Computational Journal of Mathematical and Statistical Sciences
2(2), 303–327
DOI:10.21608/CJMSS.2023.240287.1020
https://cjmss.journals.ekb.eg/

Discrete Exponentiated Generalized Family of Distributions

A. E., Abd EL-Hady[1], M. A., Hegazy[1] and A. A., EL-Helbawy[2] *

[1] Department of Statistics, Faculty of Commerce, AL-Azhar University (Girls' Branch), Tafahna Al-Ashraf, Egypt
[2] Department of Statistics, Faculty of Commerce, AL-Azhar University (Girls' Branch), Cairo, Egypt

* **Correspondence:** aah_elhelbawy@hotmail.com

Abstract: In this paper, exponentiated-G family of distributions is constructed as a new family of discrete distribution, using the general approach of discretization of a continuous distribution. Some statistical properties including quantile, mean residual life, mean time to failure, mean time between failure, availability, R*é*nyi entropy, moments and order statistics are obtained. Discrete exponentiated inverted Topp–Leone distribution as a member of this family is studied in detail. Maximum likelihood approach is applied under Type-II censored sample for estimating the unknown parameters of the exponentiated inverted Topp–Leone distribution. Also, maximum likelihood estimators of the survival, hazard rate and alternative hazard rate functions are derived. The confidence intervals for the parameters, survival, hazard rate and alternative hazard rate functions are obtained. A simulation study is performed to investigate the accuracy of the theoretical results. Finally, two real data sets are analyzed to illustrate the flexibility and applicability of the proposed model for real-life applications.

Keywords: Survival discretization, Exponentiated-G family, Order statistic, Maximum likelihood.
Mathematics Subject Classification: 94A20; 20K21; 90C56; 78M05.
Received: 2 October 2023; Revised: 19 October 2023; Accepted: 28 October 2023; Online: 31 October 2023.

1. Introduction

Discretization generally arises when it is difficult to measure the life length of a product or device on a continuous scale. Such situations may arise when it is difficult to get samples from a continuous distribution in real life. The observed values are discrete because they are usually measured to only a finite number of decimal places and can't really constitute all points in a continuum. Even if the measures are taken on a continuous (ratio or interval) scale, the observations may be recorded in a way making discrete model more appropriate. Therefore, it is reasonable to consider the observations as coming from a discretized distribution generated from the continuous model. For example, in survival analysis, the survival times for leukemia or lung cancer patients survived since therapy, length of stay in observation ward or the period from remission to relapse may be recorded as a number of days or weeks. Also, in engineering systems, the number of voltages fluctuations, which an electrical or electronic item can withstand before its failure, the number of completed cycles or the number of times it operated before failure, the life of weapon is measured by the number of rounds fired prior to failure, or the number of times a device is switched on/off. Therefore, it is reasonable and appropriate to model such situations by a suitable discrete distribution. Although, classical discrete distribution is accessible to model such situations, for example, geometric and negative binomial distributions which are the discrete versions for the exponential and gamma distributions, respectively, but it is well known that they have monotonic hazard rate function (hrf) and thus they are unsuitable for some situations.

There are few discrete distributions which can provide accurate models for both count and time. Poisson distribution is a model used to count but not time. Also, binomial distribution is not considered to be popular model for reliability, failure times and counts. It can be approximated to Poisson distribution under suitable conditions. In addition to that, these discrete distributions only cater to positive integers along with zero, but in some analysis the variable of interest can take either zero, positive or negative values. In many situations the interest may be in the difference of two discrete random variables (drvs) each having integer support $(0, \infty)$. The resulting difference will be another drv with integer support $(-\infty, \infty)$, see Chakraborty and Chakravorty [10]. Thus, there is a need to introduce more flexible discrete distributions, especially those arising from the discretization of continuous distributions to handle more sophisticated real-life phenomena.

2. General Approach of Discretization

There are several techniques to construct discrete distributions from the continuous ones, among these is the survival discretization method. One of the advantages of applying this methodology is that the developed discrete model keeps the same form of the survival function (sf) as of the sf for the continuous counterpart distribution. Therefore, many reliability characteristics of the distribution remain unchanged. Thus, this approach of discretization of continuous lifetime model is interesting and simple approach to derive a discrete lifetime model corresponding to the continuous one.

If the underlying continuous failure time X has the sf, $S(x) = P(X \geq x)$ and times are grouped into unit intervals so that the discrete observed variable is discrete X $(dX) = \lfloor X \rfloor$, the largest integer is less than or equal to X, the probability mass function (pmf) of dX can be written as

$$\begin{aligned} P(x) &= P(dX = x) = P(x \leq dX < x + 1) \\ &= S(x) - S(x+1), x = 0, 1, 2, \ldots. \end{aligned} \tag{1}$$

Therefore, for any continuous distribution, it is possible to construct corresponding discrete distribution using (1).

Many researchers used this approach to develop several discrete distributions. For instance, Nakagawa and Osaki [37] introduced the discrete Weibull distribution with some important reliability properties. Kemp [31] presented the discrete normal distribution that is characterized by maximum entropy specified mean and variance and integer support on $(-\infty, \infty)$.

Roy [42] introduced discrete normal distribution, Inusah and Kozubowski [29] obtained a discrete version of the Laplace distribution and discussed some of its properties and statistical issues of estimation. Also, Gomez-Deniz and Calderin-Ojeda [20] considered discrete Lindley distribution. AL-Huniti and AL-Dayian [4] presented the discrete Burr Type III distribution; they discussed some important properties and estimated the parameters based on the maximum likelihood (ML) and Bayesian approaches. Para and Jan [38] derived the discrete Burr-Type XII distribution. Alamatsaz et al. [3] proposed the discrete generalized Rayleigh distribution, and the reliability properties were discussed, they also estimated the parameters and analyzed two real data sets to investigate the suitability of the distribution in modeling count data.

Hussain et al. [27] discussed a two-parameter discrete Lindley distribution which is considered as a new generalization of the geometric distribution. Chakraborty and Chakravorty [10] derived the discrete logistic distribution and applied it to model real life count data. Ahmad et al. [2] presented a discrete flexible model: Geeta- Kumaraswamy distribution. They discussed some important properties and estimated the parameters based on the ML criteria.

Hegazy et al. [23] presented the discrete Gompertz distribution, they discussed some statically properties of the distribution and estimated the parameters based on the ML method. Maiti et al. [34] proposed the discrete X-Gamma distribution. Helmy [26] presented the discrete Burr Type II distribution; she discussed some important properties and estimated the parameters based on the ML and Bayesian approaches.

Hegazy et al. [24] introduced the discrete inverted Kumaraswamy distribution. They obtained some of its important distributional and reliability properties; they also used the moments and ML methods to estimate the model parameters. Elmorshedy and Eliwa [17] presented a new two-parameter exponentiated discrete Lindley distribution; they discussed some statistical properties of distribution. They used the ML method to estimate the parameters of the distribution. Almetwally and Ibrahim [5] proposed the discrete alpha power inverse Lomax distribution with Application of COVID-19 data and discussed some statistical properties of the distribution. They derived the ML estimators and confidence intervals for the distribution. Eliwa et al. [14] introduced discrete Gompertz-G Family of distributions for over-and under-dispersed data, they also studied some of its distributional properties and reliability characteristics. They also used the ML method for estimating the family parameters. Chotedelok and Bodhisuwan [11] obtained the discrete exponentiated Pareto distribution with properties and application. Tyagi et al. [43] presented inferences on discrete Rayleigh distribution under Type-II censored data.

Hegazy et al. [25] introduced the Bayesian estimation and prediction of discrete Half Logistic distribution. Ibrahim et al. [28] proposed the discrete analogue of the Weibull-G family. They studied its properties and estimation of the parameters using Bayesian and non- Bayesian estimation methods. El-Deep et al. [13] presented the discrete analog of inverted Topp-Leone Distribution. El-Morshedy et al. [18] introduced the discrete odd Weibull -G family of distributions, they studied the discrete odd Weibull – Geometric and discrete odd Weibull – inverse Weibull as special models in detail. Eliwa et al. [15] derived the discrete exponential generalized G family of distributions, they studied some of its important properties and the properties of the discrete exponential generalized Weibull. Altun et al. [7] obtained a new one parameter called discrete Bilal distribution, they studied its statistical properties and estimated the model parameter by using the ML and moment methods. Eliwa et al. [16] constructed new continuous and discrete odd DAL-G family of distributions, and they studied some properties of the special models called the new odd DAL-Weibull and discrete new odd DAL-geometric distributions.

Although there are several discrete distributions in statistical literature, there is still a lot of space left to develop new discrete distributions that are suitable under different conditions. Therefore, in this paper, a flexible discrete generator of distributions, called discrete *E-G* (DE-G) family of distributions is introduced. In practice the motivation for obtaining generalized family are:

1. Generating a new family with different shapes of hrf.
2. Presenting a new family with different shapes of pmf.
3. To provide consistently better fits than other generated models under the same baseline distribution and other well-known models in the statistical literature.
4. To improve the characteristics and flexibility of the existing distributions.
5. To introduce the extended version of the baseline distribution having closed form of cumulative density function (cdf), sf as well as hrf.

There is an increasing interest in constructing new generated families of univariate continuous distributions by adding additional shape parameter(s) to a baseline model due to the desirable properties of the new models. Such an addition of parameters makes the resulting distribution richer and more flexible for modeling data. Some of the well- known generated families are: exponentiated- G by Gupta et al. [21], beta–G by Eugene et al. [19], Nadarajah [36] discussed the exponentiated Gumbel distribution with application. Kumaraswamy-G by Cordeiro and de Castro [12]. Alzaghal et al. [8] derived and studied the exponentiated T-X family of distributions with applications. Korkmaz et al. [32] studied a family of distributions called the exponential Lindely odd log-logistic-G family. Roa and Mbwambo [40] derived the exponentiated inverse Rayleigh distribution.

Gupta et al. [21] introduced the general class of exponentiated distributions which is de?ned by powering a positive real number α to the cdf, i.e., if the random variable X has the cdf $G(x)$, then

$$F(x;\alpha) = [G(x)]^{\alpha}, \qquad \alpha>0, \qquad x\epsilon\mathbb{R}. \tag{2}$$

The probability density function (pdf), and sf, of the exponentiated-G (E-G) family are, respectively, given by

$$f(x;\alpha) = \alpha g(x)[G(x)]^{\alpha-1}, \quad \alpha>0, \qquad x\epsilon\mathbb{R}, \tag{3}$$

and

$$S(x;\alpha) = 1 - [G(x)]^{\alpha}, \qquad \alpha>0, \qquad x\epsilon\mathbb{R}. \tag{4}$$

The rest of this paper is organized as follows: in Section 3 , a discrete exponentiated-G (DE-G) family of distributions and some of its properties are studied. In Section 4, a member of the discrete G-family of distributions is presented. The ML estimation for the parameters of the distribution is discussed in Section 5. In Section 6, two real data sets are analyzed to demonstrate how the results can be used in practice. Finally, concluding remarks are given in Section 7.

