**Pak. J. Statist.**
**2016 Vol. 32(1), 1-20**

# DISCRETE GENERALIZED RAYLEIGH DISTRIBUTION

**M.H. Alamatsaz**[1], **S. Dey**[2§], **T. Dey**[3] and **S. Shams Harandi**[4]

[1] Department of Statistics, University of Isfahan, Isfahan, Iran
Naghshejahan Institute of Higher Education, Baharestan, Isfahan, Iran. Email: alamatho@sci.ui.ac.ir

[2] Department of Statistics, St. Anthony's College, Shillong, Meghalaya, India. Email: sanku_dey2k2003@yahoo.co.in

[3] Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, Ohio, USA. Email: tanujit.dey@gmail.com

[4] Department of Statistics, Qom University, Qom, Iran
Email: salimeh.shams2013@yahoo.com

[§] Corresponding author

## ABSTRACT

In this paper, a new two-parameter discrete distribution is introduced which is in fact an analogue of the continuous generalized Rayleigh distribution, called discrete generalized Rayleigh distribution. The one parameter discrete Rayleigh distribution is obtained as a particular case. Several important distributional and reliability properties such as survival and hazard rate functions, moments, order statistics, unimodality and infinite divisibility of the proposed distribution are examined. A related characterization of the distribution is also provided. Parameter estimation, using different methods, namely methods of moments, maximum likelihood and proportion, is discussed. Performance of the different estimation methods are compared by means of a Monte Carlo simulation. Finally, two real data sets are analyzed to investigate the suitability of the proposed distribution in modeling count data.

## KEYWORDS

Maximum likelihood estimation, Method of moments, Method of Proportion, Observed Fisher information matrix, Unimodality, Infinitely divisibility.

## 1.INTRODUCTION

In scientific research, we frequently come across variables that are discrete in nature. In life testing or reliability experiments, it is often difficult to quantify the life length of a device on a continuous scale. For example, in measuring reliability of an on/off-switching device, the lifetime of the switch is a discrete random variable. Similarly, in survival analysis, one may be interested in recording the number of days that a patient has survived since therapy, or the number of days taken from remission to relapse. In all such cases the lifetimes are not measured on continuous scale but are simply counted and hence are discrete random variables.

 

Not many of the known discrete distributions can provide accurate models for both times and counts. For example, Poisson distribution is used to model counts but not times. Binomial and negative binomial distributions are not considered to be popular models for reliability, failure times, counts, etc. This is partly because they are not defined over the set of all non-negative integers. Besides, binomial and negative binomial distributions can be approximated well by Poisson distribution under suitable conditions. Further, the applicability of conventional discrete distributions like geometric, Poisson, etc. has limited use as models for reliability, failure times, counts, etc. This has led to the development of some new discrete distributions based on popular continuous models for reliability, failure times, etc. Among these new distributions, the discrete Weibull distribution is the most popular. Besides Weibull, another newly developed distribution is the discrete gamma distribution which has received significant attention in applications. It was first used by Yang (1994, Equations (8)-(12)) in the area of molecular biology and evolution. Since then, many researchers in molecular biology and evolution have used the discrete gamma distribution.

More recently, Kemp (2008) studied the discrete half-normal distribution. Krishna and Pundir (2009) constructed discrete analogues of the continuous Burr and Pareto distributions. Aghababaei Jazi et al. (2010) developed a discrete analogue of the continuous inverse Weibull distribution. Nekoukhou et al. (2013) introduced the discrete generalized exponential distribution. Furthermore, Gómez-Déniz and Calderin-Ojeda (2011), Al-Huniti and Al-Dayian (2012), Hossein and Ahmad (2014), Anwar and Ahmad (2014) are related references.

For any continuous distribution on $\Re^+ = [0, +\infty)$ with probability density function (pdf) $f$, one can construct a discrete counterpart supported on the set of integers $N_0 = \{0, 1, 2, \ldots\}$, whose probability mass function (pmf) is of the form

$$p_x = P(X = x) = S(x) - S(x+1), \qquad x = 0, 1, 2, \ldots, \tag{1}$$

where S is the survival function of $f$.

The generalized Rayleigh (GR) distribution is one of the well-known lifetime distributions with pdf

$$f(x) = 2\alpha\lambda^2 x e^{-(\lambda x)^2} \left(1 - e^{-(\lambda x)^2}\right)^{\alpha-1}, \qquad x > 0 \tag{2}$$

and survival function

$$S(x) = 1 - \left(1 - e^{-(\lambda x)^2}\right)^{\alpha}, \qquad x \geq 0 \tag{3}$$

Here α and λ are the shape and scale parameters, respectively. The two-parameter GR distribution is denoted by GR(α, λ). It is observed by Raqab and Kundu (2003) that for $\alpha \leq \frac{1}{2}$ the pdf of a GR distribution is a decreasing function and it is a right skewed unimodal function for $\alpha > \frac{1}{2}$. It is also observed that the hazard function of a GR distribution can be either bathtub type or an increasing function, depending on the value of α.

In this paper, we shall propose a discrete version of the generalized Rayleigh distribution (DGR) using relation (1). Note that the two-parameter generalized Rayleigh distribution is a distinct member of the class of exponentiated Weibull distributions, originally proposed by Mudholkar and Srivastava (1993) and Mudholkar et al. (1995).

The rest of the paper is organized as follows. Section 2 introduces the DGR distribution and discusses several of its characteristics. We shall settle two important structural properties of the distribution, i.e., its unimodality and infinite divisibility. Several mathematical properties of the distribution, specifically quantile function, hazard rate function, order statistics, moments and association to other distributions are investigated. Using the method of moments, method of maximum likelihood and the method of proportion, estimator of parameters of the distribution are derived in Section 3. In Section 4, performance of the different estimation methods are compared via a simulation study. Two real life data sets are also analyzed showing applicability of the new distribution in this section. The paper ends with a brief conclusion in Section 5.

## 2. DISCRETE GENERALIZED RAYLEIGH DISTRIBUTION

The DGR distribution can be defined as a non-negative integer valued distribution with pmf

$$p_x = P(X = x) = \begin{cases} \left(1 - p^{(x+1)^2}\right)^{\alpha} - \left(1 - p^{x^2}\right)^{\alpha} & x = 0,1,2,\ldots \\ 0 & ow. \end{cases}$$
$$= \begin{cases} \sum_{j=1}^{\infty} \binom{\alpha}{j} (-1)^{j+1} p^{x^2 j} \left(1 - p^{j(2x+1)}\right) & x = 0,1,2,\ldots \\ 0 & ow. \end{cases} \qquad (4)$$

where $0 < p = e^{-\lambda^2} < 1, \ \ \alpha > 0$ . Note that when α is an integer number, the summation in (4) stops at α (see, e.g., Nekoukhou et al. (2013)).

