Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/130119

Discrete Half-Logistic Distribution: Statistical Properties, Estimation, and Application

Abd-Elmonem A. M. Teamah, Ahmed A. Elbanna, and Ahmed M. Gemeay[*]

Department of Mathematics, Faculty of Science, Tanta University, Tanta 31527, Egypt

Received: 7 Jun. 2023, Revised: 25 Jun. 2023, Accepted: 20 Jul. 2023
Published online: 1 Jan. 2024

Abstract: This article presented a novel discrete distribution with one parameter derived by the discretization approach and called the discrete half-logistic distribution. Its probability mass function and hazard function have different shapes. A variety of its statistical properties, including moments, probability generating function, incomplete moments, and order statistics, were determined mathematically. Maximum likelihood, moments, and proportion estimation methods were used to estimate its parameter. A simulation study conducts to check the various estimating method's performance. By using a real data set, its flexibility is assessed. Lastly, it can model count data sets in a way that is compared with other distributions that are already in the scientific literature.

Keywords: Discrete distribution; Half-logistic distribution; Estimation; Simulation

1 Introduction

Statistical models are vital in analyzing count data across diverse subfields within the applied sciences, encompassing domains such as ecology, environmental studies, and the insurance sector. Count data refers to the numerical representation of discrete events or occurrences, focusing on the frequency or magnitude of these events rather than their continuous measurement. These models provide a framework for understanding the underlying patterns and relationships within count data, enabling researchers to extract meaningful insights and make informed decisions in their fields of study. In ecology and environmental studies, statistical models facilitate the examination of various ecological phenomena, such as species abundance, population growth, and biodiversity assessments. These models aid in elucidating the complex interactions between organisms and their environment, allowing researchers to quantify ecological processes and investigate the impact of environmental factors on species distribution and ecosystem dynamics. Moreover, statistical modeling finds applications in studying environmental risks, monitoring pollution, and assessing the effectiveness of conservation strategies.

In the insurance sector, statistical models are fundamental for risk assessment, pricing, and underwriting purposes. By analyzing historical data on insurance claims, policyholders' characteristics, and external factors, these models help insurance companies quantify and manage risks associated with potential losses. Actuarial modeling, for instance, employs statistical techniques to estimate the likelihood and severity of specific events, such as accidents, natural disasters, or health-related claims. Such analyses enable insurers to develop appropriate pricing strategies, determine reserves, and optimize risk portfolios.

It is required to propose new discrete distributions that have properties that are wanted and desirable qualities. Most occurrences in nature and other scientific domains each have their unique set of distinguishing traits. Modeling events such as earthquakes, car accidents, the number of persons killed by a disease, and the number of landslides that occur may be done using discrete probability distributions. Researchers have developed more flexible distributions to lower the number of estimate mistakes caused by modeling these data sets.

Traditional discrete distributions like geometric and Poisson and others like them have restricted property sets. As a result, certain discrete distributions based on well-known continuous models have been developed. Consequently, discrete distributions have garnered great interest in the statistical literature. As an example, Nakagawa and Osaki [1], El-Morshedy et al. [2], Kundu and Nekoukhou [3], Roy [4], Krishna and Pundir [5], Nekoukhou and Bidram [6], El-Morshedy et al.

* Corresponding author e-mail: ahmed.gemeay@science.tanta.edu.eg

[7], Chakraborty et al. [8], Eldeeb et al. [9], Para and Jan [10], Almetwally and Ibrahim [11], Hassan et al. [12], Barbiero and Hitaj [13], Opone et al. [14], and references cited therein.

Even though the literate provides a number of different distributions that can be used to evaluate count data, there is still a need to develop a more adaptable and appropriate distribution for use in various conditions. This research aims to present a flexible version of discrete distributions called the discrete half-logistic distribution (DHLD). It has closed forms for both the survival function (SF) and the hazard function (HF), the HF of the DHLD has multiple shapes, and as a result, the parameters of the underlying distribution may be modified to fit the majority of different data sets. It has been shown that this model performs better than both the conventional models (geometric and Poisson) and some of the recently created models in the literature.

This paper is organized as follows. In Section 2, we explain the method used to derive the PMF, cumulative distribution function (CDF), and HF of the proposed model along with presenting its PMF and its HF graphically. Section 3 contains different statistical properties of the proposed model. Three different estimation methods were discussed in detail to estimate the proposed model parameter in Section 4, while the performance of these methods was tested by a randomly generated data set in Section 5. A real data set was used to explain the flexibility of the proposed model to other compared models in Section 6. Finally, concluding remarks for this work were presented in Section 7.

2 The DHLD formulation

The survival function (SF) of the continuous version of the half-logistic distribution is defined as follows

$$S(x;a) = 2^a \left(a\left(e^x - 1\right) + 2\right)^{-a}, \; x > 0. \tag{1}$$

Let X be a discrete random variable, then the PMF of X is defined as follows

$$P(X = x) = S(x) - S(x+1), \;\; x = 0, 1, 2, \ldots \tag{2}$$

By using both Equations (2) and (1), the PMF of DHLD is defined as follows

$$P(X = x) = 2^a a^{-a} \left\{ e^{-ax} \left[1 - \left(\frac{a-2}{a}\right) e^{-x}\right]^{-a} - e^{-a(x+1)} \left[1 - \left(\frac{a-2}{a}\right) e^{-(x+1)}\right]^{-a} \right\}, \tag{3}$$

where the parameter a changes the shape of the distribution. The PMF (3) has different shapes obtained by changing values of the parameter a as we can see in Figure 1.

The CDF of DHLD is determined as follows

$$F(x) = P(X \leq x) = 1 - 2^a \left(a\left(e^{x+1} - 1\right) + 2\right)^{-a}, \tag{4}$$

so the HF of the DHLD is defined as follows

$$h(x) = \left(a\left(e^x - 1\right) + 2\right)^{-a} \left(a\left(e^{x+1} - 1\right) + 2\right)^{a} - 1, \tag{5}$$

and the possible shapes of the HZ for the DHLD are displayed in Figure 2.

3 Statistical properties

This section entails an assessment of the statistical properties inherent to the proposed model, wherein a mathematical approach is employed to derive each of these attributes. The subsequent paragraphs will discuss the statistical characteristics and their corresponding mathematical derivations.

