Article

A New Family of Discrete Distributions with Mathematical Properties, Characterizations, Bayesian and Non-Bayesian Estimation Methods

Mohamed Aboraya ¹, Haitham M. Yousof ² , G.G. Hamedani ³ and Mohamed Ibrahim ^{1,*}

¹ Department of Applied Statistics and Insurance, Faculty of Commerce, Damietta University, Damietta 34519, Egypt; mohamedaboraya17@gmail.com

² Department of Statistics, Mathematics and Insurance, Benha University, Banha 13513, Egypt; haitham.yousof@fcom.bu.edu.eg

³ Department of Mathematical and Statistical Sciences, Marquette University, Milwaukee, WI 53201-1881, USA; gholamhoss.hamedani@marquette.edu

* Correspondence: mohamed_ibrahim@du.edu.eg

Received: 21 August 2020; Accepted: 22 September 2020; Published: 24 September 2020

Abstract: In this work, we propose and study a new family of discrete distributions. Many useful mathematical properties, such as ordinary moments, moment generating function, cumulant generating function, probability generating function, central moment, and dispersion index are derived. Some special discrete versions are presented. A certain special case is discussed graphically and numerically. The hazard rate function of the new class can be “decreasing”, “upside down”, “increasing”, and “decreasing-constant-increasing (U-shape)”. Some useful characterization results based on the conditional expectation of certain function of the random variable and in terms of the hazard function are derived and presented. Bayesian and non-Bayesian methods of estimation are considered. The Bayesian estimation procedure under the squared error loss function is discussed. Markov chain Monte Carlo simulation studies for comparing non-Bayesian and Bayesian estimations are performed using the Gibbs sampler and Metropolis–Hastings algorithm. Four applications to real data sets are employed for comparing the Bayesian and non-Bayesian methods. The importance and flexibility of the new discrete class is illustrated by means of four real data applications.

Keywords: discretization; Metropolis–Hastings; Markov chain Monte Carlo; maximum likelihood; Cramér–von Mises; squared error loss function; Cramér–von Mises; Bayesian estimation

1. Introduction and Genesis

Discretization of existing continuous probability distributions have received noticeable attention in recent years. In this paper, we present and study a new discrete analogue based on the continuous Rayleigh distribution called the discrete Rayleigh G family of distributions. The relevant mathematical properties, such as moments, cumulant generating function, moment generating function, probability generating function, central moment, and dispersion index (DisIx) are derived and analyzed. A special version of the new family related to well-known Weibull model is discussed. Some classical (non-Bayesian) estimation methods such as Cramér–von Mises estimation (CVME), the ordinary least squared estimation (OLSE), the maximum likelihood estimation (MLE), and the weighted least squared estimation (WLSE) are described and considered. Since the conditional posteriors of the parameters cannot be obtained in any standard forms, using a hybrid Markov chain Monte Carlo for drawing sample from the joint posterior of the parameters is suggested. The Bayesian estimation procedure under the squared error loss function (SELF) is also presented. Markov chain Monte Carlo (MCMC) simulations for comparing non-Bayesian and Bayesian estimations are performed. Gibbs sampler and

Metropolis–Hastings (M–H) algorithm are employed and applied. The flexibility and importance of the new family is illustrated by four real data sets. The new family provides a better fit than sixteen competitive distributions. Many special member distributions could be considered and studied in the future. As a future work, we may consider the bivariate and the multivariate extensions of this new family.

In probability theory, the Rayleigh distribution is a continuous probability model for nonnegative-valued random variables (NNRVs). A Rayleigh distribution can often be observed when the overall magnitude of a vector is related to its directional components. A RV X is said to have the continuous Rayleigh distribution if its cumulative distribution function (CDF) is given by $F_{\alpha}(x) = 1 - \exp[-(\alpha x)^2]$ $\Big|_{(x \geq 0 \text{ and } \alpha > 0)}$. Recently, [1] defined the continuous CDF of the Rayleigh G (RG) family as $F_{\alpha, \underline{\phi}}(x) = 1 - \exp[-\alpha \nabla_{\underline{\phi}}^2(x)]$ $\Big|_{(x \in \mathbb{R} \text{ and } \nabla_{\underline{\phi}}(\cdot) = G_{\underline{\phi}}(\cdot) / [1 - G_{\underline{\phi}}(\cdot)])}$ where $G_{\underline{\phi}}(x)$ refers to the continuous CDF of any base line model. Let $\alpha = -\log(\pi)$ then, CDF of the discrete RG family (DRG) can be expressed as

$$F_{\pi, \underline{\phi}}(y) = 1 - \pi \nabla_{\underline{\phi}}^2(y+1) \Big|_{(\pi \in (0,1) \text{ and } y \in N^* = N \cup \{0\})}, \quad (1)$$

The corresponding reliability function (RF) can be written as

$$S_{\pi, \underline{\phi}}(y) = \pi \nabla_{\underline{\phi}}^2(y+1) \Big|_{(\pi \in (0,1) \text{ and } y \in N^*)}, \quad (2)$$

The probability mass function (PMF) of the discrete analogue of DRG family corresponding to (2) can be expressed as $\mathcal{F}_{\pi, \underline{\phi}}(y) = S_{\pi, \underline{\phi}}(y) - S_{\pi, \underline{\phi}}(y+1)$. Therefore, the PMF can be written as

$$\mathcal{F}_{\pi, \underline{\phi}}(y) = \pi \nabla_{\underline{\phi}}^2(y-1) - \pi \nabla_{\underline{\phi}}^2(y) \Big|_{(\pi \in (0,1) \text{ and } y \in N^*)}. \quad (3)$$

where $\nabla_{\underline{\phi}}^2(y)$ refers to the squared odd ratio function of the NNRV Y . The hazard rate function (HRF) of the DRG family can be written as

$$h_F(y) = \mathcal{F}_{\pi, \underline{\phi}}(y) / S_{\pi, \underline{\phi}}(y-1). \quad (4)$$

In the statistical literature, many discrete versions of continuous distributions have been proposed and studied such as a generalization of the Poisson distribution ([2]), the discrete Weibull (DW) ([3]), the discrete Rayleigh (DR) ([4]), discrete half-normal distribution ([5]), discrete Pareto (DPa) ([6]), a new discrete geometric (DGc) ([7]), the discrete inverse Weibull (DIW) ([8]), the discrete Lindley (DLi) ([9]), the discrete generalized exponentiated type II (DGE-II) ([10]), the discrete inverse Rayleigh (DIR) ([11]), exponentiated discrete Weibull (EDW) ([12]), the discrete Lindley type II (DLi-II) ([13]), discrete Lomax (DLx) ([14]), the discrete log-logistic (DLL) ([15]), the discrete Burr type XII (DBXII) ([14]), and the exponentiated discrete Lindley (EDLi) ([16]), among others.

Table 1 presents some new sub-models based on the DRG family. In Section 3, however, a comprehensive graphical and numerical analysis is given for the discrete Rayleigh Weibull model. Many useful discretization processes can be applied to many existing continuous families (see [17–24]) and continuous probability distributions (see [25–30]).

This paper is organized as follows. Some mathematical properties of the DRG family are derived and analyzed in Section 2. Some useful characterization results are derived and presented in Section 3. Estimation and inference procedures are presented in Section 4. Markov chain Monte Carlo simulations for comparing Bayesian and non-Bayesian estimation methods are performed in Section 5. Four real data applications for comparing Bayesian and non-Bayesian estimation methods are presented in Section 6. Four applications for comparing the competitive discrete models are considered in Section 7. Finally, Section 8 offers some concluding remarks.

Table 1. Some new sub-models based on the discrete Rayleigh G family (DRG) family.

