Pak. J. Statist.
2019 Vol. 35(1), 75-95

DISCRETE MODIFIED INVERSE RAYLEIGH DISTRIBUTION

Noor Shahid[1] and **Rashida Raheel**[2]
Department of Statistics, Lahore College for Women University,
Lahore, Pakistan.
Email: [1]saminashahid253@yahoo.com
[2]rashidaraheel@yahoo.com

ABSTRACT

In this paper, we have presented a discrete version of Modified Inverse Rayleigh Distribution. We proposed Discrete Modified Inverse Rayleigh distribution. We have derived the reliability properties including survival function, residual reliability function hazard rate function, second rate of failure along with their graphical portray. Further distributional properties like quantile function, skewness, kurtosis, mean deviation and the distribution of order statistics of the proposed distribution are also presented. The estimation of parameters has been done by method of proportion, maximum likelihood method and method of moments. A simulation study has been conducted by these methods for different sample sizes. The suitability is shown by applying the new model on the real data sets. Goodness of the new model on over-dispersed data set is also shown.

KEY WORDS

Discrete Concentration Approach, Modified Inverse Rayleigh Distribution, Maximum likelihood Estimation, Order Statistics, Simulation.

1. INTRODUCTION

Reliability is an important dimension to shed light. Life testing experiments and test of reliability should be done before the shipment of the product. We have come across situations where the lifetime of a product is related to count data. For example life of an electric circuit is measured by the number of breakdown occur during a month, life of a certain drug is measured by the number of days it is best before expire, survival time of a patient suffering from brain hemorrhage is measured by the length of time he stay under observation. In case of on/off of an engine, lifetime is measured as count data. In such situations the observed values are discrete in nature and hence cannot get all points in continuum. Sometimes it can also happen that the variable is continuous but the measurements are observed in a way where discrete model is more appropriate. Researchers are trying to find out models for discrete data sets realized from continuous scenario. Therefor discrete version of the continuous lifetime distributions is proposed. Discrete lifetime phenomenon provides the bases for such discretized versions. These discretize distributions provide precise results for small discrete data set. The discrete versions have extended the scope of reliability analysis. The reliability of count data is observed on the bases of number of times of success/failure. A wide range of distributions are used to model discrete lifetime data such as Negative binomial distribution, Geometric distribution and Generalized Poisson distribution. But a demerit to use these distributions is that these provide imprecise results for the small data set. However, there is still need to find new discrete distributions which are suitable under different conditions. The discrete versions are playing their role efficiently in modeling small data set as well as large data set to some extent.

Various discretization approaches are proposed to present the discrete version of continuous distributions but the most extensively used in the past literature is Discrete Concentration Approach. Nakagawa and Osaki (1975) present the discrete version of Weibull distribution called “Discrete Weibull distribution”. Later on Khan (1989) estimates the parameters of Discrete Weibull distribution. Inusah (2006) obtained the discrete analogue of continuous Laplace distribution. Later, Inusah and Kozubowski (2006) presented the discrete analogue of Skew Laplace distribution. Krishna and Pundir (2007) introduced Discrete Maxwell distribution. Krishna and Pundir (2009) introduced the Discrete Burr and Pareto distribution. Jazi (2010) presented the Discrete Inverse Weibull distribution. Ojeda (2011) developed the Discrete Lindley distribution. Chakraborty and Chakravarty (2012) present the Discrete Gamma distribution, Al-Huniti (2012) presented the Discrete Burr 3 Distribution, Al-Masood (2013) presented the discrete analogues of continuous distributions. Hussain and Ahmad (2014) introduced the discrete Inverse Rayleigh distribution. Alamatsaz et al. (2016) proposed a Discrete Generalized Rayleigh distribution.

Discrete concentration approach is widely under different circumstances. For example When the support of continuous random variable is $x \geq 0$ then the discrete time space will be $N = 0,1,2,...$ with the survival function $P(X \geq x)$. If the support is $x > 0$ then $N = 1,2,...$ provide bases for the discretize random variable with the survival function $P(X > x)$.

The Geometric distribution is the discrete version of Exponential distribution considered as simplest model by using the discrete concentration approach. The probability mass function and survival function of Geometric distribution is given as,

$$P(X = x) = S(x) - S(x+1);\ x = 0,1,2,... \tag{1}$$

$$P(x) = \theta^{x} - \theta^{x+1};\ x = 0,1,2,3....$$

which is obtained by discretizing Exponential distribution with survival function

$$S(x) = e - \lambda x;\ \lambda,\ x \geq 0$$

Here $\theta = e - \lambda\ ,(0 < \theta < 1)$

Many researchers consider the Geometric distribution as the discrete version of Exponential distribution.