3. Construction of Discrete Exponentiated- G Family of Distributions

Applying the survival discretization approach in (1), $X = dX$ can be viewed as the discrete analogue to the continuous E-G family random variable X, and is commonly said to have DE-G family distribution with parameter α, denoted by DE-G family (α) distribution, where the corresponding pmf of X can be written as:

$$P(x;\alpha) = [G(x+1)]^{\alpha} - [G(x)]^{\alpha}, \qquad x = 0,\ 1,2,\ldots \qquad \alpha>0, \tag{5}$$

the cdf, sf, hrf and alternative hrf (ahrf) can be formulated as:

$$F(x;\alpha) = [G(x+1)]^{\alpha}, \qquad x = 0,1,2,\ldots \quad \alpha>0, \tag{6}$$

$$S(x;\alpha) = 1 - [G(x)]^{\alpha}, \qquad x = 0,1,2,\ldots \quad \alpha>0, \tag{7}$$

$$h(x;\alpha) = \frac{[G(x+1)]^{\alpha} - [G(x)]^{\alpha}}{1-[G(x)]^{\alpha}}, \qquad x = 0,1,2,\ldots \quad \alpha>0, \tag{8}$$

and

$$h_1(x;\alpha) = ln\left[\frac{1-[G(x)]^{\alpha}}{1-[G(x+1)]^{\alpha}}\right], \qquad x = 0,1,2,\ldots \quad \alpha>0. \tag{9}$$

3.1. Some statistical properties of discrete exponentiated- G family of distributions

This subsection is devoted to obtain some important statistical properties of DE-G family of distributions, such as the Quantile, r^{th} moments and order statistics.

3.1.1. Quantile function

The quantile function plays an important role in probability distribution theory, and it is also known as the inverse cdf. It is used to generate random numbers from a given distribution. The u^{th}quantile of a drv X, x_u, satisfies $p(X \le x_u) \ge u$ and $p(X \ge x_u) \ge 1-u$, i.e.,
$F(x_u - 1) < u \le F(x_u)$. [See Rohatgi and Saleh [41]].
The u^{th} quantile, x_u, of DE -G family is given by
$p(X \le x_u) \ge u$ from (8)

$$u = [G(x+1)]^{\alpha}, u^{\frac{1}{\alpha}} = G(x+1), 0 < u < 1\ . \tag{10}$$

3.1.2. Rényi entropy

Entropy refers to the amount of uncertainty associated with the random variable X. It has many applications in several fields such as econometrics, quantum information; information theory, survival analysis, and computer science [see Rényi (1961)].
The Rényi entropy which is a measure of variation of the uncertainty of the drv X can be expressed as

$$\begin{aligned} I_{\eta}(x) &= \frac{1}{1-\eta} log \sum_x f^{\eta}(x;\xi). \qquad x = 0,1,2,\ldots \\ &= \frac{1}{1-\eta} log \sum_x [[G(x+1)]^{\alpha} - [G(x)]^{\alpha}]^{\eta}. \quad x = 0,1,2,\ldots. \end{aligned} \tag{11}$$

The Shannon entropy can be defined by

$$I(X) = E\left[-log f(x_i;\xi)\right] = -log f(x_i;\xi)\sum_x f(x_i;\xi)$$
$$= -log\left[[G(x+1)]^{\alpha} - [G(x)]^{\alpha}]\right]\sum_x \left[[G(x+1)]^{\alpha} - [G(x)]^{\alpha}\right], \qquad x = 0,1,2,\ldots. \tag{12}$$

The Shannon entropy can be derived as a particular case of the Rényi entropy when $\eta \to 1$.

3.1.3. Mean residual lifetime function, mean time to failure, mean time between failure, and availability

The m ean residual life (MRL) function is an important characteristic in various fields such as reliability engineering, survival analysis, and actuarial studies. It has been extensively studied in the literature especially for binary systems, that is, when there are only two possible states for the system as either working or failed.
The MRL is the expected remaining life given that the item has survived to time x_0 [see, Kemp [30]] and is defined by

$$MRL(x_0) = \frac{\sum_{k=x_0+1}^{\infty} s(k)}{s(x_0)} = \frac{1}{\{1-[G(x_0)]^{\alpha}\}}\sum_{k=x_0+1}^{\infty}\{1-[G(k)]^{\alpha}\} \tag{13}$$

Mean Time to Failure (MTTF) , Mean Time between Failure (MTBF) , and Availability (Av) are reliability terms based on methods and procedures for lifecycle prediction for a product. MTTF, MTBF and Av are ways of providing a numeric value based on a compilation of data to quantify a failure rate and the resulting time of expected performance. In addition, in request to design and manufacture a maintainable system, it is necessary to predict the MTTF, MTBF, and Av. [see Eliwa et al. [14]].
The MTBF is given by

$$MTBF = \frac{-x}{\log[S(x)]} = \frac{-x}{\log[1-[G(x)]^{\alpha}]}, \qquad x > 0. \tag{14}$$

Then the MTTF is

$$MTTF = \sum_{x=1}^{\infty} S(x) = \sum_{x=1}^{\infty}\left[1-[G(x)]^{\alpha}\right], \quad x>0. \tag{15}$$

The Av is considered as being the probability that the component is successful at time t,

$$\text{Av} = \frac{MTTF}{MTBF}. \tag{16}$$

3.2. *Order statistics*

Order statistics play an important role in various fields of statistical theory and practice. The cdf of the ith order statistic for a random sample $X_1, X_2, \ldots, X_n$, from DE-G family of distributions is given by

$$F_{i:n}(x;\theta,\alpha) = \sum_{r=i}^{n}\binom{n}{r} F\left[(x;\theta,\alpha)\right]^{r}\left[1-F(x;\theta,\alpha)\right]^{n-r}, \tag{17}$$

Using the binomial expansion for $[1 - F(x;\theta,\alpha)]^{n-r}$ and substituting (8) in (17) it follows that

$$F_{i:n}(x;\theta,\alpha) = \sum_{r=i}^{n} \binom{n}{r} [F(x;\theta,\alpha)]^{r} \sum_{j=0}^{n-r} \binom{n-r}{j} (-1)^{j} [F(x;\theta,\alpha)]^{j}.$$

Hence

$$F_{i:n}(x;\theta,\alpha) = \sum_{r=i}^{n} \binom{n}{r} \sum_{j=0}^{n-r} \binom{n-r}{j} (-1)^{j} \left[[G(x+1)]\right]^{\alpha(r+j)}. \tag{18}$$

The corresponding pmf of the i^{th} order statistics can be expressed as

$$P(X_{(i)} = x) = \frac{n!}{(i-1)!(n-i)!} \int_{F(x-1)}^{F(x)} v^{i-1}(1-v)^{n-i} dv, \tag{19}$$

where v is a random variable.
Using the binomial expansion for $(1-v)^{n-i}$, then, the pmf of (19) is

$$\begin{aligned} P(X_{(i)} = x) &= \frac{n!}{(i-1)!(n-i)!} \sum_{j=0}^{n-i} \binom{n-i}{j} (-1)^{j} \int_{F(x-1)}^{F(x)} v^{i+j-1} dv \\ &= \frac{n!}{(i-1)!(n-i)!} \sum_{j=0}^{n-i} \binom{n-i}{j} (-1)^{j} \left(\frac{1}{i+j}\right) \\ &\times \left[[1-G(x+1)]^{\alpha(r+j)} - [1-G(x)]^{\alpha(r+j)}\right]. \end{aligned} \tag{20}$$

See Arnold et al. [9].

4. Discrete Exponentiated Inverted Topp – Leone Distribution

The inverted distributions have many applications in different fields such as biological science, life testing problems, engineering science environmental studies and econometrics.
Hassan and Elgarhy [22] introduced the Inverted Topp – Leone (ITL) distribution with the following cdf and pdf

$$F_{ITL}(x;\theta) = 1 - \left\{\frac{(1+2x)^{\theta}}{(1+x)^{2\theta}}\right\}; \quad x \geq 0; \ \theta > 0, \tag{21}$$

$$f_{ITL}(x,\theta) = 2\theta x(1+x)^{-2\theta-1}(1+2x)^{\theta-1}, \ x \geq 0; \quad \theta > 0, \tag{22}$$

and the sf of the ITL is

$$S(x,\theta) = (1+x)^{-2\theta}(1+2x)^{\theta}, \quad x \geq 0; \ \theta > 0. \tag{23}$$

Substituting the cdf of the ITL given in (21); as the baseline cdf, in (5). Then, the pmf of the DE-ITL distribution can be expressed as:

$$P(x;\theta,\alpha) = \left[1 - \frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{\alpha} - \left[1 - \frac{(1+2x)^{\theta}}{(1+x)^{2\theta}}\right]^{\alpha}. \ x = 0,1,2,.., \quad \alpha,\theta{>}0, \tag{24}$$

the cdf, sf, hrf and ahrf of the DE-ITL distribution are given by

$$F\left(x;\theta,\alpha\right)=\left[1-\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{\alpha},\qquad x=0,1,2,\ldots,\qquad \alpha,\theta>0, \tag{25}$$

$$S\left(x;\theta,\alpha\right)=1-\left[1-(1+2x)^{\theta}(1+x)^{-2\theta}\right]^{\alpha},\quad x=0,1,2,\ldots,\ \alpha,\theta>0, \tag{26}$$

$$h\left(x;\theta,\alpha\right)=\frac{\left[1-\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{\alpha}-\left[1-\frac{(1+2x)^{\theta}}{(1+x)^{2\theta}}\right]^{\alpha}}{1-\left[1-(1+2x)^{\theta}(1+x)^{-2\theta}\right]^{\alpha}}\ ,\qquad x=0,1,2,\ldots,\quad \alpha,\theta>0, \tag{27}$$

and

$$h_1\left(x;\theta,\alpha\right)=ln\left[\frac{1-\left[1-(1+2x)^{\theta}(1+x)^{-2\theta}\right]^{\alpha}}{1-\left[1-(1+2(x+1))^{\theta}(1+(x+1))^{-2\theta}\right]^{\alpha}}\right],\quad x=0,1,2,\ldots,\quad \alpha,\theta>0. \tag{28}$$

The relationship between $h_1(x)$ and $h(x)$, is given by:

$$h\left(x\right)=1-e^{-h_1(x)}.$$

The two concepts $h(x)$ and $h_1(x)$ have the same monotonic property, i.e., $h_1(x)$ is increasing (decreasing) only and only if $h(x)$ is increasing (decreasing).
The plots of the pmf, hrf and ahrf of the DE-ITL distribution are presented, respectively, in Figures 1, 2, 3, for some selected values of the parameters.