If α = 1, we have the discrete Rayleigh distribution as introduced by Roy (2004). Let $[X]$ be the greatest integer value of $X$. If $X$ is a random variable from generalized Rayleigh distribution (2), then $[X]$ is distributed as (4). The cumulative distribution function (cdf) of a random variable $X$ following a DGR(α, p) distribution is given by

$$F(x) = \left(1 - p^{(x+1)^2}\right)^{\alpha}, \qquad x \in N_0. \qquad (5)$$

Furthermore, the quantile function of the DGR(α, p) distribution, say Q(u), is obtained from

$$\left(1 - p^{(Q(u)+1)^2}\right) = u,$$

where $0 < u < 1$. By solving the equation, we have

$$Q(u) = \left[ 1 - \left( \frac{\log\left(1 - u^{1/\alpha}\right)}{\log(p)} \right)^{1/2} \right].$$

Figure 1 illustrates the behavior of the pmf of DGR(α, p) for several values of p and α. The unimodality property of the DGR(α, p) distribution is consistent with that of the continuous generalized Rayleigh distribution (Raqab and Kundu (2006)).

**Proposition 2.1**

DGR(α, p) distribution is unimodal for all values of α and p. In particular, the mass probability function is non-increasing when $\alpha \leq \frac{1}{2}$ .

Another important structural property of a distribution is its infinite divisibility. We refer to the monograph of Steutel and Van Harn (2004) for a good and complete introduction of the issue at hand. We first recall the following interesting result from the above mentioned monograph (page 56).

**Lemma 2.1**

If $p_k$ , $k \in Z_+$, is infinitely divisible, then we have $p_k \leq e^{-1}$, for all $k \in N$ .

**Proposition 2.2**

A DGR(α, p) distribution is not infinitely divisible in general.

**Proof:**

By Lemma 2.1, we show that $p_k > 1/e$ for some values of $k \in N$ , α and $p$ . For this, we take $\alpha = 1$, $p = 0.8$ and $k = 1$. Then, we see that $p_1 = 0.3904 > e^{-1} = 0.3679$. As seen in Figure 2, in fact DGR(α, p) distribution is not infinitely divisible for α = 1 and all $0.4 < p < 0.8$.

## 2.1 Hazard Rate Function

The survival and hazard rate functions of the DGR(α, p) distribution are given by

$$S(x;\alpha, p) = P(X > x) = 1 - \left(1 - p^{([x]+1)^2}\right)^{\alpha}, \qquad x \geq 0 \tag{6}$$

and

$$h(x;\alpha, p) = \frac{\left(1 - p^{(x+1)^2}\right)^{\alpha} - \left(1 - p^{x^2}\right)^{\alpha}}{1 - \left(1 - p^{(x+1)^2}\right)^{\alpha}}, \qquad x \in N_0, \tag{7}$$

respectively.

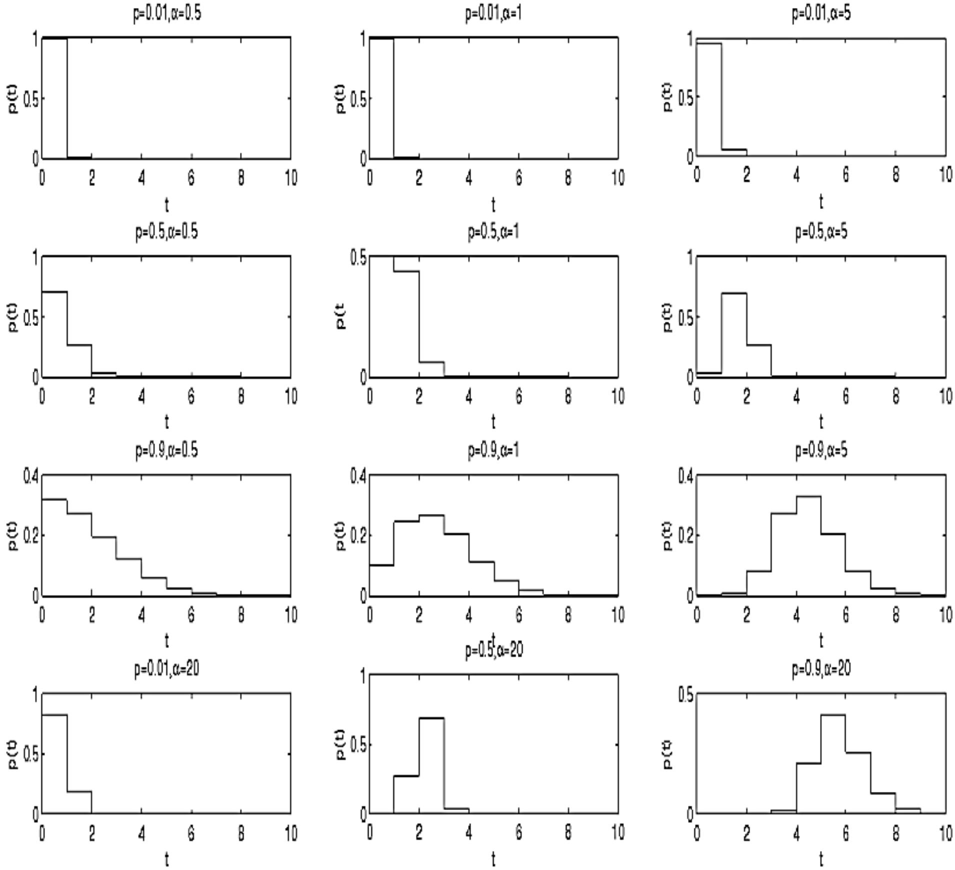


**Figure 1: Illustration of the Probability Mass Function of DGR(α, p) for Different Values of p and α.**

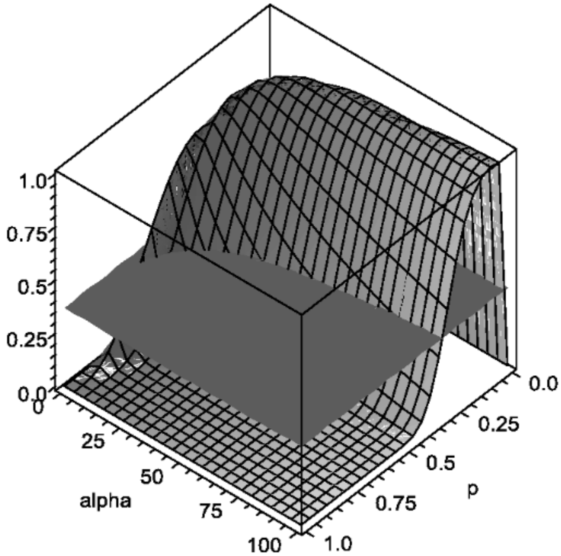


**Figure 2: Illustration of the Probability Mass Function $p_1$ of DGR(α, p) as a Function of p and α against the Uniform Plane $e^{-1}$.**

