2021

A New Discrete Distribution: Properties, Characterizations, Modeling Real Count Data, Bayesian and Non-Bayesian Estimations

Haitham M. Yousof
Benha University

Christophe Chesneau
University of Caen Normandie

Gholamhossein Hamedani
Marquette University, gholamhoss.hamedani@marquette.edu

Mohamed Ibrahim
Damietta University

STATISTICA, anno LXXXI, n. 2, 2021

A NEW DISCRETE DISTRIBUTION: PROPERTIES, CHARACTERIZATIONS, MODELING REAL COUNT DATA, BAYESIAN AND NON-BAYESIAN ESTIMATIONS

Haitham Mosad Yousof
Department of Statistics, Mathematics and Insurance, Benha University, Egypt
Christophe Chesneau [1]
LMNO, University of Caen Normandie, Caen 14032, France
Gholamhossein Hamedani
Department of Mathematical and Statistical Sciences, Marquette University, USA
Mohamed Ibrahim
Department of Applied, Mathematical and Actuarial Statistics, Faculty of Commerce, Damietta University, Damietta, Egypt

1. Introduction

Discretization of existing continuous probability distributions has received wide attention in recent years. In this article, we define a new two–parameter discrete distribution which includes the discrete Burr-Hatke (DBH) distribution (see El-Morshedy *et al.*, 2020a). The new distribution, thus, can be considered as a new generalization of the Burr-Hatke distribution. In the statistical literature, many discrete versions of continuous distributions have been defined, studied and used in the modeling of count data, such as the poisson-Lindley (PLi) distribution (see Sankaran, 1970), discrete Weibull (DW) distribution (see Nakagawa and Osaki, 1975), discrete half-normal (DHN) distribution (see Kemp, 2008), discrete Rayleigh (DR) distribution (see Roy, 2004), discrete Pareto (DPa) distribution (see Krishna and Pundir, 2009), discrete inverse Weibull (DIW) distribution (see Jazi *et al.*, 2010), generalized geometric (GGc) distribution (see Gomez-Déniz, 2010), discrete Lindley (DLi) distribution (see Gomez-Déniz and Caldern-Ojeda, 2011), two-parameter discrete generalized exponential (DGE-II) distribution (see Nekoukhou *et al.*, 2013), discrete inverse Rayleigh (DR) distribution (see Hussain and Ahmad, 2014), discrete log-logistic (DLL) distribution (see Para and Jan, 2016a), discrete

[1] Corresponding Author. E-mail: christophe.chesneau@unicaen.fr

Lomax (DLx) distribution (see Para and Jan, 2016b), discrete Burr type XII (DBXII) distribution (see Para and Jan, 2016b), two-parameter discrete Lindley (DLi-II) distribution (see Hussain *et al.*, 2016), discrete linear failure rate (DLFR) distribution (see Kumar *et al.*, 2017) and exponentiated discrete Lindley (EDLi) distribution (see El-Morshedy *et al.*, 2020b). A discrete family of distributions based on the Rayleigh distribution, called the discrete Rayleigh-G (DRG) family, was recently introduced along with many new discrete sub-distributions, such as the DR Weibull (DRW), DR exponential (DRE), DR log-logistic (DRLL), DR Lomax (DRLx), DR Rayleigh (DRR), DR Burr XII (DRBXII), DR Fréchet (DRF), DR inverse Rayleigh (DRIR), DR inverse exponential (DRIE), DR inverse Lomax (DR ILx), DR half-logistic (DRHL), DR Gumbel (DRGu), DR Lindley (DRLi), DR Nadarajah-Haghighi (DRNH), DR Gompertz (DRGz), DR Dagum (DRD), DR inverse flexible W (DRIFW), DR Burr X (DRBX) and DR inverse Gompertz (DRIGz) distributions (see Aboraya *et al.*, 2020).

Recently, Yousof *et al.* (2018) defined and studied a new continuous family of distributions based on the Burr-Hatke (BH) distribution. A random variable (RV) Z is said to have the generalized Burr-Hatke (GBH) distribution if its cumulative distribution function (CDF) is given by

$$F_{\alpha,\beta}(z) = P(Z \leq z) = 1 - \frac{1}{z+1} \exp\left[-(\alpha z)^{\beta}\right]\Big|_{(z>0,\text{ and } \alpha,\beta>0)}. \tag{1}$$

For $\beta = 1$, the GBH distribution reduces to the one parameter BH distribution first introduced by Maniu and Voda (2008). Then, the CDF of the discrete generalized Burr-Hatke (DGBH) distribution can be expressed as

$$\mathbf{F}_{\pi,\beta}(z) = 1 - \frac{\pi^{(z+1)^{\beta}}}{z+2} \quad \Big|_{(0<\pi<1 \text{ and } z\in\mathbb{N})}, \tag{2}$$

where $\mathbb{N} = \{0, 1, 2, \ldots\}$. For $\beta = 1$, the DGBH distribution reduces to the DBH distribution as introduced by El-Morshedy *et al.* (2020a). Moreover, the following first-order dominance holds: for $\beta \geq 1$, we have $\mathbf{F}_{\pi,1}(z) \leq \mathbf{F}_{\pi,\beta}(z)$, and the reverse inequality holds for $\beta < 1$. In this sense, the parameter β makes of the former DBH distribution more pliant. Further role of this parameter will be discussed later.

The corresponding reliability function (RF) due to Steutel and van Harn (2004) can be written as

$$\mathbf{S}_{\pi,\beta}(z) = \overline{\mathbf{F}}_{\pi,\beta}(z) = \frac{\pi^{(z+1)^{\beta}}}{z+2} \quad \Big|_{(0<\pi<1 \text{ and } z\in\mathbb{N})}. \tag{3}$$

The probability mass function (PMF) of the DGBH distribution corresponding to (2) and (3) can be expressed as $\mathscr{F}_{\pi,\beta}(0) = \mathbf{F}_{\pi,\beta}(0) = 1 - \pi/2$, and

$$\mathscr{F}_{\pi,\beta}(z) = \mathbf{S}_{\pi,\beta}(z-1) - \mathbf{S}_{\pi,\beta}(z) \quad \Big|_{(0<\pi<1 \text{ and } z\in\mathbb{N}^*=\mathbb{N}/\{0\})}. \tag{4}$$

That is

$$\mathscr{F}_{\pi,\beta}(z) = \frac{\pi^{z^{\beta}}}{z+1} - \frac{\pi^{(z+1)^{\beta}}}{z+2} \quad \Big|_{(0<\pi<1 \text{ and } z\in\mathbb{N}^*)}. \tag{5}$$

After a graphical analysis, one can show that the sequence $u_z = \mathscr{F}_{\pi,\beta}(z+1)/\mathscr{F}_{\pi,\beta}(z)$ can be increasing, decreasing or non-monotonic, depending on π and β, contrary to the DBH distribution corresponding to $\beta = 1$ for which u_z is increasing, implying that the related PMF is always decreasing function in z. This illustrates the flexibility of the proposed two-parameter DGBH distribution and the importance of the parameter β in this regard. The hazard rate function (HRF) can be written as

$$\mathscr{H}_{\pi,\beta}(z) = \frac{\mathscr{F}_{\pi,\beta}(z)}{\mathbf{S}_{\pi,\beta}(z-1)} = 1 - \frac{z+1}{z+2}\pi^{(z+1)^\beta - z^\beta} \quad |_{(0<\pi<1 \text{ and } z\in\mathbb{N})}. \tag{6}$$

Now, let Z be a RV following the DGBH distribution. Then, the probability generating function (PGF) of Z is given by

$$\pi(s) = \mathbb{E}(s^Z) = 1 + (s-1)\sum_{z=1}^{\infty} s^{z-1}\frac{\pi^{z^\beta}}{z+1} \quad |_{(0<s<1)}. \tag{7}$$

The r-th ordinary moments of Z is given by

$$\mathbb{E}(Z^r) = \sum_{z=1}^{\infty}[z^r - (z-1)^r]\frac{\pi^{z^\beta}}{z+1}. \tag{8}$$

Therefore, the mean and variance of the DGBH distribution do not have analytical forms. We can, however, express them as the following series expansions:

$$\mathbb{E}(Z) = \sum_{z=1}^{\infty}\frac{\pi^{z^\beta}}{z+1}, \tag{9}$$

and

$$\mathbb{V}(Z) = \sum_{z=1}^{\infty}(2z-1)\frac{\pi^{z^\beta}}{z+1} - \left(\sum_{z=1}^{\infty}\frac{\pi^{z^\beta}}{z+1}\right)^2. \tag{10}$$

Based on (9) and (10), the index of dispersion is

$$\mathbb{D}(Z) = \frac{1}{\mathbb{E}(Z)}\mathbb{V}(Z). \tag{11}$$

Similarly, we can express the first four moments of Z, allowing us to define the following skewness and kurtosis measures:

$$\mathbb{S}(Z) = \frac{\mathbb{E}(Z^3) - 3\mathbb{E}(Z)\mathbb{E}(Z^2) + 2[\mathbb{E}(Z)]^3}{[\mathbb{V}(Z)]^{3/2}} \tag{12}$$

and

$$\mathbb{K}(Z) = \frac{\mathbb{E}(Z^4) - 4\mathbb{E}(Z)\mathbb{E}(Z^2) + 6[\mathbb{E}(Z)]^2\mathbb{E}(Z^2) - 3[\mathbb{E}(Z)]^4}{[\mathbb{V}(Z)]^2}. \tag{13}$$

All these measures can be determined numerically with the help of any mathematical software. Also, one can remark that $\mathscr{F}_{0.9,4}(1) = 0.3882327 > e^{-1}$, implying that the DGBH distribution is not infinitely divisible (see Steutel and van Harn, 2004,pp. 56).

