.0	7.5	(2.55, 16.45)	(1.16,5.74)	(0.68,3.59)
5.0	10.0	(1.39, 8.02)	(0.15,3.55)	(-0.24,3.13)

r*th* incomplete Moments

The r*th* incomplete moment of the *EP* random variable *Z* can be represented as

$$m_{r,Z} = \int_{\theta}^{w} z^r h(z)\,dz = \alpha\beta\theta^{\alpha}\int \theta w z^{-(\alpha+1)} z^r \left[1-\left(\frac{z}{\theta}\right)^{-\alpha}\right]^{\beta-1},\ z \geq \theta > 0, \alpha > 0, \beta > 0$$

$t = 1-\left(\frac{z}{\theta}\right)^{-\alpha}$ and $dt = \alpha\theta^{\alpha}z^{-(\alpha+1)}dz$, borders of integral are $z = \theta \Rightarrow t = 0$ and

$z = w \Rightarrow t = 1-\left(\frac{z}{\theta}\right)^{-\alpha} = t_w.$

$$m_{r,Z} = \alpha\beta\theta^{\alpha}\int_{\theta}^{t*}\frac{1}{\alpha\theta^{\alpha}}\theta^r(1-t)^{-\frac{r}{\alpha}}t^{\beta-1}\,dt = \beta\theta^r B_{t_w}\left(\beta, 1-\frac{r}{\alpha}\right),\ r \leq \alpha, \tag{23}$$

where $B_w(x,y) = \int_0^w t^{(x-1)}(1-t)^{(y-1)}dt$ is the incomplete beta function.

Accordingly, the r*th* incomplete moment of the *OBu-P* distribution is expressed as:

$$m_{r,X}(w) = \int_{\theta}^{w} z^r f_{OB-P}(z)dz = \int_{\theta}^{w} z^r \sum_{k=0}^{\infty} n_{k+1} h_{k+1}(z)dz = \sum_{k=0}^{\infty} n_{k+1} \int_{\theta}^{w} z^r h_{k+1}(z)dz$$

$$\sum_{k=0}^{\infty} n_{k+1} (k+1)\theta^r B_{t_w}\left(k+1, 1-\frac{r}{\alpha}\right) \quad r \leq \alpha \tag{24}$$

Moment Generating Function

The moment generating function (mgf) corresponding to the *EP* random variable is found (Tahir et al. 2016) as follows:

$$M_{EP}(t) = \alpha\beta \sum_{i=0}^{\infty} (-1)^i \theta^{(i+1)} \binom{\beta-1}{i} J(\theta, (i+1)\alpha+1, t)$$

where *J* function is provided as follows:

$$J(q,p,t) = \int_{q}^{\infty} x^{-p} e^{tx} dx = (-t)^p q \left[-\frac{\pi\, csc(p\pi)}{tq\Gamma(p)} - \frac{p\Gamma(-p)}{qt} + \frac{e^{tq}}{(-t)^{p+1} q^{p+1}} + \frac{p\Gamma(-p,-tq)}{tq}\right]$$

where $\Gamma(\alpha, t) = \int_t^{\infty} x^{\alpha-1} e^{-x} dx$ is incomplete gamma function. With the help of the equation (18) the mgf of the *OBu-P* distribution is obtained as follows:

$$M(t) = \alpha \sum_{k,i=0}^{\infty} n_{k+1}(k+1)\theta^{(i+1)\alpha} \binom{k}{i} J(\theta, (i+1)\alpha+1, t). \tag{25}$$

The equation (25) is the main result of this part.

Mean and Deviation

The deviation from the mean and deviation from the median are alternative measures for spread in a population (Akinsete et al., 2008; Nassar and Elmasry, 2012, Arik, 2018; Arik and Kantar 2019). Let *X* be a random variable from the *OBu-P* distribution with the mean $\mu = E(X)$. The mean deviation from the mean and the mean deviation from the median are defined, respectively, as follows:

$$D(\mu) = E[|X-\mu|] = \int_{\theta}^{\infty} |X-\mu| f(x)\, dx$$

and

$$D(M) = E[|X-M|] = \int_{\theta}^{\infty} |x-M| f(x) dx$$

The mean deviation from the mean can be derived in the following form:

$$\begin{aligned} D(\mu) &= \int_{\theta}^{\mu} (\mu-x) f(x) dx + \int_{\mu}^{\infty} (x-\mu) f(x) dx \\ &= 2\int_{\theta}^{\mu} (\mu-x) f(x) dx \\ &= 2\int_{\theta}^{\mu} \mu f(x) dx - 2\int_{\theta}^{\mu} x f(x) dx \\ &= 2\mu F(\mu) - 2\int_{\theta}^{\mu} x\, f(x)\, dx \end{aligned}$$

$$D(\mu) = 2\mu F(\mu) - 2m_{1,X}(\mu), \tag{26}$$

where $m_{1,X}(\mu)$is the first incomplete moment equation of the *OBu-P* distribution given in equation (24). Let *M* is the median of the random variable *X* and the mean deviation from the median is calculated as follows:

$$\begin{aligned} D(M) &= \int_{\theta}^{M}(M-x)f(x)dx + \int_{M}^{\infty}(x-M)f(x)dx \\ &= \int_{\theta}^{\infty}(x-M)f(x)dx + 2\int_{\theta}^{M}(M-x)f(x)dx \\ &= E(X-M) + 2\left[\int_{\theta}^{M}Mf(x)dx - \int_{\theta}^{M}xf(x)dx\right] \\ &= \mu - M + 2MF(M) - 2m_{1,X}(M) \end{aligned}$$

$$D(M) = \mu - 2m_{1,X}(M). \tag{27}$$

Order Statistics

Let $\{X_i\}_{i=1}^{n}$ be a sample of size *n* from the *OBu–P* distribution. After sorting the observations in increasing order, we obtain the sequence of order statistics $\{X_{(k)}: k = 1, 2, \dots, n\}$, where $X_{(k)}$ denotes the *k*-th smallest value. The pdf corresponding to the *k*-th order statistic, $X_{(k)}$, is given by,

$$f_{x_{(k)}}(x) = \frac{n!}{(n-k)!(k-1)!}f(x)[F(x)]^{k-1}[1-F(x)]^{n-k} = \frac{n!}{(n-k)!(k-1)!}\sum_{i=0}^{n-k}(-1)^i\binom{n-k}{i}[F(x)]^{k+i-1}f(x). \tag{28}$$

Thus, $f_{x_{(k)}}(x)$ can be written as,

$$f_{x_{(k)}}(x) = \frac{n!}{(n-k)!(k-1)!}\sum_{i=0}^{n-k}(-1)^i\binom{n-k}{i}\left[\sum_{l=0}^{\infty}(l+1)b_{l+1}g(x)G(x)^l\right]\left[\sum_{t=0}^{\infty}b_tG(x)^t\right]^{k+i-1} \tag{29}$$

$$f_{x_{(k)}}(x) = \frac{n!}{(n-k)!(k-1)!}\sum_{i=0}^{n-k}(-1)^i\binom{n-k}{i}\left[\sum_{l=0}^{\infty}(l+1)b_{l+1}g(x)G(x)^l\right]\left[\sum_{t=0}^{\infty}a_tG(x)^t\right] \tag{30}$$

is obtained (Gradshteyn & Ryzhik (2002)). Here $a_0 = b_0^{k+i-1}$ and for $t \geq 1, a_m = (mb_0)^{-1}\sum_{t=1}^{m}[t(k+i) - m]b_t a_{m-t}$. $X_{(k)}$pdf is expressed as,

$$f_{x_{(j)}}(x) = \sum_{i=0}^{n-k}\sum_{t,l=0}^{\infty}\frac{n!}{(n-k)!(k-1)!}(-1)^i\binom{n-k}{i}(l+1)b_{l+1}a_tg(x)G(x)^{l+t} = \sum_{i=0}^{n-k}\sum_{t,l=0}^{\infty}\frac{n!}{(n-k)!(k-1)!}(-1)^i\binom{n-k}{i}\frac{(l+1)b_{l+1}a_t}{(l+t+1)}h(x)^{l+t+1}. \tag{31}$$

Setting,

$$d_{i,t,l} = \frac{n!}{(n-k)!(k-1)!}(-1)^i\binom{n-k}{i}\frac{(l+1)b_{l+1}a_t}{(l+t+1)} \tag{32}$$

Finally, $X_{(k)}$'s pdf is:

$$f_{x_{(k)}}(x) = \sum_{i=0}^{n-k}\sum_{t,l=0}^{\infty}d_{i,t,l}h_{(l+t+1)}(x), \tag{33}$$

Here, *h(x)* denotes the *EP* pdf with parameters $(l + t + 1)$, α and θ and the equation (33) shows that the *OBu-P* order statistics pdf reduces to a linear combination of three EP terms (Arik and Kantar 2019).