No.	Baseline Model	$\Delta_{\Phi}(y) _{y \in N_0, \pi \in (0,1)}$	Sub Model
1	Log-logistic (LL)	$y^{\theta} _{\theta>0}$	DRLl
2	Log-logistic (LL)	$(y/\theta)^{\beta} _{\theta, \beta>0}$	DRLl
3	Weibull (W)	$\exp(y^{\theta}) - 1 _{\theta>0}$	DRW
4	Exponential (E)	$\exp(\beta y) - 1 _{\beta>0}$	DRE
5	Lomax (Lx)	$(1+y)^{\beta} - 1 _{\beta>0}$	DRLx
6	Rayleigh (R)	$\exp(\beta y)^2 - 1 _{\beta>0}$	DRR
7	Weibull (W)	$\exp(\beta y)^{\theta} - 1 _{\theta, \beta>0}$	DRW
8	Lomax (Lx)	$(1+y/\beta)^{\theta} - 1 _{\theta, \beta>0}$	DRLx
9	Burr XII (BXII)	$(1+y^{\theta})^{\beta} - 1 _{\theta, \beta>0}$	DRBXII
10	Fréchet (F)	$[\exp(y^{-\theta}) - 1]^{-1} _{\theta>0}$	DRF
11	Inverse Rayleigh (IR)	$[\exp(\beta y^{-2}) - 1]^{-1} _{\beta>0}$	DRIR
12	Inverse Exponential (IE)	$[\exp(\beta y^{-1}) - 1]^{-1} _{\beta>0}$	DRIE
13	Inverse Lomax (ILx)	$[(1+\beta y^{-1})^{\theta} - 1]^{-1} _{\theta, \beta>0}$	DRILx
14	Inverse Lomax (ILx)	$[(1+y^{-1})^{\theta} - 1]^{-1} _{\theta>0}$	DRILx
15	Burr XII (BXII)	$[(1+(y/\lambda)^{\theta})^{\beta} - 1]^{-1} _{\theta, \beta, \lambda>0}$	DRBXII
16	Fréchet (F)	$[\exp(\beta^{\theta} y^{-\theta}) - 1]^{-1} _{\beta, \theta>0}$	DRF
17	Half-logistic (HL)	$\left\{ \left[\frac{1-\exp(-y)}{1+\exp(-y)} \right]^{-1} - 1 \right\}^{-1}$	DRHL
18	Standard Gumbel (Gu)	$(\exp\{\exp[-(y)]\} - 1)^{-1}$	DRSGu
19	Lindley (Li)	$\exp(\beta y)[(1+\beta+\beta y)/(1+\beta)]^{-1} - 1 _{\beta>0}$	DRLi
20	Nadarajah-Haghighi (NH)	$\exp[(1+\beta x)^{\theta} - 1] - 1 _{\beta, \theta>0}$	DRNH
21	Half-logistic (HL)	$\left\{ \left[\frac{1-\exp(-\beta y)}{1+\exp(-\beta y)} \right]^{-1} - 1 \right\}^{-1} _{\beta>0}$	DRHL
22	Gompertz	$\exp\{\theta[\exp(\beta y) - 1]\} - 1 _{\beta, \theta>0}$	DRGz
23	Dagum (D)	$[(1+(y/\lambda)^{-\theta})^{\beta} - 1]^{-1} _{\theta, \beta, \lambda>0}$	DRD
24	Inverse Flexible W (IFW)	$(\exp\{\exp[\beta/y - \theta y]\} - 1)^{-1} _{\theta, \beta>0}$	DRIFW
25	Gumbel (Gu)	$(\exp\{\exp[-(y-\theta)/\beta]\} - 1)^{-1} _{\beta>0, \theta \in \mathbb{R}}$	DRGu
26	Burr X (BX)	$\left(\left[1 - \exp[-(\beta y)^2] \right]^{-\theta} - 1 \right)^{-1} _{\beta, \theta>0}$	DRBX
27	Inverse Gompertz	$(\exp\{-\theta[\exp(\beta y) - 1]/\beta\} - 1)^{-1} _{\beta, \theta>0}$	DRIGz

2. Properties

2.1. Ordinary Moments

Theorem 1. Let Y be an NNRV, where $Y \sim \text{DRG}(\pi, \underline{\Phi})$ family, then the r^{th} moment of Y can be expressed as

$$\mu'_r = E(Y^r) = \sum_{y=1}^{\infty} [y^r - (y-1)^r] \pi^{\nabla_{\underline{\Phi}}^2(y)} \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=1,2,3,\dots)}. \quad (5)$$

Proof. See Appendix A.1. $\square$

Then, using Equation (5), the mean (μ'_1), and μ'_2 can be respectively written as

$$\mu_1 = \mu'_1 = E(Y) = \sum_{y=1}^{\infty} \pi^{\nabla_{\underline{\Phi}}^2(y)} \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=1)},$$

and

$$\mu'_2 = E(Y^2) = \sum_{y=1}^{\infty} (2y-1) \pi^{\nabla_{\underline{\Phi}}^2(y)} \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=2)}.$$

2.2. Central Moments

The r^{th} central moment of Y , say μ_r , is

$$\mu_r = E(Y - \mu'_1)^r = \sum_{w=0}^r (-\mu'_1)^w \binom{r}{w} \mu'_{r-w} \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=1,2,3,\dots)}.$$

So, the variance ($V(X)$) from

$$E(Y - \mu'_1)^2 = \mu \sum_{w=0}^r (-\mu'_1)^w \binom{r}{w} \mu'_{2-w} \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=2)},$$

can be derived using the ordinary moments

$$\mu_2 = V(Y) = \mu'_2 - (\mu'_1)^2 = \sum_{y=1}^{\infty} (2y-1) \pi^{\nabla_{\underline{\Phi}}^2(y)} - (\mu'_1)^2 \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=2)}.$$

The DisIx or the variance to mean ratio (VMR) of the DRG family can be derived as

$$\text{DisIx}(Y) = \frac{\sum_{y=1}^{\infty} (2y-1) \pi^{\nabla_{\underline{\Phi}}^2(y)}}{\sum_{y=1}^{\infty} \pi^{\nabla_{\underline{\Phi}}^2(y)}} - \sum_{y=1}^{\infty} \pi^{\nabla_{\underline{\Phi}}^2(y)} \Big|_{(y \in N^{\bullet} \text{ and } \pi \in (0,1))}.$$

The DisIx is used to quantify whether a specific set of observations are clustered or “dispersed” compared to a specific statistical model, so, the DisIx is used to determine whether the observed real data set can be modeled via a Poisson process. For any real data set, when the DisIx is <1 , the data set is said to be “under-dispersed” data. A numerical analysis with its related calculations for the DisIx(Y) is presented in Table 2.

2.3. The Moment and Cumulant Generating Function (MGF & CGF)

Theorem 2. Let Y be an NNRV, where $Y \sim \text{DRG}(\pi, \underline{\Phi})$ family, then the MGF of Y can be obtained as

$$M_Y(t) = 1 + \sum_{y=1}^{\infty} \{ \exp(tY) - \exp[t(Y-1)] \} \pi_{\underline{\Phi}}^{\nabla^2}(y) \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=1,2,3,\dots)} \quad (6)$$

Proof. See Appendix A.2. $\square$

The first r derivatives of Equation (6), with respect to $t|t = 0$, yield the first r raw moments, i.e.,

$$\mu'_r = E(Y^r) = \frac{d^r}{dt^r} M_Y(t) \Big|_{(t=0 \text{ and } r=1,2,3,\dots)},$$

where

$$\mu'_1 = E(Y) = \frac{d}{dt} M_Y(t) \Big|_{t=0} = \sum_{y=1}^{\infty} \pi_{\underline{\Phi}}^{\nabla^2}(y) \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=1)},$$

$$\mu'_2 = E(Y^2) = \frac{d^2}{dt^2} M_Y(t) \Big|_{(t=0)} = \sum_{y=1}^{\infty} (2y-1) \pi_{\underline{\Phi}}^{\nabla^2}(y) \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=2)},$$

$$\mu'_3 = E(Y^3) = \frac{d^3}{dt^3} M_Y(t) \Big|_{(t=0)} = \sum_{y=1}^{\infty} [3y(y-1) + 1] \pi_{\underline{\Phi}}^{\nabla^2}(y) \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=3)},$$

and

$$\mu'_4 = E(Y^4) = \frac{d^4}{dt^4} M_Y(t) \Big|_{(t=0)} = \sum_{y=1}^{\infty} [y^4 - (y-1)^4] \pi_{\underline{\Phi}}^{\nabla^2}(y) \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=4)}.$$

