Mathematics

AIMS Mathematics, 11(1): 1145–1174.
DOI: 10.3934/math.2026049
Received: 18 September 2025
Revised: 19 November 2025
Accepted: 28 November 2025
Published: 15 Juanary 2026

https://www.aimspress.com/journal/Math

Research article

An innovative discrete distribution for modeling genotoxicity data: Applications in pharmacology and environmental health

Amal S. Hassan[1], Eslam Abdelhakim Seyam[2,*], Asma Ahmad Alzahrani[3], Omar A. Saudi[1,4]

[1] Faculty of Graduate Studies for Statistical Research, Cairo University, 5 Dr. Ahmed Zewail Street, Giza, 12613, Egypt
[2] Department of Insurance and Risk Management, College of Business, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia
[3] Department of Mathematics, Faculty of Science, Al-Baha University, Al-Baha 65779, Saudi Arabia
[4] Department of Basic Sciences, Higher Institute of Management Sciences (HIMS), El Katameya, New Cairo 3,11936, Egypt

* **Correspondence:** Email: isiam@imamu.edu.sa.

Abstract: This paper proposes a flexible probability mass function for modeling count data, particularly over-dispersed and asymmetric observations. A novel two-parameter discrete distribution, the discrete power inverted Topp–Leone distribution, is presented using a survival discretization technique. The following statistical characteristics are examined: factorial moments, probability-generating function, quantiles, mean, variance, mean residual life, and entropy measures. The best estimators of the unknown parameters were obtained using various techniques, such as maximum likelihood, moments, least squares, Anderson–Darling, and Cramér-von Mises. A simulation study showed that the accuracy of the estimates improves with larger samples, although higher parameter values may affect precision. The findings indicate that the efficiency of these estimation methods varies under different conditions. Applications to liver lesions, chromatid aberrations, and criminal sociology datasets confirm the model's usefulness for discrete count data in fields such as social sciences, pharmacology, and environmental health. Finally, for modeling count data, the new probabilistic model can be employed as a competitive alternative to other existing distributions.

Keywords: count distribution; discretization; survival function; Anderson-Darling; factorial moments
Mathematics Subject Classification: 60E05, 62E10, 62F10, 62N05, 62P10

1. Introduction

In a variety of disciplines, such as applied science, epidemiology, agriculture, sociology, public health, and medicine, count data modeling is crucial. Although many probability distributions have been developed to model count data, more flexible discrete distributions are required to handle count datasets with high over-dispersion. The widely used traditional discrete probability models, such as Poisson, geometric, and negative binomial distributions, have limitations because of their unique behaviors. For example, the Poisson distribution works best with datasets with equidispersion, while the negative binomial distribution works best with datasets that are overly dispersed. Flexible discrete distributions always require a good resolution because real-life datasets can be either over- or under-dispersed. There are several ways to discretize continuous distributions, such as infinite series methods, mixed-Poisson models, and survival discretization, the most widely applied of these [1]. More specifically, if W is a continuous random variable with survival function (SF), $S(w) = P(W \geq w)$, then $P(W = w) = P(w \leq W < w + 1) = P(W \geq w) - P(W \geq w + 1) = S(w) - S(w + 1)$, w = 0, 1, 2,..., gives the probability mass function (PMS) of its analogue discrete random variable W with support for the set of nonnegative integer numbers.

The discretization of continuous probability distributions has received a lot of attention in recent decades. Additionally, numerous authors presented new discrete models, some of which are listed as follows: discrete Weibull distribution [2], discrete Burr (DBurr) and Pareto distributions [3], discrete inverse Weibull (DIW) distribution [4], discrete generalized Burr-XII distribution [5], discrete log-logistic distribution (DLL) [6], discrete extended Weibull distribution [7], discrete Burr-Hatke distribution [8], discrete Xgamma distribution [9], discrete Bilal distribution [10], discrete Teissier distribution [11], discrete inverted Topp–Leone distribution [12], discrete Nadharajah-Haghighi distribution [13], discrete two-parameter Lindley distribution [14], discrete Marshall-Olkin-Weibull distribution [15], discrete Gumbel distribution [16], discrete Ramos-Louzada distribution [17], discrete quasi Xgamma distribution [18], discrete new XLindley distribution [19], discrete inverted Kumaraswamy (DIKU) distribution [20], discrete Marshall-Olkin Burr XII distribution[21], and discrete Muth distribution [22].

Even though there are many different lifetime data distributions available, there is still a need for more adaptable models that can capture complex data behavior in a range of scenarios. The continuous power inverted Topp-Leone (PITL) distribution introduced by [23] has proven useful in a variety of reliability scenarios. The cumulative distribution function (CDF) and SF of the PITL model are given, respectively, by:

$$F(w;\varsigma) = 1 - \left[\frac{(1+2w^{\eta})^{\kappa}}{(1+w^{\eta})^{2\kappa}}\right]; \eta, \kappa, w > 0, \tag{1}$$

and

$$S(w;\varsigma) = \frac{(1+2w^{\eta})^{\kappa}}{(1+w^{\eta})^{2\kappa}},$$

where κ and η are shape parameters and $\varsigma = (\kappa, \eta)^T$ is the parameter set. The CDF (1) reduces to inverted Topp-Leone distribution as proposed by [24]. As mentioned by [23], the PITL distribution has proven useful in the engineering field and has multiple applications in various fields, as investigated by numerous researchers (see, e.g., El-Saeed et al. [25], Nassr et al. [26], Ibrahim et al. [27], and Hassan

and Almetwally [28]).

In many fields, count data are frequently too dispersed, which makes traditional statistical models difficult to use. We provide the discrete PITL (DPITL) distribution as a solution to this problem. In order to efficiently model over-dispersed discrete data, we primarily contribute by introducing the DPITL distribution, which is derived from the survival discretization method. It should be noted that the discrete ITL (DITL) distribution, which was first presented by Eldeeb et al. [12], can be considered as a particularly special model from the novel DPITL distribution. The following are the primary reasons to investigate this novel's distribution.

- The DPITL distribution has a decreasing and upside-down bathtub hazard rate function (HRF) and can handle over-dispersed count data.
- To determine various statistical characteristics, such as moments, quantile function (QF), median, factorial moments, moment generating function (MGF), dispersion index, coefficient of variation, mean residual lifetime, probability-generating function (PGF), and entropy measures.
- To use the maximum likelihood (ML), moments (MM), least squares (LS), Anderson–Darling (AD), and Cramér–von Mises (CM) estimation approaches to accurately estimate the parameters of the DPITL distribution. Additionally, a comprehensive simulation analysis is conducted to assess the robustness of the proposed estimators.
- The applicability of the proposed model in comparison to the others is demonstrated using three real-world datasets from social sciences, environmental health, and pharmacology.

The remainder of this paper is divided into seven sections. In Section 2, the DPITL distribution's construction is derived. Some of its fundamental traits are explained in Section 3. In Section 4, the ML, CM, LS, AD, and MM techniques are used to estimate the DPITL distribution parameters. In Section 5, the performance of the recently proposed estimates is assessed using Monte Carlo simulation results. The practical applications of DPITL distribution are examined in Section 6. The key findings and implications from this study's examination of the DPITL distribution are summed up in Section 7.

2. The discrete power inverted Topp–Leone distribution

This section provides the probability mass function (PMF) of the DPITL distribution, CDF, HRF, and reversed HRF. The PMF of a random variable W has the DPITL distribution provided by using the survival discretization method as follows:

$$\rho(W=w;\varsigma)=P(W\geq w)-P(W\geq w+1)=\frac{(1+2w^{\eta})^{\kappa}}{(1+w^{\eta})^{2\kappa}}-\frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}};\ w\in N_0, \tag{2}$$

where $N_0=0,1,2,\ldots$ and $\eta,\kappa>0$. The CDF of DPITL distribution is given by

$$\beta(w;\varsigma)=P(W\leq w)=1-P(W>w)=1-S(w+1)=1-\frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}};\ w\in N_0. \tag{3}$$

So, the SF and HRF of the DPITL distribution are obtained as follows:

$$S(w;\varsigma)=P(W\geq w)=1-F(w)+\rho(W=w)=1-F(w+1)=\frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}},$$

$$\psi(w;\varsigma)=\frac{\rho(W=w;\varsigma)}{S(w;\varsigma)}=\frac{(1+2w^{\eta})^{\kappa}(1+(w+1)^{\eta})^{2\kappa}}{(1+w^{\eta})^{2\kappa}(1+2(w+1)^{\eta})^{\kappa}}-1. \tag{4}$$

To show the shapes of the PMF (2) and HRF (4), Figures 1 and 2 depict their plots for varying values of the parameters κ and η. The PMF curves in Figure 1 are right-skewed and decreasing, which highlights the DPITL distribution's effective handling of skewed data. The HRF in Figure 2 is also very flexible, with useful shapes like decreasing and upside-down curves that make it a great choice for various count models.

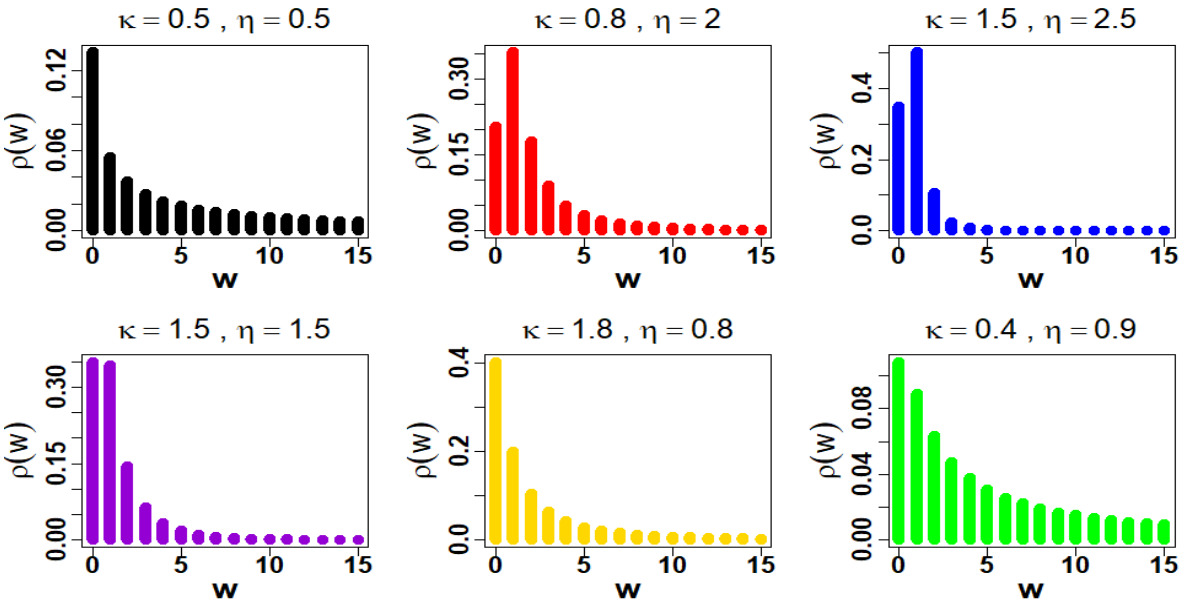

Figure 1. Visualizing the PMF of the DPITL distribution.

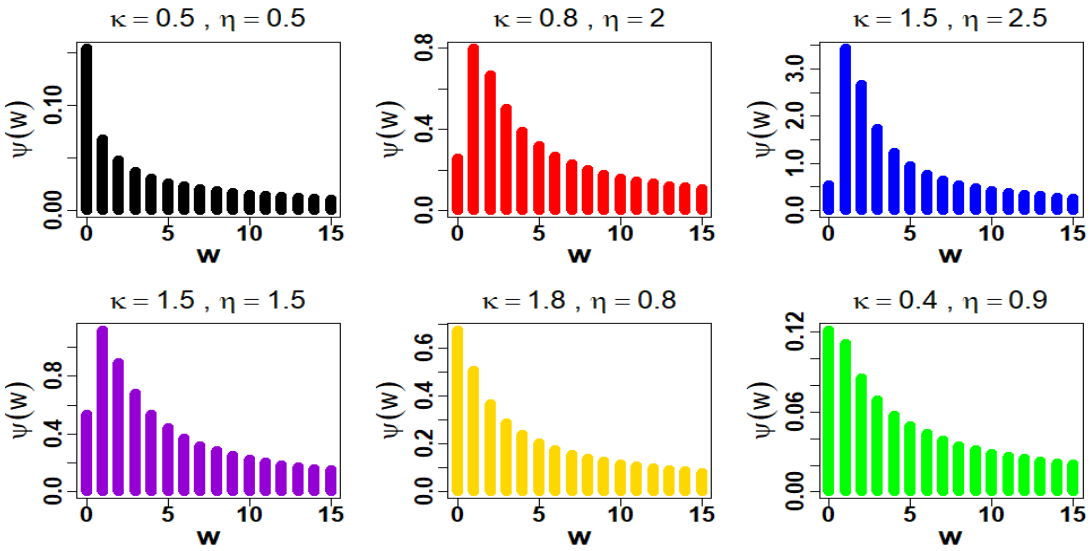

Figure 2. Visualizing the HRF of the DPITL distribution.