3.1 Moments with related measures

The DHLD ordinary moments are defined as follows

$$\mu_r' = 2^a a^{-a} \sum_{k=0}^{\infty} \binom{a+k-1}{a-1} \left(\frac{a-2}{a}\right)^k \left(1 - e^{-(a+k)}\right) \sum_{x=0}^{\infty} x^r e^{-x(a+k)}. \tag{6}$$

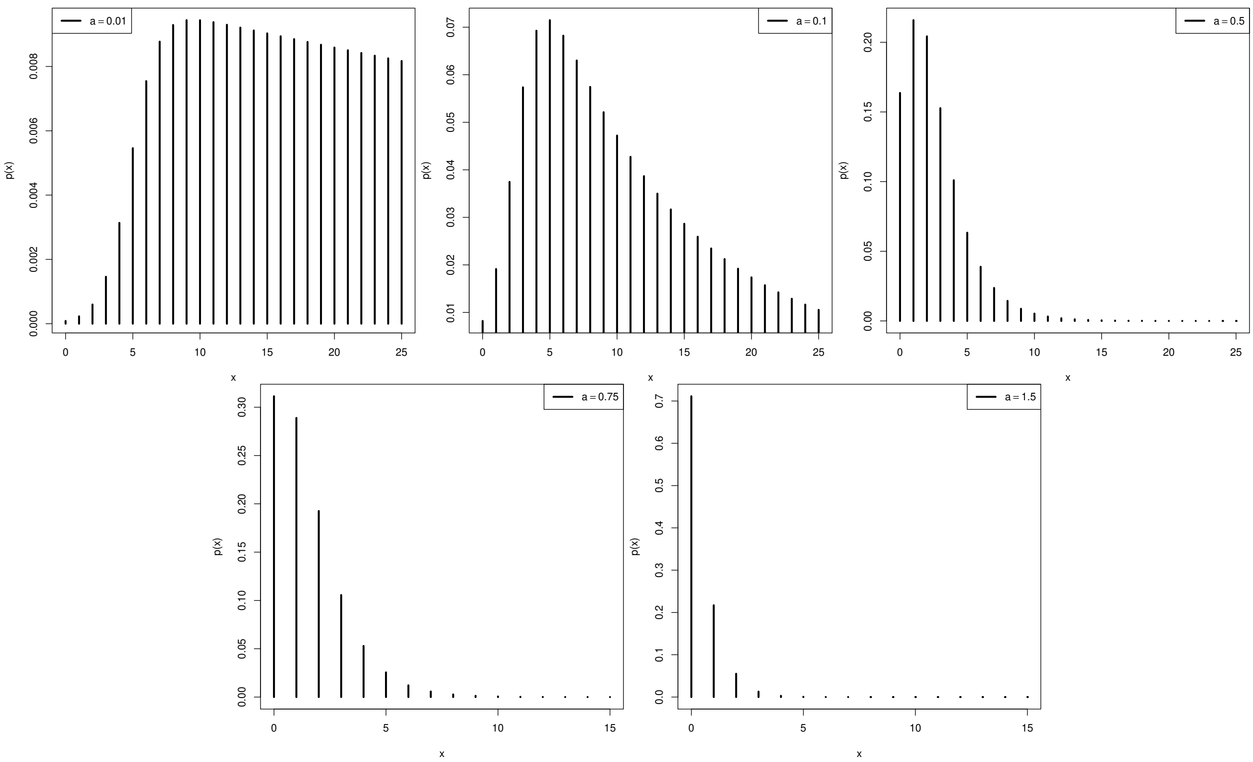

Fig. 1: Plots of the PMF of the DHLD by different parametric values.

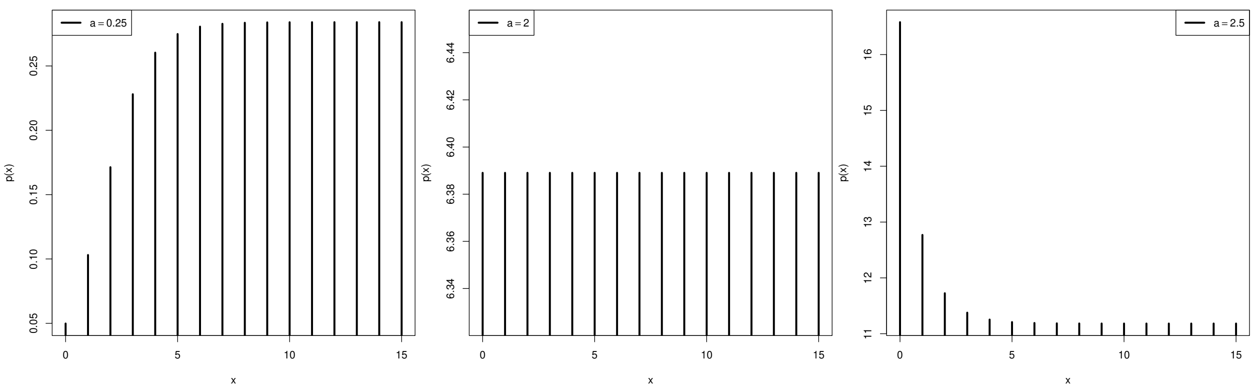

Fig. 2: Plots of the HF of the DHLD by different parametric values.

By using Equation (6), we have the first four moments of the DHLD as follows

$$\mu_1' = 2^a a^{-a} \sum_{k=0}^{\infty} \frac{\left(\frac{a-2}{a}\right)^k \binom{a+k-1}{a-1}}{e^{a+k}-1},$$

$$\mu_2' = 2^a a^{-a} \sum_{k=0}^{\infty} \frac{\left(\frac{a-2}{a}\right)^k \left(e^{a+k}+1\right) \binom{a+k-1}{a-1}}{\left(e^{a+k}-1\right)^2},$$

$$\mu_3' = 2^a a^{-a} \sum_{k=0}^{\infty} \frac{\left(\frac{a-2}{a}\right)^k \left(4e^{a+k} + e^{2(a+k)} + 1\right) \binom{a+k-1}{a-1}}{\left(e^{a+k}-1\right)^3},$$

$$\mu_4' = 2^a a^{-a} \sum_{k=0}^{\infty} \frac{\left(\frac{a-2}{a}\right)^k \left(11e^{a+k} + 11e^{2(a+k)} + e^{3(a+k)} + 1\right) \binom{a+k-1}{a-1}}{\left(e^{a+k}-1\right)^4},$$

which are used to determine variance, coefficient skewness, and coefficient kurtosis, respectively, by the following relation

$$\sigma^2 = \mu_2' - (\mu_1')^2,$$
$$\delta = \frac{2(\mu_1')^3 - 3\mu_2'\mu_1' + \mu_3'}{\left(\mu_2' - (\mu_1')^2\right)^{3/2}},$$
$$\varepsilon = \frac{-3(\mu_1')^4 + 6\mu_2'(\mu_1')^2 - 4\mu_3'\mu_1' + \mu_4'}{\left(\mu_2' - (\mu_1')^2\right)^2}.$$