The CGF is the logarithm of the MGF. Thus, r th cumulant, say κ_r , can be obtained from $\kappa_r = \frac{d^r}{dt^r} \log[M_Y(t)] \Big|_{(t=0, \text{ and } r=1,2,3,\dots)}$. The first cumulant is the mean, the second cumulant is the variance, and the third cumulant is the same as the third central moment $\kappa_3 = \mu_3$ that being said $\kappa_1 = \mu'_1 = E(Y)$, $\kappa_2 = \mu_2 = \mu'_2 - \mu_1'^2$ and $\kappa_3 = \mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2\mu_1'^3$. However, the fourth and higher order cumulants are not equal to central moments. In some cases, theoretical treatments of problems in terms of cumulants are simpler than those using moments. In particular, when two or more RVs are statistically independent, the r th order cumulant of their sum is equal to the sum of their r th order cumulants. Moreover, the cumulants can be also obtained from $\kappa_r \Big|_{r \geq 1} = \mu'_r - \sum_{w=0}^{r-1} \binom{r-1}{w-1} \mu'_{r-w} \kappa_w$. The skewness ($S(Y)$) and kurtosis measures ($K(Y)$) of Y follow from the well-known relationships. A numerical analysis for the $S(Y)$ and $K(Y)$ for a special member of the new family is presented in Table 2 with useful comments. The probability generating function can be written as $P_Y(s) = 1 + \sum_{y=1}^{\infty} (1 - \frac{1}{s}) s^y \pi_{\underline{\Phi}}^{\nabla^2}(y) \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } r=1,2,3,\dots)}$. For the Weibull base line model, we have $\nabla_{\underline{\Phi}}(y) = \exp(y^{\theta}) - 1 \Big|_{y \in N^{\bullet}, \pi \in (0,1), \theta > 0}$. Then based on (3), the PMF of the DRW model (see Table 1) can be expressed as $\mathcal{F}_{\pi, \theta}(y) = \pi^{[\exp(y^{\theta})-1]^2} - \pi^{[\exp((y+1)^{\theta})-1]^2} \Big|_{(y \in N^{\bullet}, \pi \in (0,1) \text{ and } \theta > 0)}$. Clearly, when $\theta = 1$, the DRW model reduced to the DRE model. Figure 1 gives some plots of the PMF of the DRW model for various values of the parameters. Figure 2 shows some plots of the HRF of the DRW model for various values of the parameters. Based on Figure 1, we conclude that the PMF of the DRW can be “heavy tail right skewed PMF” and “left skewed PMF”. Based on Figure 2, we see that the HRF of the DRW can be “decreasing” ($\pi = 0.5, \theta = 5$), “upside down” ($\pi = 0.9, \theta = 0.75$), “increasing” ($\pi = 0.992, \theta = 0.263$) and “decreasing-constant-increasing” ($\pi = 0.99, \theta = 0.25$). Based on Table 2, it is noted that $E(Y)$ of the DRW model increases as π increases and decreases as θ increases. The $S(y)$ of the DRW distribution can be positive and negative. The spread for $K(y)$ is ranging from 1.9 to ∞ . The $\text{DisI}_x(Y)$ of the DRW distribution can range in the interval (0.1) or more than 1 or equal 1

(for $\pi = 0.001, \theta = 5$) like Poisson distribution. So, the new DRW distribution could be useful in modeling “under-dispersed” or “over-dispersed” count data.

Figure 1. The probability mass function (PMF) of the DRW for different values of the parameters.

Figure 2. The hazard rate function (HRF) of the DRW for different values of the parameters.

Table 2. E(Y), V(Y), S(Y), and K(Y) of the DRW distribution.

π	θ	E(Y)	V(Y)	S(Y)	K(Y)	DisIx(Y)
0.001	5	1.3884×10^9	1.3884×10^{-9}	26,837.29	720,240.2	1
0.100		0.0011156	0.0011144	29.88949	894.3817	0.9988844
0.250		0.0166887	0.0164102	7.545712	57.93777	0.9833113
0.500		0.1291847	0.1124960	2.211154	5.889202	0.8708153
0.750		0.4276804	0.2447699	0.2923527	1.085470	0.5723196
0.900		0.7326581	0.1958702	−1.051390	2.105422	0.2673419
0.990		0.9707624	0.0283828	−5.588618	32.23265	0.0292376
0.999		0.9970504	0.0029409	−18.33114	337.0307	0.0029496
0.5	0.001	804.5731	401004320	33.20704	1242.543	498406.30
	0.1	0.526052	7.049518	14.38802	426.5162	13.400800
	0.5	0.1303937	0.1158098	2.286966	6.577818	0.8881549
	5	0.1291847	0.112496	2.211154	5.889202	0.8708153

3. Characterizations of DRG Distribution

The problem of characterizing a distribution is an important problem in applied sciences, where an investigator is vitally interested to know 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 DRG distribution. These characterizations are based on the conditional expectation of certain function of the random variable and in terms of the hazard function.

Proposition 1. Let $Y : \Omega \rightarrow N^*$ be a random variable. The PMF of Y is Equation (3) if and only if

$$E \left\{ \left[\pi^{\nabla_{\Phi}^2(Y)} + \pi^{\nabla_{\Phi}(Y+1)} \right] \middle| X > k \right\} = \pi^{\nabla_{\Phi}^2(k+1)} \Big|_{k \in N^*}. \quad (7)$$

Proof. If Y has PMF Equation (3), then the left-hand side of Equation (7), will be

$$\left[1 - F_{\pi, \Phi}(k) \right]^{-1} \sum_{y=k+1}^{\infty} \left\{ \pi^{2\nabla_{\Phi}^2(y)} - \pi^{2\nabla_{\Phi}^2(y+1)} \right\} = \pi^{-\nabla_{\Phi}^2(k+1)} \left[\pi^{2\nabla_{\Phi}^2(k+1)} \right] = \pi^{\nabla_{\Phi}^2(k+1)} \Big|_{k \in N^*}.$$

Conversely, if Equation (7) holds, then

$$\sum_{y=k+1}^{\infty} \left\{ \left[\pi^{\nabla_{\Phi}^2(y)} + \pi^{\nabla_{\Phi}^2(y+1)} \right] f(y) \right\} = [1 - F(k)] \pi^{\nabla_{\Phi}^2(k+1)} = \left\{ \frac{[1 - F_{\pi, \Phi}(k+1)]}{+ \mathcal{F}_{\pi, \Phi}(k+1)} \right\} \pi^{\nabla_{\Phi}^2(k+1)}. \quad (8)$$

From Equation (8), we also have

$$\sum_{y=k+2}^{\infty} \left\{ \left[\pi^{\nabla_{\Phi}^2(y)} + \pi^{\nabla_{\Phi}^2(y+1)} \right] \mathcal{F}_{\pi, \Phi}(y) \right\} = [1 - F_{\pi, \Phi}(k+1)] \pi^{\nabla_{\Phi}^2(k+2)}. \quad (9)$$

Now, subtracting Equation (9) from Equation (8), we arrive at

$$\left[\frac{\pi^{\nabla_{\Phi}^2(k+1)}}{+ \pi^{\nabla_{\Phi}^2(k+2)}} \right] \mathcal{F}_{\pi, \Phi}(k+1) = \left[\pi^{\nabla_{\Phi}^2(k+1)} - \pi^{\nabla_{\Phi}^2(k+2)} \right] [1 - F_{\pi, \Phi}(k+1)] + \pi^{\nabla_{\Phi}^2(k+1)} \mathcal{F}_{\pi, \Phi}(k+1).$$

or

$$\pi^{\nabla_{\Phi}^2(k+2)} \mathcal{F}_{\pi, \Phi}(k+1) = \left[\pi^{\nabla_{\Phi}^2(k+1)} - \pi^{\nabla_{\Phi}^2(k+2)} \right] [1 - F_{\pi, \Phi}(k+1)],$$

or

$$\frac{\mathcal{F}_{\pi, \underline{\Phi}}(k+1)}{1 - F_{\pi, \underline{\Phi}}(k+1)} = \frac{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)} - \pi \frac{\nabla_{\underline{\Phi}}^2(k+2)}{\nabla_{\underline{\Phi}}^2(k+2)}}{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}} = \frac{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}}{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}} - 1.$$

From the last equality, we have

$$h_F(k+1) = \frac{\mathcal{F}_{\pi, \underline{\Phi}}(k+1)}{1 - F_{\pi, \underline{\Phi}}(k+1)} = \frac{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}}{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}} - 1,$$

which implies that Y has PMF Equation (3). $\square$

Proposition 2. Let $Y : \Omega \rightarrow N^*$ be a random variable. The PMF of Y is Equation (3) if and only if its hazard function, h_F , satisfies the following difference equation

$$h_F(k+1) - h_F(k) = \frac{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}}{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}} - \frac{\pi \frac{\nabla_{\underline{\Phi}}^2(k)}{\nabla_{\underline{\Phi}}^2(k+1)}}{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}} \Bigg|_{k \in N^*}, \quad (10)$$

with the initial condition $h_F(0) = \frac{1}{\pi \frac{\nabla_{\underline{\Phi}}^2(1)}{\nabla_{\underline{\Phi}}^2(2)}} - 1$.