We also observe that the reversed hazard rate function of this distribution is a non-decreasing function, because

$$\bar{h}_X(x) = \frac{f(x)}{F(x)} = \frac{\left(1-p^{(x+1)^2}\right)-\left(1-p^{x^2}\right)^{\alpha}}{\left(1-p^{(x+1)^2}\right)^{\alpha}} = 1 - \frac{\left(1-p^{x^2}\right)^{\alpha}}{\left(1-p^{(x+1)^2}\right)^{\alpha}}, \qquad x \in N_0$$

and $\dfrac{\left(1-p^{x^2}\right)^{\alpha}}{\left(1-p^{(x+1)^2}\right)^{\alpha}}$ is a non-increasing function of x, we have the required result.

## 2.2 Order Statistics of the DGR distribution

Let $F_i\left(x;\ \alpha,\ p\right)$ be the cumulative distribution function of the ith order statistic for a random sample $X_1, X_2, ..., X_n$ from X ~ DGR(α, p). Since

$$F_i(x;\alpha,p) = \sum_{k=i}^{n} \binom{n}{k} \left\{F(x;p,\alpha)\right\}^k \left\{1-F(x;p,\alpha)\right\}^{n-k},$$

using the binomial expansion for $\left\{1-F(x;p,\alpha)\right\}^{n-k}$, we obtain

$$F_i(x;\alpha,p) = \sum_{k=i}^{n}\sum_{j=0}^{n-k}\binom{n}{k}\binom{n-k}{j}(-1)^j \left\{F(x;p,\alpha)\right\}^{k+j},$$

$$= \sum_{k=i}^{n}\sum_{j=0}^{n-k}\binom{n}{k}\binom{n-k}{j}(-1)^j \left\{1-p^{([x]+1)^2}\right\}^{\alpha(k+j)},$$

$$= \sum_{k=i}^{n}\sum_{j=0}^{n-k}\binom{n}{k}\binom{n-k}{j}(-1)^j F\left(x;p,\alpha(k+j)\right).$$

Note that $F_i\left(x;\alpha,p\right)$ is a linear combination of a finite number of DGR$\left(\alpha(k+j),p\right)$ distributions. This helps one to obtain certain interesting properties of the order statistics, such as the moment generating function and the moments, from those of the corresponding DGR distribution.

Suppose that $X_1, X_2, ..., X_n$ is a random sample from DGR(α, p) distribution. Let $X_{1:n}, X_{2:n}, ..., X_{n:n}$ denote the corresponding order statistics. Then, the pmf of $X_{i:n}$, is given by

$$P\left(X_{i:n}=x\right) = \frac{n!}{(i-1)!(n-i)!}\int_{F(x-1)}^{F(x)} u^{i-1}(1-u)^{n-i}\,du,$$

$$= \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}\binom{j}{n-i}(-1)^{j/(i+j)}\int_{F(x-1)}^{F(x)} u^{i+j-1}du,$$

$$= \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}\binom{j}{n-i}(-1)^{j/(i+j)}\left[\left(1-p^{(x+1)^2}\right)^{\alpha(i+j)} - \left(1-p^{x^2}\right)^{\alpha(i+j)}\right],$$

$$= \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}\binom{j}{n-i}(-1)^{j}\left[F\left(x;p,\alpha(i+j)\right)-F\left(x-1;p,\alpha(i+j)\right)\right].$$

**Proposition 2.3**

Let $X_i$'s be independent random variables from DGR$\left(\alpha_i,\, p\right)$ distribution, $i=1,2,...,n,$ respectively. Then, $V=\max\left(X_1, X_2, ..., X_n\right)$ follows a $DGR\left(\sum_{i=1}^{n}\alpha_i, p\right)$ distribution.

**Proof:**

We simply have

$$F_V(v) = P(V \le v) = \prod_{i=1}^{n}\left(1-p^{\left([v]+1\right)^2}\right)^{\alpha_i} = \left(1-p^{\left([v]+1\right)^2}\right)^{\sum_{i=1}^{n}\alpha_i}$$

as required.

**Proposition 2.4**

Let α be an integer value. If $X_i$'s, $i = 1, 2, ..., \alpha$, are independent and identical random variables from a geometric distribution with parameter p, then $\left[\max_{i=1}^{\alpha}\left(\sqrt{X_i}\right)\right]$ has DGR(α, p) distribution.

**Proof:**

$$P\left(\left[\max_{i=1}^{\alpha}\left(\sqrt{X_i}\right)\right] = z\right) = P\left(z \le \max_{i=1}^{\alpha}\left(\sqrt{X_i}\right) < z+1\right),$$

$$= P\left(z-1 < \max_{i=1}^{\alpha}\left(\sqrt{X_i}\right) \le z\right),$$

$$= P\left(\max_{i=1}^{\alpha}\left(\sqrt{X_i}\right) \le z\right) - P\left(\max_{i=1}^{\alpha}\left(\sqrt{X_i}\right) \le z-1\right),$$

$$= \left[P\left(\sqrt{X_i} \le z\right)\right]^{\alpha} - \left[P\left(\sqrt{X_i} \le z-1\right)\right]^{\alpha}.$$

Since Roy (2004) proved that if X be a random variable from geometric distribution with parameter p, then [X] has discrete Rayleigh distribution, this clearly completes the proof.

**Proposition 2.5**

Let X be a continuous random variable and t be a positive constant.

Then, $X_t = \left[\frac{X}{t}\right]$ is DGR if, and only if, X has a generalized Rayleigh (GR) distribution.

Proof:

Let X ~ GR(α, λ) and S(x) be its survival function. We have

$$P\left(X_t = x\right) = P\left(x \le X / t < x+1\right) = S(xt) - S\left((x+1)t\right).$$

Accordingly,

$$\begin{aligned} S_t(x) = P\left(X_t > x\right) = S\left((x+1)t\right) &= 1 - \left(1 - e^{-\left(\lambda t(x+1)\right)^2}\right)^{\alpha} \\ &= 1 - \left(1 - p_t^{(x+1)^2}\right)^{\alpha}, \qquad t > 0 \end{aligned} \tag{8}$$

where $p_t = e^{-(\lambda t)^2}$. Thus, $X_t$ follows a $\text{DGR}\left(\alpha,\ p_t\right)$ distribution for every $t > 0$.

