ISSN 1995-0802, Lobachevskii Journal of Mathematics, 2020, Vol. 41, No. 6, pp. 945–955. © Pleiades Publishing, Ltd., 2020.

Discrete Generalized Odd Lindley–Weibull Distribution with Applications

Sirinapa Aryuyuen[1*], Winai Bodhisuwan[#2], and Andrei Volodin[3***]**

(Submitted by A. M. Elizarov)

[1]*Department of Mathematics and Computer Science, Rajamangala University of Technology Thanyaburi, Pathum Thani, 12110 Thailand*

[2]*Department of Statistics, Faculty of Science, Kasetsart University, Chatuchak, Bangkok, 10900 Thailand*

[3]*Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, S4S 0A2 Canada*

Received November 15, 2019; revised January 10, 2020; accepted February 8, 2020

Abstract—In this paper, we introduce a new probability mass function by discretizing the continuous failure model of the generalized odd Lindley–Weibull distribution, which is called the discrete generalized odd Lindley–Weibull (DGOL-W) distribution. This new probability mass function is characterized by a very flexible probability function: reverse J-shape, right-skewed shape, left-skewed shape, and close to symmetric shape. The proposed distribution has five special models, i.e., the discrete generalized odd Lindley-exponential, discrete generalized odd Lindley–Rayleigh, discrete odd Lindley–Weibull, discrete odd Lindley-exponential, and discrete odd Lindley–Rayleigh distributions. Some properties of the proposed distribution are introduced. The maximum likelihood estimation is used to estimate the unknown parameters of the DGOL-W distribution. Applications are illustrated, which show that the model is suited for use in various data sets, i.e., the mean and variance of the count data are equal, over-dispersion count data, and under-dispersion count data. Based on the results, we have shown that the DGOL-W distribution provides a better fit compared to the Poisson, discrete Lindley and four sub-models of DGOL-W distribution for count data.

DOI: 10.1134/S1995080220060037

Keywords and phrases: *discrete generalized odd Lindley–Weibull, quantile function, count data, probability distribution, estimation.*

1. INTRODUCTION

In the field of reliability theory, modeling of lifetime data is very important. Statistical distributions such as exponential, lognormal, Weibull, Lindley, Rayleigh, etc., are available for modeling lifetime data. However, in many practical areas, these distributions do not provide adequate fit in modeling data, and there is a clear need for the extended version of the distributions (Ahmad et al. (2018) [1]). Recently, Afify et al. (2019) [2] proposed the generalized odd Lindley–Weibull (GOL-W) distribution, which is a new class of the generalized odd Lindley-G family of distribution. The GOL-W distribution has the probability density function (pdf) and the cumulative density function (cdf) as follows:

$$g_{\text{GOL-W}}(x) = \frac{\alpha\beta^2\lambda\theta x^{\theta-1}p_\lambda^{x^\theta}\left(1-p_\lambda^{x^\theta}\right)^{\alpha-1}}{(1+\beta)\left[1-\left(1-p_\lambda^{x^\theta}\right)^{\alpha}\right]^3}\exp\left\{-\frac{\beta\left(1-p_\lambda^{x^\theta}\right)^{\alpha}}{1-\left(1-p_\lambda^{x^\theta}\right)^{\alpha}}\right\}, \tag{1.1}$$

*E-mail: sirinapa_a@rmutt.ac.th
**E-mail: fsciwnb@ku.ac.th
***E-mail: andrei.stat@gmail.com
#Corresponding author.

$$G_{\text{GOL-W}}(x) = 1 - \left[1 + \frac{\beta}{1+\beta}\frac{\left(1-p_\lambda^{x^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{x^\theta}\right)^\alpha}\right]\exp\left\{-\beta\frac{\left(1-p_\lambda^{x^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{x^\theta}\right)^\alpha}\right\}; \quad x>0, \tag{1.2}$$

where $0 < p_\lambda = \exp(-\lambda) < 1$ and $\alpha, \beta, \lambda, \theta > 0$. The GOL-W distribution has four special models. That is, if $\theta = 1$ and $\theta = 2$ then the GOL-W distribution reduces to the generalized odd Lindley-exponential and generalized odd Lindley-Rayleigh distributions, respectively. When $\alpha = 1$ the GOL-W distribution reduces to the odd Lindley-exponential distribution for $\theta = 1$ and the odd Lindley—Rayleigh distribution for $\theta = 2$.

Generally, one associates the lifetime of a product with continuous non-negative lifetime distributions, however, in practice, the lifetime can be best described through non-negative integer-valued random variables. Although continuous lifetime distributions are playing their roles in reliability analysis very well, in certain scenarios, when measured data is discrete and realized from a continuous setup, an alternative is needed. For this purpose, researchers developed the discretized version of continuous lifetime distributions. This development is generally based on discrete lifetime phenomena that are expressed through grouping or finite precision measurement of continuous-time phenomena. Therefore, the inference is based on observed discrete values that are only indicative of the intervals to the which unobserved continuous variable belongs, but not its true values. Hence, this is a case where one makes use of a discretization of the underlying continuous variable. For this purpose, the discretized version of continuous lifetime distributions was developed. This development is generally based on discrete lifetime phenomena that are expressed through grouping or finite precision measurement of continuous-time phenomena. In survival analysis, the survival function may be a function of a count random variable that is a discrete version of the underlying continuous random variable. From these examples, it is clear that the continuous lifetime may not necessarily always be measured on a continuous scale, but may often be counted as discrete random variables (see Ahmad et al. (2018) [1], Chakraborty (2015) [3]). The characterization of a probability distribution plays an important role in statistics and mathematical sciences. Many researchers developed a new distributions (Ahsanullah et al. (2015) [4]).

Let X be a random variable that has a lifetime distribution on $[0, \infty)$ with the pdf $g(x)$ and the cdf $G(x)$; one can construct a discrete counterpart supported on the set of integers $0, 1, 2, \ldots,$ whose probability mass function (pmf) is (see Roy (2003), [5], Roy (2004) [6], and Alamatsaz et al. (2016) [7]) given by:

$$f(x) = P(X = x) = S_G(x) - S_G(x+1), \quad x = 0, 1, 2, \ldots, \tag{1.3}$$

where $S(x)$ is a survival function of the lifetime distribution; that is, $S(x) = P(X > x) = 1 - G(x)$.

