IRAQI Academic Scientific Journals

University of Kirkuk Journal For Administrative and Economic Science (2025) 15 (3) Part (1): 191-201

UKJAES

University of Kirkuk Journal For Administrative and Economic Science

ISSN:2222-2995 E-ISSN:3079-3521
University of Kirkuk Journal For Administrative and Economic Science

Latif Zainab Tariq Abdel & Ahmed Rikan A. . Transmuted the Exponentiated Pareto-I Distribution Using Ranked Set Sampling. *University of Kirkuk Journal For Administrative and Economic Science* (2025) 15 (3) Part (1) :191-201.

Transmuted the Exponentiated Pareto-I Distribution Using Ranked Set Sampling

Zainab Tariq Abdel Latif [1], Rikan A. Ahmed [2]

[1,2] *Department of Statistics and Informatics, University of Mosul, Mosul, Iraq*

zainab.23csp134@student.uomosul.edu.iq [1], rikan.ahmed@uomosul.edu.iq [2]

Abstract: Probability distributions are extensively employed in the analysis of natural phenomena. Numerous distributions have been defined and examined over time; however, their capabilities are limited, rendering them unsuitable for every situation in representing the phenomena under study, as the characteristics of these phenomena evolve over time. Consequently, statistical researchers encounter various challenges when initiating statistical analysis, with the primary difficulty being the identification of the suitable probability distribution for the data of the examined phenomenon, owing to the intricate nature of real data by significant skewness, curvature, or the presence of outliers.

This paper aims to suggest a strategy for generalizing the exponentiated Pareto-I distribution by characterizing ranked set sampling and studying the properties of this distribution in modelling the phenomenon and deriving an equivalent model.

A follow-up was conducted on prior studies concerning the topic, providing a comprehensive account of the rank sample as a contemporary methodology employed. Subsequently, a theoretical exploration of the research subject was undertaken, examining the attributes of the generalized distribution and the estimation of its parameters through the maximum likelihood method, accompanied by an elucidation of both experimental and practical applications.

The simulation results showed how different sample sizes affected the estimated values of the model parameters, variance, bias, and MSE, along with how the alpha and lambda parameters influenced some averages and skewness. The When tested on empirical data, the model under consideration proved more suitable than the exponential distribution and the exponentiated Pareto-I distribution.

The results indicate that using a bigger ranked sample size and an increased number of cycles r yields more accurate for estimations. From the result of the application, we determined that the proposed distribution, CTEP-1, adequately fits the ranked sampling drawn from the unimodal leptokurtic and right-skewed data, in addition to the unimodal mesokurtic data.

Keywords: Rank sample, CTEP-1 distribution, Exponentiated Pareto-I, Maximum likelihood, Parameter estimation.

تحويل توزيع باريتو-I الأسي باستخدام مجموعة العينات المرتبة

الباحثة: زينب طارق عبد اللطيف[1]، أ.م.د. ريكان عبد العزيز احمد[2]

[1،2] قسم الإحصاء والمعلوماتية-جامعة الموصل، الموصل، العراق

المستخلص: تُستخدم التوزيعات الاحتمالية في تحليل الظواهر الطبيعية، لكن قدراتها محدودة، مما يتطلب اختيار توزيع مناسب للبيانات المعقدة. يهدف هذا البحث إلى تعميم توزيع باريتو-١ الأسي من خلال دراسة العينات المرتبة وخصائص هذا التوزيع. تم استخدام طريقة الاحتمالية القصوى لتقدير المعالم، مع توضيح التطبيقات العملية. أظهرت نتائج المحاكاة أن حجم العينة يؤثر على دقة التقديرات، وأن التوزيع المقترح، CTEP-1، يتناسب بشكل جيد مع البيانات ذات القمة الأحادية وذيول طويلة ومنحرفة نحو اليمين.

الكلمات المفتاحية: عينة مرتبة، توزيعCTEP-1، الأسي باريتو-١، الاحتمالية العظمى، تقدير المعلمات.

Corresponding Author: E-mail: zainab.23csp134@student.uomosul.edu.iq

Introduction

The expense of data collection is a critical concern in statistics, particularly when the subjects of investigation are costly or labor-intensive to assess. In the early 1950s, McIntyre [1] introduced a sampling technique known as Rank Sample (RSS). McIntyre's initial concept was expanded, rendering the method relevant to a broader array of disciplines, and other adaptations and implications of the original notion have been suggested and examined.

This data collection approach is antiquated, although it gained widespread utilization only in the final decade of the previous century. The predominant method of data collection employs simple random sampling from the population under study. Even so, the rank sample technique enhances the identification of the most representative measurements, yielding superior results. As a result, compared to simple random sampling, rank sample provides a valuable alternative with improved efficiency. because it lowers the variance of the estimator while keeping the same level of accuracy with a smaller sample size than simple random sampling [2].

Conversely, researchers encounter numerous challenges in conducting statistical analyses, particularly in identifying the suitable distribution for the data pertaining to the phenomenon under investigation. This complexity arises from the intricate behavior of real data, often characterized by significant skewness, curvature, or the presence of outliers. Probability distributions are extensively employed in the examination of natural phenomena, and although many distributions have been defined and analyzed over the years, their applicability is constrained. Consequently, they cannot universally represent the phenomenon under study, as the characteristics of the phenomenon evolve. Gupta et al [3]. presented the two-parameter Pareto exponential distribution, demonstrating its efficacy in the analysis of various age data sets.