3. MLE

Let $X_1, \dots, X_n$ be random variables from the *OBu-P* distribution with the parameters *a, b,* α and θ, thus, the corresponding log-likelihood function is,

$$\log L\,(a,b,\alpha,\theta) = n\log(ab\alpha) + nab\alpha\log\theta - (ab\alpha+1)\sum_{i=1}^{n}\log x_i$$
$$+(a-1)\sum_{i=1}^{n}\log\left[1-\left(\frac{x_i}{\theta}\right)^{-\alpha}\right] - (b+1)\sum_{i=1}^{n}\log\left\{\left[1-\left(\frac{x_i}{\theta}\right)^{-\alpha}\right]^{a} + \left(\frac{x_i}{\theta}\right)^{-\alpha a}\right\}. \quad (34)$$

The score functions of parameters are respectively presented as:

$$\frac{\partial \log L}{\partial a} = \frac{n}{a} + nb\alpha\log\theta - b\alpha\sum_{i=1}^{n}\log x_i + \sum_{i=1}^{n}\log\left[1-\left(\frac{x_i}{\theta}\right)^{-\alpha}\right]$$
$$-(b+1)\sum_{i=1}^{n}\frac{\left[1-\left(\frac{x_i}{\theta}\right)^{-\alpha}\right]^{a}\log\left[1-\left(\frac{x_i}{\theta}\right)^{-\alpha}\right] - \alpha\log\left(\frac{x}{\theta}\right)\left(\frac{x}{\theta}\right)^{-a\alpha}}{\left[1-\left(\frac{x_i}{\theta}\right)^{-\alpha}\right]^{a} + \left(\frac{x}{\theta}\right)^{-a\alpha}} = 0 \quad (35)$$

$$\frac{\partial \log L}{\partial b} = \frac{n}{b} + na\alpha\log\theta - a\alpha\sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n}\log\left\{\left[1-\left(\frac{x_i}{\theta}\right)^{-\alpha}\right]^{a} + \left(\frac{x_i}{\theta}\right)^{-\alpha a}\right\} = 0 \quad (36)$$

$$\frac{\partial \log L}{\partial \alpha} = \frac{n}{\alpha} + nab\log\theta - ab\sum_{i=1}^{n}\log x_i + (a-1)\sum_{i=1}^{n}\frac{\left(\frac{x_i}{\theta}\right)^{-\alpha}\log\left(\frac{x_i}{\theta}\right)}{\left[1-\left(\frac{x_i}{\theta}\right)^{-\alpha}\right]}$$
$$-(b+1)\sum_{i=1}^{n}\frac{a\log\left(\frac{x_i}{\theta}\right)\left[1-\left(\frac{x_i}{\theta}\right)^{-\alpha}\right]^{a-1}\left(\frac{x_i}{\theta}\right)^{-\alpha} - a\log\left(\frac{x_i}{\theta}\right)\left(\frac{x_i}{\theta}\right)^{a\alpha}}{\left(\frac{x_i}{\theta}\right)^{-a\alpha} + \left[1-\left(\frac{x_i}{\theta}\right)^{-\alpha}\right]^{a}} = 0 \quad (37)$$

The parameter θ is estimated by the sample $\min\{x_1, \dots, x_n\}$, while the remaining parameters are obtained by maximizing the log-likelihood function after excluding this observation from the sample. The resulting likelihood equations are solved numerically to obtain consistent estimates for the parameters, where a combination of the Newton–Raphson method and grid search is employed to ensure convergence to the global optimum (Aljarrah et al. 2015).