Proof. Clearly, if Y has PMF Equation (3), then Equation (10) holds. Now, if Equation (10) holds, then

$$\sum_{y=0}^k [h_F(y+1) - h_F(y)] = \sum_{y=0}^k \left[\frac{\pi \frac{\nabla_{\underline{\Phi}}^2(y+1)}{\nabla_{\underline{\Phi}}^2(y+2)}}{\pi \frac{\nabla_{\underline{\Phi}}^2(y+1)}{\nabla_{\underline{\Phi}}^2(y+2)}} - \frac{\pi \frac{\nabla_{\underline{\Phi}}^2(y)}{\nabla_{\underline{\Phi}}^2(y+1)}}{\pi \frac{\nabla_{\underline{\Phi}}^2(y+1)}{\nabla_{\underline{\Phi}}^2(y+2)}} \right],$$

or

$$h_F(k+1) - h_F(0) = -\frac{1}{\pi \frac{\nabla_{\underline{\Phi}}^2(1)}{\nabla_{\underline{\Phi}}^2(2)}} + \frac{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}}{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}},$$

or, in view of the initial condition $h_F(0) = \frac{1}{\pi \frac{\nabla_{\underline{\Phi}}^2(1)}{\nabla_{\underline{\Phi}}^2(2)}} - 1$ we have

$$h_F(k+1) = \frac{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}}{\pi \frac{\nabla_{\underline{\Phi}}^2(k+1)}{\nabla_{\underline{\Phi}}^2(k+2)}} - 1 \Bigg|_{k \in N^*},$$

which is the hazard function corresponding to the PMF Equation (3). $\square$

4. Estimation and Inference

In this subsection, non-Bayesian and Bayesian estimation methods are considered. In the first subsection, we will consider the MLE method, CVME method, OLSE method, and WLSE method. In second subsection, the Bayesian estimation method under the SELF. 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 $Y_1, Y_2, \dots, Y_n$ be a random sample (RS) from the DRG distribution. The log-likelihood function for $\underline{V}$ where $\underline{V} = (\pi, \underline{\Phi}_j)^T \Big|_{j=1,2,\dots,p}$, is given by

$$\ell = \ell(\underline{V}) = \sum_{i=0}^n \ln \left[\pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n})} - \pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n}+1)} \right] \Bigg|_{(\pi \in (0,1) \text{ and } y_{i:n} \in N^*)}$$

which can be maximized either using the statistical programs or by solving the nonlinear system obtained from $\ell(\underline{V})$ via differentiation. The score vector $U(\underline{V}) = \left(\partial \ell(\underline{V}) / \partial \pi, \partial \ell(\underline{V}) / \partial \underline{\Phi}_j \right)^T$, are

$$\partial \ell(\underline{V}) / \partial \pi = \sum_{i=0}^n \frac{\nabla_{\underline{\Phi}_j}^2(y_{i:n}) \pi^{[\nabla_{\underline{\Phi}_j}^2(y_{i:n})]-1} - \nabla_{\underline{\Phi}_j}^2(y_{i:n}+1) \pi^{[\nabla_{\underline{\Phi}_j}^2(y_{i:n}+1)]-1}}{\pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n})} - \pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n}+1)}},$$

and

$$\partial \ell(\underline{V}) / \partial \underline{\Phi}_j = \sum_{i=0}^n \frac{\frac{\partial \nabla_{\underline{\Phi}_j}^2(y_{i:n})}{\partial \underline{\Phi}_j} \pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n})} \ln \pi - \frac{\partial \nabla_{\underline{\Phi}_j}^2(y_{i:n}+1)}{\partial \underline{\Phi}_j} \pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n}+1)} \ln \pi}{\pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n})} - \pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n}+1)}} \Bigg|_{(j=1,2,\dots,p)},$$

where

$$\partial \nabla_{\underline{\Phi}_j}^2(y_{i:n}) / \partial \underline{\Phi}_j = 2 \nabla_{\underline{\Phi}_j}(y_{i:n}) \frac{\partial}{\partial \underline{\Phi}_j} \nabla_{\underline{\Phi}_j}(y_{i:n}),$$

and

$$\partial \nabla_{\underline{\Phi}_j}^2(y_{i:n}+1) / \partial \underline{\Phi}_j = 2 \nabla_{\underline{\Phi}_j}(y_{i:n}+1)^{2-1} \frac{\partial}{\partial \underline{\Phi}_j} \nabla_{\underline{\Phi}_j}(y_{i:n}+1).$$

Setting $0 = \partial \ell(\underline{V}) / \partial \pi = \partial \ell(\underline{V}) / \partial \underline{\Phi}_j$ and solving them simultaneously yields the MLEs for the parameters of the DRG family. The Newton–Raphson algorithm is employed for obtaining the numerical solutions in such cases.

4.1.2. The CVME Method

The CVME of the parameter vector $\underline{V} = (\pi, \underline{\Phi}_j)^T$ are obtained via minimizing the following expression with respect to (WRT) to π and $\underline{\Phi}_j$ respectively, where

$$CVM_{(\underline{V})} = \frac{1}{12} n^{-1} + \sum_{i=1}^n \left[F_{\pi, \underline{\Phi}_j}(y_{i:n}) - \xi_{(i,n)}^{[1]} \right]^2 \Bigg|_{(\pi \in (0,1) \text{ and } y_{i:n} \in N^*)},$$

where $\xi_{(i,n)}^{[1]} = \frac{2i-1}{2n}$ and

$$CVM_{(\underline{V})} = \sum_{i=1}^n \left[1 - \pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n}+1)} - \xi_{(i,n)}^{[1]} \right]^2.$$

The, CVMEs are obtained by solving the following two non-linear equations

$$\sum_{i=1}^n \left(1 - \pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n}+1)} - \xi_{(i,n)}^{[1]} \right) D_{(\pi)}(y_{i:n}+1, \underline{V}) = 0,$$

and

$$\sum_{i=1}^n \left(1 - \pi^{\nabla_{\underline{\Phi}_j}^2(y_{i:n}+1)} - \xi_{(i,n)}^{[1]} \right) D_{(\underline{\Phi}_j)}(y_{i:n}+1, \underline{V}) = 0,$$

where $D_{(\pi)}(y_{i:n}+1, \underline{V}) = \partial F_{\pi, \underline{\Phi}_j}(y_{i:n}) / \partial \pi$ and $D_{(\underline{\Phi}_j)}(y_{i:n}+1, \underline{V}) = \partial F_{\pi, \underline{\Phi}_j}(y_{i:n}) / \partial \underline{\Phi}_j$ are the first partial derivatives of the CDF of DRG distribution WRT to π and $\underline{\Phi}_j$ respectively.

4.1.3. OLSE Method

Let $F_{\pi, \underline{\Phi}_j}(y_{i:n})$ denote the CDF of DRG model and let $Y_1 < Y_2 < \dots < Y_n$ be the n ordered RS. The OLSEs are obtained upon minimizing

$$OLSE_{(\underline{V})} = \sum_{i=1}^n \left[F_{\pi, \underline{\Phi}_j}(y_{i:n}) - \xi_{(i,n)}^{[2]} \right]^2,$$

then, we have

$$OLSE_{(\underline{V})} = \sum_{i=1}^n \left[1 - \pi^{\nabla_{\underline{\Phi}_j}^2}(y_{i:n}+1) - \xi_{(i,n)}^{[2]} \right]^2,$$

where $\xi_{(i,n)}^{[2]} = \frac{i}{n+1}$. The LSEs are obtained via solving the following non-linear equations

$$0 = \sum_{i=1}^n \left[1 - \pi^{\nabla_{\underline{\Phi}_j}^2}(y_{i:n}+1) - \xi_{(i,n)}^{[2]} \right] D_{(\pi)}(y_{i:n} + 1, \underline{V}),$$

and

$$0 = \sum_{i=1}^n \left[1 - \pi^{\nabla_{\underline{\Phi}_j}^2}(y_{i:n}+1) - \xi_{(i,n)}^{[2]} \right] D_{(\underline{\Phi}_j)}(y_{i:n} + 1, \underline{V}),$$

where $D_{(\pi)}(y_{i:n} + 1, \underline{V})$ and $D_{(\underline{\Phi}_j)}(y_{i:n} + 1, \underline{V})$ defined above.