The index of dispersion (ID) and the coefficient of variation (CV) are, respectively, determined as follows

$$ID = \frac{\sigma^2}{\mu_1'},\; CV = \frac{\sigma}{\mu_1'}.$$

Table 1 displays the DHLD's numerical values for mean ($\mu = \mu_1'$), σ^2, δ, ε, ID, and CV. These numerical values are displayed for a variety of a values. From Table 1, we conclude that

–μ and σ^2 of the DHLD both tend to increase as the parameter a moves closer and closer to 1.
–The DHLD model is suitable for use in the modeling of data that is right-skewed.
–Both of δ and ε has an increasing-shaped.
–The nature of the DHLD may be described as leptokurtic.
–The ID has an inverse J-shape with an increasing value of a.
–The CV has an increasing shape with an increasing value of a.

Table 1: Some numerical values for the DHLD.

measure↓ a ⟶	1	1.3	1.6	1.9	2.2	2.5	2.8	3.2	3.5
μ	0.928327	0.543267	0.320358	0.187628	0.108307	0.0613647	0.0340556	0.0150201	0.00792773
σ^2	1.33308	0.711178	0.393556	0.219879	0.122355	0.0673358	0.0365155	0.0157349	0.00819946
δ	2.71545	3.71115	5.4011	8.21271	12.9647	21.2024	35.8895	76.2665	139.312
ε	6.83365	7.91816	9.84765	13.0467	18.3611	27.3871	43.1661	85.6778	151.071
ID	1.43601	1.30908	1.22849	1.17189	1.1297	1.09731	1.07223	1.04759	1.03428
CV	1.24373	1.5523	1.95825	2.49916	3.22963	4.22868	5.61112	8.35139	11.422

The nth central moment of the DHLD is determined as follows

$$\mu_n = E(x-\mu)^n = \sum_{k=0}^{\infty} (-1)^k \binom{n}{k} \mu_1'^k \mu_{n-k}'$$

3.2 Probability generating function

The probability generating function of DHLD is determined as follows

$$PG(z) = \sum_{x=0}^{\infty} z^x p(x) = 2^a a^{-a} \sum_{k=0}^{\infty} \binom{a+k-1}{a-1} \left(\frac{a-2}{a}\right)^k \left(1 - e^{-(a+k)}\right) \sum_{x=0}^{\infty} z^x e^{-x(a+k)}$$
$$= 2^a a^{-a} \sum_{k=0}^{\infty} \binom{a+k-1}{a-1} \left(\frac{a-2}{a}\right)^k \left(1 - e^{-(a+k)}\right) \frac{e^{a+k}}{e^{a+k} - z},$$

by replacing z with e^t and e^{it}, we have both the moment generating function and the characteristic function of the DHLD.

3.3 Incomplete moments with related measures

The rth incomplete moments of the DHLD are defined as follows

$$\begin{aligned}\omega_r(t) &= \sum_{x=0}^{t} x^r p(x) = 2^a a^{-a} \sum_{k=0}^{\infty} \binom{a+k-1}{a-1} \left(\frac{a-2}{a}\right)^k \left(1-e^{-(a+k)}\right) \sum_{x=0}^{t} x^r e^{-x(a+k)} \\ &= 2^a a^{-a} \sum_{k=0}^{\infty} \binom{a+k-1}{a-1} \left(\frac{a-2}{a}\right)^k \left(e^{a+k}-1\right) e^{(t+2)(-(a+k))} \\ &\times \left(e^{(t+1)(a+k)} \mathrm{Li}_{-m}\left(e^{-a-k}\right) - \Phi\left(e^{-a-k}, -m, t+1\right)\right), \end{aligned} \tag{7}$$

where $\Phi(z,s,a) = \sum_{k=0}^{\infty} \frac{z^k}{(a+k)^s}$ and $\mathrm{Li}_n(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^n}$.

The mean deviation of the DHLD about the mean and about the median (M) are, respectively, determined as follows

$$\Omega_1 = 2\mu_1' F(\mu_1') - 2\omega_1(\mu_1'),\ \Omega_2 = \mu - 2\omega_1(M),$$

where $F(\cdot)$ is the CDF (4) and $\omega_1(\cdot)$ is the first incomplete moments.

3.4 Order statistics

The density function of the ith order statistic (OS) of the DHLD is determined as follows

$$\begin{aligned} f_{i:n}(x) &= \sum_{r=i}^{n} \binom{n}{r} \left[F(x)^r(1-F(x))^{n-r} - F(x-1)^r(1-F(x-1))^{n-r}\right] = 2^{a(n-i)} \binom{n}{i} \\ &\times \left[\left(\left(a\left(e^{x+1}-1\right)+2\right)^a - 2^a\right)^i \left(a\left(e^{x+1}-1\right)+2\right)^{-an} H_1 \right. \\ &\left. - \left(\left(a\left(e^{x}-1\right)+2\right)^a - 2^a\right)^i \left(a\left(e^{x}-1\right)+2\right)^{-an} H_2\right]. \end{aligned} \tag{8}$$

and its corresponding CDF is determined as follows

$$\begin{aligned} F_{i:n}(x) &= \sum_{r=i}^{n} \tbinom{n}{r} (F(x))^r (1-F(x))^{n-r} \\ &= \binom{n}{i} \left(1 - \left(\frac{1}{2}a\left(e^{x+1}-1\right)+1\right)^{-a}\right)^i \left(\frac{1}{2}a\left(e^{x+1}-1\right)+1\right)^{-a(n-i)} H_1, \end{aligned} \tag{9}$$

where $H_1 = {}_2F_1\left(1, i-n; i+1; 1-\left(\frac{1}{2}a\left(-1+e^{x+1}\right)+1\right)^a\right)$ and $H_2 = {}_2F_1\left(1, i-n; i+1; 1-\left(\frac{1}{2}a\left(-1+e^{x}\right)+1\right)^a\right)$ are hyper geometric functions.