$$I(\hat{\beta}) = \begin{bmatrix} -\frac{\partial^2 \log L}{\partial a^2} & -\frac{\partial^2 \log L}{\partial a \partial b} & -\frac{\partial^2 \log L}{\partial a \partial \alpha} \\ -\frac{\partial^2 \log L}{\partial b \partial a} & -\frac{\partial^2 \log L}{\partial b^2} & -\frac{\partial^2 \log L}{\partial b \partial \alpha} \\ -\frac{\partial^2 \log L}{\partial a \partial \alpha} & -\frac{\partial^2 \log L}{\partial b \partial \alpha} & -\frac{\partial^2 \log L}{\partial \alpha^2} \end{bmatrix}^{-1} \quad (38)$$

where $\hat{\beta} = (\hat{a}, \hat{b}, \hat{\alpha})$. The 100(1-α)% confidence interval of the parameters is expressed as follows

$$\hat{a} \pm z_{\frac{\alpha}{2}}\sqrt{I_{11}}$$
$$\hat{b} \pm z_{\frac{\alpha}{2}}\sqrt{I_{22}}$$
$$\hat{\alpha} \pm z_{\frac{\alpha}{2}}\sqrt{I_{33}}$$

See Appendix for matrix.

4. Simulation

This section presents a simulation study evaluating the MLE performance for the *OBu-P* distribution, where random samples from (10) of sizes $n = 50, 100, 500$ are generated and each scenario is replicated 5000 times under different parameter settings.

Since $x \geq \theta$, the scale parameter θ is estimated as minimum x value. For each case, the bias and mean square error (MSE) values of the parameter estimates are calculated.

The results, summarized in Table 4, indicate that as the sample size n increases, both the bias and MSE values decrease. The systematic decrease in bias and mean squared error (MSE) with increasing n, is an indicator of consistency. These findings demonstrate that the MLE method provides reliable and efficient parameter estimates for the *OBu-P* distribution, particularly with larger sample sizes.

Table 4. MSE and Bias values

				BIAS				*MSE*			
n	a	b	α	$\hat{a}$	$\hat{b}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{a}$	$\hat{b}$	$\hat{\alpha}$	$\hat{\theta}$
50	0.5	0.5	0.5	-0.0723	-0.1485	-0.1301	-0.0075	0.0218	0.3586	2.2202	0.0004
	0.5	0.5	1	-0.0770	-0.1429	-0.2300	-0.0032	0.0231	0.2453	3.2005	0.0001
	0.5	0.5	2	-0.0681	-0.1261	-0.4452	-0.0019	0.0188	0.2944	6.3371	0.0000
	0.5	1	0.5	-0.0447	-0.1349	-0.1108	-0.0016	0.0101	0.4553	0.2526	0.0000
	0.5	1	1	-0.0480	-0.1509	-0.1789	-0.0009	0.0099	0.4364	0.9482	0.0000
	0.5	1	2	-0.0440	-0.1479	-0.3103	-0.0004	0.0094	0.4415	2.5061	0.0000
	0.5	2	0.5	-0.0276	-0.0001	-0.2180	-0.0004	0.0049	0.7441	0.3174	0.0000
	0.5	2	1	-0.0261	-0.0112	-0.4327	-0.0002	0.0050	0.8103	1.2746	0.0000
	0.5	2	2	-0.0267	-0.0786	-0.8549	-0.0001	0.0050	0.9207	10.6907	0.0000
100	0.5	0.5	0.5	-0.0450	-0.0947	-0.0327	-0.0017	0.0097	0.1912	0.1854	0.0000
	0.5	0.5	1	-0.0466	-0.0787	-0.0341	-0.0009	0.0096	0.1539	0.4484	0.0000
	0.5	0.5	2	-0.0455	-0.0688	-0.1040	-0.0004	0.0090	0.0840	2.0006	0.0000
	0.5	1	0.5	-0.0226	-0.0922	-0.0288	-0.0004	0.0041	0.1981	0.0651	0.0000
	0.5	1	1	-0.0266	-0.1094	-0.0274	-0.0002	0.0042	0.2076	0.2391	0.0000
	0.5	1	2	-0.0229	-0.0851	-0.1035	-0.0001	0.0042	0.1654	1.0007	0.0000
	0.5	2	0.5	-0.0146	-0.0465	-0.0997	-0.0001	0.0021	0.4907	0.0944	0.0000
	0.5	2	1	-0.0172	-0.0394	-0.2162	-0.0001	0.0026	0.5030	0.4600	0.0000
	0.5	2	2	-0.0155	-0.0398	-0.4361	0.0000	0.0023	0.4987	1.7891	0.0000
500	0.5	0.5	0.5	-0.0134	-0.0110	0.0016	-0.0001	0.0016	0.0087	0.0131	0.0000
	0.5	0.5	1	-0.0106	-0.0122	0.0049	0.0000	0.0013	0.0075	0.0466	0.0000
	0.5	0.5	2	-0.0126	-0.0146	0.0262	0.0000	0.0014	0.0076	0.1960	0.0000
	0.5	1	0.5	-0.0078	-0.0247	0.0018	0.0000	0.0008	0.0248	0.0116	0.0000
	0.5	1	1	-0.0060	-0.0159	-0.0067	0.0000	0.0007	0.0229	0.0425	0.0000
	0.5	1	2	-0.0083	-0.0161	-0.0037	0.0000	0.0008	0.0222	0.1617	0.0000
	0.5	2	0.5	-0.0050	-0.0354	-0.0102	0.0000	0.0004	0.1066	0.0159	0.0000
	0.5	2	1	-0.0043	-0.0323	-0.0294	0.0000	0.0005	0.1008	0.0641	0.0000
	0.5	2	2	-0.0039	-0.0245	-0.0656	0.0000	0.0005	0.1147	0.2668	0.0000