In Section 2 of this article, a new discretization of a continuous distribution, the GOL-W distribution, is proposed. Properties and application of the proposed distribution are discussed in Sections 3 and 4, respectively. A conclusion is presented in Section 5.

2. A NEW DISCRETIZATION OF A CONTINUOUS DISTRIBUTION

In this section, we provide a new discretization of the GOL-W distribution called the discrete generalized odd Lindley—Weibull (DGOL-W) distribution. From the cdf in equation (1.2), we have the survival function of the GOL-W distribution (see Afify et al. (2019) [2]), i.e.,

$$S_{\text{GOL-W}}(x) = \left[1 + \frac{\beta}{1+\beta}\frac{\left(1-p_\lambda^{x^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{x^\theta}\right)^\alpha}\right]\exp\left\{-\beta\frac{\left(1-p_\lambda^{x^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{x^\theta}\right)^\alpha}\right\}; \quad x>0. \tag{2.1}$$

As a result of the equation (2.1) we have Definition 1, which is presents the pmf of the DGOL-W distribution.

Definition 1. Let X be a DGOL-W distributed random variable with the parameters α, β, λ, and θ, denoted as $X \sim \text{DGOL-W}(\alpha, \beta, \lambda, \theta)$. The pmf of X is then defined by

$$f_{\text{DGOL-W}}(x) = \left[1 + \frac{\beta}{1+\beta}\frac{\left(1-p_\lambda^{x^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{x^\theta}\right)^\alpha}\right]\exp\left\{-\beta\frac{\left(1-p_\lambda^{x^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{x^\theta}\right)^\alpha}\right\}$$

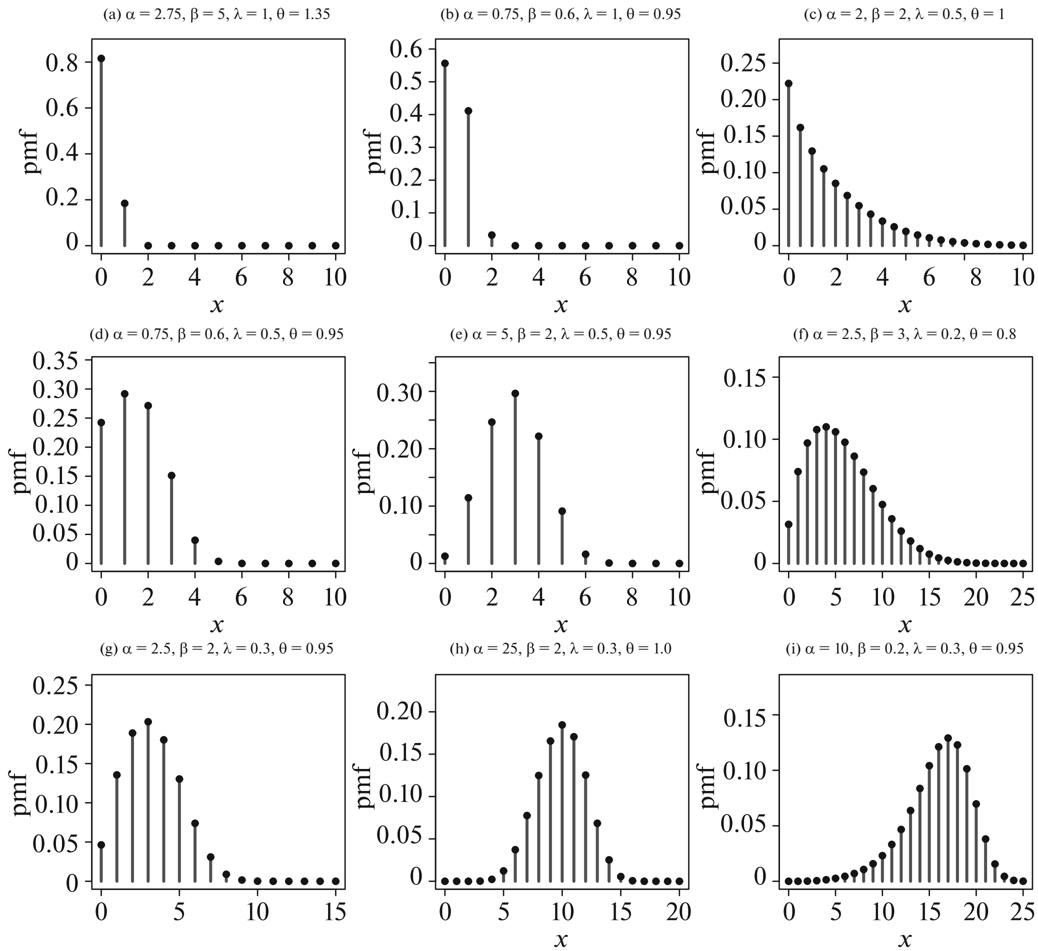

Fig. 1. The pmf plot of the $DGOL\text{-}W$ distribution with the specified parameters of α, β, λ, and θ.

$$-\left[1+\frac{\beta}{1+\beta}\frac{\left(1-p_{\lambda}^{(x+1)^{\theta}}\right)^{\alpha}}{1-\left(1-p_{\lambda}^{(x+1)^{\theta}}\right)^{\alpha}}\right]\exp\left\{-\beta\frac{\left(1-p_{\lambda}^{(x+1)^{\theta}}\right)^{\alpha}}{1-\left(1-p_{\lambda}^{(x+1)^{\theta}}\right)^{\alpha}}\right\}, \tag{2.2}$$

where $x = 0, 1, 2, \ldots$ and $0 < p_{\lambda} = \exp(-\lambda) < 1$.

Figure 1 illustrates the pmf behaviors of the DGOL-W distribution for several values of α, β, λ and θ. The DGOL-W pmf has several behaviors such as the reverse J-shape (see Figure 1 (a)–(c)), the unimodality (see Figure 1 (d)–(i)), the right-skewed shape (see Figure 1 (a)–(d), (f)–(g)), and the left-skewed shape (see Figure 1 (i)). Figure 1 (e), (h) shows the DGOL-W pmf that is close to a symmetric distribution. In addition, the mode of the distribution shifts towards the right when α increases.