The PDFs and CDFs of the minimum OS, (L_n), and maximum OS, (Z_n), can be determined from Equations (8) and (9) with $i=1$ and $i=n$, respectively. The limiting distributions of (L_n) and (Z_n) are, respectively, determined by Theorem 2.1.5 and 2.1.1 in [15] as follows

$$\lim_{n\to+\infty} P(L_n < d_n\, x) = 1 - e^{-1},\ d_n = F^{-1}\left(-\frac{1}{n}\right),$$

and

$$\lim_{n\to+\infty} P(Z_n < K_n\, x) = \begin{cases} 0, & x=0, \\ e^{-1}, & x=1, \\ 1, & x=2,3,\ldots, \end{cases} \qquad K_n = F^{-1}\left(1-\frac{1}{n}\right).$$

4 Estimation methods

In this section, we will examine maximum likelihood estimation (MLE), method of moment estimation (MME), and proportion estimation (PE) as potential solutions to the problem of determining the values of the DHLD's undetermined parameters.

Consider $X_1,\dots,X_n$ is a random sample from the DHLD, the corresponding log-likelihood function of PMF (3) is defined as follows

$$L=\sum_{i=1}^{n}\log\left(2^a\left(a\left(e^{x_i}-1\right)+2\right)^{-a}-2^a\left(a\left(e^{x_i+1}-1\right)+2\right)^{-a}\right), \tag{10}$$

then by differentiating Equation (10) with respect to a and equating it by zero, we have

$$\begin{aligned}&\sum_{i=1}^{n}\frac{1}{2^a\left(a\left(e^{x_i}-1\right)+2\right)^{-a}-2^a\left(a\left(e^{x_i+1}-1\right)+2\right)^{-a}}\\&\times\Big\{2^a\log(2)\left(a\left(e^{x_i}-1\right)+2\right)^{-a}-2^a\log(2)\left(a\left(e^{x_i+1}-1\right)+2\right)^{-a}\\&-2^a\Big[-ae^{x_i+1}\left(a\left(e^{x_i+1}-1\right)+2\right)^{-a-1}+a\left(a\left(e^{x_i+1}-1\right)+2\right)^{-a-1}\\&-\left(a\left(e^{x_i+1}-1\right)+2\right)^{-a}\log\left(a\left(e^{x_i+1}-1\right)+2\right)\Big]+2^a\Big[-ae^{x_i}\left(a\left(e^{x_i}-1\right)+2\right)^{-a-1}\\&+\ a\left(a\left(e^{x_i}-1\right)+2\right)^{-a-1}-\left(a\left(e^{x_i}-1\right)+2\right)^{-a}\log\left(a\left(e^{x_i}-1\right)+2\right)\Big]\Big\}=0,\end{aligned}$$

we have a solution of a by solving the last equation, but unfortunately, there is no explicit form for this equation, so we can use a numerical solution that direct maximizes the log-likelihood function.

For estimating the DHLD unknown parameter a by MME, we should first equate the population mean to the corresponding sample mean as follows

$$\frac{1}{n}\sum_{i=1}^{n}x_i-2^a a^{-a}\sum_{k=0}^{\infty}\frac{\left(\frac{a-2}{a}\right)^k\binom{a+k-1}{a-1}}{e^{a+k}-1}=0, \tag{11}$$

by solving the previous equation, we have the estimated parameter a by MME.

Consider $x_1,\dots,x_n$ be a random sample from the DHLD, which has only one parameter, and the one indicator function is defined as follows

$$I(x_i)=\begin{cases}1, & x_i=0\\ 0, & otherwise,\end{cases} \tag{12}$$

where $B=\sum_{i=1}^{n}I(x_i)$ denotes the count of zero observations in the sample.

Then by using Equations (4) and (12), we have

$$P(X\le 0)=\frac{B}{n},$$

then

$$1-2^a((e-1)a+2)^{-a}=\frac{B}{n},$$

we have the estimated parameter a by PE by solving the previous equation.

5 Numerical simulation

This section demonstrates the performance of the offered estimate methodologies in Section 4 by using simulated data sets to estimate the parameter of the DHLD. We generate different random samples ($N=1000$) from the DHLD of different sample sizes $(20,80,150,300,500)$ by using different parameter values as initial values for our generation ($a=(0.1,0.5,0.8,1.2,1.5,2.5)$). After each generation, we determine the estimated parameter ($\hat{a}$) along with its bias (B) and mean square error (MSE), and mean relative estimates (MRE). Finally, we determine the average of these measures, which are presented in Tables 2–4; from these tables, we conclude the following results

–The estimated parameter $\hat{a}$ has the consistency property.
–For the three estimation methods, the B, MSE, and MRE have a decreasing shape as the sample size increase.
–For fixed sample size, the B, MSE, and MRE of the three different estimation methods increase as the parameter values increase.
–The three different estimation methods' B, MSE, and MRE have large values if the sample size is small ($n = 20$) and the parameter value is large ($a = 2.5$).

Table 2: Numerical values of the average of both $\hat{a}$, B, MSE, and MRE.

		$a = 0.1$			$a = 0.5$		
n	Est.	MLE	MME	PE	MLE	MME	PE
20	$A_{\hat{a}}$	0.10381032	0.04248311	0.10496382	0.51167466	0.48488312	0.51055588
	A_B	0.01658751	0.11072029	0.01753844	0.06980073	0.13541293	0.07008139
	A_{MSE}	0.00047015	0.01307526	0.00050746	0.00807306	0.03022653	0.00771704
	A_{MRE}	0.16587505	1.10720287	0.17538444	0.13960147	0.27082587	0.14016278
80	$A_{\hat{a}}$	0.1005217	0.06972104	0.10122178	0.50359118	0.49734124	0.50047697
	A_B	0.00816773	0.07042358	0.00842866	0.03268806	0.06127153	0.03265338
	A_{MSE}	0.00010716	0.00627185	0.00011544	0.00169184	0.00583382	0.00173353
	A_{MRE}	0.0816773	0.70423582	0.08428658	0.06537612	0.12254305	0.06530677
150	$A_{\hat{a}}$	0.10027133	0.07828372	0.10078011	0.50041998	0.49773315	0.50387915
	A_B	0.00605981	0.0499898	0.00599685	0.02332841	0.04388985	0.02619673
	A_{MSE}	5.806e-05	0.00403472	5.783e-05	0.00084789	0.00302352	0.00108147
	A_{MRE}	0.06059813	0.499898	0.05996854	0.04665682	0.0877797	0.05239346
300	$A_{\hat{a}}$	0.10037704	0.09262336	0.10014466	0.50030236	0.49935219	0.50074009
	A_B	0.00417221	0.03188567	0.00432398	0.01710425	0.0315892	0.01677044
	A_{MSE}	2.798e-05	0.00177986	2.867e-05	0.00046645	0.00159547	0.00043449
	A_{MRE}	0.04172213	0.31885674	0.04323977	0.03420849	0.06317839	0.03354089
500	$A_{\hat{a}}$	0.10016362	0.09637877	0.10018917	0.50078277	0.49911329	0.49980601
	A_B	0.00320219	0.02063464	0.00335311	0.01316718	0.02454106	0.01276683
	A_{MSE}	1.651e-05	0.00078133	1.784e-05	0.00026732	0.00093949	0.00025158
	A_{MRE}	0.03202186	0.20634642	0.03353114	0.02633436	0.04908212	0.02553366