4.1. Some distributional properties of discrete exponentiated inverted Topp -Leone distribution

This subsection focuses on obtaining some important statistical properties of DE-ITL distribution, such as the Quantile, r^{th} moments and order statistics.

4.1.1. Quantile function

The u^{th} quantile x_u, of the DE-ITL distribution is given by

$$x_u=\left\lfloor\left(\frac{1}{\left(1-\left(1-u^{\frac{1}{\alpha}}\right)^{\frac{1}{\theta}}\right)^{\frac{-1}{2}}-1}\right)-1\right\rfloor,\qquad 0<u<1,$$

where $\lfloor X\rfloor$denotes the smallest integer greater than or equal to X.

Proof

$p\left(X\leq x_u\right)\geq u$, from Equation (25), one gets

$$\left[1-\frac{(1+2(x_u+1))^{\theta}}{(1+(x_u+1))^{2\theta}}\right]^{\alpha}\geq u,$$

$$\left[1-\frac{(1+2(x_u+1))^{\theta}}{(1+(x_u+1))^{2\theta}}\right]\geq u^{\frac{1}{\alpha}},\left(1-u^{\frac{1}{\alpha}}\right)\geq\frac{(1+2(x_u+1))^{\theta}}{(1+(x_u+1))^{2\theta}}$$

$$\left(1-u^{\frac{1}{\alpha}}\right)^{\frac{1}{\theta}}\geq\frac{(1+2(x_u+1))}{(1+(x_u+1))^{2}}\ ,\qquad \left(1-u^{\frac{1}{\alpha}}\right)^{\frac{1}{\theta}}\geq\frac{(1+2(x_u+1))+(x_u+1)^2-(x_u+1)^2}{(1+(x_u+1))^{2}},$$

$$\left(1-u^{\frac{1}{\alpha}}\right)^{\frac{1}{\theta}}\geq 1-\frac{(x_u+1)^2}{(1+(x_u+1))^{2}},$$

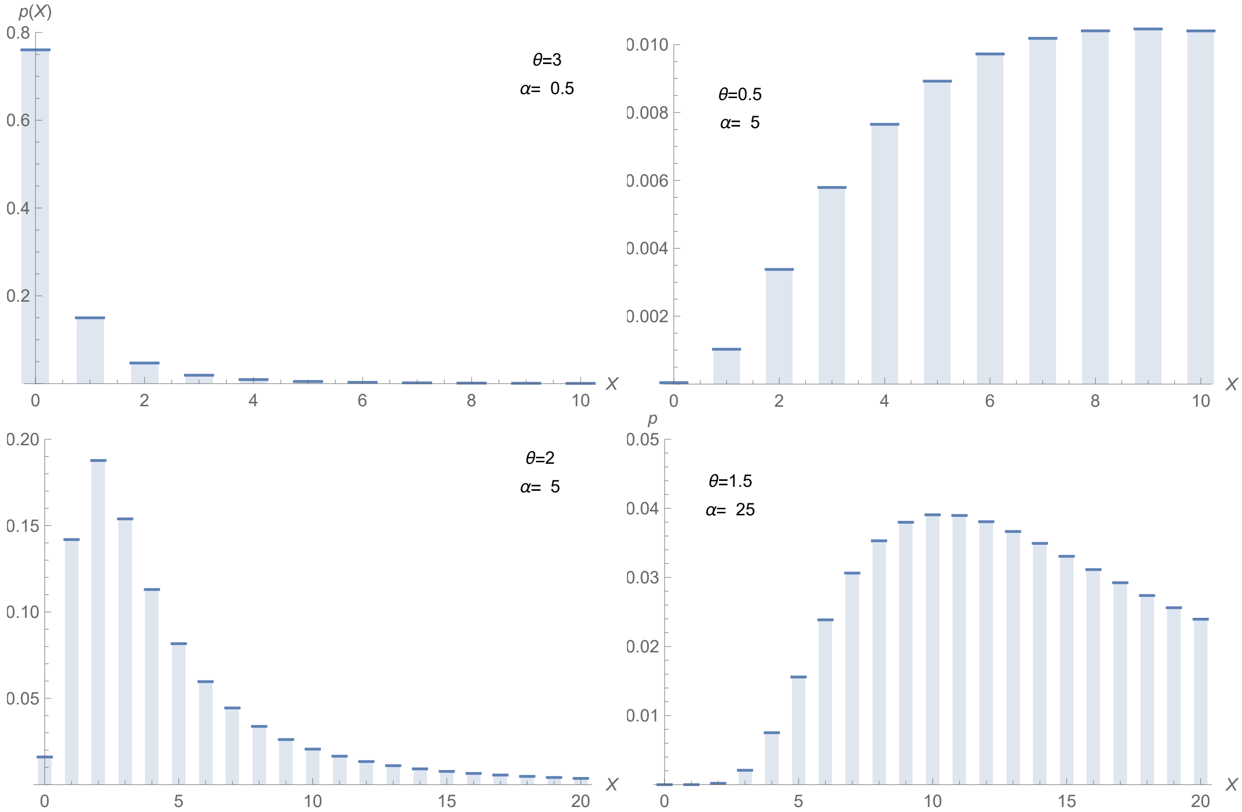

Figure 1. Plots of the pmf of DE-ITL (α, θ) for different values of (α, θ)

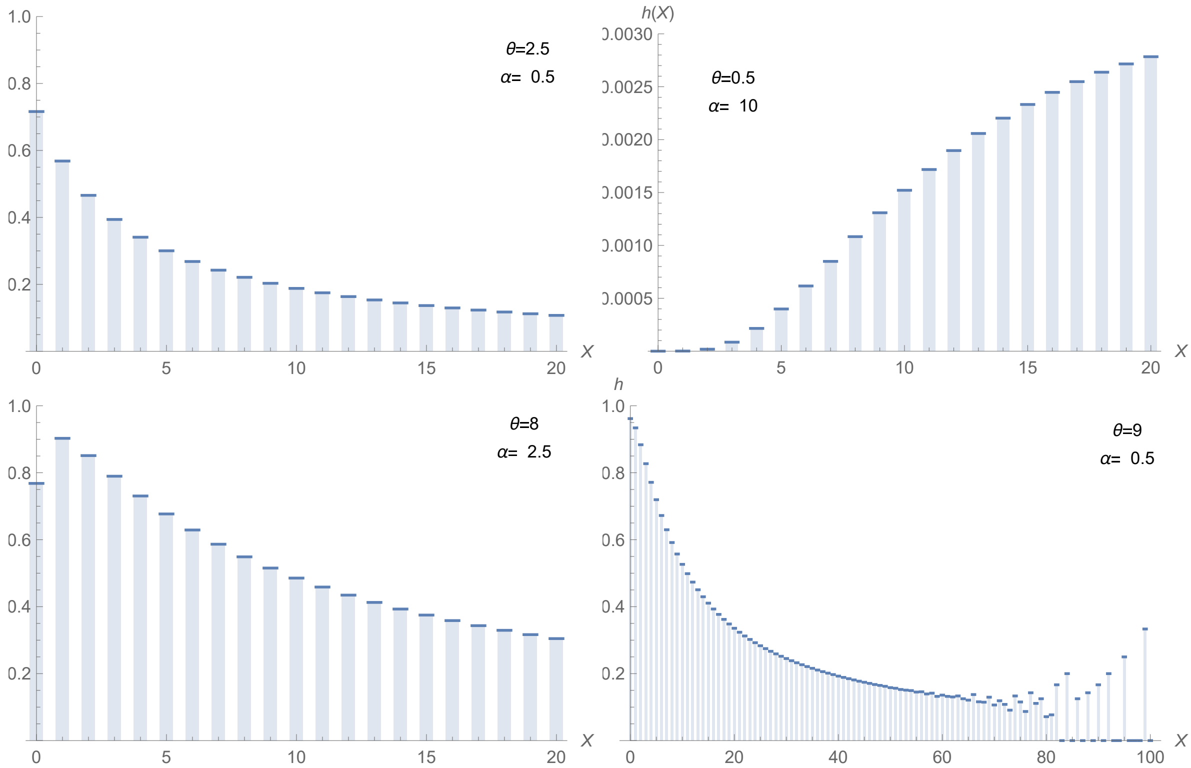

Figure 2. Plots of the hrf of DE-ITL (α, θ) for different values of (α, θ)

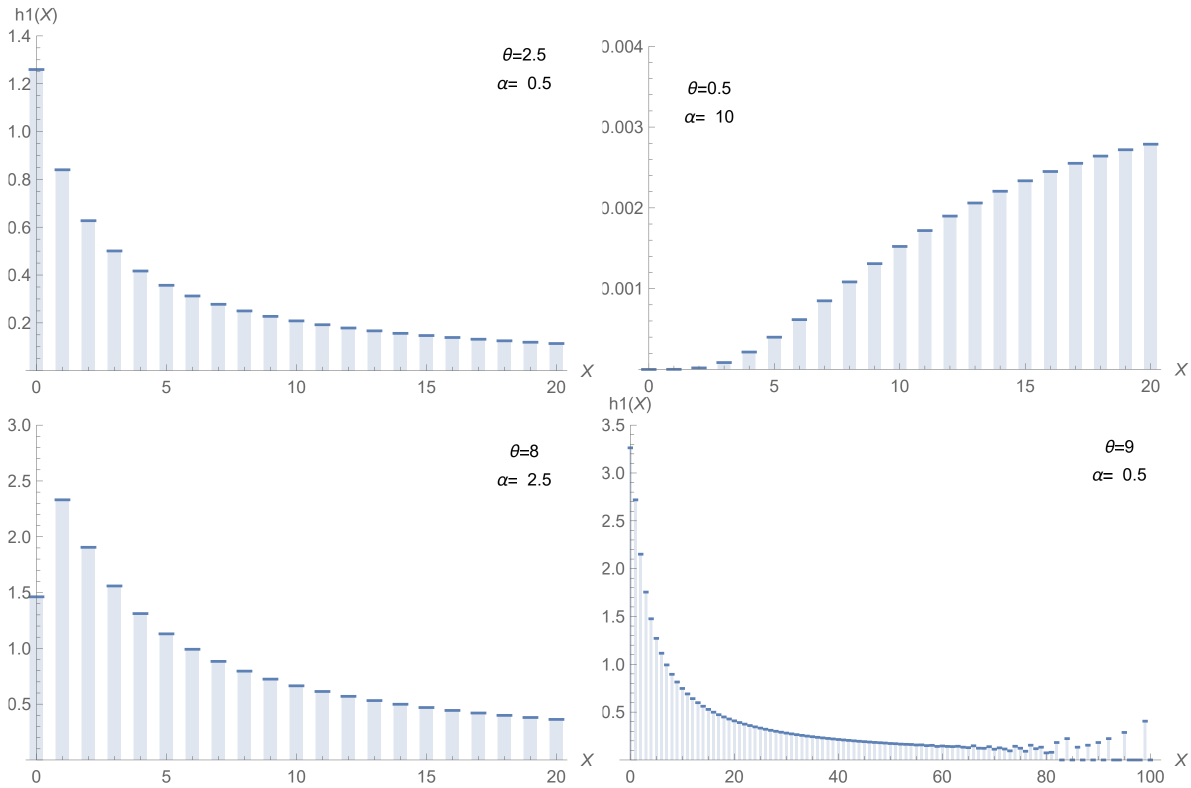

Figure 3. Plots of the ahrf of DE-ITL (α, θ) for different values of (α, θ)