Altering the standard distribution frequently leads to a composite distribution that is more effective and adaptable. This facilitates the analysis of data from two distinct ranges and enhances the realism of data modelling compared to conventional single distributions.

This study uses a broader version of the transformed distribution that Al-Kadim [4] made to show a new broader version of the first broader exponential Pareto distribution. This variant of the initial Pareto distribution is referred to as the cubic transmuted Pareto (CTEP)[5]. The generalization of probability distributions seeks to enhance their applicability, rendering them more adaptable for expressing diverse events in both practical and statistical contexts and facilitating the development of a more accurate model for addressing intricate real-world data [6].

This research is structured as follows: Section 2 delineates the ranked set sampling approach; Section 3 introduces the distribution used in the study; Section 4 articulates the composite distribution employing ranked set sampling; Section 5 illustrates the generalization map; Section 6 presents the generalized distribution model via ranked set sampling and examines some of its properties; Section 7 estimates the parameters of the model under investigation using the maximum likelihood (ML) approach; Section 8 showcases the simulation results; Section 9 reveals the application results pertaining to the real data set; finally, Section 10 discusses the conclusion and concludes the study.

1st: RANK SAMPLE

Rank sample is the methodology of obtaining observations from a population by categorizing it into ranked strata and thereafter selecting a random sample from each stratum according to its size and significance Wolfe, D.A. [7], Zamanzade, E. and A.I. Al-Omari [8]. When compared to simple random sampling, rank sample gives a more accurate picture of the population because it is more likely to include all possible values with a smaller sample size.

1- To obtain a sample via the rank sample method, the subsequent procedures must be executed:

A. Utilize simple random sampling to select an initial sample from a statistical population.
B. Evaluate this sample depending on the relevant attribute by employing a set of procedures to derive the ranking, utilizing auxiliary considerations, expert opinions, and visual comparisons.
C. Arrange k simple random samples that are separate and have a size of k. Rely on the total count of units in the population to arrange these samples, which completes the cycle of one ordered set.
D. The unit employed for found to be the smallest after the ranking is called the smallest ranking statistic, and it is denoted by $x_{[1]}$. The reason for using a square bracket is used in place of the conventional parentheses is that the smallest ranking statistic may or may not be the smallest of the k^{th} groups in simple random sampling.
E. The process is repeated by taking a second random sample (independent of the first sample) of size k from the population, and this group is arranged in the same way as the first group.
F. The item ranked as the second smallest element of the group k is selected and is termed the second degree of judgment statistic and is denoted by $x_{[2]}$.
G. The operation is iterated k^{th} times until a sample size is k obtained where this process produces the measured observations $x_{[1]1}, x_{[2]1}, \ldots, x_{[k]1}$, and this process is termed a cycle as k it represents the size of the group.
H. Arrange k simple random samples that are separate and have a size of k. Use the total number of k^2 units in the population to arrange these samples, which completes the cycle of one ordered set.
I. The process is done r^{th} times to get r separate cycles that add up to the needed total number of observations. Doing this task r times will give you a balanced sample size $m = r \times k$, and if the totals k are equal, you have a balanced sample (BRSS).

2- Below is an example illustrating the procedures for sampling ordered sets utilizing five observations of the variable x with a number of cycles equal to

1. All samples are randomly selected.

Selected samples →

1	2	3	4	5
x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
x_{21}	x_{22}	x_{23}	x_{24}	x_{25}
x_{31}	x_{32}	x_{33}	x_{34}	x_{35}
x_{41}	x_{42}	x_{43}	x_{44}	x_{45}
x_{51}	x_{52}	x_{53}	x_{54}	x_{55}

2. The chosen samples are organized.

Arranged samples

1	2	3	4	5
$x_{(11)}$	$x_{(12)}$	$x_{(13)}$	$x_{(14)}$	$x_{(15)}$
$x_{(21)}$	$x_{(22)}$	$x_{(23)}$	$x_{(24)}$	$x_{(25)}$
$x_{(31)}$	$x_{(32)}$	$x_{(33)}$	$x_{(34)}$	$x_{(35)}$
$x_{(41)}$	$x_{(42)}$	$x_{(43)}$	$x_{(44)}$	$x_{(45)}$
$x_{(51)}$	$x_{(52)}$	$x_{(53)}$	$x_{(54)}$	$x_{(55)}$

3. Each of the elements of the principal diagonal is chosen to represent the sample of the ranking sets.

1	2	3	4	5
$x_{(11)}$	$x_{(12)}$	$x_{(13)}$	$x_{(14)}$	$x_{(15)}$
$x_{(21)}$	$x_{(22)}$	$x_{(23)}$	$x_{(24)}$	$x_{(25)}$
$x_{(31)}$	$x_{(32)}$	$x_{(33)}$	$x_{(34)}$	$x_{(35)}$
$x_{(41)}$	$x_{(42)}$	$x_{(43)}$	$x_{(44)}$	$x_{(45)}$
$x_{(51)}$	$x_{(52)}$	$x_{(53)}$	$x_{(54)}$	$x_{(55)}$