Now, let us establish some general relations regarding the order statistics of the DGBH distribution. The CDF of the i^{th} order statistic from the DGBH distribution is given by

$$\mathbf{F}_{\pi,\beta,i:n}(z) = \sum_{x=i}^{n} \binom{n}{x} \left[\mathbf{F}_{\pi,\beta}(z)\right]^{x} \left[\mathbf{S}_{\pi,\beta}(z)\right]^{n-x}. \tag{14}$$

Applying the binomial formula to $\left[\mathbf{F}_{\pi,\beta}(z)\right]^{x} = \left[1-\mathbf{S}_{\pi,\beta}(z)\right]^{x}$, we have

$$\begin{aligned}\mathbf{F}_{\pi,\beta,i:n}(z) &= \sum_{x=i}^{n}\sum_{j=0}^{x} \binom{n}{x}\binom{x}{j}(-1)^{j} \left[\mathbf{S}_{\pi,\beta}(z)\right]^{j+n-x} \\ &= \sum_{x=i}^{n}\sum_{j=0}^{x} \binom{n}{x}\binom{x}{j}(-1)^{j} \frac{\pi^{(j+n-x)(z+1)^{\beta}}}{(z+2)^{j+n-x}}, \quad |_{(0<\pi<1 \text{ and } z\in\mathbb{N})}. \end{aligned} \tag{15}$$

Also, the corresponding PMF is obtained as

$$\begin{aligned}\mathscr{F}_{\pi,\beta,i:n}(z) &= \mathbf{F}_{\pi,\beta,i:n}(z) - \mathbf{F}_{\pi,\beta,i:n}(z-1) \\ &= \sum_{x=i}^{n}\sum_{j=0}^{x} \binom{n}{x}\binom{x}{j}(-1)^{j} \left[\frac{\pi^{(j+n-x)(z+1)^{\beta}}}{(z+2)^{j+n-x}} - \frac{\pi^{(j+n-x)z^{\beta}}}{(z+1)^{j+n-x}}\right], \quad |_{(0<\pi<1 \text{ and } z\in\mathbb{N})}. \end{aligned} \tag{16}$$

From this PMF, several measures and functions can be derived, as done for the former DGBH distribution.

2. Graphical and numerical analysis

In this Section, we analyze the effect of adding the new additional parameter β on the PMF, HRF, skewness and kurtosis measures. The superiority of the new DGBH distribution is illustrated as well. Figure 1 gives some plots of the PMF of the DGBH distribution. Figure 2 gives some plots of the HRF of the DGBH distribution. According to Figure 1, it can be seen that the shape of the PMF can be "right skewed" with different shapes, bimodal, and "uniform-PMF". Based on Figure 2, the HRF of the new model can be "monotonically decreasing", "upside down", "monotonically increasing", "upside down increasing", and "upside down-constant-increasing".

The shapes of the PMF and HRF of the DGBH model can also be described analytically but we must deal with some complicated equations from the mathematical point of view. In order to support this claim, if we focus on the mode nature of the DGBH distribtuion only, a mode z is determined by the following inequality: for any $k \in \mathbb{N}$, $\mathscr{F}_{\pi,\beta}(k) \leq \mathscr{F}_{\pi,\beta}(z)$. When the mode differs from $z = 0$, the situation becomes

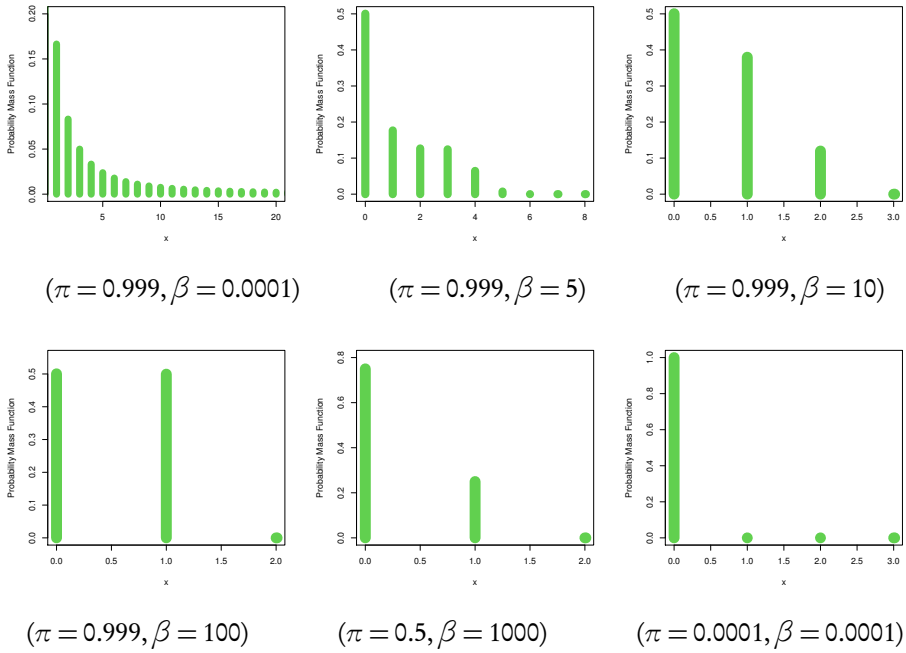

Figure 1 – The PMF plots of the DGBH distribution.

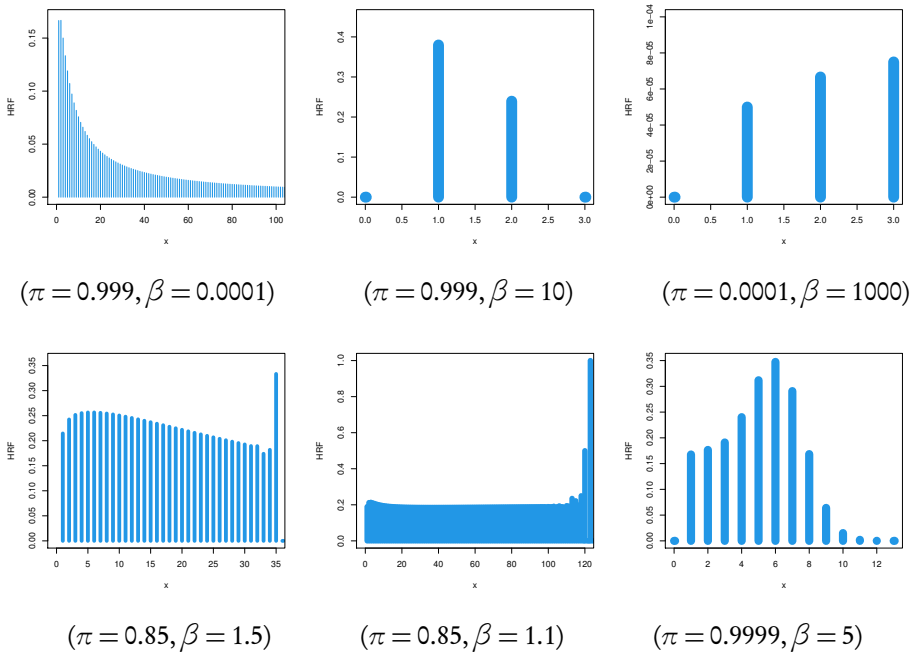

Figure 2 – The HRF plots of the DGBH distribution.

inextricable, necessitating the solution of the following inequality:

$$\frac{\pi^{k^\beta}}{k+1}-\frac{\pi^{(k+1)^\beta}}{k+2}\leq\frac{\pi^{z^\beta}}{z+1}-\frac{\pi^{(z+1)^\beta}}{z+2}, \tag{17}$$

which depends on the conjoint actions of π and β. To the best of our knowledge, only the use of a computer algebra system is possible to determine the numerical value of such a z. This explains why we have proposed a graphical approach to demonstrate the complexity of the situations and the unimodal and bimodal natures of the DGBH distribution. So, for the shapes of the PMF and HRF, using most of computer algebra systems, we can examine the local maximums and minimums and inflexion points. Otherwise, we have to examine the new PMF and its corresponding HRF graphically as presented in Figures 1 and 2, respectively.