Hence

$$x_u = \left\lceil \left(\frac{1}{\left(1 - \left(1 - u^{\frac{1}{\alpha}}\right)^{\frac{1}{\theta}}\right)^{\frac{-1}{2}} - 1} \right) - 1 \right\rceil. \tag{29}$$

Similarly, if $p\,(X \geq x_u) \geq 1 - u$, one obtains

$$x_u = \left\lceil \frac{1}{\left(1 - \left(1 - u^{\frac{1}{\alpha}}\right)^{\frac{1}{\theta}}\right)^{\frac{-1}{2}} - 1} \right\rceil. \tag{30}$$

Combining (29) and (30), one gets

$$\left\lceil \left(\frac{1}{\left(1 - \left(1 - u^{\frac{1}{\alpha}}\right)^{\frac{1}{\theta}}\right)^{\frac{-1}{2}} - 1} \right) - 1 \right\rceil = x_u = \left\lceil \frac{1}{\left(1 - \left(1 - u^{\frac{1}{\alpha}}\right)^{\frac{1}{\theta}}\right)^{\frac{-1}{2}} - 1} \right\rceil.$$

Hence, x_u is an integer value given by

$$x_u = \left\lceil \left(\frac{1}{\left(1 - \left(1 - u^{\frac{1}{\alpha}}\right)^{\frac{1}{\theta}}\right)^{\frac{-1}{2}} - 1} \right) - 1 \right\rceil. \tag{31}$$

By putting $u = 0.5$ in (31), one gets the median of DE-ITL as follows:

$$x_{0.5} = \left[\left(\frac{1}{\left(1-\left(1-u^{\frac{1}{\alpha}}\right)^{\frac{1}{\theta}}\right)^{\frac{-1}{2}} - 1} \right) - 1 \right]. \tag{32}$$

4.1.2. The moments of the discrete exponentiated inverted Topp – Leone distribution

a. The non-central moments

Calculating a probability distribution's mean, variance, skewness, kurtosis, and other properties involves using the distribution's moments. The non-central moments of DE-ITL distribution are

$$\acute{\mu_r} = \sum_x x^r \left[1 - \frac{(1+2(x+1))^\theta}{(1+(x+1))^{2\theta}}\right]^\alpha - \left[1 - \frac{(1+2x)^\theta}{(1+x)^{2\theta}}\right]^\alpha. \qquad x = 1, 2, \ldots$$

In particular, the mean μ is given by:

$$\acute{\mu_1} \equiv \mu = \sum_x x \left[1 - \frac{(1+2(x+1))^\theta}{(1+(x+1))^{2\theta}}\right]^\alpha - \left[1 - \frac{(1+2x)^\theta}{(1+x)^{2\theta}}\right]^\alpha, \tag{33}$$

the variance of DE-ITL distribution is

$$\begin{aligned} \mu_2 = & \sum_x x^2 \left[\left[1 - \frac{(1+2(x+1))^\theta}{(1+(x+1))^{2\theta}}\right]^\alpha - \left[1 - \frac{(1+2x)^\theta}{(1+x)^{2\theta}}\right]^\alpha\right] \\ & - \left[\sum_x x \left[1 - \frac{(1+2(x+1))^\theta}{(1+(x+1))^{2\theta}}\right]^\alpha - \left[1 - \frac{(1+2x)^\theta}{(1+x)^{2\theta}}\right]^\alpha\right]^2 \end{aligned} \tag{34}$$

b. The standard moments

The r^{th} standard moments can be obtained as follows:

$$\alpha_r = E\left(\frac{X-\mu}{\sigma}\right)^r \tag{35}$$

The skewness and kurtosis are, respectively, given by

$$\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}} \quad \text{and} \quad \alpha_4 = \frac{\mu_4}{\mu_2^2}, where\ \ \mu_r = E(X-\mu)^r, r = 1, 2, \ldots. \tag{36}$$

The ratio of the variance to the mean is known as the index of dispersion (ID), a method for detecting whether data is uniformly, unequally, or overly spread is the ID, such as if
$ID > 1$, it refers to over-dispersion, when $ID < 1$;it refers to the under-dispersion, and if
$ID = 1$, then it refers to equi-dispersion. The ID of the DE-ITL distribution can be calculated as follows:
$ID = \frac{V(x)}{E(x)}$, where $V(x)$ and $E(x)$ are the variance and mean of X, respectively.
The mean, ID, variance, skewness and kurtosis of the DE-ITL distribution for different values of α and θ are calculated numerically and displayed in Table 1 using (33) - (36).

From Table 1, one can observe that both values of the mean and variance of the DE-ITL distribution decrease when the value of the parameter θ increases. Also, the values of the mean and variance increase and then decrease when the parameter α increases. The DE-ITL distribution can be used to model positively skewed data. It can be used to model leptokurtic (kurtosis > 3) data. It is suitable for modeling over- and under-dispersed datasets where the $ID > (<) 1$.

Table 1. Some descriptive measures for different values of the parameters of DE-ITL distribution

θ	α	Mean	Variance	ID	Skewness	Kurtosis
0.2		2.9444	31.2254	10.6050	3.0016	11.2373
0.5		2.9900	29.5553	9.8847	2.7944	10.7078
0.7	0.5	2.6522	25.194	9.4993	2.9693	12.275
0.9		2.2725	20.5076	9.0242	3.2563	14.8058
0.2		3.2715	34.9489	10.6828	3.0461	10.6652
0.5		4.4729	39.2696	8.7794	2.3641	7.6277
0.7	0.99	4.2797	36.123	8.4405	2.3103	7.8226
0.9		3.8267	31.1176	8.1317	2.4300	8.8669
0.2		2.7920	33.7637	12.0930	3.2830	11.9422
0.5		5.1775	43.1468	8.3335	2.3461	7.0173
0.7	1.5	5.3469	41.4819	7.7581	2.1337	6.5730
0.9		4.9875	37.3306	7.4848	2.1235	6.9708
5		1.6188	1.4269	0.8815	2.9362	24.2061
10		0.4705	0.1317	0.2799	0.8380	3.7801
15	10	0.1274	0.1143	0.8972	2.3630	7.1392
20		0.0313	0.0305	0.9744	5.4250	30.7887

4.1.3. The order statistics of the discrete exponentiated inverted Topp-Leone distribution

The cdf of the i^{th} order statistic of the DE-ITL is given by

$$F_{i:n}(x;\theta,\alpha) = \sum_{r=i}^{n}\binom{n}{r}\sum_{j=0}^{n-r}\binom{n-r}{j}(-1)^{j}\left[1-\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{\alpha(r+j)}. \tag{37}$$

Special cases

Case I: if $i = 1$ in (37) one can obtain the distribution function of the first order statistics, as given below

$$\begin{aligned} F_1(x;\theta,\alpha) &= 1-[1-F(x;\theta,\alpha)]^{n} \\ &= 1-\left[1-\left[1-\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{\alpha}\right]^{n}, \end{aligned} \tag{38}$$

Case II: if $i = n$ in (37) one can get the distribution function of the largest order statistics, as follows:

$$F_n(x;\theta,\alpha) = [F(x;\theta,\alpha)]^{n} = \left[1-\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{n\alpha}. \tag{39}$$

the pmf of the i^{th} order statistics, is

$$
\begin{aligned}
P\left(X_{(i)}=x\right) &= \frac{n!}{(i-1)!\,(n-i)!}\sum_{j=0}^{n-i}\binom{n-i}{j}(-1)^{j}\int_{F(x-1)}^{F(x)} v^{i+j-1}dv \\
&= \frac{n!}{(i-1)!\,(n-i)!}\sum_{j=0}^{n-i}\binom{n-i}{j}(-1)^{j}\left(\frac{1}{i+j}\right) \\
&\times\left[\left[1-\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{\alpha(r+j)}-\left[1-\frac{(1+2(x))^{\theta}}{(1+x)^{2\theta}}\right]^{\alpha(r+j)}\right].
\end{aligned} \tag{40}
$$

The pmf of the smallest order statistics is obtained by substituting $i=1$ in (40) as follows:

$$
\begin{aligned}
P\left(X_{(1)}=x\right) &= \frac{n!}{(1-1)!(n-1)!}\int_{F(x-1)}^{F(x)} v^{1-1}(1-v)^{n-1+1}dv \\
&= \left[\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{n\alpha}-\left[\frac{(1+2(x))^{\theta}}{(1+x)^{2\theta}}\right]^{n\alpha},
\end{aligned} \tag{41}
$$

also, the pmf of the largest order statistic is obtained by substituting $i=n$ in (40) as follows:

$$
P\left(X_{(n)}=x\right)=\left[1-\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{n\alpha}-\left[1-\frac{(1+2(x))^{\theta}}{(1+x)^{2\theta}}\right]^{n\alpha} \tag{42}
$$

4.1.4. Rényi entropy

The Rényi entropy and Shannon entropy for the DE-ITL can be obtained using (11) and (12) as given bellow

$$
I_{\eta}(x)=\frac{1}{1-\eta}log\sum_{x}\left[\left[1-\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{\alpha}-\left[1-\frac{(1+2x)^{\theta}}{(1+x)^{2\theta}}\right]^{\alpha}\right]^{\eta},\quad x=0,1,2,\ldots, \tag{43}
$$

and

$$
I(X)=-log\left[\left[1-\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{\alpha}-\left[1-\frac{(1+2x)^{\theta}}{(1+x)^{2\theta}}\right]^{\alpha}\right]\sum_{x}\left[\left[1-\frac{(1+2(x+1))^{\theta}}{(1+(x+1))^{2\theta}}\right]^{\alpha}-\left[1-\frac{(1+2x)^{\theta}}{(1+x)^{2\theta}}\right]^{\alpha}\right],
$$

$$
x=0,1,2,\ldots. \tag{44}
$$

The Shannon entropy is a special case of the Rényi entropy when $\eta\to 1$.

