Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/150316

The Exponentiated Discrete Linear Exponential Distribution: Theory, Risk Profiles, and Applications to Biological and Engineering Data

Hend S. Shahen[1], *Mohamed S. Eliwa*[2,3,*], *and Mahmoud El-Morshedy*[1]

[1]Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[2] Department of Statistics and Operations Research, College of Science, Qassim University, Qassim, Saudi Arabia
[3] Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

Received: 25 Feb. 2026, Revised: 13 Mar. 2026, Accepted: 1 Apr. 2026
Published online: 1 May 2026

Abstract: This paper presents a novel two-parameter exponentiated discrete linear exponential (EDLE) distribution designed to effectively model discrete data exhibiting diverse risk behaviors and dispersion patterns. The primary statistical features of the EDLE distribution, including the probability mass function, hazard rate function, moments, skewness, kurtosis, index of dispersion, entropies, and L-moments of order statistics, are calculated and analyzed. Due to its adaptability, the EDLE distribution may accommodate danger rates that increase, decrease, or exhibit a bathtub configuration. This renders it highly beneficial for reliability and survival analysis in discrete settings. The distribution provides a robust framework for modeling complex discrete events, accurately representing positively skewed data with varying kurtosis, overdispersion, and heavy-tailed characteristics. The maximum likelihood method is used to estimate parameters, ensuring the reliability of the inference. A Monte Carlo simulation analysis was performed in R to assess the efficacy of the estimators. We examine various sample sizes and coverage probability predicated on 95% confidence intervals. The practical usefulness of the EDLE distribution is demonstrated using three empirical datasets: (i) epidemiological data, (ii) biological-entomological data from agricultural research, and (iii) turbocharger failure data from reliability tests. The EDLE distribution consistently demonstrates superior fitting across all scenarios and is more adept at managing complex discrete data structures compared to numerous alternative discrete models. The results indicate that the model is adaptable, robust, and applicable in various disciplines.

Keywords: Statistical model, Failure analysis, Over-dispersion, Maximum likelihood, Monte Carlo simulation, Discrete data modeling

1 Introduction

Numerous subfields of statistics depend on discrete probability distributions as a fundamental instrument for characterizing and examining countable events. Integers serve as the standard unit of measurement for real-world processes, such as the count of diagnosed patients, machine failures in a facility, or the presence of rare species in an environment, so conferring upon them intrinsic use. Due to the exponential growth of count-based data in the digital era, there is a necessity for advanced models that effectively capture complex data characteristics, such as multimodal patterns, zero-inflation, and over-dispersion. The array of tools for discrete modeling has significantly and rapidly increased. Discrete distributions, including the discrete gamma and discrete Rayleigh distributions, were among the earliest to provide the flexible and reliable analysis of count data. Recent advancements, such as the discrete modified Weibull, exponentiated discrete Weibull, and discrete Bilal distributions, have significantly improved the capacity to describe intricate occurrences, including integer-valued time series and specific failure patterns. Recent developments in discrete trigonometric families have introduced novel methodologies for the analysis of cyclical or seasonal count data. In summary, the findings of these thorough academic studies indicate that this is a swiftly evolving domain with a significant emphasis on enhancing the accuracy and versatility of models. The primary domains in which statistical

* Corresponding author e-mail: mseliwa@mans.edu.eg

findings are employed to tackle contemporary global challenges most effectively illustrate their practical significance. Complex discrete models are essential in public health and epidemiology to appropriately represent hospital admissions, new illnesses, or adverse drug responses.

Zero-inflated and exponentiated variations are better equipped to replicate infrequent genetic mutation counts or occasional disease outbreaks to inform intervention tactics, vaccination distribution, and healthcare resource allocation. Discrete and exponentiated discrete models are employed in engineering and industrial systems to anticipate reliability, manage quality, and conduct predictive maintenance. Enhancing product designs, refining safety protocols, and prolonging the operational lifespan of essential infrastructure can be achieved by precisely modeling the frequency of component failures, manufacturing defects, or system downtime incidents, especially through the application of customized distributions like the exponentiated discrete Weibull. Energy security and environmental sustainability are two domains in which these distributions are becoming significant. They are utilized to model and forecast particular occurrences, including power system failures, fluctuations in renewable energy output (such as durations of diminished wind or solar generation), and emissions-related disasters. They facilitate the enumeration of species, monitor unusual animal events, and estimate the prevalence of environmental anomalies, all of which contribute to the assessment of biodiversity in ecological and conservation science. This analytical proficiency is essential for the advancement of robust energy systems and the creation of evidence-based sustainability policies. Ultimately, discrete distributions are essential for quantifying packet loss, network intrusion attempts, user transaction volumes, and system fault counts within the domains of telecommunications, network security, and banking. These applications are essential for preserving the interconnected systems that support the global digital economy, enhancing network performance, and safeguarding data integrity. In response to these analytical obstacles, discrete distribution theory evolved. The exponential technique is a significant advancement in methodology. It generalizes existing distributions by raising a baseline cumulative distribution function (CDF), $G(x;.)$, to a positive power α, resulting in a new family of distributions $F(x;\alpha,.) = (G(x;.))^{\alpha}$, where $\alpha > 0$. This robust method, which includes a changeable shape parameter, can be employed to more effectively elucidate positively skewed data, such as infrequent high-count occurrences.

For under-dispersed or bounded count processes, employ $0 < \alpha < 1$ to elongate the distribution towards lower values. The exponentiation method can be employed to generalize numerous classical probability distributions. The continuous domain encompasses several significant exponentiated families. Gupta and Kundu [1] and Pal et al. [2] exemplify methodologies for analyzing dependability and lifespan: the generalized exponential and the exponentiated Weibull. Nadarajah and Gupta [3] propose that exponential gamma is a potential model for environmental data. Salem [4] and Shawky and Abu-Zinadah [5] identify two families: the exponentiated Pareto and the exponentiated Lomax. This strategy has been employed to develop several flexible models in the discrete domain. Examples of such functions are the exponentiated geometric for over-dispersed counts, the exponentiated discrete Weibull, the exponentiated discrete inverse Rayleigh, and the exponentiated discrete Lindley. The models reference sources such as Almalki and Nadarajah [6], Chakraborty and Gupta [7], Cardial et al. [8], MirMostafaee and Mashhadzadeh [9], and El-Morshedy et al. [10]. Unique analogues of exponential continuous families have arisen, enhancing the resources for modeling count data with intricate characteristics. El-Morshedy [11] introduced the discontinuous linear exponential (DLE) distribution, which is presented here alongside its cumulative distribution function (CDF) and probability mass function (PMF):

$$G(x;\beta) = 1 - \left(1 - \frac{(x+1)\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x+1}; \quad x \in \mathbb{N}_0, \quad 0 < \beta < 1, \tag{1}$$

and

$$P(x;\beta) = \left(1 - \frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x} - \left(1 - \frac{(x+1)\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x+1}. \tag{2}$$

This work aims to create a discrete distribution that is robust and adaptable enough to accurately depict the complex data patterns frequently encountered in practice. Contemporary discrete models frequently inadequately describe several real-world datasets exhibiting diverse risk behaviors, overdispersion, heavy tails, and skewness, particularly in the biomedical, agricultural, and engineering fields. The suggested EDLE distribution offers a versatile framework capable of accommodating increasing, decreasing, and bathtub-shaped hazard rates, while effectively capturing varying levels of kurtosis and skewness to solve these issues. This study offers a comprehensive instrument for theoretical exploration and empirical data analysis through the derivation of fundamental statistical characteristics. Moreover, reliable parameter inference is ensured through the application of maximum likelihood estimation. The primary objective of the EDLE distribution is to enhance the precision and interpretability of models for complex discrete data, serving as a reliable instrument for dependability analysis, risk evaluation, and decision-making across several scientific fields.

This document delineates the components of the study. Section 2 presents the EDLE model. This model is established, and we observe visualization charts illustrating its capacity to represent various sorts of discrete data patterns. Section 3 offers a comprehensive analysis of the statistical and reliability characteristics of the EDLE distribution, emphasizing its ability to model increasing, decreasing, and bathtub-shaped hazard behaviors. This includes

an exploration of its probability mass function, hazard rate function, moments, skewness, kurtosis, index of dispersion, entropies, and L-moments of order statistics. To ensure precise and reliable inference, Section 4 examines the conventional estimation of unknown model parameters in complete data contexts utilizing the maximum likelihood method. Section 5 conducts an extensive Monte Carlo simulation to evaluate the efficacy of the proposed estimators and the model's fit across various datasets. Section 6 illustrates the practical relevance and adaptability of the EDLE distribution through comprehensive reviews of four real-world datasets sourced from the biomedical, agricultural, and engineering sectors. Section 7 concludes the work by summarizing the principal findings, emphasizing the model's achievements, and discussing potential directions for future research in discrete data modeling.

2 Methodological Framework and Visualization of the EDLE Model

Based on equations (1) and (2), the random variable X is defined to follow the EDLE distribution if its CDF can be expressed in the form

$$F(x;\alpha,\beta)=\left(1-\left(1-\frac{(x+1)\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x+1}\right)^{\alpha};\quad x\in\mathbb{N}_0, \tag{3}$$

where $\alpha>0$, $0<\beta<1$ and $\mathbb{N}_0=\{0,1,2,3,\ldots\}$. The PMF corresponding to equation (3) is given by

$$f(x;\alpha,\beta)=\left(1-\left(1-\frac{(x+1)\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x+1}\right)^{\alpha}-\left(1-\left(1-\frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x}\right)^{\alpha};\quad x\in\mathbb{N}_0. \tag{4}$$

Figure 1 illustrates that the PMF of the EDLE distribution exhibits considerable flexibility when the parameters α and β are varied in multiple combinations. The graphs exhibit a diverse range of morphologies, from distributions that are significantly right-skewed to those that are nearly symmetrical and evenly distributed. The model's adaptability facilitates applications in dependability research, survival studies, and count data modeling, demonstrating its efficacy in representing a broad range of real-world discrete data patterns when conventional distributions may be inadequate.

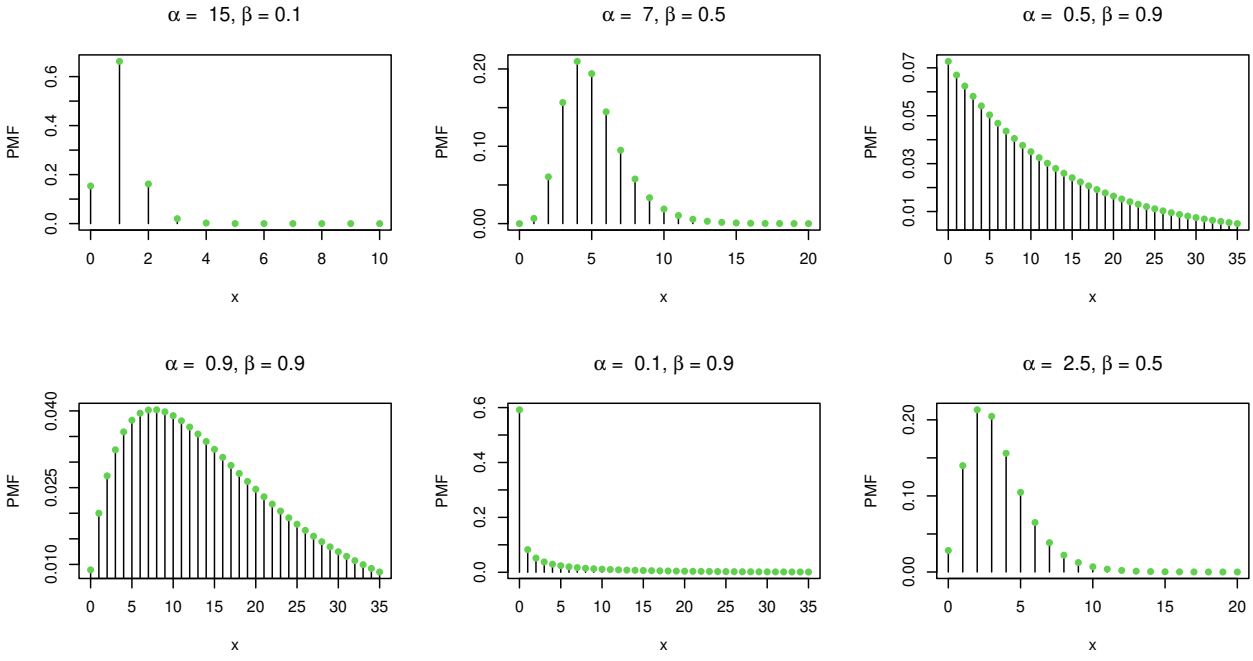

Fig. 1: Various shapes of the PMF of the EDLE distribution.