Based on $F(x) = P(X \leq x) = 1 - S_G(x) + P(X = x)$ (see Alamatsaz et al. (2016) [7], Jayakumar and Babu (2019) [8]), we have the cdf of the DGOL-W distribution in equation (2.3),

$$F_{\text{DGOL-W}}(x) = 1-\left[1+\frac{\beta}{1+\beta}\frac{\left(1-p_{\lambda}^{(x+1)^{\theta}}\right)^{\alpha}}{1-\left(1-p_{\lambda}^{(x+1)^{\theta}}\right)^{\alpha}}\right]\exp\left\{-\beta\frac{\left(1-p_{\lambda}^{(x+1)^{\theta}}\right)^{\alpha}}{1-\left(1-p_{\lambda}^{(x+1)^{\theta}}\right)^{\alpha}}\right\}, \tag{2.3}$$

where $x = 0, 1, 2, \ldots$. Some cdf plots of X are shown in Figure 2.

We have five special models of the DGOL-W distribution as follows.

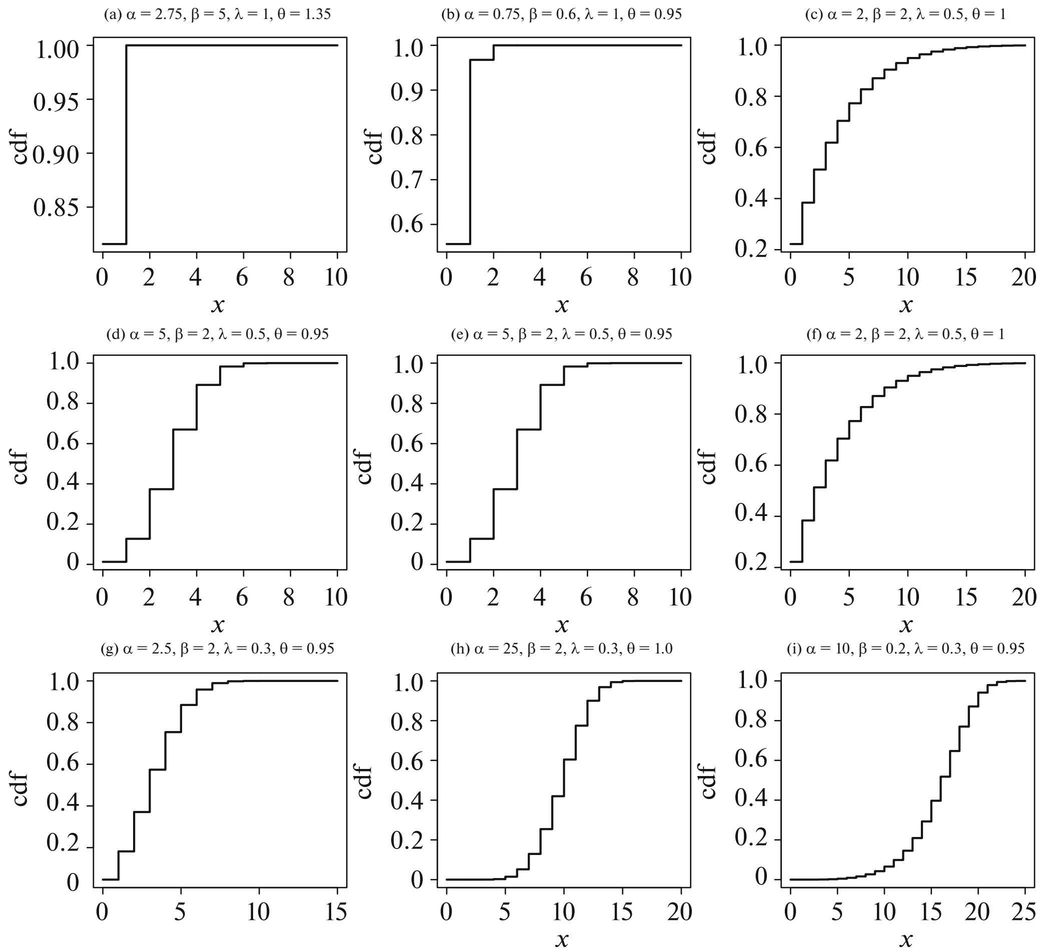

Fig. 2. The cdf plot of the *DGOL-W* distribution with the specified parameters of α, β, λ, and θ.

2.1. The Discrete Generalized Odd Lindley-Exponential (DGOL-E) Distribution

If $X \sim$ DGOL-W$(\alpha,\beta,\lambda,\theta)$ with the pmf in equation (2.2), where $\theta = 1$ then the DGOL-W distribution reduces to the DGOL-E distribution with the pmf

$$f_{\text{DGOL-E}}(x) = \left[1 + \frac{\beta}{1+\beta}\frac{(1-p_\lambda^x)^\alpha}{1-(1-p_\lambda^x)^\alpha}\right]\exp\left\{-\beta\frac{(1-p_\lambda^x)^\alpha}{1-(1-p_\lambda^x)^\alpha}\right\}$$
$$-\left[1 + \frac{\beta}{1+\beta}\frac{(1-p_\lambda^{x+1})^\alpha}{1-(1-p_\lambda^{x+1})^\alpha}\right]\exp\left\{-\beta\frac{[1-p_\lambda^{x+1}]^\alpha}{1-[1-p_\lambda^{x+1}]^\alpha}\right\}, \tag{2.4}$$

where $x = 0,1,2,\ldots$ and $\alpha,\beta,\lambda > 0$.

2.2. The Discrete Generalized odd Lindley-Rayleigh (DGOL-R) Distribution

The DGOL-W$(\alpha,\beta,\lambda,\theta)$ distribution reduces to the DGOL-R distribution for $\theta = 2$. We have the pmf of the DGOL-R distribution as follows

$$f_{\text{DGOL-R}}(x) = \left[1 + \frac{\beta}{1+\beta}\frac{\left(1-p_\lambda^{x^2}\right)^\alpha}{1-\left(1-p_\lambda^{x^2}\right)^\alpha}\right]\exp\left\{-\beta\frac{\left(1-p_\lambda^{x^2}\right)^\alpha}{1-\left(1-p_\lambda^{x^2}\right)^\alpha}\right\}$$
$$-\left\{1 + \frac{\beta}{1+\beta}\frac{\left[1-p_\lambda^{(x+1)^2}\right]^\alpha}{1-\left[1-p_\lambda^{(x+1)^2}\right]^\alpha}\right\}\exp\left\{-\beta\frac{\left[1-p_\lambda^{(x+1)^2}\right]^\alpha}{1-\left[1-p_\lambda^{(x+1)^2}\right]^\alpha}\right\}, \tag{2.5}$$

where $x = 0,1,2,\ldots,$ and $\alpha,\beta,\lambda > 0$.