4.1.5. Mean residual lifetime function, mean time to failure, mean time between failure, and availability

The mean residual lifetime function, mean time to failure, mean time between failure, and availability for the DE-ITL can be obtained from (13)-(16), respectively, as follows:

$$
MRL(x)=\frac{1}{1-\left[1-(1+2x_0)^{\theta}(1+x_0)^{-2\theta}\right]^{\alpha}}\sum_{k=x_0+1}^{\infty}1-\left[1-(1+2k)^{\theta}(1+k)^{-2\theta}\right]^{\alpha}, \tag{45}
$$

the MTBF is

$$MTBF \;=\; \frac{-x}{\log\left[1-\left[1-(1+2x)^{\theta}(1+x)^{-2\theta}\right]^{\alpha}\right]}.\; x=1,2,\ldots. \quad , \tag{46}$$

then the MTTF is given by

$$MTTF = \sum_{x=1}^{\infty} 1-\left[1-(1+2x)^{\theta}(1+x)^{-2\theta}\right]^{\alpha}. \qquad x=1,2,\ldots. \quad , \tag{47}$$

and Av is considered as being the probability that the component is successful at time t,

$$\text{Av} = \frac{MTTF}{MTBF} \;=\; \frac{\sum_{x=1}^{\infty} 1-\left[1-(1+2x)^{\theta}(1+x)^{-2\theta}\right]^{\alpha}\log\left[1-\left[1-(1+2x)^{\theta}(1+x)^{-2\theta}\right]^{\alpha}\right]}{-x}. \tag{48}$$

5. Maximum likelihood estimation

his section is devoted to estimate the vector of the unknown parameters, $\underline{\varphi} = (\theta, \alpha)$, sf, hrf and ahrf of the DE-ITL (θ, α) distribution, under Type II censored samples, also the confidence intervals of the parameters α, θ, sf, hrf and ahrf are derived.

Suppose that $X_1, X_2, \ldots, X_r$ is a Type II censored sample of size r obtained from a life-test of n items whose lifetimes have a DE-ITL (θ, α) distribution. Then the likelihood function is

$$L\left(\underline{\varphi};\underline{x}\right) \propto \left\{\prod_{i=1}^{r} P\left(x_i\right)\right\}\left[S\left(x_{(r)}\right)\right]^{n-r}, \tag{49}$$

where $P(x)$ and $S(x)$ are given respectively by (24) and (26).

$$L\left(\underline{\varphi};\underline{x}\right) \propto \left\{\prod_{i=1}^{r}\left[1-\frac{(1+2(x_i+1))^{\theta}}{(1+(x_i+1))^{2\theta}}\right]^{\alpha}-\left[1-\frac{(1+2x_i)^{\theta}}{(1+x_i)^{2\theta}}\right]^{\alpha}\right\}$$

$$\times\left[1-\left[1-\frac{(1+2x_r)^{\theta}}{(1+x_r)^{2\theta}}\right]^{\alpha}\right]^{n-r}, \tag{50}$$

$$L\left(\underline{\varphi};\underline{x}\right) \propto \left\{\prod_{i=1}^{r}\left[u_{i1}\right]^{\alpha}-\left[u_{i2}\right]^{\alpha}\right\}\times\left[1-\left[u_r\right]^{\alpha}\right]^{n-r},$$

where

$u_{i1} = \left[1-\frac{(1+2(x_i+1))^{\theta}}{(1+(x_i+1))^{2\theta}}\right]$, $u_{i2} = \left[1-\frac{(1+2x_i)^{\theta}}{(1+x_i)^{2\theta}}\right]$, and $u_r = \left[1-\frac{(1+2x_r)^{\theta}}{(1+x_r)^{2\theta}}\right]$.

The natural logarithm of the likelihood function is given by

$$\ell \equiv lnL\left(\underline{\varphi};\underline{x}\right) \propto ln\left\{\prod_{i=1}^{r}\left[1-\frac{(1+2(x_i+1))^{\theta}}{(1+(x_i+1))^{2\theta}}\right]^{\alpha}-\left[1-\frac{(1+2x_i)^{\theta}}{(1+x_i)^{2\theta}}\right]^{\alpha}\right\}$$

$$+\;(n-r)\,ln\left[1-\left[1-\frac{(1+2x_r)^{\theta}}{(1+x_r)^{2\theta}}\right]^{\alpha}\right], \tag{51}$$

then

$$\ell \equiv lnL\left(\underline{\varphi};\underline{x}\right) \propto \sum_{i=1}^{r} ln\left[\left[1-\frac{(1+2(x_i+1))^{\theta}}{(1+(x_i+1))^{2\theta}}\right]^{\alpha}-\left[1-\frac{(1+2x_i)^{\theta}}{(1+x_i)^{2\theta}}\right]^{\alpha}\right]$$

$$+ (n-r)\, ln\left[1 - \left[1 - \frac{(1+2x_r)^{\theta}}{(1+x_r)^{2\theta}}\right]^{\alpha}\right]$$

$$\propto \sum_{i=1}^{r} ln\left[[u_{i1}]^{\alpha} - [u_{i2}]^{\alpha}\right] + (n-r)\, ln\left[1 - [u_r]^{\alpha}\right], \tag{52}$$

where
$u_{i1} = \left[1 - \frac{(1+2(x_i+1))^{\theta}}{(1+(x_i+1))^{2\theta}}\right]$, $u_{i2} = \left[1 - \frac{(1+2x_i)^{\theta}}{(1+x_i)^{2\theta}}\right]$, and $u_r = \left[1 - \frac{(1+2x_r)^{\theta}}{(1+x_r)^{2\theta}}\right]$
Considering the two parameters, θ and α are unknown and differentiating (59), with respect to α and θ, one obtains

$$\frac{\partial \ell}{\partial a} = \sum_{i=1}^{r} \left\{ \frac{[u_{i1}]^{\alpha} ln\,(u_{i1}) - [u_{i2}]^{\alpha} ln\,(u_{i2})}{[[u_{i1}]^{\alpha} - [u_{i2}]^{\alpha}]} \right\} - (n-r)\, \frac{\{[u_r]^{\alpha} ln\,[u_r]\}}{[1 - [u_r]^{\alpha}]}, \tag{53}$$

and

$$\frac{\partial \ell}{\partial ?} = \sum_{i=1}^{r} \left\{ \frac{[\alpha\, [u_{i1}]^{\alpha-1} \acute{u}_{i1}] - [\alpha\, [u_{i2}]^{\alpha-1} \acute{u}_{i2}]}{[[u_{i1}]^{\alpha} - [u_{i2}]^{\alpha}]} \right\} - (n-r)\, \frac{[\alpha\, [u_r]^{\alpha-1} \acute{u}_r]}{1 - (u_r)}, \tag{54}$$

where

$$\acute{u}_{i1} = -\frac{\left\{\left[(1+(x_i+1))^{2\theta}\right] * \left[(1+2(x_i+1))^{\theta}\right] * ln\,[(1+2(x_i+1))]\right\} - \left\{\left[(1+2(x_i+1))^{\theta}\right] * \left[(1+(x_i+1))^{2\theta}\right] * ln\,[(1+(x_i+1))] * 2\right\}}{\left[(1+(x_i+1))^{2\theta}\right]^2},$$

$$\acute{u}_{i2} = -\frac{\left\{\left[(1+(x_i))^{2\theta}\right] * \left[(1+2(x_i))^{\theta}\right] * ln\,[(1+2(x_i))]\right\} - \left\{\left[(1+2(x_i)^{\theta}\right] * \left[(1+(x_i))^{2\theta}\right] * ln\,[(1+(x_i))] * 2\right\}}{\left[(1+(x_i))^{2\theta}\right]^2},$$

and

$$\acute{u}_r = -\frac{\left\{\left[(1+x_r)^{2\theta}\right] * \left[(1+2x_r)^{\theta}\right] * ln\,[(1+2x_r)]\right\} - \left\{\left[(1+2x_r)^{\theta}\right] * \left[(1+x_r)^{2\theta}\right] * ln\,[(1+x_r)] * 2\right\}}{\left[(1+x_r)^{2\theta}\right]^2}, \tag{55}$$

Then the ML estimates of the parameters, denoted by $\widehat{\alpha}$ and $\widehat{\theta}$ can be derived by equating (53) and (54) to zeros and solving numerically.
Depending on the invariance property, the ML estimators of $S\,(x)$, $h(x)$ and $h_1(x)$can be obtained by replacing α and θ with their corresponding ML estimators $\widehat{\alpha}$ and $\widehat{\theta}$, respectively, in (26) - (28), as given below

$$\hat{S}_{ML}\,(x;\theta,\alpha) = 1 - \left[1 - (1+2x)^{\widehat{\theta}}(1+x)^{-2\widehat{\theta}}\right]^{\widehat{\alpha}}, \qquad x = 0, 1, 2, \ldots, \tag{56}$$

$$\hat{h}_{ML}\,(x;\theta,\alpha) = \frac{\left[1 - \frac{(1+2(x+1))^{\widehat{\theta}}}{(1+(x+1))^{2\widehat{\theta}}}\right]^{\widehat{\alpha}} - \left[1 - \frac{(1+2x)^{\widehat{\theta}}}{(1+x)^{2\widehat{\theta}}}\right]^{\widehat{\alpha}}}{1 - \left[1 - (1+2x)^{\widehat{\theta}}(1+x)^{-2\widehat{\theta}}\right]^{\widehat{\alpha}}}, \qquad x = 0, 1, 2, \ldots, \tag{57}$$

and

$$\hat{h}_{1ML}(x;\theta,\alpha) = ln\left[\frac{1-\left[1-(1+2x)^{\widehat{\theta}}(1+x)^{-2\widehat{\theta}}\right]^{\widehat{\alpha}}}{1-\left[1-(1+2(x+1))^{\widehat{\theta}}(1+(x+1))^{-2\widehat{\theta}}\right]^{\widehat{\alpha}}}\right]. \qquad x = 0,1,2,\ldots\ . \tag{58}$$

When the sample size is large and the regularity conditions are satisfied, the asymptotic distribution of the ML estimators is $\widehat{\underline{\varphi}} \sim Bivariate\ Normal\left(\underline{\varphi}, I^{-1}\underline{x}\left(\underline{\varphi}\right)\right)$, where $\underline{\varphi} = (\theta,\alpha)$, $\widehat{\underline{\varphi}} = \left(\widehat{\theta},\widehat{\alpha}\right)$, and $I^{-1}\left(\underline{\varphi}\right)$ is the asymptotic variance covariance matrix of the ML estimators α and θ, which is the inverse of the asymptotic observed Fisher information matrix. The asymptotic observed Fisher information matrix can be obtained as follows:

$$I\left(\underline{\varphi}\right) \approx -\left[\frac{\partial^2\ell}{\partial\omega_i\partial\omega_j}\right], \qquad i,\ j = 1,2,\ \ldots, \tag{59}$$

The asymptotic 100 $(1-\alpha)$% confidence interval for θ,α, $S_{ML}(x)$, $h_{ML}(x)$ and $h_{1ML}(x)$ are given, respectively, by

$L_\omega = \widehat{\omega} - Z_{\frac{\alpha}{2}}\ \sigma_{\widehat{\omega}}$. and $U_\omega = \widehat{\omega} + Z_{\frac{\alpha}{2}}\ \sigma_{\widehat{\omega}}$. (59)

where L_ω and U_ω are the lower and upper limits respectively, $\widehat{\omega}$ is $\widehat{\theta},\widehat{\alpha}$, $\hat{S}(x)$, $\hat{h}(x)$ or $\hat{h}_1(x)$, where Z is the 100 % $\left(1-\frac{\alpha}{2}\right)$ standard normal percentile, $(1-\alpha)$ is the confidence coefficient and $\sigma_{\widehat{\omega}}$ is the standard deviation.