3. Some structural properties

This section assesses several statistical properties of the proposed model, each of which is derived mathematically. In the following paragraphs, the statistical features and their mathematical derivations will be discussed.

3.1. Quantile function

According to Rohatgi and Saleh [29], the qth quantile function of a discrete random variable W, say w_q, satisfies the following conditions:

$$P(W \leq w_q) \geq q \text{ and } P(W \geq w_q) \geq 1 - q.$$

This can be expressed as follows using the CDF $\beta(w)$ as:

$$\beta(w_q - 1) < q \leq \beta(w_q).$$

One can find the DPITL distribution's qth quantile by:

$$w_{(q)} = \left(1 - (1-q)^{1/\kappa} + \sqrt{(1-(1-q)^{1/\kappa})^2 - (1-(1-q)^{1/\kappa})}\right)^{1/\eta} - 1. \tag{5}$$

Based on Eq (5), the median is given by:

$$w_{(0.5)} = \left(1 - (0.5)^{1/\kappa} + \sqrt{(1-(0.5)^{1/\kappa})^2 - (1-(0.5)^{1/\kappa})}\right)^{1/\eta} - 1.$$

Similarly, the first quartile ($w_{(0.25)}$) and third quartile ($w_{(0.75)}$) can be produced by setting q = 0.25 and 0.75 in Eq (5). From a computational perspective, formula (5) can be readily adapted by using the inverse transform method to produce pseudo-random data from the DPITL distribution. Specifically, a pseudo-random value from the DPITL distribution will be obtained by evaluating w_q with a pseudo-random value q taken from the uniform distribution (0,1).

3.2. Probability-generation function

Here, the PGM of the DPITL model is obtained. Given the set of nonnegative integers {0, 1, 2,... }, let W be a discrete random variable. The PGF of the DPITL distribution is given as follows:

$$G_W(t) = \sum_{w=0}^{\infty} t^w \rho(W = w; \varsigma) = 1 + (t-1)\sum_{w=1}^{\infty} t^{w-1} S(w; \varsigma)$$

$$= 1 + (t-1)\sum_{w=1}^{\infty} t^{w-1} \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}. \tag{6}$$

Furthermore, the factorial moments of the DPITL distribution can be determined from the PGF by using the following relationship: $\mu'_m(t) = \frac{d^m}{dt^m} G_W(t)|_{t=1}$. Hence, the mean of the DPITL distribution is the

first factorial moment, which can be determined as follows:

$$G'_W(t) = \sum_{w=1}^{\infty} t^{w-1} \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}} + (t-1) \sum_{w=1}^{\infty} (w-1) t^{w-2} \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}.$$

Hence, setting t =1 in the previous equation, the mean of distribution is given by:

$$\mu'_1 = E(W) = G'_W(1) = \sum_{w=1}^{\infty} \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}.$$

The second differentiation $G''_W(t)$ at t =1 provides the second factorial moments

$$G''_W(1) = \sum_{w=1}^{\infty} (w-1) \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}.$$

The $(G'''(w))$ and fourth $(G''''(w))$ differentiations differentiations at t =1 provide the third and fourth factorial moments

$$G'''(1) = \sum_{w=1}^{\infty} (w-1)(w-2) \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}.$$

and,

$$G''''(1) = \sum_{w=1}^{\infty} (w-1)(w-2)(w-3) \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}.$$

3.3. Non-central moments

Here, the non-central moments of DPITL model are provided based on MGF. It is worth noting that MGF can take a similar formula of Eq (6) by replacing it with e^t (i.e., $t = e^t$)

$$M_W(t) = 1 + (e^t - 1) \sum_{w=1}^{\infty} (e^t)^{w-1} \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}. \quad (7)$$

Differentiating (7) with respect to t and setting t = 0, we determine the mean of the DPITL distribution

$$\frac{dM_W(t)}{dt} = (e^t - 1) \sum_{w=1}^{\infty} (w-1)(e^t)^{w-2} \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}} + \sum_{w=1}^{\infty} (e^t)^{w} \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}. \quad (8)$$

By setting t=0 in (8), the first moment of the DPITL distribution is obtained as follows:

$$\mu'_1 = E(W) = \frac{dM_W(t)}{dt} | t = 0 = \sum_{w=1}^{\infty} \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}. \quad (9)$$

The second derivative of (9) with respect to t, when evaluated at t=0, provides the second non-central moment of the DPITL distribution

$$E(W^2) = \frac{d^2 M_W(t)}{dt^2} | t = 0 = \sum_{w=1}^{\infty} (2w-1) \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}. \quad (10)$$

The variance of the DPITL distribution is given by

$$\sigma^2 = E(W^2) - (E(W))^2 = \sum_{w=1}^{\infty}(2w-1)\frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}} - \left[\sum_{w=1}^{\infty}\frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}\right]^2.$$

The third non-central moment of the DPITL distribution is obtained by getting the third derivative of (7) with respect to t and setting t=0:

$$E(W^3) = \frac{d^3M_W(t)}{dt^3}|t=0 = \sum_{w=1}^{\infty}(3w^2-3w+1)\frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}.$$

The fourth non-central moment of the DPITL distribution is obtained by getting the fourth derivative of (7) with respect to t and setting t=0:

$$E(W^4) = \frac{d^4M_W(t)}{dt^4}|t=0 = \sum_{w=1}^{\infty}(4w^3-6w^2+4w-1)\frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}.$$

According to the aforementioned moments, the measure of skewness (λ_1) is defined by $\lambda_1 = \frac{E(W-\mu_1')^3}{\sigma^3}$, kurtosis ($\lambda_2$) is given by $\lambda_2 = \frac{E(W-\mu_1')^4}{\sigma^4}$, and coefficient of variation is given by $CV = \left(\frac{\sigma}{\mu}\right)\times 100$, where σ is the standard deviation and dispersion index formulated as $ID = \frac{\sigma^2}{\mu_1'}$. Table 1 presents a comprehensive statistical summary of the DPITL distribution. It lists the μ_1', σ^2, λ_1, λ_2, CV, and ID key properties that change when one changes the parameter values to help understand the distribution's behavior and for evaluating its performance under different conditions.

Table 1. Statistical exploration of DPITL distribution under varying parameters.

κ	η	μ'_1	σ^2	λ_1	λ_2	CV	ID
	0.5	1.35932	5.59691	1.97402	6.14751	1.74041	4.11744
1.5	1.5	1.28721	2.58857	2.17410	9.08805	1.24991	2.01099
	3	0.77279	0.50632	1.48094	10.60068	0.92077	0.65518
	0.5	1.17580	4.53568	2.22520	7.57907	1.81129	3.85752
2.5	1.5	0.71045	0.97402	2.48998	13.90659	1.38916	1.37100
	3	0.50883	0.29685	0.47918	2.71612	1.07077	0.58340

From Table 1, the following key observations are revealed:

- As κ increases, holding η constant, the values for μ_1', σ^2, and ID decrease, while those for CV increases.
- As η increases, holding κ constant, the values of μ_1', σ^2, CV, and ID decrease.
- The DPITL distribution is exceptionally versatile, offering robust modeling capability for both under-dispersed ($\sigma^2 < \mu_1'$) and over-dispersed ($\sigma^2 > \mu_1'$) count data across its full parameter space.
- Parameter selection allows the DPITL distribution to achieve flexible dispersion behavior.

- Modeling a lower dispersion ($ID < 1$) suggests suitability for modeling failures and a high degree of predictability.
- Lifetime data with high variability is best suited for reliability models that account for higher dispersion ($ID > 1$).
- Positive λ_1 values indicate that this distribution is right-skewed, and as the skewness value approaches zero, the PMF becomes more symmetric.
- Elevated kurtosis, or leptokurtic distributions, indicates greater tail risk and a higher probability of extreme outliers compared to normal distribution. Conversely, low kurtosis, or platykurtic distributions, suggests that values are more tightly clustered around the mean, resulting in a lower probability of extreme outcomes.

Figure 3 visually represents these patterns through 3D plots of μ_1', σ^2, λ_1, λ_2, CV, and ID.

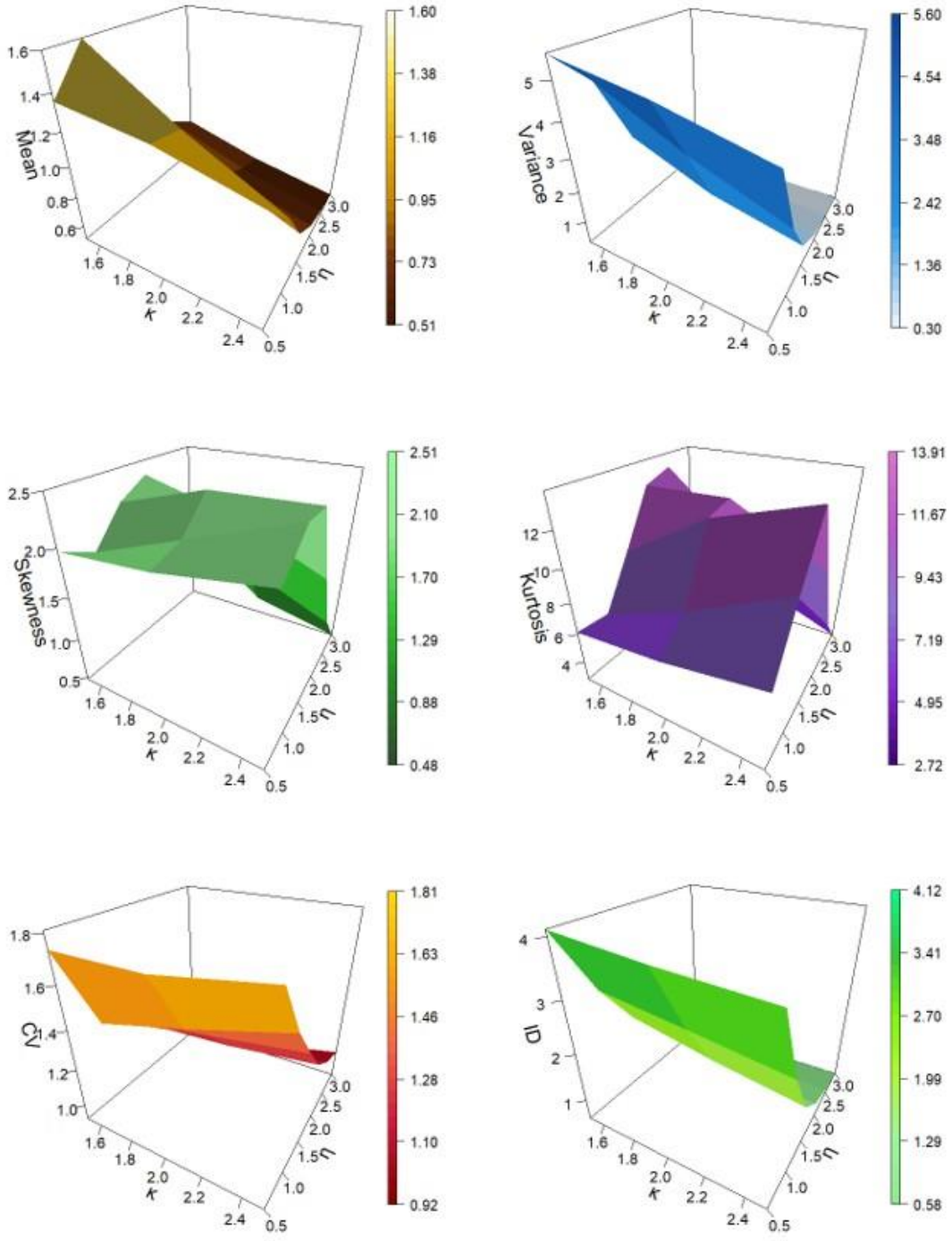

Figure 3. 3D visualization of DPITL distribution properties.

