STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
*Stat., Optim. Inf. Comput.*, Vol. **10**, March 2022, pp 352–371.
Published online in International Academic Press (www.IAPress.org)

# The Discrete Inverse Burr Distribution with Characterizations, Properties, Applications, Bayesian and Non-Bayesian Estimations

Christophe Chesneau [1,*], Haitham M. Yousof [2], G.G. Hamedani [3] and Mohamed Ibrahim[4]

[1] *Université de Caen Normandie, LMNO, Campus II, Science 3, 14032, Caen, France*
[2] *Department of Statistics, Mathematics and Insurance, Benha University, Benha 13513, Egypt*
[3] *Department of Mathematical and Statistical Sciences, Marquette University, USA*
[4] *Department of Applied, Mathematical and Actuarial Statistics, Faculty of Commerce, Damietta University, Damietta, Egypt*

**Abstract** A new one-parameter heavy-tailed discrete distribution with an infinite mean is defined and studied. The probability mass function of the new distribution can be "unimodal and right-skewed" and its failure rate can be monotonically decreasing. Some of its relevant properties are discussed. Some characterizations based on: (i) the conditional expectation of a certain function of the random variable and (ii) in terms of the reversed hazard function are presented. Different Bayesian and non-Bayesian estimation methods are described and compared using simulations, and two real data applications are given. The new model is used to model carious teeth data and counts of cysts in kidney datasets, and it outperforms many well-known competitive discrete models.

**Keywords** Discretization, Characterization, Weighted Least Square Estimation, Maximum Likelihood Estimation, Count Data, Bayesian Estimation, Ordinary Least Square Estimation

**AMS 2010 subject classifications** 62E15

**DOI:** 10.19139/soic-2310-5070-1393

## 1. Introduction

A continuous failure time distribution can be employed for generating a corresponding discrete distributions. If a continuous failure time random variable (RV) $W$ has the survival function (SF) $\mathbf{S}_W(w) = \Pr[W > w]$, the probability mass function (PMF) of the largest integer less than or equal to $W$, i.e., $\lfloor W \rfloor$, can be written as $\mathcal{F}_W(0) = \Pr(\lfloor W \rfloor = 0)$, and

$$\mathcal{F}_W(w) = \Pr(\lfloor W \rfloor = w) = \Pr(w \leq W \leq 1 + w) = \mathbf{S}_W(w) - \mathbf{S}_W(1 + w)\,|_{w \in \mathbb{N}^*},$$

where $\mathbb{N}^* = \mathbb{N}/\{0\}$ and $\mathbb{N} = \{0, 1, 2, \ldots\}$. Thus, discretization of a continuous RV is an interesting and simple approach to derive a discrete RV corresponding to the continuous one. A plethora of discrete distributions has been introduced as a result of this approach. The most relevant of them include the poisson-Lindley (PLi) distribution by Sankaran [29], discrete Weibull distribution by Nakagawa and Osaki [23], another discrete Weibull distribution by Stein and Dattero [30], discrete Rayleigh (DR) distribution by Roy [28], discrete half-normal distribution by Kemp [19], discrete Pareto (DPa) distribution by Krishna and Pundir [21], discrete inverse Weibull (DIW) distribution by Jazi et al. [17], discrete generalized geometric (DGGc) distribution by Gomez-Déniz [11], discrete Lindley (DLi) distribution by Gommez-Déniz and Calderin-Ojeda [12], discrete generalized exponential (DGE-II) distribution

*Correspondence to: Christophe Chesneau (Email: christophe.chesneau@gmail.com). Université de Caen Normandie, LMNO, Campus II, Science 3, 14032, Caen, France.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

by Nekoukhou et al. [24], discrete additive Weibull (DAW) distribution by Bebbington et al. [3], discrete inverse Rayleigh (DIR) distribution by Hussain and Ahmad [13], discrete Lomax (DLx) distribution by Para and Jan [25], discrete log-logistic (DLL) distribution by Para and Jan [25], discrete Burr-XII (DBXII) distribution by Para and Jan [26], discrete Lindley (DLi-II) distribution by Hussain and Aslam [14], discrete linear failure rate (DLFR) distribution by Kumar et al. [22], discrete weighted Lindley (DWLi) distribution by Bodhisuwan and Sangpoom [3], exponentiated discrete Lindley (EDLi) distribution by El-Morshedy et al. [9], discrete generalized Burr-Hatke (DGBH) distribution by El-Morshedy et al. [10], discrete generalized Burr-Hatke (DGBH) distribution by Yousof et al. [31], discrete Rayleigh G (DRG) family of distributions by Aboraya et al. [1] which contains many sub-distributions such as the discrete Rayleigh Weibull, discrete Rayleigh exponential, discrete Rayleigh Log-logistic, discrete Rayleigh Lomax, discrete Rayleigh Rayleigh, discrete Rayleigh Burr-XII, discrete Rayleigh Fréchet, discrete Rayleigh inverse Rayleigh, discrete Rayleigh inverse exponential, discrete Rayleigh inverse Lomax, discrete Rayleigh half-logistic, discrete Rayleigh Gumbel, discrete Rayleigh Lindley, discrete Rayleigh Nadarajah-Haghighi, discrete Rayleigh Gompertz, discrete Rayleigh Dagum, discrete Rayleigh inverse flexible Weibull, discrete Rayleigh inverse Gompertz, and the discrete analogue of the Weibull G family by Ibrahim et al. [15], among others.

In reliability analysis, classification of lifetime distributions is defined in terms of their SFs and other related reliability characteristics. For example, the class of the increasing (decreasing) hazard rate IHR (DHR), the class of the new better (worse) than used NBU (NWU), the class of the new better (worse) than used in expectation NBUE (NWUE), the class of the increasing (decreasing) hazard rate average IHRA (DHRA) and the class of the increasing (decreasing) mean residual lifetime IMRL (DMRL), etc. (See Kemp [19], Kemp [18] and Krishna and Pundir [21]). In this work, we derive a discrete inverse Burr (DIB) distribution using the general approach of discretizing a continuous RV.

The rest of this paper is organized as follows. In Section 2, the discrete analogue of inverse Burr distribution is developed and some of its distributional properties are studied. In Section 3, we present certain characterizations of DIB distribution based on: $(i)$ the conditional expectation of certain function of a RV and $(ii)$ in terms of the reversed hazard function. A graphical and numerical analysis is presented in Section 4. In Section 5, non-Bayesian and Bayesian estimation methods are considered. Section 6 deals with some numerical simulation results for comparing Bayesian and non-Bayesian estimation methods. Two application under the carious teeth and counts of cysts of kidneys data sets are considered in Section 7 for comparing Bayesian and non-Bayesian estimation methods. In Section 8, the DIB model is considered for modeling the carious teeth and counts of cysts of kidneys data sets using the maximum likelihood estimation method and also compared with many competitive models. Some concluding remarks are given in Section 9.

## 2. The new model

A RV $X$ is said to have the inverse Burr (IB) distribution with parameter $\widetilde{\pi}$ if its cumulative distribution function (CDF) is given by

$$F_{\widetilde{\pi}}(x)\,|_{(\widetilde{\pi}>0 \text{ and } x>0)} = \left(1+\frac{1}{x}\right)^{-\widetilde{\pi}}.$$

Based on this function, the CDF of the discrete inverse Burr (DIB) distribution with parameter $\widetilde{\pi}$ can be expressed as

$$\mathbf{F}_{\widetilde{\pi}}(x)\,|_{(\widetilde{\pi}>0 \text{ and } x\in\mathbb{N})} = F_{\widetilde{\pi}}(1+x) = \left(1+\frac{1}{1+x}\right)^{-\widetilde{\pi}} = \left(\frac{1+x}{2+x}\right)^{\widetilde{\pi}}. \tag{1}$$

An immediate property of this CDF is the following first-order stochastic ordering property: for any $\widetilde{\pi}_1 \leq \widetilde{\pi}_2$, the following inequality is fulfilled: $\mathbf{F}_{\widetilde{\pi}_2}(x) \leq \mathbf{F}_{\widetilde{\pi}_1}(x)$, meaning the DIB distribution with parameter $\widetilde{\pi}_2$ first-order stochastically dominates the DIB distribution with parameter $\widetilde{\pi}_1$. As another remark, we can write the CDF under

the following power function form:

$$\mathbf{F}_{\widetilde{\pi}}(x) = \pi^{g(x)},$$

where $\pi = \exp(-\widetilde{\pi}) \in (0,1)$ and $g(x) = \log[1 + 1/(1+x)]$. Following the "geometric distribution spirit" and for practical reasons, in the rest of the paper, we will sometimes refer to the parameter $\pi$ instead of $\widetilde{\pi}$.