Table 1 gives some numerical results related to $\mathbb{E}(Z)$, $\mathbb{V}(Z)$, $\mathbb{S}(Z)$, $\mathbb{K}(Z)$ and $\mathbb{D}(Z)$ of the DGBH and DBH distributions, respectively.

Based on the numerical results given in Tables 1 and 2, it is noted that the mean of the DGBH distribution increases as π increases. For comparison purposes, let us now put the distributional name "DGBH"or "DBH" in index of the corresponding measures $\mathbb{S}(Z)$, $\mathbb{K}(Z)$ and $\mathbb{D}(Z)$. The $\mathbb{S}_{\text{DGBH}}(Z)$ is positive and can range in the interval $(2.0\times 10^{-4}, 158220.3)$ whereas $\mathbb{S}_{\text{DBH}}(Z)$ can only range in the interval $(3.208, 106.612)$, the spread for its $\mathbb{K}_{\text{DGBH}}(Z)$ ranges from 1 to $\simeq\infty$, whereas $\mathbb{K}_{\text{DBH}}(Z)$ ranges from 17.95 to 20159.1. The $\mathbb{D}_{\text{DGBH}}(Z)\in(0.5, 80680.57)$. So, $\mathbb{D}_{\text{DGBH}}(Z)\in(0.1)$ or $\mathbb{D}_{\text{DGBH}}(Z)>1$. Thus, the new DGBH distribution could be useful in modeling "under-dispersed" or "over-dispersed" count data, whereas $\mathbb{D}_{\text{DBH}}(Z)\in(1.001, 2424.225)$.

3. Characterizations of the DGBH distribution

The problem of characterizing a distribution is an important problem in applied sciences, where an investigator is vitally interested in knowing if their model follows the right distribution. To this end, the investigator relies on conditions under which their model would follow specifically the chosen distribution. In this Section, we present certain characterizations of the DGBH distribution. These characterizations are based on: (i) the conditional expectation of certain function of the random variable; and (ii) in terms of the HRF.

Proposition 1. *Let $Z:\Omega\to\mathbb{N}$ be a RV. The PMF of Z is indicated as (5) if and only if*

$$\mathbb{E}\left\{\left[\frac{\pi^{Z^\beta}}{Z+1}+\frac{\pi^{(Z+1)^\beta}}{Z+2}\right]\middle|\{Z>x\}\right\}=\frac{\pi^{(x+1)^\beta}}{x+2},\quad x\in\mathbb{N}. \tag{18}$$

Proof. If Z has the PMF given by (5), then, for any $x\in\mathbb{N}$, the left-hand side of

TABLE 1
$\mathbb{E}(Z)$, $\mathbb{V}(Z)$, $\mathbb{S}(Z)$, $\mathbb{K}(Z)$ *and* $\mathbb{D}(Z)$ *of the DGBH distribution.*

π	β	$\mathbb{E}(Z)$	$\mathbb{V}(Z)$	$\mathbb{S}(Z)$	$\mathbb{K}(Z)$	$\mathbb{D}(Z)$
1×10^{-5}	1×10^{-5}	1.24×10^{-4}	9.9865	158220.3	$\simeq \infty$	80639.88
0.001		0.012	999.146	15818.13	$\simeq \infty$	80659.84
0.05		0.620	49980.01	2236.565	6669461	80675.96
0.10		1.239	99967.44	1581.443	3334511	80678.31
0.15		1.859	149957.2	1291.225	2222933	80679.42
0.25		3.098	249939.9	1000.166	1333715	80680.37
0.35		4.337	349924.4	845.293	952639.9	80680.57
0.55		6.816	549893.7	674.315	606223.1	80680.02
0.95		11.773	949819.4	513.094	350982.6	80677.40
0.999		12.391	999708.3	500.129	333468.8	80677.00
0.001	100	5×10^{-4}	0.00050	44.68781	1998.001	0.9995
0.25		0.125	0.109	2.268	6.143	0.875
0.50		0.250	0.188	1.155	2.333	0.750
0.75		0.375	0.234	0.516	1.267	0.625
0.95		0.475	0.249	0.100	1.010	0.525
0.9999		0.450	0.250	2×10^{-4}	1.000	0.500
0.1	20	0.500	0.048	4.129	18.053	0.950
0.15	50	0.075	0.069	3.227	11.414	0.925
0.25	200	0.125	0.109	2.268	6.143	0.875
0.75	1000	0.375	0.234	0.516	1.267	0.625
0.95	5000	0.475	0.249	0.100	1.010	0.525
0.99	10000	0.495	0.250	0.020	1.000	0.505
0.999	100000	0.499	0.250	0.002	1.000	0.501

TABLE 2
$\mathbb{E}(Z)$, $\mathbb{V}(Z)$, $\mathbb{S}(Z)$, $\mathbb{K}(Z)$ *and* $\mathbb{D}(Z)$ *of the DBH distribution.*

π	$\mathbb{E}(Z)$	$\mathbb{V}(Z)$	$\mathbb{S}(Z)$	$\mathbb{K}(Z)$	$\mathbb{D}(Z)$
0.001	0.001	0.001	44.762	2010.005	1.001
0.10	0.054	0.059	4.899	30.525	1.092
0.15	0.083	0.096	4.186	24.158	1.145
0.25	0.151	0.192	3.552	19.519	1.272
0.35	0.231	0.331	3.297	18.098	1.435
0.45	0.329	0.543	3.208	17.948	1.652
0.55	0.452	0.885	3.227	18.668	1.958
0.65	0.615	1.491	3.351	20.361	2.423
0.75	0.848	2.735	3.618	23.715	3.224
0.85	1.232	6.120	4.192	31.398	4.968
0.95	2.153	26.903	6.144	65.712	12.493
0.999	5.915	1945.273	34.457	2089.719	328.890
0.9999	8.211	19905.94	106.612	20159.06	2424.225

(18), will be

$$\left[1-\mathbf{F}_{\pi,\beta}(x)\right]^{-1}\sum_{z=x+1}^{\infty}\left[\left(\frac{\pi^{z^{\beta}}}{z+1}\right)^{2}-\left(\frac{\pi^{(z+1)^{\beta}}}{z+2}\right)^{2}\right]=\frac{x+2}{\pi^{(x+1)^{\beta}}}\left(\frac{\pi^{(x+1)^{\beta}}}{x+2}\right)^{2}$$
$$=\frac{\pi^{(x+1)^{\beta}}}{x+2}. \tag{19}$$

Conversely, if (18) holds, assuming that the distribution of Z is unknown but adopting the related notation for convenience, we get

$$\sum_{z=x+1}^{\infty}\left\{\left[\frac{\pi^{z^{\beta}}}{z+1}+\frac{\pi^{(z+1)^{\beta}}}{z+2}\right]\mathscr{F}_{\pi,\beta}(z)\right\}=\left[1-\mathbf{F}_{\pi,\beta}(x)\right]\frac{\pi^{(x+1)^{\beta}}}{x+2}$$
$$=\left[1-\mathbf{F}_{\pi,\beta}(x+1)+\mathscr{F}_{\pi,\beta}(x+1)\right]\frac{\pi^{(x+1)^{\beta}}}{x+2}. \tag{20}$$

From (20), we also have

$$\sum_{z=x+2}^{\infty}\left\{\left[\frac{\pi^{z^{\beta}}}{z+1}+\frac{\pi^{(z+1)^{\beta}}}{z+2}\right]\mathscr{F}_{\pi,\beta}(z)\right\}=\left[1-\mathbf{F}_{\pi,\beta}(x+1)\right]\frac{\pi^{(z+2)^{\beta}}}{x+3}. \tag{21}$$