2.3. The Discrete Odd Lindley–Weibull (DOL-W) Distribution

The DGOL-W distribution with the parameters α, β, λ and θ where $\alpha = 1$ reduces to the DOL-W distribution with the pmf

$$f_{\text{DOL-W}}(x) = \left(1 + \frac{\beta}{1+\beta}\frac{1-p_\lambda^{x^\theta}}{p_\lambda^{x^\theta}}\right)\exp\left(-\beta\frac{1-p_\lambda^{x^\theta}}{p_\lambda^{x^\theta}}\right)$$
$$-\left(1 + \frac{\beta}{1+\beta}\frac{1-p_\lambda^{(x+1)^\theta}}{p_\lambda^{(x+1)^\theta}}\right)\exp\left[-\beta(\frac{1-p_\lambda^{(x+1)^\theta}}{p_\lambda^{(x+1)^\theta}})\right], \quad (2.6)$$

where $x = 0, 1, 2, \ldots$, and β, λ, $\theta > 0$.

2.4. The Discrete Odd Lindley-exponential (DOL-E) Distribution

Let $X \sim$ DGOL-W$(\alpha, \beta, \lambda, \theta)$ for $\alpha = \theta = 1$, we obtain the DOL-E distribution with the positive parameters β and λ. The pmf of the DOL-E distribution is

$$f_{\text{DOL-E}}(x) = \left(1 + \frac{\beta}{1+\beta}\frac{1-p_\lambda^{x}}{p_\lambda^{x^\theta}}\right)\exp\left(-\beta\frac{1-p_\lambda^{x}}{p_\lambda^{x^\theta}}\right)$$
$$-\left(1 + \frac{\beta}{1+\beta}\frac{1-p_\lambda^{(x+1)}}{p_\lambda^{(x+1)}}\right)\exp\left[-\beta(\frac{1-p_\lambda^{x+1}}{p_\lambda^{x+1}})\right]; \quad x = 0, 1, 2, \ldots. \quad (2.7)$$

2.5. The Discrete Odd Lindley–Rayleigh (DOL-R) Distribution

Let $X \sim$ DGOL-W$(\alpha, \beta, \lambda, \theta)$ for $\alpha = 1$ and $\theta = 2$, we have the DOL-R distribution with the positive parameters β and λ. The pmf of DOL-R is given by

$$f_{\text{DOL-R}}(x) = \left(1 + \frac{\beta}{1+\beta}\frac{1-p_\lambda^{x^2}}{p_\lambda^{x^2}}\right)\exp\left(-\beta\frac{1-p_\lambda^{x^2}}{p_\lambda^{x^\theta}}\right)$$
$$-\left(1 + \frac{\beta}{1+\beta}\frac{1-p_\lambda^{(x+1)^2}}{p_\lambda^{(x+1)^2}}\right)\exp\left[-\beta(\frac{1-p_\lambda^{(x+1)^2}}{p_\lambda^{(x+1)^2}})\right]; \quad x = 0, 1, 2, \ldots. \quad (2.8)$$

3. CHARACTERISTIC PROPERTIES OF THE DGOL-W DISTRIBUTION

3.1. Quantile Function

Let X be a random variable that Weibull distribution with positive parameters λ and θ. Then, the quantile function of X is

$$G_{\text{Weibull}}(x) = 1 - \exp^{-\lambda x^\theta} \quad \text{and} \quad G_{\text{Weibull}}^{-1}(u) = \left[-\frac{1}{\lambda}\log(1-u)\right]^{1/\theta}.$$

If X is a GOL-W random variable, then the quantile function is

$$Q_{\text{GOL-W}}(u) = G_{\text{Weibull}}^{-1}\left[\left(\frac{\beta + \mathrm{W}_{-1}\left[-(1-u)(\beta+1)\exp(-\beta-1)\right]+1}{W_{-1}\left[-(1-u)(\beta+1)\exp(-\beta-1)\right]+1}\right)^{1/\alpha}\right]$$
$$= \left\{-\frac{1}{\lambda}\log\left[1 - \left(\frac{\beta + \mathrm{W}_{-1}\left[-(1-u)(\beta+1)\exp(-\beta-1)\right]+1}{W_{-1}\left[-(1-u)(\beta+1)\exp(-\beta-1)\right]+1}\right)^{1/\alpha}\right]\right\}^{1/\theta}, \quad (3.1)$$

where $W(z)$ is the negative branch of the Lambert function. The branches of this function are defined by $z = W[z\exp(z)]$. It is a two-valued function on the interval $[-1/\exp(1), 0\ )$. For $W(z) \leq -1$, the function is denoted $W_{-1}(z)$ and is called the negative branch. For $W(z) > -1$, the function is called the principal branch of the W function. The Lambert function cannot be expressed in terms of elementary functions (see Afify et al. (2019) [2]).

From the cdf of the DGOL-W distribution in equation (2.3), the quantile function of the DGOL-W distribution denoted by $Q_F(u)$ can be obtained by inverting its distribution function of it. When u is on [0,1] then $F_{\text{DGOL-W}}(Q(u)) = u$. The quantile function of the DGOL-W distribution, i.e., $Q_{\text{DGOL-W}}(u) = Q_{\text{F}}(u)$ is given by

$$Q_{\text{F}}(u) = \left\lfloor \left\{ -\frac{1}{\lambda} \log \left[1 - \left(\frac{\beta + \text{W}_{-1}\left[-(1-u)(\beta+1)\exp(-\beta-1)\right]+1}{W_{-1}\left[-(1-u)(\beta+1)\exp(-\beta-1)\right]+1} \right)^{1/\alpha} \right] \right\}^{1/\theta} - 1 \right\rfloor, \quad (3.2)$$

where $\lfloor\cdot\rfloor$ is the floor function.

To generate a random variable X from the DGOL-W$(\alpha, \beta, \lambda, \theta)$, which is based on generating random data by inverting equation (3.2), one can use the following algorithm:

(1) Set the values of α, β, λ and θ.

(2) Set the sample size of n.

(3) Generate U_i according to the uniform distribution on interval (0,1) where $i = 1, 2, \dots, n$.

(4) Make the transformation $X_i = \left\lfloor \left\{ -\frac{1}{\lambda} \log \left[1 - \left(\frac{\beta+\text{W}_{-1}[-(1-u)(\beta+1)\exp(-\beta-1)]+1}{W_{-1}[-(1-u)(\beta+1)\exp(-\beta-1)]+1} \right)^{1/\alpha} \right] \right\}^{1/\theta} - 1 \right\rfloor$.