As another function of importance, the associated SF can be written as

$$\mathbf{S}_{\widetilde{\pi}}(x)\,|_{(\widetilde{\pi}>0 \text{ and } x\in\mathbb{N})} = 1 - \mathbf{F}_{\widetilde{\pi}}(x) = 1 - \left(\frac{1+x}{2+x}\right)^{\widetilde{\pi}}.$$

Also, the PMF of the DIB distribution satisfies: $\mathcal{F}_{\widetilde{\pi}}(0) = \mathbf{F}_{\widetilde{\pi}}(0) = (1/2)^{\widetilde{\pi}}$ and

$$\mathcal{F}_{\widetilde{\pi}}(x) = \mathbf{S}_{\widetilde{\pi}}(x-1) - \mathbf{S}_{\widetilde{\pi}}(x) \quad |_{(\widetilde{\pi}>0 \text{ and } x\in\mathbb{N}^*)},$$

which can be written as

$$\mathcal{F}_{\widetilde{\pi}}(x)\,|_{(\widetilde{\pi}>0 \text{ and } x\in\mathbb{N}^*)} = \left(\frac{1+x}{2+x}\right)^{\widetilde{\pi}} - \left(\frac{x}{1+x}\right)^{\widetilde{\pi}}. \tag{2}$$

Let us now discuss the mode of the DIB distribution. For determining the mode of the DIB distribution, the following two inequalities are useful:

$$\mathcal{F}_{\widetilde{\pi}}(1+x)/\mathcal{F}_{\widetilde{\pi}}(x) \leq 1, \quad \mathcal{F}_{\widetilde{\pi}}(x)/\mathcal{F}_{\widetilde{\pi}}(x-1) \geq 1.$$

If $x = \varkappa_{\widetilde{\pi}}(1+x, x)$ is the minimum value of $x$ satisfying the first inequality, and $x = \varkappa_{\widetilde{\pi}}(x, x-1)$ the maximum value of $x$ satisfying the second inequality, respectively, then the mode of the DIB distribution, say $\mathbb{M}_x$, can be obtained from $\varkappa_{\widetilde{\pi}}(1+x, x) \leq \lfloor \mathbb{M}_x \rfloor \leq \varkappa_{\widetilde{\pi}}(x, x-1)$. Clearly, it is not easy to to find $\varkappa_{\widetilde{\pi}}(1+x, x)$ and $\varkappa_{\widetilde{\pi}}(x, x-1)$ theoretically. The quantities $\varkappa_{\widetilde{\pi}}(1+x, x)$ and $\varkappa_{\widetilde{\pi}}(x, x-1)$, however, depend on the parameter $\widetilde{\pi}$ which enables the numerical and graphical solutions to work. Hence, the numerical and the graphical approaches are required for obtaining the mode for different parameter values.

The identifiability of the DIB distribution is also an interesting property that deserves discussion. Theoretically, it is difficult to ensure. Ideally, to prove this property, we need to prove that, for any $x \in \mathbb{N}$, the equality $\mathcal{F}_{\widetilde{\pi}^*}(x) = \mathcal{F}_{\widetilde{\pi}}(x)$ implies that $\widetilde{\pi}^* = \widetilde{\pi}$. This equality is clear for $x = 0$. However, for the other cases, the complexity of the PMF is a significant handicap to demonstrating it in full rigor. Our practical investigations, however, have revealed no problem of this kind, but the rigorous proof still remains a strong mathematical challenge in many cases.

As a more "dynamic and fingerprint function" of the DIB distribution, the hazard rate function (HRF) is specified by

$$h_{\widetilde{\pi}}(x)\,|_{(\widetilde{\pi}>0 \text{ and } x\in\mathbb{N}^*)} = \frac{\mathcal{F}_{\widetilde{\pi}}(x)}{\mathbf{S}_{\widetilde{\pi}}(x-1)} = \frac{[(1+x)/(2+x)]^{\widetilde{\pi}} - [x/(1+x)]^{\widetilde{\pi}}}{1 - [x/(1+x)]^{\widetilde{\pi}}}.$$

The quantile function (QF) of the DIB distribution is defined as

$$\begin{aligned}\mathbf{Q}_{\widetilde{\pi}}(\pi)\,|_{(\widetilde{\pi}>0 \text{ and } \pi\in(0,1))} &= \inf\{x \in \mathbb{N} :\ \mathbf{F}_{\widetilde{\pi}}(x) \geq \pi\} \\ &= \inf\left\{x \in \mathbb{N} :\ x \geq \frac{2\pi^{1/\widetilde{\pi}} - 1}{1 - \pi^{1/\widetilde{\pi}}}\right\} = \left\lceil \frac{2\pi^{1/\widetilde{\pi}} - 1}{1 - \pi^{1/\widetilde{\pi}}} \right\rceil.\end{aligned}$$

Thus, the QF can be calculated in any situation

Let us now emphasize with the moment properties of the DIB distribution. First, we can note that, as $x \to +\infty$, we have $\mathcal{F}_{\widetilde{\pi}}(x) \sim \widetilde{\pi}/x^2$. Therefore, the ordinary moments of $X$ do not exist by the Riemann series theorem. The same holds for the negative moments since the support of the DIB distribution is $\mathbb{N}$. However, for a real number

$a \in (0,1)$, $\mathbb{E}(X^a)$ exists and can be theoretically expanded as

$$\mathbb{E}(X^a) = \sum_{x\in\mathbb{N}} x^a \mathcal{F}_{\widetilde{\pi}}(x) = \sum_{x\in\mathbb{N}} [x^a - (1+x)^a] \left(\frac{1+x}{2+x}\right)^{\widetilde{\pi}}.$$

The DIB distribution is also a heavy-tailed discrete distribution; for any $t > 0$, since $\mathcal{F}_{\widetilde{\pi}}(x) \sim \widetilde{\pi}/x^2$ when $x \to +\infty$, we have

$$\sum_{x\in\mathbb{N}} e^{tx} \mathcal{F}_{\widetilde{\pi}}(x) = +\infty.$$

The discrete heavy-tailed property is particularly demanded for the analysis of data with an underlying PMF which appears to decay very slowly when the value of the variable increases.

Also, the probability generating function (PGF) is convergent here, and can be defined as

$$\mathbf{G}_{\widetilde{\pi}}(s)\,|_{(\widetilde{\pi}>0 \text{ and } s\in(0,1))} = \mathbb{E}(s^X) = \sum_{x\in\mathbb{N}} s^x \mathcal{F}_{\widetilde{\pi}}(x) = (1-s)\sum_{x\in\mathbb{N}} s^x \left(\frac{1+x}{2+x}\right)^{\widetilde{\pi}}.$$

Therefore, moments of the form $s^X$ are quite manageable as series expansions, provided that $s \in (0,1)$.

Now, some distributional properties of the order statistics of the DIB distribution are presented. We recall that the $i^{\text{th}}$ order statistic is defined as the $i^{\text{th}}$ greatest RV among $n$ independent and identically distributed RVs $X_1, \ldots, X_n$. In particular, by using a well-established general result, the CDF of the $i^{\text{th}}$ order statistic of the DIB distribution can be expressed as

$$\begin{aligned}
\mathbf{F}_{\widetilde{\pi},i:n}(x)|_{(\widetilde{\pi}>0 \text{ and } x\in\mathbb{N})} &= \sum_{\kappa=i}^{n} \binom{n}{\kappa} [\mathbf{F}_{\widetilde{\pi}}(x)]^{\kappa} [\mathbf{S}_{\widetilde{\pi}}(x)]^{n-\kappa} \\
&= \sum_{\kappa=i}^{n}\sum_{j=0}^{n-\kappa} \binom{n}{\kappa}\binom{n-\kappa}{j}(-1)^j [\mathbf{F}_{\widetilde{\pi}}(x)]^{\kappa+j} \\
&= \sum_{\kappa=i}^{n}\sum_{j=0}^{n-\kappa} \binom{n}{\kappa}\binom{n-\kappa}{j}(-1)^j \left(\frac{1+x}{2+x}\right)^{\widetilde{\pi}(\kappa+j)}.
\end{aligned}$$

The associated PMF immediately follows as

$$\begin{aligned}
&\mathcal{F}_{\widetilde{\pi},i:n}(x)|_{(\widetilde{\pi}>0 \text{ and } x\in\mathbb{N})} = \mathbf{F}_{\widetilde{\pi},i:n}(x) - \mathbf{F}_{\widetilde{\pi},i:n}(x-1) \\
&= \sum_{\kappa=i}^{n}\sum_{j=0}^{n-\kappa} \binom{n}{\kappa}\binom{n-\kappa}{j}(-1)^j \left[\left(\frac{1+x}{2+x}\right)^{\widetilde{\pi}(\kappa+j)} - \left(\frac{x}{1+x}\right)^{\widetilde{\pi}(\kappa+j)}\right].
\end{aligned}$$

This expression is the basis for more properties of the order statistics of the DIB distribution. The next section is about important mathematical characterizations of the DIB distribution.

## 3. Characterizations

Characterizations of distributions is an important part of distribution theory which has attracted the attention of a good number of researchers in applied sciences, where an investigator is interested in knowing if their model follows the right distribution. Therefore, the investigator relies on conditions under which their model would follow, specifically the chosen distribution. In this section, certain characterizations of the DIB distribution are presented in terms of $(i)$ the conditional expectation of certain function of the RV and $(ii)$ the reversed HRF.