5. Real Life Application

The bladder cancer remission data, consisting of 128 observations, have also been analyzed by Zea et al. (2012) and Aljarrah et al. (2015). In this study, the OBu-P distribution is compared with the Pareto (PD), Weibull–Pareto (WPD; Aljarrah et al., 2015), Beta–Pareto (BPD; Akinsete et al., 2008), and Beta–Exponentiated Pareto (BEPD; Zea et al., 2012) distributions, as discussed in Arik (2018). Table 5 presents the parameter estimates $\hat{\theta}$ and $\hat{\beta}$ as scale parameters, others as shape parameters, together with the Negative Log-Likelihood (−LL), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Kolmogorov–Smirnov (K-S) statistics, along with the corresponding standard errors in parentheses. As seen in Table 5, the *OBu-P* model yields the lowest values of −LL, AIC, BIC, and K-S, confirming its superior fit compared to the alternative Pareto-type models. Furthermore, Figure 6 illustrates that the *OBu-P* distribution provides a noticeably better fit than PD, WPD, BPD, and BEPD for the remission time data.

Table 5. Parameter estimates and criteria results for bladder cancer data

	PD	WPD	BPD	BEPD	*OBu-P*
MLE	$\hat{\alpha} = 0.232$ (0.021)	$\hat{c} = 4.136$ (0.118)	$\hat{a} = 4.805$ (0.055)	$\hat{a} = 0.348$ (0.097)	$\hat{a} = 4.277$ (0.338)
	$\hat{\theta} = 0.080$	$\hat{\gamma} = 0.436$ (0.088)	$\hat{b} = 100.502$ (0.251)	$\hat{b} = 159831$ (183.7)	$\hat{b} = 6.674$ (7.938)
		$\hat{k} = 0.077$ (0.013)	$\hat{k} = 0.011$ (0.001)	$\hat{k} = 0.051$ (0.019)	$\hat{\alpha} = 0.107$ (0.028)
		$\hat{\theta} = 0.080$	$\hat{\theta} = 0.080$	$\hat{\alpha} = 8.611$ (2.093)	$\hat{\theta} = 0.080$
				$\hat{\beta} = 0.080$	
-logL	539.591	407.370	480.446	432.410	406.965
AIC	1081.182	820.740	966.893	872.819	819.930
BIC	1084.026	829.273	975.425	884.197	828.463
K-S	0.425	0.043	0.217	0.142	0.039

Note: The values of negative log-likelihood, AIC and K-S statistics for WPD, BPD and BEPD are taken from (Aljarrah et al., 2015). The criteria results for PD and OBu-PD are calculated by using MATLAB software. Standard errors are given in parentheses. This application is taken from from the PhD thesis of Ibrahim Arık (2018).