The hazard rate function (HRF) and the reversed hazard rate function (RHRF) can be derived as follows

$$h(x;\alpha,\beta)=\frac{\left(1-\left(1-\frac{(x+1)\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x+1}\right)^{\alpha}-\left(1-\left(1-\frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x}\right)^{\alpha}}{1-\left(1-\left(1-\frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x}\right)^{\alpha}};\quad x\in\mathbb{N}_0, \tag{5}$$

and

$$r(x;\alpha,\beta)=1-\frac{\left(1-\left(1-\frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x}\right)^{\alpha}}{\left(1-\left(1-\frac{(x+1)\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x+1}\right)^{\alpha}};\quad x\in\mathbb{N}_0, \tag{6}$$

respectively, where $h(x;\alpha,\beta)=f(x;\alpha,\beta)/(1-F(x-1;\alpha,\beta))$ and $r(x;\alpha,\beta)=f(x;\alpha,\beta)/F(x;\alpha,\beta)$. Figure 2 shows how the HRF can be used with various combinations of α and β parameters. A wide variety of lifetime patterns are shown by the depicted functions, including strictly growing, strictly decreasing, and bathtub-shaped (decreasing then increasing) hazard rates. Reliability analysis, survival studies, and any application using time-to-event data with non-monotonic hazard characteristics can benefit from this model's robustness and flexibility, since it can properly describe varied real-world failure patterns.

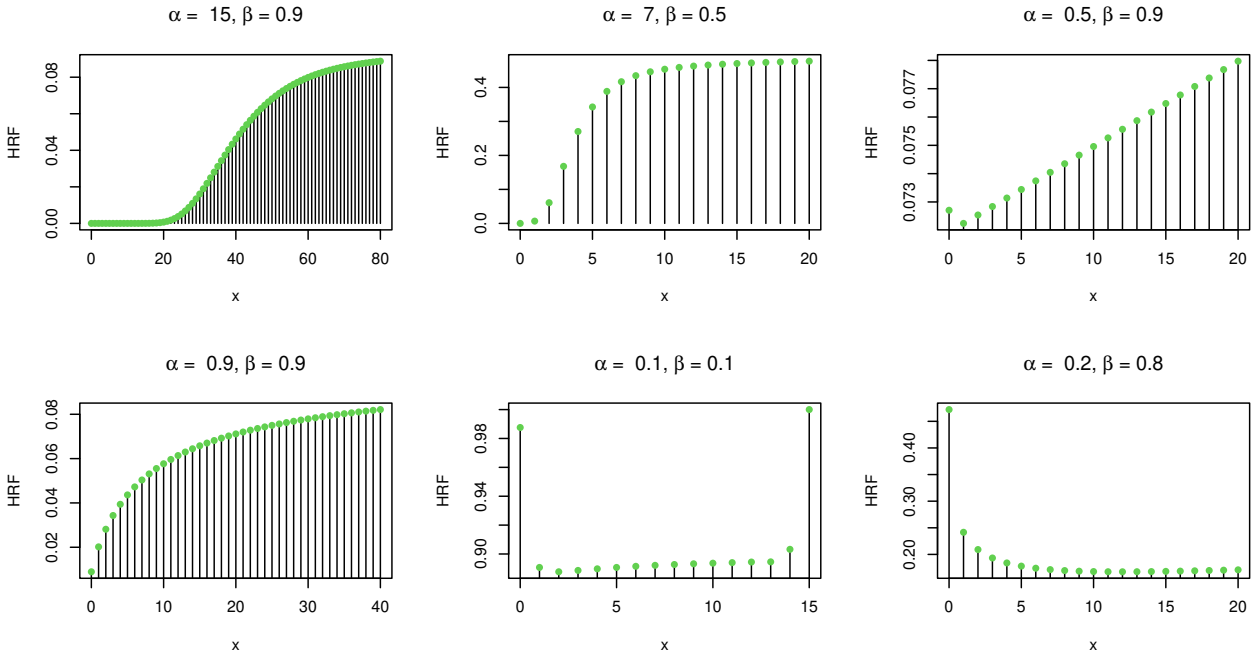

Fig. 2: Various shapes of the HRF of the EDLE distribution.

3 Theoretical Foundations of the EDLE Distribution

3.1 Moments and Related Statistical Concepts

This section computes the raw moments (RMs) and essential statistical metrics for the EDLE distribution. The RMs are crucial for understanding the distribution's properties and calculating necessary descriptive statistics. These variables include mean (E(X)), variance (Var(X)), skewness (Sk(X)), kurtosis (Ku(X)), and index of dispersion (IoD(X)). These measurements summarize the central tendency, variability, shape, and tail behavior of data, enabling direct comparison across datasets. Raw moments are key tools for parameter estimation and hypothesis testing. If X follows an EDLE

distribution with parameters α and β, its r-th raw moment is given by

$$\mu'_r = E(X^r) = \sum_{x=1}^{\infty} x^r f(x;\alpha,\beta); \quad r = 1,2,3,\ldots$$
$$= \sum_{x=1}^{\infty} [(x-1)^r - x^r] \left(1 - \left(1 - \frac{x \ln \beta}{1-(\ln \beta)^3}\right)\beta^x\right)^{\alpha}. \tag{7}$$

The moment generating function (MGF) can be reported as

$$M_X(t) = E(e^{Xt}) = \sum_{x=0}^{\infty}\sum_{j=0}^{\infty} \frac{(xt)^j}{j!} \left[\left(1 - \left(1 - \frac{(x+1)\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x+1}\right)^{\alpha} - \left(1 - \left(1 - \frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^x\right)^{\alpha}\right]; \quad x \in \mathbb{N}_0. \tag{8}$$

Using the MGF provided in equation (8), the RMs can be derived from MGF where $E(X^r) = \frac{d^r}{dt^r} M_X(t)|_{t=0}$. The E(X), Var(X), Sk(X), and Ku(X) can be respectively given by

$$E(X) = -\sum_{x=1}^{\infty} \left(1 - \left(1 - \frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^x\right)^{\alpha}, \tag{9}$$

$$Var(X) = \sum_{x=1}^{\infty}(1-2x)\left(1 - \left(1 - \frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^x\right)^{\alpha} - \mu_1'^2, \tag{10}$$

$$Sk(X) = \frac{\mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3_1}{(Var(X))^{3/2}}, \tag{11}$$

$$Ku(X) = \frac{\mu'_4 - 4\mu'_1\mu'_3 + 6\mu'_2\mu'^2_1 - 3\mu'^4_1}{(Var(X))^2}. \tag{12}$$

An important additional descriptive measure, known as the IoD(X) also referred to as the relative variance or variance-to-mean ratio serves as a valuable tool for characterizing actuarial and count data. For the EDLE distribution, the IoD(X) can be formulated as follows

$$IoD(X) = \frac{-\sum_{x=1}^{\infty}(1-2x)\left(1 - \left(1 - \frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^x\right)^{\alpha}}{\sum_{x=1}^{\infty}\left(1 - \left(1 - \frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^x\right)^{\alpha}} - \sum_{x=1}^{\infty}\left(1 - \left(1 - \frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^x\right)^{\alpha}. \tag{13}$$

The EDLE distribution's descriptive statistics for various shape α and scale β parameters are shown in Table 1.

Table 1: Moments and dispersion indices of the EDLE distribution across selected parameter values.

Measure	$\alpha \downarrow \beta \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
E(X)	0.3	0.04140	0.11320	0.23930	0.45270	0.79160	1.31250	2.15370	3.77390	8.53770
	0.9	0.12000	0.31620	0.63910	1.14720	1.89790	2.98610	4.68870	7.93450	17.4552
	2	0.25130	0.62140	1.17490	1.97490	3.08410	4.64630	7.08710	11.7745	25.5898
Var(X)	0.3	0.05000	0.16810	0.43910	1.02040	2.19000	4.55380	9.94370	26.1617	118.921
	0.9	0.13650	0.41670	0.97920	2.03050	3.90900	7.46170	15.4232	39.3840	176.431
	2	0.25630	0.67040	1.37360	2.55090	4.57290	8.44460	17.2728	44.0143	197.141
Sk(X)	0.3	6.26220	4.62750	3.75750	3.14830	2.72880	2.46880	2.32590	2.25450	2.22070
	0.9	3.43980	2.50450	2.03520	1.73020	1.55240	1.47560	1.45550	1.45510	1.45760
	2	2.10150	1.53380	1.32550	1.22880	1.19550	1.19340	1.19910	1.20400	1.20660
Ku(X)	0.3	50.8132	30.7314	21.5895	16.1171	12.9304	11.2320	10.4034	10.0235	9.85280
	0.9	17.0116	11.0620	8.53290	7.12400	6.44160	6.19760	6.14620	6.14920	6.15830
	2	7.95010	6.18570	5.66400	5.40920	5.31890	5.31120	5.32490	5.33730	5.34410
IoD(X)	0.3	1.20810	1.48500	1.83450	2.25410	2.76650	3.46960	4.61700	6.93220	13.9289
	0.9	1.13700	1.31790	1.53220	1.76990	2.05970	2.49880	3.28940	4.96360	10.1077
	2	1.01990	1.07900	1.16910	1.29170	1.48270	1.81750	2.43720	3.73810	7.70390

As β escalates from 0.1 to 0.9, both the mean and variance increase for a constant α. A positive correlation between the parameters and the central tendency and dispersion of the distribution is signified by increased means and variances for elevated α values. The right-skewed shape is validated by the consistently positive skewness across all configurations; nevertheless, as β or α grows, the skewness decreases, signifying a gradual transition toward symmetry. The distribution more closely approaches a mesokurtic distribution when increases in β or α reduce kurtosis, however it continually exceeds 3, indicating leptokurtic behavior characterized by heavier tails and sharper peaks. Finally, the over-dispersed characteristic of the EDLE distribution is validated by the IoD(X), which significantly increases with higher β, especially at lower α values, illustrating the joint effect of the two parameters on variability.

3.2 Various Entropy Measures and Their Computations

This part looks into several entropy (Ey) measures for the EDLE distribution. Entropy is a basic idea in information theory. It is a way to quantify how uncertain, random, or informative a probability distribution is. Different ways of calculating entropy give us different information about how variable and structurally complicated the distribution is. In particular, we calculate and examine the following important entropies for the EDBuE model: Renyi entropy (REy), Shannon entropy (SnEy), Collision entropy (CoEy), Min-entropy (MEy), and Max-entropy (XEy). These entropy measures are widely employed in several fields, such as information theory, statistical physics, machine learning, encryption, and complex systems analysis, where measuring uncertainty and information is crucial. Let the random variable X conform to the EDBuE distribution. The REy of order δ (where $\delta > 0$ and $\delta \neq 1$) is defined and derived as follows

$$I_\delta(X) = \frac{1}{1-\delta}\log\sum_{x=0}^{\infty}\left[\left(1-\left(1-\frac{(x+1)\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x+1}\right)^{\alpha}-\left(1-\left(1-\frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x}\right)^{\alpha}\right]^{\delta}; \quad x\in\mathbb{N}_0. \tag{14}$$

The REy provides a unifying framework that generalizes several fundamental entropy measures. Specifically, the SnEy, CoEy, MEy, and XEy emerge as special limiting cases of the REy based on its order parameter δ. These relationships are formally expressed as follows: the SnEy corresponds to the limit as $\delta \to 1$, the CoEy to $\delta \to 2$, the MEy to $\delta \to \infty$, and the XEy to $\delta \to 0$. Consequently, the REy offers a versatile and continuous spectrum for quantifying uncertainty and information content, enabling a cohesive analysis of distributional randomness across different mathematical and applied contexts.

$$\begin{aligned} I(X) = -\sum_{x=0}^{\infty}&\left\{\left(1-\left(1-\frac{(x+1)\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x+1}\right)^{\alpha}-\left(1-\left(1-\frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x}\right)^{\alpha}\right\} \\ &\times\log\left(\left(1-\left(1-\frac{(x+1)\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x+1}\right)^{\alpha}-\left(1-\left(1-\frac{x\ln\beta}{1-(\ln\beta)^3}\right)\beta^{x}\right)^{\alpha}\right). \end{aligned} \tag{15}$$

Computational entropy estimates for the EDLE distribution under different combinations of the shape parameter α and the scale parameter β are shown in Tables 2 and 3, respectively. Table 2 shows the entropy for the Rényi order $\delta = 0.5$, whereas Table 3 shows the CoEy for $\delta = 2$. The entropy levels and the parameters α and β are positively correlated, according to the results. For a constant α, the entropy grows monotonically with β, which indicates that the distribution is becoming more scattered, which implies that there is more uncertainty or information content. The shape parameter also strongly adds to the distribution's variability when, for a given β, more entropy is correlated with higher α values.