*Proposition 1*
Let $X:\Omega\to\mathbb{N}$ be a RV. The PMF of $X$ is (2) if and only if

$$\mathbb{E}\left\{\left[\left(\frac{1+X}{2+X}\right)^{\widetilde{\pi}}-\left(\frac{X}{1+X}\right)^{\widetilde{\pi}}\right]|X\leq k\right\}=\left(\frac{1+k}{2+k}\right)^{\widetilde{\pi}},\quad k\in\mathbb{N}. \tag{3}$$

**Proof.** If $X$ has PMF (2), then we have

$$\begin{aligned}&\mathbb{E}\left\{\left[\left(\frac{1+X}{2+X}\right)^{\widetilde{\pi}}-\left(\frac{X}{1+X}\right)^{\widetilde{\pi}}\right]|X\leq k\right\}=\frac{1}{\mathbf{F}_{\widetilde{\pi}}(k)}\sum_{x=0}^{k}\left\{\left[\left(\frac{1+x}{2+x}\right)^{2\widetilde{\pi}}-\left(\frac{x}{1+x}\right)^{2\widetilde{\pi}}\right]\right\}\\&=\left(\frac{2+k}{1+k}\right)^{\widetilde{\pi}}\sum_{x=0}^{k}\left\{\left[\left(\frac{1+x}{2+x}\right)^{2\widetilde{\pi}}-\left(\frac{x}{1+x}\right)^{2\widetilde{\pi}}\right]\right\}=\left(\frac{2+k}{1+k}\right)^{\widetilde{\pi}}\left(\frac{1+k}{2+k}\right)^{2\widetilde{\pi}}=\left(\frac{1+k}{2+k}\right)^{\widetilde{\pi}}.\end{aligned}$$

Conversely, if (3) holds, then the following equalities hold:

$$\begin{aligned}&\sum_{x=0}^{k}\left\{\left[\left(\frac{1+x}{2+x}\right)^{\widetilde{\pi}}-\left(\frac{x}{1+x}\right)^{\widetilde{\pi}}\right]\mathcal{F}_{\widetilde{\pi}}(x)\right\}=\mathbf{F}_{\widetilde{\pi}}(k)\left(\frac{1+k}{2+k}\right)^{\widetilde{\pi}}\\&=\left[\mathbf{F}_{\widetilde{\pi}}(1+k)-\mathcal{F}_{\widetilde{\pi}}(1+k)\right]\left(\frac{1+k}{2+k}\right)^{\widetilde{\pi}}.\end{aligned} \tag{4}$$

From (4), we also have

$$\sum_{x=0}^{1+k}\left\{\left[\left(\frac{1+x}{2+x}\right)^{\widetilde{\pi}}-\left(\frac{x}{1+x}\right)^{\widetilde{\pi}}\right]\mathcal{F}_{\widetilde{\pi}}(x)\right\}=\mathbf{F}_{\widetilde{\pi}}(1+k)\left(\frac{2+k}{3+k}\right)^{\widetilde{\pi}}. \tag{5}$$

Now, subtracting (4) from (5), we arrive at

$$\begin{aligned}&\mathbf{F}_{\widetilde{\pi}}(1+k)\left[\left(\frac{2+k}{3+k}\right)^{\widetilde{\pi}}-\left(\frac{1+k}{2+k}\right)^{\widetilde{\pi}}\right]+\left(\frac{1+k}{2+k}\right)^{\widetilde{\pi}}\mathcal{F}_{\widetilde{\pi}}(1+k)\\&=\left[\left(\frac{2+k}{3+k}\right)^{\widetilde{\pi}}+\left(\frac{1+k}{2+k}\right)^{\widetilde{\pi}}\right]\mathcal{F}_{\widetilde{\pi}}(1+k),\end{aligned}$$

or

$$\mathbf{F}_{\widetilde{\pi}}(1+k)\left[\left(\frac{2+k}{3+k}\right)^{\widetilde{\pi}}-\left(\frac{1+k}{2+k}\right)^{\widetilde{\pi}}\right]=\left(\frac{2+k}{3+k}\right)^{\widetilde{\pi}}\mathcal{F}_{\widetilde{\pi}}(1+k).$$

From the last equality, we have

$$r_{\mathbf{F}_{\widetilde{\pi}}}(1+k)=\frac{\mathcal{F}_{\widetilde{\pi}}(1+k)}{\mathbf{F}_{\widetilde{\pi}}(1+k)}=1-\left(\frac{(1+k)(3+k)}{(2+k)^2}\right)^{\widetilde{\pi}},$$

which is the reversed HRF related to the PMF in (2). The proof of Proposition 1 is now complete. □

*Proposition 2*
Let $X:\Omega\to\mathbb{N}$ be a RV. The PMF of $X$ is (2) if and only if its reversed HRF, $r_{\mathbf{F}_{\widetilde{\pi}}}(k)$, satisfies the following

difference equation:

$$r_{\mathbf{F}_{\widetilde{\pi}}}(1+k)-r_{\mathbf{F}_{\widetilde{\pi}}}(k)=\left(\frac{k(2+k)}{(1+k)^2}\right)^{\widetilde{\pi}}-\left(\frac{(1+k)(3+k)}{(2+k)^2}\right)^{\widetilde{\pi}},\quad k\in\mathbb{N}^*, \tag{10}$$

with the initial condition $r_{\mathbf{F}_{\widetilde{\pi}}}(0)=1$.

**Proof.** Clearly, if $X$ has PMF (2), then (10) holds. Now, if (10) holds, then

$$\sum_{x=0}^{k}\left\{r_{\mathbf{F}_{\widetilde{\pi}}}(1+x)-r_{\mathbf{F}_{\widetilde{\pi}}}(x)\right\}=\sum_{x=0}^{k}\left\{\left[\frac{x(2+x)}{(1+x)^2}\right]^{\widetilde{\pi}}-\left[\frac{(1+x)(3+x)}{(2+x)^2}\right]^{\widetilde{\pi}}\right\},$$

or

$$r_{\mathbf{F}_{\widetilde{\pi}}}(1+k)-r_{\mathbf{F}_{\widetilde{\pi}}}(0)=-\frac{1}{(2+k)^{2\widetilde{\pi}}}(1+k)^{\widetilde{\pi}}(3+k)^{\widetilde{\pi}},$$

or, in view of the initial condition $r_{\mathbf{F}_{\widetilde{\pi}}}(0)=1$, we have

$$r_{\mathbf{F}_{\widetilde{\pi}}}(1+k)=1-\frac{1}{(2+k)^{2\widetilde{\pi}}}(1+k)^{\widetilde{\pi}}(3+k)^{\widetilde{\pi}},$$

which is the reversed HRF corresponding to PMF in (2). Proposition 2 is established. □

## 4. Graphical and numerical analysis

To avoid complicated theoretical treatments, we analyze the effect of adding the new additional parameter $\widetilde{\pi}|_{\widetilde{\pi}=-\ln(\pi)}$ to the PMF, HRF, skewness and kurtosis measures. Figure 1 below gives some plots of the PMF of the DIB distribution. Some plots of the HRF of the DIB distribution are shown in Figure 2. These plots show the relationship between the parameter and the measure. Following the spirit of the distribution with no moments, we consider a "highly truncated" RV and investigates sime of its moment properties. Hence, for the truncated RV $X_*=X$ for $X\leq 1000000$, and $X_*=0$ otherwise, Table 1 gives the associated mean $(\mathbb{E}(X_*))$, variance $(\mathbb{V}(X_*))$, skewness $(\mathbb{S}(X_*))$, kurtosis $(\mathbb{K}(X_*))$ and Index of dispersion $(\mathbb{D}_{\text{index}}(X_*))$. Based on Figure 1, it is noted that the shape of the PMF can be "unimodal right skewed" with different shapes. Based on Figure 2, the corresponding HRF can be "monotonically decreasing" with different shapes. According to Table 1, $\mathbb{S}(X_*)$ is positive for all possible parameter values and can range in the interval $(\simeq 49.4, \simeq \infty)$, the spread for its $\mathbb{K}(X_*)$ is ranging from 3255.456 to $\simeq\infty$. As anticipated, we have $\mathbb{D}_{\text{index}}(X_*)>1$ for all possible parameter values. Thus, the new DIB distribution could be useful in modeling the over-dispersed count data. Furthermore, it is noted that the mean, variance, skewness and kurtosis of $X_*$ decrease as $\pi$ increases. This claim is naturally useless for the DIB distribution since moments of order superior or equal to 1 dont exist.

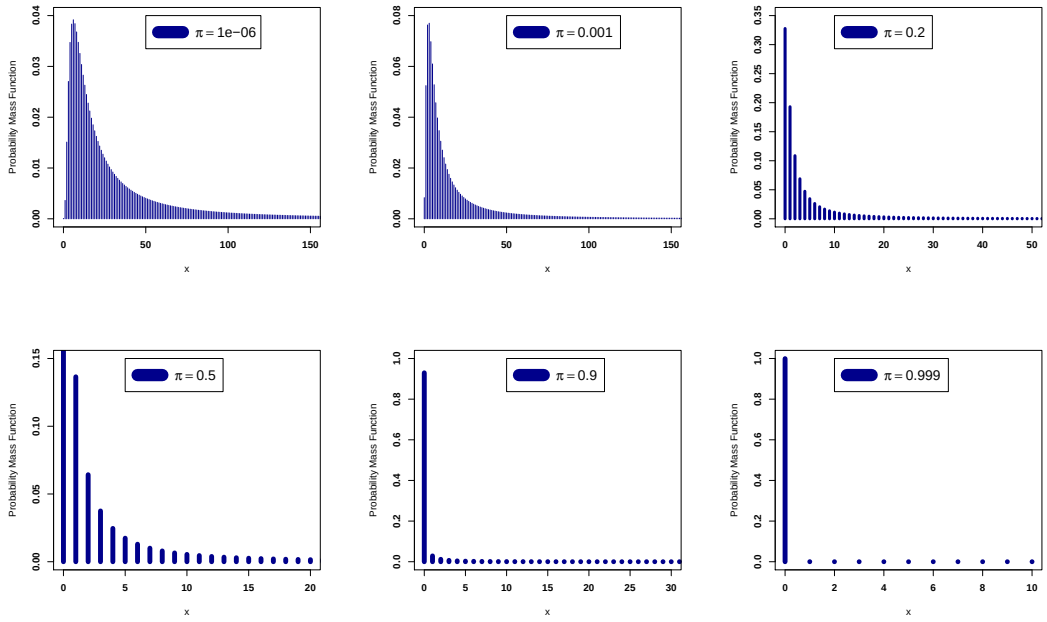


Figure 1. The PMF plots of the DIB distribution.