Conversely, let $X_t \sim \text{DGR}\left(\alpha,\ p_t\right)$ with survival function

$$S_t(x) = P\left(X_t > x\right) = 1 - \left(1 - p_t^{(x+1)^2}\right)^{\alpha}, \qquad x = 0, 1, 2, ...;$$

from (8) we have

$$S\left((x+1)t\right)=P\left(X\geq(x+1)t\right)=1-\left(1-p_t^{(x+1)^2}\right)^{\alpha},\qquad x=0,1,2,\ldots,t\geq 0,\qquad (9)$$

where $p_t=e^{-(\lambda t)^2}$ with $\lambda>0$. If $t=0$, $S\left((x+1)t\right)=S(0)=1$ and relation (8) is trivially true. Now, let F(t) be the distribution function of $Y=X^2$. We have

$$F(t)=1-S\left(\sqrt{t}\right),\ t\geq 0.$$

and, thus,

$$1-F\left((x+1)^2t^2\right)=1-\left(1-\left(1-F^{1/\alpha}\left(t^2\right)\right)^{(x+1)^2}\right)^{\alpha},\qquad t\geq 0.$$

As a result, we obtain

$$1-F^{1/\alpha}\left((x+1)^2t^2\right)=\left(1-F^{1/\alpha}\left(t^2\right)\right)^{(x+1)^2},\qquad t\geq 0.$$

Now, let $\beta=1/\alpha$ and $\mathrm{G(y)}=\mathrm{F}^{\beta}(y)$. Then,

$$1-G\left((x+1)^2t^2\right)=\left(1-G\left(t^2\right)\right)^{(x+1)^2},\qquad t\geq 0.$$

Hence, using Result 4 of Roy (2004), the distribution function G is an exponential distribution function, and consequently $Y\sim F=G^{\alpha}$ is a generalized exponential random variable. Therefore, we have

$$P(X\leq x)=P\left(\sqrt{Y}\leq x\right)=P\left(Y\leq x^2\right)=\left(1-e^{-\lambda x^2}\right)^{\alpha},$$

which is the cdf of the GR distribution.

**2.3 Moments**

We have the first and the second moments of the distribution (3) as

$$E(X)=\sum_{x=0}^{\infty}\left(1-\left(1-p^{x^2}\right)^{\alpha}\right)$$

and

$$E\left(X^2\right)=2\sum_{x=0}^{\infty}x\left(1-\left(1-p^{x^2}\right)^{\alpha}\right)+E(X).$$

The above expressions are infinite series and cannot be written in closed forms. From these equations and a result of Jazi et al. (2010), we have

$$\sum_{x=1}^{\infty}\left(1-\left(1-p^{x^2}\right)^{\alpha}\right)<\int_0^{\infty}\left(1-\left(1-p^{x^2}\right)^{\alpha}\right)dx<\sum_{x=0}^{\infty}\left(1-\left(1-p^{x^2}\right)^{\alpha}\right).$$

Thus, the mean of the continuous generalized Rayleigh distribution satisfies $\mu_d-1<\mu_c<\mu_d$, where $\mu_d$ and $\mu_c$ are the means of discrete and continuous Rayleigh distributions, re-spectively.

### 2.3.1 Index of Dispersion

The index of dispersion (ID) is defined as variance divided by the mean of a distribution.

If ID value is greater than one, the corresponding distribution is overdispersed, and if it is less than one, the distribution is underdispersed. Figure 3 represents the ID plot for the DGR distribution for different values of p and α. It is observed that for the values of α greater than 1, the distribution is always overdispersed but for the values of α less than 1, the distribution may be over or underdispersed. The ID seems to increase with the increase of α for fixed p.

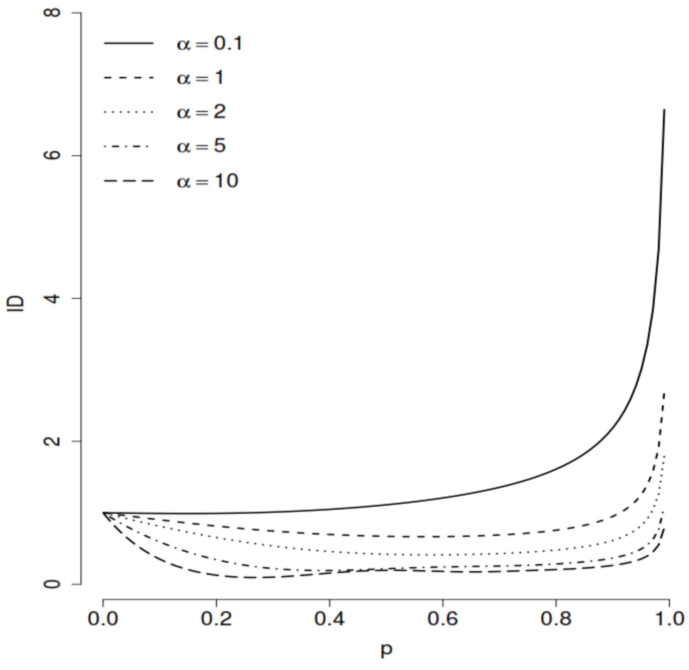


**Figure 3: Index of Dispersion Plot of DGR Distribution for Different Values of p and α.**

## 3. PARAMETER ESTIMATION

### 3.1 Method of Proportions

Khan et al. (1989) proposed a method of proportions to estimate the parameters of the discrete Weibull distribution, that Jazi et al. (2010) implemented to estimate the parameters of the discrete inverse Weibull distribution. Here, we use the same technique to estimate the parameters of discrete generalized Rayleigh distribution. Let $x_1, x_2, ..., x_n$ be $n$ observations from distribution (4). Define

$$I_1(x_i) = \begin{cases} 1 & x_i = 0 \\ 0 & ow. \end{cases}$$

Then, $Z = \sum_{i=1}^{n} I_1(x_i)$ denotes the number of 0′s in the sample. It is likely that $p_0 = (1-p)^{\alpha}$ can be estimated with $Z/n$ . But, since we have two parameters, we define another indicator function as follows:

$$I_2(x_i) = \begin{cases} 1 & x_i = 1 \\ 0 & ow. \end{cases}$$

Akin to the first case, it is apparent that the estimator for $p_1 = \left(1-p^4\right)^{\alpha} - (1-p)^{\alpha}$ can be obtained by equating it to $W/n = \sum_{i=1}^{n} I_2(x_i)/n$ . Subsequently, we have simultaneous equations

$$\begin{cases} (1-p)^{\alpha} = Z/n \\ \left(1-p^4\right)^{\alpha} = (W+Z)/n. \end{cases}$$

Thus, $\hat{p} = 1-(Z/n)^{1/\hat{\alpha}}$ and $\hat{\alpha} = \log\left(\{W+Z\}/n\right) / \log\left(1-\hat{p}^4\right)$. Since $Z/n$ and $(W+Z)/n$ are unbiased and consistent empirical estimators of probabilities $P(X \le 0)$ and $P(X \le 1)$, the above estimators of the parameters are also consistent.