6 Real data analysis

This section aims to explore the adaptability of the DHLD in terms of fitting real data set. The analyzed data set consists of $n = 64$ observations representing the Waiting time between eruptions (in seconds), and it is available on www.statsci.org/data/oz/kiama.html. The use of the real data set demonstrates the adaptability of the DHLD in comparison to a variety of well-known distributions like Poisson distribution (PD), Geometric distribution (GD), discrete Lindley distribution (DLD) [16], discrete gamma distribution (DGD) [17], discrete Weibull distribution (DWD) [1], discrete Ramos-Louzada distribution (DRLD) [18], discrete Rayleigh distribution (DRD) [4], discrete mixture of gamma and exponential distribution (DMGED) [19], discrete Burr distribution (DBD) [5], discrete Pareto distribution (DPD) [5], discrete Burr-Hatke distribution (DBHD) [20] and discrete Lomax distribution (DLoD) [21].

To compare between models, we used different measures such as Akaike IC (A), the correct Akaike information criterion (C), the Bayesian information criterion (B), and the Hannan information criterion (H) with Kolmogorov–Smirnov (KS) statistics and its corresponding p-value. Table 5 presents parameter estimates (standard error) and $-L$ for the investigated data set, in addition to the comparison calculated measures obtained by utilizing MLE. Compared to other competing distributions, the DHLD provides a close fit to the modeled data set. This is seen by the fact that the values of all the DHLD measures in this table are small, except for the p-value, which has the highest value.

The log-likelihood function behavior with the DHLD estimated parameter $\hat{a}$ is provided graphically in Figure 3 (left panel); we see that it has a unimodal shape with maximum values of this curve at the estimated parameter $\hat{a}$ value, we

Table 3: Numerical values of the average of both $\hat{a}$, B, MSE, and MRE.

		$a=0.8$			$a=1.2$		
n	Est.	MLE	MME	PE	MLE	MME	PE
20	$A_{\hat{a}}$	0.81429289	0.8077769	0.82169913	1.22293254	1.22271265	1.22595772
	A_B	0.10075187	0.13734235	0.09920438	0.15468604	0.17356367	0.14747972
	A_{MSE}	0.01656792	0.03041603	0.01676007	0.04117207	0.04834595	0.0364793
	A_{MRE}	0.12593984	0.17167794	0.12400547	0.12890503	0.14463639	0.12289976
80	$A_{\hat{a}}$	0.80141829	0.80359507	0.80497585	1.21353299	1.20591599	1.20772017
	A_B	0.05052432	0.06992358	0.05074923	0.07464638	0.08598832	0.07239286
	A_{MSE}	0.00400899	0.00776049	0.00410343	0.00905316	0.01211432	0.0086314
	A_{MRE}	0.0631554	0.08740447	0.06343654	0.06220532	0.07165693	0.06032738
150	$A_{\hat{a}}$	0.80436003	0.79752283	0.79946251	1.20373332	1.20311528	1.20418113
	A_B	0.036321	0.05026319	0.03623703	0.05430534	0.0659743	0.05302421
	A_{MSE}	0.00214021	0.00397953	0.00206915	0.00457814	0.00679678	0.00450734
	A_{MRE}	0.04540125	0.06282898	0.04529629	0.04525445	0.05497858	0.04418684
300	$A_{\hat{a}}$	0.8008973	0.8000342	0.80148819	1.2012397	1.2027132	1.20266068
	A_B	0.02560719	0.03690633	0.0260917	0.03716314	0.04397916	0.03718448
	A_{MSE}	0.0010386	0.00214611	0.00106173	0.0021773	0.00303913	0.00210858
	A_{MRE}	0.03200898	0.04613291	0.03261463	0.03096929	0.0366493	0.03098706
500	$A_{\hat{a}}$	0.80034359	0.80260392	0.79992791	1.20149104	1.20233165	1.20159212
	A_B	0.0202126	0.02906592	0.01974468	0.02872143	0.03297201	0.03024494
	A_{MSE}	0.00063984	0.001327	0.00061399	0.00127439	0.00172791	0.00141086
	A_{MRE}	0.02526575	0.03633239	0.02468086	0.02393453	0.02747668	0.02520412

conclude that we have a global maximum value of our estimated parameter. We use the Nmaximize function in Wolfram Mathematica Software, which attempts to find the global maximum solution, and that reveals with the result displayed in Figure 3 (left panel). In the same figure (right panel), we graphically provide the existence and uniqueness of the estimated parameter $\hat{a}$. The probability-probability (PP) plots of the proposed model with other compared models are presented in Figure 4, which support the results reported in Table 5. The estimated CDFs were presented graphically in Figure 5, which also support the results reported in Table 5.

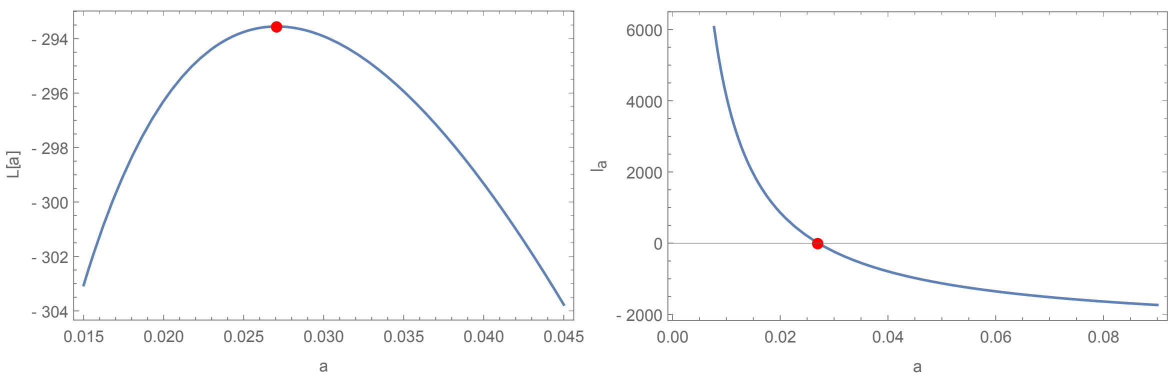

Fig. 3: Plots of log-likelihood function with existence and uniqueness of estimated DHLD parameter for the eruptions data set.