4.1.4. WLSE Method

The WLSE are obtained by minimizing the function $WLSE_{(\underline{V})}$ WRT π and $\underline{\Phi}_j$

$$WLSE_{(\underline{V})} = \sum_{i=1}^n \xi_{(i,n)}^{[3]} \left[F_{\pi, \underline{\Phi}_j}(y_{i:n}) - \xi_{(i,n)}^{[2]} \right]^2,$$

where

$$\xi_{(i,n)}^{[3]} = \left[(1+n)^2(2+n) \right] / [i(1+n-i)].$$

The WLSEs are obtained by solving

$$0 = \sum_{i=1}^n \xi_{(i,n)}^{[3]} \left[1 - \pi^{\nabla_{\underline{\Phi}_j}^2}(y_{i:n}+1) - \xi_{(i,n)}^{[2]} \right] D_{(\pi)}(y_{i:n} + 1, \underline{V}),$$

and

$$0 = \sum_{i=1}^n \xi_{(i,n)}^{[3]} \left[1 - \pi^{\nabla_{\underline{\Phi}_j}^2}(y_{i:n}+1) - \xi_{(i,n)}^{[2]} \right] D_{(\underline{\Phi}_j)}(y_{i:n} + 1, \underline{V}),$$

where $D_{(\pi)}(y_{i:n} + 1, \underline{V})$ and $D_{(\underline{\Phi}_j)}(y_{i:n} + 1, \underline{V})$ are defined above.

4.2. Bayesian Estimation

Assume the beta and uniform priors for the parameters π and $\underline{\Phi}_j$ respectively. Then,

$$p_{1,(\Phi_1, \psi_1)}(\pi) \sim \text{beta}(\Phi_1, \psi_1),$$

and

$$p_{2,(\Phi_2, \psi_2)}(\underline{\Phi}_j) \sim \text{Uniform}(\Phi_2, \psi_2).$$

Assume that the parameters are independently distributed. The joint prior distribution $p_{(\Phi_i, \psi_i)}(\pi, \underline{\Phi}_j)$ is given by

$$p_{(\Phi_i, \psi_i)}(\pi, \underline{\Phi}_j) = \frac{1}{(\psi_2 - \Phi_2)B(\Phi_1, \psi_1)} \pi^{\Phi_1} (1 - \pi)^{\psi_1},$$

where $B(\cdot, \cdot)$ is the beta function. The posterior distribution $p(\pi, \underline{\Phi}_j | \underline{z})$ of the parameters is defined as

$$p(\pi, \underline{\Phi}_j | \underline{z}) \propto \text{likelihood function} \times p_{(\Phi_i, \psi_i)}(\pi, \underline{\Phi}_j).$$

Under SELF, the Bayesian estimators of π and $\underline{\Phi}_j$ are the means of their marginal. It is not possible to obtain the Bayesian estimates through the above formulae. So, the numerical approximations are needed. We propose the use of MCMC techniques, namely the Gibbs sampler and M–H algorithm (see [31–34] for more details). Since the conditional posteriors of the parameters π and $\underline{\Phi}_j$ cannot be obtained in any standard forms, using a hybrid MCMC for drawing sample from the marginal posterior of the parameters is suggested. Then, the full conditional posteriors of π and $\underline{\Phi}_j$ can be easily derived. The simulation algorithm is given by:

1. Provide the initial values, say π and $\underline{\Phi}_j$ then at i^{th} stage;
2. Using M–H algorithm, generate $[\pi]_{(i)} \sim p_1([\pi]_{(i)} | [\pi]_{(i-1)}, [\underline{\Phi}_j]_{(i-1)}, \underline{z})$;
3. Using M–H algorithm, generate $[\underline{\Phi}_j]_{(i)} \sim p_2([\underline{\Phi}_j]_{(i)} | [\pi]_{(i)}, [\underline{\Phi}_j]_{(i-1)}, \underline{z})$;
4. Repeat steps 1 – 3, $M = 100,000$ times to obtain the sample of size M from the corresponding posteriors of interest. Obtain the Bayesian estimates of π and $\underline{\Phi}_j$ using the following formulae

$$\hat{\pi}_{\text{Bayes}} = \frac{1}{M - M_0} \sum_{h=1+M_0}^M \pi^{[h]} \text{ and } \hat{\underline{\Phi}}_{j\text{Bayes}} = \frac{1}{M - M_0} \sum_{h=1+M_0}^M \underline{\Phi}_j^{[h]} \text{ respectively, where } M_0 (\approx 50,000) \text{ is the burn-in period of the generated MCMC.}$$

5. Simulations for Comparing Non-Bayesian and Bayesian Estimation Methods

For the DRW case, 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 DRW distribution, where $n = 50, 150, 300, 500$. The MSEs are obtained and listed in Tables 3–5. Based on Tables 3–5 we note that all methods perform well. Though the Bayesian method is better in some cases. The performance of all estimation methods improves and tends to 0 when $n \rightarrow \infty$.

Table 3. Mean squared errors (MSEs) for $\pi = 0.607$ and $\theta = 0.1$.

n		MLE	OLS	WLS	Bayesian
50	π	0.00449	0.00287	0.00289	0.00255
	θ	0.00073	0.00261	0.00235	0.00045
150	π	0.00176	0.00090	0.00092	0.00069
	θ	0.00017	0.00073	0.00046	0.00036
300	π	0.00073	0.00045	0.00047	0.00041
	θ	0.00006	0.00024	0.00012	0.00011
500	π	0.00043	0.00027	0.00028	0.00032
	θ	0.00004	0.00015	0.00007	0.00003

MLE: maximum likelihood estimation; OLS: ordinary least squared; WLS: weighted least squared.

Table 4. MSEs for $\pi = 0.9$ and $\theta = 0.8$.

n		MLE	OLS	WLS	Bayesian
50	π	0.00039	0.00039	0.00054	0.00022
	θ	0.01086	0.00956	0.01566	0.00376
150	π	0.00017	0.00010	0.00020	0.00008
	θ	0.00299	0.00143	0.00143	0.00196
300	π	0.00010	0.00004	0.00011	0.00004
	θ	0.00181	0.00062	0.00062	0.00069
500	π	0.00007	0.00002	0.00006	0.00002
	θ	0.00132	0.00038	0.00039	0.00036

Table 5. MSEs for $\pi = 0.5$ and $\theta = 0.3$.

N		MLE	OLS	WLS	Bayesian
50	π	0.00448	0.00435	0.00423	0.00359
	θ	0.33939	0.38439	0.38022	0.20457
150	π	0.00120	0.00129	0.00126	0.00116
	θ	0.06961	0.07194	0.06676	0.00282
300	π	0.00056	0.00067	0.00061	0.00068
	θ	0.00793	0.00679	0.00656	0.00168
500	π	0.00032	0.00038	0.00039	0.00057
	θ	0.00182	0.00218	0.00118	0.00096

6. Real Data Modeling for Comparing Bayesian and Non-Bayesian Estimation Methods

In this Section, four examples of real data sets are introduced for comparing the Bayesian and non-Bayesian estimation methods. We consider the Akaike information criterion (AIC) and Correct Akaike IC (CAIC) statistic for comparing the Bayesian and non-Bayesian estimation methods.

6.1. Failure Times Data of 50 Devices

This data represents the failure times of 50 devices (in weeks) put on a certain life test (see [35]). The data are: 0.1, 0.2, 1, 1, 1, 1, 2, 3, 6, 7, 11, 12, 18, 18, 18, 18, 18, 21, 32, 36, 40, 45, 46, 47, 50, 55, 60, 63, 63, 67, 67, 67, 67, 72, 75, 79, 82, 82, 83, 84, 84, 84, 85, 85, 85, 85, 85, 86, 86. Table 6 gives the estimators under Bayesian and non-Bayesian estimation methods, AIC, and CAIC statistics for data set I. Figure 3 presents the frequency distribution using Bayesian estimation for data set I.

Figure 3. Frequency distribution using Bayesian estimation for data set I.

1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5, 5, 6, 6, 7, 7, 8, 9, 10, 11, 11. Table 8 gives the estimators under Bayesian and non-Bayesian estimation methods, AIC, and CAIC statistics for data set III. Figure 5 gives the frequency distribution using Bayesian estimation for data set III.

Table 8. Estimators under Bayesian and non-Bayesian estimation methods, AIC and CAIC statistics for data set III.

Method	$\hat{\pi}$	$\hat{\theta}$	AIC	CAIC
MLE	0.745	0.183	338.9	338.9
OLS	0.740	0.171	339.4	339.5
WLS	0.739	0.188	339.2	339.3
Bayesian	0.737	0.183	339.0	339.2

Figure 5. Frequency distribution using Bayesian estimation for data set III.

Based on Table 7, the **MLE** method is the best method with $AIC = 338.9$ and $CAIC = 338.9$. However, all other methods perform very well.

6.4. Number of European Corn-Borer Larvae Parasites

[illegible]

Table 9. Estimators under Bayesian and non-Bayesian estimation methods, AIC and CAIC statistics for data set IV.