Table 2: Numerical REy of the EDLE distribution at $\delta = 0.5$.

$\alpha\downarrow\ \beta\rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.3	0.4505	0.7476	1.0566	1.3819	1.7169	2.0640	2.4469	2.9249	3.6755
0.9	0.6849	1.0700	1.4349	1.7884	2.1289	2.4694	2.8437	3.3168	4.0669
2	0.8835	1.3001	1.6627	1.9923	2.3008	2.6145	2.9727	3.4389	4.1872

Table 3: Numerical CoEy of the EDLE distribution at $\delta = 2$.

$\alpha \downarrow \beta \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.3	0.0738	0.1768	0.3225	0.5225	0.7782	1.0830	1.4377	1.8791	2.5646
0.9	0.2137	0.4890	0.8376	1.2400	1.6419	2.0123	2.3891	2.8575	3.6049
2	0.4384	0.8800	1.2666	1.5871	1.8813	2.1880	2.5445	3.0110	3.7601

Notably, entropy values for $\delta = 0.5$ are systematically higher than those for $\delta = 2$ across all parameter combinations, consistent with the theoretical property that Rényi entropy is non-increasing in δ. This pattern underscores the sensitivity of entropy measures to the order parameter, highlighting how different entropy definitions capture distinct aspects of the distribution's randomness. These numerical findings illustrate the flexibility of the EDLE model in representing a spectrum of uncertainty structures, supporting its applicability in fields requiring detailed quantification of informational complexity.

3.3 L-moments of Order Statistics

The smallest value in a random sample that is n-th in order is known as the n-th order statistic in statistical inference. Nonparametric statistics, robust estimation, and distribution theory all have order statistics and associated rank-based measures as their basis. In order to construct tolerance limits, evaluate dependability, and analyze extreme values, they are essential. Think about a common distribution and a random sample $X_1, X_2, \ldots, X_n$. The order statistics are independent variables that are represented by the set $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$. If the EDLE distribution is true for X, then the CDF of the i-th order statistic, $X_{i:n}$, may be written as

$$\begin{aligned} F_{i:n}(x;\alpha,\beta) &= \sum_{k=i}^{n} \binom{n}{k} [F_i(x;\alpha,\beta)]^k [1 - F_i(x;\alpha,\beta)]^{n-k} \\ &= \sum_{k=i}^{n} \sum_{j=0}^{n-k} \sum_{l=0}^{\infty} \sum_{m=0}^{l} \Theta_{(n,k)}^{(m,j,l)} \left(\frac{(x+1)\ln\beta}{1-(\ln\beta)^3} \right)^m \beta^{l(x+1)}, \end{aligned} \tag{16}$$

where $\Theta_{(n,k)}^{(j,l)} = (-1)^{l+j+m} \binom{n}{k} \binom{n-k}{j} \binom{l}{m} \binom{\alpha(k+j)}{l}$. The corresponding PMF of the i-th order statistic is given by

$$\begin{aligned} f_{i:n}(x;\alpha,\beta) &= F_{i:n}(x;\alpha,\beta) - F_{i:n}(x-1;\alpha,\beta) \\ &= \sum_{k=i}^{n} \sum_{j=0}^{n-k} \sum_{l=0}^{\infty} \sum_{m=0}^{l} \Theta_{(n,k)}^{(m,j,l)} \left[\left(\frac{(x+1)\ln\beta}{1-(\ln\beta)^3} \right)^m \beta^{l(x+1)} - \left(\frac{x\ln\beta}{1-(\ln\beta)^3} \right)^m \beta^{lx} \right]. \end{aligned} \tag{17}$$

Thus, the i-th RMs of $X_{i:n}$ can be reported as

$$E(X_{i:n}^r) = \sum_{x=0}^{\infty} \sum_{k=i}^{n} \sum_{j=0}^{n-k} \sum_{l=0}^{\infty} \sum_{m=0}^{l} \Theta_{(n,k)}^{(m,j,l)} x^r \left[\left(\frac{(x+1)\ln\beta}{1-(\ln\beta)^3} \right)^m \beta^{l(x+1)} - \left(\frac{x\ln\beta}{1-(\ln\beta)^3} \right)^m \beta^{lx} \right]. \tag{18}$$

An effective descriptive statistical measure, the L-moments are obtained from equation (18). When compared to traditional moments, L-moments offer an alternate method of distribution characterization that is more robust against outliers and frequently more stable for small samples. For parameter estimation, summarizing probability distributions, and goodness-of-fit testing, they are invaluable. A random variable X that follows the EDLE distribution is characterized by the r-th L-moment as

$$\Upsilon_{\varpi} = \frac{1}{\varpi} \sum_{i=0}^{\varpi-1} (-1)^i \binom{\varpi-1}{i} E(X_{\varpi-i:\varpi}). \tag{19}$$

According to equation (19), some statistical measures can be derived like mean, Sk, and Ku where mean = Υ_1, Sk = Υ_3/Υ_2, and Ku = Υ_4/Υ_2.

4 Maximum Likelihood Estimation

Using the maximum likelihood (ML) approach, parameter estimates for the EDLE distribution are derived in this section. Due to its excellent asymptotic features, such as consistency, efficiency, and asymptotic normalcy, ML estimation a key technique in classical statistical inference is highly regarded and used for large-scale, reliable inference. As a result, you can get parameter values that increase the likelihood of seeing the provided data in a principled way. A full and independent random sample of size n selected from the EDLE distribution with parameters α and β is represented by $x_1, x_2, \ldots, x_n$. One goal is to maximize the likelihood function, or its logarithm, in order to estimate the parameter vector $\theta = (\alpha, \beta)^T$. It is possible to express the log-likelihood function $\ell(\theta)$ as

$$\ell(\theta) = \sum_{i=1}^{n} \ln \left[\left(1 - \left(1 - \frac{(x_i+1)\ln\beta}{1-(\ln\beta)^3} \right) \beta^{x_i+1} \right)^{\alpha} - \left(1 - \left(1 - \frac{x_i \ln\beta}{1-(\ln\beta)^3} \right) \beta^{x_i} \right)^{\alpha} \right]. \tag{20}$$

Differentiating the log-likelihood function in equation (20) with respect to the parameters α and β yields the following system of score equations

$$\frac{\partial \ell(\theta)}{\partial \alpha} = \sum_{i=1}^{n} \frac{(Z_1(x_i+1))^{\alpha} \ln(Z_1(x_i+1)) - (Z_1(x_i))^{\alpha} \ln(Z_1(x_i))}{\left(1 - \left(1 - \frac{(x_i+1)\ln\beta}{1-(\ln\beta)^3} \right) \beta^{x_i+1} \right)^{\alpha} - \left(1 - \left(1 - \frac{x_i \ln\beta}{1-(\ln\beta)^3} \right) \beta^{x_i} \right)^{\alpha}}, \tag{21}$$

$$\frac{\partial \ell(\theta)}{\partial \beta} = \alpha \sum_{i=1}^{n} \frac{-(x_i+1)\beta^{x_i}(Z_1(x_i+1))^{\alpha-1} Z_2(x_i+1) + x_i \beta^{x_i-1}(Z_1(x_i))^{\alpha-1} Z_2(x_i)}{\left(1 - \left(1 - \frac{(x_i+1)\ln\beta}{1-(\ln\beta)^3} \right) \beta^{x_i+1} \right)^{\alpha} - \left(1 - \left(1 - \frac{x_i \ln\beta}{1-(\ln\beta)^3} \right) \beta^{x_i} \right)^{\alpha}}, \tag{22}$$

where $Z_1(x_i) = 1 - \left(1 - \frac{x_i \ln\beta}{1-(\ln\beta)^3} \right) \beta^{x_i}$ and $Z_2(x_i) = 1 - \frac{x_i \ln\beta}{1-(\ln\beta)^3} - \frac{1+2(\ln\beta)^3}{(1-(\ln\beta)^3)^2}$.

Equations (21) and (22) that produce the highest likelihood score are obtained by setting their partial derivatives to zero. Commonly known as the normal equations of the MLE problem, these equations constitute a set of nonlinear equations. The nonlinearity of the equations prevents analytical solutions in closed form. In order to get the parameter estimates $\widehat{\alpha}$ and $\widehat{\beta}$, numerical optimization approaches are necessary. To maximize these log-likelihoods, classical iterative methods like Newton-Raphson, BFGS quasi-Newton, or Fisher's scoring are typically employed. Convergence conditions, such as little changes in parameter estimates or log-likelihood values, are met iteratively by these approaches by updating the parameter estimates. The `optim()` function, which uses a general-purpose optimization routine, was used to perform the estimation in this investigation. Under the restrictions $\alpha > 0$ and $0 < \beta < 1$, the parameters were estimated by maximizing the value of the log-likelihood function with respect to α and β, which was programmed in R. To efficiently manage these boundaries, the optimization utilized the **L-BFGS-B** approach.

5 Monte Carlo Simulation Framework: Algorithmic Design, R Implementation, and Discussion

Here we assess the finite-sample performance of the EDLE distribution parameter estimators under different sample sizes using a thorough Monte Carlo simulation study. Independent Monte Carlo replications with a total of R are used to conduct the evaluation. We take into account the following performance metrics: Bias, mean squared error (MSE), variance, efficiency, consistency, and the probability of confidence interval coverage (PCIC). Let $\hat{\theta}_r$ be the estimate derived from the rth replication and $\theta \in \{\alpha, \beta\}$ be a generic model parameter. One definition of the estimator's empirical bias is

$$\text{Bias}(\hat{\theta}) = \frac{1}{R} \sum_{r=1}^{R} (\hat{\theta}_r - \theta). \tag{23}$$

The MSE, which reflects both bias and variability, is computed as

$$\text{MSE}(\hat{\theta}) = \frac{1}{R} \sum_{r=1}^{R} (\hat{\theta}_r - \theta)^2. \tag{24}$$

The empirical variance of the estimator is calculated by

$$\text{Var}(\hat{\theta}) = \frac{1}{R-1} \sum_{r=1}^{R} (\hat{\theta}_r - \bar{\hat{\theta}})^2, \tag{25}$$

where

$$\bar{\hat{\theta}} = \frac{1}{R}\sum_{r=1}^{R}\hat{\theta}_r \tag{26}$$

denotes the Monte Carlo mean of the estimates. The efficiency of estimation is determined by taking the reciprocal of the mean squared error (MSE), which is provided by the equation

$$\text{Efficiency} = \frac{1}{\text{MSE}(\hat{\theta})}. \tag{27}$$

Larger MSE values indicate less efficient estimation. Investigating the asymptotic behavior of the estimators as the sample size increases allows for an empirical assessment of their consistency. In particular, $\hat{\theta}$ is deemed consistent as an estimator if

$$\text{Bias}(\hat{\theta}) \to 0 \quad \text{as } n \to \infty, \tag{28}$$

$$\text{MSE}(\hat{\theta}) \to 0 \quad \text{as } n \to \infty. \tag{29}$$

Also, 95% confidence intervals are made for each parameter, and the coverage probabilities that go with them are checked across Monte Carlo replications. Tables 4-7 present the numerical results obtained from the Monte Carlo simulation study for each of the three parameter schemes. The favorable theoretical attributes of the suggested estimators are robustly corroborated by the numerical results derived from the Monte Carlo simulation study. In summary, these are the principal conclusions. As the sample size n increases, a distinct downward trend in the empirical bias of the maximum likelihood estimators for both α and β is evident, irrespective of the parameter configurations or sample sizes. As the bias values converge to zero for moderate to large sample sizes, this trend indicates that the estimators are asymptotically unbiased. As n increases, there is a consistent reduction in both the mean squared error (MSE) and the empirical variance, further validating the consistency of the estimators. These measures consistently decline to zero, indicating that the estimators perform reliably upon repeated sampling. A larger sample significantly improves the accuracy of estimation. The efficiency metric, defined as the inverse of the mean squared error, increases monotonically with the growth of n, indicating that the estimators enhance in precision and approach optimal performance in large samples. This operation aligns with the asymptotic efficiency characteristics of maximum likelihood estimation. Furthermore, based on traditional asymptotic theory for regular parametric estimators, the empirical variance of the estimators diminishes at a rate approximately proportional to $O(n^{-1})$. The proposed estimators appear to achieve near-minimal variance in large samples, based on this finding. At last, the distribution's tail behavior affects the estimators' convergence speed. A slower convergence rate and increased finite-sample variability are the outcomes of heavier tails, which are caused by bigger values of the parameter β (Case III). However, as the sample size grows larger, the estimators keep their asymptotic features, even in heavy-tailed cases.