### 3.2 Method of Moments

To apply the method of moments for estimating the parameters p and α of DGR distribution, we need to equate the population moments to the corresponding sample moments and subsequently solve the two equations simultaneously. Since the moments of the discrete generalized Rayleigh distribution cannot be obtained in closed forms, the equations can not be solved via ordinary techniques. So, we resort to a method of pseudo-moment by minimizing

$$S(\alpha, p) = \left(M_1 - E(X)\right)^2 + \left(M_2 - E\left(X^2\right)\right)^2$$

with respect to p and α, where $M_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $M_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ are the first and second sample moments.

### 3.3 Maximum Likelihood Estimation

One popular method of estimating parameters of distributions is the method of maximum likelihood. To apply this method for estimating p and α, assume that $x = \left(x_1,\ x_2,\ ...,\ x_n\right)^T$ is a random sample of size n from a DGR(α, p) distribution. The log-likelihood function becomes

$$\ell = \sum_{i=1}^{n} \log\left[\left(1 - p^{\left(x_i+1\right)^2}\right)^{\alpha} - \left(1 - p^{x_i^2}\right)^{\alpha}\right]. \tag{10}$$

Hence, the normal equations are

$$\frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^{n} \frac{\left(1 - p^{\left(x_i+1\right)^2}\right)^{\alpha} \log\left(1 - p^{\left(x_i+1\right)^2}\right) - \left(1 - p^{x_i^2}\right)^{\alpha} \log\left(1 - p^{x_i^2}\right)^{\alpha}}{\left(1 - p^{\left(x_i+1\right)^2}\right)^{\alpha} - \left(1 - p^{x_i^2}\right)^{\alpha}} = 0 \tag{11}$$

and

$$\frac{\partial \ell}{\partial p} = \alpha \sum_{i=1}^{n} \frac{\left(1 - p^{x_i^2}\right)^{\alpha-1} p^{x_i^2-1} x_i - \left(1 - p^{\left(x_i+1\right)^2}\right)^{\alpha-1} p^{\left(x_i+1\right)^2-1} \left(x_i + 1\right)^2}{\left(1 - p^{\left(x_i+1\right)^2}\right)^{\alpha} - \left(1 - p^{x_i^2}\right)^{\alpha}} = 0. \tag{12}$$

The solutions of likelihood equations (11) and (12) present the maximum likelihood estimators (MLEs) of α and p, which can be obtained via the numerical method of two dimensional Newton-Raphson type procedure.

Since the MLE of the vector of unknown parameters $\theta = (\alpha, p)^T$ cannot be derived in closed forms, it is therefore hard to derive the exact distribution of the MLEs. Hence, we can not find the exact bounds for the parameters. However, using large sample approximation, we see that the DGR family satisfies the regularity conditions of the parameters in the interior of the parameter space but not on the boundary (see, e.g., Ferguson, 1996, pp. 121). It is known that the asymptotic distribution of the MLE $\hat{\theta}$ is

$$\left(\hat{\theta} - \theta\right) \to N\left(0, I^{-1}(\theta)\right), \tag{13}$$

(see Lawless (1982) ), where $I^{-1}(\theta)$ is the inverse of the Fisher's information matrix of the unknown parameters $\theta = (\alpha, p)^T$ as follows:

$$I_X(\theta) = \begin{bmatrix} -E\left(\frac{\partial^2 \ell}{\partial \alpha^2}\right) & -E\left(\frac{\partial^2 \ell}{\partial \alpha \partial p}\right) \\ -E\left(\frac{\partial^2 \ell}{\partial \alpha \partial p}\right) & -E\left(\frac{\partial^2 \ell}{\partial p^2}\right) \end{bmatrix}. \tag{14}$$

On the other hand, the Fisher's information matrix can be computed using the approximation

$$I_X\left(\hat{\theta}\right) \approx \begin{bmatrix} -\dfrac{\partial^2 \ell}{\partial \alpha^2}\Big|_{(\hat{\alpha},\hat{p})} & -\dfrac{\partial^2 \ell}{\partial \alpha \partial p}\Big|_{(\hat{\alpha},\hat{p})} \\ -\dfrac{\partial^2 \ell}{\partial \alpha \partial p}\Big|_{(\hat{\alpha},\hat{p})} & -\dfrac{\partial^2 \ell}{\partial p^2}\Big|_{(\hat{\alpha},\hat{p})} \end{bmatrix}, \tag{15}$$

where α and p are the MLEs of α and p, respectively.

The second partial derivatives of the log-likelihood function (10) are given below:

$$\frac{\partial^2 \ell}{\partial \alpha^2} = \sum_{i=1}^{n} \left\{ \frac{\left(1-p^{(x_i+1)^2}\right)^{\alpha} \left[\log\left(1-p^{(x_i+1)^2}\right)\right]^2 - \left(1-p^{x_i^2}\right)^{\alpha} \left[\log\left(1-p^{x_i^2}\right)\right]^2}{\left(1-p^{(x_i+1)^2}\right)^{\alpha} - \left(1-p^{x_i^2}\right)^{\alpha}} \right.$$

$$\left. - \frac{\left[\left(1-p^{(x_i+1)^2}\right)^{\alpha} \log\left(1-p^{(x_i+1)^2}\right) - \left(1-p^{x_i^2}\right)^{\alpha} \log\left(1-p^{x_i^2}\right)\right]^2}{\left[\left(1-p^{(x_i+1)^2}\right)^{\alpha} - \left(1-p^{x_i^2}\right)^{\alpha}\right]^2} \right\},$$

$$\frac{\partial^2 \ell}{\partial p^2} = \alpha \left[ \sum_{i=1}^{n} \left\{ \frac{x_i^2 \left(x_i^2-1\right)\left(1-p^{x_i^2}\right)^{\alpha-1} p^{x_i^2-2} - (\alpha-1) x_i^4 \left(1-p^{x_i^2}\right)^{\alpha-2} p^{2x_i^2-2}}{\left(1-p^{(x_i+1)^2}\right)^{\alpha} - (1-p^{x_i^2})^{\alpha}} \right. \right.$$

$$- \frac{(x_i+1)^2((x_i+1)^2-1)\left(1-p^{(x_i+1)^2}\right)^{\alpha-1} p^{(x_i+1)^2-1}}{\left(1-p^{(x_i+1)^2}\right)^{\alpha} - (1-p^{x_i^2})^{\alpha}}$$