Method	$\hat{\pi}$	$\hat{\theta}$	AIC	CAIC
MLE	0.849	0.298	406.1	406.2
OLS	0.855	0.330	409.6	409.7
WLS	0.860	0.320	407.2	407.3
Bayesian	0.841	0.292	406.3	406.4

Figure 6. Frequency distribution using Bayesian estimation for data set IV.

Based on Table 7, the MLE method is the best method with $AIC = 406.1$ and $CAIC = 406.2$. However, all other methods perform very well.

7. Real Data Modeling for Comparing the Competitive Discrete Models Using the MLE Method

We illustrate the flexibility and the importance of the DRW distributions using four real data applications. The fitted distributions (see Table 10) are analyzed and compared using the log-likelihood function (ℓ), AIC, CAIC, Chi-square (χ^2_V) with degree of freedom (d.f), and its p -value ($P_{[V]}$), Kolmogorov–Smirnov ($K - S$) and its $P_{[V]}$. Table 6 below gives the competitive models.

Table 10. The competitive models.

	Model	Abbreviation
1	discrete Pareto	DPa
2	discrete Lomax	DLx
3	discrete Lindley	DLi
4	discrete Weibull	DW
5	discrete Rayleigh	DR
6	discrete log-logistic	DLL
7	discrete exponential	DE
8	discrete Burr type XII	DBXII
9	discrete Lindley type II	DLi-II
10	discrete inverse Rayleigh	DIR
11	Poisson	Poisson
12	discrete inverse Weibull	DIW
13	Exponentiated discrete Weibull	EDW
14	Exponentiated discrete Lindley	EDLi
15	discrete generalized E type II	DGE-II
16	Negative binomial	NB

7.1. Failure Times Data of 50 Devices

We compare the fits of the DRW model with some competitive models, such as DW, EDW, DIW, EDLi, DPa, DLi-II, and DLL. The MLEs along with their corresponding standard errors (SEs), and the goodness of fit (GOF) test statistics are listed in Tables 11 and 12, respectively. Statisticians have developed a remarkably powerful set of tools for analyzing normally distributed data. The most popular one is the “normal quantile-quantile (Q-Q) plot”. If the distribution of the data matched the normal distribution perfectly, all the quantile points would lie between the two blue lines. Figure 7 (left plot) gives the Q-Q plot for dataset I. Figure 7 (right) gives the box for the failure time data (data set I). In the applications, the shape of the HRF can help in selecting a particular model. For this aim, the total time on test (TTT) plot (see [39]) can be used. The TTT plot is obtained by plotting $T(r/n)$ against r/n , where $T(r/n) = [\sum_{i=1}^n y_{(i)} + (n-r)y_{(r)}] / \sum_{i=1}^n y_{(i)}$, $r = 1, \dots, n$ and the $y_{(i)}$ ’s are

the order statistics of the sample. It is “convex shape” for the “monotonically decreasing HRF” and is “concave shape” for the “monotonically increasing HRF”. When solid line matches dashed line, the HRF of the data is “constant”. Figure 8 shows the TTT plot (left panel) and estimated HRF (EHRF) (right panel) for the DRW model for data set I. Based on Table 8, the DRW provides the best fits against all competitive models.

Table 11. MLEs (SEs) for data set I.

	π	θ	λ
Model	MLE (SE)	MLE (SE)	MLE (SE)
DRW	0.992 (0.004)	0.234 (0.019)	
EDW	0.989 (0.164)	1.139 (3.227)	0.784 (3.053)
DW	0.981 (0.011)	1.023 (0.131)	
DIW	0.018 (0.013)	0.582 (0.061)	
DLi-II	0.969 (0.005)	0.058 (0.027)	
EDLi	0.972 (0.005)	0.480 (0.087)	
DLL	1.000 (0.321)	0.439 (0.062)	
DPa	0.739 (0.032)		

Table 12. The goodness of fit (GOF) statistics for data set I.

Model↓	Statistic				
	$-\ell$	AIC	CAIC	K-S	$P_{[V]}$
DRW	237.5	478.9	479.2	0.179	0.081
EDW	240.2	486.7	487.2	0.195	0.045
DW	241.6	487.2	487.5	0.187	0.061
DIW	261.9	527.8	528.1	0.258	0.003
DLi-II	240.6	485.2	485.4	0.186	0.064
EDLi	240.3	484.6	484.8	0.195	0.045
DLL	294.9	593.8	594.0	0.535	<0.001
DPa	275.9	553.7	553.8	0.335	<0.001

Figure 7. The quantile-quantile (Q-Q) plot and box for the failure time data (data set I).

Figure 8. The total time on test (TTT) plot and estimated hazard rate function (EHRF) for the DRW model for dataset I.

7.2. Failure Times of 15 Electronic Components

For this application, we compare the fits of the DRW model with some competitive models such as DGE-II, DLx, DEx, DIR, DR, DIW, DPa, and DBXII. The MLEs with their SEs, and the GOF statistics are reported in Tables 13 and 14, respectively. Figure 9 gives the Q-Q plot and box for the failure times data (data set II). Figure 10 shows the total time test plots (TTT) plot and estimated HRF (EHRF) for the DRW model for data set II. Based on Table 14, the DRW provides the best fits against all competitive models.

Table 13. MLEs (SEs) for dataset II.

	π	θ
Model	MLE (SE)	MLE (SE)
DRW	0.992 (0.006)	0.264 (0.033)
DGE-II	0.956 (0.013)	1.491 (0.535)
DIW	2.2×10^{-4} (7.8×10^{-4})	0.875 (0.164)
DLx	0.012 (0.039)	104.506 (84.409)
DBXII	0.975 (0.051)	13.367 (27.785)
DR	0.999 (2.58×10^{-4})	
DIR	1.8×10^{-7} (0.055)	
DPa	0.720 (0.061)	
DE	0.965 (0.009)	

Table 14. The GOF statistics for dataset II.

Model↓	Statistic				
	$-\ell$	AIC	CAIC	K-S	$P_{[V]}$
DRW	64.0	132.1	133.1	0.116	0.988
DE	65.0	134.0	136.3	0.177	0.673
DGE-II	64.4	134.8	135.8	0.129	0.937
DR	66.4	134.8	136.1	0.216	0.433
DIR	89.1	180.2	180.5	0.698	<0.001
DIW	68.7	141.4	142.4	0.209	0.482
DLx	65.9	135.7	136.7	0.205	0.491
DB-XII	75.7	155.5	156.5	0.388	0.015
DPa	77.4	156.8	157.1	0.405	0.009

Q-Q Plot **Box Plot****Figure 9.** Q-Q plot and box for the failure time data (dataset II). **TTT Plot** **EHRF for the DRW model****Figure 10.** TTT plot and EHRF for the DRW model for dataset II.

7.3. Counts of Cysts of Kidneys

For this real data set, we compare the fits of the DRW distribution with DW, DR, DIW, DE, DLx, DLI-II, DLI, and Poisson. The MLEs with their SEs are listed in Table 15. Table 16 gives the GOF statistics. Figure 11 gives the TTT plot, Q-Q plot and Box plot versus the EHRFs for data set III. Figure 12 presents the fitted PMFs and EHRF for data set III. Based on Table 16, the DRW provides the best fits against all competitive models.

Table 15. MLEs (SEs) for data set III.

Model	π	θ
	MLE (SE)	MLE (SE)
DRW	0.745 (0.028)	0.183 (0.019)
DW	0.750 (0.084)	0.431 (0.340)
DIW	0.581 (0.048)	1.049 (0.146)
DLi-II	0.581 (0.045)	0.001 (0.058)
DLx	0.150 (0.098)	1.830 (0.951)
DR	0.901 (0.009)	
DE	0.581 (0.030)	
DLi	0.436 (0.026)	
Poisson	1.390 (0.112)	

Table 16. The GOF statistics for dataset III.