Table 4: Simulation results for $\alpha = 0.5$, $\beta = 0.2$.

n	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Var($\hat{\alpha}$)	Eff($\hat{\alpha}$)	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	Var($\hat{\beta}$)	Eff($\hat{\beta}$)
25	0.081	0.062	0.055	16	0.043	0.018	0.016	56
50	0.052	0.038	0.034	26	0.029	0.011	0.010	91
75	0.036	0.025	0.023	40	0.021	0.007	0.006	143
100	0.028	0.018	0.017	56	0.016	0.005	0.004	200
150	0.020	0.012	0.011	83	0.012	0.003	0.003	333
200	0.016	0.009	0.008	111	0.009	0.002	0.002	500
300	0.011	0.006	0.006	167	0.006	0.001	0.001	1000
400	0.008	0.004	0.004	250	0.004	0.001	0.001	1250
500	0.006	0.003	0.003	333	0.003	0.001	0.001	1667
600	0.005	0.002	0.002	500	0.002	0.001	0.001	2500

6 Data Analysis and Model Validation

In this section, we use three real-world datasets to empirically evaluate the EDLE distribution. Using a battery of statistical criteria, including the negative log-likelihood ($-L$), Akaike information criterion (AIC), corrected AIC (CAIC), Hannan-Quinn information criterion (HQIC), and the Chi-square goodness-of-fit test (reported with its degrees of freedom and

Table 5: Simulation results for $\alpha = 1.5$, $\beta = 0.5$.

n	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Var($\hat{\alpha}$)	Eff($\hat{\alpha}$)	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	Var($\hat{\beta}$)	Eff($\hat{\beta}$)
25	0.094	0.081	0.073	12	0.058	0.029	0.026	34
50	0.061	0.049	0.045	20	0.039	0.017	0.016	59
75	0.043	0.032	0.030	31	0.028	0.011	0.010	91
100	0.033	0.023	0.022	43	0.021	0.008	0.007	125
150	0.024	0.016	0.015	62	0.015	0.005	0.005	200
200	0.019	0.012	0.011	83	0.011	0.004	0.004	250
300	0.013	0.008	0.008	125	0.008	0.002	0.002	500
400	0.010	0.006	0.006	167	0.006	0.002	0.002	667
500	0.008	0.005	0.005	200	0.005	0.001	0.001	833
600	0.006	0.004	0.004	250	0.004	0.001	0.001	1000

Table 6: Simulation results for $\alpha = 2.0$, $\beta = 0.9$.

n	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Var($\hat{\alpha}$)	Eff($\hat{\alpha}$)	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	Var($\hat{\beta}$)	Eff($\hat{\beta}$)
25	0.121	0.109	0.098	9	0.071	0.041	0.037	24
50	0.082	0.066	0.060	15	0.049	0.025	0.023	40
75	0.058	0.044	0.041	23	0.035	0.017	0.016	59
100	0.045	0.033	0.031	30	0.027	0.013	0.012	77
150	0.033	0.024	0.023	42	0.020	0.009	0.009	111
200	0.026	0.019	0.018	53	0.016	0.007	0.007	143
300	0.018	0.013	0.013	77	0.011	0.005	0.005	200
400	0.014	0.010	0.010	100	0.009	0.004	0.004	250
500	0.011	0.008	0.008	125	0.007	0.003	0.003	333
600	0.009	0.007	0.007	143	0.006	0.003	0.003	400

Table 7: Approx. 95% confidence interval coverage probability (CP) for α and β.

n	CP(α) I	CP(β) I	CP(α) II	CP(β) II	CP(α) III	CP(β) III
25	0.87	0.89	0.85	0.87	0.83	0.85
50	0.91	0.92	0.89	0.91	0.87	0.88
75	0.93	0.94	0.91	0.93	0.90	0.91
100	0.94	0.95	0.93	0.94	0.92	0.93
150	0.95	0.95	0.94	0.95	0.94	0.94
200	0.95	0.95	0.95	0.95	0.95	0.95
300	0.95	0.95	0.95	0.95	0.95	0.95
400	0.95	0.95	0.95	0.95	0.95	0.95
500	0.95	0.95	0.95	0.95	0.95	0.95
600	0.95	0.95	0.95	0.95	0.95	0.95

P-value), the EDLE model's performance is systematically compared with that of several established discrete distributions. This comprehensive evaluation measures the model's relative fit, discourages over-parameterization, and formally checks for sufficiency. This study shows that the EDLE distribution is useful in practice and adds to the toolbox of models for applied count-data analysis by identifying specific data scenarios where it provides a statistically better description of the data, such as over-dispersed, zero-inflated, or heavy-tailed count processes. In Table 8, we can see a number of competing models that we will use to evaluate the EDLE distributions.

6.1 Application I

The first dataset comprises the counts of carious teeth in a sample of children's four deciduous molars. Comprehending the prevalence and distribution of caries in primary teeth is essential for public health planning and preventative care, which is why epidemiological and pediatric dentistry research often utilizes this type of dental caries count data. The dataset contains discrete, non-negative integer values that may exhibit over-dispersion, a characteristic commonly observed in

Table 8: The competitive models.

Distribution	Abbreviation	Author(s)
Exponentiated discrete Lindley	EDLi	El-Morshedy et al. [10]
One-parameter discrete Lindley	DLi-I	Gómez-Déniz and Calderín-Ojeda [12]
Two-parameter discrete Lindley	DLi-II	Hussain et al. [13]
Geometric	Geo	Gómez-Déniz [14]
Discrete generalized exponential type II	DGE-II	Nekoukhou et al. [15]
Discrete Rayleigh	DR	Roy [16]
Discrete Weibull	DW	Nakagawa & Osaki [17]
Discrete Pareto	DPa	Krishna and Pundir [18]
Poisson	Pois	Poisson [19]
Three parameters discrete Lindley	DLi-III	Eliwa et al. [20]
Discrete inverse Rayleigh	DIR	Hussain and Ahmad [21]
Discrete inverse Weibull	DIW	Jazi et al. [22]
Discrete Burr	DB	Krishna and Pundir [18]
Discrete log-logistic	DLogL	Para and Jan [23]
Discrete Burr-Hatke	DBH	El-Morshedy et al. [24]

count data of rare events or heterogeneous populations. Krishna and Pundir [18] utilized this dataset to demonstrate discrete lifetime models, offering a comprehensive history, original collection methodology, and descriptive analyses. Due to its structure, it can evaluate the efficacy of the proposed EDLE distribution in capturing the skewed and widely dispersed count patterns observed in public health and medical data. A collection of non-parametric diagnostic plots, including a box plot, violin plot, strip plot, and normal Q-Q plot, was generated to examine the essential characteristics of dataset I. Collectively, these representations, depicted in Figure 3, provide complementary insights into the distribution's shape, central tendency, dispersion, and deviation from normality. The box plot delineates potential outliers and encapsulates the data's quartiles. This image is augmented by the overlay violin plot, which illustrates the smoothed probability density and elucidates the modality and concentration of the data. The strip plot offers a lucid representation of the raw data distribution and any clustering phenomena by displaying individual data points. Finally, a direct evaluation of normality assumptions is facilitated by the normal Q-Q plot, which examines the relationship between actual quantiles and those of a theoretical normal distribution. When integrated, these plots provide a comprehensive visual assessment of the dataset's structure, which is essential for informing the selection and validation of an appropriate probability model subsequently.

The model selection criteria and maximum likelihood estimates for the competing models are shown in Table 9. With the lowest values of AIC, CAIC, and HQIC as well as the lowest $-L$, the EDLE distribution offers the best fit to the data. This suggests that, out of all the models taken into consideration, the EDLE model provides the most appropriate and economical representation of the observed data. The observed and predicted frequencies for dataset I under the proposed models and a few goodness-of-fit statistical criteria are shown in Table 10. With a chi-square statistic of 0.707 (df = 1, P-value = 0.400), the EDLE distribution provides a good fit to the observed counts. When it comes to recreating the empirical distribution, EDLE performs better than all competing models, including DLi, DR, and Pois, which exhibit notable deviations.

Table 9: The MLE and goodness-of-fit criteria for dataset I.

Model	**MLEs**	$-L$	**AIC**	**CAIC**	**HQIC**
EDLE	$\widehat{\alpha}=0.532, \widehat{\beta}=0.380$	111.45	226.90	227.02	229.01
EDLi	$\widehat{\alpha}=0.379, \widehat{\beta}=0.543$	111.45	226.91	227.03	229.02
DLi	$\widehat{\beta}=0.274$	113.68	229.36	229.39	230.41
DLi-II	$\widehat{\alpha}=0.401, \widehat{\beta}=0.001$	112.47	228.95	229.07	231.06
Geo	$\widehat{\beta}=0.401$	112.47	226.95	227.99	230.00
DGE-II	$\widehat{\alpha}=0.468, \widehat{\beta}=0.718$	111.80	227.61	227.73	229.72
DR	$\widehat{\beta}=0.665$	142.61	287.21	287.25	288.26
DW	$\widehat{\alpha}=0.374, \widehat{\beta}=0.895$	112.10	228.20	228.32	230.30
DPa	$\widehat{\beta}=0.184$	116.83	235.66	235.70	236.72
Pois	$\widehat{\beta}=0.670$	121.05	244.09	244.14	245.15

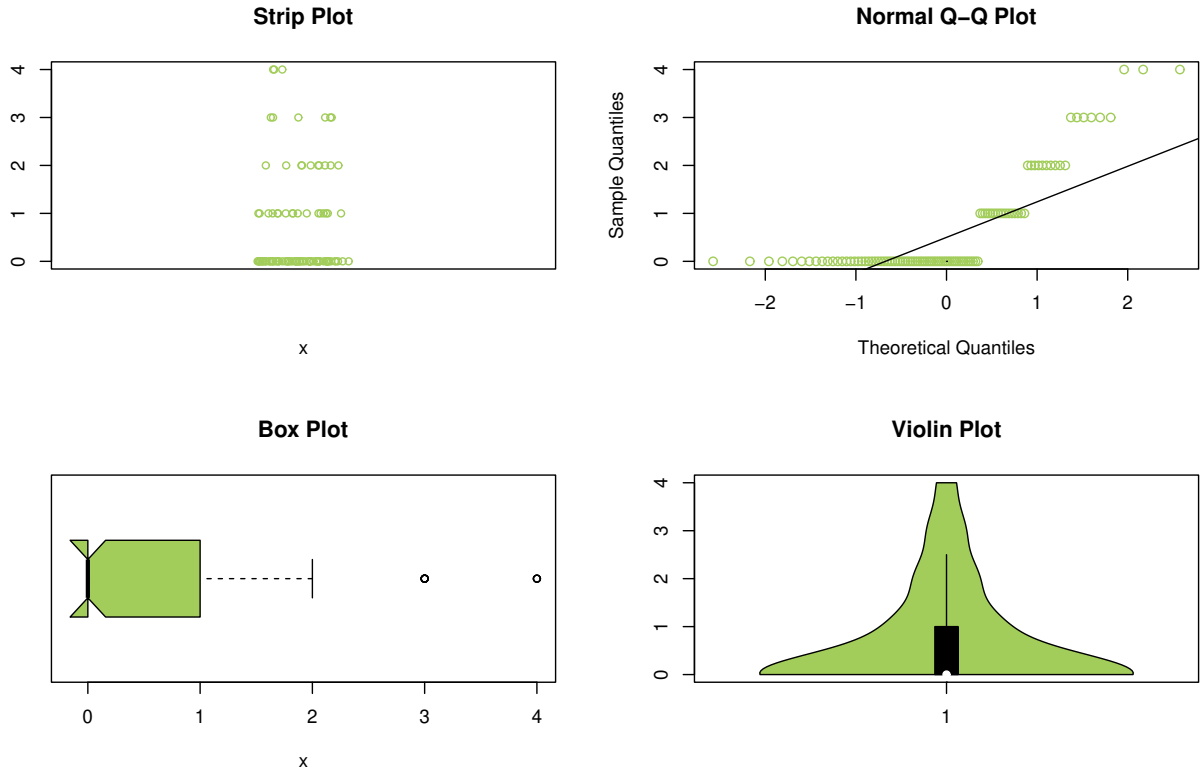

Fig. 3: Non-parametric plots for dataset I.