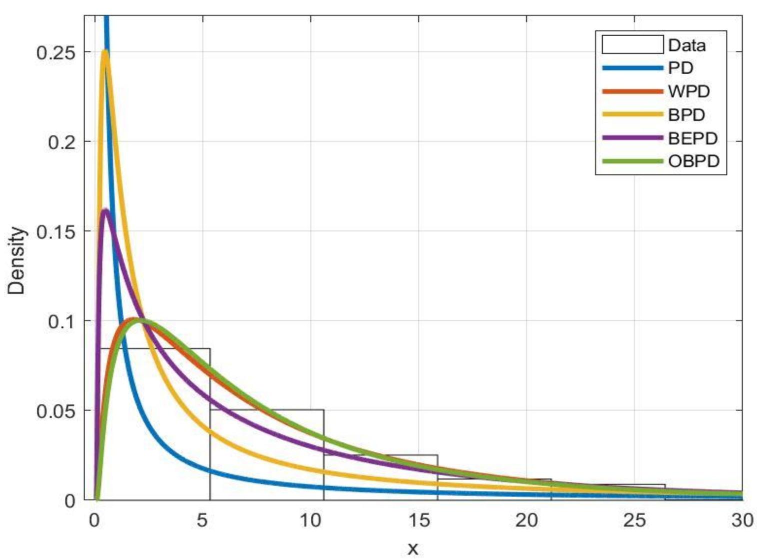

Figure 6. Pdfs of PD, WPD, BPD, BEPD, and OBu-P and the Histogram for Bladder Cancer Data

6. Conclusion

In this study, we introduce a novel four-parameter *OBu-P* distribution. We explore its statistical properties, including the hazard rate, survival and quantile functions, moments, incomplete moments, moment generating function, mean

 DOI: https://doi.org/10.18187/pjsor.v21i3.4950

deviation, and order statistics. We also derive the MLE estimates for the *OBu-P* distribution and demonstrated its consistency through a simulation study.

The *OBu-P* distribution is applied to a real-world dataset (Arik 2018), where it was shown to outperform other generalized and/or extended Pareto distributions in terms of fit. Notably, the *OBu-P* distribution proves to be a highly flexible model for datasets containing statistical uncertainty. Its capacity to capture the behaviour of peak data makes it particularly well-suited for applications in insurance, survival analysis, and fields where understanding extreme or rare events is critical. Therefore, the *OBu-P* distribution emerges as a promising candidate for modelling and analysing data that involves extreme or peak values, offering significant improvements in prediction and risk assessment.

Appendix:

The elements of the Hessian matrix, i.e., the second-order partial derivatives of the log-likelihood function for the *OBu-P* distribution, are given by:

$$\frac{\partial^2 \log L}{\partial a^2} = -\frac{n}{a^2} - (b+1)\sum_{i=1}^{n} \frac{\left(\frac{x_i}{\theta}\right)^{\alpha a} {u_i}^{a}\left\{\log^2 u_i + 2\alpha \log\left(\frac{x_i}{\theta}\right)\log u_i + \alpha^2 \log^2\left(\frac{x_i}{\theta}\right)\right\}}{\left\{\left(\frac{x_i}{\theta}\right)^{\alpha a} {u_i}^{a} + 1\right\}^2},$$

$$\frac{\partial^2 \log L}{\partial b^2} = -\frac{n}{b^2},$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = -\frac{n}{\alpha^2} - (a-1)\sum_{i=1}^{n}\left[\log^2\left(\frac{x_i}{\theta}\right)(1-u_i){u_i}^{-1} + \log^2\left(\frac{x_i}{\theta}\right)(1-u_i)^2 {u_i}^{-2}\right]$$
$$-(b+1)\sum_{i=1}^{n}\left\{\begin{array}{c}\frac{-a\log^2\left(\frac{x_i}{\theta}\right){u_i}^{a-1}(1-u_i) + (a-1)\log^2\left(\frac{x_i}{\theta}\right){u_i}^{a-2}(1-u_i)^2 - a^2\left(\frac{x_i}{\theta}\right)^{a\alpha}\log^2\left(\frac{x_i}{\theta}\right)}{(1-u)^a + u^a} \\ \frac{\left[a\log\left(\frac{x_i}{\theta}\right){u_i}^{a-1}(1-u_i) - a\log\left(\frac{x_i}{\theta}\right)(1-u_i)^a\right]\left[a\log\left(\frac{x_i}{\theta}\right){u_i}^{a-1}(1-u_i) - a\left(\frac{x_i}{\theta}\right)^{a\alpha}\log\left(\frac{x_i}{\theta}\right)\right]}{[(1-u_i)+{u_i}^a]^2}\end{array}\right\},$$

$$\frac{\partial^2 \log L}{\partial a \partial b} = n\alpha \log\theta - \alpha\sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n}\frac{{u_i}^a \log u_i - \alpha\log\left(\frac{x_i}{\theta}\right)(1-u_i)^a}{{u_i}^a + (1-u_i)^a},$$