Now, subtracting (21) from (20), we arrive at

$$\left[\frac{\pi^{(x+1)^{\beta}}}{x+2}+\frac{\pi^{(z+2)^{\beta}}}{x+3}\right]\mathscr{F}_{\pi,\beta}(x+1)$$
$$=\left[1-\mathbf{F}_{\pi,\beta}(x+1)\right]\left[\frac{\pi^{(x+1)^{\beta}}}{x+2}-\frac{\pi^{(x+2)^{\beta}}}{x+3}\right]+\frac{\pi^{(x+1)^{\beta}}}{x+2}\mathscr{F}_{\pi,\beta}(x+1), \tag{22}$$

implying that

$$\left[\frac{\pi^{(z+2)^{\beta}}}{x+3}\right]\mathscr{F}_{\pi,\beta}(x+1)=\left[1-\mathbf{F}_{\pi,\beta}(x+1)\right]\left[\frac{\pi^{(x+1)^{\beta}}}{x+2}-\frac{\pi^{(z+2)^{\beta}}}{x+3}\right] \tag{23}$$

or

$$\frac{\mathscr{F}_{\pi,\beta}(x+1)}{1-\mathbf{F}_{\pi,\beta}(x+1)}=\frac{\frac{\pi^{(x+1)^{\beta}}}{x+2}-\frac{\pi^{(x+2)^{\beta}}}{x+3}}{\frac{\pi^{(x+2)^{\beta}}}{x+3}}=\frac{(x+3)\,\pi^{(x+1)^{\beta}}}{(z+2)\pi^{(z+2)^{\beta}}}-1, \tag{24}$$

which is the HRF in (6), corresponding to the PMF in (5), taking at $z=x+1$. □

PROPOSITION 2. *Let $Z:\Omega\to\mathbb{N}$ be a RV. The PMF of Z is (5) if and only if its HRF, say $h_F(x)$, satisfies the following difference equation*

$$h_F(x+1)-h_F(x)=\frac{(x+3)\pi^{(x+1)^\beta}}{(x+2)\pi^{(x+2)^\beta}}-\frac{(z+2)\pi^{x^\beta}}{(x+1)\pi^{(x+1)^\beta}},\quad x\in\mathbb{N}, \tag{25}$$

with the initial condition $h_F(0)=2/\pi-1$.

PROOF. Clearly, if Z has PMF specified in (5), then (25) holds. Now, if (25) holds, then

$$\sum_{x=0}^{z-1}\{h_F(x+1)-h_F(x)\}=\sum_{x=0}^{z-1}\left\{\frac{(x+3)\pi^{(x+1)^\beta}}{(x+2)\,\pi^{(z+2)^\beta}}-\frac{(x+2)\pi^{x^\beta}}{(x+1)\pi^{(x+1)^\beta}}\right\}, \tag{26}$$

or

$$h_F(z)-h_F(0)=\frac{2}{\pi}+\frac{(z+2)\,\pi^{z^\beta}}{(z+1)\pi^{(z+1)^\beta}}, \tag{27}$$

or, in view of the initial condition $h_F(0)=2/\pi-1$, we have

$$h_F(z)=\frac{(z+2)\,\pi^{z^\beta}}{(z+1)\pi^{(z+1)^\beta}}-1,\quad z\in\mathbb{N}, \tag{28}$$

which is the HRF in (6), corresponding to the PMF in (5). □

4. ESTIMATION

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

4.1. Non-Bayesian estimation methods

4.1.1. The MLE method

Let $Z_1,Z_2,\ldots,Z_n$ be a random sample (RS) of size n from the DGBH distribution, and $z_1,z_2,\ldots,z_n$ be observed values of them. The log-likelihood function is given by

$$\ell=\ell(\pi,\beta)=\sum_{i=1}^{n}\ln\left[\frac{\pi^{z_i^\beta}}{z_i+1}-\frac{\pi^{(z_i+1)^\beta}}{z_i+2}\right]\Big|_{(\pi\in(0,1)\text{ and }z_i\in\mathbb{N})}, \tag{29}$$

which can be maximized either using statistical programs or by solving the nonlinear system obtained from $\ell(\pi,\beta)$ by differentiation. The components of the score vector $\mathbf{U}(\pi,\beta)=(\partial\ell(\pi,\beta)/\partial\pi,\partial\ell(\pi,\beta)/\partial\beta)^{\intercal}$ are easily derived. Setting $\partial\ell(\pi,\beta)/\partial\pi=\partial\ell(\pi,\beta)/\partial\beta=0$ and solving them simultaneously yields the MLEs for the DGBH parameters. The Newton-Raphson algorithms are employed for the numerically solving in such cases.

As usual, under regularity conditions, the properties of consistency and asymptotic normality are satisfied. Especially, the asymptotic distribution behind the MLEs is mutlivariate normal, with mean (π,β) and covariance matrix derived to the inverse of the expected Fisher covariance matrix. This asymptotic distribution is useful to construct confidence intervals (CIs), confidence regions, and various kinds of likelihood test (see Casella and Berger, 1990).

Theoretically, the issue of identifiability is difficult to ensure. Ideally, to prove this property, we need to prove that, for any $k\in\mathbb{N}$, the equality $\mathscr{F}_{\pi^*,\beta^*}(k)=\mathscr{F}_{\pi,\beta}(k)$ implies that $\pi^*=\pi$ and $\beta^*=\beta$. This equality is clear for $k=0$. However, for the other cases, the complexity of $\mathscr{F}_{\pi^*,\beta^*}(k)$ is a significant barrier to demonstrating that in full rigor. Our practical investigations, however, have revealed no problem of this kind, but the rigorous proof remains a strong mathematical challenge.

4.1.2. OLSE method

Let $z_{1:n}\leq z_{2:n}\leq\ldots\leq z_{n:n}$ be the n ordered observed values. The OLSEs are obtained upon minimizing

$$OLSE_{(\pi,\beta)}=\sum_{i=1}^{n}\left[\mathbf{F}_{\pi,\beta}(z_{i:n})-\varsigma_{(i,n)}^{(1)}\right]^2=\sum_{i=1}^{n}\left[1-\frac{\pi^{(z_{i:n}+1)^{\beta}}}{z_{i:n}+2}-\varsigma_{(i,n)}^{(1)}\right]^2, \tag{30}$$

with respect to π and β, where $\varsigma_{(i,n)}^{(1)}=i/(n+1)$. In an equivalent manner, the OLSEs are obtained via solving the following non-linear equations

$$0=\sum_{i=1}^{n}\left[1-\frac{\pi^{(z_{i:n}+1)^{\beta}}}{z_{i:n}+2}-\varsigma_{(i,n)}^{(1)}\right]\nu_{(\pi)}(z_{i:n};\pi,\beta), \tag{31}$$

and

$$0=\sum_{i=1}^{n}\left[1-\frac{\pi^{(z_{i:n}+1)^{\beta}}}{z_{i:n}+2}-\varsigma_{(i,n)}^{(1)}\right]\nu_{(\beta)}(z_{i:n};\pi,\beta), \tag{32}$$

where

$$\nu_{(\pi)}(z;\pi,\beta)=-\frac{(z+1)^{\beta}}{z+2}\pi^{(z+1)^{\beta}-1} \tag{33}$$

and

$$\nu_{(\beta)}(z;\pi,\beta)=-\frac{(z+1)^{\beta}}{z+2}\log(\pi)\log(z+1)\pi^{(z+1)^{\beta}}. \tag{34}$$

4.1.3. WLSE method

The WLSEs are obtained by minimizing the function $WLSE_{(\pi,\beta)}$ with respect to π and β

$$WLSE_{(\pi,\beta)}=\sum_{i=1}^{n}\varsigma_{(i,n)}^{(2)}\left[\mathbf{F}_{\pi,\beta}(z_{i:n})-\varsigma_{(i,n)}^{(1)}\right]^2=\sum_{i=1}^{n}\varsigma_{(i,n)}^{(2)}\left[1-\frac{\pi^{(z_{i:n}+1)^{\beta}}}{z_{i:n}+2}-\varsigma_{(i,n)}^{(1)}\right]^2, \tag{35}$$

where $\varsigma_{(i,n)}^{(2)}=\left[(1+n)^2(2+n)\right]/\left[i(1+n-i)\right]$. The WLSEs are obtained by solving

$$0=\sum_{i=1}^{n}\varsigma_{(i,n)}^{(2)}\left[1-\frac{\pi^{(z_{i:n}+1)^{\beta}}}{z_{i:n}+2}-\varsigma_{(i,n)}^{(1)}\right]\nu_{(\pi)}(z_{i:n};\pi,\beta), \tag{36}$$

and

$$0=\sum_{i=1}^{n}\varsigma_{(i,n)}^{(2)}\left[1-\frac{\pi^{(z_{i:n}+1)^{\beta}}}{z_{i:n}+2}-\varsigma_{(i,n)}^{(1)}\right]\nu_{(\beta)}(z_{i:n};\pi,\beta), \tag{37}$$

where $\nu_{(\pi)}(z;\pi,\beta)$ and $\nu_{(\beta)}(z;\pi,\beta)$ are defined by (33) and (34), respectively.