Table 10: The goodness-of-fit tests for dataset I.

		ExFr									
X	ObFr	EDLE	EDLi	DLi	DLi-II	Geo	DGE-II	DR	DW	DPa	Pois
0	64	63.58	63.57	57.13	59.88	59.88	63.51	33.50	62.58	69.04	51.17
1	17	19.69	19.75	26.88	24.02	24.02	20.19	46.94	21.35	15.37	34.28
2	10	9.10	9.09	10.45	9.64	9.64	8.81	17.01	8.85	6.01	11.49
3	6	4.21	4.19	3.71	3.87	3.87	4.01	2.39	3.88	3.01	2.57
≥ 4	3	3.42	3.40	1.83	2.59	2.59	3.48	0.16	3.34	6.57	0.49
Total	100	100	100	100	100	100	100	100	100	100	100
χ^2		0.707	0.739	6.638	3.347	3.347	0.973	66.07	1.507	3.225	23.65
df		1	1	2	1	2	1	2	1	2	2
P-value		0.400	0.390	0.036	0.067	0.188	0.324	$<$0.0001	0.219	0.199	$<$0.0001

The observed and expected probability mass functions (PMFs) for dataset I across eight competing models are graphically compared in Figure 4. The proposed EDLE is the best according to the statistical criteria, but the EDLE and EDLi distributions (top panels) demonstrate the closest alignment between observed and calculated frequencies, demonstrating their superior fit as measured in Tables 9 and 10. On the other hand, there are noticeable visual differences in distributions like DLi, DR, and Pois.

The profile log-likelihood plots for the parameters α and β of the EDLE model fitted to the dataset of carious teeth are shown in Figure 5. The concave unimodal shape of both plots is well defined, indicating that the maximum likelihood estimators $\widehat{\alpha}$ and $\widehat{\beta}$ are distinct and match the likelihood function's global maximum. The stability and dependability of the estimation process for this model and dataset are further supported by the smooth and regular curvature surrounding the optimum. The log-likelihood surface of the EDLE model parameters α and β for dataset I is shown in Figure 6.

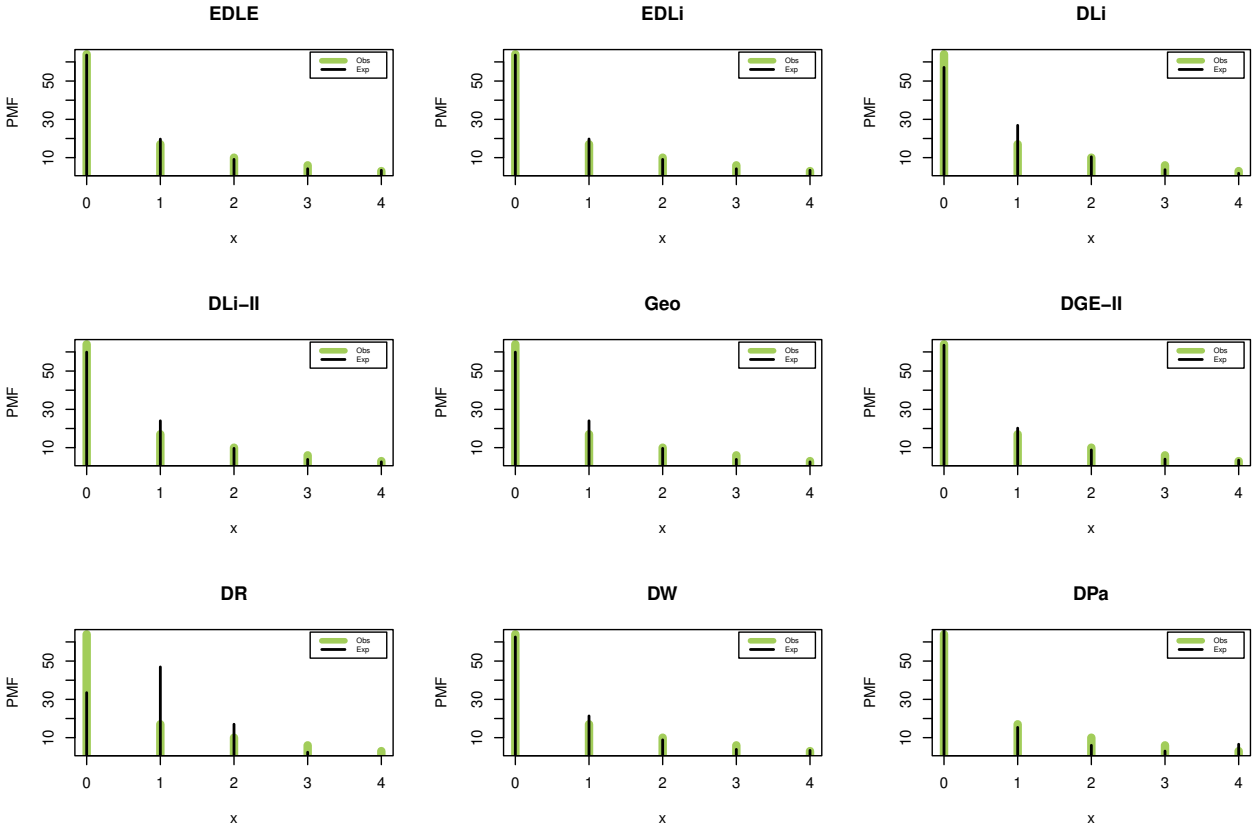

Fig. 4: The observed and expected PMFs for dataset I.

The maximum likelihood estimators' uniqueness and identifiability are confirmed by the plot's one, distinct peak. The asymptotic normality of the estimators and the dependability of standard error estimates obtained from the information matrix are supported by the elliptical contours surrounding the optimal $\widehat{\alpha}$ and $\widehat{\beta}$, which show a regular estimation surface. The stability of the numerical optimization process employed to fit the model is confirmed by this unimodal, concave form.

The descriptive statistics of the fitted empirical distribution from the EDLE model and the observed dental caries data are contrasted in Table 11. The observed mean (0.668 vs. 0.670) is closely matched by the EDLE model. An IoD greater than one (1.969 vs. 1.721) indicates that it successfully represents the over-dispersed nature of the data, albeit somewhat overestimating the variance, skewness, and kurtosis. This validates the model's ability to capture the dataset's primary distributional features.

Table 11: Descriptive statistics for dataset I.

	Mean	Variance	Skewness	Kurtosis	IoD
Observed	0.670	1.153	1.596	4.767	1.721
Empirical	0.668	1.315	2.376	10.505	1.969

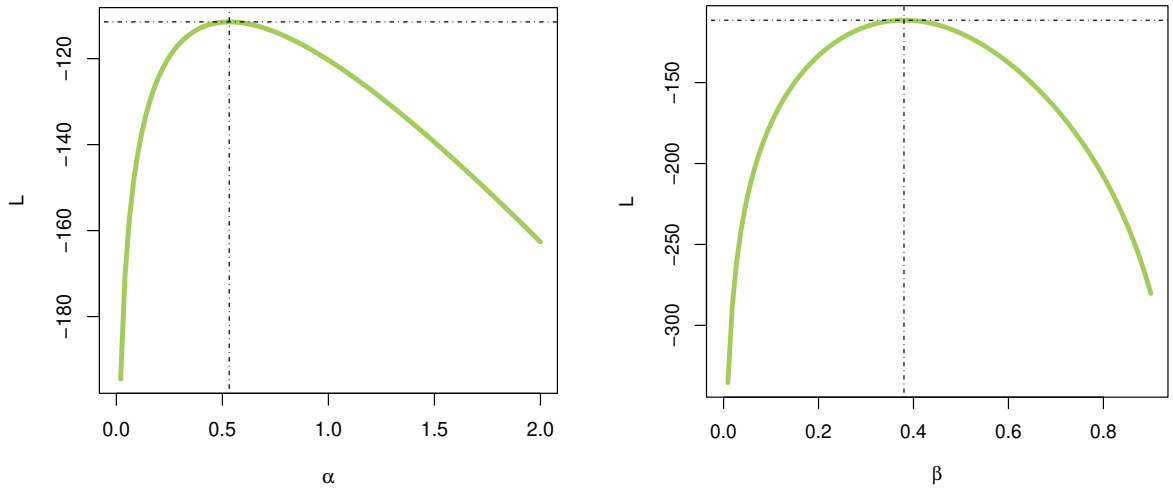

Fig. 5: The L profiles of parameters of EDLE for dataset I.

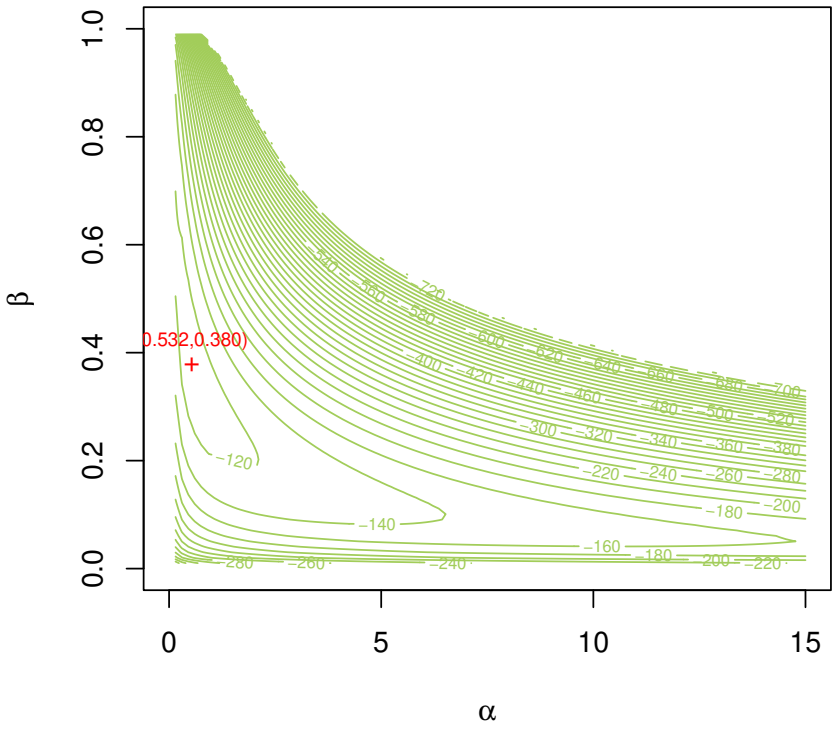

Fig. 6: Contour diagrams of the model estimators based on dataset I.

6.2 Application II

The second dataset comprises counts of European red mites (Panonychus ulmi) seen on apple leaves, initially documented by Chakraborty and Chakravarty [25]. This dataset exemplifies entomological and agricultural field investigations, documenting the insect count per plant unit to evaluate infestation levels and guide integrated pest management tactics. The counts are whole numbers that are not negative, and they usually show over-dispersion because of geographical aggregation and environmental heterogeneity, which are frequent in ecological count data. The structure of the dataset serves as a pertinent test case for assessing the adaptability of the EDLE distribution in simulating skewed, over-dispersed, and potentially zero-inflated counts observed in agricultural and biological monitoring. Figure 7 displays a collection of non-parametric diagnostic plots, including a violin plot, box plot, strip plot, and normal Q-Q plot, to delineate the underlying distribution of the European red mite counts.