Consistent with the numerical data presented in Table 1, Figure 3 effectively illustrates the flexibility of the DPITL distribution. In this visual representation, the (X, Y) axes display the actual values of the parameters used in the numerical data, while the (Z) axis represents the value of the measure being studied; additionally, the accompanying color key confirms the precise numerical output for the

surface. This comprehensive visual evidence confirms that its constantly positive skewness and the potential for high and low kurtosis underscore the versatility of its PMF and HRF in modeling diverse data shapes. The mean residual life function for a discrete random variable W with a SF is expressed as follows:

$$M(l;\varsigma) = E(W - l|W \geq l) = \frac{1}{S(l-1;\varsigma)}\sum_{i=l+1}^{\infty} 1 - S(i-1;\varsigma). \tag{11}$$

Based on (11), the following represents the mean residual life at w of a random variable W with the DPITL distribution

$$M(l;\varsigma) = \frac{(1+2l^{\eta})^{\kappa}}{(1+l^{\eta})^{2\kappa}}\sum_{i=l+1}^{\infty}\frac{(1+2i^{\eta})^{\kappa}}{(1+i^{\eta})^{2\kappa}}.$$

3.4. Entropy measures

The degree of uncertainty attached to a random variable is known as entropy. It is widely used in a variety of disciplines, including computer science, survival analysis, econometrics, quantum information, and information theory (see Rényi [30]). It can be stated as

$$\vartheta_b = \frac{1}{1-b} log\{\sum_{w=0}^{\infty}\rho^b(W=w;\varsigma)\}, b \neq 1, b > 0, \tag{12}$$

where the W~DPITL(ς) distribution can be derived by substituting PMF (2) in (12) as follows:

$$\vartheta_b = \frac{1}{1-b} log\left\{\sum_{w=0}^{\infty}\left[\frac{(1+2w^{\eta})^{\kappa}}{(1+w^{\eta})^{2\kappa}} - \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}\right]^b\right\}.$$

The Tsallis entropy presented by [31] for the DPITL distribution has the following expression:

$$T_b = \frac{1}{b-1}\{1 - \sum_{w=0}^{\infty}\rho^b(W=w;\varsigma)\}, b \neq 1, b > 0. \tag{13}$$

The Tsallis entropy of the DPITL distribution can be derived by substituting PMF (2) in (13) as follows:

$$T_b = \frac{1}{b-1}\left\{1 - \sum_{w=0}^{\infty}\left[\frac{(1+2w^{\eta})^{\kappa}}{(1+w^{\eta})^{2\kappa}} - \frac{(1+2(w+1)^{\eta})^{\kappa}}{(1+(w+1)^{\eta})^{2\kappa}}\right]^b\right\}.$$

The entropy of the DPITL distribution is challenging to obtain in a closed form using simple computations but can be used to numerically solve complicated equations using a range of parameter values, as shown in Table 2.

The entropy values for the DPITL distribution shown in Table 2 are detailed as follows:

- ϑ_b obtains its smallest value with respect to T_b when b = 0.5. On the other hand, T_b has a lower value than ϑ_b when b = 1.5.
- As κ increases while b and η remain constant, entropy measurements decrease.
- As η increases while κ and b remain constant, entropy measurements decrease.
- As b increases while κ and η remain constant, entropy measurements decrease.

Table 2. Entropy measurements for DPITL distribution.

Parameters			Measure	
k	η	b	ϑ_b	T_b
0.8	2	0.5	3.23794	8.09579
	5		1.29933	1.82981
1.8	2		1.60341	2.45867
	5		0.81092	0.99998
2.5	2		1.24812	1.73297
	5		0.73738	0.89168
0.8	2	1.5	1.69625	1.14357
	5		0.73630	0.61598
1.8	2		1.06104	0.82340
	5		0.68466	0.57977
2.5	2		0.87374	0.70789
	5		0.69516	0.58721

The glyph designs in Figure 4 are used to visually represent the uncertainty of measurements for the DPITL distribution.

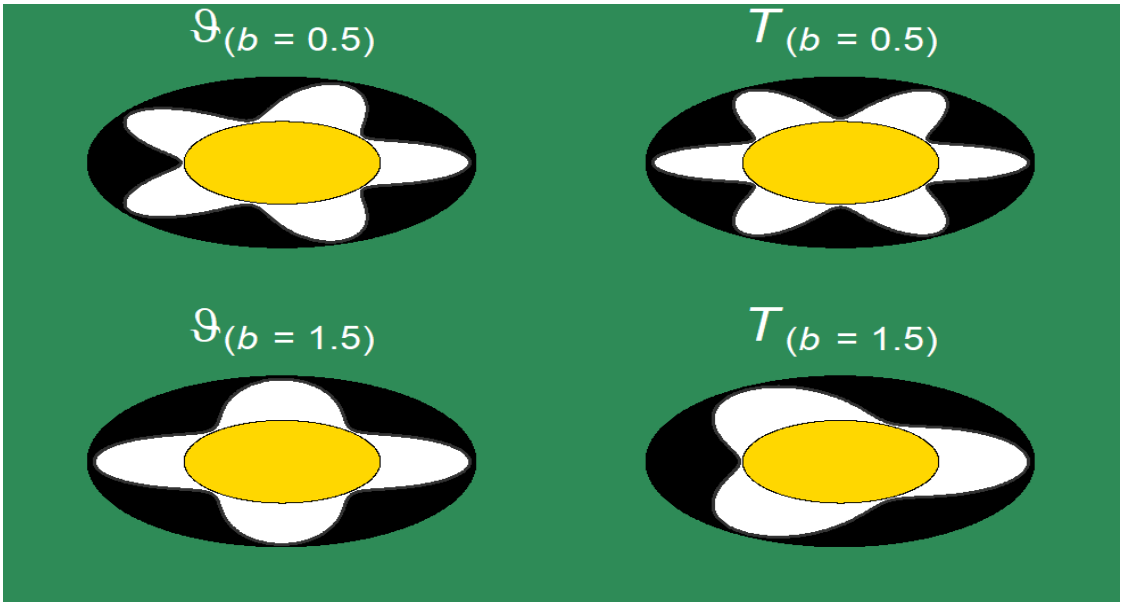

Figure 4. Visual entropy of glyph designs for DPITL distribution.

Figure 4 illustrates the use of two colors in the glyphs to convey information: the yellow center represents the average entropy, while the surrounding white area reflects the degree of uncertainty. A smaller white area indicates lower uncertainty and thus better performance. From this visualization, performance depends on the value of b. When b = 0.5, measure ϑ_b performs better with smaller white areas, whereas at b =1.5, measure T_b shows superior results with reduced uncertainty. These graphical insights align perfectly with the numerical results reported in Table 2.

4. Estimation methods

To estimate the unknown parameters of DPITL distribution, five methods are employed: ML, MM, LS, CM, and AD.

4.1. Maximum likelihood estimation

Given a random sample w_1, w_2,…,w_n drawn independently from the DPITL distribution, the likelihood function can be expressed as:

$$L(\varsigma) = \prod_{j=1}^{n}\left[\frac{(1+2w_j^{\eta})^{\kappa}}{(1+w_j^{\eta})^{2\kappa}} - \frac{(1+2(w_j+1)^{\eta})^{\kappa}}{(1+(w_j+1)^{\eta})^{2\kappa}}\right]. \tag{14}$$

We start by determining the log-likelihood function in order to get the ML estimator of $\hat{\kappa}$ of κ and $\hat{\eta}$ of η. This is accomplished by taking Eq (14)'s natural logarithm, denoted by $log\,\ell$, which produces the following expression:

$$log\,\ell = \sum_{j=1}^{n} log\big[A(w_j,\varsigma) - B(w_j+1,\varsigma)\big], \tag{15}$$

where $A(w_j,\varsigma) = \frac{(1+2w_j^{\eta})^{\kappa}}{(1+w_j^{\eta})^{2\kappa}}$ and $B(w_j+1,\varsigma) = \frac{(1+2(w_j+1)^{\eta})^{\kappa}}{(1+(w_j+1)^{\eta})^{2\kappa}}$.

The first partial derivatives of (15), with respect to the κ and η are

$$\frac{\partial\, log\,\ell}{\partial\kappa} = \sum_{j=1}^{n}\frac{A_{\kappa}^{'}(w_j,\varsigma) - B_{\kappa}^{'}(w_j+1,\varsigma)}{\big[A(w_j,\varsigma) - B(w_j+1,\varsigma)\big]}, \tag{16}$$

$$\frac{\partial\, log\,\ell}{\partial\eta} = \sum_{j=1}^{n}\frac{A_{\eta}^{'}(w_j,\varsigma) - B_{\eta}^{'}(w_j+1,\varsigma)}{\big[A(w_j,\varsigma) - B(w_j+1,\varsigma)\big]}, \tag{17}$$

where

$$A_{\kappa}^{'}(w_j,\varsigma) = \frac{(1+2w_j^{\eta})^{\kappa}\big[ln(1+2w_j^{\eta}) - 2\,ln(1+w_j^{\eta})\big]}{(1+w_j^{\eta})^{2\kappa}}, \tag{18}$$

and

$$A_{\eta}^{'}(w_j,\varsigma) = \frac{2\kappa w_j^{\eta}\,ln\,w_j\big[(1+w_j^{\eta})^{2\kappa}(1+2w_j^{\eta})^{\kappa-1} - 2\kappa(1+2w_j^{\eta})^{\kappa}(1+w_j^{\eta})^{2\kappa-1}\big]}{(1+w_j^{\eta})^{4\kappa}}. \tag{19}$$

Note that $B_{\kappa}'(w_j,\varsigma)$ and $B_{\eta}'\big(w_j+1,\varsigma\big)$ have the same expressions as provided in (18) and (19) by replacing w_j with w_j+1. The ML estimators $\hat{\kappa}$ and $\hat{\eta}$ are obtained after setting the right-hand side of Eqs (16) and (17) to zero. Since there is no analytical solution for this system, a numerical approach such as Nelder-Mead is required to solve them numerically.

4.2. Moments method

Here, the moment estimators $\hat{\kappa}_2$ of κ and $\hat{\eta}_2$ of η are discussed. In order to apply this method, we usually equate the population moments to the corresponding sample moments and solve the resulting equations of moments to obtain $\hat{\kappa}_2$ and $\hat{\eta}_2$. To be more precise, we would equate the population moments $E(W)$ and $E(W^2)$ (from Eqs (9) and (10), respectively) to the sample moments $m_1 = \frac{1}{n}\sum_{j=1}^{n} w_j$ and $m_2 = \frac{1}{n}\sum_{j=1}^{n} w_j^2$. Then, we would solve these two equations for κ and η simultaneously. It can be challenging to find a direct numerical solution for this system of two nonlinear equations in two unknowns; instead, Khan et al. [32] presented a method of pseudo-moment by minimizing the following function with respect to κ and η.

$$H(\kappa, \eta) = (m_1 - E(W))^2 + (m_2 - E(W^2))^2.$$

For instance, optimization functions such as *optim* can be used in the R environment to complete the task.

4.3. Least squares method

Here, the LS estimators $\hat{\kappa}_3$ of κ and $\hat{\eta}_3$ of η are determined. The LS estimators $\hat{\kappa}_3$ and $\hat{\eta}_3$ are obtained by minimizing the following function:

$$ls(\varsigma) = \sum_{i=1}^{n}\left[\beta(w_{(i)};\varsigma) - \frac{i}{n+1}\right]^2,$$

where $\beta(w_{(i)};\varsigma)$ denotes the CDF of the DPITL distribution evaluated at the i-th order statistic. Alternatively, the LS estimators $\hat{\kappa}_3$ and $\hat{\eta}_3$ can be acquired by working out the following equations, with respect to κ and η:

$$\frac{\partial ls(\varsigma)}{\partial \kappa} = \sum_{i=1}^{n}\left[\beta(w_{(i)};\varsigma) - \frac{i}{n+1}\right]\delta_{\kappa}(w_{(i)}|\varsigma) = 0, \tag{20}$$

$$\frac{\partial ls(\varsigma)}{\partial \eta} = \sum_{i=1}^{n}\left[\beta(w_{(i)};\varsigma) - \frac{i}{n+1}\right]\delta_{\eta}(w_{(i)}|\varsigma) = 0, \tag{21}$$

where $\delta_{\kappa}(w_{(i)}|\varsigma), \delta_{\eta}(w_{(i)}|\varsigma)$ have the same expressions of Eqs (18) and (19) with negative sign and the ordered sample $w_{(i)}$. Since the solution to Eqs (20) and (21) cannot be found in closed form, numerical optimization techniques can be used to obtain $\hat{\kappa}_3$ and $\hat{\eta}_3$.