6. Numerical Illustration

This section aims to investigate the precision of the theoretical results based on simulated and real data**.**

6.1. Simulation study

In this subsection, a simulation study is conducted to illustrate the performance of the presented ML estimates based on generated data from the DE-ITL distribution. The ML averages of the parameters, sf, hrf and ahrf based on complete sample and Type II censoring are computed. Moreover, confidence intervals of the parameters, sf, hrf and ahrf are calculated. The simulation study is performed using Mathematica11.

Tables 2 displays the averages, relative absolute biases (RABs), mean squared error (MSE) and variances, for the parameters, sf, hrf and ahrf estimates, also 95% confidence intervals under three levels of $\frac{r}{n} \times 100$ percentage of uncensored observations Type II censoring 60%, 80% and 100%. Table 3 presents the same computational results, but for different population parameter values from the DE-ITL distribution for different samples of size n=30, 60 and 120 and number of replications (NR) = 5000.

The RABs, variances of the ML estimates of the parameters, sf, hrf and ahrf are computed as follows:

Table 2: ML averages, relative absolute biases, mean squared errors, variances of ML estimates, 95% confidence intervals of the parameters, survival, hazard rate and alternative hazard rate functions at $(X_0 = 1)$ from DE-ITL distribution for different samples sizes n, censoring size r $(NR = 5000, \theta = 0.99, \alpha = 1.5)$.

Table 3: ML averages, relative absolute biases, mean squared errors, variances of ML estimates, 95% confidence intervals of the parameters, survival, hazard rate and alternative hazard rate functions

at ($x_0 = 1$) from DE-ITL distribution for different samples size n, censoring size r ($NR = 5000, \theta = 0.6, \alpha = 1.5$)

6.2. Concluding remarks

From Tables 2 and 3 of simulation study one can observe that:

1. The RABs, MSEs, and variances of the ML averages of the ML estimates of the parameters, sf, hrf, and ahrf perform better as n increases, i.e., the RABs, MSEs, and variances decrease when the sample size n increases, as expected. This is indicative of the fact that the estimates are consistent and approach the population parameter values as the sample size increases. Also, the lengths of the confidence intervals get shorter when the sample size increases.
2. The RABs, MSEs, and variances of the ML estimates of the parameters, sf, hrf, and ahrf decrease when the level of censoring decreases, which is expected since decreasing the level of censoring means that more information is provided by the sample and hence increases the accuracy of the estimates.
3. In general, all the results of RABs, MSEs and variances perform better when n and r are larger; RABs, MSEs and variances obtained for complete sample sizes, are less than the corresponding results for censored samples.

6.3. Applications

In this subsection, the importance and applicability of the DE-ITL distribution is discussed by using two real data sets. The proposed distribution is compared with different competitive distributions such as discrete Marshall-Olkin inverted Topp-Leone (DMOITL) introduced by Almetwally et al. [6], discrete Half Logistic (DHL) presented by Hegazy et al. [23], discrete generalized inverted exponential (DGIE) proposed by AbdElaziz et al. [1] and discrete generalized Rayleigh (DGR) obtained by Alamatsaz et al. [3]. The fitted probability distributions are compared using some criteria, Akaike Information Criterion (AIC), Akaike Information Criterion with Correction (AICC), Bayesian Information Criterion (BIC) and Hannon-Quinn Information Criterion (HQIC). The best distribution corresponds to the highest p-value and the lowest values of AIC, AICC, BIC and HQIC.
Where $AIC = -2\log L + 2k$, $AICC = AIC + \frac{2k(k+1)}{n-k-1}$, $BIC = -2logL + k\log n$,
and
$HQIC = -2\log L + 2k\ log\,(log\,(n))$, where k is the number of the parameters, n is the sample size and L is the maximized value of the likelihood function for the estimated model. Tables 4 and 5 display the values of p-value, AIC, BIC, AICC and HQIC for the first and second data sets.
Kolmogorov-Smirnov (K-S) goodness of fit test is applied to check the validity of the fitted model. The p-values are respectively 0.1674 and 0.5099. It shows that DE-ITL fits the data very well.
Data set I: The first data set was given by Lawless [33]. It represents remission times, in weeks, for 20 leukemia patients randomly assigned to a specific treatment.
The data is: 1, 3, 3, 6, 7, 7, 10, 12, 14, 15, 18, 19, 22, 26, 28, 29, 34, 40, 48, 49.
Table 4 presents the ML estimates and corresponding standard errors (SEs), p-value, $-2lnL$, AIC, BIC, AICC and HQIC. It is observed that all models fit the data set. However, the proposed distribution has highest p-value and smallest values of $-2lnL$, AIC, BIC, AICC. Hence, the DE-ITL distribution is the best fit for this data compared with other distributions considered here.

Table 2. Some resumes of ML for parameters, survival, hazard rate and alternative hazard rate functions: ($\theta = 0.99, \alpha = 1.5$)

n	r	parameter	Average	RAB	MSE	Variance	UL	LL	Length
30	18	θ	1.1334	0.1449	0.0911	0.0705	1.6541	0.6128	1.0412
		α	2.1878	0.4585	0.5168	0.0437	2.5976	1.7779	0.8197
		$S(x_0)$	0.9390	0.0711	0.0040	0.0001	0.9643	0.9136	0.0507
		$h(x_0)$	0.1556	0.1957	0.0029	0.0015	0.2318	0.0794	0.1524
		$h_1(x_0)$	0.1702	0.2084	0.0041	0.0021	0.2609	0.0795	0.1813
	24	θ	1.1088	0.1200	0.0655	0.0514	1.5532	0.6643	0.8889
		α	2.1675	0.4451	0.4810	0.0356	2.5375	1.7979	0.7395
		$S(x_0)$	0.9398	0.0720	0.0041	0.0001	0.9620	0.9172	0.0451
		$h(x_0)$	0.1521	0.2136	0.0028	0.0010	0.2171	0.0872	0.1299
		$h_1(x_0)$	0.1658	0.2288	0.0039	0.0015	0.2425	0.0892	0.1533
	30	θ	1.0762	0.0871	0.0486	0.0411	1.4739	0.6787	0.7951
		α	2.1338	0.4225	0.4330	0.0312	2.4803	1.7872	0.6932
		$S(x_0)$	0.9403	0.0726	0.0041	0.0001	0.9623	0.9182	0.0441
		$h(x_0)$	0.1483	0.2333	0.0029	0.0009	0.2072	0.0895	0.1176
		$h_1(x_0)$	0.1612	0.2503	0.0041	0.0012	0.2302	0.0922	0.1380
60	36	θ	1.1216	0.1329	0.0564	0.0391	1.5093	0.7339	0.7754
		α	2.1719	0.4479	0.4739	0.0224	2.4653	1.8785	0.5868
		$S(x_0)$	0.9390	0.0712	0.0039	0.0000	0.9574	0.9207	0.0364
		$h(x_0)$	0.1543	0.2025	0.0023	0.0008	0.2112	0.0974	0.1137
		$h_1(x_0)$	0.1682	0.2179	0.0033	0.0011	0.2355	0.1009	0.1346
	48	θ	1.0980	0.1091	0.0385	0.0269	1.4195	0.7761	0.6430
		α	2.1528	0.4352	0.4437	0.0174	2.4119	1.8939	0.5180
		$S(x_0)$	0.9395	0.0721	0.0040	0.0000	0.9560	0.9237	0.0322
		$h(x_0)$	0.1510	0.2195	0.0023	0.0005	0.1982	0.1037	0.0945
		$h_1(x_0)$	0.1641	0.2368	0.0034	0.0008	0.2198	0.1084	0.1113
	60	θ	1.0780	0.0889	0.0284	0.0206	1.3598	0.7963	0.5635
		α	2.1301	0.4201	0.4121	0.0149	2.3695	1.8904	0.4794
		$S(x_0)$	0.9400	0.0723	0.0040	0.0000	0.9554	0.9245	0.0308
		$h(x_0)$	0.1488	0.2307	0.0024	0.0004	0.1907	0.1069	0.0837
		$h_1(x_0)$	0.1614	0.2491	0.0035	0.0006	0.2107	0.1123	0.0981
120	72	θ	1.1187	0.1300	0.0372	0.0206	1.4006	0.8368	0.5638
		α	2.1632	0.4424	0.4522	0.0118	2.3770	1.9501	0.4264
		$S(x_0)$	0.9387	0.0708	0.0039	0.0000	0.9518	0.9252	0.0263
		$h(x_0)$	0.1543	0.2024	0.0019	0.0004	0.1955	0.1130	0.0825
		$h_1(x_0)$	0.1679	0.2191	0.0028	0.0006	0.2167	0.1191	0.0975
	96	θ	1.0963	0.1079	0.0253	0.0139	1.3284	0.8652	0.4631
		α	2.1460	0.4306	0.4262	0.0088	2.3303	1.9617	0.3686
		$S(x_0)$	0.9394	0.0717	0.0039	0.0000	0.9510	0.9278	0.0232
		$h(x_0)$	0.1512	0.2181	0.0020	0.0003	0.1853	0.1172	0.0681
		$h_1(x_0)$	0.1642	0.2361	0.0030	0.0004	0.2043	0.1241	0.0802
	120	θ	1.0744	0.0853	0.0174	0.0103	1.2736	0.8752	0.3982
		α	2.1220	0.4146	0.3945	0.0074	2.2914	1.9522	0.3385
		$S(x_0)$	0.9392	0.0720	0.0040	0.0000	0.9508	0.9286	0.0222
		$h(x_0)$	0.1487	0.2314	0.0022	0.0002	0.1785	0.1189	0.0595
		$h_1(x_0)$	0.1611	0.2505	0.0032	0.0003	0.1961	0.1262	0.0698