2nd: THE EXPONINTIATED PARETO-I DISTRIBUTION (EP-I)

This distribution serves as a versatile analytical instrument for characterizing data with significant variability. It is utilized across various disciplines, including statistics, economics, and data science, to model data exhibiting specific characteristics, such as a significant likelihood of outliers and a concentration around minimal values Eledum, H. and S. Ansari [9], Eledum, H. and S. I Ansari [10], Eledum, H [11].
Suppose a random variable x is distributed according to a complex Exponentiated Pareto-I distribution.
Its distribution function (pdf) is defined as follows:

$$g(x) = \alpha b^{\alpha} e^{-\alpha x} ; x \in (\ln b, \infty); \alpha, b > 0 \quad (1)$$

Where α represents the shape parameter and b represents the location parameter,
The probability function (CDF) is given by:

$$G(x) = 1 - b^{\alpha} e^{-\alpha x} ; x \in (\ln b, \infty); \alpha, b > 0 \quad (2)$$

3rd: USING RANK SAMPLE IN DISTRIBUTIONAL STUDIES

The ranked set sampling method is utilized to improve the estimate of distribution parameters, attain increased accuracy in results, and reduce the impact of outliers. This benefit arises from its focus on the natural or logical organization of values, hence diminishing the probability of errors or outliers inducing confusion.
Assuming that the observations $(x_{11}, x_{22}, \ldots, x_{mm})$ represent a balanced random sample (BRSS), the probability function (PDF) for the random variable x_i is the ordered statistics function defined as follows:

$$h(x_i) = \binom{k}{i-1} (G(x_i))^{i-1} (1 - G(x_i))^{k-i} g(x_i)$$

If the sample is taken from a population that follows an Exponentiated Pareto-I distribution, the probability function for the random variable x_i is defined as follows:

$$h(x / \alpha, b) = \binom{k}{i-1} (1 - b^{\alpha} e^{-\alpha x})^{i-1} (1 - (1 - b^{\alpha} e^{-\alpha x}))^{k-i} \alpha b^{\alpha} e^{-\alpha x}$$

$$= \binom{k}{i-1} (1 - b^{\alpha} e^{-\alpha x})^{i-1} (1 - 1 + b^{\alpha} e^{-\alpha x})^{k-i} \alpha b^{\alpha} e^{-\alpha x}$$

For $i = 1, 2, \ldots k ; j = 1, 2, \ldots r$

$$h(x_{[i]j} / \alpha, b) = \frac{1}{r} \sum_{j=1}^{r} \sum_{i=1}^{k} \binom{k}{i-1} (1 - b^{\alpha} e^{-\alpha x_{[i]j}})^{i-1} (1 - 1 + b^{\alpha} e^{-\alpha x_{[i]j}})^{k-i} \alpha b^{\alpha} e^{-\alpha x_{[i]j}} \quad (3)$$

The probability function (CDF) is given by:

$$H(x_{[i]j} / \alpha, b) = \int_{\ln b}^{\infty} \frac{1}{r} \sum_{j=1}^{r} \sum_{i=1}^{k} \binom{k}{i-1} (1 - b^{\alpha} e^{-\alpha x_{[i]j}})^{i-1} (1 - 1 + b^{\alpha} e^{-\alpha x_{[i]j}})^{k-i} \alpha b^{\alpha} e^{-\alpha x_{[i]j}} \quad (4)$$

To prove that $h(x_{[i]j} / \alpha, b)$ is a PDF, both of the following must be proven:

1) $h(x_{[i]j}/\alpha,b)\geq 0$

And to prove that $h(x_{[i]j}/\alpha,b)\geq 0$ we substitute $x=\ln b$ in the equation (3)

$$h(x_{[i]j}/\alpha,b)=\frac{1}{r}\sum_{j=1}^{r}\sum_{i=1}^{k}\binom{k}{i-1}(1-b^{\alpha}e^{-\alpha x_{[i]j}})^{i-1}(1-1+b^{\alpha}e^{-\alpha x_{[i]j}})^{k-i}\alpha b^{\alpha}e^{-\alpha x_{[i]j}}$$

Let $u=i-1$

$$\frac{1}{r}\sum_{j=1}^{r}\sum_{u=0}^{k-1}\binom{k-1}{u}(1-b^{\alpha}e^{-\alpha x_{[i]j}})^{u}(1-1+b^{\alpha}e^{-\alpha x_{[i]j}})^{k-(u+1)}=1$$

$$h(x_{[i]j}/\alpha,b)=\alpha b^{\alpha}e^{-\alpha\ln b}$$

$$=\alpha b^{\alpha}e^{\ln b^{-\alpha}}=\alpha b^{\alpha}b^{-\alpha}=\alpha$$

Since $\alpha>0$

And by compensation $x=\infty$ in the equation (3) be $f(x_{[i]j}/\alpha,b)=0$

2) $\int_{-\infty}^{\infty}h(x_{[i]j}/\alpha,b)dx_{[i]j}=1$

$$\int_{\ln b}^{\infty}\alpha b^{\alpha}e^{-\alpha x_{[i]j}}dx_{[i]j}$$

$$=\alpha b^{\alpha}\int_{\ln b}^{\infty}e^{-\alpha x_{[i]j}}dx_{[i]j}=\alpha b^{\alpha}\left[\frac{e^{-\alpha x_{[i]j}}}{-\alpha}\right]_{\ln b}^{\infty}$$

$$=\alpha b^{\alpha}\frac{e^{-\alpha\ln b}}{\alpha}=\alpha b^{\alpha}\frac{b^{-\alpha}}{\alpha}=1$$