4.4. Cramér-von Mises method

Here, the CM estimators $\hat{\kappa}_4$ of κ and $\hat{\eta}_4$ of η are obtained. The CM estimators $\hat{\kappa}_4$ and $\hat{\eta}_4$are obtained by minimizing the following function:

$$C(\varsigma) = \frac{1}{12n} + \sum_{i=1}^{n}\left[\beta(w_{(i)};\varsigma) - \frac{2i-1}{2n}\right]^2. \tag{22}$$

The CM estimators $\hat{\kappa}_4$ and $\hat{\eta}_4$ can be obtained by solving the following equations, with respect to κ and η, as an alternative to Eq (22).

$$\frac{\partial C(\varsigma)}{\partial \kappa} = \sum_{i=1}^{n}\left[\beta(w_{(i)};\varsigma) - \frac{2i-1}{2n}\right]\delta_{\kappa}(w_{(i)}|\varsigma) = 0, \tag{23}$$

$$\frac{\partial C(\varsigma)}{\partial \eta} = \sum_{i=1}^{n}\left[\beta(w_{(i)};\varsigma) - \frac{2i-1}{2n}\right]\delta_{\eta}(w_{(i)}|\varsigma) = 0, \tag{24}$$

where $\delta_{\kappa}(w_{(i)}|\varsigma), \delta_{\eta}(w_{(i)}|\varsigma)$ have the same expressions of Eqs (18) and (19) with negative sign and the ordered sample $w_{(i)}$. Since the solution to Eqs (23) and (24) cannot be found in closed form, numerical optimization techniques can be used to obtain $\hat{\kappa}_4$ and $\hat{\eta}_4$.

4.5. Anderson darling method

The AD estimators $\hat{\kappa}_5$ of κ and $\hat{\eta}_5$ of η are obtained by minimizing the following function:

$$A(\varsigma) = -n - \frac{1}{n} + \sum_{i=1}^{n}(2i-1)\left[log\,\beta\,(w_{(i)};\varsigma) + log\left[1-\beta(w_{(n-i+1)};\varsigma)\right]\right]. \tag{25}$$

The AD estimators $\hat{\kappa}_5$ and $\hat{\eta}_5$ can be obtained by solving the following equations, with respect to $\kappa$and η, as an alternative to Eq (25):

$$\frac{\partial A(\varsigma)}{\partial \kappa} = \sum_{i=1}^{n}\frac{(2i-1)\delta_k(w_{(i)}|\varsigma)}{\beta(w_{(i)};\varsigma)} - \frac{(2i-1)\delta_k(w_{(n-i+1)}|\varsigma)}{1-\beta(w_{(n-i+1)};\varsigma)}, \tag{26}$$

$$\frac{\partial A(\varsigma)}{\partial \eta} = \sum_{i=1}^{n}\frac{(2i-1)\delta_\eta(w_{(i)}|\varsigma)}{\beta(w_{(i)};\varsigma)} - \frac{(2i-1)\delta_\eta(w_{(n-i+1)}|\varsigma)}{1-\beta(w_{(n-i+1)};\varsigma)}, \tag{27}$$

Where $\delta_k(w_{(i)}|\varsigma)$, $\delta_\eta(w_{(i)}|\varsigma)$, $\delta_k(w_{(n-i+1)}|\varsigma)$, $\delta_\eta(w_{(n-i+1)}|\varsigma)$ have the same expressions of Eqs (18) and (19) by replacing the random sample with the corresponding ordered sample. Since the solution to Eqs (26) and (27) cannot be found in closed form, numerical optimization techniques can be used to obtain $\hat{\kappa}_5$ and $\hat{\eta}_5$.

5. Numerical simulation

This section presents the performance of ML, MM, LS, CM, and AD methods for estimating the parameters of DPITL distribution. A simulation study is conducted under different sample sizes and selected parameter values. The steps of the simulation algorithm are outlined as follows:

1) Set the number of replications (N) to 1000.
2) Choose the sample size n as (15,25,50,100).
3) Select parameters as Set1=$(\kappa = 0.9, \eta = 1.2)$, Set2=$(\kappa = 1.8, \eta = 0.5)$, and Set3 =$(\kappa = 1.3, \eta = 1.3)$.
4) Generate u_i for i =1, 2 ,…, n as independent observations from the uniform distribution $U(0,1)$.

5) Random samples are then drawn from the DPITL distribution based on Eq (5).
6) Estimate the DPITL distribution parameters using the chosen methods.
7) Repeat steps 4–6 N times.
8) Evaluate the accuracy of the estimators by calculating the mean squared error (MSE) and the mean bias (MB) of the estimated parameters.

R version 4.5.0 was used for all computations, and the *bbmle* and *MASS* packages were utilized to estimate the parameters. Table 3 and Figure 5 show the findings. The results of the simulation offer important new information on how well the estimators work in various scenarios. Below is an overview of the key findings:

- Increasing n makes low MSEs and thus various estimates more accurate, as shown in Figure 5.
- When parameters κ or η get higher values, results are less accurate and there is a bigger MSE, as shown in Figure 5 and Table 3.
- For the η parameter, the optimal estimation method is ML, while for the parameter κ, the AD method is preferred, as shown in Figure 5 and Table 3.
- As Figure 5 shows, the estimation of parameter η was more accurate in Set 2, while parameter κ was estimated more accurately in Set 1.
- The CM method demonstrates superior performance for κ, outperforming both ML and MM methods in most cases, as shown in Table 3.
- The AD method demonstrates superior performance for η, outperforming both CM and LS methods in most cases, as shown in Table 3.

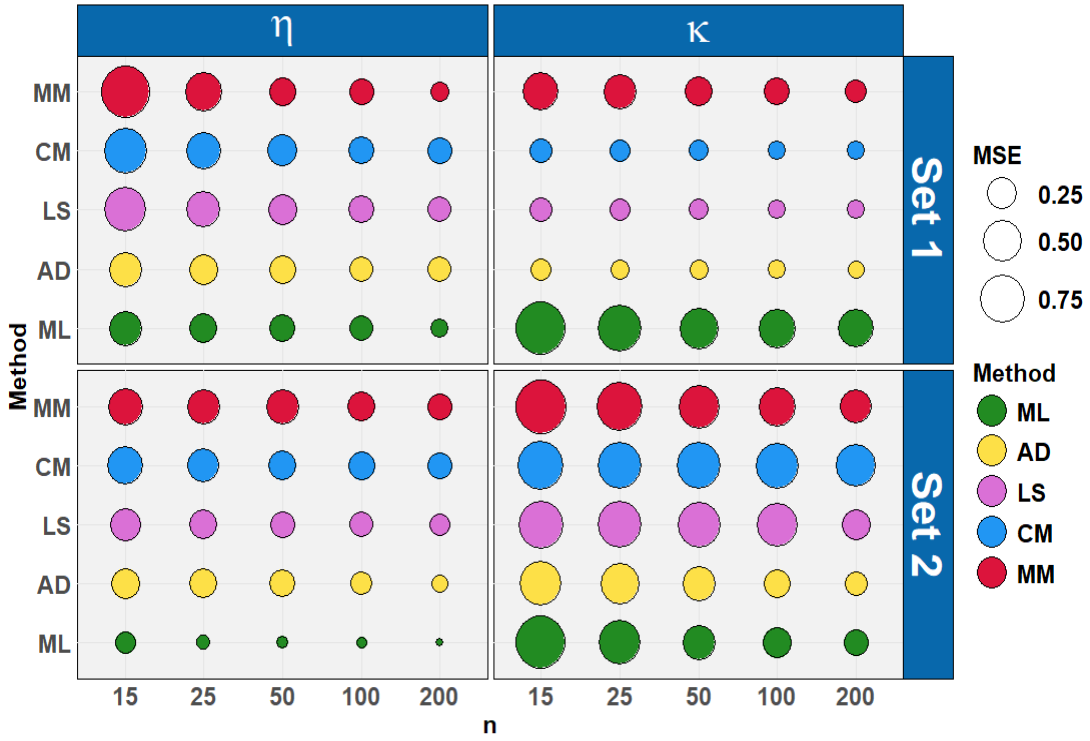

Figure 5. Bubble plot MSE performance for DPITL distribution methods.

Table 3. Numerical simulation results of DPITL distribution with MBs and MSEs for different parameter values.

n	True value	Set1				Set2				Set3			
	Estimate	$\hat{\kappa}$		$\hat{\eta}$		$\hat{\kappa}$		$\hat{\eta}$		$\hat{\kappa}$		$\hat{\eta}$	
	Methods	MB	MSE	MB	MSE	MB	MSE	MB	MSE	MB	MSE	MB	MSE
15	ML	0.67251	0.93345	0.08445	0.34531	0.42206	0.92955	0.00061	0.09481	0.98479	0.92358	0.31364	0.26987
	AD	0.12353	0.09059	0.30157	0.34534	0.74087	0.62627	0.39947	0.24561	0.25817	0.16431	0.34487	0.36204
	LS	0.20218	0.11652	0.51952	0.63224	0.80793	0.74491	0.43141	0.27883	0.32685	0.21057	0.53182	0.49971
	CM	0.20861	0.11801	0.55917	0.65202	0.82487	0.76199	0.53639	0.40521	0.33414	0.21460	0.57836	0.52204
	MM	0.18736	0.42261	0.35193	0.90917	0.31935	0.99356	0.58872	0.38010	0.19414	0.93862	0.73131	0.92279
25	ML	0.65943	0.69091	0.35475	0.21212	0.40634	0.61817	0.03548	0.01574	0.95730	0.79578	0.42705	0.10989
	AD	0.14988	0.06717	0.29734	0.23021	0.76035	0.52670	0.40756	0.21368	0.27971	0.13055	0.36232	0.21304
	LS	0.20089	0.08476	0.42437	0.35851	0.80069	0.69913	0.39860	0.20323	0.32264	0.15921	0.46264	0.32494
	CM	0.20444	0.08671	0.45554	0.39324	0.81226	0.71058	0.48699	0.29573	0.32768	0.16210	0.49967	0.34797
	MM	0.21443	0.33078	0.07360	0.44560	0.36767	0.76107	0.56701	0.33567	0.25988	0.91579	0.27833	0.82626
50	ML	0.61549	0.51337	0.39087	0.18067	0.37187	0.33685	0.05567	0.00794	0.97811	0.48627	0.44875	0.04079
	AD	0.18663	0.05953	0.32547	0.18339	0.78308	0.33510	0.40127	0.18113	0.32169	0.12075	0.40134	0.20268
	LS	0.21466	0.07049	0.38699	0.23798	0.80188	0.66902	0.36587	0.14881	0.34393	0.14676	0.44960	0.25530
	CM	0.21536	0.07125	0.40634	0.25471	0.80971	0.67811	0.43923	0.21529	0.34649	0.14798	0.47056	0.26720
	MM	0.29080	0.21056	0.16245	0.20140	0.38454	0.57643	0.56531	0.33074	0.36862	0.83306	0.00740	0.65072
100	ML	0.62258	0.45430	0.41420	0.13753	0.34390	0.22619	0.05633	0.00566	0.99405	0.12629	0.47469	0.01626
	AD	0.19367	0.05016	0.32783	0.14381	0.79972	0.20163	0.40910	0.10699	0.32318	0.10332	0.41059	0.18856
	LS	0.20963	0.05636	0.35760	0.16702	0.80918	0.56899	0.36738	0.14251	0.33421	0.11850	0.42770	0.20809
	CM	0.20828	0.05612	0.36742	0.17430	0.81316	0.67347	0.42790	0.19324	0.33574	0.12696	0.44496	0.21883
	MM	0.37944	0.17321	0.23149	0.14901	0.37507	0.46047	0.55457	0.22054	0.64526	0.51676	0.32841	0.38747
200	ML	0.61467	0.41040	0.42037	0.04786	0.44887	0.15675	0.09445	0.00094	0.98703	0.09790	0.49024	0.00714
	AD	0.20179	0.04708	0.33513	0.13090	0.78445	0.12152	0.48574	0.04298	0.32914	0.08492	0.41487	0.01499
	LS	0.21181	0.05106	0.35060	0.14262	0.78657	0.22586	0.42933	0.08994	0.33446	0.09868	0.41533	0.18374
	CM	0.20916	0.05013	0.35482	0.14524	0.79107	0.53199	0.49682	0.15404	0.33544	0.09151	0.43193	0.19614
	MM	0.49905	0.09651	0.21485	0.05954	0.43111	0.31059	0.50304	0.15560	0.88097	0.45821	0.58165	0.18550

6. Application of count data

This section demonstrates the effectiveness of DPITL distribution by applying it to a dataset derived from real-world data. Its efficacy is tested in a comprehensive evaluation comparison with other distributions, namely DIKU, DBurr, DIW, DLL, negative binomial (NB), and discrete Lomax (DLO) [33]. Various metrics of goodness-of-fit, such as the negative log-likelihood ($-log\,\ell$), Akaike information criterion (Λ_1), Bayesian information criterion (Λ_2), Hannan-Quinn information criterion (Λ_3), consistent Akaike information criterion (Λ_4), and chi-square (χ^2) with its corresponding p-value (Λ_5), are employed to compare the fitted models. The ML method was employed to parameterize all competing distributions. For DPITL distribution, its parameters were also estimated using the MM, LS, CM, and AD techniques. All parameter estimates were reported along with their respective standard errors (SEs).