$$+ \frac{(\alpha-1)(x_i+1)^4 p^{2(x_i+1)^2-2} \left(1-p^{(x_i+1)^2}\right)^{\alpha-2}}{\left(1-p^{(x_i+1)^2}\right)^{\alpha} - \left(1-p^{x_i^2}\right)^{\alpha}}$$

$$\left. \left. - \frac{\left[x_i^2 p^{x_i^2-1}\left(1-p^{x_i^2}\right)^{\alpha-1} - (x_i+1)^2 p^{(x_i+1)^2-1}\left(1-p^{(x_i+1)^2}\right)^{\alpha-1}\right]^2}{\left[\left(1-p^{(x_i+1)^2}\right)^{\alpha} - \left(1-p^{x_i^2}\right)^{\alpha}\right]^2} \right\} \right]$$

and

$$\frac{\partial^2 \ell}{\partial\alpha\partial p} = \sum_{i=1}^{n}\left\{\frac{y_i^2\, p^{x_i^2-1}\left(1-p^{x_i^2}\right)^{\alpha-1}\left[\alpha\log\left(1-p^{x_i^2}\right)+1\right]}{\left(1-p^{(x_i+1)^2}\right)^{\alpha}-\left(1-p^{x_i^2}\right)^{\alpha}}\right.$$

$$-\frac{(x_i+1)^2\left(1-p^{(x_i+1)^2}\right)^{\alpha-1}\, p^{(x_i+1)^2-1}\left[\alpha\log\left(1-p^{(x_i+1)^2}\right)+1\right]}{\left(1-p^{(x_i+1)^2}\right)^{\alpha}-\left(1-p^{x_i^2}\right)^{\alpha}}$$

$$-\alpha\frac{x_i^2\, p^{x_i^2}\left(1-p^{x_i^2}\right)^{\alpha-1}\left(1-p^{(x_i+1)^2}\right)^{\alpha}\log\left(1-p^{(x_i+1)^2}\right)}{\left[\left(1-p^{(x_i+1)^2}\right)^{\alpha}-\left(1-p^{x_i^2}\right)^{\alpha}\right]^2}$$

$$+\alpha\frac{\left(x_i+1\right)^2\, p^{(x_i+1)^2-1}\left(1-p^{(x_i+1)^2}\right)^{2\alpha-1}\log\left(1-p^{(x_i+1)^2}\right)}{\left[\left(1-p^{(x_i+1)^2}\right)^{\alpha}-\left(1-p^{x_i^2}\right)^{\alpha}\right]^2}$$

$$+\alpha\frac{\left(x_i+1\right)^2\, p^{(x_i+1)^2-1}\left(1-p^{(x_i+1)^2}\right)^{\alpha}\left(1-p^{x_i^2}\right)^{\alpha-1}\log\left(1-p^{x_i^2}\right)}{\left[\left(1-p^{x_i+1}\right)^{\alpha}-\left(1-p^{x_i}\right)^{\alpha}\right]^2}$$

$$\left.-\alpha\frac{x_i^2\, p^{x_i^2-1}\left(1-p^{x_i^2}\right)^{2\alpha-1}\log\left(1-p^{x_i^2}\right)}{\left[\left(1-p^{x_i+1}\right)^{\alpha}-\left(1-p^{x_i}\right)^{\alpha}\right]^2}\right\}.$$

## 4. NUMERICAL EXPERIMENTS

### 4.1 Simulation Study

We perform a simulation study to compare different methods of estimation for the parameters of DGR distribution. We have considered different sample sizes; $n = 20$, 50, 80, different values of p=0.5, 0.8 and different values of α=1, 2. The performance of the estimators are compared based on the bias and the mean squared error (MSE) of the estimators using three different methods. In Table 1, M OM stands for the method of moments technique, M OP represents the method of proportions technique, and finally M LE corresponds to the method of maximum likelihood technique. All results are based on averaging over 10,000 replications.

From Table 1, we see that the performance of the MLE method is better than the other two methods. Under the MLE method, the estimator of p is slightly negative biased and so is the case under the MOM method. The performance of the MOP method is inferior with respect to the other methods; one reason being that it uses only the information of 0's and 1's from the samples and discards all other information. This might be the reason behind increased biases and MSEs. It is also observed that given a fixed value of α, as p increases, the precision of the estimates also increases.

**4.2 An Example**

We are using two real life count data sets (1) to illustrate several estimation procedures proposed in this article; and (2) to investigate how well the proposed DGR model works in comparison to other existing distributions. The first data set in Table 2 represents the number of European red mites on apple leaves (Chakraborty (2010)). Chakraborty and Chakravarty (2012), deem this an over-dispersed data set. The second data set in Table 3 represents the number of outbreaks of strikes in UK coal mining industries in four successive week periods during 1948-59 (Ridout and Besbeas (2004)). Chakraborty and Chakravarty (2012) mentioned that this is an underdispersed data set. Our goal is to illustrate how DGR can be useful in modeling two unique and distinct data sets.

In evaluating DGR with other distributions, we took into consideration Discrete Burr (Krishna and Pundir (2009)), Discrete generalized exponential (Nekoukhoua et al (2013)) and Discrete Weibull (Nakagawa and Osaki (1975), Khan et al (1989)) distributions. The comparison has been performed by using $\chi^2$ goodness-of-fit test along with two information criteria: Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). These criteria are well known for implementing model selection. In the set of competing models, a model is selected as the best model that has the smallest AIC and BIC values.

For the first data set, the MLEs of the parameters of DGR distribution are given in Table 2. The MOM estimates are α = 0.2807 and p = 0.9238, while the MOP estimates are α = 1.0323 and p = 0.5221. The variance-covariance matrix of the MLEs is given by

$$\begin{pmatrix} 0.001684215 & -0.000334180 \\ -0.000334180 & 0.0001869777 \end{pmatrix}.$$