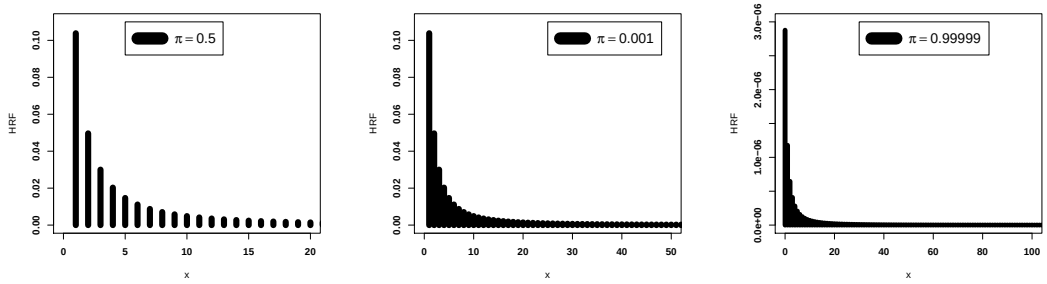


Figure 2. The HRF plots of the DIB distribution.

Table 1. $\mathbb{E}(X_*)$, $\mathbb{V}(X_*)$, $\mathbb{S}(X_*)$, $\mathbb{K}(X_*)$ and $\mathbb{D}_{\text{index}}(X_*)$ for a RV $X_*$ which is a "hightly truncated" version of a RV with the DIB distribution.

| $\pi$ | $\mathbb{E}(X_*)$ | $\mathbb{V}(X_*)$ | $\mathbb{S}(X_*)$ | $\mathbb{K}(X_*)$ | $\mathbb{D}_{\text{index}}(X_*)$ |
|---|---|---|---|---|---|
| $10^{-24}$ | 889.8635 | 102719803 | 49.48966 | 3255.456 | 115433.2 |
| $10^{-12}$ | 273.0858 | 27547603 | 95.39052 | 12116.52 | 100875.3 |
| $10^{-6}$ | 145.6234 | 13791817 | 134.7449 | 24188.76 | 94708.80 |
| $10^{-4}$ | 100.4884 | 9199035 | 164.9534 | 36257.79 | 91543.23 |
| $10^{-2}$ | 52.94871 | 4601997 | 233.1594 | 72458.62 | 86914.25 |
| 0.15 | 22.95495 | 1896500 | 363.1388 | 175794.8 | 82618.36 |
| 0.25 | 16.99379 | 1385946 | 424.7752 | 240544.6 | 81556.03 |
| 0.50 | 8.671137 | 693048.1 | 600.6561 | 481009.6 | 79925.86 |
| 0.75 | 3.648086 | 287660.3 | 932.2919 | 1158835 | 78852.38 |
| 0.95 | 0.6560745 | 51291.5 | 2207.799 | 6498994 | 78179.38 |
| 0.99 | 0.1287509 | 10050.06 | 4987.652 | 33168152 | 78058.13 |

## 5. Estimation

In this section, non-Bayesian and Bayesian estimation methods are considered. In the first subsection, we will consider the maximum likelihood estimation (MLE) method, ordinary least squared estimation (OLSE) method, and weighted least squared estimation (WLSE) method. In the second subsection, the Bayesian estimation method under the squared error loss function (SELF) is considered. All non-Bayesian estimation methods are discussed in the statistical literature in more detail.

### *5.1. Non-Bayesian estimation methods*

*5.1.1. The MLE method* Let $x_1, x_2, \ldots, x_n$ be $n$ observations from a random sample (RS) from the DIB distribution. The log-likelihood function is given by

$$\ell = \ell_n(\widetilde{\pi}) = \sum_{i=1}^{n} \ln \left[ \left( \frac{1+x_i}{2+x_i} \right)^{\widetilde{\pi}} - \left( \frac{x_i}{1+x_i} \right)^{\widetilde{\pi}} \right] |_{(\pi \in (0,1) \text{ and } x_i \in \mathbb{N})}$$

which can be maximized either using the statistical programs or by solving the nonlinear system obtained from $\ell(\widetilde{\pi})$ by differentiation. The components of the score vector $\mathbf{U}(\widetilde{\pi}) = (\partial \ell(\widetilde{\pi})/\partial \widetilde{\pi})^{\intercal}$ can be easily derived. Setting $\partial \ell(\widetilde{\pi})/\partial \widetilde{\pi} = 0$ and solving it yields the MLE for the DIB parameter. The Newton-Raphson algorithm is employed for numerically solving in such cases. As is customary under regularity conditions, the properties of consistency and asymptotic normality, among others, are satisfied. Last but not least, note that the the MLE of $\widetilde{\pi}$ implies the MLE of $\pi$ via the invariance property; all the established theory on consistency and asymptotic normality can be applied to the MLE of $\pi$ with slight modifications. We refer the reader to Casella and Berger [5].

*5.1.2. OLSE method* Let $\mathbf{F}_{\widetilde{\pi}}(x)$ denotes the CDF of the DIB distribution as described in (1), and let $x_1 < x_2 < \cdots < x_n$ be the $n$ ordered values of a RS from the DIB distribution. The OLSE is obtained upon minimizing

$$\mathbf{OLSE}_{(\widetilde{\pi})} = \sum_{i=1}^{n} \left[ \mathbf{F}_{\widetilde{\pi}}(x_i) - \varsigma_{(i,n)}^{[1]} \right]^2,$$

with respect to $\widetilde{\pi}$. Note that we have

$$\mathbf{OLSE}_{(\widetilde{\pi})} = \sum_{i=1}^{n} \left[ \left( \frac{1+x_i}{2+x_i} \right)^{\widetilde{\pi}} - \varsigma_{(i,n)}^{[1]} \right]^2,$$

where $\varsigma_{(i,n)}^{[1]} = i/(n+1)$. The OLSE is obtained via solving the following non-linear equation:

$$0 = \sum_{i=1}^{n} \left[ \left( \frac{1+x_i}{2+x_i} \right)^{\widetilde{\pi}} - \varsigma_{(i,n)}^{[1]} \right] d_{(\widetilde{\pi})}(x_i),$$

where

$$d_{(\widetilde{\pi})}(x_i) = \log\left( \frac{1+x_i}{2+x_i} \right) \left( \frac{1+x_i}{2+x_i} \right)^{\widetilde{\pi}}.$$

*5.1.3. WLSE method* The WLSE is obtained by minimizing the function $\mathbf{WLSE}_{(\widetilde{\pi})}$ with respect to $\widetilde{\pi}$, where

$$\mathbf{WLSE}_{(\widetilde{\pi})} = \sum_{i=1}^{n} \varsigma_{(i,n)}^{[2]} \left[ \mathbf{F}_{\widetilde{\pi}}(x_i) - \varsigma_{(i,n)}^{[1]} \right]^2,$$

where $\varsigma_{(i,n)}^{[2]} = \left[(1+n)^2(2+n)\right] / \left[i(1+n-i)\right]$. The WLSE is obtained by solving the following non-linear equation:

$$0 = \sum_{i=1}^{n} \varsigma_{(i,n)}^{[2]} \left[ \left( \frac{1+x_i}{2+x_i} \right)^{\widetilde{\pi}} - \varsigma_{(i,n)}^{[1]} \right] d_{(\widetilde{\pi})}(x_i).$$

### *5.2. Bayesian estimation*

Assume the beta and uniform priors for the parameter, respectively. Then

$$p_{(\xi_1,\xi_2)}(\widetilde{\pi}) \sim \text{Gamma}(\xi_1,\xi_2)|_{\widetilde{\pi}=-\ln(\pi)}.$$

Then, the prior distribution $p_{(\xi_1,\xi_2)}(\widetilde{\pi})$ is given by

$$p_{(\xi_1,\xi_2)}(\widetilde{\pi}) = \frac{\xi_2^{\xi_1}}{\Gamma(\xi_1)} \widetilde{\pi}^{\xi_1-1} \exp(-\widetilde{\pi}\xi_2),$$

where $\mathrm{B}(\cdot,\cdot)$ is the beta function. The posterior distribution $p(\widetilde{\pi}|\underline{x})$ of the parameter is defined as $p(\widetilde{\pi}|\underline{x}) \propto \ell_n(\widetilde{\pi})$ $\times p_{(\xi_1,\xi_2)}(\widetilde{\pi})$, where $\underline{x} = (x_1,\ldots,x_n)$. Under SELF, the Bayesian estimators of $\widetilde{\pi}$ are the means of their marginal posteriors. They are defined as

$$\widehat{\widetilde{\pi}}_{(\text{Bayesian})} = \int \widetilde{\pi} p(\widetilde{\pi}|\underline{x})\, d\widetilde{\pi}.$$

Because the Bayesian estimates cannot be obtained using the above formula, numerical approximations are required. We propose the use of Markov chain Monte Carlo (MCMC) techniques, namely Gibbs sampler and M-H algorithm (see Cai [4], Chib and Greenberg [7] and Korkmaz et al. [20] for more details). Since the conditional posteriors of the parameters $\widetilde{\pi}$ cannot be obtained in any standard forms, using a hybrid MCMC for drawing sample from the marginal posterior of the parameters is suggested. The theory beyond the Bayes estimators can be found in [16] and [6]. Then, the full conditional posteriors of $\widetilde{\pi}$ can be easily derived. The simulation algorithm is given by:

1. Provide the initial values, say $\widetilde{\pi}$, then at $i^{\text{th}}$ stage;

2. Using M-H algorithm, generate

$$\widetilde{\pi}_{(i)} \sim p_2\left(\widetilde{\pi}_{(i)}|\widetilde{\pi}_{(i-1)}, \underline{x}\right);$$

3. Repeat steps 1 and 2, 100000 times to obtain the sample of size $\mathbf{M}$ from the corresponding posteriors of interest. Obtain the Bayesian estimates of $\widetilde{\pi}$ using the following formulae

$$\widehat{\widetilde{\pi}}_{(\text{Bayesian})} = \frac{1}{\mathbf{M}-\mathbf{M}_0}\sum_{h=\mathbf{M}_0+1}^{\mathbf{M}} \widetilde{\pi}^{[h]},$$

respectively, where $\mathbf{M}_0(\approx 50000)$ is the burn-in period of the generated MCMC.