6.1. Corticosteroid kidney data

The dataset comprises observations on liver lesions induced by steroid exposure. The study aimed to investigate the impact of these agents on the formation of lesions in rat embryos. [34]. The basic descriptive measurements of corticosteroid kidney data are shown graphically in Figure 6.

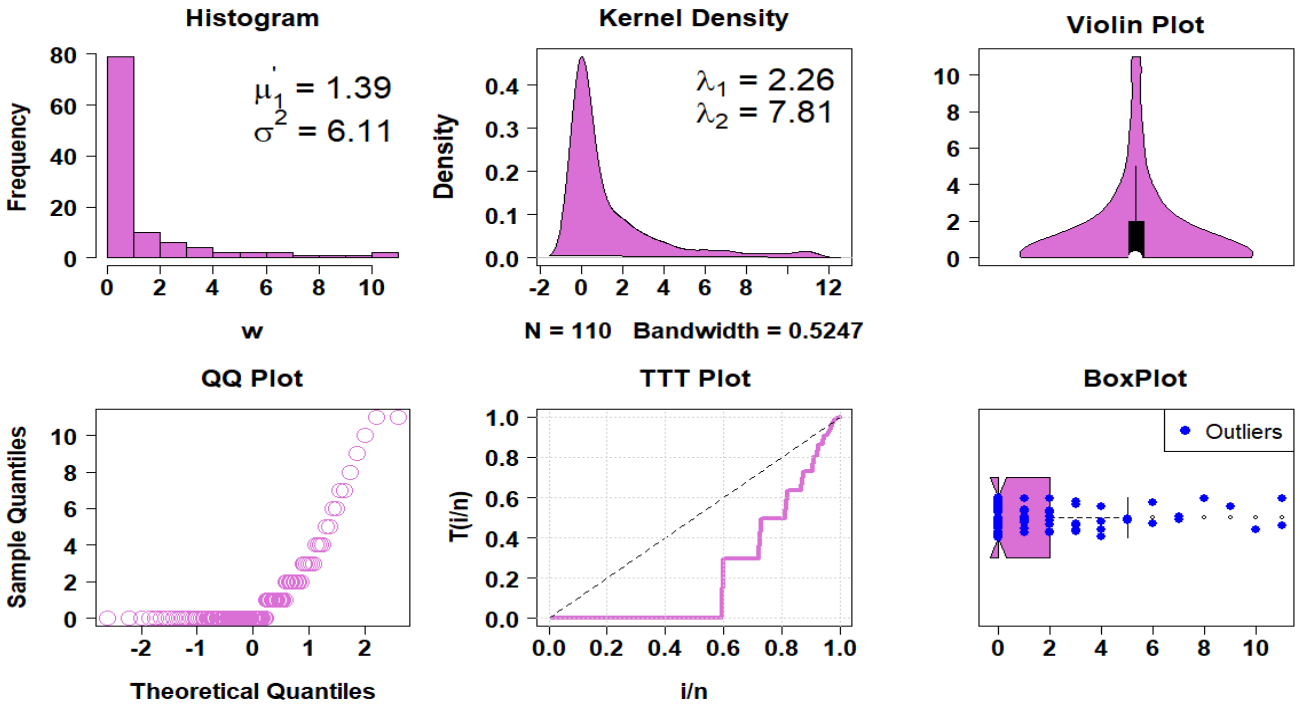

Figure 6. Descriptive visual representation of corticosteroid kidney data.

A look at Figure 6 confirms the DPITL distribution's excellent fit for the data, based on its relationship to the total test time (TTT) plot and HRF. Moreover, the calculation of its parameters confirms that the data are indeed over-dispersed, which is consistent with the characteristics of the DPITL distribution. PMF analysis then reveals the asymmetrical nature of the data, leading us to use quantile-quantile (QQ) plots and box plots to identify any outliers, where the outliers' values stand out

as blue rings. Table 4 provides a comprehensive overview of the fitted discrete distributions and their respective goodness-of-fit measures, including $-log\,\ell$, Λ_1, Λ_2, Λ_3, Λ_4, χ^2, and Λ_5.

Table 4. Observed vs. expected frequencies and model fit for corticosteroid kidney data.

w	Observed	Expected						
	Distribution	DPITL	DIKU	DBurr	DIW	DLL	DLO	NB
0	65	64.24978	64.00033	64.73913	63.96051	63.18818	61.54119	64.24497
1	14	19.07262	19.72589	19.18476	20.68863	20.07872	21.03007	16.77790
2	10	8.74196	8.70531	8.48552	8.04508	8.64045	9.69959	9.00464
3	6	4.87816	4.73183	4.63222	4.22855	4.65721	5.28547	5.65981
4	4	3.04729	2.91023	2.86267	2.59534	2.86608	3.20519	3.81736
5	2	2.05058	1.94231	1.91985	1.75079	1.92272	2.09428	2.67987
6	2	1.45520	1.37388	1.36441	1.25875	1.37020	1.44595	1.93055
7	2	1.07488	1.01496	1.01254	0.94749	1.02102	1.04159	1.41608
8	1	0.81925	0.77554	0.77705	0.73835	0.78736	0.77599	1.05265
9	1	0.64039	0.60879	0.61249	0.59119	0.62390	0.59416	0.79055
10	1	0.51111	0.48855	0.49344	0.48377	0.50538	0.46539	0.59855
11	2	3.45880	3.72237	3.91592	4.71155	4.33878	2.82114	2.02707
	Total	110						
	$-log\,\ell$	170.29343	171.04999	171.13916	172.93488	171.71715	170.48061	170.54405
	Λ_1	344.58686	346.09998	346.27832	349.86977	347.43431	344.96122	345.08810
	Λ_2	349.98783	351.50094	351.67928	355.27073	352.83527	350.36218	350.48906
	Λ_3	346.58686	348.09998	348.27832	351.86977	349.43431	346.96122	347.08810
	Λ_4	346.77752	348.29064	348.46898	352.06043	349.62497	347.15188	347.27876
	χ^2	4.42097	5.51000	5.55726	8.28466	6.20214	5.14089	5.35152
	Λ_5	0.88159	0.78778	0.78328	0.50573	0.71953	0.82186	0.79811

Table 4 conclusively demonstrates the superior fit of the DPITL distribution to the data. This is quantified by its consistently lowest values across all fitted metrics and the highest Λ_5 when compared to all other competing distributions, establishing DPITL distribution as the optimal choice. Furthermore, Figure 7 offers a visual confirmation of the estimation quality.

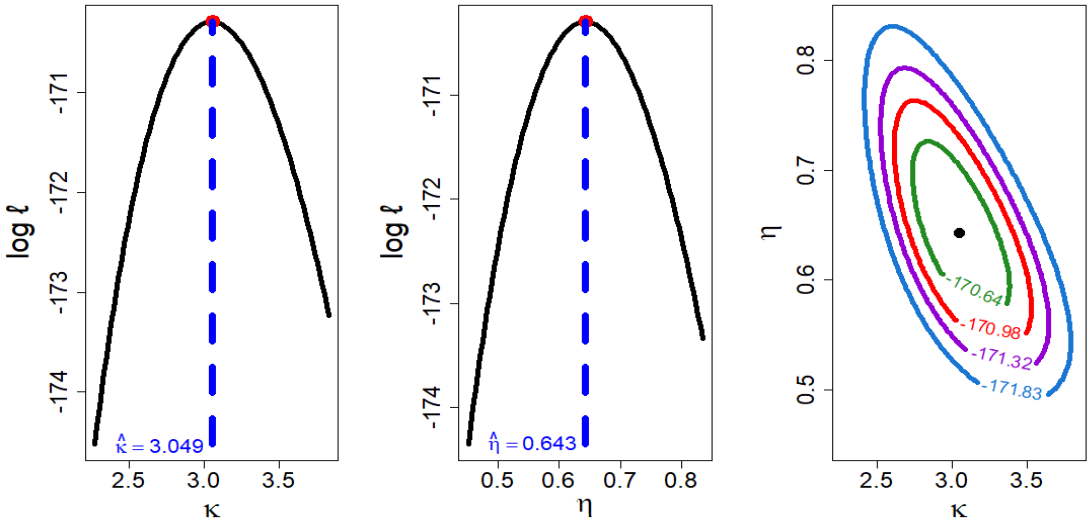

Figure 7. Contour and profile plots for DPITL distribution on corticosteroid kidney data.

Figure 7 utilizes contour plots and profile log-likelihood functions to effectively illustrate the behavior of the estimators and demonstrate how they successfully maximize the fit of the DPITL distribution to the corticosteroid kidney data. A visual representation of the estimated PMFs is shown in Figure 8, demonstrating how the fitted distributions apply to the observed and expected frequencies of the corticosteroid kidney data.

Figure 8. Observed and expected frequencies of fitted PMFs for corticosteroid kidney data**.**

Following the visual confirmation of the fit provided by Figure 8, the precise results of the different estimation methods (MM, LS, CM, AD, ML) and their SEs are quantified in Table 5, with Figure 9 offering a final visual comparison to determine the best method.

Table 5. Estimators and SEs from different distributions for corticosteroid kidney data.

Distribution	DPITL					DIKU	DBurr	DIW	DLL	DLO	NB
Methods	LS	AD	CM	MM	ML						
$\hat{\kappa}$	1.23860	1.22998	1.22818	0.12103	3.04949	1.15826	1.28117	0.55809	1.20633	1.88192	0.32153
$\hat{\eta}$	1.49317	1.48915	1.51802	0.03668	0.64278	1.42086	1.05364	1.04949	0.77984	1.83186	0.18776
SE ($\hat{\kappa}$)	0.43285	0.19714	0.43128	0.56968	0.39074	0.31923	0.16290	0.09733	0.15927	0.64673	0.07419
SE ($\hat{\eta}$)	0.78488	0.30790	0.79620	0.14930	0.09616	0.24068	0.16735	0.14595	0.13570	0.95453	0.04525

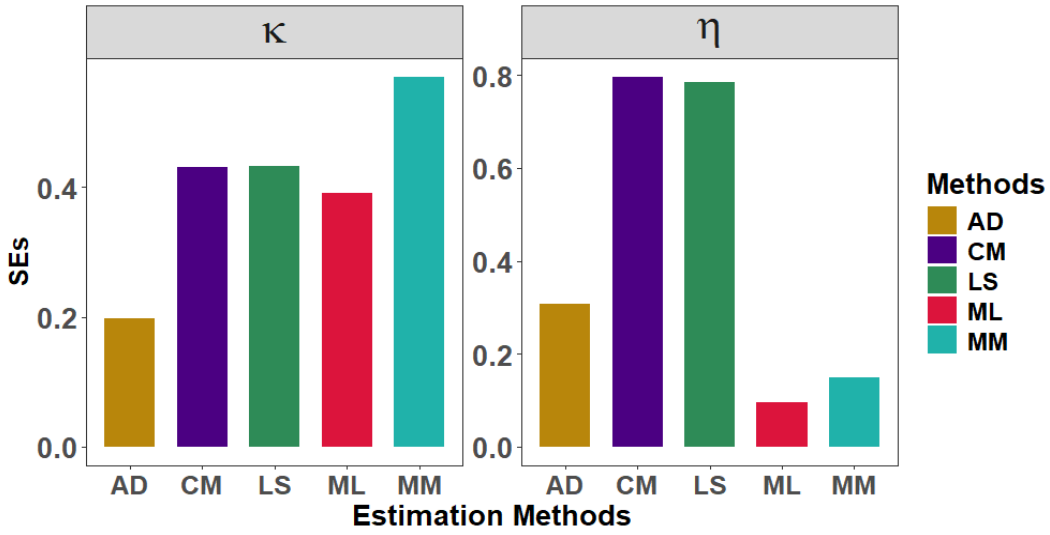

Figure 9. Visual comparison of SEs for DPITL distribution of corticosteroid kidney data.

Consistent with the simulation outcomes, both Table 5 and Figure 9 show that AD is the most reliable method for estimating parameter κ, while ML is the most reliable for parameter η.

6.2. Leukocyte chromatid data

A culture of human leukocytes was treated with 0.2 g of quinone 1 over a 24-hour period. Chinone is known for its genotoxic effects. This exposure was found to cause chromatid aberrations per cell [35]. The basic descriptive measurements of the leukocyte chromatid data are visually represented in Figure 10.