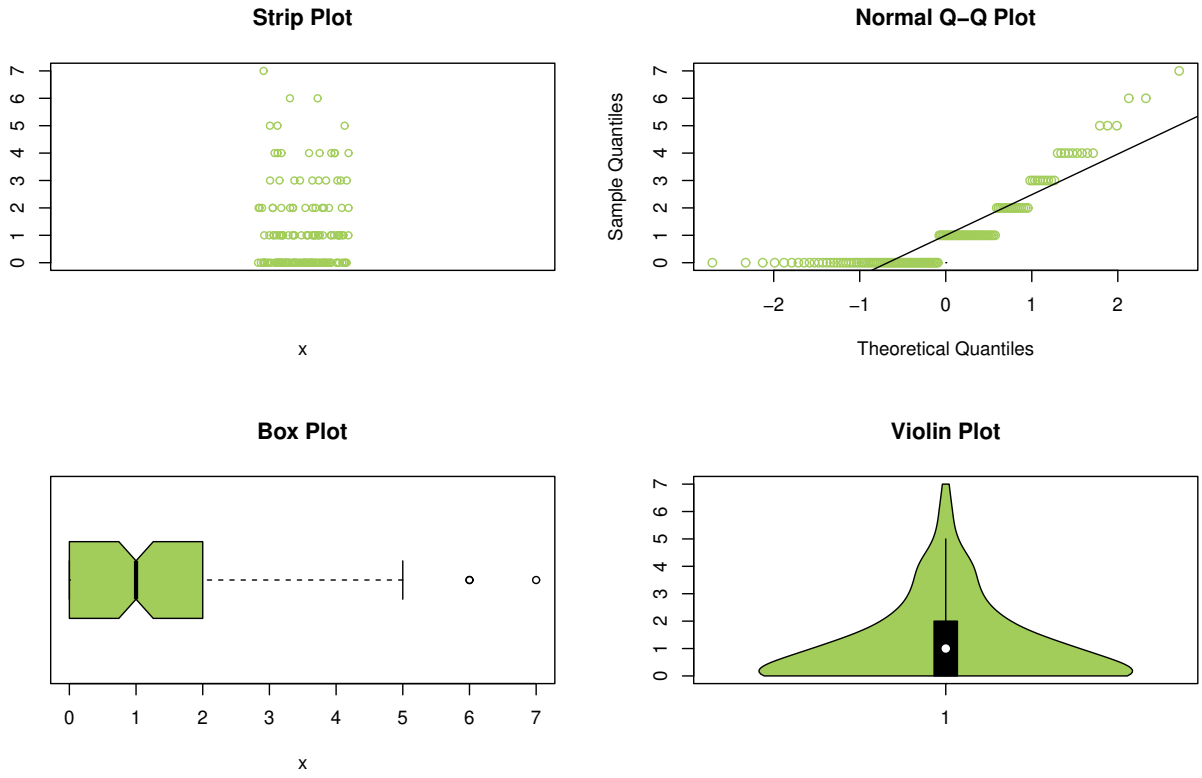

Fig. 7: Non-parametric plots for dataset II.

The violin plot shows a kernel density estimate and a box plot at the same time. It shows a distribution that is substantially right-skewed and unimodal, with a lot of counts of zero. The box plot shows that there are multiple upper-tail outliers, which means that the data is too spread out. The strip plot shows all of the individual observations, making it evident that the data is discrete and non-negative, and that there are a lot of zeros. Finally, the normal Q-Q plot displays a clear difference from the expected diagonal line, especially in the upper tail. This means that the assumption of normality is not true. These visuals together provide a case for using a flexible, over-dispersed discrete model like the EDLE distribution for this kind of ecological count data. Table 12 shows that the EDLE model fits the European red mite dataset the best because it meets the lowest values of all information criteria among the competing distributions.

Table 13 shows the expected and observed frequencies for dataset II based on the candidate models. The EDLE distribution fits the data best since it closely matches the observed numbers and has a chi-square statistic of 2.373 (df = 3, P-value = 0.499), which shows that it fits very well. Other models, such as DIR, DIW, DPa, DBH, Pois, DB and DLogL, exhibit substantial discrepancies and inadequately replicate the empirical distribution. So, the EDLE is the best model for dataset II.

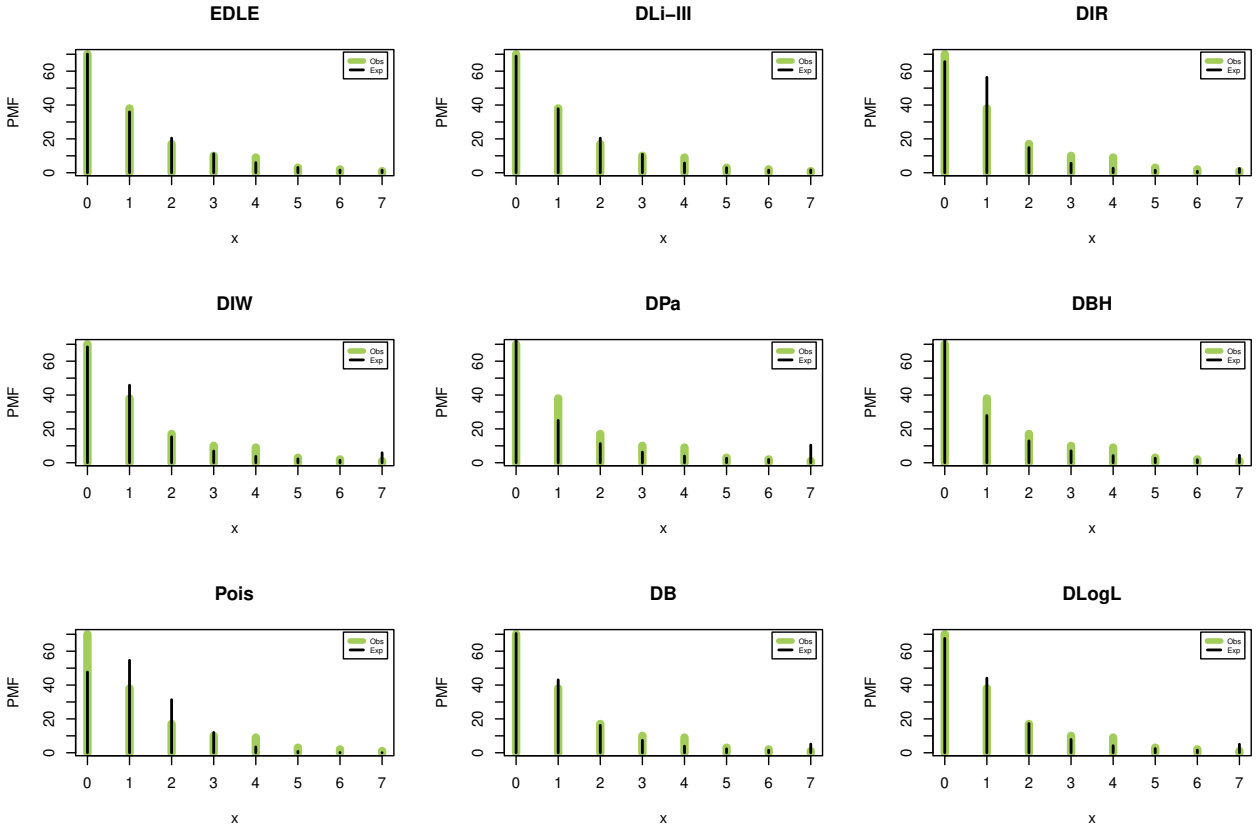

Fig. 9: The observed and expected PMFs for dataset II.

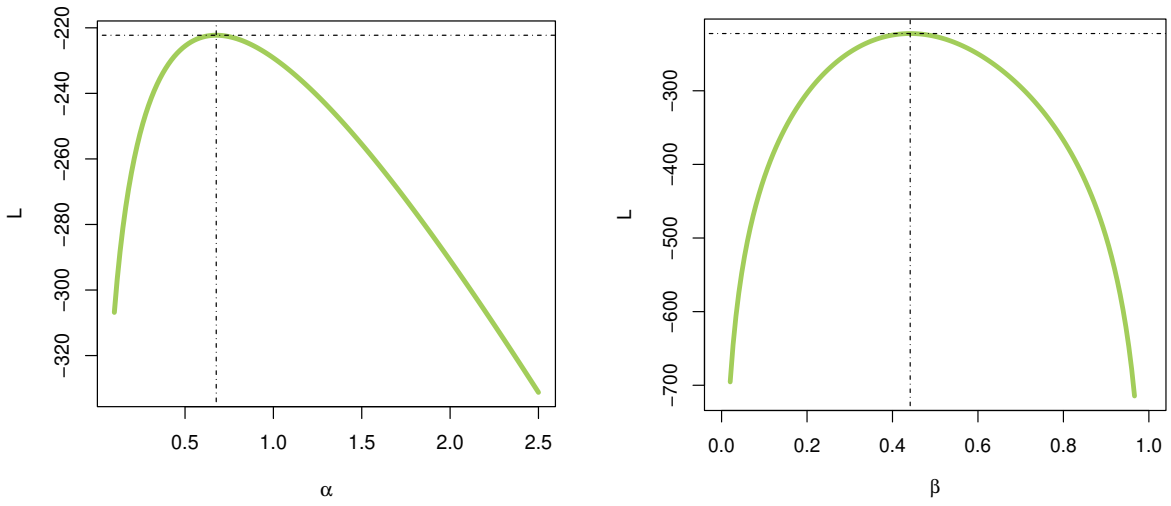

Fig. 8: The L profile of parameters of the EDLE for dataset II.

Table 12: The MLE and goodness-of-fit criteria for dataset II.

Model	**MLEs**	$-L$	**AIC**	**CAIC**	**HQIC**
EDLE	$\widehat{\alpha}=0.677, \widehat{\beta}=0.441$	222.265	448.529	448.611	450.975
DLi-III	$\widehat{\alpha}=1.468, \widehat{\beta}=0.228, \widehat{\lambda}=0.479$	222.383	450.765	450.390	454.435
DIR	$\widehat{\theta}=0.438$	233.142	468.284	468.311	469.507
DIW	$\widehat{\alpha}=0.456, \widehat{\beta}=1.527$	229.333	462.666	462.747	465.112
DPa	$\widehat{\alpha}=0.278$	238.832	479.663	479.690	480.886
DBH	$\widehat{\alpha}=0.814$	230.552	463.103	463.130	464.326
Pois	$\widehat{\alpha}=1.147$	242.809	487.619	487.647	488.843
DB	$\widehat{\alpha}=0.400, \widehat{\beta}=1.882$	227.727	459.454	459.536	461.901
DLogL	$\widehat{\alpha}=1.116, \widehat{\beta}=1.829$	227.265	458.531	458.613	460.977

Table 13: The goodness-of-fit tests for dataset II.

		ExFr								
X	ObFr	EDLE	DLi-III	DIR	DIW	DPa	DBH	Pois	DB	DLogL
0	70	70.159	68.902	65.658	68.411	88.308	88.938	47.654	70.469	67.527
1	38	35.911	37.819	56.351	45.814	25.005	27.919	54.643	43.053	44.099
2	17	20.458	20.423	14.835	15.307	11.314	12.905	31.329	16.214	17.266
3	10	11.232	10.889	5.608	6.935	6.312	7.056	11.975	7.364	7.874
4	9	5.983	5.746	2.673	3.777	3.972	4.238	3.433	3.924	4.167
5	3	3.111	3.007	1.473	2.311	2.705	2.702	0.787	2.338	2.458
6	2	1.586	1.562	0.895	1.530	1.948	1.795	0.150	1.509	1.569
7	1	1.560	1.652	2.507	5.915	10.436	4.447	0.029	5.129	5.040
Total	150	150	150	150	150	150	150	150	150	150
χ^2		2.373	2.515	17.376	11.306	26.916	15.573	26.646	8.829	7.843
df		3	2	3	3	4	4	2	2	2
P-value		0.499	0.284	≤ 0.001	0.010	≤ 0.001	0.004	≤ 0.001	0.012	≤ 0.001

Figure 9 shows the profile log-likelihood functions for α and β. Each one has a clear, unimodal peak, which shows that the MLEs are unique and stable. The elliptical contours centered at $(\widehat{\alpha}, \widehat{\beta}) = (0.677, 0.441)$ show that the estimation surface is regular and concave, which supports the asymptotic normality of the estimators and the accuracy of the numerical optimization. These figures together show that the method for estimating parameters for the EDLE model using this dataset is strong.

Table 14 shows a comparison of the descriptive statistics for the European red mite dataset with the fitted EDLE model. The EDLE distribution fits the mean 1.146 vs. 1.147 very well and accurately represents the over-dispersed variance, as shown by an index of dispersion more than one 2.058 vs. 1.983. The fitted model predicts increased skewness and kurtosis, but it does a good job of reproducing the basic distributional features of the ecological count data that were seen.

Table 14: Descriptive statistics for dataset II.