Z	OF	DRW	DW	DIW	DR	DEx	DLi	DLi-II	DLx	Poisson
0	65	63.83	59.01	63.91	11.00	46.09	40.25	46.03	61.89	27.42
1	14	16.53	19.84	20.70	26.83	26.78	29.83	26.77	21.01	38.08
2	10	9.306	10.78	8.05	29.55	15.56	18.36	15.57	9.65	26.47
3	6	5.900	6.26	4.23	22.23	9.04	10.35	9.05	5.24	12.26
4	4	3.970	4.19	2.60	12.49	5.25	5.53	5.27	3.17	4.26
5	2	2.76	2.01	1.75	5.42	3.05	2.86	3.06	2.06	1.18
6	2	1.967	1.99	1.26	1.85	1.77	1.44	1.78	1.42	0.27
7	2	1.427	1.32	0.95	0.52	1.03	0.71	1.04	1.02	0.05
8	1	1.049	0.99	0.74	0.11	0.60	0.35	0.60	0.76	0.01
9	1	0.779	0.86	0.59	0.02	0.35	0.17	0.35	0.58	0.00
10	1	0.584	0.76	0.48	0.00	0.20	0.08	0.20	0.46	0.00
11	2	0.441	1.99	4.74	0.00	0.28	0.07	0.28	2.74	0.00
$-\ell$		167.441	170.14	172.93	277.78	178.77	189.1	178.8	170.48	246.21
AIC		338.883	344.28	349.87	557.56	359.53	380.2	361.5	344.96	494.42
CAIC		338.995	344.39	349.98	557.59	359.57	380.3	361.6	345.07	494.46
χ^2		0.77951	3.125	6.463	321.07	22.88	43.48	22.89	3.316	294.10
d.f		3	3	3	4	4	4	3	3	4
$P_{[V]}$		0.85436	0.373	0.091	<0.0001	0.0001	<0.0001	<0.0001	0.345	<0.0001

Figure 11. TTT plot, Q-Q plot, and Box plot versus the EHRFs for data set III.

Figure 12. The fitted PMFs and EHRF for data set III.

7.4. Number of European Corn-Borer Larvae Parasites

We shall compare the fits of the DRW distributions with DIW, DGIW, DIR, DP_a, DR, DBXII, negative binomial (NB), and Poisson distributions. The MLEs with their corresponding SEs are listed in Table 17. Table 18 gives the GOF statistics. Figure 13 gives the TTT plot, Q-Q plot and box plot versus the EHRFs for data set IV. Figure 14 presents the fitted PMFs and EHRF for data set IV. Based on Table 18, the DRW provides the best fits against all competitive models

Table 17. MLEs (SEs) for dataset IV.

	π	θ	λ	β
Model	MLE	MLE	MLE	MLE
DRW	0.849 (0.019)	0.298 (0.022)		
DGW	0.0450 (0.429)	2.539 (4.703)	2.159 (2.698)	0.479 (0.466)
DIW	0.345 (0.043)	1.541 (0.156)		
DBXII	0.519 (0.051)	2.358 (0.366)		
NB	0.870 (0.036)	9.956 (0.096)		
DIR	0.319 (0.042)			
DR	0.867 (0.012)			
DPa	0.329 (0.034)			
Poisson	1.483 (0.025)			

Table 18. The GOF statistics for dataset IV.

Z	OF	DRW	DIW	DBXII	DIR	DR	NB	DPa	Poisson
0	43	46.103	41.37	43.84	38.28	15.92	30.12	64.45	27.23
1	35	28.007	41.85	39.61	51.90	36.17	38.87	20.15	40.38
2	17	18.637	15.42	15.62	15.51	34.58	27.61	9.69	29.95
3	11	11.82	7.17	7.20	6.04	21.03	14.26	5.65	14.81
4	5	7.107	3.94	3.91	2.91	8.89	5.99	3.68	
5	4	4.048	2.42	2.37	1.61	2.70	2.17	2.58	1.63
6	1	2.186	1.61	1.56	0.98	0.60	0.70	1.90	0.40
7	2	1.119	1.13	1.09	0.64	0.09	0.21	1.46	0.09
8	2	0.543	5.09	4.80	2.14	0.02	0.06	10.44	0.02
$-\ell$		201.058	204.810	204.293	208.440	235.23	211.52	220.63	219.19
AIC		406.115	413.621	412.587	418.881	472.45	427.05	443.24	440.38
CAIC		406.218	413.723	412.689	418.915	472.49	427.14	443.27	440.41
χ^2		2.223	5.511	4.664	14.274	70.688	20.367	32.462	38.478
d.f		3	3	3	4	4	3	4	4
$P_{[V]}$		0.527	0.138	0.198	<0.0001	<0.0001	0.0001	<0.0001	<0.0001

Figure 13. TTT plot, Q-Q plot, and box plot versus the EHRFs for data set IV.

Figure 14. The fitted PMFs and EHRF for data set IV.

8. Concluding Remarks

In this work, we presented and studied a new discrete analogue based on the continuous Rayleigh distribution called the discrete Rayleigh G (DRG) family of distributions. Some of its statistical properties, such as moments, cumulant generating function, moment generating function, probability generating function, central moment, and DisIx are derived. A special discrete version of the DRG family related to a Weibull distribution is discussed. The HRF of the new family can be “decreasing”, “upside down”, “increasing” and “decreasing-constant-increasing (U-HRF)”. The PMF of the DRG can be “heavy tail right skewed PMF” and “left skewed PMF”. Some classical (non-Bayesian) estimation methods, such as Cramér–von Mises estimation, the maximum likelihood estimation, the weighted least squared estimation, and the ordinary least squared estimation are described and considered.

Some useful characterization results based on the conditional expectation of certain function of the random variable and in terms of the hazard function are derived and presented.

The Bayesian procedure under the SELF is also presented. Since the conditional posteriors of the parameters cannot be obtained in any standard forms, using a hybrid MCMC for drawing sample from the joint posterior of the parameters is suggested. MCMC simulations for comparing non-Bayesian and Bayesian estimation are performed. Gibbs sampler and M-H algorithm are employed and applied. The Bayesian method is the best for all sample sizes with the minimum mean squared errors. The non-Bayesian estimation methods perform well but not better than the Bayesian method. The performance of all estimation methods improves when $n \rightarrow \infty$.

Four applications to real data sets are employed for comparing the Bayesian and non-Bayesian methods. The importance and flexibility of the new discrete class is illustrated by means of four real data applications. Many special member distributions could be considered and studied in the future in separate papers. As a future work, we may consider the bivariate and the multivariate extensions of the DRG family. We hope that the DRG family will attract wider applications in reliability, engineering, and other areas of research.

Author Contributions: M.A. (writing—review and editing), H.M.Y. (writing the original draft preparation, validation, review and editing, and conceptualization), G.G.H. (characterization, writing—review and editing), M.I. (writing the original draft preparation, validation, review and editing, and conceptualization). All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors gratefully acknowledge, with thanks, the very thoughtful and constructive comments and suggestions of the four reviewers, which resulted in a much improved paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1

Proof of Theorem 1. Since

$$\mu'_r = E(Y^r) = \sum_{y=0}^{\infty} y^r S_{\pi, \underline{\phi}}(y).$$

Then,

$$\begin{aligned} \mu'_r = E(Y^r) &= \sum_{y=0}^{\infty} y^r \left[\pi^{\nabla_{\underline{\phi}}^2(y)} - \pi^{\nabla_{\underline{\phi}}^2(y+1)} \right] = \sum_{y=1}^{\infty} [y^r - (y-1)^r] S_{\pi, \underline{\phi}}(y-1) \\ &= \sum_{y=1}^{\infty} [y^r - (y-1)^r] \pi^{\nabla_{\underline{\phi}}^2(y)} \Big|_{(y \in \mathbb{N}^{\bullet}, \pi \in (0,1) \text{ and } r=1,2,3,\dots)}. \end{aligned}$$

The MGF of our NNRV Y can be derived from

$$M_Y(t) = \sum_{y=0}^{\infty} \exp(tY) \pi^{\nabla_{\underline{\phi}}^2(y)}.$$

□

Appendix A.2

Proof of Theorem 2. Using Equation (3) we have

$$M_Y(t) = \sum_{y=0}^{\infty} \exp(tY) \left[\pi^{\nabla_{\underline{\phi}}^2(y)} - \pi^{\nabla_{\underline{\phi}}^2(y+1)} \right],$$

then,

$$M_Y(t) = \pi \frac{\nabla^2_{\underline{\Phi}}(0)}{\underline{\Phi}} + \left[\pi \frac{\nabla^2_{\underline{\Phi}}(1)}{\underline{\Phi}} - \pi \frac{\nabla^2_{\underline{\Phi}}(1)}{\underline{\Phi}} \right] + \left[\pi \frac{\nabla^2_{\underline{\Phi}}(2)}{\underline{\Phi}} - \pi \frac{\nabla^2_{\underline{\Phi}}(2)}{\underline{\Phi}} \right] + \left[\pi \frac{\nabla^2_{\underline{\Phi}}(3)}{\underline{\Phi}} - \pi \frac{\nabla^2_{\underline{\Phi}}(3)}{\underline{\Phi}} \right] + \dots,$$

finally

$$M_Y(t) = 1 + \sum_{y=1}^{\infty} \{ \exp(tY) - \exp[t(Y-1)] \} \pi \frac{\nabla^2_{\underline{\Phi}}(y)}{\underline{\Phi}} \Bigg|_{(y \in \mathbb{N}^*, \pi \in (0,1) \text{ and } r=1,2,3,\dots)}.$$