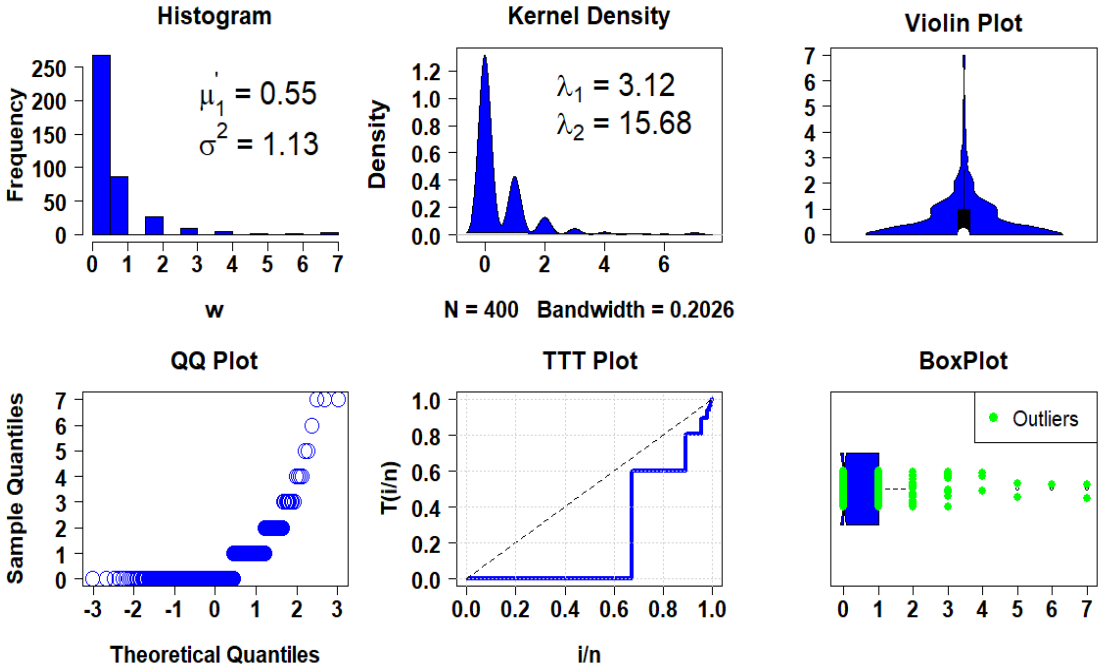

Figure 10. Descriptive visual representation of leukocyte chromatid data.

Table 6. Observed vs. expected frequencies and model fit for leukocyte chromatid data.

w	Observed	Expected						
Distribution		DPITL	DIKU	DBurr	DIW	DLL	DLO	NB
0	268	267.98513	267.75762	268.03832	267.35656	267.05105	270.39691	270.18834
1	87	87.26737	88.73520	88.66910	92.15893	91.10951	80.90346	78.54283
2	26	25.67757	24.53958	24.42941	21.33613	22.78113	28.19664	29.83460
3	9	9.55303	9.09453	8.94648	8.00163	8.38642	11.05799	12.21935
4	4	4.21370	4.09845	4.02150	3.84329	3.91256	4.76436	5.18624
5	2	2.10008	2.10950	2.07998	2.14002	2.11808	2.21652	2.24743
6	1	1.14627	1.19435	1.18893	1.31449	1.26829	1.09920	0.98727
7	3	2.05685	2.47078	2.62628	3.84896	3.37296	1.36492	0.79395
Total		400						
$-log\,\ell$		398.54026	398.76146	398.89533	400.26411	399.60105	398.88132	399.85684
Λ_1		801.08051	801.52293	801.79066	804.52822	803.20209	801.76265	803.71368
Λ_2		809.06344	809.50586	809.77359	812.51115	811.18502	809.74558	811.69661
Λ_3		803.08051	803.52293	803.79066	806.52822	805.20209	803.76265	805.71368
Λ_4		804.24186	804.68427	804.95200	807.68957	806.36343	804.92399	806.87502
χ^2		0.20323	0.27507	0.21911	1.71242	0.79496	3.14624	8.69786
Λ_5		0.99901	0.89908	0.98887	0.88749	0.91739	0.67715	0.12174

The analysis of the leukocyte chromatid data begins with Figure 10, which uses a TTT plot to illustrate the excellent fit of the DPITL distribution. This choice is supported numerically, as the calculation of the parameters confirms the data's over-dispersed nature, aligning with the DPITL characteristics. Because the PMF shows an asymmetrical distribution, QQ plots and box plots were used

to detect outliers, marked by green rings. To evaluate the model's overall performance, Table 6 summarizes the goodness-of-fit metrics, while Figure 11 provides a visual confirmation of how the estimators successfully maximize the fit to the data.

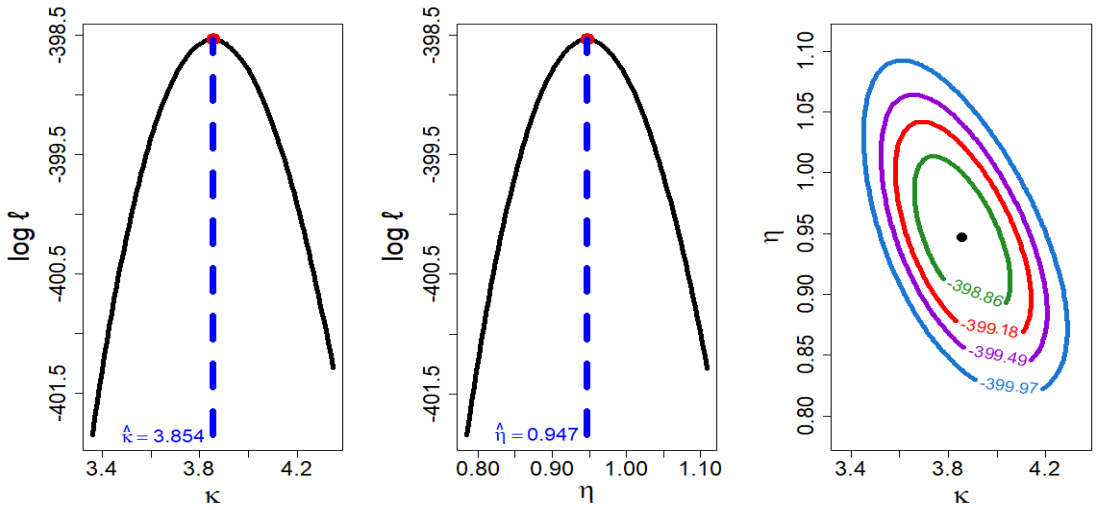

Figure 11. Contour and profile plots for DPITL distribution on leukocyte chromatid data.

The analysis definitively establishes DPITL distribution as the optimal choice for the leukocyte chromatid data. Table 6 quantifies this superiority by showing that DPITL achieves the lowest metrics and the highest Λ_5 among all competitors. This high quality of estimation is visually supported by Figure 11, which confirms how the estimators successfully maximize the fit of the DPITL model. Furthermore, the effectiveness of the chosen distribution is visibly demonstrated in Figure 12, where the estimated PMFs align perfectly with the observed data frequencies.

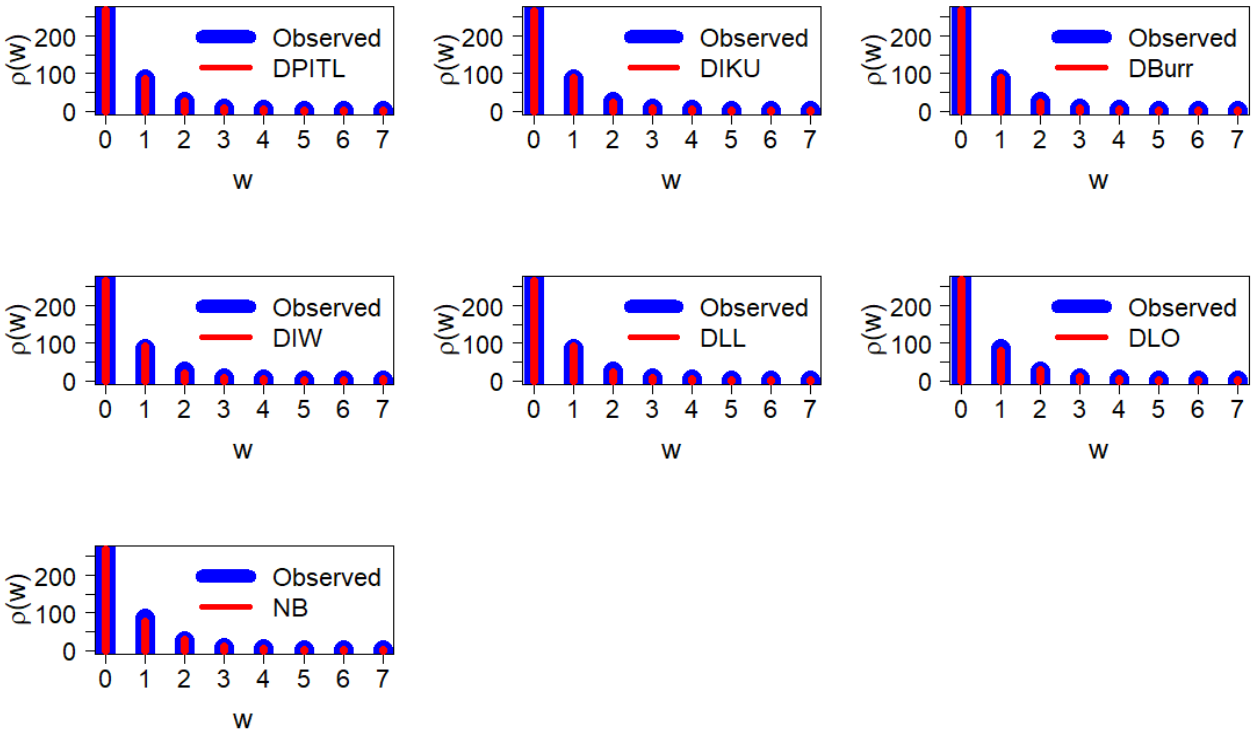

Figure 12. Observed and expected frequencies of fitted PMFs for leukocyte chromatid data.

The visual confirmation of the fit provided by Figure 12 is followed by the quantification of the model parameters and their SEs in Table 7, with Figure 13 offering a final visual comparison of method accuracy.

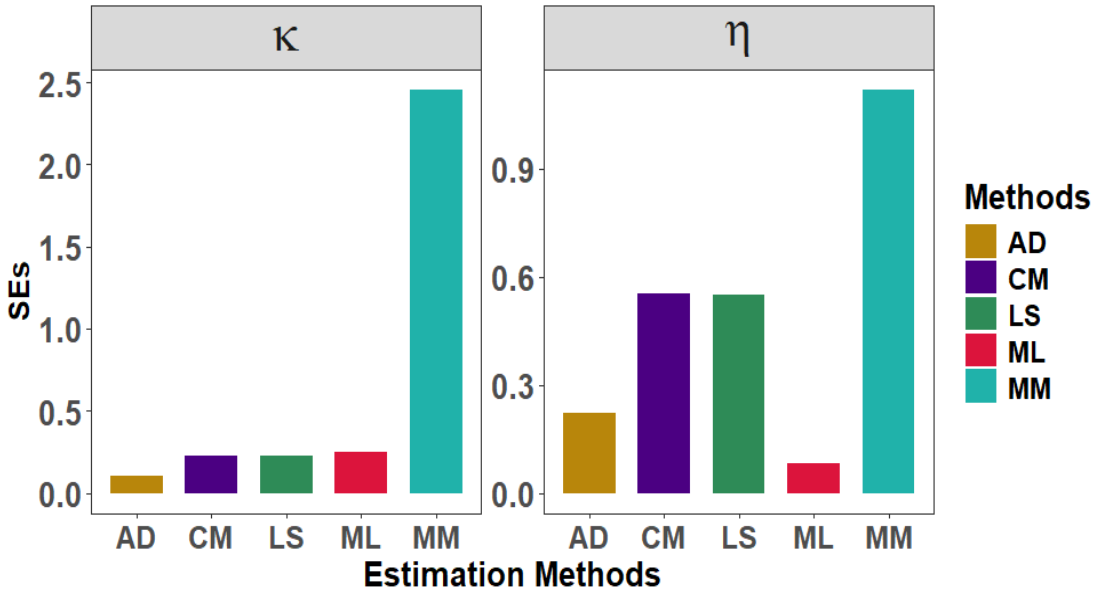

Figure 13. Visual comparison of SEs for DPITL distribution of leukocyte chromatid data.

Table 7. Estimators and SEs from different distributions for leukocyte chromatid data.