3.2. Order Statistics

Let $X_1, X_2, \dots, X_n$ be n independent and identically distributed (iid) random variables, i.e., $X_i \sim$ DGOL-W$(\alpha, \beta, \lambda, \theta)$, each with cdf and pmf in equations (2.3) and (2.2), respectively. Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote these random variables rearranged in non-descending order of magnitude. Then, the pmf of the rth order statistic can always be expressed as

$$\begin{aligned} f_{\text{DGOL-W}_{(r)}}(x) &= \frac{n!}{(r-1)!(n-r)!} f_{\text{DGOL-W}}(x) [F_{\text{DGOL-W}}(x)]^{r-1} [1 - F_{\text{DGOL-W}}(x)]^{n-r} \\ &= \frac{n!}{(r-1)!(n-r)!} \left\{ \left[1 + \frac{\beta}{1+\beta} \frac{\left(1 - p_\lambda^{x^\theta}\right)^\alpha}{1 - \left(1 - p_\lambda^{x^\theta}\right)^\alpha} \right] \exp\left[-\beta \frac{\left(1 - p_\lambda^{x^\theta}\right)^\alpha}{1 - \left(1 - p_\lambda^{x^\theta}\right)^\alpha} \right] \right. \\ &\quad \left. - \left[1 + \frac{\beta}{1+\beta} \frac{\left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha}{1 - \left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha} \right] \exp\left[-\beta \frac{\left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha}{1 - \left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha} \right] \right\} \\ &\quad \times \left\{ 1 - \left[1 + \frac{\beta}{1+\beta} \frac{\left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha}{1 - \left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha} \right] \exp\left[-\beta \frac{\left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha}{1 - \left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha} \right] \right\}^{r-1} \\ &\quad \times \left\{ \left[1 + \frac{\beta}{1+\beta} \frac{\left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha}{1 - \left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha} \right] \exp\left[-\beta \frac{\left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha}{1 - \left(1 - p_\lambda^{(x+1)^\theta}\right)^\alpha} \right] \right\}^{n-r}, \end{aligned} \quad (3.3)$$

where $x = 0, 1, 2, \dots$ and $r = 1, 2, \dots, n$.

3.3. Index of Dispersion

The index of dispersion (ID) for any distribution indicates whether the distribution is over-dispersed (ID > 1) or under-dispersed (ID < 1) (see Chakraborty and Chakravarty, 2012 [9]). The ID for any distribution is defined as the ratio between variance to mean. From the first and second moments of the distribution in equation (2.2), the mean and variance of the DGOL-W distribution are given:

$$\text{E}(X) = \sum_X x p_x(\alpha, \beta, \lambda, \theta) \quad \text{and} \quad \text{V}(X) = \sum_X x^2 p_x(\alpha, \beta, \lambda, \theta) - \text{E}^2(X); \quad x = 0, 1, 2, \dots,$$

where $p_x(\alpha, \beta, \lambda, \theta)$ is the pmf as in equation (2.2). The above expressions are infinite series and cannot be written in closed forms. However, the value of the mean, variance, and ID are shown in Table 1 for examples under some specified values of parameters α, β, λ, and θ from Fig. 1.

Table 1. The mean, variance and ID values of the DGOL-W distribution for different value of α, β, λ, and θ

Figure	α	β	λ	θ	E(X)	V(X)	ID
1a	2.75	5	1	1.35	0.1843	0.1503	0.8157
1b	0.75	0.6	1	0.95	0.4762	0.3144	0.6602
1c	2	2	0.5	0.5	3.3984	12.1689	3.5807
1d	0.75	0.6	0.5	0.95	1.4660	1.3198	0.9003
1e	5	2	0.5	0.95	2.9429	1.5600	0.5301
1f	2.5	3	0.2	0.8	5.7382	12.9131	2.2504
1g	2.5	2	0.3	0.95	3.2420	3.2373	0.9986
1h	25	2	0.3	1	9.8865	4.1936	0.4242
1i	10	0.2	0.3	0.95	15.9686	10.9512	0.6858

3.4. The Parameter Estimation of the DGOL-W Distribution

In this section, we present the maximum likelihood estimation (MLE) of the parameters for the DGOL-W distribution. Let $X_1, X_2, \ldots, X_n$ be n iid random variables with DGOL-W$(\alpha, \beta, \lambda, \theta)$ distribution, then the likelihood function of the DGOL-W distribution is given by:

$$L(\alpha,\beta,\lambda,\theta) = \prod_{i=1}^{n} \left\{ \left[1 + \frac{\beta}{1+\beta} \frac{\left(1-p_\lambda^{x_i^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{x_i^\theta}\right)^\alpha} \right] \exp\left(-\beta \frac{\left(1-p_\lambda^{x_i^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{x_i^\theta}\right)^\alpha} \right) \right.$$
$$\left. - \left[1 + \frac{\beta}{1+\beta} \frac{\left(1-p_\lambda^{(x_i+1)^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{(x_i+1)^\theta}\right)^\alpha} \right] \exp\left(-\beta \frac{\left(1-p_\lambda^{(x_i+1)^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{(x_i+1)^\theta}\right)^\alpha} \right) \right\}.$$

The corresponding log-likelihood equation is

$$\ell(\alpha,\beta,\lambda,\theta) = \sum_{i=1}^{n} \log \left\{ \left[1 + \frac{\beta}{1+\beta} \frac{\left(1-p_\lambda^{x_i^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{x_i^\theta}\right)^\alpha} \right] \exp\left(-\beta \frac{\left(1-p_\lambda^{x_i^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{x_i^\theta}\right)^\alpha} \right) \right.$$
$$\left. - \left[1 + \frac{\beta}{1+\beta} \frac{\left(1-p_\lambda^{(x_i+1)^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{(x_i+1)^\theta}\right)^\alpha} \right] \exp\left(-\beta \frac{\left(1-p_\lambda^{(x_i+1)^\theta}\right)^\alpha}{1-\left(1-p_\lambda^{(x_i+1)^\theta}\right)^\alpha} \right) \right\}.$$

To estimate the unknown parameters α, β, λ and θ, we take the partial derivatives of the log-likelihood function $\ell(\alpha, \beta, \lambda, \theta)$ with respect to α, β, λ and θ and equate them to zero, i.e.

$$\frac{\partial \ell(\alpha,\beta,\lambda,\theta)}{\partial \alpha} = 0, \frac{\partial \ell(\alpha,\beta,\lambda,\theta)}{\partial \beta} = 0, \frac{\partial \ell(\alpha,\beta,\lambda,\theta)}{\partial \lambda} = 0, \frac{\partial \ell(\alpha,\beta,\lambda,\theta)}{\partial \theta} = 0. \tag{3.4}$$

The equations (3.4) cannot be solved in closed form. We obtain the solutions of the maximum likelihood estimators (MLEs) of α, β, λ and θ from the equation (3.4) by using the numerical method of the four-dimensional Newton-Raphson type procedure. Solving the score equations simultaneously using the *nlm* function in the statistical software package R (R Development Core Team (2018) [10]) is used to estimate the parameters.