## 6. Simulations for comparing Bayesian and non-Bayesian estimation methods

To assess and compare the performance of non-Bayesian and Bayesian estimations, numerical MCMC simulation studies are performed. The numerical assessment is performed based on the mean squared errors (MSEs). First, we generated 1000 samples of the DIB distribution, where $n = 50$, 150, 300 and 500. The MSEs are obtained and listed in Tables 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11. Ten combinations of initial values are considered. Based on Tables 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11, we note that all methods perform well. All estimation methods improve their performance as $n \to \infty$.

Table 2. MSEs for for comparing methods | $\pi_0 = 0.5$.

| $n$ | MLE | OLS | WLS | Bayesian |
|---|---|---|---|---|
| 50 | 0.01793 | 0.01748 | 0.01621 | 0.01763 |
| 100 | 0.00897 | 0.00820 | 0.00873 | 0.00896 |
| 200 | 0.00429 | 0.00467 | 0.00444 | 0.00402 |
| 300 | 0.00283 | 0.00283 | 0.00303 | 0.00328 |

Table 3. MSEs for for comparing methods | $\pi_0 = 0.9$.

| $n$ | MLE | OLS | WLS | Bayesian |
|---|---|---|---|---|
| 50 | 0.03676 | 0.03267 | 0.04099 | 0.03421 |
| 100 | 0.01972 | 0.01746 | 0.01790 | 0.01843 |
| 200 | 0.00841 | 0.00973 | 0.00863 | 0.00814 |
| 300 | 0.00558 | 0.00580 | 0.00611 | 0.00590 |

Table 4. MSEs for for comparing methods | $\pi_0 = 1.2$.

| $n$ | MLE | OLS | WLS | Bayesian |
|---|---|---|---|---|
| 50 | 0.05534 | 0.05237 | 0.05947 | 0.04594 |
| 100 | 0.02797 | 0.02517 | 0.02526 | 0.02812 |
| 200 | 0.01178 | 0.01331 | 0.01288 | 0.01206 |
| 300 | 0.00860 | 0.00847 | 0.00916 | 0.00928 |

Table 5. MSEs for for comparing methods | $\pi_0 = 1.5$.

| $n$ | MLE | OLS | WLS | Bayesian |
|---|---|---|---|---|
| 50 | 0.07254 | 0.07908 | 0.07472 | 0.07003 |
| 100 | 0.03501 | 0.03495 | 0.03567 | 0.03770 |
| 200 | 0.01672 | 0.01879 | 0.02004 | 0.02869 |
| 300 | 0.01170 | 0.01164 | 0.01298 | 0.01335 |

Table 6. MSEs for for comparing methods | $\pi_0 = 2$.

| $n$ | MLE | OLS | WLS | Bayesian |
|---|---|---|---|---|
| 50 | 0.11304 | 0.12914 | 0.11705 | 0.10569 |
| 100 | 0.05421 | 0.05311 | 0.06190 | 0.06064 |
| 200 | 0.02676 | 0.02937 | 0.02842 | 0.03144 |
| 300 | 0.01848 | 0.02073 | 0.02056 | 0.16231 |

Table 7. MSEs for for comparing methods | $\pi_0 = 2.5$.

| $n$ | MLE | OLS | WLS | Bayesian |
|---|---|---|---|---|
| 50 | 0.16339 | 0.18367 | 0.15629 | 0.10376 |
| 100 | 0.07900 | 0.08475 | 0.08324 | 0.07924 |
| 200 | 0.04143 | 0.04313 | 0.04707 | 0.05397 |
| 300 | 0.02550 | 0.03009 | 0.02056 | 0.04015 |

Table 8. MSEs for for comparing methods | $\pi_0 = 4$.

| $n$ | MLE | OLS | WLS | Bayesian |
|---|---|---|---|---|
| 50 | 0.37701 | 0.47502 | 0.36365 | 0.43232 |
| 100 | 0.18426 | 0.21137 | 0.21528 | 0.21894 |
| 200 | 0.08837 | 0.11353 | 0.10920 | 0.12608 |
| 300 | 0.05371 | 0.07778 | 0.08354 | 0.04283 |

Table 9. MSEs for for comparing methods | $\pi_0 = 8$.

| $n$ | MLE | OLS | WLS | Bayesian |
|---|---|---|---|---|
| 50 | 1.51837 | 1.73538 | 1.44175 | 1.95487 |
| 100 | 0.69834 | 0.90560 | 0.70421 | 0.44893 |
| 200 | 0.34222 | 0.43632 | 0.37434 | 0.12406 |
| 300 | 0.20064 | 0.29674 | 0.27200 | 0.12169 |

Table 10. MSEs for for comparing methods | $\pi_0 = 15$.

| $n$ | MLE | OLS | WLS | Bayesian |
|---|---|---|---|---|
| 50 | 4.65915 | 6.02921 | 5.01556 | 1.79753 |
| 100 | 2.47947 | 2.87890 | 2.57184 | 0.75391 |
| 200 | 1.10139 | 1.63056 | 1.32381 | 0.31169 |
| 300 | 0.75640 | 0.97191 | 0.94531 | 0.29139 |

Table 11. MSEs for for comparing methods | $\pi_0 = 25$.

| $n$ | MLE | OLS | WLS | Bayesian |
|---|---|---|---|---|
| 50 | 13.39616 | 16.36931 | 14.67991 | 2.64481 |
| 100 | 6.55614 | 7.82160 | 6.85395 | 1.86572 |
| 200 | 3.20947 | 4.10796 | 3.39501 | 1.42769 |
| 300 | 2.20396 | 2.69540 | 2.48062 | 0.84161 |

## 7. Application for comparing Bayesian and non-Bayesian estimation methods

In this section, two examples of real data sets are introduced and analyzed for comparing the Bayesian and non-Bayesian estimation methods. We consider the Kolmogorov-Smirnov test ($ks$) and P-value $(pv)$ statistic for comparing the Bayesian and non-Bayesian estimation methods.

### *7.1. Application 1: Carious teeth data*

The first data set consists of the number of carious teeth among the four deciduous molars. The sample size is 100. Figure 3 gives the Kaplan–Meier survival plots for carious teeth data. Table 12 shows the estimations for the carious teeth data under Bayesian and non-Bayesian methods, Kolmogorov-Smirnov test ($ks$) and P-value ($pv$) statistics. Based on Table 12, the Bayesian method is the best model with lowest $ks = 4.07325$ and biggest $pv = 0.25366$.

Table 12. Comparison of the methods for based on the carious teeth data, with rank under brackets.

| Method↓ estimations and statistics→ | $\widehat{\pi}$ | $ks$ | $pv$ |
|---|---|---|---|
| MLE[2] | 0.607838490 | 4.73624 | 0.19216 |
| OLS[3] | 0.606884929 | 4.76582 | 0.18977 |
| WLS[4] | 0.573977811 | 5.92970 | 0.11508 |
| Bayesian[1] | 0.631423018 | 4.07325 | 0.25366 |

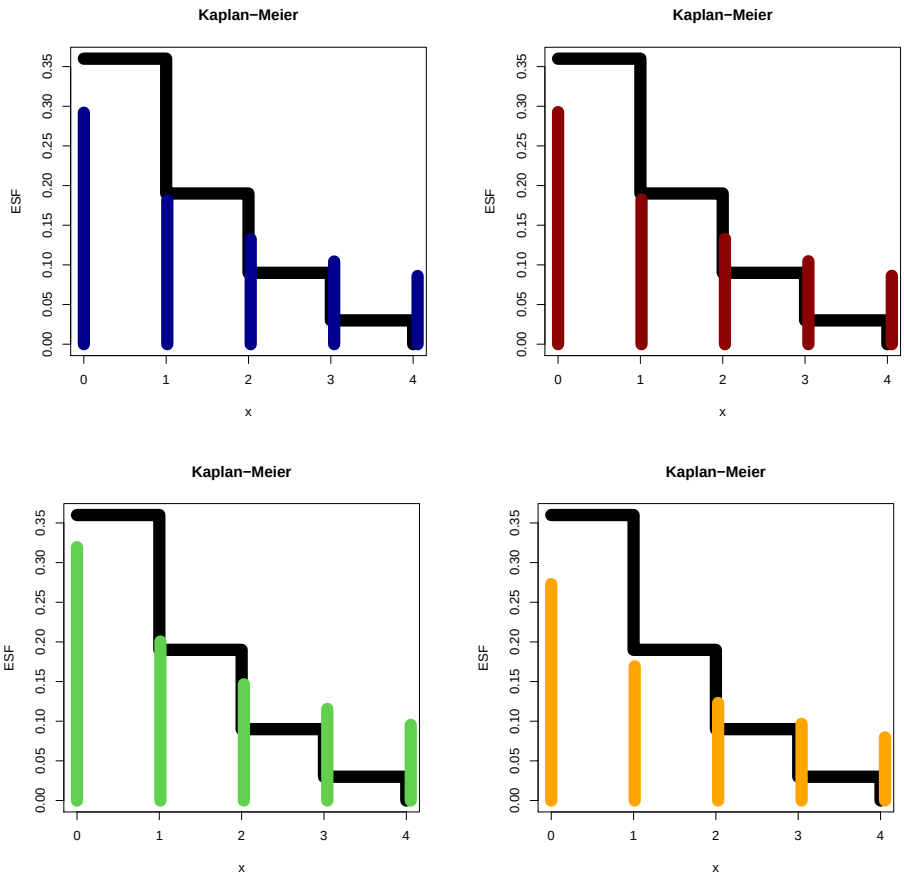


Figure 3. Kaplan–Meier survival plots for the carious teeth data

### *7.2. Application 2: Counts of cysts of kidneys data*

Due to Chan et al. [8], the second data set represents the counts of cysts of corticosteroid-induced kidney dysmorphogenesis which is associated with deregulated expression of known cystogenic molecules as well as Indian hedgehog. Figure 4 gives the Kaplan–Meier survival plots for counts of cysts of kidneys data. Table 13 gives the estimations under Bayesian and non-Bayesian methods, $ks$ and $pv$ statistics for counts of cysts of kidneys data. According to Table 13, the Bayesian method is the best model with lowest $ks = 3.34854$ and biggest $pv = 0.50128$.