4.2. Bayesian estimation

Assume the beta and uniform priors for the parameters π and β, respectively. Then,

$$p_{1,(c_1,d_1)}(\pi)\sim\text{beta}(c_1,d_1)=\frac{1}{\mathrm{B}(c_1,d_1)}\pi^{c_1}(1-\pi)^{d_1}, \tag{38}$$

and

$$p_{2,(c_2,d_2)}(\beta)\sim\text{Gamma}(c_2,d_2)=\frac{d_2^{c_2}}{\Gamma(c_2)}\beta^{c_2-1}\exp(-\beta d_2), \tag{39}$$

where $\mathrm{B}(a,b)$ is the beta function and $\Gamma(a)$ is the gamma function. Assume that the parameters are independently distributed. The joint prior distribution $p_{(c_i,d_i)}(\pi,\beta)$ is given by

$$p_{(c_i,d_i)}(\pi,\beta)=\frac{1}{\mathrm{B}(c_1,d_1)\Gamma(c_2)}d_2^{c_2}\beta^{c_2-1}\pi^{c_1}(1-\pi)^{d_1}\exp(-\beta d_2). \tag{40}$$

The posterior distribution $p(\pi,\beta|\underline{z})$ of the parameters is defined as $p(\pi,\beta|\underline{z})\propto$ likelihood function $\times\, p_{(c_i,d_i)}(\pi,\beta)$, where $\underline{z}=(z_1,\ldots,z_n)$. If we consider the SELF, the

Bayesian estimators of π and β are the means of their marginal posteriors and defined as

$$\widehat{\pi}_{(\text{Bayesian})} = \int_{\pi,\beta} \pi p\left(\pi,\beta|\underline{z}\right)\, d\beta d\pi, \tag{41}$$

and

$$\widehat{\beta}_{(\text{Bayesian})} = \int_{\beta,\pi} \beta p\left(\pi,\beta|\underline{z}\right) d\pi d\beta, \tag{42}$$

respectively. It is not possible to obtain the Bayesian estimates through the above formulae. So, numerical approximations are needed. We propose the use of Markov chain Monte Carlo (MCMC) techniques, namely Gibbs sampler and M-H algorithm (for more details see Cai, 2010; Chib and Greenberg, 1995; Korkmaz *et al.*, 2018; Aboraya *et al.*, 2020).

It is worth mentioning that the MLEs and the Bayesian estimates have equivalent asymptotic properties (for more details see Ibragimov, 1962; Chao, 1970). Also, since the determination of the MLE is independent of the loss function and the prior measure, the asymptotic properties behind the Bayesian estimates hold for all priors and loss functions in a certain class.

Since the conditional posteriors of the parameters π and β cannot be obtained in any standard forms, using a hybrid MCMC for drawing samples from the marginal posterior of the parameters is suggested. Then, the full conditional posteriors of π and β can be easily derived. The simulation algorithm is given by:

1. Provide the initial values, say π and β, then at i^{th} stage;
2. Using M-H algorithm, generate

$$\pi_{(i)} \sim p_1\left(\pi_{(i)}|\pi_{(i-1)},\beta_{(i-1)},\underline{z}\right); \tag{43}$$

3. Using M-H algorithm, generate

$$\beta_{(i)} \sim p_2\left(\beta_{(i)}|\pi_{(i)},\beta_{(i-1)},\underline{z}\right); \tag{44}$$

4. Repeat steps $1-3$, 100000 times to obtain the sample of size $\mathbf{M}$ from the corresponding posteriors of interest. Obtain the Bayesian estimates of π and β using the following formulae:

$$\widehat{\pi}_{(\text{Bayesian})} = \frac{1}{\mathbf{M}-\mathbf{M}_0}\sum_{h=\mathbf{M}_0+1}^{\mathbf{M}} \pi^{[h]}, \text{ and } \widehat{\beta}_{(\text{Bayesian})} = \frac{1}{\mathbf{M}-\mathbf{M}_0}\sum_{h=\mathbf{M}_0+1}^{\mathbf{M}} \beta^{[h]}, \tag{45}$$

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

5. Simulations for comparing Bayesian and non-Bayesian estimation methods

A numerical MCMC simulation studies are performed for assessing and comparing the performance of non-Bayesian and Bayesian estimations. The numerical assessment is performed based on the mean squared errors (MSEs). First, we generated 1000 samples of the DGBH distribution, where $n = 50, 150, 300, 500$. The MSEs are obtained and listed in Tables 3, 4 and 5. Three combinations of initial values are considered (**I**: $\pi_0 = 0.7$ and $\beta_0 = 2$; **II**: $\pi_0 = 0.243$ and $\beta_0 = 0.500$ and **III**: $\pi_0 = 0.585$ and $\beta_0 = 0.9$). Based

TABLE 3
*MSEs for combination **I**.*

n		MLE	OLS	WLS	Bayesian
50	$\pi_0 = 0.7$	0.007	0.012	0.013	0.006
	$\beta_0 = 2.0$	0.989	0.807	0.761	0.113
150	$\pi_0 = 0.7$	0.002	0.004	0.005	0.003
	$\beta_0 = 2.0$	0.041	0.0833	0.066	0.037
300	$\pi_0 = 0.7$	0.001	0.002	0.003	0.001
	$\beta_0 = 2.0$	0.0182	0.0321	0.023	0.024
500	$\pi_0 = 0.7$	0.001	0.001	0.002	0.001
	$\beta_0 = 2.0$	0.010	0.021	0.014	0.013

on Tables 3, 4 and 5, we note that all methods perform well. The performance of all estimation methods improves as $n \to \infty$.

6. 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 (K-S) test and P-value $\left(\mathbf{P}_{[V]}\right)$ statistics in this regard.

6.1. 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 the

TABLE 4
*MSEs for combination **II**.*

n		MLE	OLS	WLS	Bayesian
50	$\pi_0 = 0.243$	0.003	0.004	0.008	0.008
	$\beta_0 = 0.500$	0.500	0.274	0.097	0.031
150	$\pi_0 = 0.243$	0.002	0.003	0.003	0.006
	$\beta_0 = 0.500$	0.032	0.074	0.035	0.024
300	$\pi_0 = 0.243$	0.001	0.001	0.001	0.001
	$\beta_0 = 0.500$	0.0129	0.028	0.019	0.022
500	$\pi_0 = 0.243$	0.001	0.001	0.001	0.001
	$\beta_0 = 0.500$	0.0067	0.016	0.013	0.006

TABLE 5
*MSEs for combination **III**.*

n		MLE	OLS	WLS	Bayesian
50	$\pi_0 = 0.585$	0.008	0.014	0.014	0.014
	$\beta_0 = 0.900$	0.109	0.265	0.094	0.044
150	$\pi_0 = 0.585$	0.003	0.005	0.005	0.004
	$\beta_0 = 0.900$	0.024	0.064	0.035	0.024
300	$\pi_0 = 0.585$	0.002	0.002	0.003	0.002
	$\beta_0 = 0.900$	0.009	0.031	0.021	0.016
500	$\pi_0 = 0.585$	0.001	0.001	0.002	0.001
	$\beta_0 = 0.900$	0.006	0.017	0.016	0.006

carious teeth data. Table 6 gives the estimations and the values of the considered statistics under Bayesian and non-Bayesian methods.

TABLE 6
Comparing methods under carious teeth data.

Method↓ estimations & statistics→	$\widehat{\pi}$	$\widehat{\beta}$	K-S	$\mathbf{P}_{[V]}$
MLE	0.764	1.309	1.204	0.548
OLS	0.724	1.006	0.757	0.685
WLS	0.721	1.190	1.411	0.494
Bayesian	0.741	1.297	1.622	0.444

Based on Table 6, the OLS method is the best with K-S $= 0.757$ and $\mathbf{P}_{[V]} = 0.685$.

6.2. *Application 2: Counts of cysts of kidneys data*

Due to Chan *et al.* (2009), these data represent the counts of cysts of corticosteroid-induced kidneys dysmorphogenesis which associated with deregulated expression of known cystogenic molecules as well as Indian hedgehog. Figure 4 gives the Kaplan–Meier survival plots for the counts of cysts of kidneys data. Table 7 gives the estimations and the values of the considered statistics under Bayesian and non-Bayesian methods.

TABLE 7
Comparing methods under kidneys data.

Method↓ estimations & statistics→	$\widehat{\pi}$	$\widehat{\beta}$	K-S	$\mathbf{P}_{[V]}$
MLE	0.993	2.276	3.491	0.32192
OLS	0.825	0.452	3.140	0.37050
WLS	0.821	0.930	5.030	0.16962
Bayesian	0.985	1.222	2.833	0.418

Based on Table 7, the Bayesian method is the best with K-S $= 2.833$ and $\mathbf{P}_{[V]} = 0.418$.