Table 3. Some resumes of ML for parameters, survival, hazard rate and alternative hazard rate functions: ($\theta = 0.6, \alpha = 1.5$)

n	r	parameter	Average	RAB	MSE	Variance	UL	LL	Length
30	18	θ	0.7449	0.2415	0.0533	0.0323	1.0974	0.3923	0.7050
		α	2.1252	0.4168	0.4553	0.0643	2.6225	1.6280	0.9944
		$S(x_0)$	0.9695	0.0348	0.0011	0.0000	0.9812	0.9577	0.0235
		$h(x_0)$	0.0829	0.2141	0.0008	0.0003	0.1188	0.0470	0.0718
		$h_1(x_0)$	0.0867	0.2219	0.0010	0.0003	0.1259	0.0477	0.0782
	24	θ	0.7082	0.1804	0.0344	0.0514	1.0034	0.4130	0.5903
		α	2.0627	0.3751	0.3691	0.0356	2.5144	1.6111	0.9033
		$S(x_0)$	0.9690	0.0343	0.0010	0.0001	0.9819	0.9561	0.0257
		$h(x_0)$	0.0806	0.2363	0.0008	0.0010	0.1117	0.0494	0.0623
		$h_1(x_0)$	0.0841	0.2452	0.0010	0.0015	0.1180	0.0503	0.0677
	30	θ	0.6805	0.1341	0.0233	0.0168	0.9348	0.4261	0.5086
		α	2.0084	0.3389	0.3048	0.0463	2.4302	1.5865	0.8436
		$S(x_0)$	0.9681	0.0334	0.0010	0.0001	0.9834	0.9529	0.0304
		$h(x_0)$	0.0794	0.2476	0.0009	0.0002	0.1090	0.0497	0.0592
		$h_1(x_0)$	0.0828	0.2570	0.0010	0.0002	0.1150	0.0507	0.0643
60	36	θ	0.7405	0.2345	0.0364	0.0166	0.9932	0.4881	0.5051
		α	2.1088	0.4059	0.4013	0.0306	2.4518	1.7659	0.6858
		$S(x_0)$	0.9691	0.0344	0.0010	0.0000	0.9771	0.9611	0.0159
		$h(x_0)$	0.0831	0.2121	0.0006	0.0001	0.1094	0.0568	0.0525
		$h_1(x_0)$	0.0869	0.2207	0.0008	0.0002	0.1155	0.0583	0.0572
	48	θ	0.7022	0.1704	0.0221	0.0116	0.9141	0.4903	0.4234
		α	2.0456	0.3637	0.3247	0.0269	2.3676	1.7236	0.6439
		$S(x_0)$	0.9688	0.0341	0.0010	0.0000	0.9777	0.9598	0.0177
		$h(x_0)$	0.0805	0.2370	0.0007	0.0001	0.1029	0.0581	0.0447
		$h_1(x_0)$	0.0840	0.2466	0.0009	0.0001	0.1083	0.0597	0.0486
	60	θ	0.6773	0.1289	0.0144	0.0084	0.8577	0.4969	0.3607
		α	1.9939	0.3293	0.2677	0.0237	2.2959	1.6919	0.6039
		$S(x_0)$	0.9679	0.0332	0.0009	0.0000	0.9784	0.9575	0.0208
		$h(x_0)$	0.0796	0.2454	0.0007	0.0001	0.1004	0.0588	0.0416
		$h_1(x_0)$	0.0830	0.2553	0.0009	0.0001	0.1056	0.0604	0.0451
120	72	θ	0.7353	0.2251	0.0268	0.0085	0.9166	0.5540	0.3625
		α	2.0937	0.3958	0.3689	0.0164	2.3453	1.8420	0.5033
		$S(x_0)$	0.9688	0.0341	0.0010	0.0000	0.9743	0.9633	0.0110
		$h(x_0)$	0.0831	0.2120	0.0005	0.0001	0.1015	0.0647	0.0368
		$h_1(x_0)$	0.0868	0.2210	0.0007	0.0001	0.1069	0.0668	0.0401
	96	θ	0.6995	0.1659	0.0156	0.0056	0.8471	0.5520	0.2950
		α	2.0334	0.3558	0.2977	0.0128	2.2563	1.8111	0.4451
		$S(x_0)$	0.9685	0.0337	0.0010	0.0000	0.9747	0.9622	0.0125
		$h(x_0)$	0.0808	0.2344	0.0006	0.0001	0.0966	0.0649	0.0316
		$h_1(x_0)$	0.0843	0.2442	0.0008	0.0001	0.1015	0.0670	0.0344
	120	θ	0.6753	0.1255	0.0097	0.0040	0.7998	0.5507	0.2490
		α	1.9880	0.3253	0.2489	0.0107	2.1917	1.7843	0.4073
		$S(x_0)$	0.9680	0.0332	0.0009	0.0000	0.9751	0.9608	0.0142
		$h(x_0)$	0.0795	0.2460	0.0007	0.0001	0.0942	0.0648	0.0294
		$h_1(x_0)$	0.0829	0.2563	0.0008	0.0001	0.0989	0.0669	0.0319

Table 4. ML estimates, SEs and goodness of fit for various models fitted for data set I

Models	Estimates	SEs	P-Value	AIC	BIC	AICC	HQIC
DE-ITL	θ=2.1868 α=2.3121	0.1949 0.1969	0.1674	170.038	174. 744	176.03	174.038
DMOITL	θ=0.7016 α=0.3327	1.1020 1.1081	0.08	200.422	204.422	206.413	205.128
DHL	θ=7.6156	0. 9780	.0810	172.465	174.918	175.918	175.141
DGLE	λ=3.9563 α=0.7242	1.0422 1.1016	0.0793	172.465	176.465	178.457	177.171
DGR	λ=0.0177 α=0.1871	1.1147 1.1115	0.0778	170.779	174.779	176.77	175.485

Figure 4 exhibits TTT plot of data set I which indicates that this data has an increasing- shaped hazard rate. The fitted pmf, P-P and Q-Q plots indicate that the DE-ITL distribution provides the best fit for this data.

Data set II

The second dataset was presented by Lawless [33]. It represents the failure times in minutes for Epoxy Insulation Specimens at the voltage level 57.5 KV. For ease of modelling, the whole dataset is divided by 100.

The data are: 1, 1, 2, 2, 2, 2, 2, 3, 3, 4, 4, 5, 5, 5, 5, 6, 6, 7, 9, 10.

Table 5 presents the ML estimates and corresponding SEs, p-value, $-2lnL$, AIC, BIC and AICC for the data set II. It is observed that all models fit the data set. However, the DE-ITL distribution has highest p-value and the smallest values of $-2lnL$, AIC, BIC, AICC. Hence, the DE-ITL distribution is the best fit for this data compared with other distributions considered here.

Figure 5 shows that TTT plot of real data set II indicates that this data has an increasing-shaped hazard rate. The fitted pmf, P-P and Q-Q plots indicate that the DE-ITL distribution fits the data very well.

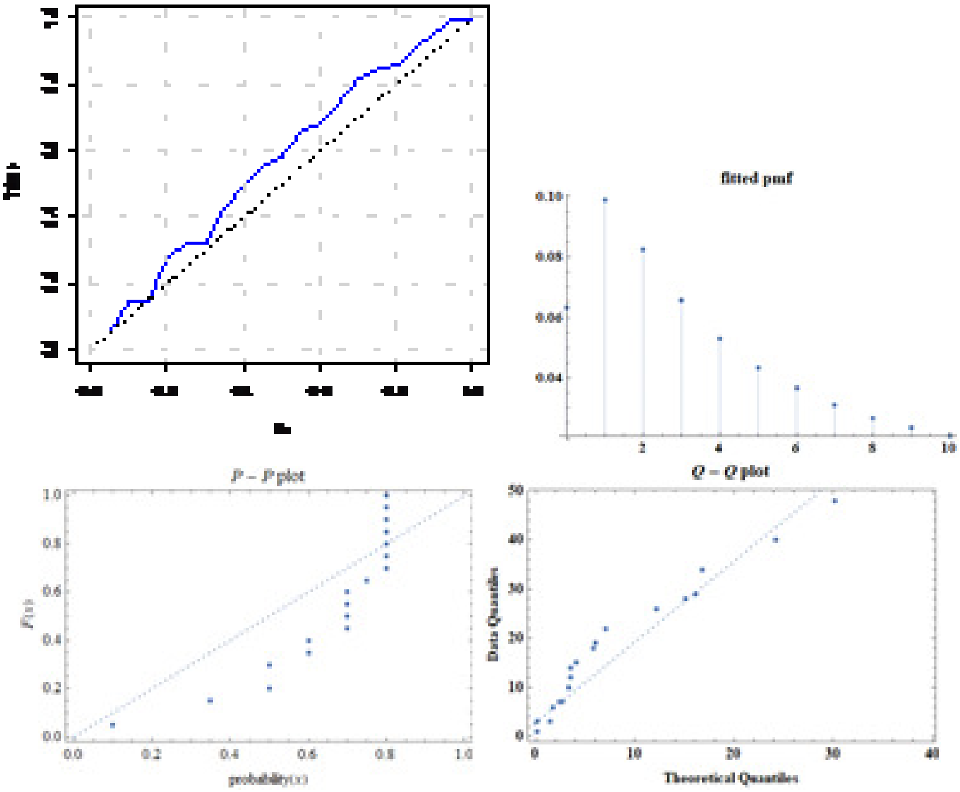

Figure 4. TTT, fitted pmf, P-P and Q-Q plots of the DE-ITL distribution for data set I

Table 5. ML estimates, SEs and goodness of fit for various models fitted for the data set II

Models	Estimates	SEs	P-Value	AIC	BIC	AICC	HQIC
DE-ITL	θ=.6221 α=2.7032	0.4276 0.3844	0.8053	98.6342	102.634	104.626	103.34
DMOITL	θ=0.4220 α=0.5520	3.498 3.4873	0.1485	121.484	125.484	127.475	126.19
DHL	θ=24.5701	1.0259	.0.3086	156.253	158.253	159.249	158.475
DGLE	λ=14.8494 α=9.5671	0.7488 0.5503	0.1587	103.236	107.236	109.228	107.942
DGR	λ=0.0993 α=0.6886	0.4953 0.4690	0.0272	102.747	106.747	108.739	107.453