Table 13. Comparison of the methods for based on the kidneys data, with rank under brackets.

| Method↓ estimations and statistics→ | $\widehat{\pi}$ | $ks$ | $pv$ |
|---|---|---|---|
| MLE[3] | 0.7523056604 | 3.45680 | 0.48448 |
| OLS[2] | 0.7570113907 | 3.36403 | 0.49885 |
| WLS[4] | 0.7384775217 | 3.75382 | 0.44035 |
| Bayesian[1] | 0.7578183582 | 3.34854 | 0.50128 |

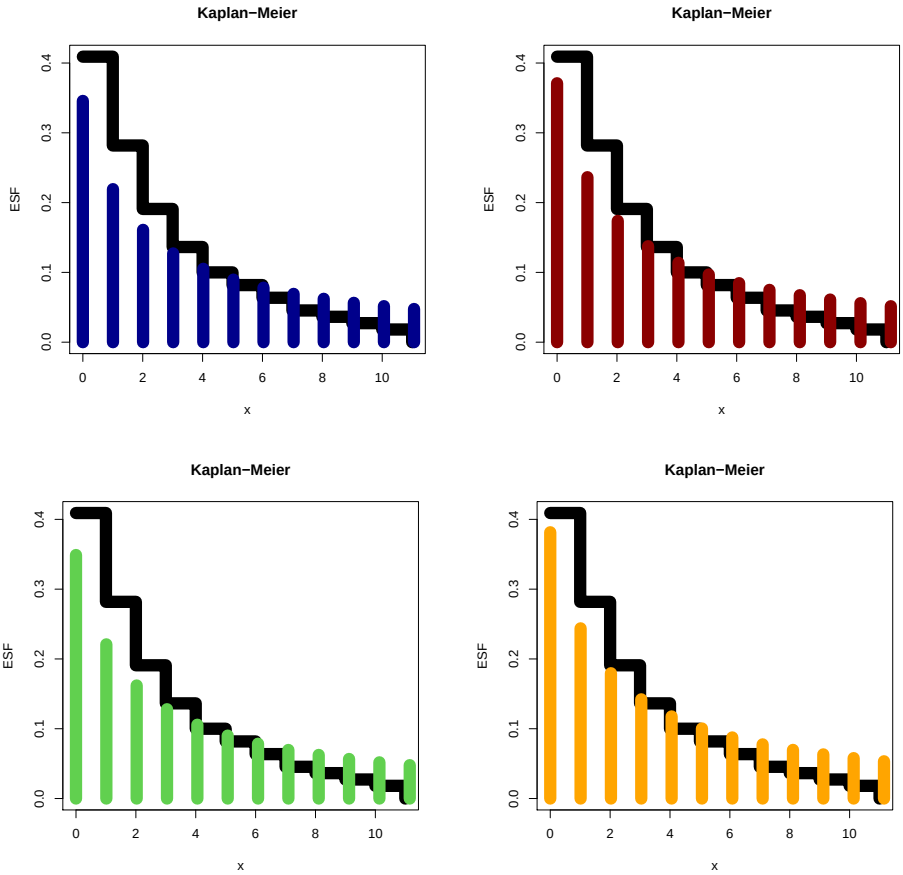


Figure 4. Kaplan–Meier survival plots for the counts of cysts of kidneys data.

## 8. Application for comparing models and testing of hypothesis

### *8.1. Application 1: Carious teeth data*

Using the carious teeth data, the fits of the DIB model is compared with some competitive models such as DBH, Poisson (P), DR, DLi, DIR, and PLi models. Table 14 gives the observed frequency (OF), expected frequency (EF), MLEs, Chi-squared test $\left(\chi^2_V\right)$ and $pv$ for these models. For exploring the carious teeth data graphically, we present some useful plots such as box plot, quantile-quantile (Q-Q) plot, total time in test (TTT) plot and the initial histogram. Based on Figure 5 (top left panel and top right panel), it is noted that the carious teeth data have some extremes. Based on Figure 5 (bottom left panel), it is noted that the underlying HRF of the data is decreasing. From Figure 5 (bottom right panel), we see that the data are right skewed. Figure 6 gives the fitted PMF, estimated HRF (EHRF), estimated SF (ESF) (Kaplan-Meier plot) and estimated CDF (ECDF). Based on Table 14 and Figure 6, the DIB model provides the best fits against all competitive models with $\chi^2_V = 4.73624$, $pv = 0.19216$ and $df = 3$; the $pv$ is the best. As expected, the DIB distribution is useful in modeling the over-dispersed count data. By comparing the initial HRF (Figure 5 bottom left panel) of the carious teeth data with the estimated HRF (Figure 6 top right panel) using the estimated parameter, we conclude that the both shapes means monotonically decreasing HRF.

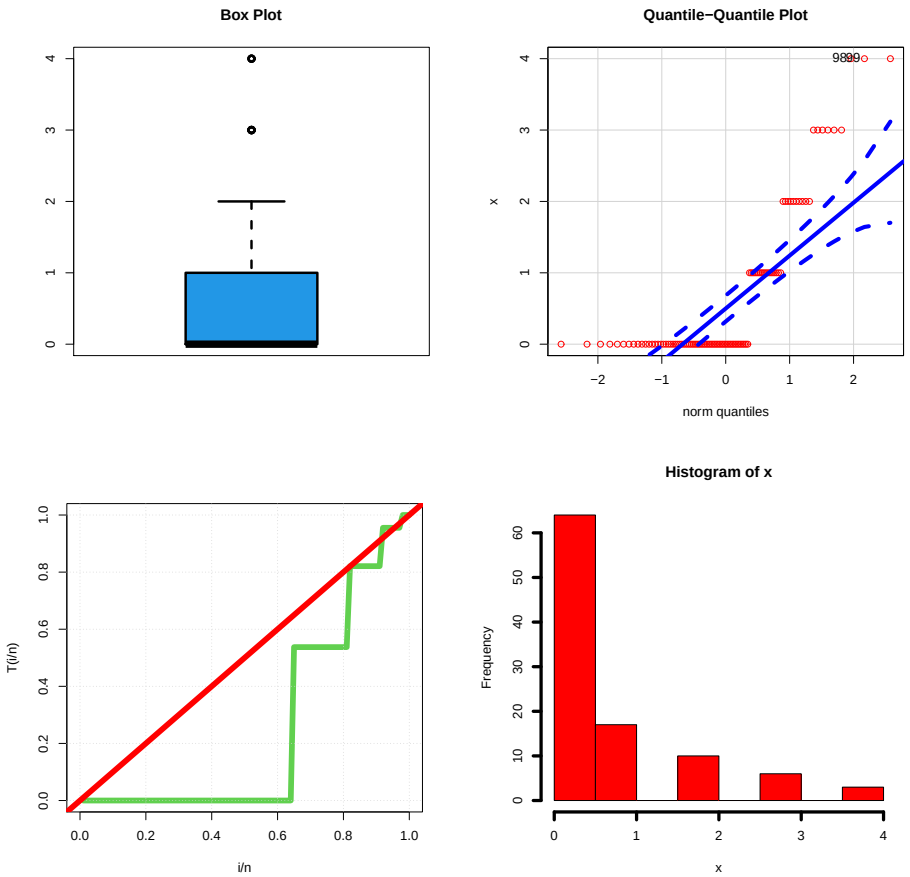


Figure 5. Box plot, Q-Q plot, TTT plots and histogram for the carious teeth data.