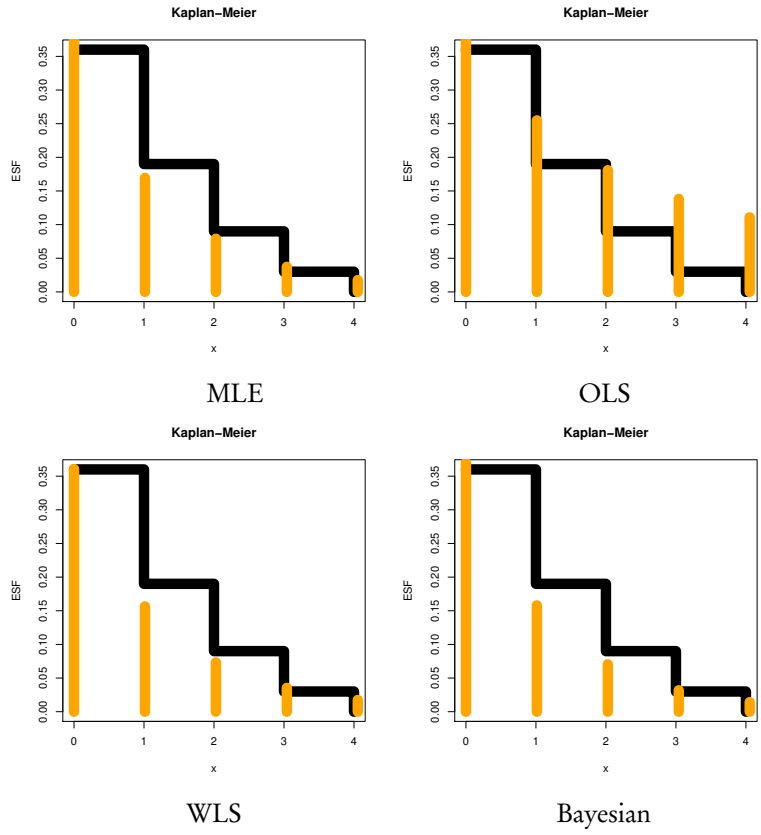

Figure 3 – Kaplan–Meier survival plots for the carious teeth data.

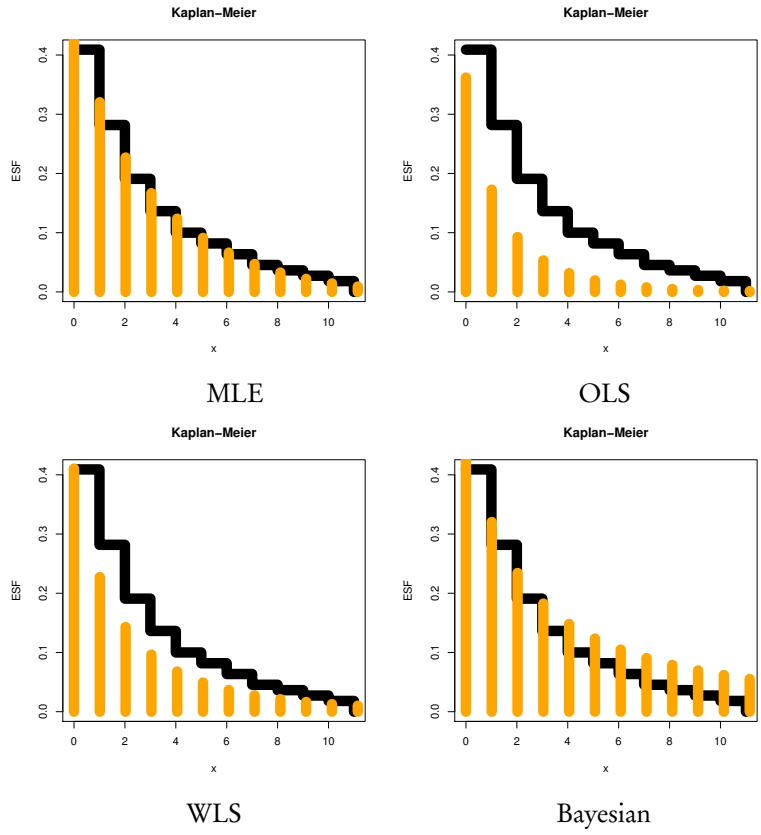

Figure 4 – Kaplan–Meier survival plots for Counts of cysts of kidneys data.

7. Application for comparing models and testing of hypothesis

7.1. Application 1: Carious teeth data

Using the carious teeth data, the fits of the DGBH model are compared with those of some competitive models such as the DBH, Poisson (P), Geometric (Gc), DPa, DR, DLi, DIR, PLi, EDLi, DW, DLi-II, DIW, GGc, DLL, DLFR and DGE-II models. Tables 8 and 9 give the observed frequency (OF), expected frequency (EF), MLEs, standard errors (SEs), 95%CIs, Chi-squared test $\left(\chi^2_V\right)$ and $\mathbf{P}_{[V]}$ for the competitive models for the carious teeth data. Figure 5 gives the box plot, quantile-quantile (Q-Q) plot and total time in test (TTT) plot. Figure 6 presents the fitted PMF, estimated HRF (EHRF), estimated SF (ESF) (Kaplan-Meier plot) and estimated CDF (ECDF).

TABLE 8
OF, EF, MLE, SEs, 95%CIs, χ^2_V and $\mathbf{P}_{[V]}$ for carious teeth data.

Z	OF	EF								
		DGBH	DBH	P	Gc	DPa	DR	DLi	DIR	PLi
0	64	61.791	66.44	51.17	59.88	59.88	33.50	57.13	62.50	37.50
1	17	21.087	18.54	34.28	24.02	24.02	46.94	26.88	26.41	25.00
2	10	9.066	7.460	11.49	9.640	9.640	17.01	10.45	5.990	15.63
≥ 3	9	8.056	7.560	3.060	6.460	6.460	2.550	5.450	5.100	21.87
$\sum$	100	100	100	100	100	100	100	100	100	100
π	MLE	0.764	0.671	0.670	0.401	0.184	0.665	0.625	0.625	1.998
	SEs	0.097	0.062	0.082	0.038	0.032	0.029	0.049	0.049	0.263
	95%LowerCI	0.573	0.549	0.509	0.327	0.121	0.608	0.529	0.529	1.481
	95%UpperCI	0.955	0.792	0.831	0.475	0.247	0.722	0.721	0.721	2.514
β	MLE	1.309	-	-	-	-	-	-	-	-
	SEs	0.329	-	-	-	-	-	-	-	-
	95%LowerCI	0.663	-	-	-	-	-	-	-	-
	95%UpperCI	1.954	-	-	-	-	-	-	-	-
χ^2_V		0.967	1.357	23.65	3.347	3.325	66.7	6.638	9.056	30.899
d.f		2	2	2	2	2	2	2	2	2
$\mathbf{P}_{[V]}$		0.548	0.507	<0.001	0.188	0.199	<0.001	.036	0.011	<0.001

TABLE 9
OF, EF, MLE, SEs, 95%CIs, χ^2_V and $\mathbf{P}_{[V]}$ for the carious teeth data (continued).

Z	OF	EF								
		DGBH	EDLi	DW	DLi-II	DIW	GGc	DLL	DLFR	DGE-II
0	64	61.791	63.57	62.58	59.88	63.30	62.73	62.73	59.90	63.51
1	17	21.087	19.75	21.35	24.02	22.48	21.36	22.42	24.01	20.19
2	10	9.066	9.090	8.85	9.640	4.440	8.760	7.010	9.630	8.81
≥ 3	9	8.056	7.230	7.22	6.460	7.780	7.150	7.840	6.460	7.49
$\sum$	100	100	100	100	100	100	100	100	100	100
π	MLE	0.764	0.379	0.374	0.401	0.633	0.467	0.745	0.402	0.468
	SEs	0.097	0.065	0.049	0.269	0.049	0.089	0.101	0.056	0.072
	95%LowerCI	0.573	0.252	0.278	0.000	0.537	0.293	0.546	0.291	0.327
	95%UpperCI	0.955	0.506	0.470	0.928	0.729	0.641	0.944	0.511	0.609
β	MLE	1.309	0.543	0.895	0.478	1.576	0.678	1.768	1.000	0.718
	SEs	0.329	0.158	0.119	0.529	0.251	0.302	0.267	0.044	0.206
	95%LowerCI	0.663	0.234	0.662	0.000	1.084	0.086	1.244	0.912	0.324
	95%UpperCI	1.954	0.852	1.128	1.514	2.067	1.270	2.292	1.080	1.122
χ^2_V		0.967	0.739	1.507	3.347	3.503	1.57	2.783	3.340	0.973
d.f		2	1	1	1	1	1	1	1	1
$\mathbf{P}_{[V]}$		0.548	0.390	0.219	0.067	0.061	0.210	0.095	0.068	0.324

Based on Tables 8 and 9, and Figure 6, the DGBH model provides the best fits against all competitive models with $\mathbf{P}_{[V]} = 0.548$. For $\widehat{\pi} = 0.764$ and $\widehat{\beta} = 1.309$, we have $\mathbb{E}(Z) = 2.424$, $\mathbb{V}(Z) = 17.244$, $\mathbb{S}(Z) = 10.263$, $\mathbb{K}(Z) = 231.096$ and $\mathbb{D}(Z) = 7.114 > 1$.