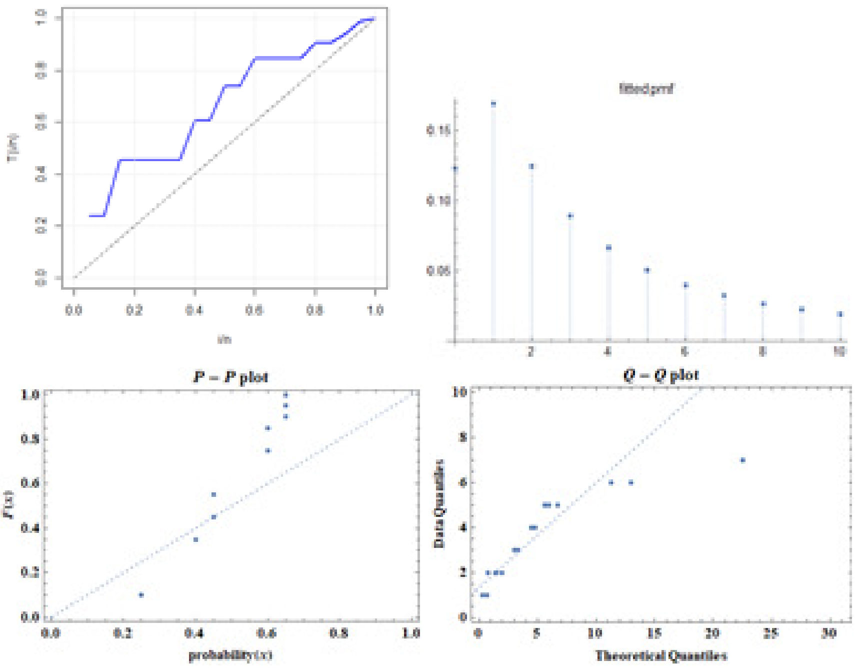

Figure 5. The fitted pdf, PP-plot, QQ-plot and TTT-plot for the sconed data set

7. Conclusion

In this paper, a new family of discrete distributions called discrete exponentiated-G family of distributions is proposed. Several important statistical characteristics are investigated. Discrete exponentiated inverted Topp – Leone distribution as a sub model of the proposed family is studied in detail. It is explored that the discrete exponentiated-G family can be used for modeling count and lifetime data. The discrete inverted Topp – Leone distribution can model a negatively skewed or a positively skewed and the hrf can take different shapes. Further, it is appropriate for modeling both over- and under-dispersed data. The ML method was used for estimating the family parameters. A simulation study was carried out to evaluate the performance and accuracy of the estimators. Finally, the flexibility and applicability of the discrete exponentiated-G family in real life was illustrated by applying to real datasets. The discrete exponentiated inverted Topp – Leone distribution appears to be more suitable for modeling real data sets and is a better alternative to some other distributions. We wish the proposed model is applied to a wider range of applications in medicine, engineering, and other fields of research fields.

References

1. AbdElaziz, M., Abdelrahman, N. S., Hassan, A. N. and Mohamed, M. O. (2022). discrete generalized inverted exponential distribution. Mathematical Problems in Engineering. https://doi.org/10.1155/2022/3029932.

2. Ahmad, Z., Rashid, A. and Jan, T. R. (2017). A new discrete flexible model: Geeta - Kum-

maraswamy distribution. *International Journal of A defense Research Science and Engineering*, 6(3), 2319-8254.

3. Alamatsaz, M. H., Dey, S., Dey, T. and Shams Harandi, S. (2016). Discrete generalized Rayleigh distribution. Pakistan Journal of Statistics, 32(1), 1-20.
4. AL-Huniti, A. A. and AL-Dayian, G. R. (2012). Discrete Burr Type III distribution. American Journal of Mathematics and Statistics. 2(2), 145-152.
5. Almetwally, E. M. and Ibrahim, G. M. (2020). Discrete Alpha Power Inverse Lomax distribution with application of COVID-19 Data. International Journal of Applied Mathematics & Statistical Science, 9(6), 2319-3972.
6. Almetwally, E. M., Abdo, D. A., Hafez, E. H., Jawa, T. M., Sayed-Ahmed, N., & Almongy, H. M. (2022). The new discrete distribution with application to COVID-19 Data. Results in Physics, 32, 104987.
7. Altun, E., El-Morshedy, M. and Eliwa, M. S. (2022). A study on discrete Bilal distribution with properties and applications on integer valued autoregressive process. REVSTAT-Statistical Journal, 20(2), 501-528.
8. Alzaghal, A., Famoye, F. and Lee, C. (2013). Exponentiated T-X family of distributions with some applications. International Journal of Statistics and Probability, 2(3), 31–49.
9. Arnold, B., Balakrishnan, N. and Najaraja, H. N. (2008). A First Course in Order Statistics. *John-Wiley and Sons*, New York.
10. Chakraborty, S. and Chakravorty, D. (2016). A New Discrete Probability distribution with integer support on. Communication in Statistics – Theory and Methods, 45(2), 492-505.
11. Chotedelok, Y. and Bodhisuwan, W. (2020). The discrete exponentiated Pareto distribution: Its properties and application. Journal of Applied Science, 19(2), 24-36.
12. Cordeiro, G., M. and de Castro, M. (2011). A new family of generalized distribution. Journal of Statistical Computation and Simulation, 81(7), 883-898.
13. El-Deep, A. S. Ul -Haq, M. and Babar, A. (2021). A discrete analog of inverted Topp-Leone distribution: properties, estimation and applications. International Journal of Analysis and Applications. 19(5), 695-708.
14. Eliwa, M. S., Alhussain, Z. A., & El-Morshedy, M. (2020). Discrete Gompertz-G family of distributions for over-and under-dispersed data with properties, estimation, and applications. Mathematics, 8(3), 358.
15. Eliwa, M. S., El-Morshedy, M., & Yousof, H. M. (2022). A discrete exponential generalized-G family of distributions: Properties with Bayesian and non-Bayesian estimators to model medical, engineering and agriculture data. Mathematics, 10(18), 3348.
16. Eliwa, M. S., Tahir, M. H., Hussain, M. A., Almohaimeed, B., Al-Bossly, A. and El-Morshedy, M. (2023). Univariate probability-G classes for scattered samples under different forms of hazard: continuous and discrete version with their inferences tests. Mathematics, 11(13), 2929.
17. El-Morshedy, M., Eliwa, M. S., & Nagy, H. (2020). A new two-parameter exponentiated discrete Lindley distribution: properties, estimation and applications. Journal of applied statistics, 47(2), 354-375.

18. El-Morshedy, M., Eliwa, M. S., & Tyagi, A. (2022). A discrete analogue of odd Weibull-G family of distributions: properties, classical and Bayesian estimation with applications to count data. Journal of Applied Statistics, 49(11), 2928-2952.

19. Eugene, N., Lee, C., & Famoye, F. (2002). Beta-normal distribution and its applications. Communications in Statistics-Theory and methods, 31(4), 497-512.

20. Gómez-Déniz, E., & Calderin-Ojeda, E. (2011). The discrete Lindley distribution: properties and applications. Journal of statistical computation and simulation, 81(11), 1405-1416.

21. 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.

22. Hassan, A. S., Elgarhy, M., & Ragab, R. (2020). Statistical properties and estimation of inverted Topp-Leone distribution. J. Stat. Appl. Probab, 9(2), 319-331.

23. Hegazy, M. A., EL-Helbawy, A. A., Abd EL-Kader, R. E., & AL-Dayian, G. R. (2018). Discrete Gompertz distribution: properties and estimation. Journal of Biostatistics and Biometric Applications, 3(3), 307.

24. Hegazy, M., Abd EL-Kader, R., Gannat, A. D., & EL-Helbawy, A. A. A. (2022). Discrete inverted Kumaraswamy distribution: Properties and estimation. Pakistan Journal of Statistics and Operation Research, 297-328.

25. Hegazy, M. A., EL-Helbawy, A. A., Abd EL-Kader, R. E. and AL-Dayian, G. R. (2020). Bayesian estimation and prediction of discrete Half Logistic distribution. Annual International conference on statistics and modeling in human and social science.

26. Helmy, N. (2018). Discrete Burr Type II distribution. An Academic Periodical Referreed Journal, AL-Azhar University, 20(2), 1-21.

27. Hussain, T., Aslam, M., & Ahmad, M. (2016). A two parameter discrete Lindley distribution. Revista Colombiana de Estadistica, 39(1), 45-61.

28. Ibrahim, M., Ali, M. M., & Yousof, H. M. (2021). The discrete analogue of the Weibull G family: properties, different applications, Bayesian and non-Bayesian estimation methods. Annals of Data Science, 1-38. To appear

29. Inusah, S., & Kozubowski, T. J. (2006). A discrete analogue of the Laplace distribution. Journal of statistical planning and inference, 136(3), 1090-1102.

30. Kemp, A. W. (2004) Classes of discrete lifetime distributions. Communications in Statistics Theory and Methods. 33(12), 3069-3093.

31. Kemp, A. W. (1997). Characterizations of a discrete normal distribution. Journal of Statistical Planning and Inference, 63(2), 223-229.

32. Korkmaz, M. C., Yousof, H. M., & Hamedani, G. G. (2018). The exponential Lindley odd log-logistic-G family: properties, characterizations and applications. Journal of Statistical Theory and Applications. 17(3), 554-571.

33. Lawless, J. F. (2011). Statistical models and methods for lifetime data. John Wiley & Sons.

34. Maiti, S. S., Dey, M. and MONDAL, S. S. (2018). Discrete X-Gamma distribution: properties, estimation, and an application to the collective risk model. Journal of Reliability and Statistical Studies, 11, 117-132.

35. Murthy, D. N. P., Xie, M. and Jiang, R. (2004). Weibull models. John Wily & Sons, New York.
36. Nadarajah, S. (2005). The exponentiated Gumbel distribution with climate application. Environ Metrics; 17: Wiley Inter Science. https://doi.org/10.1002/env.739.
37. Nakagawa, T. and Osaki, S. (1975). The discrete Weibull distribution. IEEE Transactions on Reliability, 24(5), 300-301.
38. Para, B. A. and Jan, T. R. (2014). Discrete generalized Burr-Type XII distribution. Journal of Modern Applied Statistical Method, 13(2), 244-258.
39. Rényi, A. (1961, January). On measures of entropy and information. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics (Vol. 4, pp. 547-562). University of California Press.
40. Roa, G. S. and Mbwambo, S. (2019). Exponentiated inverse Rayleigh distribution and an application to coating weights of iron sheets data. Journal of Probability and Statistics, 2019(1), 1-13.
41. Rohatgi, V. K. and Saleh, E. A. K. (2001). An Introduction to probability and statistics. 2nd Edition, John- Wiley and Sons, New York.
42. Roy, D. (2003). Discrete normal distribution. *Communication in Statistics- Theory and Methods*, 32(10), 1871-1883.
43. Tyagi, A., Choudhary, N. and Singh, B. (2020). Inferences on discrete Rayleigh distribution under Type-II censored data using a tie-run approach, Int. Journal Agriculture. Stat. Sci. 16, 939–951.

SASAR
Open Access Journal