4th: CUBIC RANKING TRANSMUTATION MAP

The researcher AL-Kadim [4] devised a generalization for the transmuted distribution, wherein the probability function (CDF) and the distribution function (PDF) are delineated as follows:

$$f(x)=(1+\lambda)g(x)-4\lambda g(x)G(x)+3\lambda G2(x)g(x) \quad (5)$$

$$F(x)=(1+\lambda)G(x)-2\lambda G2(x)+\lambda G3(x),|\lambda|\leq 1,x\in(\ln b,\infty);\alpha,b>0 \quad (6)$$

where G(x) and g(x) refer to the CDF and pdf of the base distribution, respectively, α represents the shape parameter, b represents the location parameter and λ represent the transmuted parameter.

5th: CUBIC TRANSMUTED EXPONENTIATED PARETO-I DISTRIBUTION BY USING RANK SAMPLE

This part talks about the generalized distribution using rank set sampling. It goes into detail about the probability density function (pdf), the cumulative distribution function (CDF), the reliability function, other statistical features.

The plots of the probability density Functions in Figures 1 and 2 indicate that, as the value of the shape parameter (α) grows larger, the skewness of the distribution stabilizes, as corroborated by Table 2. Conversely, an increase in the transmuted parameter λ results in heightened skewness, as corroborated by Table 3. Additionally, Figure 3 illustrates that b serves as the location parameter. Furthermore, depending on the value of the transmuted parameter, the distribution displays moderate to significant right skewness.

Let $x_{[i]j}$ be order statistics with CTEP-I distribution. The (PDF) and (CDF) are specified, respectively, as:

$$f(x_{[i]j})=\alpha b\alpha e^{-\alpha x_{[i]j}}-2\lambda\alpha b^{2\alpha}e^{-2\alpha x_{[i]j}}+3\lambda\alpha b^{3\alpha}e^{-3\alpha x_{[i]j}}$$

$$f(x_{[i]j})=\alpha b^{\alpha}e^{-\alpha x_{[i]j}}(1-2\lambda b^{\alpha}e^{-\alpha x_{[i]j}}+3\lambda b^{2\alpha}e^{-2\alpha x_{[i]j}}) \quad (7)\; x_{[i]j}\in[\ln b,\infty);\ \alpha,b>0,\ |\lambda|\leq 1$$

$$F(x_{[i]j})=1-b^{\alpha}e^{-\alpha x_{[i]j}}(1-\lambda b^{\alpha}e^{-\alpha x_{[i]j}}+\lambda b^{2\alpha}e^{-2\alpha x_{[i]j}}) \quad (8)$$

To prove $f(x_{[i]j})$ is a pdf, we need to prove $f(x_{[i]j}) \geq 0$ and $\int_{-\infty}^{\infty} f(x_{[i]j})dx_{[i]j} = 1$

$Proof\ of\ f(x) \geq 0$

Substitute $x_{[i]j} = lnb$ in Eq. (7) we get

$$f\left(x_{[i]j}\right) = \alpha b^{\alpha} e^{-\alpha lnb}(1-2\lambda b^{\alpha} e^{-\alpha lnb} + 3\lambda b^{2\alpha} e^{-2\alpha lnb}$$

$$= \alpha b^{\alpha} e^{lnb^{-\alpha}} (1-2\lambda b^{\alpha} e^{lnb^{-\alpha}} + 3\lambda b^{2\alpha} e^{lnb^{-2\alpha}})$$

$$= \alpha b^{\alpha} b^{-\alpha} (1-2\lambda b^{\alpha} b^{-\alpha} + 3\lambda b^{2\alpha} b^{-2\alpha})$$

$$= \alpha(1+\lambda)\ Since\ \alpha > 0, |\lambda| \leq 1\ then\ \alpha(1+\lambda) \geq 0.$$

Similarly, substitute $x_{[i]j} = \infty$ in Eq. (7) we get, $f\left(x_{[i]j}\right) = 0$