$$\frac{\partial^2 \log L}{\partial a \partial \alpha} = nb\log\theta - b\sum_{i=1}^{n}\log x_i + \sum_{i=i}^{n}\log\left(\frac{x_i}{\theta}\right)(1-u_i){u_i}^{-1}$$
$$-(b+1)\sum_{i=1}^{n}\left\{\begin{array}{l}\frac{a\log^2\left(\frac{x_i}{\theta}\right)\alpha\left(\frac{x_i}{\theta}\right)^{-a\alpha} + a\log\left(\frac{x_i}{\theta}\right)\log u_i\, {u_i}^{a-1}(1-u_i) + \log\left(\frac{x_i}{\theta}\right){u_i}^{a-1}(1-u_i) - \log\left(\frac{x_i}{\theta}\right)(1-u_i)^a}{(1-u_i)^a + {u_i}^a} \\ -\frac{\left[a\log\left(\frac{x_i}{\theta}\right){u_i}^{a-1}(1-u_i) - a\log\left(\frac{x_i}{\theta}\right)(1-u_i)^a\right]\left[\log u_i\, {u_i}^{a-1} - \alpha\log\left(\frac{x_i}{\theta}\right)(1-u_i)^a\right]}{[(1-u_i)^a + {u_i}^a]^2}\end{array}\right\}$$

$$\frac{\partial^2 \log L}{\partial \alpha \partial b} = na\log\theta - a\sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n}\frac{a\log\left(\frac{x_i}{\theta}\right){u_i}^{a-1}(1-u_i) - a\log\left(\frac{x_i}{\theta}\right)\left(\frac{x_i}{\theta}\right)^{a\alpha}}{(1-u_i)^a + {u_i}^a}$$

where,

$$u_i = 1 - \left(\frac{x_i}{\theta}\right)^{-\alpha}.$$

In the Newton–Raphson method for maximum likelihood estimation, the parameter vector is updated iteratively as

$$\theta^{(m+1)} = \theta^{(m)} - H^{-1}\left(\theta^{(m)}\right)U\left(\theta^{(m)}\right)$$

where $U(\theta)$ is the score vector (first derivatives of the log-likelihood / Eqs, 35-37) and $H(\theta)$ is the Hessian matrix (second derivatives). At each iteration, the Hessian provides curvature information, scaling the score vector so that updates move efficiently toward the maximum. A negative definite Hessian at convergence confirms that the solution is indeed a local maximum of the log-likelihood function.

Acknowledgement This article was produced from a part of Ph.D. thesis entitled "Arik, I., New distribution families and their statistical properties for survival analysis, Anadolu University, Graduate School of Sciences, Eskisehir", Turkey.