4. APPLICATIONS

In this section, the DGOL-W distribution is applied to three real data sets. The first data set (Data I) is the number of deaths due to horse kicks (X) in the Prussian army between 1875 and 1894 (see Klugman et al. (2012) [11]). The mean, variance, and ID values of Y are 0.6100, 0.6100 and 1.0016,

Table 2. Comparisons of the observed and expected values of Data I, and criteria values of each distribution

X	Observed frequency	Expected frequency							
		Poisson	DL	DOL-R	DOL-E	DOL-W	DGOL-R	DGOL-E	DGOL-W
0	109	108.67	119.42	85.12	110.37	110.16	113.10	110.82	109.02
1	65	66.29	52.46	93.17	62.47	62.52	56.91	60.78	65.42
2	22	20.22	19.01	20.37	22.43	22.78	25.61	24.00	20.83
3	3	4.11	6.29	1.31	4.35	4.19	4.26	4.18	4.15
4	1	0.63	1.97	0.03	0.37	0.34	0.11	0.23	0.53
MLEs	$\hat{\mu}$	0.6100	–	–	–	–	–	–	–
	$\hat{\alpha}$	–	–	–	–	–	0.0081	0.0119	11.6588
	$\hat{\beta}$	–	–	693.7769	2.7067	0.7331	0.0452	0.0244	1.0149
	$\hat{\lambda}$	–	1.3647	0.0008	0.3285	1.0628	0.0524	0.3994	3.0187
	$\hat{\theta}$	–	–	–	–	0.6136	–	–	0.2899
	$-2\hat{L}$	206.1067	209.9637	217.0148	206.3491	206.3639	208.0777	206.8113	206.0931
	AIC	414.2134	421.9274	438.0296	416.6982	418.7278	422.1554	419.6226	420.1862
	BIC	417.5117	425.2257	444.6262	423.2948	428.6228	432.0504	429.5176	433.3795
	K-S	0.0048	0.0521	0.1194	0.0069	0.0066	0.0204	0.0120	0.0036

Table 3. Comparison of the observed and expected values of Data III, and criteria values of each distribution

Y	Observed frequency	Expected frequency							
		Poisson	DL	DOL-R	DOL-E	DOL-W	DGOL-R	DGOL-E	DGOL-W
0	46	57.76	72.18	46.86	53.46	48.23	54.58	51.91	46.70
1	76	57.39	43.89	71.85	59.30	68.22	54.15	59.44	73.27
2	24	28.51	22.12	31.11	35.15	31.62	38.90	36.73	28.06
3	9	9.44	10.19	5.69	7.72	7.05	8.23	7.62	6.65
4	1	2.35	4.45	0.47	0.36	0.82	0.13	0.29	1.15
MLEs	$\hat{\mu}$	0.9936	–	–	–	–	–	–	–
	$\hat{\alpha}$	–	–	–	–	–	0.0005	0.0070	55.3489
	$\hat{\beta}$	–	–	358.0744	1.0320	0.0044	0.0016	0.0072	5.5369
	$\hat{\lambda}$	–	1.0318	0.0010	0.5319	5.5417	0.0777	0.5437	3.0603
	$\hat{\theta}$	–	–	–	–	0.2099	–	–	0.2562
	$-2\hat{L}$	191.9362	206.8300	188.4519	191.6568	188.2743	195.5790	192.0915	187.3904
	AIC	385.8724	415.6600	380.9038	387.3136	382.5486	397.1580	390.1830	382.7808
	BIC	388.9223	418.7099	387.0035	393.4133	391.6982	406.3076	399.3326	394.9802
	K-S	0.0754	0.1678	0.0245	0.0592	0.0355	0.0851	0.0682	0.0130

respectively (200 observations). Since the mean and variance of Data I are equal, we can use a Poisson distribution to fit the data.

The second data set (Data II) is the number of strikes (Y) in the UK coal mining industries (156 observations) in four successive week periods during 1948-1959 (see Ridout and Besbeas (2004) [12]). The mean, variance, and ID values of Y are 0.9936, 0.7419 and 0.7467, respectively. Since the ID value of Data II is less than 1, Y should be modeled by an under-dispersed distribution for count data.

Table 4. Comparison of the observed and expected values of Data III, and criteria values of each distribution

Z	Observed frequency	Expected frequency							
		Poisson	DL	DOL-R	DOL-E	DOL-W	DGOL-R	DGOL-E	DGOL-W
0	3541	3277.13	3356.31	594.52	1399.98	3577.61	783.07	3577.61	3545.42
1	599	970.03	839.14	1431.34	1518.06	513.11	849.95	513.11	582.61
2	176	143.56	171.73	1430.12	1057.22	200.73	1099.96	200.73	177.17
3	48	14.17	32.01	762.23	378.60	76.79	1050.61	76.79	61.78
4	20	1.05	5.65	176.60	50.58	26.78	537.23	26.78	23.23
5	12	0.06	0.96	11.11	1.56	8.23	83.82	8.23	9.19
6	5	0.00	0.16	0.08	0.01	2.17	1.37	2.17	3.77
7	1	0.00	0.03	0.00	0.00	0.48	0.00	0.48	1.59
8	4	0.00	0.00	0.00	0.00	0.09	0.00	0.09	0.69
MLEs	$\hat{\mu}$	0.2960	–	–	–	–	–	–	–
	$\hat{\alpha}$	–	–	–	–	0.0571	0.0607	0.0069	19.6439
	$\hat{\beta}$	–	–	5.0002	1.1385	3.9702	0.1340	0.0468	0.2203
	$\hat{\lambda}$	–	1.9413	0.0341	0.4497	0.1335	0.0461	0.1149	5.5717
	$\hat{\theta}$		–	–	–	–	–	–	0.0991
	$-2\hat{L}$	3304.51	3117.35	8402.32	5489.91	3009.77	7711.52	3031.42	3009.10
	AIC	6611.02	6236.71	16808.64	10983.81	6025.54	15429.05	6068.85	6026.21
	BIC	6617.41	6243.10	16821.42	10996.60	6044.71	15448.22	6088.02	6051.77
	K-S	0.0599	0.0419	0.6687	0.4859	0.0035	0.6260	0.0112	0.0027