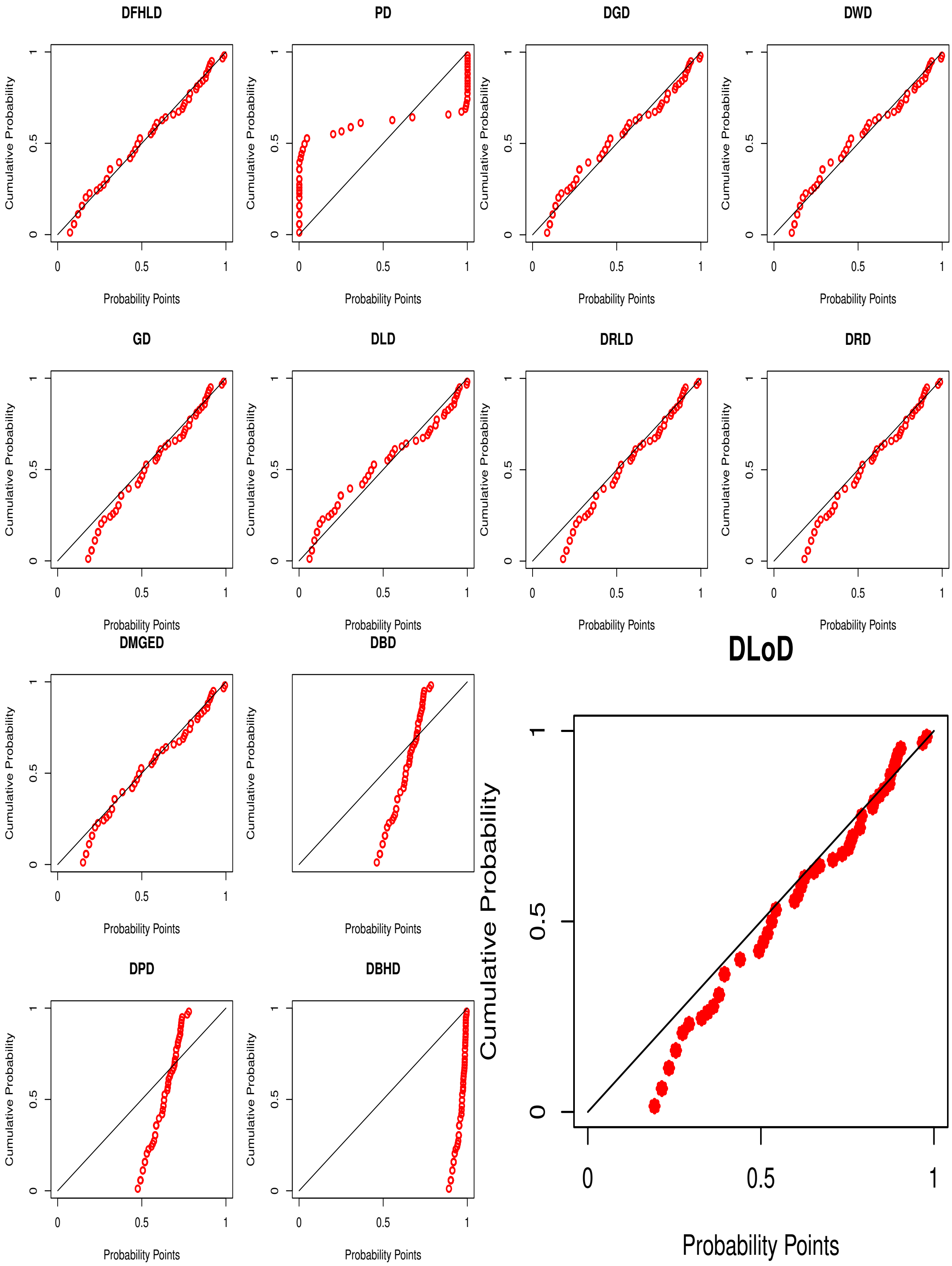

Fig. 4: P-P plot of the DHLD and other compared models for the eruptions data set.