**Table 1**
**The Average Bias and the Mean Squared Error (within Parenthesis) of the Estimates for Three Different Methods**

| $p$ | $\alpha$ | Sample Size | MLE | | MOM | | MOP | |
|---|---|---|---|---|---|---|---|---|
| | | | $\hat{p}$ | $\hat{\alpha}$ | $\hat{p}$ | $\hat{\alpha}$ | $\hat{p}$ | $\hat{\alpha}$ |
| 0.5, | 1 | 20 | -0.0285 | 0.2186 | -0.0809 | 0.3412 | 0.1257 | 0.4586 |
| | | | (0.0186) | (0.3879) | (0.0509) | (0.3956) | (0.0389) | (0.2490) |
| | | 50 | -0.0148 | 0.0848 | -0.0283 | 0.3288 | 0.1246 | 0.4351 |
| | | | (0.0068) | (0.1290) | (0.0133) | (0.3645) | (0.0260) | (0.2031) |
| | | 80 | -0.0027 | 0.0339 | -0.0103 | 0.0992 | 0.1215 | 0.4332 |
| | | | (0.0043) | (0.0630) | (0.0066) | (0.1968) | (0.0228) | (0.1954) |
| 0.8, | 1 | 20 | -0.0188 | 0.1501 | -0.0253 | 0.2428 | 0.1891 | 0.5124 |
| | | | (0.0047) | (0.2783) | (0.0057) | (0.4257) | (0.0312) | (0.2895) |
| | | 50 | -0.0082 | 0.0545 | -0.0101 | 0.0801 | 0.1691 | 0.4621 |
| | | | (0.0014) | (0.0490) | (0.0017) | (0.0726) | (0.0296) | (0.2645) |
| | | 80 | -0.0038 | 0.0272 | -0.0053 | 0.0438 | 0.1552 | 0.4421 |
| | | | (0.0008) | (0.0296) | (0.0004) | (0.0405) | (0.0212) | (0.2614) |
| 0.5, | 2 | 20 | -0.0151 | 0.2971 | -0.0465 | 0.2589 | 0.4412 | 0.1978 |
| | | | (0.0104) | (0.5489) | (0.0255) | (0.5812) | (0.1912) | (0.1206) |
| | | 50 | -0.0056 | 0.0821 | -0.0124 | 0.1969 | 0.4323 | 0.1896 |
| | | | (0.0038) | (0.3007) | (0.0054) | (0.5318) | (0.1896) | (0.1101) |
| | | 80 | -0.0043 | 0.0540 | -0.0098 | 0.1415 | 0.4213 | 0.1157 |
| | | | (0.0022) | (0.1432) | (0.0035) | (0.3088) | (0.1625) | (0.0664) |
| 0.8, | 2 | 20 | -0.0080 | 0.1790 | -0.0193 | 0.2314 | 0.1996 | 0.1856 |
| | | | (0.0021) | (0.7455) | (0.0037) | (0.3679) | (0.0399) | (0.1106) |
| | | 50 | -0.0016 | 0.0389 | -0.0048 | 0.0844 | 0.1826 | 0.1785 |
| | | | (0.0006) | (0.1559) | (0.0009) | (0.0756) | (0.0321) | (0.1028) |
| | | 80 | -0.0029 | 0.0729 | -0.0007 | 0.0354 | 0.1679 | 0.1054 |
| | | | (0.0003) | (0.0971) | (0.0005) | (0.0478) | (0.0256) | (0.0589) |

**Table 2**
**Distribution of Number of European Red Mites on Apple Leaves and Goodness of fit Tests**

| # European red mites | Frequency | DGR | DB | DGE | DW |
|---|---|---|---|---|---|
| 0 | 70 | 71.08538 | 70.45071 | 68.06997 | 63.225 |
| 1 | 38 | 32.08335 | 43.07854 | 38.87909 | 37.08962 |
| 2 | 17 | 20.76224 | 16.21801 | 20.60835 | 21.34886 |
| 3 | 10 | 12.87537 | 7.363343 | 10.78056 | 12.21733 |
| 4 | 9 | 7.251933 | 3.922487 | 5.610971 | 6.967468 |
| 5 | 3 | 3.601875 | 2.336814 | 2.91364 | 3.963802 |
| 6 | 2 | 1.544784 | 1.50786 | 1.511293 | 2.250772 |
| 7 | 1 | 0.5639898 | 1.032287 | 0.7834622 | 1.276109 |
| Total | 150 | 149.7689 | 145.91 | 149.1573 | 148.2705 |
| | | $\hat{\alpha} = 0.2939$ | $\hat{\theta} = 0.4005$ | $\hat{\alpha} = 1.0823$ | $\hat{\beta} = 1.0135$ |
| | | $\hat{p} = 0.9212$ | $\hat{\alpha} = 1.8836$ | $\hat{p} = 0.5181$ | $\hat{p} = 0.4215$ |
| | | $\log L = -221.2375$ | $\log L = -227.7271$ | $\log L = -222.5334$ | $\log L = -222.4289$ |
| | | $\chi^2 = 2.867601$ | $\chi^2 = 8.414688$ | $\chi^2 = 2.930197$ | $\chi^2 = 2.902079$ |
| | | $DF = 5$ | $DF = 5$ | $DF = 5$ | $DF = 5$ |
| | | $p$-value $= 0.7204$ | $p$-value $= 0.1348$ | $p$-value $= 0.7107$ | $p$-value $= 0.7150$ |
| | | $AIC = 446.4750$ | $AIC = 459.4542$ | $AIC = 449.0668$ | $AIC = 448.8578$ |
| | | $BIC = 452.4963$ | $BIC = 465.4755$ | $BIC = 455.0881$ | $BIC = 454.8791$ |

In the Table, DGR is abbreviated for DGR($\alpha$, p); DB is abbreviated for Discrete Burr($\theta$, $\alpha$); DGE is abbreviated for Discrete Generalized Exponential Type-II($\alpha$, p); and DW is abbreviated for Discrete Weibull($\beta$, p).

This provides a 95% confidence interval for $\alpha$ as (0.2135, 0.3744) and for p as (0.8944, 0.9480). As per the comparison with other distributions, Table 2 reveals that DGR gives the best fit compared to others with respect to $\chi^2$ goodness-of-fit test statistic as well as AIC and BIC. It is noteworthy mentioning that binomial, Poisson and negative binomial distributions are often used to model count data; Chakraborty and Chakravarty (2012) used discrete gamma distribution to model this data set and compared it with negative binomial and generalized Poisson distributions. They showed that discrete gamma performed better than the other two distributions. Comparing the result provided in Table 6 of Chakraborty and Chakravarty (2012), DGR performs even better than discrete gamma distribution.

**Table 3**
**Distribution of Number of Outbreak of Strikes and Goodness of fit Tests**

| # outbreak | Frequency | DGR | DB | DGE | DW |
|---|---|---|---|---|---|
| 0 | 46 | 47.50587 | 47.27603 | 43.74631 | 48.5004 |
| 1 | 76 | 69.30403 | 80.22819 | 76.69157 | 68.69698 |
| 2 | 24 | 31.80576 | 17.64411 | 26.92559 | 31.0997 |
| 3 | 9 | 6.664544 | 5.435168 | 6.6624 | 6.841383 |
| 4 | 1 | 0.6836565 | 2.260804 | 1.528991 | 0.8063582 |
| Total | 156 | 155.9639 | 152.8443 | 155.5549 | 155.9448 |
| | | $\hat{\alpha} = 0.9414$ | $\hat{\theta} = 0.5940$ | $\hat{\alpha} = 4.9983$ | $\hat{\beta} = 1.9017$ |
| | | $\hat{p} = 0.7172$ | $\hat{\alpha} = 4.6526$ | $\hat{p} = 0.2246$ | $\hat{p} = 0.3109$ |
| | | $\log L = -187.3296$ | $\log L = -192.2095$ | $\log L = -187.5343$ | $\log L = -188.1832$ |
| | | $\chi^2 = 3.567344$ | $\chi^2 = 3.853685$ | $\chi^2 = 0.8395505$ | $\chi^2 = 3.249548$ |
| | | $DF = 3$ | $DF = 3$ | $DF = 3$ | $DF = 3$ |
| | | $p$-value $= 0.3121$ | $p$-value $= 0.2777$ | $p$-value $= 0.8399$ | $p$-value $= 0.3547$ |
| | | $AIC = 378.6592$ | $AIC = 388.4190$ | $AIC = 379.0686$ | $AIC = 380.3664$ |
| | | $BIC = 384.7589$ | $BIC = 394.5187$ | $BIC = 385.1683$ | $BIC = 386.4661$ |