□

References

- Bourguignon, M.; Silva, R.B.; Cordeiro, G.M. The Weibull-G family of probability distributions. *J. Data Sci.* **2014**, *12*, 53–68.
- Consul, P.C.; Jain, G.C. A generalization of the Poisson distribution. *Technometrics* **1973**, *15*, 791–799.
- Nakagawa, T.; Osaki, S. The discrete Weibull distribution. *IEEE Trans. Reliab.* **1975**, *24*, 300–301.
- Roy, D. Discrete Rayleigh distribution. *IEEE Trans. Reliab.* **2004**, *53*, 255–260.
- Kemp, A.W. The discrete half-normal distribution. In *Advances in Mathematical and Statistical Modeling*; Birkhauser: Basel, Switzerland, 2008; pp. 353–365.
- Krishna, H.; Pundir, P.S. Discrete Burr and discrete Pareto distributions. *Stat. Methodol.* **2009**, *6*, 177–188.
- Gomez-Déniz, E. Another generalization of the geometric distribution. *Test* **2010**, *19*, 399–415.
- Jazi, A.M.; Lai, D.C.; Alamatsaz, H.M. Inverse Weibull distribution and estimation of its parameters. *Stat. Methodol.* **2010**, *7*, 121–132.
- Gomez-Déniz, E.; Caldern-Ojeda, E. The discrete Lindley distribution: Properties and applications. *J. Stat. Comput. Simul.* **2011**, *81*, 1405–1416.
- Nekoukhou, V.; Alamatsaz, M.H.; Bidram, H. Discrete generalized exponential distribution of a second type. *Statistics* **2013**, *47*, 876–887.
- Hussain, T.; Ahmad, M. Discrete inverse Rayleigh distribution. *Pak. J. Stat.* **2014**, *30*, 203–222.
- Nekoukhou, V.; Bidram, H. The exponentiated discrete Weibull distribution. *Stat. Oper. Res. Trans.* **2015**, *39*, 127–146.
- Hussain, T.; Aslam, M.; Ahmad, M.A. Two parameter discrete Lindley distribution. *Rev. Colomb. Estad.* **2016**, *39*, 45–61.
- Para, B.A.; Jan, T.R. On discrete three-parameter Burr type XII and discrete Lomax distributions and their applications to model count data from medical science. *Biom. Biostat. Int. J.* **2016**, *4*, 1–15.
- Para, B.A.; Jan, T.R. Discrete version of log-logistic distribution and its applications in genetics. *Int. J. Mod. Math. Sci.* **2016**, *14*, 407–422.
- El-Morshedy, M.; Eliwa, M.S.; Nagy, H. A new two-parameter exponentiated discrete Lindley distribution: Properties, estimation and applications. *J. Appl. Stat.* **2020**, *47*, 354–375. [\[CrossRef\]](#)
- Aryal, G.R.; Yousof, H.M. The exponentiated generalized-G Poisson family of distributions. *Econ. Qual. Control.* **2017**, *32*, 1–17.
- Hamedani, G.G.; Yousof, H.M.; Rasekhi, M.; Alizadeh, M.; Najibi, S.M. Type I general exponential class of distributions. *Pak. J. Stat. Oper. Res.* **2017**, *XIV*, 39–55.
- Hamedani, G.G.; Altun, E.; Korkmaz, M.C.; Yousof, H.M.; Butt, N.S. A new extended G family of continuous distributions with mathematical properties, characterizations and regression modeling. *Pak. J. Stat. Oper. Res.* **2018**, *14*, 737–758.
- Hamedani, G.G.; Rasekhi, M.; Najibi, S.M.; Yousof, H.M.; Alizadeh, M. Type II general exponential class of distributions. *Pak. J. Stat. Oper. Res.* **2019**, *XV*, 503–523.
- Korkmaz, M.C.; Altun, E.; Yousof, H.M.; Hamedani, G.G. The odd power Lindley generator of probability distributions: Properties, characterizations and regression modeling. *Int. J. Stat. Probab.* **2019**, *8*, 70–89.
- Nascimento, A.D.C.; Silva, K.F.; Cordeiro, G.M.; Alizadeh, M.; Yousof, H.M. The odd Nadarajah-Haghighi family of distributions: Properties and applications. *Stud. Sci. Math. Hung.* **2019**, *56*, 1–26. [\[CrossRef\]](#)
- Cordeiro, G.M.; Yousof, H.M.; Ramires, T.G.; Ortega, E.M.M. The Burr XII system of densities: Properties, regression model and applications. *J. Stat. Comput. Simul.* **2018**, *88*, 432–456. [\[CrossRef\]](#)

24. Korkmaz, M.C.; Yousof, H.M.; Hamedani, G.G.; Ali, M.M. The Marshall–Olkin generalized G Poisson family of distributions. *Pak. J. Stat.* **2018**, *34*, 251–267.
25. Khalil, M.G.; Hamedani, G.G.; Yousof, H.M. The Burr X exponentiated Weibull model: Characterizations, mathematical properties and applications to failure and survival times data. *Pak. J. Stat. Oper. Res.* **2019**, *XV*, 141–160. [[CrossRef](#)]
26. Ibrahim, M.; Yousof, H.M. Transmuted Topp–Leone Weibull Lifetime Distribution: Statistical Properties and Different Method of Estimation. *Pak. J. Stat. Oper. Res.* **2020**, *16*, 501–515. [[CrossRef](#)]
27. Ibrahim, M.; Mohammed, W.S.; Yousof, H.M. Bayesian and Classical Estimation for the One Parameter Double Lindley Model. *Pak. J. Stat. Oper. Res.* **2020**, *16*, 409–420.
28. Korkmaz, M.Ç.; Yousof, H.M. The one-parameter odd Lindley exponential model: Mathematical properties and applications. *Stoch. Qual. Control* **2017**, *32*, 25–35. [[CrossRef](#)]
29. Yousof, H.M.; Korkmaz, M.Ç.; Hamedani, G.G. The odd Lindley Nadarajah–Haghighi distribution. *J. Math. Comput. Sci.* **2017**, *7*, 864–882.
30. Gad, A.M.; Hamedani, G.G.; Salehabadi, S.M.; Yousof, H.M. The Burr XII–Burr XII distribution: Mathematical properties and characterizations. *Pak. J. Stat.* **2019**, *35*, 229–248.
31. Cai, L. Metropolis–Hastings Robbins–Monro algorithm for confirmatory item factor analysis. *J. Educ. Behav. Stat.* **2010**, *35*, 307–335. [[CrossRef](#)]
32. Chib, S.; Greenberg, E. Understanding the metropolis–hastings algorithm. *Am. Stat.* **1995**, *49*, 327–335.
33. Korkmaz, M.C.; Yousof, H.M.; Rasekhi, M.; Hamedani, G.G. The Odd Lindley Burr XII Model: Bayesian Analysis, Classical Inference and Characterizations. *J. Data Sci.* **2018**, *16*, 327–353.
34. Rasekhi, M.; Saber, M.M.; Yousof, H.M. Bayesian and classical inference of reliability in multicomponent stress–strength under the generalized logistic model. *Commun. Stat. Theory Methods* **2020**, 1–12. [[CrossRef](#)]
35. Bodhisuwan, W.; Sangpoom, S. The discrete weighted Lindley distribution. In Proceedings of the International Conference on Mathematics, Statistics, and Their Applications, Banda Aceh, Indonesia, 4–6 October 2016.
36. Lawless, J.F. *Statistical Models and Methods for Lifetime Data*; Wiley: New York, NY, USA, 2003.
37. Chan, S.; Riley, P.R.; Price, K.L.; McElduff, F.; Winyard, P.J. Corticosteroid-induced kidney dysmorphogenesis is associated with deregulated expression of known cystogenic molecules, as well as Indian hedgehog. *Am. J. Physiol. Ren. Physiol.* **2009**, *298*, 346–356. [[CrossRef](#)]
38. Bebbington, M.; Lai, C.D.; Wellington, M.; Zitikis, R. The discrete additive Weibull distribution: A bathtub-shaped hazard for discontinuous failure data. *Reliab. Eng. Syst. Saf.* **2012**, *106*, 37–44. [[CrossRef](#)]
39. Aarset, M.V. How to identify a bathtub hazard rate. *IEEE Trans. Reliab.* **1987**, *36*, 106–108. [[CrossRef](#)]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).