$Proof\ of\ \int_{-\infty}^{\infty} f\left(x_{[i]j}\right)dx = 1$

$$\int_{\ln b}^{\infty} f\left(x_{[i]j};\alpha,b,\lambda\right)dx_{[i]j}$$

$$= \int_{\ln b}^{\infty} (\alpha b^{\alpha} e^{-\alpha x_{[i]j}} - 2\lambda\alpha b^{2\alpha} e^{-2\alpha x_{[i]j}} + 3\lambda\alpha b^{3\alpha} e^{-3\alpha x_{[i]j}}\ dx_{[i]j}$$

$$= \alpha b^{\alpha} \int_{\ln b}^{\infty} e^{-\alpha x_{[i]j}} dx_{[i]j} - 2\lambda\alpha b^{2\alpha} \int_{\ln b}^{\infty} e^{-2\alpha x_{[i]j}} dx_{[i]j} + 3\lambda\alpha b^{3\alpha} \int_{\ln b}^{\infty} e^{-3\alpha x_{[i]j}} dx_{[i]j}$$

$$= 1 - \lambda + \lambda - 1$$

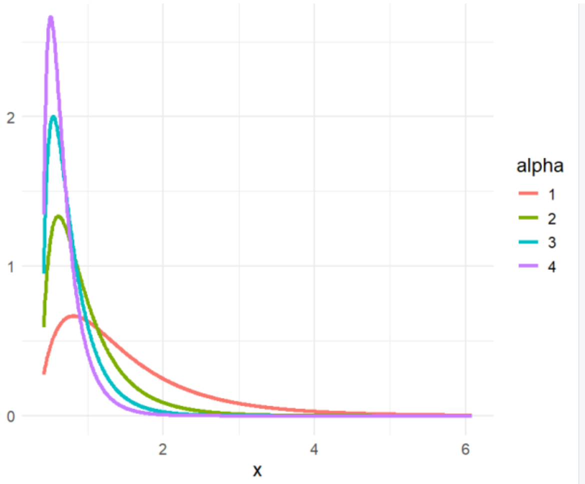

Fig. (1) The $f(x_{[i]j})$ of CTEP-I at α =1,2,3,4, λ = −0.75 and b = 1.5.

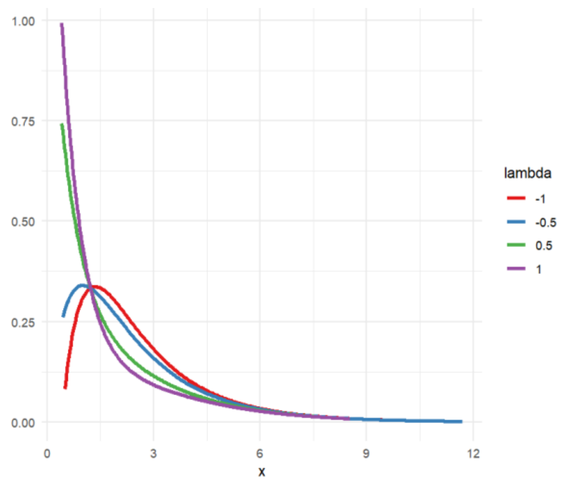

Fig. (2) The $f(x_{[i]j})$ of CTEP-I at λ = −1, −0.5,0.5, 1, α = 0.5 and b = 1.5

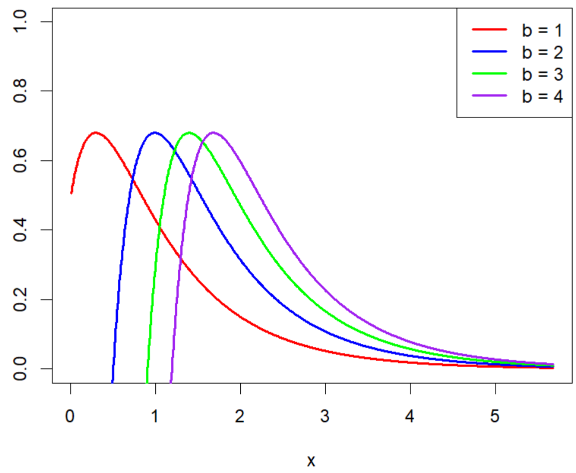

Fig. (3) The $f\ (x_{[i]j})$ of CTEP-I at λ = −0.5, α = 1 and b =1,2,3,4,5

1- Statistical Properties

The statistical characteristics of a generalized distribution are delineated, encompassing moments, moment-generating functions, geometric means, and harmonic means. The properties of the distribution were adopted from [9].

A. The moments

If the order statistics x is derived from (RSS) that is distributed according to a CTEP-I distribution, the moment of inertia of order r^{th} for x around the zero point is given by:

$$E(x_{[i]j}^r) = b^{\alpha}\alpha^{-r}\Gamma(r+1,\alpha\ln b) - \lambda b^{2\alpha}(2\alpha)^{-r}\Gamma(r+1,2\alpha\ln b) + \lambda b^{3\alpha}(3\alpha)^{-r}\Gamma(r+1,3\alpha\ln b) \quad (9)$$

B. Survival and Hazard Functions

If the order statistics x obtained via (]RSS) adheres to a CTEP-I distribution , then the survival (reliability) function for $x_{[i]j}$ is given by:

$$S(x) = 1 - F(x)$$

$$S(x_{[i]j}) = b^{\alpha}e^{-\alpha x_{[i]j}}\ (1 - \lambda b^{\alpha}e^{-\alpha x_{[i]j}} + \lambda b^{2\alpha}e^{-2\alpha x_{[i]j}})\ (10)$$

C. The moment-generating function

The moment-generating function for each random variable x is defined by the following relationship: $Mx_{(t)} = E(e^{tx})$