Table 14. OF, EF, MLE, $\chi^2_V$ and $pv$ for the carious teeth data.

| $X$ | OF | EF | | | | | |
|---|---|---|---|---|---|---|---|
| | | **DIB** | P | DR | DLi | DIR | PLi |
| 0 | 64 | **66.45** | 51.17 | 33.50 | 57.13 | 62.50 | 37.50 |
| 1 | 17 | **16.60** | 34.28 | 46.94 | 26.88 | 26.41 | 25.00 |
| 2 | 10 | **7.80** | 11.49 | 17.01 | 10.45 | 5.990 | 15.63 |
| $\geq 3$ | 9 | **9.15** | 3.060 | 2.550 | 5.450 | 5.100 | 21.87 |
| $\sum$ | 100 | **100** | 100 | 100 | 100 | 100 | 100 |
| $\widehat{\pi}$ | MLE | **0.60784** | 0.670 | 0.665 | 0.625 | 0.625 | 1.998 |
| $\chi^2_V$ | | **4.73624** | 23.65 | 66.7 | 6.638 | 9.056 | 30.899 |
| $df$ | | **3** | 2 | 2 | 2 | 2 | 2 |
| $pv$ | | **0.19216** | <0.001 | <0.001 | 0.036 | 0.011 | <0.001 |

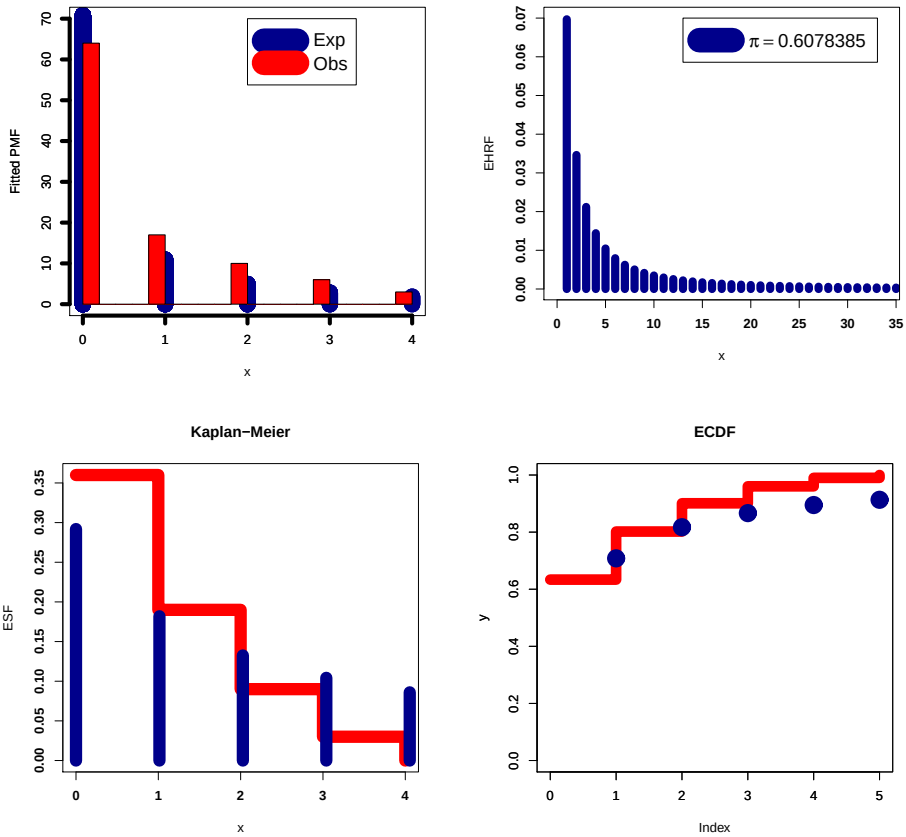


Figure 6. The fitted PMF, EHRF, ESF and ECDF for the carious teeth data.

### *8.2. Application 2: Counts of cysts of kidneys data*

For the counts of cysts of kidneys data set, the same methodology is employed; we compare the fits of the DIB model with the DLi-II, DIW, DR, DIR, DLi, PLi, P and geometric (Gc) models. Table 15 gives the observed frequency (OF), expected frequency (EF), MLEs, $\chi^2_V$ and $pv$ for the competitive models. For exploring counts of cysts of kidneys data graphically, we present the box plot, the Q-Q plot, the TTT plot and the initial histogram for the data. Based on Figure 7 (top left panel and top right panel), it is noted that the counts of cysts of kidneys data have some extremes. According to Figure 7 (bottom left panel), it is noted that the underlying HRF is decreasing. Based on Figure 7 (bottom right panel), it is noted that the counts of cysts of kidneys data are heavy tail right skewed data. Figure 8 gives the fitted PMF, EHRF, ESF and ECDF. Based on Table 15 and Figure 8, the DIB model provides the best fits against all the competitive models with $\chi^2_V = 3.45680$, $pv = 0.48448$ and $df = 4$; the $pv$ is the best. By comparing the initial HRF (Figure 7 bottom left panel) of the counts of cysts of kidneys data with the estimated HRF (Figure 8 top right panel) using the estimated parameter, we conclude that the both shapes means monotonically decreasing HRF.

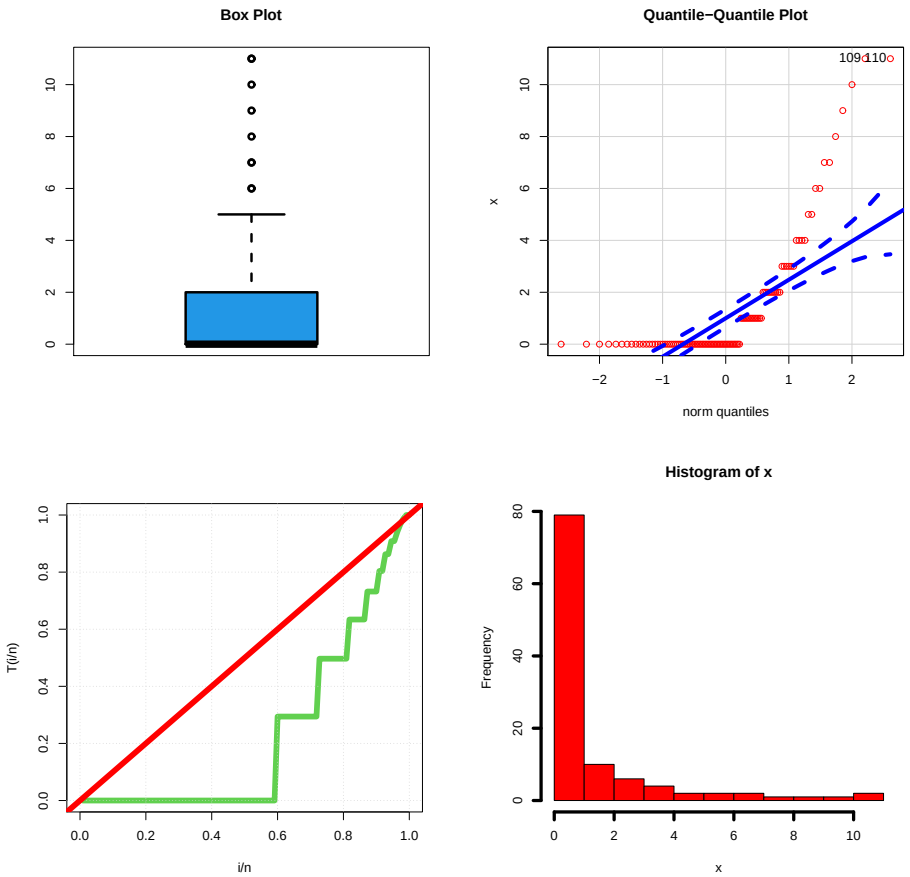


Figure 7. Box plot, Q-Q plot, TTT plots and histogram for the counts of cysts of kidneys data

Table 15. OF, EF, MLE, $\chi^2_V$ and $pv$ for the counts of cysts of kidneys data.

| $X$ | OF | EF | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **DIB** | DLi-II | DIW | Gc | P | DR | DIR | DLi | PLi |
| 0 | 65 | **66.30** | 46.03 | 63.91 | 45.98 | 27.42 | 11.00 | 60.94 | 40.25 | 44.14 |
| 1 | 14 | **16.78** | 26.77 | 20.70 | 26.67 | 38.08 | 26.83 | 33.96 | 29.83 | 28.00 |
| 2 | 10 | **8.51** | 15.57 | 8.05 | 15.57 | 26.47 | 29.55 | 8.11 | 18.36 | 17.70 |
| 3 | 6 | **5.41** | 9.05 | 4.23 | 9.060 | 12.26 | 22.23 | 3.00 | 10.35 | 9.57 |
| 4 | 4 | **3.9** | 5.27 | 2.60 | 5.280 | 4.26 | 12.49 | 1.42 | 5.53 | 5.34 |
| 5 | 2 | **3.05** | 3.06 | 1.75 | 3.070 | 1.18 | 5.42 | 0.87 | 2.86 | 2.92 |
| 6 | 2 | **1.63** | 1.78 | 1.26 | 1.79 | 0.27 | 1.85 | 0.47 | 1.44 | 1.57 |
| 7 | 2 | **1.29** | 1.04 | 0.95 | 1.04 | 0.05 | 0.52 | 0.31 | 0.71 | 0.84 |
| 8 | 1 | **0.95** | 0.60 | 0.74 | 0.62 | 0.01 | 0.11 | 0.21 | 0.35 | 0.44 |
| 9 | 1 | **0.87** | 0.35 | 0.59 | 0.35 | 0.00 | 0.02 | 0.15 | 0.17 | 0.23 |
| 10 | 1 | **0.74** | 0.20 | 0.48 | 0.21 | 0.00 | 0.00 | 0.11 | 0.08 | 0.12 |
| 11 | 2 | **0.57** | 0.28 | 4.74 | 0.28 | 0.00 | 0.00 | 0.54 | 0.07 | 0.13 |
| $\sum$ | 110 | **110** | 110 | 110 | 110 | 110 | 110 | 110 | 110 | 110 |
| $\widehat{\pi}$ | MLE | **0.75231** | 0.581 | 0.581 | 0.582 | 1.390 | 0.901 | 0.554 | 0.436 | 1.087 |
| $\widehat{\theta}$ | MLE | **-** | 0.001 | 1.049 | - | - | - | - | - | - |
| $\chi^2_V$ | | **3.45680** | 22.89 | 24.135 | 22.84 | 294.10 | 321.07 | 51.047 | 43.48 | 31.151 |
| $df$ | | **4** | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 |
| $pv$ | | **0.48448** | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |

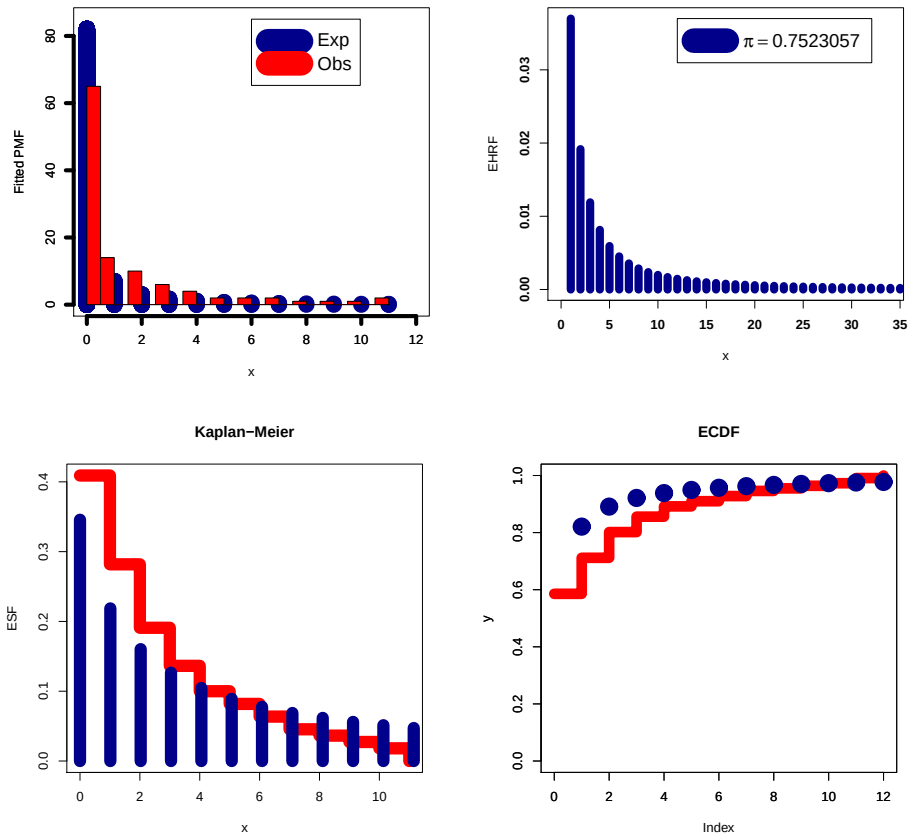


Figure 8. The fitted PMF, EHRF, ESF and ECDF for the counts of cysts of kidneys data.

## 9. Conclusions

A new one-parameter decreasing failure rate discrete distribution, called the discrete inverse Burr (DIB) distribution, is defined and studied. The probability mass function of the new distribution can be "unimodal and right skewed" with various shapes. Also, the DIB distribution appears to be a heavy-tailed discrete distribution with an infinite mean. Some of its relevant properties are discussed. Certain characterizations of the DIB distribution involving (i) the conditional expectation of a certain function of the random variable and (ii) the reversed hazard function are presented. The new distribution could be useful in modeling the over-dispersed count data. Different Bayesian and non-Bayesian estimation methods (maximum likelihood, ordinary and weighted least squared) are described and compared using MCMC simulations and two real data applications. The Bayesian estimation is considered under the squared error loss function. The DIB model is applied for modeling carious teeth data and counts of cysts of kidneys data sets. The results show that it provides the best fits compared to many well-known competitive discrete models.

## Acknowledgement

We thank the two reviewers for their helpful comments on the former version of the article.

## REFERENCES


1. M. Aboraya, M. H. Yousof, G. G. Hamedani, and M. Ibrahim. *A new family of discrete distributions with mathematical properties, characterizations, Bayesian and non-Bayesian estimation methods*, Mathematics, vol. 8, no. 10, 1648, 2020.
2. M. Bebbington, C. D. Lai, M. Wellington, and R. Zitikis. *The discrete additive Weibull distribution: A bathtub-shaped hazard for discontinuous failure data*, Reliability Engineering & System Safety, vol. 106, pp. 37–44, 2012.
3. W. Bodhisuwan, and S. Sangpoom. *The discrete weighted Lindley distribution*, In Proceedings of the International Conference on Mathematics, Statistics, and Their Applications, Banda Aceh, Indonesia, 4–6 October 2016, 2016.
4. L. Cai. *Metropolis-Hastings Robbins-Monro algorithm for confirmatory item factor analysis*, Journal of Educational and Behavioral Statistics, vol. 35, no. 3, pp. 307–335, 2010.
5. G. Casella, and R. L. Berger, *Statistical Inference*, Brooks/Cole Publishing Company, California, 1990.
6. M. T. Chao, *The asymptotic behavior of Bayes' estimators*, The Annals of Mathematical Statistics, vol. 41, no. 2, pp. 601–608, 1970.
7. S. Chib, and E. Greenberg. *Understanding the metropolis-hastings algorithm*, The american statistician, vol. 49, no. 4, pp. 327–335, 1995.
8. S. Chan, P. R. Riley, K. L. Price, F. McElduff, and P. J. Winyard. *Corticosteroid-induced kidney dysmorphogenesis is associated with deregulated expression of known cystogenic molecules, as well as Indian hedgehog*, The American Journal of Physiology - Renal Physiology, vol. 298, pp. 346–356, 2009.
9. M. El-Morshedy, M. S. Eliwa, and H. Nagy. *A new two-parameter exponentiated discrete Lindley distribution: Properties, estimation and applications*, Journal of Applied Statistics, vol. 47, pp. 354–375, 2020.
10. M. El-Morshedy, M. S. Eliwa, and E. Altun. *Discrete Burr-Hatke distribution with properties, estimation methods and regression model*, IEEE Access, vol. 8, pp. 74359–74370, 2020.
11. E. Gomez-Déniz. *Another generalization of the geometric distribution*, Test, vol. 19, no. 2, pp. 399–415, 2010.
12. E. Gomez-Déniz, and E. Caldern-Ojeda. *The discrete Lindley distribution: Properties and applications*, Journal of Statistical Computation and Simulation, vol. 81, no. 11, pp. 1405–1416, 2011.
13. T. Hussain, and M. Ahmad. *Discrete inverse Rayleigh distribution*, Pakistan Journal of Statistics, vol. 30, no. 2, pp. 203–222, 2014.
14. T. Hussain, M. Aslam, and M. A. Ahmad. *Two parameter discrete Lindley distribution*, Revista Colombiana de Estadística, vol. 39, no. 1, pp. 45–61, 2016.
15. M. Ibrahim, M. M. Ali, and H. M. Yousof, *The discrete analogue of the Weibull G family: properties, different applications, Bayesian and non-Bayesian estimation methods*, Annals of Data Science, forthcoming, 2021.
16. I. A. Ibragimov. *Some limit theorems for stationary processes*, Theory of Probability & Its Applications, vol. 7, no. 4, pp. 349-382, 1962.
17. A. M. Jazi, D. C. Lai, and H. M. Alamatsaz. *Inverse Weibull distribution and estimation of its parameters*, Statistical Methodology, vol. 7, no. 2, pp. 121–132, 2010.
18. A. W. Kemp. *Classes of discrete lifetime distributions*, Communications in Statistics - Theory and Methods, vol. 33, no. 12, pp. 3069–3093, 2004.
19. A. W. Kemp. *The discrete half-normal distribution*, In: Advances in mathematical and statistical modeling, Birkhauser, Basel, pp. 353–365, 2008.
20. M. C. Korkmaz, H. M. Yousof, M. Rasekhi, and G. G. Hamedani. *The Odd Lindley Burr XII Model: Bayesian Analysis, Classical Inference and Characterizations*, Journal of Data Science, vol. 16, no. 2, pp. 327–353, 2018.
21. H. Krishna, and P. S. Pundir. *Discrete Burr and discrete Pareto distributions*, Statistical Methodology, vol. 6, no. 2, pp. 177–188, 2009.
22. C. Kumar, Y. M. Tripathi, and M. K. Rastogi. *On a discrete analogue of linear failure rate distribution* American journal of mathematical and management sciences, vol. 36, no. 3, pp. 229–246, 2017.
23. T. Nakagawa, and S. Osaki. *The discrete Weibull distribution*, IEEE Transactions on Reliability, vol. 24, no. 5, pp. 300–301, 1975.
24. V. Nekoukhou, M. H. Alamatsaz, and H. Bidram. *Discrete generalized exponential distribution of a second type*, Statistics, vol. 47, no. 4, pp. 876–887, 2013.
25. B. A. Para, and T. R. Jan. *Discrete version of log-logistic distribution and its applications in genetics*, International Journal of Modern Mathematical Sciences, vol. 14, pp. 407–422, 2016.
26. B. A. Para, and T. R. Jan. *On discrete three-parameter Burr type XII and discrete Lomax distributions and their applications to model count data from medical science*, Biometrics & Biostatistics International Journal, vol. 4, pp. 1–15, 2016.
27. S. D. Poisson. *Probabilité des Jugements en Matiére Criminelle et en Matiére Civile, Précédées des RéGles Génerales du Calcul des Probabilitiés*, Bachelier: Paris, France, pp. 206–207, 1837.
28. D. Roy. *Discrete Rayleigh distribution*, IEEE Transactions on Reliability, vol. 53, pp. 255–260, 2004.
29. M. Sankaran. *The discrete poisson-lindley distribution*, Biometrics, vol. 26, no. 1, 145-149, 1970.
30. W. E. Stein, and R. A. Dattero. *A new discrete Weibull distribution*, IEEE Transactions on Reliability, vol. R-33, no. 2, pp. 196–197, 1984.
31. H. M. Yousof, C. Chesneau, G. G. Hamedani, and M. Ibrahim. *A new discrete distribution: properties, characterizations, modeling real count data, Bayesian and non-Bayesian estimations*, Statistica, vol. 81, no. 2, pp. 135-162, 2021.