7.2. Application 2: Counts of cysts of kidneys data

For counts of cysts of kidneys data set, we compare the fits of the DGBH model with those of the DLi-II, DIW, DR, DIR, DLi, PLi, P and Gc models. Table 10 gives the observed frequency (OF), expected frequency (EF), MLEs, standard errors (SEs), 95%CIs, χ^2_V and $\mathbf{P}_{[V]}$ for the competitive models for the counts of cysts of kidneys data. Figure 7 gives the box plot, Q-Q plot and TTT plot, and Figure 8 displays the fitted PMF, EHRF, ESF and ECDF.

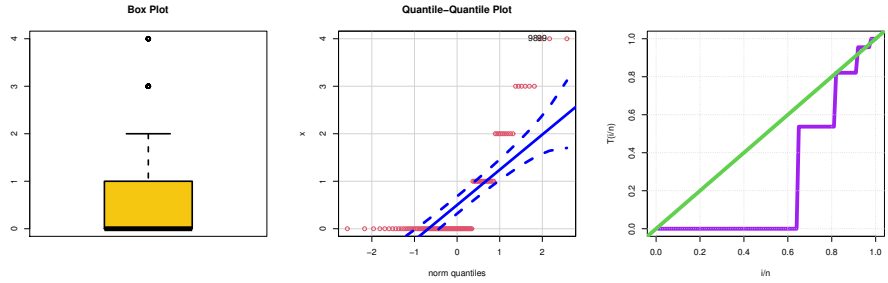

Figure 5 – Box plot, Q-Q plot and TTT plot for the carious teeth data.

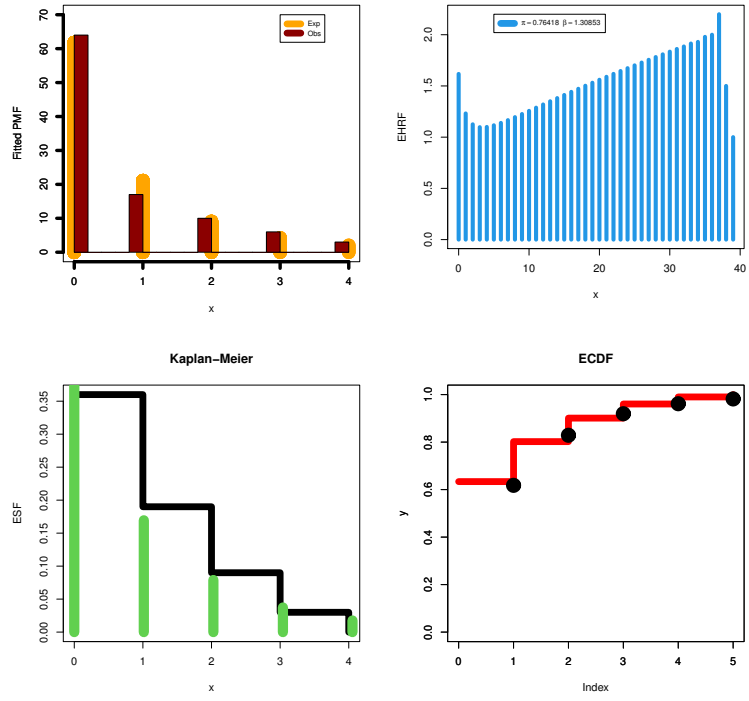

Figure 6 – The fitted PMF, EHRF, ESF and ECDF for the carious teeth data.

TABLE 10
OF, EF, MLE, SEs, 95%CIs, χ^2_V and $\mathbf{P}_{[V]}$ for the counts of cysts of kidneys data.

Z	OF	EF								
		DGBH	DLi-II	DIW	Gc	P	DR	DIR	DLi	PLi
0	65	55.297	46.03	63.91	45.98	27.42	11.00	60.94	40.25	44.14
1	14	19.079	26.77	20.70	26.67	38.08	26.83	33.96	29.83	28.00
2	10	10.156	15.57	8.05	15.57	26.47	29.55	8.11	18.36	17.70
3	6	6.602	9.05	4.23	9.060	12.26	22.23	3.00	10.35	9.57
4	4	4.777	5.27	2.60	5.280	4.26	12.49	1.42	5.53	5.34
5	2	3.648	3.06	1.75	3.070	1.18	5.42	0.87	2.86	2.92
6	2	2.842	1.78	1.26	1.79	0.27	1.85	0.47	1.44	1.57
7	2	2.208	1.04	0.95	1.04	0.05	0.52	0.31	0.71	0.84
8	1	1.684	0.60	0.74	0.62	0.01	0.11	0.21	0.35	0.44
9	1	1.247	0.35	0.59	0.35	0.00	0.02	0.15	0.17	0.23
10	1	0.891	0.20	0.48	0.21	0.00	0.00	0.11	0.08	0.12
11	2	0.610	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
π	MLE	0.993	0.581	0.581	0.582	1.390	0.901	0.554	0.436	1.087
	SEs	0.013	0.045	0.048	0.030	0.112	0.009	0.049	0.026	1.109
	95%LowerCI	0.968	0.494	0.489	0.523	1.170	0.882	0.458	0.385	0.873
	95%UpperCI	1.018	0.669	0.675	0.641	1.601	0.918	0.649	0.487	1.301
β	MLE	2.276	0.001	1.049	-	-	-	-	-	-
	SEs	0.762	0.058	0.146	-	-	-	-	-	-
	95%LowerCI	0.782	0.000	0.763	-	-	-	-	-	-
	95%UpperCI	3.769	0.115	1.335	-	-	-	-	-	-
χ^2_V		0.967	22.89	24.135	22.84	294.10	321.07	51.047	43.48	31.151
d.f		4	3	3	4	4	4	4	4	4
$P_{[V]}$		0.322	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

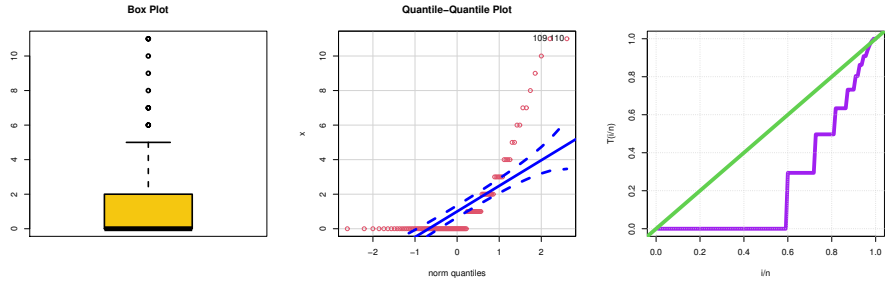

Figure 7 – Box plot, Q-Q plot and TTT plot for the counts of cysts of kidneys data.

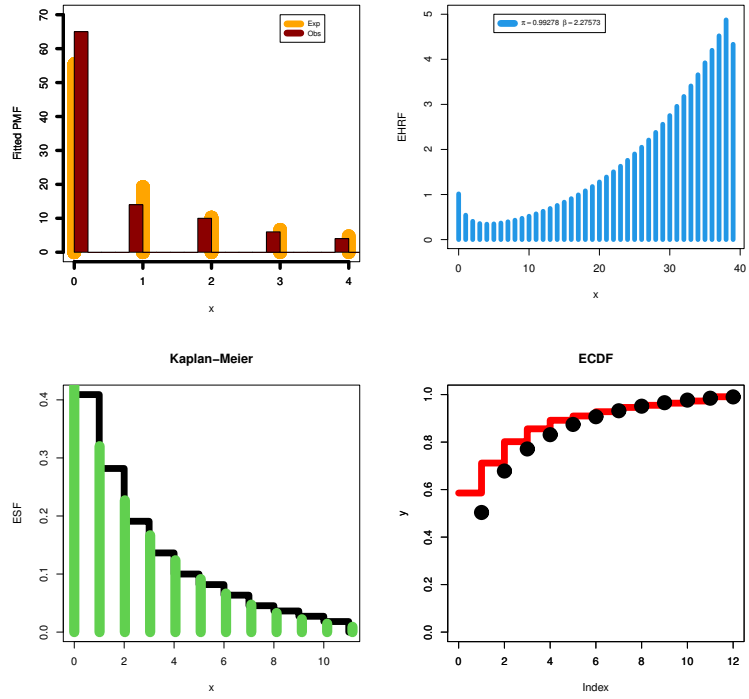

Figure 8 – The fitted PMF, EHRF, ESF and ECDF for the counts of cysts of kidneys data.