If the order statistics x derived from (RSS) is distributed according to a CTEP-I distribution, then the moment-generating function for $x_{[i]j}$ is given by:

$$Mx_{(t)} = \frac{\alpha b^t}{\alpha - t} - \frac{2\lambda\alpha b^t}{2\alpha - t} + \frac{3\lambda\alpha b^t}{3\alpha - t} \quad (11)$$

D. The Geometric mean

If the order statistics x derived from (RSS) is distributed according to a CTEP-I distribution, then the Geometric mean for $x_{[i]j}$ is given by:

$$\log G = \ln(\ln b) + b^{\alpha}\Gamma(0,\alpha\ln b) - \lambda b^{2\alpha}\left[\Gamma(0,2\alpha\ln b) + b^{\alpha}\Gamma(0,3\alpha\ln b)\right]$$
$$G = Anti\ \log\left(\ln(\ln b) + b^{\alpha}\Gamma(0,\alpha\ln b) - \lambda b^{2\alpha}\left[\Gamma(0,2\alpha\ln b) + b^{\alpha}\Gamma(0,3\alpha\ln b)\right]\right) \quad (12)$$

E. The Harmonic mean

If the random variable x derived from (RSS) is distributed according to a CTEP-I distribution, then the Harmonic mean for $x_{[i]j}$ is given by:

$$\frac{1}{H}=\alpha b^{\alpha}\Gamma(0,\alpha\ln b)-2\lambda\alpha b^{2\alpha}\Gamma(0,2\alpha\ln b)+3\lambda\alpha b^{3\alpha}\Gamma(0,\alpha\ln b)$$

$$H=(\alpha b^{\alpha}\Gamma(0,\alpha\ln b)-2\lambda\alpha b^{2\alpha}\Gamma(0,2\alpha\ln b)+3\lambda\alpha b^{3\alpha}\Gamma(0,\alpha\ln b))^{-1} \quad (13)$$

6th: PARAMETERS ESTIMATION

The distribution parameters have been estimated using the maximum likelihood (ML) method, an important method for statistical parameter estimation.
It relies on selecting values that optimize the likelihood of the observed data manifesting.
Suppose that $x_{[1]1},\ldots x_{[i]j}$ is order Statistics derived from (RSS) is distributed according to a CTEP-I distribution, The likelihood function is given by:

$$L=\prod_{r=1}^{n}(\alpha b^{\alpha}e^{-\alpha x_{[i]j}}(1-2\lambda b^{\alpha}e^{-\alpha x_{[i]j}}+3\lambda b^{2\alpha}e^{-2\alpha x_{[i]j}}))$$

$$\ln L=n\ln\alpha+n\alpha\ln b-\alpha\sum_{r=1}^{n}x_{[i]jr}+\sum_{r=1}^{n}\ln(1-2\lambda b^{\alpha}e^{-\alpha x_{[i]jr}}+3\lambda b^{2\alpha}e^{-2\alpha x_{[i]jr}})$$

$$\frac{d\ln L}{d\alpha}=\frac{n}{\alpha}+n\ln b-\sum_{r=1}^{n}x_{[i]jr}+\sum_{r=1}^{n}\frac{(-2\lambda b^{\alpha}\ln b e^{-\alpha x_{[i]jr}}+2\lambda b^{\alpha}x_{[i]jr}e^{-\alpha x_{[i]jr}}+6\lambda b^{2\alpha}\ln b e^{-2\alpha x_{[i]jr}}-6\lambda b^{2\alpha}x_{[i]jr}e^{-2\alpha x_{[i]jr}})}{1-2\lambda b^{\alpha}e^{-\alpha x_{[i]jr}}+3\lambda b^{2\alpha}e^{-2\alpha x_{[i]jr}}}=0 \quad (14)$$

$$\frac{d\ln L}{db}=\frac{n\alpha}{b}+\sum_{r=1}^{n}\frac{(-\frac{2\lambda b^{\alpha}\alpha e^{-\alpha x_{[i]jr}}}{b}+\frac{6\lambda b^{2\alpha}\alpha e^{-2\alpha x_{[i]jr}}}{b})}{1-2\lambda b^{\alpha}e^{-\alpha x_{[i]jr}}+3\lambda b^{2\alpha}e^{-2\alpha x_{[i]jr}}}=0 \quad (15)$$

$$\frac{d\ln L}{d\lambda}=\sum_{r=1}^{n}\frac{-2b^{\alpha}e^{-\alpha x_{[i]jr}}+3b^{2\alpha}e^{-2\alpha x_{[i]jr}}}{1-2\lambda b^{\alpha}e^{-\alpha x_{[i]jr}}+3\lambda b^{2\alpha}e^{-2\alpha x_{[i]jr}}}=0 \quad (16)$$

The estimated values of the three parameters α, b and λ are found by solving the nonlinear system of equations (14), (15) and (16) using one of the numerical methods such as the Quazi-Newton or Newton-Raphson to numerically maximize the log likely-hood function.