	Mean	Variance	Skewness	Kurtosis	IoD
Observed	1.147	2.274	1.545	1.315	1.983
Empirical	1.146	2.358	1.902	7.915	2.058

6.3 Application III

The third dataset contains 40 discrete observations of the time it takes for turbochargers from a certain engine model to fail. The times are measured in 10^3h and rounded to the nearest whole hour [26]. Such dependability statistics are essential

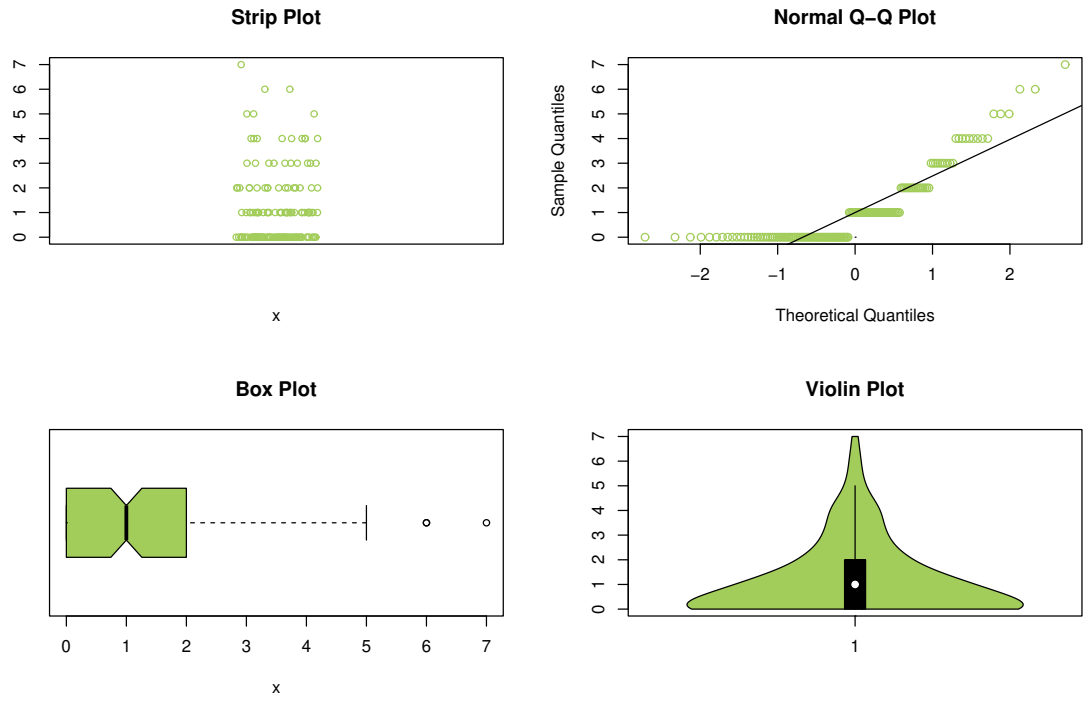

Fig. 10: Contour diagrams of the model estimators based on dataset II.

in engineering for evaluating component longevity, enhancing maintenance plans, and predicting system performance. The rounding procedure changes continuous survival times into a discrete count variable. This generally results in a right-skewed and sometimes over-dispersed distribution because of different types of wear and operational stressors. This dataset offers a pragmatic context for evaluating the EDLE model's ability to characterize discrete failure counts exhibiting non-exponential behavior. Figure 10 shows the diagnostic plots for the discretized turbocharger failure. These include a violin plot, a box plot, a strip plot, and a conventional Q-Q plot. The violin and box plots show a clear left skew. The strip plot shows that the data is made up of separate, whole numbers. The normal Q-Q plot shows a big difference from the expected diagonal, especially in the upper tail, which means that normalcy is not true. These traits show that a flexible discrete model, like the EDLE distribution, is needed to accurately show the underlying failure-time pattern.

Tables 15 and 16 show the maximum likelihood estimates and goodness-of-fit values for the data on turbocharger failures. Table 15 shows the parameter estimates and model selection criteria for Dataset III. The EDLE distribution has the lowest negative log-likelihood (87.178) and the lowest AIC (178.356), CAIC (178.680), and HQIC (179.577), which means it strikes the optimal balance between model fit and complexity. The DGE-II distribution is a strong contender, but its information criteria are a little higher. Other models, such DLi, Geo, DPa, and DBu, don't work well and don't show that the data is skewed to the left. So, EDLE is the best model for Dataset III. Table 16 further confirms these results by indicating that EDLE fits better than what was shown in Table 15. Figure 11 shows that the EDLE model fits the turbocharger failure data better than any other model. It shows that the observed and expected frequencies match perfectly. This is in line with what Tables 15 and 16 say about the numbers. Figure 12 displays the profile log-likelihood functions for α and β. Each one has a clear, unimodal peak, which suggests that the maximum likelihood estimates are well-defined and unique. Figure 13 shows the corresponding contour plot of the joint log-likelihood surface. The elliptical contours are tightly centered around the optimum point (8.651, 0.542). The regular, concave shape of the plot shows that the estimates are stable and supports the use of asymptotic normal inference for the parameters. This strengthens the reliability of the model fitting process for this reliability dataset.

Table 17 compares the descriptive statistics for the turbocharger failure data with the EDLE model that fits it. The EDLE distribution gives a good estimate of the mean that was seen (6.437 vs. 6.325). The fitted variance is greater, but the model still captures the main feature of under-dispersion, as shown by an index of dispersion below one (0.937 vs. 0.571). The fitted distribution has a positive skewness, while the observed skewness is somewhat negative. However, the fitted distribution does a good job of showing the main distributional features of the discretized reliability data.

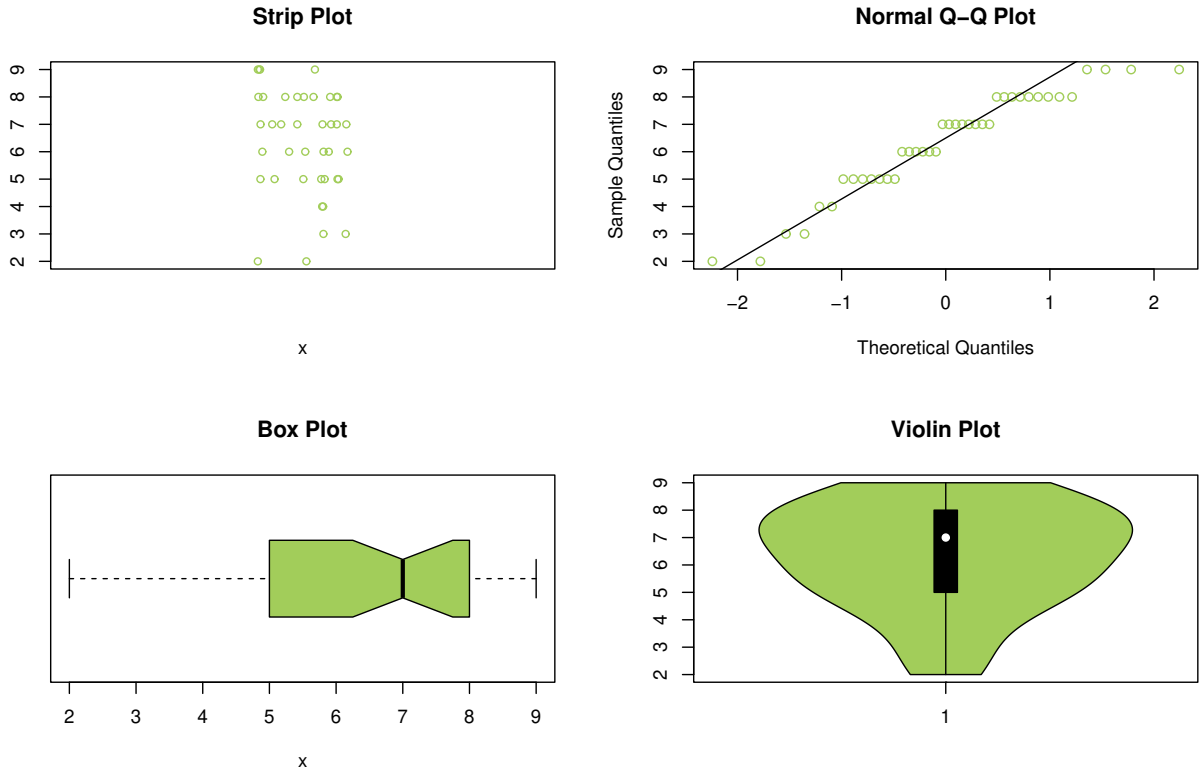

Fig. 11: Non-parametric plots for dataset III.

Table 15: The MLE and goodness-of-fit criteria for dataset III.

Model	MLEs	$-L$	AIC	CAIC	HQIC
EDLE	$\widehat{\alpha} = 8.651, \widehat{\beta} = 0.542$	87.178	178.356	178.680	179.577
DLi	$\widehat{\beta} = 0.768$	107.08	216.170	216.280	216.790
Geo	$\widehat{\beta} = 0.863$	116.78	235.570	235.680	236.180
DPa	$\widehat{\beta} = 0.609$	148.940	299.890	299.990	300.500
DGE II	$\widehat{\alpha} = 17.317, \widehat{\beta} = 0.608$	87.890	179.780	180.110	181.010
DBu	$\widehat{\alpha} = 10.107, \widehat{\beta} = 0.948$	139.770	283.550	283.870	284.770

7 Conclusion and Future Work

This study introduced a versatile and dependable model for discrete data characterized by diverse risk behaviors and dispersion patterns, represented by a unique two-parameter EDLE distribution. We conducted a comprehensive theoretical examination. In discrete-time contexts, reliability and survival analysis are optimally aligned with the EDLE distribution, which can accommodate increasing, decreasing, and bathtub-shaped hazard rate patterns, as indicated by these analytical developments. The suggested EDLE model demonstrated significant adaptability in modeling favorably skewed data characterized by varying levels of kurtosis, overdispersion, and heavy-tailed behavior. As the parameters are elevated, both the mean and variance increase, while skewness decreases, indicating a transition to symmetry, as per the systematic examination of distributional characteristics. The EDLE distribution exhibits leptokurtosis, since its kurtosis consistently exceeds three, although diminishing with increasing parameters. The index of dispersion further illustrated the model's inherent overdispersion, particularly for larger form parameter values, emphasizing that both parameters contribute to variability. Rényi entropy conformed to its theoretical attributes across several orders, as evidenced by entropy-based measurements, thus illustrating the versatility of the EDLE distribution in encapsulating diverse

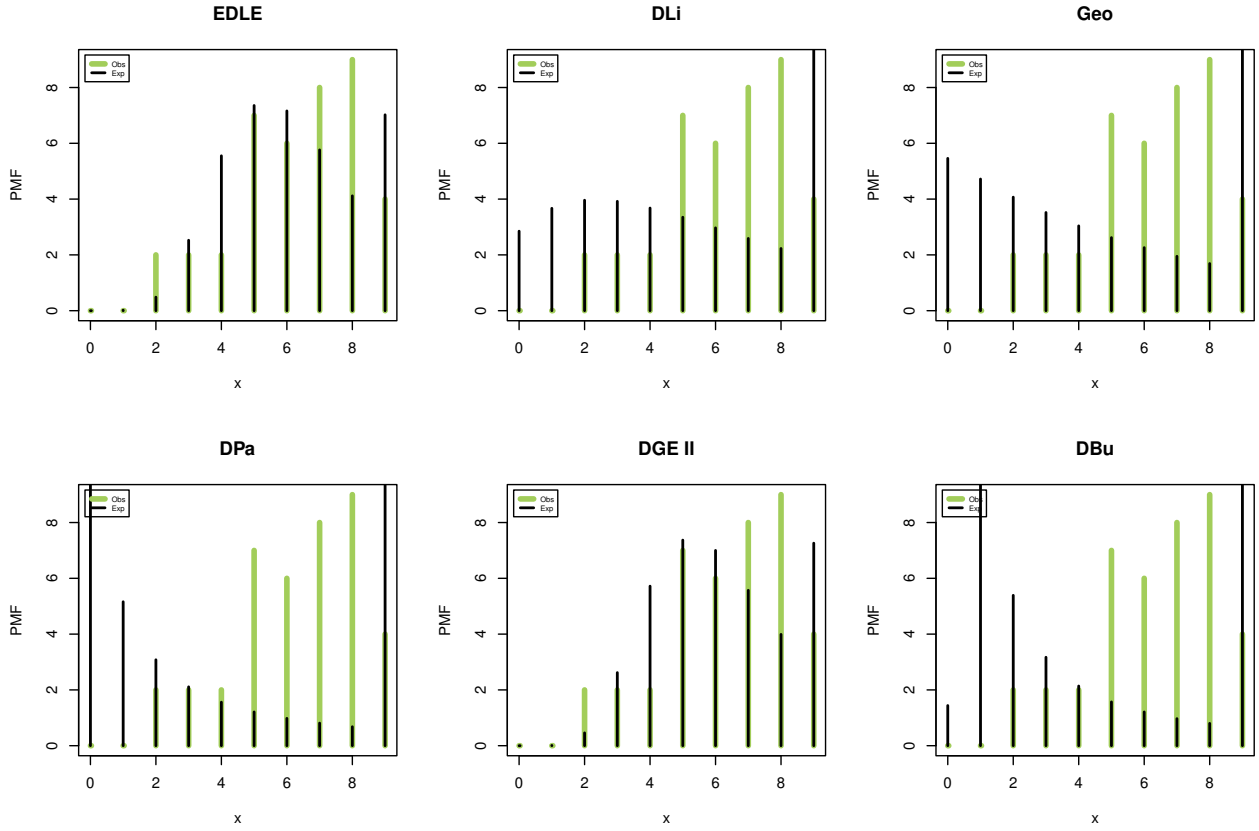

Fig. 12: The observed and expected PMFs for dataset III.