Based on Table 10 and Figure 8, the DGBH provides the best fits against all competitive models with $\mathbf{P}_{[V]} = 0.322$. For $\widehat{\pi} = 0.993$ and $\widehat{\beta} = 2.276$, we have $\mathbb{E}(Z) = 2.647$, $\mathbb{V}(Z) = 6.784$, $\mathbb{S}(Z) = 2.192$, $\mathbb{K}(Z) = 8.318$ and $\mathbb{D}(Z) = 2.563 > 1$.

8. Conclusions

A new discrete distribution which includes the discrete Burr-Hatke distribution is defined and studied. The probability mass function of the new model can be "right skewed" with different shapes, bimodal and "uniformed". The hazard rate function of the new model can be "monotonically decreasing", "upside down", "monotonically increasing", "upside down increasing", and "upside down-constant-increasing". Relevant statistical properties, such as the probability generating function, ordinary moments, index of dispersion and order statistics are derived. Based on a numerical analysis for the mean, variance, skewness, kurtosis and the index of dispersion of the discrete extended Burr-Hatke (DGBH) and discrete Burr-Hatke (DBH) distributions, it is noted that the mean of the DGBH distribution increases as the parameter π increases. The skewness of the DGBH distribution is positive and can range in the interval $(2.0 \times 10^{-4}, 158220.3)$ whereas the skewness of the DBH distribution can only range in the interval $(3.208, 106.612)$, the spread for kurtosis of the DGBH distribution ranges from 1 to $\approx \infty$, whereas kurtosis of the DBH distribution ranges from 17.94759 to 20159.06. The index of dispersion of the DGBH distribution belongs to the interval $(0.5, 80680.57)$. So, the index of dispersion of the DGBH distribution can be in $(0, 1)$ or be greater than 1. Thus, the new DGBH distribution could be useful in modeling "under-dispersed" or "over-dispersed" count data, whereas index of dispersion of the DBH distribution can onlyrange in the interval $(1, 2424.225)$. We presented certain characterizations of the DGBH distribution. These characterizations are based on: (i) the conditional expectation of certain function of the RV and (ii) in terms of the hazard rate function. Bayesian (under the squared error loss function) and non-Bayesian estimation methods (maximum likelihood estimation, ordinary least squares and weighted least squares) are considered. Numerical simulations for comparing Bayesian and non-Bayesian estimation methods are performed. Moreover, the DGBH model is applied for modeling the carious teeth data and counts of cysts of kidneys data. The DGBH model provides the best fits against many well-known competitive models. We hope that the new distribution will attract wider applications in reliability, engineering and other areas of research.

Acknowledgements

The authors would like to thank the reviewer for the thorough comments on the manuscript, improving it in several aspects.

References

M. Aboraya, H. Yousof, G. Hamedani, M. Ibrahim (2020). *A new family of discrete distributions with mathematical properties, characterizations,* B*ayesian and non-*B*ayesian estimation methods*. Mathematics, 8, no. 10.

L. Cai (2010). M*etropolis-*H*astings* R*obbins-*M*onro algorithm for confirmatory item factor analysis*. Journal of Educational and Behavioral Statistics, 35, no. 3, pp. 307–335.

G. Casella, R. Berger (1990). *Statistical Inference*. Brooks/Cole Publishing Company, California.

S. Chan, P. Riley, K. Price, F. McElduff, P. Winyard (2009). *Corticosteroid-induced kidney dysmorphogenesis is associated with deregulated expression of known cystogenic molecules, as well as* I*ndian hedgehog*. American Journal of Physiology: Renal Physiology, 298, pp. 346–356.

M. Chao (1970). *The asymptotic behavior of* B*ayes' estimators*. The Annals of Mathematical Statistics, 41, no. 2, pp. 601–608.

S. Chib, E. Greenberg (1995). *Understanding the* M*etropolis-*H*astings algorithm*. The American Statistician, 49, no. 4, pp. 327–335.

M. El-Morshedy, M. Eliwa, E. Altun (2020a). *Discrete* B*urr-*H*atke distribution with properties*. IEEE Access, 8, pp. 74359–74370.

M. El-Morshedy, M. Eliwa, H. Nagy (2020b). *A new two-parameter exponentiated discrete* L*indley distribution: Properties, estimation and applications*. Journal of Applied Statistics, 47, pp. 354–375.

E. Gomez-Déniz (2010). *Another generalization of the geometric distribution*. TEST, 19, pp. 399–415.

E. Gomez-Déniz, E. Caldern-Ojeda (2011). *The discrete* L*indley distribution: Properties and applications*. Journal of Statistical Computation and Simulation, 81, pp. 1405–1416.

T. Hussain, M. Ahmad (2014). *Discrete inverse* R*ayleigh distribution*. Pakistan Journal of Statistics, 30, pp. 203–222.

T. Hussain, M. Aslam, M. Ahmad (2016). *Two parameter discrete* L*indley distribution*. Revista Colombiana de Estadística, 39, pp. 45–61.

I. Ibragimov (1962). *Some limit theorems for stationary processes*. Theory of Probability & Its Applications, 7, no. 4, pp. 349–382.

A. Jazi, D. Lai, H. Alamatsaz (2010). *Inverse* W*eibull distribution and estimation of its parameters*. Statistical Methodology, 7, pp. 121–132.

A. KEMP (2008). *The discrete half-normal distribution*. Advances in mathematical and statistical modeling, Birkhauser, Basel, pp. 353–365.

M. KORKMAZ, H. YOUSOF, M. RASEKHI, G. HAMEDANI (2018). *The odd Lindley Burr XII model: Bayesian analysis, classical inference and characterizations*. Journal of Data Science, 16, no. 2, pp. 327–353.

H. KRISHNA, P. PUNDIR (2009). *Discrete Burr and discrete Pareto distributions*. Statistical Methodology, 6, no. 2, pp. 177–188.

C. KUMAR, Y. TRIPATHI, M. RASTOGI (2017). *On a discrete analogue of linear failure rate distribution*. American journal of mathematical and management sciences, 36, no. 3, pp. 229–246.

A. MANIU, V. VODA (2008). *Generalized Burr-Hatke equation as generator of a homographic failure rate*. Journal of applied quantitative methods, 3, pp. 215–222.

T. NAKAGAWA, S. OSAKI (1975). *The discrete Weibull distribution*. IEEE Transactions on Reliability, 24, no. 5, pp. 300–301.

V. NEKOUKHOU, M. ALAMATSAZ, H. BIDRAM (2013). *Discrete generalized exponential distribution of a second type*. Statistics, 47, no. 4, pp. 876–887.

B. PARA, T. JAN (2016a). *Discrete version of log-logistic distribution and its applications in genetics*. International Journal of Modern Mathematical Sciences, 14, pp. 407–422.

B. PARA, T. JAN (2016b). *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, 4, pp. 1–15.

D. ROY (2004). *Discrete Rayleigh distribution*. IEEE Transactions on Reliability, 53, pp. 255–260.

M. SANKARAN (1970). *The discrete Poisson-Lindley distribution*. Biometrics, 26, no. 1, pp. 145–149.

F. STEUTEL, K. VAN HARN (2004). *Infinite Divisibility of Probability Distributions on the Real Line*. Marcel Dekker, New York, NY, USA.

H. YOUSOF, E. ALTUN, T. RAMIRES, M. ALIZADEH, M. RASEKHI (2018). *A new family of distributions with properties, regression models and applications*. Journal of Statistics and Management Systems, 21, no. 1, pp. 163–188.

SUMMARY

In this work, a new discrete distribution which includes the discrete Burr-Hatke distribution is defined and studied. Relevant statistical properties are derived. The probability mass function of the new distribution can be "right skewed" with different shapes, bimodal and "uniformed". Also, the corresponding hazard rate function can be "monotonically decreasing", "upside down", "monotonically increasing", "upside down increasing", and "upside down-constant-increasing". A numerical analysis for the mean, variance, skewness, kurtosis and the index of dispersion is presented. The new distribution could be useful in the modeling of "under-dispersed" or "over-dispersed" count data. Certain characterizations of the new distribution are presented. These characterizations are based on the conditional expectation of a certain function of the random variable and in terms of the hazard rate function. Bayesian and non-Bayesian estimation methods are considered. Numerical simulations for comparing Bayesian and non-Bayesian estimation methods are performed. The new model is applied for modeling carious teeth data and counts of cysts of kidneys data.

Keywords: Discretization; Characterizations; Discrete Burr-Hatke distribution; Bayesian estimation; Metropolis Hastings; Markov Chain Monte Carlo; Maximum likelihood; Cramér-von-Mises.