7th: SIMULATION

Simulation is a technique employed to depict the behave aur of a realistic system or process by constructing a model analogous to the actual one and using a mathematical or computational framework to analyze its performance under varying conditions.
A simulation model constructed using realistic equations or data is developed to represent the system, followed by the execution of virtual experiments on it, with the findings analyzed and extrapolated to the original model.
We employed a Monte Carlo simulation method to generate generalized distribution CTEP-I tracking data using the inverse generation technique. A dataset comprising 2500 views was created, from which a ranked sample of 500 was extracted. The number of sets, k, was set to (10,6) and the number of cycles, r, was established at (50,70,80,30). This sample was later used to estimate the parameters of the generalized distribution.

Table (1): shows the effect of different rank sample sizes on the estimated values of parameters, bias, variance, and mean square error (MSE) by using Simulation.

When k=10, r=50				
Parameter estimates	**Mean estimators**	**Bias**	**Variance**	**MSE**
α	0.01233786	-0.9876621	0.002966205	0.9784427
B	0.12319190	-0.8768081	0.087137936	0.8559304
λ	-0.07501035	-1.0750103	0.041206580	1.1968538
When k=10, r=80				
Parameter estimates	**Mean estimators**	**Bias**	**Variance**	**MSE**
α	0.002443017	-0.9975570	4.451429e-05	0.9951644
B	0.052768214	-0.9472318	3.059120e-02	0.9278393
λ	-0.084143784	-1.0841438	1.139386e-02	1.1867616
When k=10, r=70				
Parameter estimates	**Mean estimators**	**Bias**	**Variance**	**MSE**
α	0.003452228	-0.9965478	6.621642e-05	0.9931737
B	0.080377583	-0.9196224	4.611385e-02	0.8918192
λ	-0.081181863	-1.0811819	2.495134e-02	1.1939056
When k=6, r=30				
Parameter estimates	**Mean estimators**	**Bias**	**Variance**	**MSE**
α	0.02945191	-0.9705481	0.006384275	0.9483479
B	0.55501818	-0.4449818	4.817341295	5.0153501
λ	-0.04023584	-1.0402358	0.017835399	1.0999260

The best method of parameter estimation was found to be the likelihood function (ML) method, The simulation was run 500 times, and effect of different ranked sample sizes and how the number of cycles affects the metrics in the table on the estimated values of parameters, bias, variance, and in addition to the error metric known as MSE (Average Squared Error), as presented in Table 1.

Table (2): Effect of different alpha values on some measures of central tendency and skewness with λ =-0.75 and b=1.5

α	**Mean**	**Median**	**Mode**	**Skewness**
1	1.552664	1.308411	0.913236	1.599662
2	1.552664	1.308411	0.913236	1.599662
3	1.552664	1.308411	0.913236	1.599662
4	1.552664	1.308411	0.913236	1.599662

Table (3): Effect of different lambda values on some measures of central tendency and skewness with α =0.5 and b=1.5

λ	Mean	Median	Mode	Skewness
-0.5	2.642150	2.300145	1.26	1.504346
-1.5	2.512058	1.969116	0.71	2.078333
'0.5	2.225060	1.565448	0.49	1.918985
1	1.977567	1.208913	0.43	1.998643

The effect of the parameters α and λ on the model was studied as four different initial values were taken for each parameter and some measures of central tendency and skewness were calculated as shown in Tables 2, and 3.
The findings presented in Table 2 show that as we increase the shape parameter, the skewness of the distribution is fixed, while increasing the transformation parameter causes the skewness to increase, as indicated in Table 3. This effect can be seen by plotting the PDF (distribution function) associated with the distribution in Figures 1 and 2.

8th: APPLICATIONS

This section presents an application of the suggested distribution, CTEP-1, utilizing various sample sizes, where ordered sets of differing sizes are employed to estimate the parameters of the generalized distribution CTEP-1 through ranked set sampling. We additionally compare the present model with alternative models, including the exponential distribution, exponentiated Pareto-I(EP-I).

Table (4): Kevlar49/epoxy strand failure time data (pressureat90%).

0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.04
0.05	0.06	0.07	0.07	0.08	0.09	0.09	0.10
0.10	0.11	0.11	0.12	0.13	0.18	0.19	0.20
0.23	0.24	0.24	0.29	0.34	0.35	0.36	0.38
0.40	0.42	0.43	0.52	0.54	0.56	0.60	0.60
0.63	0.65	0.67	0.68	0.72	0.72	0.72	0.73
0.79	0.79	0.80	0.80	0.83	0.85	0.90	0.92
0.95	0.99	1.00	1.01	1.02	1.03	1.05	1.10
1.10	1.11	1.15	1.18	1.20	1.29	1.31	1.33
1.34	1.40	1.43	1.45	1.50	1.51	1.52	1.53
1.54	1.54	1.55	1.58	1.60	1.63	1.64	1.80
1.80	1.81	2.02	2.05	2.14	2.17	2.33	3.03
3.03	3.34	4.20	4.69				

Table 4, as published by Barlow et al [12], presents the failure time of Kevlar49/epoxy strands under a pressure corresponding to 90% of the stress level (FTK). The data exhibits leptokurtic characteristics, is unimodal, and demonstrates right skewness (skewness = 1.626, kurtosis = 6.699), as analyzed by Andrews and Herzberg [13], Cordeiro and Lemonte [14], and Al-Aqtash et al [15].

Table (5): ML parameter estimates, models election criteria for the ranked sample drawn from FTK data set.