The third data set is the number of hospital stays (Z) for individuals age 66 and over (4406 observations), this data set was obtained from the National Medical Expenditure Survey in 1987 and 1988 (see Flynn et al. (2009) [13] and Deb and Trivedi (1997) [14]). The mean, variance, and ID values of Z are 0.2960, 0.5571 and 1.8824, respectively. Since the ID value of Data II is more than 1, Z should be modeled by an over-dispersed distribution for count data.

To estimate the parameters of each distribution, we compare the proposed DGOL-W model works in comparison to the other models, such as the Poisson distribution, discrete Lindley (DL) distribution (Gómez-Déniz, Emilio and Calderín-Ojeda, Enrique (2011) [12]) and the sub-models of the DGOL-W distributions by using the minimum values of the criterion such how as the AIC (Akaike information criterion) and the BIC (Bayesian information criterion) as indicators of the relative quality of the statistical models for the given set of data. Given a collection of models for the data, these criteria estimate the quality of each model, relative to each of the other models.

Suppose that we have a statistical model of some data. Let k be the number of estimated parameters in the model, and n be the sample size. Let $\hat{L}(\tilde{\omega})$ be the maximum value of the likelihood function for the models. Then the AIC and BIC values of the model are the following:

$$\text{AIC} = -2\hat{L}(\tilde{\omega}) + 2k \quad \text{and} \quad \text{BIC} = -2\hat{L}(\tilde{\omega}) + k\log n. \tag{4.1}$$

Testing of the Kolmogorov–Smirnov (K-S) test is used to compare fitting distributions, where the smaller values of these statistics give the best fit for the data. If the hypothesized distribution is $F_0(x)$, and the empirical (sample) cumulative distribution function is $F_n(x)$, where $F_n(x)$ is based on n independent and identically distributed ordered observations $X_{(i)}$, where $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ is defined as $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{(-\infty,x)}[X_{(i)}]$, where $I_{(-\infty,x)}[X_{(i)}]$ is the indicator function, which is equal

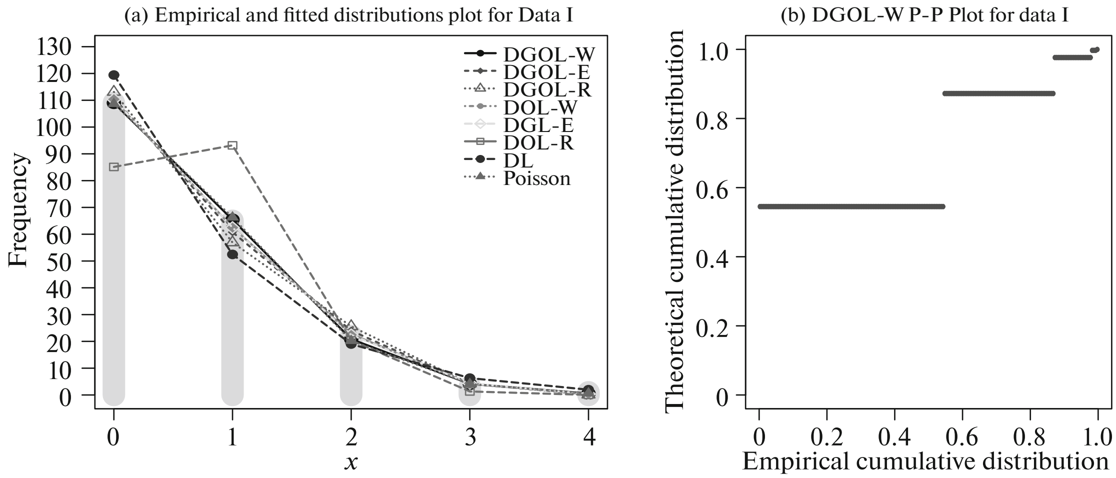

Fig. 3. The empirical and fitted distribution plots and the DGOL-W P-P plot for Data I.

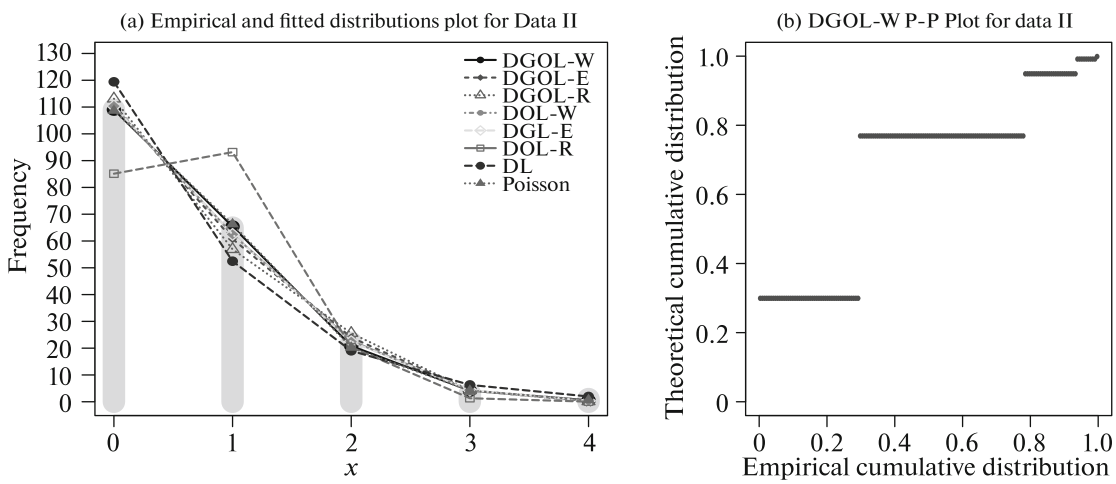

Fig. 4. The empirical and fitted distribution plots and the DGOL-W P-P plot for Data II.

to 1 if $X_{(i)} \leq x$ and equal to 0 otherwise. The values of the K-S test are

$$\text{K-S} = \max |F_n(x) - F_0(x)|. \tag{4.2}$$

From the results in Tables 2–4, for the goodness of fit K-S test the DGOL-W distribution gives a smaller value of the test statistics compared to the other distributions, i.e., Poisson, DL, DGOL-E, DGOL-R, DOL-W, DOL-E, and DOL-R. Note that the smaller K-S value gives better fit for the data. Thus, the DGOL-W distribution is the best model to fit these data sets (see Figures 3, 4, and 5).