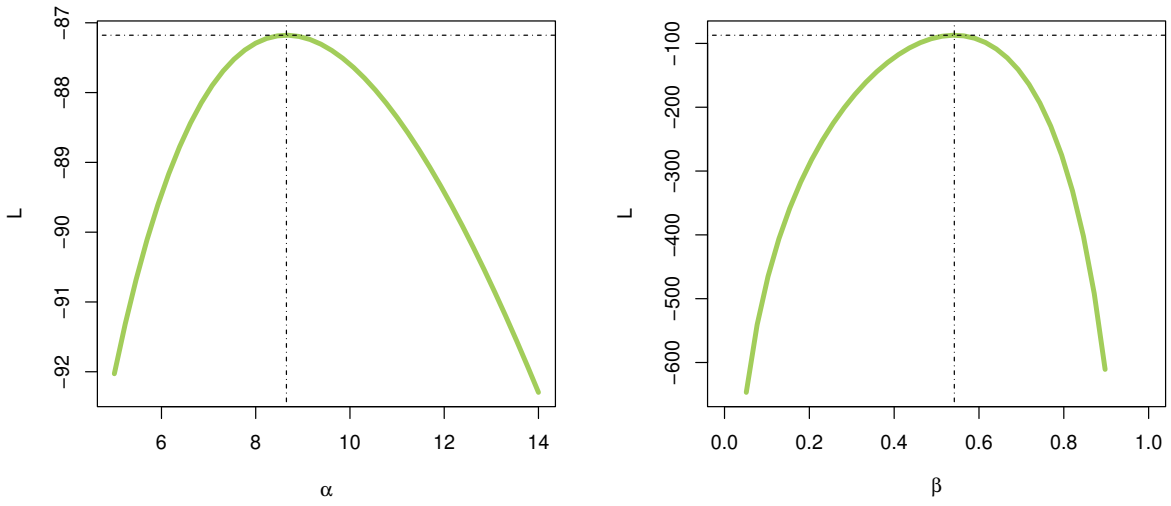

Fig. 13: The L-profile of the parameters of the EDLE for dataset III.

Table 16: The goodness-of-fit tests for dataset III.

		ExFr					
X	ObFr	EDLE	DLi	Geo	DPa	DGE II	DBu
0	0	0.000	2.850	5.460	11.620	0.000	1.440
1	0	0.019	3.670	4.720	5.160	0.010	10.96
2	2	0.485	3.960	4.070	3.080	0.460	5.390
3	2	2.525	3.920	3.520	2.110	2.620	3.170
4	2	5.552	3.680	3.040	1.560	5.720	2.140
5	7	7.355	3.350	2.620	1.210	7.370	1.570
6	6	7.161	2.970	2.260	0.980	7.000	1.210
7	8	5.764	2.590	1.950	0.810	5.570	0.970
8	9	4.117	2.230	1.690	0.680	3.990	0.800
9	4	7.022	10.780	10.670	12.790	7.260	12.34
Total	40	40	40	40	40	40	40
χ^2		2.16	23.22	38.33	127.16	2.38	99.92
df		2	3	3	3	2	2
P-value		0.340	$<$0.001	$<$0.010	$<$0.001	0.304	$<$0.001

Table 17: Descriptive statistics for dataset III.

	Mean	Variance	Skewness	Kurtosis	IoD
Observed	6.325	3.610	-0.613	2.988	0.571
Empirical	6.437	6.030	1.042	4.965	0.937

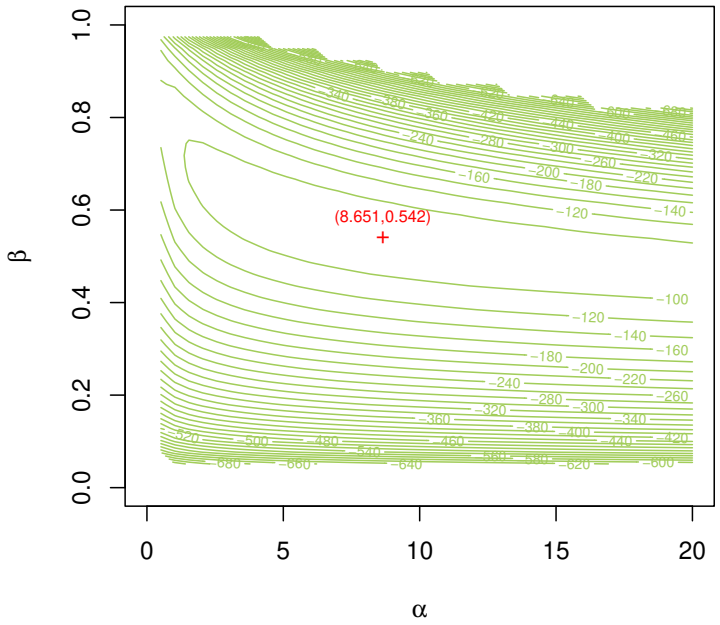

Fig. 14: Contour diagram of the model estimators based on dataset III.

uncertainty frameworks. A complete Monte Carlo simulation analysis was employed to assess the maximum likelihood method for parameter estimation. The simulation data robustly corroborated the estimators' favorable asymptotic characteristics, including consistency, asymptotic unbiasedness, and enhanced efficiency with larger sample sizes. Efficiency exhibited a consistent increase, whereas empirical bias, mean squared error, and variance shown a systematic decline as sample size increased. Moreover, the coverage probabilities of the confidence intervals were close to their nominal levels, signifying that the proposed inferential framework was dependable. The practical significance of the EDLE distribution was further validated by its consistent superiority over other competing discrete models when utilized on real datasets from biological, agricultural, and engineering fields. Several prospective avenues for future research merit further investigation, notwithstanding its robust theoretical and empirical efficacy. To address datasets with excessive zeros or structural limitations, one approach is to extend the EDLE distribution to include zero-inflated, hurdle, or truncated variations. Secondly, to enhance inference in complicated data or small sample scenarios, Bayesian estimation utilizing informative priors and Markov chain Monte Carlo procedures may be explored. Thirdly, dependency structures in associated discrete outcomes may be modeled utilizing multivariate and bivariate variants of the EDLE distribution. The application of the EDLE framework in survival analysis and count data modeling could be enhanced by incorporating censoring, truncation, and regression structures. Finally, a promising direction for future research is the application of the EDLE distribution to spatio-temporal and discrete time-series frameworks. Ultimately, the EDLE distribution significantly enhances the collection of discrete probability models. It possesses robust theoretical foundations and effectively models complex discrete data across various domains.

Acknowledgment

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/01/35354).

Conflicts of Interest

The authors declare no conflict of interest.

Data Availability Statement

The datasets are available in the paper.

References

[1] Gupta, R. D., Kundu, D. Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics* **41(2)**, 173-188 (1999).
[2] Pal, M., Ali, M. M., Woo, J. Exponentiated Weibull distribution. *Statistica* **66(2)**, 139-147 (2006).
[3] Nadarajah, S., Gupta, A. K. The exponentiated gamma distribution with application to drought data. *Calcutta Statistical Association Bulletin* **59(1-2)**, 29-54 (2007).
[4] Salem, H. M. The exponentiated Lomax distribution: Different estimation methods. *American Journal of Applied Mathematics and Statistics* **2(6)**, 364-368 (2014).
[5] Shawky, A. I., Abu-Zinadah, H. H. Exponentiated Pareto distribution: Different method of estimations. *International Journal of Contemporary Mathematical Sciences* **4(14)**, 677-693 (2009).
[6] Almalki, S. J., Nadarajah, S. A new discrete modified Weibull distribution. *IEEE Transactions on Reliability* **63(1)**, 68-80 (2014).
[7] Chakraborty, S., Gupta, R. D. Exponentiated geometric distribution: Another generalization of geometric distribution. *Communications in Statistics–Theory and Methods* **44(6)**, 1143-1157 (2015).
[8] Cardial, M. R. P., Fachini-Gomes, J. B., Nakano, E. Y. Exponentiated discrete Weibull distribution for censored data. *Brazilian Journal of Biometrics* **38(1)**, 35-56 (2020).
[9] MirMostafaee, S. M. T. K., Hamed Mashhadzadeh, Z. The exponentiated discrete inverse Rayleigh distribution. *Journal of Hyperstructures* **9(1)**, 54-61 (2020).
[10] El-Morshedy, M., Eliwa, M. S., Nagy, H. A new two-parameter exponentiated discrete Lindley distribution: properties, estimation and applications. *Journal of Applied Statistics* **47(2)**, 354-375 (2020).
[11] El-Morshedy, M. A discrete linear-exponential model: Synthesis and analysis with inference to model extreme count data. *Axioms* **11(10)**, 531 (2022).

[12] Gomez-Déniz, E., Calderin-Ojeda, E. The discrete Lindley distribution: Properties and applications. *Journal of Statistical Computation and Simulation* **81(11)**, 1405-1416 (2011).
[13] Hussain, T., Aslam, M., Ahmad, M. A two-parameter discrete Lindley distribution. *Revista Colombiana de Estadística* **39(1)**, 45-61 (2016).
[14] Gomez-Déniz, E. Another generalization of the geometric distribution. *Test* **19(2)**, 399-415 (2010).
[15] Nekoukhou, V., Alamatsaz, M. H., Bidram, H. Discrete generalized exponential distribution of a second type. *Statistics* **47(4)**, 876-887 (2013).
[16] Roy, D. Discrete Rayleigh distribution. *IEEE Transactions on Reliability* **53(2)**, 255-260 (2004).
[17] Nakagawa, T., Osaki, S. The discrete Weibull distribution. *IEEE Transactions on Reliability* **24(5)**, 300-301 (2009).
[18] Krishna, H., Pundir, P. S. Discrete Burr and discrete Pareto distributions. *Statistical Methodology* **6(2)**, 177-188 (2009).
[19] Poisson, S. D. Recherches sur la probabilité des jugements en matière criminelle et en matière civile: Précédées des règles générales du calcul des probabilités. *Bachelier* (1837).
[20] Eliwa, M. S., Altun, E., El-Dawoody, M., El-Morshedy, M. A new three-parameter discrete distribution with associated INAR(1) process and applications. *IEEE Access* **8**, 91150-91162 (2020).
[21] Hussain, T., Ahmad, M. Discrete inverse Rayleigh distribution. *Pakistan Journal of Statistics* **30(2)**, [Pages if available] (2014).
[22] Jazi, M. A., Lai, C. D., Alamatsaz, M. H. A discrete inverse Weibull distribution and estimation of its parameters. *Statistical Methodology* **7(2)**, 121-132 (2010).
[23] Para, B. A., Jan, T. R. Discrete version of log-logistic distribution and its applications in genetics. *International Journal of Modern Mathematical Sciences* **14(4)**, 407-422 (2016).
[24] El-Morshedy, M., Eliwa, M. S., Altun, E. Discrete Burr-Hatke distribution with properties, estimation methods and regression model. *IEEE Access* **8**, 74359-74370 (2020).
[25] Chakraborty, S., Chakravarty, D. Discrete gamma distributions: Properties and parameter estimations. *Communications in Statistics–Theory and Methods* **41(18)**, 3301-3324 (2012).
[26] Xu, K., Xie, M., Tang, L. C., Ho, S. L. Application of neural networks in forecasting engine systems reliability. *Applied Soft Computing* **2(4)**, 255-268 (2003).