Distribution	Parameter estimates	−2log(L)	AIC	BIC
Exponential	$\lambda = 0.2088847$	107.0096	109.0096	111.5094
EP-I	$\alpha = 1.015182$	106.6627	110.6627	115.6623
CTEP-I(proposed)	$\alpha = 0.01283322$ $\lambda = 0.02314449$	21.6385	27.6385	35.13793

Table 5 presents the outcomes from applying the fit to this data to the exponential distribution and the EP-I distribution. proposed allocation, the CTEP-I.
According to -2log(L), Akaike's information criterion (AIC), and Schwarz's Bayesian information criterion (BIC), the proposed distribution, CTEP-I, outperforms the alternatives. The present instance demonstrates that the CTEP-I distribution is well-suited for leptokurtic, unimodal, and right-skewed data. Where AIC=−2log(L)+2k and BIC=−2log(L)+k ln n; k signifies the number of estimated parameters, n indicates the sample size, and log(L) indicates the negative log likelihood of the data.

9th: CONCLUSIONS

This paper introduces a novel, generalized version of the exponentiated Pareto-I distribution, termed the cubic transmuted exponentiated Pareto-I (CTEP-I) distribution, utilizing rank sample (RSS) and the generalization formula for transmuted distributions proposed by Al-Kadim (2018). Certain statistical characteristics of the distribution are obtained. The model parameters are determined using the greatest likelihood method. The simulation revealed that having a well-organized sample (RSS) impacts certain statistical measures, showing that a bigger sample size makes the model more reliable and results in better estimates, less variation, and a lower average squared error (MSE). More cycles of r' signify additional iterations of various r values, hence enhancing the stability of the statistical estimates. As r grows, a larger and more diverse sample is obtained while preserving the rank r within each group, hence improving the quality of the estimation. In general, increasing the size of the ordered sample (RSS) by adding more groups (k), more cycles (r), or both greatly improves the results of statistical estimates because the total size of the ordered sample is calculated as n=k*r.
This study presents an application of the CTEP-I distribution to empirical data using a ranked set sample (RSS), and compares it with both the exponential and the exponentiated Pareto-I distributions. we determined that the proposed distribution, CTEP-1, adequately fits the ranked sampling drawn from the unimodal leptokurtic and right-skewed data, in addition to the unimodal mesokurtic data. we recommend CTEP-1 for modelling right-skewed data and datasets exhibiting constant, inverted, anticipating substantial future applications.

References

1- Al-Aqtash, R., C. Lee, and F. Famoye, Gumbel-Weibull distribution: Properties and applications. Journal of Modern applied statistical methods, 2014. 13(2): p. 11.
2- AL-Kadim, K.A., Proposed Generalized Formula for Transmuted Distribution. Journal of University of Babylon for Pure and Applied Sciences, 2018. 26(4): p. 66-74.
3- Andrews, D., et al., Stress-rupture life of kevlar 49/epoxy spherical pressure vessels. Data: A Collection of Problems from Many Fields for the Student and Research Worker, 1985: p. 181-186.
4- Andrews, D.F. and A.M. Herzberg, Data: a collection of problems from many fields for the student and research worker. 2012: Springer Science & Business Media.
5- Cordeiro, G.M. and A.J. Lemonte, The β-Birnbaum–Saunders distribution: An improved distribution for fatigue life modeling. Computational statistics & data analysis, 2011. 55(3): p. 1445-1461.
6- Eledum, H. and S. Ansari, A Generalization Of Exponentiated Pareto-I Distribution With Applications. Journal of Applied Science and Engineering, 2024. 27(5): p. 2471-2481.
7- Eledum, H. and S. I Ansari, MG Exponentiated Pareto Distribution. Journal of Statistics Applications & Probability, 2020. 9(3): p. 507-514.
8- Eledum, H., Some cubic transmuted exponentiated pareto-1 distribution. Journal of Mathematics and Statistics, 2020. 16: p. 113.124.
9- Gupta, R.C., P.L. Gupta, and R.D. Gupta, Modeling failure time data by Lehman alternatives. Communications in Statistics-Theory and methods, 1998. 27(4): p. 887-904.
10- McIntyre, G., A method for unbiased selective sampling, using ranked sets. Australian journal of agricultural research, 1952. 3(4): p. 385-390.
11- Patil, G.P., A. Sinha, and C. Taillie, 5 Ranked set sampling. Handbook of statistics, 1994. 12: p. 167-200.
12- Rahman, M.M., B. Al-Zahrani, and M.Q. Shahbaz, A general transmuted family of distributions. Pakistan Journal of Statistics and Operation Research, 2018: p. 451-469.
13- Rahman, M.M., B. Al-Zahrani, and M.Q. Shahbaz, Cubic transmuted Pareto distribution. Annals of Data Science, 2020. 7(1): p. 91-108.
14- Wolfe, D.A., Ranked set sampling: its relevance and impact on statistical inference. International scholarly research notices, 2012. 2012(1): p. 568385.
15- Zamanzade, E. and A.I. Al-Omari, New ranked set sampling for estimating the population mean and variance. Hacettepe Journal of Mathematics and Statistics, 2016. 45(6): p. 1891-1905.