5. DISCUSSION AND CONCLUSION

Modeling count data is one of the most important issues in statistical research. In this paper, we introduce a new probability mass function (pmf) by discretizing the continuous failure model of the generalized odd Lindley–Weibull distribution, which is called the discrete generalized odd Lindley–Weibull (DGOL-W) distribution. The pmf behaviors have various shapes, i.e., inverse J-shape, right-skewed shape, left-skewed shape, and close to symmetric shape. The proposed DGOL-W distribution has five sub-models: the discrete generalized odd Lindley-exponential (DGOL-E), discrete generalized odd Lindley–Rayleigh (DGOL-R), discrete odd Lindley–Weibull (DOL-W), discrete odd Lindley-exponential (DOL-E), and discrete odd Lindley–Rayleigh (DOL-R) distributions. Some properties of the proposed distribution are introduced. The maximum likelihood estimation (MLE) is used to estimate

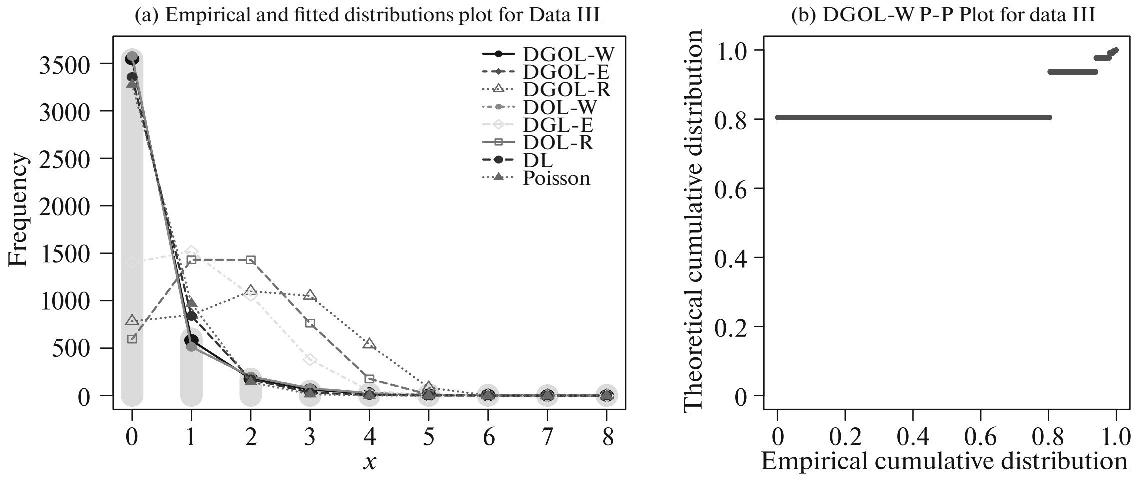

Fig. 5. The empirical and fitted distribution plots and the DGOL-W P-P plot for Data III.

the unknown parameters of the DGOL-E distribution. Applications are illustrated, which show that the model is suitable for application in various data sets, i.e., the mean and variance of the count data are equal, under-dispersion count data, and over-dispersion count data. Based on the results, we have shown that the DGOL-W distribution provides a better fit compared to the Poisson, discrete Lindley (DL) and the four sub-models for count data.

REFERENCES

1. Z. Ahmad, M. Elgarhy, and G. G. Hamedani, "A new Weibull-X family of distributions: Properties, characterizations and applications," J. Stat. Distrib. Appl. **5** (1), 5 (2018).
2. A. Z. Afify, G. M. Cordeiro, M. E. Maed, M. Alizadeh, H. Al-Moflen, and Z. M. Nofal, "The generalized odd Lindley-G family: Properties and applications," Anais Acad. Brasil. Ci. **91** (3) (2019).
3. S. Chakraborty, "Generating discrete analogues of continuous probability distributions—a survey of methods and constructions," J. Stat. Distrib. Appl. **2** (1), 6 (2015).
4. M. Ahsanullah, M. Shakil, and B.G. Kibria, "Characterizations of folded student's distribution," J. Stat. Distrib. Appl. **2** (1), 15 (2015).
5. D. Roy, "The discrete normal distribution," Commun. Stat., Theory Methods **32**, 1871–1883 (2003).
6. D. Roy, "Discrete Rayleigh distribution," IEEE Trans. Reliab. **53**, 255–260 (2004).
7. M. H. Alamatsaz, S. Dey, and S. S. Harandi, "Discrete generalized Rayleigh distribution," Pakistan J. Stat. **32**, 63–79 (2016).
8. K. Jayakumar and M. G. Babu, "Discrete additive Weibull geometric distribution," J. Stat. Theory Appl. **18**, 33–45 (2019).
9. S. Chakraborty and D. Chakravarty, "Discrete gamma distributions: Properties and parameter estimations," Commun. Stat., Theory Methods **41**, 3301–3324 (2012).
10. R Development Core Team, *R: A Language and Environment for Statistical Computing* (2018).
11. S. A. Klugman, H. H. Panjer, and G. E. Willmot, *Loss Models: From Data to Decisions,* 3rd ed., *Wiley Series in Probability and Statistics* (Wiley, Hoboken, NJ, 2012).
12. M. S. Ridout and P. Besbeas, "An empirical model for underdispersed count data," Stat. Model. **4**, 77–89 (2004).
13. M. Flynn and L. A. Francis, in *Proceedings of the Casualty Actuarial Society E-Forum, Winter, 2009*.
14. P. Deb and P. K. Trivedi, "Demand for medical care by the elderly: A finite mixture approach," J. Appl. Econometr. **12**, 313–336 (1997).
15. E. Gómez-Déniz and E. Calderín-Ojeda, "The discrete Lindley distribution: Properties and applications," J. Stat. Comput. Simul. **81**, 1405–1416 (2011).