References

1. Abramowitz, M. & Stegun, I.A. (1972). Handbook of Mathematical Functions. Dover.
2. Akinsete A., Famoye F., & Lee, C. (2008). The beta-Pareto distribution. Statistics, 42(6), 547-563.
3. Alexander C., Cordeiro GM., Ortega EMM., Sarabia JM. (2012) Generalized beta-generated distributions. Computational Statistics & Data Analysis, 56, 1880–1897.
4. Alizadeh M.., Cordeiro GM, Nascimento ADC., Lima, M. do CS, Ortega EMM. (2017). Odd-Burr generalized family of distributions with some applications. Journal of Statistical Computation and Simulation, 87(2), 367-389
5. Aljarrah, MA., Famoye F, Lee C. (2015). A new Weibull-Pareto distribution. Communications in Statistics - Theory and Methods, 44, 4077-4095.
6. Altun G., Alizadeh M., Altun E, Ozel G. (2017). Odd Burr Lindley distribution with properties and applications. Hacettepe Journal of Mathematics and Statistics, 46(2), 255-276
7. Alzaatreh A., Famoye F., Lee C. (2012). Gamma-Pareto distribution and its applications. Journal of Modern Applied Statistical Methods, 11(1), 78-94
8. Alzaatreh A., Lee C., Famoye F. (2013). A new method for generating families of continuous distributions. Metron, 71, 63–79
9. Arik I. (2018). New distribution families and their statistical properties for survival analysis PhD Dissertation, Anadolu University, Eskisehir, Turkey.
10. Arik I., Kantar YM. (2019). New Odd Burr-Rayleigh Distribution: Theory and Applications. Far East Journal of Theoretical Statistics, 55(1), 53–82.
11. Arnold BC. (2008). Pareto and Generalized Pareto Distributions in Modelling income distributions and Lorenz Curves (Chapter 7) Springer.
12. Bhatt, A. (2023). Advanced Statistics (MAT503). Uttarakhand Open University, School of Sciences, Haldwani, Uttarakhand, India, 263139.
13. Bourguignon M., Silva RB., Cordeiro GM. (2014). The Weibull-G family of probability distributions. Journal of Data Science, 12, 53–68.
14. Bourguignon M., Silva RB., Zea LM., Cordeiro GM. (2013). The Kumaraswamy Pareto distribution. Journal of Statistical Theory and Applications, 12(2), 129-144.
15. Caudill SB. (2012). A partially adaptive estimator for the censored regression model based on a mixture of normal distributions. Statistical Methods and Applications, 21, 121–137.
16. Cordeiro GM., de Castro M. (2011). A new family of generalized distributions. Journal of Statistical Computation and Simulation, 81, 883–898.
17. Eugene N., Lee C., Famoye F. (2002). Beta-normal distribution and its applications. Communications in Statistics - Theory and Methods, 31, 497–512
18. Imsap P., Bodhisuwan W., Volodin A. (2023). Alpha Power Exponentiated Pareto Distribution Lobachevskii. Journal of Mathematics, 44, 3779–3788.
19. Ihtisham S., Khalil A., Manzoor S., Khan SA., Ali A. (2019). Alpha-Power Pareto distribution: Its properties and applications. *PLOS ONE*, 14(6).
20. Mahmoudi E. (2011). The beta generalized Pareto distribution with application to lifetime data. Mathematics and Computers in Simulation, 81(11), 2414-2430
21. Godase, SS., Shirke, DT., Kashid, DN. (2017). Tolerance intervals and confidence intervals for the scale parameter of Pareto-Rayleigh distribution. East Journal of Applied Statistical Analysis, 10(1), 29–49
22. Gradshteyn IS., Ryzhik IM. (2002). *Table of integrals, series, and products.* Academic Press
23. Johnson NL., Kotz S., Balakrishnan N. (1994). Continuous univariate distributions (Vol 1, 2nd ed) Wiley
24. Kumaraswamy P. (1980). A generalized probability density function for double-bounded random processes. Journal of Hydrology, 46, 79-88.

25. Marshall AW., 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, 641–652.
26. Nadarajah S. (2005) Exponentiated Pareto distributions. Statistics, 39(3), 255–260
27. Nadarajah S., Chan S., Afuecheta CS. (2013) Double generalized Pareto distribution. Statistics & Probability Letters, 83, 2656-2663.
28. Nassar, M.M., & Elmasry, A. (2012). A study of generalized logistic distributions. *Journal of the Egyptian Mathematical Society*, 20(2), 126–133. https://doi.org/10.1016/j.joems.2012.08.011
29. Olmos NM., Gómez-Déniz E., Venegas O., Gómez-Geraldo H. (2024). A composite half-normal-Pareto distribution with applications to income and expenditure data. Mathematics, 12(11), 1631.
30. Tahir MH., Cordeiro GM., Alzaatreh A., Mansoor AM., Zubair M. (2016). A new Weibull–Pareto distribution: Properties and applications. Communications in Statistics- Simulation and Computation, 45(10), 3548–3567.
31. Pareto V. (1986). La courbe de la répartition de la richesse In C Viret-Genton (Ed), *Recueil publié par la Faculté de Droit à l'occasion de l'Exposition Nationale Suisse* (pp 373–387) Université de Lausanne
32. Smith LR. (1985). Maximum likelihood estimation in a class of nonregular cases. Biometrika, 72(1), 67–90.
33. Stoppa G.(1990). A new model for income size distributions, In C Dagum & M Zenga (Eds), Income and Wealth Distribution, Inequality and Poverty (pp 33–41) Springer.
34. Zea LM, Silva RB, Bourguignon M, Santos AM, Cordeiro, GM (2012) The beta exponentiated Pareto distribution with application to bladder cancer susceptibility International Journal of Statistics and Probability, 1(2), 8–19
35. Zinodiny S, Nadarajah S. (2021.) Matrix variate Pareto distributions Mathematica Slovaca, 71(2), 475-490.