Distribution	DPITL					DIKU	DBurr	DIW	DLL	DLO	NB
Methods	LS	AD	CM	MM	ML						
$\hat{\kappa}$	1.41983	1.41740	1.41774	2.98979	3.85406	2.91931	1.59960	0.62218	2.09107	7.45989	0.61981
$\hat{\eta}$	2.08292	2.09990	2.09160	1.29241	0.94687	2.96071	1.59249	1.91624	0.71622	6.13210	0.53099
SE ($\hat{\kappa}$)	0.22568	0.10626	0.22555	2.45210	0.24722	0.67253	0.10265	0.03970	0.17210	3.42905	0.12690
SE ($\hat{\eta}$)	0.55084	0.22346	0.55298	1.11840	0.08117	0.27721	0.14587	0.16498	0.04428	3.22225	0.05597

Table 7 and Figure 13 both agree that the estimation reliability depends on the parameter being studied. The AD method is the most reliable choice for estimating parameter κ. However, the ML method is the most reliable for estimating parameter η. This is consistent with the simulation results.

6.3. Criminology data

This third dataset provides specific details from the field of criminal sociology, featuring a collection of samples of individuals who display deviant behavior. The information was originally gathered and made available to the research community in [36]. A visual representation detailing the core descriptive measurements of this dataset is provided in Figure 14.

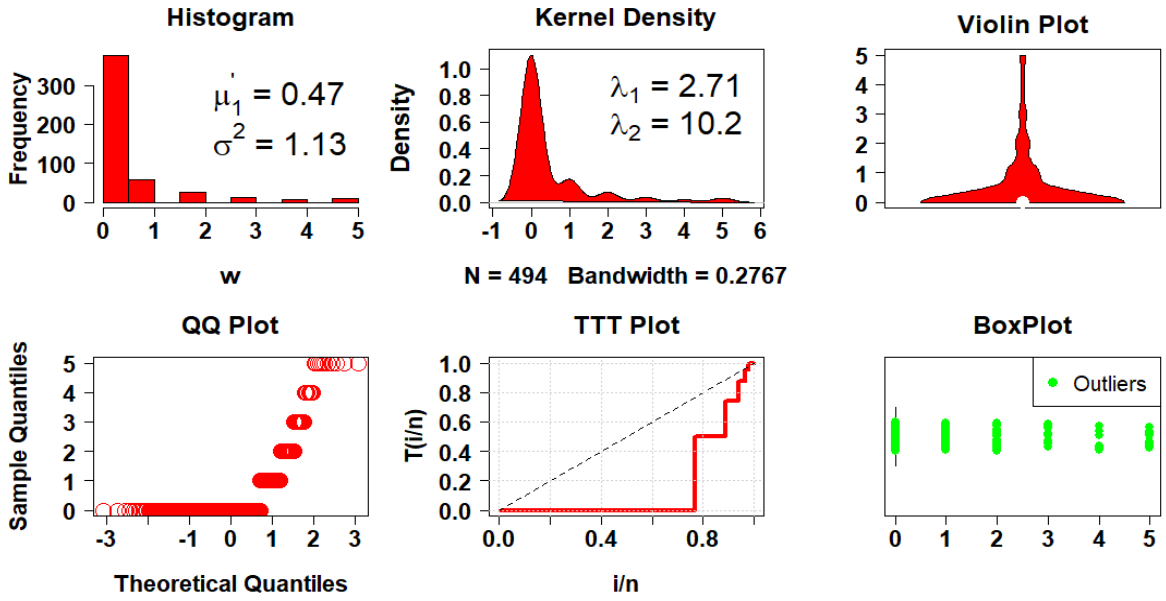

Figure 14. Descriptive visual representation of criminology data.

Figure 14 demonstrates the excellent fit of the DPITL distribution for the criminology data using a TTT plot. This selection was numerically supported because parameter calculations confirmed the data's over-dispersed nature, matching the DPITL model characteristics. Since the PMF displayed an asymmetrical distribution, QQ plots and box plots were used to find outliers, marked by green rings. Finally, Table 8 summarizes the overall performance with goodness-of-fit metrics, and Figure 15 visually shows how the estimators successfully maximized the data fit.

Table 8. Observed vs. expected frequencies and model fit for criminology data.

w	Observed	Expected						
Distribution		DPITL	DIKU	DBurr	DIW	DLL	DLO	NB
0	378	376.78024	377.07151	377.35103	376.95193	376.05933	373.31546	376.91281
1	59	67.86700	68.22705	67.91836	70.63013	70.27035	72.20826	63.79716
2	26	23.19178	22.59963	22.55391	20.11535	21.08508	23.81006	25.84724
3	13	10.40213	10.02107	10.01595	8.83166	9.24766	10.24708	12.50426
4	7	5.44737	5.25824	5.26263	4.77912	4.95317	5.18377	6.54085
5	11	10.31149	10.82249	10.89813	12.69180	12.38441	9.23537	8.39769
	Total	494						
	$-log\,\ell$	436.34625	437.42421	437.44150	441.40221	440.26395	436.88571	438.00594
	Λ_1	876.69250	878.84843	878.88300	886.80443	884.52791	877.77141	878.01188
	Λ_2	885.09757	887.25350	887.28807	895.20950	892.93298	886.17649	886.41695
	Λ_3	878.69250	880.84843	880.88300	888.80443	886.52791	879.77141	880.01188
	Λ_4	879.99233	882.14826	882.18284	890.10426	887.82774	881.07125	881.31171
	χ^2	2.63980	3.22717	3.16228	6.86441	5.48641	4.38935	3.22306
	Λ_5	0.45056	0.35791	0.36727	0.07635	0.13945	0.22237	0.34748

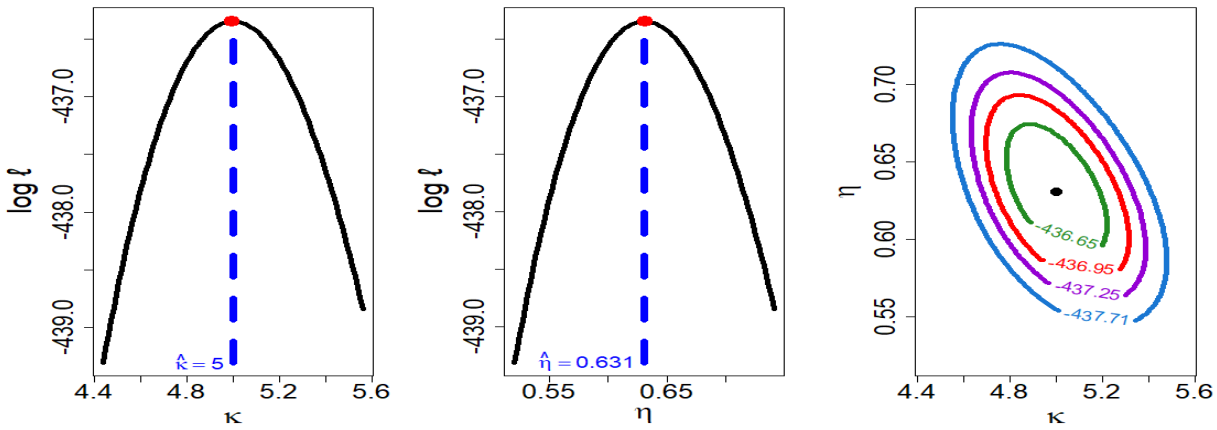

Figure 15. Contour and profile plots for DPITL distribution on criminology data.

The analysis clearly shows that DPITL distribution is the best choice for criminology data. Table 8 proves this because the DPITL model achieved the lowest metrics and highest Λ_5 scores compared to all other models. This excellent quality of estimation is visually supported by Figure 15, which confirms that the estimators successfully fit the DPITL model. Furthermore, Figure 16 clearly demonstrates the distribution's effectiveness, as the estimated PMFs perfectly match the actual observed data.

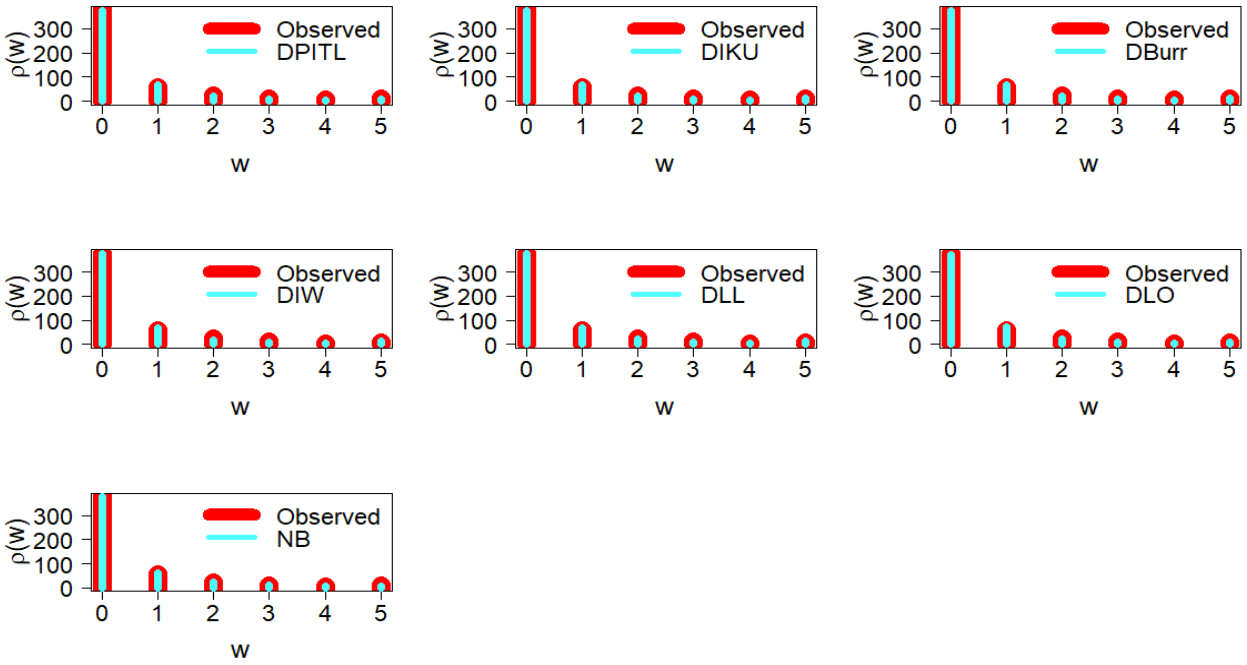

Figure 16. Observed and expected frequencies of fitted PMFs for criminology data.

Figure 16 visually confirms the model's fit. Following this, Figure 17 offers a last visual comparison to show how accurate the method is. Finally, Table 9 lists the model parameters and their SEs.

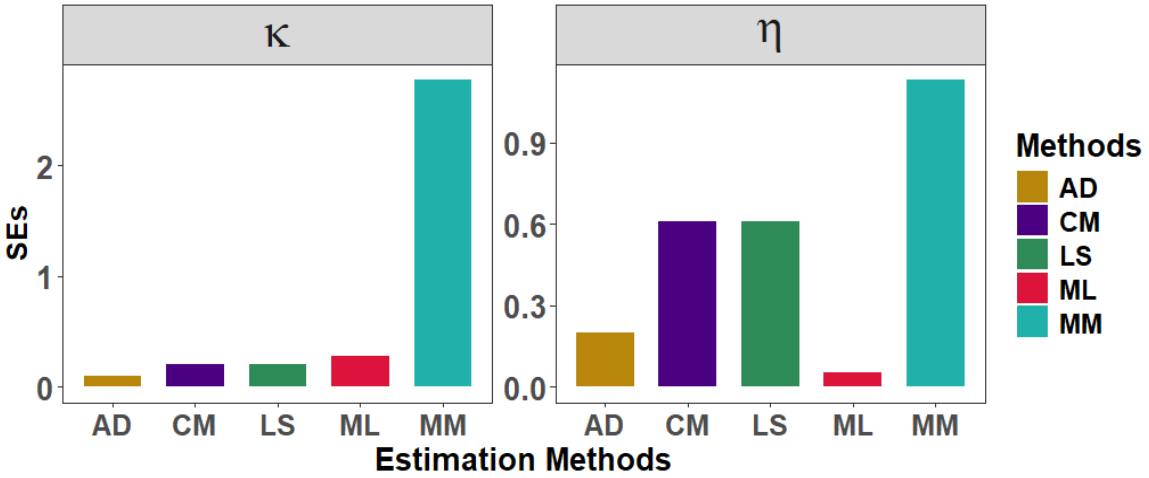

Figure 17. Visual comparison of SEs for DPITL distribution of criminology data.

Table 9. Estimates and SEs from different distributions for criminology data.