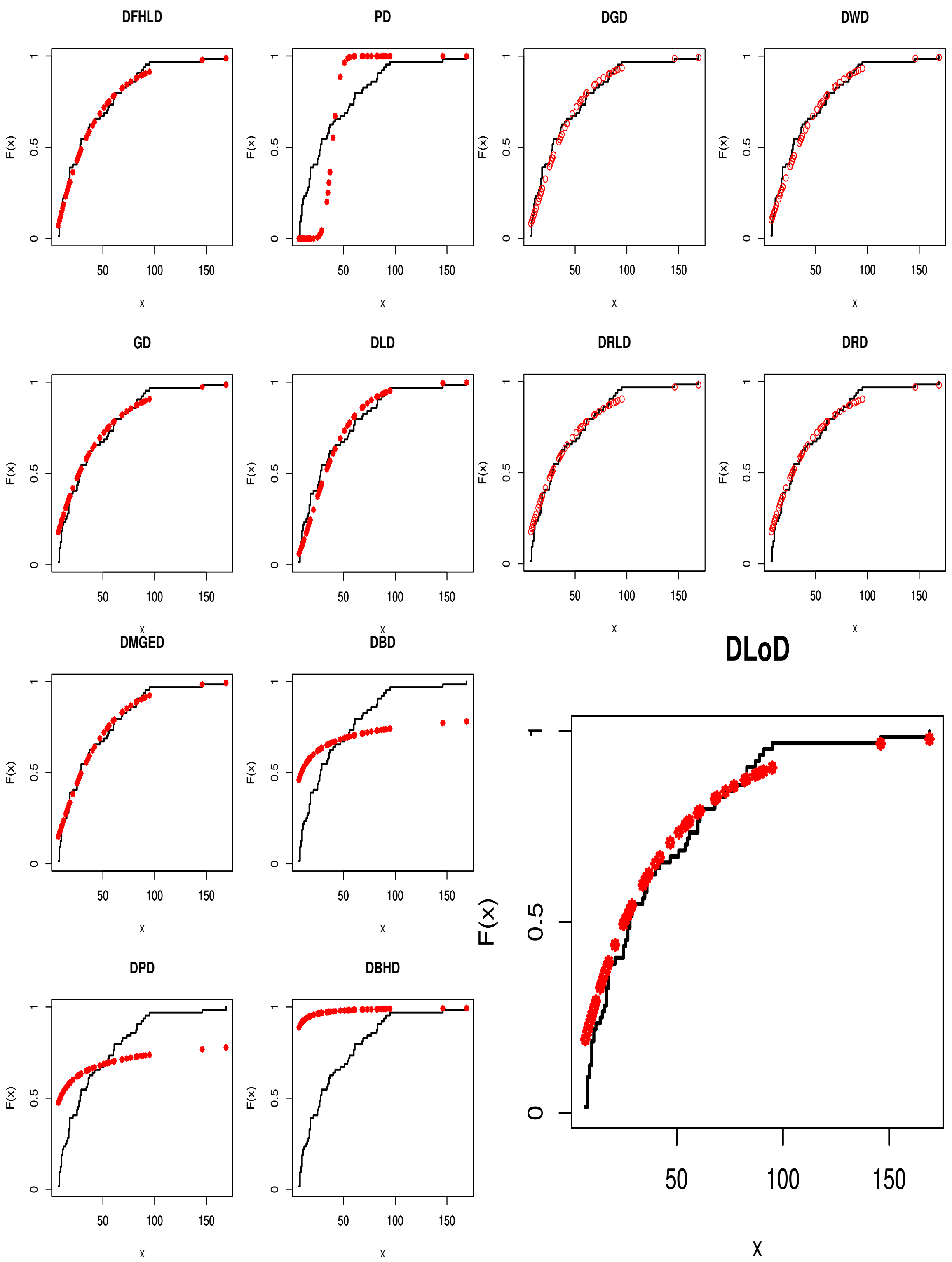

Fig. 5: Fitted CDF of the DHLD and other compared models for the eruptions data set.

Table 4: Numerical values of the average of both $\hat{a}$, B, MSE, and MRE.

		$a = 1.5$			$a = 2.5$		
n	Est.	MLE	MME	PE	MLE	MME	PE
20	$A_{\hat{a}}$	1.55036357	1.56201278	1.55516013	5.31858946	4.44659993	4.5314259
	A_B	0.20325805	0.23850044	0.20907116	3.06929026	2.19927803	2.28447832
	A_{MSE}	0.13490638	0.24254368	0.0742961	26.95421131	13.07753737	13.44406908
	A_{MRE}	0.13550536	0.15900029	0.13938077	1.2277161	0.87971121	0.91379133
80	$A_{\hat{a}}$	1.51681163	1.50151064	1.51090035	2.64972779	2.58060239	2.60369867
	A_B	0.0967442	0.10245518	0.09478967	0.29935364	0.24032052	0.25454172
	A_{MSE}	0.01467609	0.01667349	0.01457611	1.18309694	0.24213961	0.35219041
	A_{MRE}	0.06449614	0.06830345	0.06319312	0.11974146	0.09612821	0.10181669
150	$A_{\hat{a}}$	1.50633676	1.50500006	1.50380541	2.52484788	2.53116172	2.52251102
	A_B	0.06388323	0.0729854	0.06800988	0.14700026	0.15364345	0.15408628
	A_{MSE}	0.00655839	0.00861836	0.0075398	0.0353905	0.07790259	0.03908058
	A_{MRE}	0.04258882	0.04865693	0.04533992	0.05880011	0.06145738	0.06163451
300	$A_{\hat{a}}$	1.50474815	1.50090184	1.50431723	2.51388132	2.51874551	2.5084716
	A_B	0.04726563	0.05326025	0.04966837	0.10248397	0.10858419	0.10530677
	A_{MSE}	0.00366175	0.00440471	0.00383019	0.01743403	0.0195046	0.01870902
	A_{MRE}	0.03151042	0.03550683	0.03311225	0.04099359	0.04343368	0.04212271
500	$A_{\hat{a}}$	1.5021338	1.50081128	1.50225353	2.50842086	2.51028656	2.50634903
	A_B	0.03605575	0.03921868	0.03673413	0.08105533	0.0802475	0.07591874
	A_{MSE}	0.00203902	0.0024415	0.00214846	0.01044542	0.01031853	0.00925252
	A_{MRE}	0.02403717	0.02614579	0.02448942	0.03242213	0.032099	0.0303675

Table 5: Numerical analysis for the eruptions data set.

Distribution	$-L$	**A**	**C**	**B**	**H**	**KS**	p**-value**	**Estimates**
DHLD	293.552	589.104	589.169	591.263	589.955	0.103991	0.493169	$\hat{a} = 0.0270688\ (0.00333184)$
PD	952.542	1907.08	1907.15	1909.24	1907.94	0.501204	< 0.00001	$\hat{\lambda} = 39.8278\ (0.788862)$
GD	300.61	603.219	603.284	605.378	604.07	0.184405	0.0257451	$\hat{\lambda} = 0.0244929\ (0.00302389)$
DLD	336.734	675.469	675.533	677.627	676.319	0.144288	0.139175	$\hat{\alpha} = 0.0502006\ (0.00443668)$
DGD	296.224	596.448	596.645	600.766	598.149	0.113873	0.377748	$\hat{\alpha} = 1.673\ (0.0799418)$
								$\hat{\beta} = 0.0414854\ (0.00522843)$
DWD	297.327	598.654	598.851	602.972	600.355	0.10499	0.480757	$\hat{\alpha} = 0.992549\ (0.00386331)$
								$\hat{\beta} = 1.29439\ (0.12197)$
DRLD	300.603	603.206	603.27	605.364	604.056	0.184372	0.0257855	$\hat{\lambda} = 39.2831\ (5.04203)$
DRD	300.61	603.219	603.284	605.378	604.07	0.184405	0.0257451	$\hat{\lambda} = 0.975507\ (0.00302389)$
DMGED	298.983	599.967	600.031	602.126	600.817	0.151755	0.104895	$\hat{\alpha} = 0.981248\ (0.00205632)$
DBD	357.359	718.718	718.915	723.036	720.419	0.463399	< 0.00001	$\hat{\alpha} = 0.954382\ (0.247949)$
								$\hat{\beta} = 6.35577\ (35.3637)$
DPD	361.133	724.267	724.331	726.426	725.117	0.474892	< 0.00001	$\hat{\alpha} = 0.746119\ (0.0273148)$
DBHD	437.051	876.102	876.167	878.261	876.953	0.889213	< 0.00001	$\hat{\alpha} = 0.999634\ (0.00242957)$
DLoD	301.569	607.138	607.334	611.455	608.839	0.198291	0.0130403	$\hat{\alpha} = 439.016\ (134.087)$
								$\hat{\beta} = 7.06218 \times 10^{-6}\ (0.0000227432)$