In the Table, DGR is abbreviated for DGR(α, p); DB is abbreviated for Discrete Burr(θ, α); DGE is abbreviated for Discrete Generalized Exponential Type-II(α, p); and DW is abbreviated for Discrete Weibull(β, p).

For the second data set, the MLEs are reported in Table 3. The MOM estimates are $\hat{\alpha} = 0.9083$ and $\hat{p} = 0.7227$, whereas the MOP estimates are $\hat{\alpha} = 1.5791$ and $\hat{p} = 0.8546$. The variance-covariance matrix of the MLEs is

$$\begin{pmatrix} 0.01872593 & -0.0031163403 \\ -0.0031163403 & 0.0009446754 \end{pmatrix}$$

This provides a 95% confidence interval for α as (0.6732, 1.2096) and for p as (0.6569, 0.7774). As per the comparison with other distributions, Table 3 reveals that if we choose $\chi^2$ goodness-of-fit test as the tool for comparison, the generalized exponential distribution gives better fit compared to others; and the DGR performance is quite competitive. If we choose AIC or BIC to be the tool of comparison, then DGR gives better fit compared to the others. Chakraborty and Chakravarty (2012) also used discrete gamma distribution to model this data set and compared it with binomial and Poisson distributions. They showed that discrete gamma performed better than the other two distributions. Comparing the result provided in Table 7 of Chakraborty and Chakravarty (2012), DGR performs better than discrete gamma distribution.

## 5. CONCLUSION

In this paper, we proposed a discretization version of the generalized Rayleigh distribution and discussed several important distributional and structural properties of the newly defined distribution. Three different methods have been conferred to estimate the parameters of the aforementioned distribution. The performance of these methods have been weighed against each other via Monte Carlo simulation study. In conclusion, a couple of real life data sets are used to illustrate the budding prospect and potential of this newly proposed distribution.

## ACKNOWLEDGMENT

The authors are thankful to the Editor and a reviewer whose comments led to improvement of the article.

## REFERENCES

1. Al-Huniti, A.A. and Al-Dayjan, G.R. (2012). Discrete Burr type III distribution. *Amer. J. Math. and Statist.*, 2(5), 145-152.
2. Anwar M. and Ahmad M. (2014). On some properties of geometric Poisson distribution. *Pak. J. Statist.*, 30(2), 233-244.
3. Chakraborty, S. (2010). On some distributional properties of the family of weighted generalized Poisson distribution. *Commun. Statist. Theor. Meth.*, 39, 2767-2788.
4. Chakraborty, S. and Chakravarty, D. (2012). Discrete Gamma Distributions: Properties and Parameter Estimations. *Commun. Statist. Theor. Meth.*, 41, 3301-3324.
5. Ferguson T.S. (1996). *A course in large sample theory*, London: Chapman and Hall.
6. Hussain T. and Ahmad M. (2014). Discrete inverse Rayleigh distribution. *Pak. J. Statist.*, 30(2), 203-222.
7. Gómez-Déniz, E. and Calderin-Ojeda E. (2011). The discrete Lindley distribution: properties and applications. Journal of Statistical and Simulation, 81(11), 1405-1416.
8. Jazi M.A., Lai C.D. and Alamatsaz M.H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. *Stat. Methodology*, 7, 121-132.
9. Kemp, A.W. (2008). The discrete half-normal distribution. *Int. Conf. Mathemat. Statist. Model. Honor of Enriqur Castillo.*, June, 28-30.
10. Khan, M.S.A., Khalique, A. and Abouammoh, A.M. (1989). On estimation parameters in a discrete Weibull distribution. *IEEE Trans. Reliability*, 38(3), 348-350.
11. Krishna, H. and Pundir, P.S. (2009). Discrete Burr and discrete Pareto distributions. *Stat. Methodology*, 6, 177-188.
12. Lawless, J.F. (1982). *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons, New York.
13. Mudholkar, G.S. and Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing bathtub failure data. *IEEE Trans. Reliability*, 42, 299-302.
14. Mudholkar, G.S., Srivastava, D.K. and Freimer, M. (1995). The exponentiated Weibull family; A re-analysis of the bus motor failure data. *Technometrics*, 37, 436-445.
15. Nagagawa, T. and Osaki, S. (1975). The discrete Weibull distribution. *IEEE Trans. Reliability*, 24, 300-301.
16. Nekoukhou, V., Alamatsaz, M.H. and Bidram, H. (2013). Discrete generalized exponential distribution of a second type. *Statistics: A Journal of Theoretical and Applied Statistics*, 47(4), 876-887.
17. Phyo, I. (1973). Use of a chain binomial in epidemiology of caries. *Journal of Dental Research*, 52, 750-752.

18. Raqab, M.Z. and Kundu, D. (2006). Burr Type X Distribution: Revisited. *Journal of Probability and Statistical Science*, 4(2), 179-193.
19. Ridout, M.S. and Besbeas, P. (2004). An empirical model for under dispersed count data. *Statist. Model*. 4, 77-89.
20. Roy, D. (2003). The discrete normal distribution. *Commun. Statist. Theor. Meth.*, 53(10), 1871-1883.
21. Roy, D. (2004). Discrete Rayleigh distribution. *IEEE Trans. Reliability*, 53(2), 255-260.
22. Steutel, F.W. and Van Harn, K. (2004). *Infinite divisibility of probability distributions on the real line*, New york: Marcel Dekker, Inc.
23. Surles, J.G. and Padgett, W.J. (2001). Inference for reliability and stress-strength for a scaled Burr Type X distribution. *Lifetime Data Anal.*, 7, 187-200.
24. Surles, J.G. and Padgett, W.J. (2005). Some properties of a scaled Burr type X distribution. *J. Statist. Plann. Infer.*, 128(1), 271-280.