Distribution	DPITL					DIKU	DBurr	DIW	DLL	DLO	NB
Methods	LS	AD	CM	MM	ML						
$\hat{\kappa}$	1.67904	1.68196	1.67764	3.22128	5.00026	1.07831	2.08234	0.40690	1.55405	2.62558	0.26405
$\hat{\eta}$	1.96284	1.88004	1.97031	1.21884	0.63065	2.17412	1.02957	1.45442	0.47418	1.40748	0.35897
SE ($\hat{\kappa}$)	0.20499	0.09979	0.20491	2.77023	0.28035	0.22169	0.11648	0.04364	0.13686	0.57886	0.04434
SE ($\hat{\eta}$)	0.60893	0.19908	0.61146	1.13021	0.05513	0.22557	0.09366	0.13265	0.04787	0.44710	0.04611

Both Table 9 and Figure 17 agree: the best method for estimation changes depending on the parameter. The AD method is the most reliable for estimating parameter κ. However, the ML method works best for parameter η. This result is consistent with the simulation tests.

7. Concluding remarks

This article introduces a new two-parameter discrete distribution called DPITL distribution using a survival discretization technique. Factorial moments, PGF, QF, mean, variance, mean residual life, Tsallis entropy, and Rényi entropy are among the statistical features analyzed. Some estimation methods, including ML, MM, LS, AD, and CM, are used to determine the best estimators of the unknown parameters. The simulation study demonstrates the performance of different estimation methods for DPITL distribution. Results show that larger sample sizes enhance estimation accuracy, while higher parameter values may reduce precision. Moreover, the findings indicate that the most suitable method varies by parameter: the ML method performs best for the η parameter, whereas the AD method is preferable for the κ parameter. The distribution's usefulness was also confirmed through real data applications, including corticosteroid-induced liver lesions in rats, chromatid aberrations in human cells, and criminology data. These applications highlight their strong potential for modeling discrete count data, especially in fields such as social sciences, toxicology, pharmacology, and epidemiology, where it can be applied to analyze events like drug side effects, cellular changes, and the spread of new disease cases.

The present study establishes a strong foundation for DPITL distribution, yet it also suggests several compelling directions for future work. A particularly valuable extension would be to formulate a Bayesian paradigm for estimating the model's parameters. This would serve as a powerful complement to the frequentist techniques (ML, MM, LS, CM, AD) already investigated. The Bayesian approach naturally incorporates previous knowledge, and its result, a joint posterior distribution for the parameters, offers a thorough and understandable framework for probabilistic inference that goes beyond point estimates to completely describe uncertainty [37–40].

Author contributions

Amal S. Hassan: Methodology (equal); Writing–original draft (equal); Writing–review; formal analysis (equal) & editing (equal); Eslam Abdelhakim Seyam: Methodology (equal); Software (equal); Writing–original draft (equal); Formal analysis (equal); Asma Ahmad Alzahrani: Funding acquisition (equal); Methodology (equal); Project administration; Writing–review & editing (equal); Omar A. Saudi: Methodology (equal); Data curation (equal); Writing–review & editing (equal); Software (equal); Writing–original draft (equal); Formal analysis (equal). All authors have read and agreed to the

published version of the manuscript.

Funding

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2603).

Use of Generative-AI tools declaration

The authors declare they have not used Artificial intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare that they have no conflicts of interest.

References

1. D. Roy, The discrete normal distribution, *Commun. Stat. Theor. M.*, **32** (2003), 1871–1883. https://doi.org/10.1081/STA-120023256
2. T. Nakagawa, S. Osaki, The discrete Weibull distribution, *IEEE T. Reliab.*, **24** (1975), 300–301. https://doi.org/10.1109/TR.1975.5214915
3. H. Krishna, P. S. Pundir, Discrete Burr and discrete Pareto distributions, *Stat. Methodol.*, **6** (2009), 177–188. https://doi.org/10.1016/j.stamet.2008.07.001
4. M. A. Jazi, C. Lai, M. H. Alamatsaz, A discrete inverse Weibull distribution and estimation of its parameters, *Stat. Methodol.*, **7** (2010), 121–132. https://doi.org/10.1016/j.stamet.2009.11.001
5. B. A. Para, Discrete generalized Burr-Type XII distribution, *J. Mod. Appl. Stat. Meth.*, **13** (2014), 244–258. https://doi.org/10.22237/jmasm/1414815120
6. B. A. Para, T. R. Jan, Discrete version of log-logistic distribution and its application in genetics, *Int. J. Mod. Math. Sci.*, **14** (2016), 407–422.
7. J. M. Jia, Z. Z. Yan, X. Y. Peng, A new discrete extended Weibull distribution, *IEEE Access*, **7** (2019), 175474–175486. https://doi.org/10.1109/ACCESS.2019.2957788
8. M. El-Morshedy, M. S. Eliwa, E. Altun, Discrete Burr-Hatke distribution with properties, estimation methods and regression model, *IEEE Access*, **8** (2020), 74359–74370. https://doi.org/10.1109/ACCESS.2020.2988431
9. S. S. Maiti, M. Dey, S. Sarkar, Discrete Xgamma distributions: properties, estimation and an application to the collective risk model, *J. Reliab. Stat. Stud.*, **11** (2018), 117–132.
10. E. Altun, M. El-Morshedy, M. S. Eliwa, A Study on discrete Bilal distribution with properties and applications on integer-valued autoregressive process, *Revstat-Stat. J.*, **20** (2022), 501–528.
11. M. Irshad, P. Jodrá, A. Krishna, R. Maya, On the discrete analogue of the Teissier distribution and its associated INAR (1) process, *Math. Comput. Simul.*, **214** (2023), 227–245. https://doi.org/10.1016/j.matcom.2023.07.007
12. A. S. Eldeeb, M. Ahsan-Ul-Haq, A. Babar, A discrete analog of inverted Topp-Leone distribution: properties, estimation and applications, *Int. J. Anal. Appl.*, **19** (2021), 695–708. https://doi.org/10.28924/2291-8639-19-2021-695

13. M. Shafqat, S. Ali, I. Shah, S. Dey, Univariate discrete Nadarajah and Haghighi distribution: Properties and different methods of estimation, *Statistica*, **80** (2020), 301–330. https://doi.org/10.6092/issn.1973-2201/9532
14. D. Das, M. Abouelenein, B. Das, P. Hazarika, M. El-Morshedy, N. Roushdy, et al., A discrete expansion of the Lindley distribution: Mathematical and statistical characterizations with estimation techniques, simulation, and goodness-of-fit analysis, *Comput. J. Math. Stat. Sci.*, 2025. https://doi.org/10.21608/cjmss.2025.373562.1146
15. F. C. Opone, E. A. Izekor, I. U. Akata, F. E. U. Osagiede, A discrete analogue of the continuous Marshall-Olkin Weibull distribution with application to count data, *Earthline J. Math. Sci.*, **5** (2021), 415–428. https://doi.org/10.34198/ejms.5221.415428
16. S. Chakraborty, D. Chakravarty, J. Mazucheli, W. Bertoli, A discrete analog of Gumbel distribution: Properties, parameter estimation and applications, *J. Appl. Stat.*, **48** (2021), 712–737. https://doi.org/10.1080/02664763.2020.1744538
17. A. S. Eldeeb, M. Ahsan-Ul-Haq, 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 Math.*, **7** (2021), 1726–1741. https://doi.org/10.3934/math.2022099
18. J. Mazucheli, W. Bertoli, R. P. Oliveira, A. F. B. Menezes, On the discrete quasi Xgamma distribution, *Methodol. Comput. Appl. Probab.*, **22** (2020), 747–775. https://doi.org/10.1007/s11009-019-09731-7
19. R. Maya, P. Jodrá, S. Aswathy, M. R. Irshad, The discrete new XLindley distribution and the associated autoregressive process, *Int. J. Data Sci. Anal.*, **20** (2025), 1767–1793. https://doi.org/10.1007/s41060-024-00563-4
20. A. A. EL-Helbawy, M. A. Hegazy, G. R. AL-Dayian, R. E. Abd EL-Kader, A discrete analog of the inverted Kumaraswamy distribution: Properties and estimation with application to COVID-19 Data, *Pak. J. Stat. Oper. Res.*, **18** (2022), 297–328. https://doi.org/10.18187/pjsor.vl8il.3634
21. H. Fawzy, Discrete Marshall-Olkin extended Burr Type XII distribution: Properties and estimation, *Egyptian Stat. J.*, **66** (2022), 17–41.
22. H. Elsayed, M. Hussein, A new discrete analogue of the continuous Muth distribution for over-dispersed data: Properties, estimation techniques, and application, *Entropy*, **27** (2025), 409. https://doi.org/10.3390/e27040409
23. T. A. Abushal, A. S. Hassan, A. R. El-Saeed, S. G. Nassr, Power inverted Topp-Leone distribution in acceptance sampling plans, *Comput. Mater. Con.*, **67** (2021), 991–1011. https://doi.org/10.32604/cmc.2021.014620
24. A. S. Hassan, M. Elgarhy, R. Ragab, Statistical properties and estimation of inverted Topp-Leone distribution, *J. Stat. Appl. Probab.*, **9** (2020), 319–331. https://doi.org/10.18576/jsap/090212
25. A. R. El-Saeed, A. S. Hassan, N. M. Elharoun, A. Al Mutairi, R. H. Khashab, S. G. Nassr, A class of power inverted Topp-Leone distribution: Properties, different estimation methods and applications, *J. Radiat. Res. Appl. Sc.*, **16** (2023), 100643. https://doi.org/10.1016/j.jrras.2023.100643
26. S. G. Nassr, A. S. Hassan, R. Alsultan, A. R. El-Saeed, Acceptance sampling plans for the three-parameter inverted Topp-Leone model, *AIMS Math. Biosci. Eng.*, **19** (2022), 13628–13659. https://doi.org/10.3934/mbe.2022636

27. G. M. Ibrahim, A. S. Hassan, E. M. Almetwally, H. M. Almongy, Parameter estimation of alpha power inverted Topp-Leone distribution with applications, *Intell. Autom. Soft Co.*, **29** (2021), 353–371. https://doi.org/10.32604/iasc.2021.017586
28. A. S. Hassan, E. M. Almetwally, Applications to physical data using four-parameter inverted Topp-Leone model, *Thail. Statist.*, **22** (2024), 430–457.
29. V. K. Rohatgi, A. K. M. E. Saleh, *An introduction to probability and statistics*, John Wiley and Sons, 2015. https://doi.org/10.1002/9781118799635
30. A. Rényi, On measures of entropy and information, In: *Proceedings of the 4th Berkeley symposium on mathematical statistics and probability*, University of California Press, **1** (1961), 547–561.
31. C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, *J. Stat. Phys.*, **52** (1988), 479–487. https://doi.org/10.1007/BF01016429
32. M. A. Khan, A. Khalique, A. M. Abouammoh, On estimating parameters in a discrete Weibull distribution, *IEEE T. Reliab.*, **38** (1989), 348–350. https://doi.org/10.1109/24.44179
33. B. A. Para, T. R. Jan, On discrete three-parameter Burr type XII and discrete Lomax distributions and their applications to model count data from medical science, *Biom. Biostat. Int. J.*, **4** (2016), 70–82. https://doi.org/10.15406/bbij.2016.04.00092
34. S. Chan, P. R. Riley, K. L. Price, F. McElduff, P. J. Winyard, Corticosteroid-induced kidney dysmorphogenesis is associated with deregulated expression of known cystogenic molecules, as well as Indian hedgehog, *Am. J. Physiol. Renal*, **298** (2010), 346–356. https://doi.org/10.1152/ajprenal.00350.2009
35. M. Borah, J. Hazarika, Discrete Shanker distribution and its derived distributions, *Biom. Biostat. Int. J.*, **5** (2017), 146–153.
36. M. Ahsan-ul-Haq, M. N. S. Hussain, J. Talib, S. Tariq, Zero-inflated Poisson XLindley distribution for medical science modeling, *J. Stat.*, **29** (2025), 110–127. https://doi.org/10.58575/4cwngm09
37. H. M. Yousof, C. Chesneau, G. H. Hamedani, M. Ibrahim, A new discrete distribution: properties, characterizations, modeling real count data, Bayesian and non-Bayesian estimations, *Statistica*, **81** (2021), 135–162. https://doi.org/10.6092/issn.1973-2201/11635
38. O. A. Alamri, Classical and Bayesian estimation of discrete Poisson Agu-Eghwerido distribution with applications, *Alex. Eng. J.*, **109** (2024), 768–777. https://doi.org/10.1016/j.aej.2024.09.063
39. I. Alkhairy, Classical and Bayesian inference for the discrete Poisson Ramos-Louzada distribution with application to COVID-19 data, *Math. Biosci. Eng.*, **20** (2023), 14061–14080. https://doi.org/10.3934/mbe.2023628
40. J. Debastiani Neto, R. P. Oliveira, F. A. Moala, J. A. Achcar, Introducing the discrete xLindley distribution: A one-parameter model for overdispersed data, *Rev. Colomb. Estad.*, **48** (2025), 39–70. http://doi.org/10.15446/rce.v48n1.115319

AIMS Press