7 Conclusion

This article introduces a novel one-parameter discrete distribution called the discrete half-logistic distribution (DHLD), which is constructed based on the discretization approach of the continuous half-logistic distribution. The hazard function of the DHLD exhibits three possible shapes: increasing, constant, or decreasing. Several well-established mathematical expansions from the literature were employed to assess its statistical properties to calculate the proposed distribution's selected statistical features. Various estimation methods were employed to estimate the parameter of the DHLD, including maximum likelihood, moments, and proportion methods. A comprehensive simulation study was conducted to evaluate the performance of these estimation methods. The simulation study provides valuable insights into the accuracy and efficiency of each method in estimating the DHLD parameter. Furthermore, the flexibility of the DHLD was investigated using a real data set. The real data set served as a practical case study to examine the goodness of fit of the DHLD compared to other competing models. Through careful analysis and comparison, the results indicated that the DHLD outperformed other models in terms of fitting the real data set, thus highlighting its superior suitability in capturing the underlying patterns and characteristics of the observed data.

This paper highlights several promising avenues for future research, which have the potential to significantly contribute to the understanding and application of the proposed model in various fields. These directions encompass expanding the analysis to the bivariate case, exploring alternative censored methods, adopting Bayesian approaches for parameter estimation, considering various loss functions, and investigating the application of neutrosophic statistics. Expanding the analysis to the bivariate case would involve studying the behavior and properties of the proposed model when applied to two variables simultaneously.

References

[1] T. Nakagawa and S. Osaki, "The discrete weibull distribution," *IEEE transactions on reliability*, vol. 24, no. 5, pp. 300–301, 1975.
[2] M. El-Morshedy, M. S. Eliwa, and H. Nagy, "A new two-parameter exponentiated discrete lindley distribution: properties, estimation and applications," *Journal of applied statistics*, vol. 47, no. 2, pp. 354–375, 2020.
[3] D. Kundu and V. Nekoukhou, "Univariate and bivariate geometric discrete generalized exponential distributions," *Journal of Statistical Theory and Practice*, vol. 12, no. 3, pp. 595–614, 2018.
[4] D. Roy, "Discrete rayleigh distribution," *IEEE Transactions on Reliability*, vol. 53, no. 2, pp. 255–260, 2004.
[5] H. Krishna and P. S. Pundir, "Discrete burr and discrete pareto distributions," *Statistical methodology*, vol. 6, no. 2, pp. 177–188, 2009.
[6] V. Nekoukhou and H. Bidram, "The exponentiated discrete weibull distribution," *Sort*, vol. 39, pp. 127–146, 2015.
[7] M. El-Morshedy, E. Altun, and M. S. Eliwa, "A new statistical approach to model the counts of novel coronavirus cases," *Mathematical Sciences*, vol. 16, no. 1, pp. 37–50, 2022.
[8] S. Chakraborty, D. Chakravarty, J. Mazucheli, and W. Bertoli, "A discrete analog of gumbel distribution: properties, parameter estimation and applications," *Journal of Applied Statistics*, vol. 48, no. 4, pp. 712–737, 2021.
[9] A. S. Eldeeb, M. Ahsan-Ul-Haq, and A. Babar, "A discrete analog of inverted topp-leone distribution: properties, estimation and applications," *International Journal of Analysis and Applications*, vol. 19, no. 5, pp. 695–708, 2021.
[10] B. A. Para and T. R. Jan, "Discrete version of log-logistic distribution and its applications in genetics," *International Journal of Modern Mathematical Sciences*, vol. 14, no. 4, pp. 407–422, 2016.
[11] E. M. Almetwally and G. M. Ibrahim, "Discrete alpha power inverse lomax distribution with application of covid-19 data," *Int. J. Appl. Math*, vol. 9, no. 6, pp. 11–22, 2020.
[12] A. Hassan, G. A. Shalbaf, S. Bilal, and A. Rashid, "A new flexible discrete distribution with applications to count data," *Journal of Statistical Theory and Applications*, vol. 19, no. 1, pp. 102–108, 2020.
[13] A. Barbiero and A. Hitaj, "A discrete analogue of the half-logistic distribution," in *2020 International Conference on Decision Aid Sciences and Application (DASA)*, pp. 64–67, IEEE, 2020.
[14] F. C. Opone, E. A. Izekor, I. U. Akata, and F. E. U. Osagiede, "A discrete analogue of the continuous marshall-olkin weibull distribution with application to count data," *Earthline Journal of Mathematical Sciences*, vol. 5, no. 2, pp. 415–428, 2021.
[15] J. Galambos, "The asymptotic theory of extreme order statistics," *R.E. Krieger Pub. Co.*, 1987.
[16] H. S. Bakouch, M. A. Jazi, and S. Nadarajah, "A new discrete distribution," *Statistics*, vol. 48, no. 1, pp. 200–240, 2014.
[17] Z. Yang, "Maximum likelihood phylogenetic estimation from dna sequences with variable rates over sites: approximate methods," *Journal of Molecular evolution*, vol. 39, no. 3, pp. 306–314, 1994.
[18] A. S. Eldeeb, M. Ahsan-ul Haq, and M. S. Eliwa, "A discrete ramos-louzada distribution for asymmetric and over-dispersed data with leptokurtic-shaped: Properties and various estimation techniques with inference," *AIMS Mathematics*, vol. 7, no. 2, pp. 1726–1741, 2022.
[19] M. S. Eliwa and M. El-Morshedy, "A one-parameter discrete distribution for over-dispersed data: Statistical and reliability properties with applications," *Journal of Applied Statistics*, pp. 1–21, 2021.
[20] M. El-Morshedy, M. S. Eliwa, and E. Altun, "Discrete burr-hatke distribution with properties, estimation methods and regression model," *IEEE access*, vol. 8, pp. 74359–74370, 2020.
[21] B. A. Para and T. R. Jan, "On discrete three-parameter burr type xii and discrete lomax distributions and their applications to model count data from medical science," *Biometrics and Biostatistics International Journal*, vol. 4, no. 2, pp. 1–15, 2016.