Modified Inverse Rayleigh (MIR) distribution is the modified form of Inverse Rayleigh distribution presented by Khan (2014). MIR is a two parameter extension of single parameter Inverse Rayleigh distribution. As the Inverse Rayleigh distribution is a special case of Inverse Weibull distribution. Similarly the Modified Inverse Rayleigh distribution can be observed as a special case of Modified Inverse Weibull distribution proposed Khan and King(2012) i.e. $X < MIW\left(\alpha,\beta,\theta,\eta\right)$ then, for $\beta = 2$ and $\eta = 0$, $X < MIR\left(\alpha,\theta\right)$. When the parameter changes the Modified Inverse Rayleigh distribution approach to Inverse Exponential and Inverse Rayleigh distribution. Let X denote the random variable from the modified inverse rayleigh distribution, i.e. $X < MIR\left(\alpha,\ \theta\right)$. The probability density function is given as,

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

with the survival function

$$S(x) = 1 - e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2},\ x,\ \alpha,\ \theta > 0 \tag{3}$$

where $\alpha > 0$ and $\theta > 0$ are the scale parameters. The new parameter α provides more flexibility in reliability analysis. The reliability and non-reliability functions of MIR shows decreasing and increasing patterns. The failure rate of MIR has upside-down bathtub shape. It has component failure rate for lifetime data. The new distribution is capable of modeling bathtub hazard rate function. MIR distribution has wide application in reliability theory and in prediction purpose of real world problems like economic and weather.

2. DISCRETE MODIFIED INVERSE RAYLEIGH DISTRIBUTION AND ITS PROPERTIES

A random variable X is distributed as Discrete Modified Inverse Rayleigh (DMIR) Distribution with parameters α and λ, denoted by dMIR$\left(\alpha,\ \lambda\right)$, then the probability mass function(pmf) by using approach in Eq.(1) is as follow,

$$P(X = x) = e^{-\frac{\alpha}{(x+1)}-\theta\left(\frac{1}{(x+1)}\right)^2} - e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2},\ x = 0,1,2,\ \ldots. \tag{4}$$

where $S\left(x\right)$ is the survival function of Modified Inverse Rayleigh distribution for which discrete version is proposed. Discrete model is proposed from a continuous model if times are grouped in unit intervals. Therefor the discrete observed random variable $dX = \left[X\right]$ is equal to the greatest integer less than or equal to X. Hence grouping is introduced on time axis.

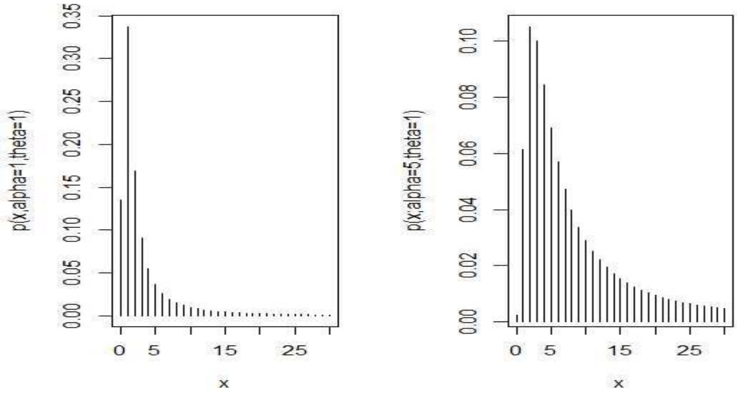

Figure 1: pmf plot of DMIR i.e. p(x,α,θ)

Figure (1) shows the probability mass function plot for the Discrete Modified Inverse Rayleigh distribution for various α and λ. It is observed from the plots that for small values of α the mode of the distribution moves toward left and for large values of α mode moves towards right. For large values of both α and λ, the distribution is positively skewed.

Infinite Divisibility

If a r.v X denotes the number of times a product fails in any given time period then $P(X=0)$ shows the probability that no failure occur in any given time period. This is also a condition for infinite divisibility of a discrete distribution (by Steutel and Van Harn (2004)). Infinite divisibility of discrete random variable implies that $P(X=0)>0$. According to general definition, infinite divisibility of X means the existence for every $k \in N$ of random variable X_k. If p_k, $k \in Z_+$ valued, is infinite divisible if $p_k \le e-1$ for all $k \in N$.

$$P(X=0)=e^{-\alpha-\theta}>0$$

In general if we take $\alpha=1, \theta=1$ then $p_1=0.3371<e^{-1}$.

So DMIR is infinite divisible. If X is shifted to zero then its possible infinite divisibility does not effect, hence it can be concluded that without loss of generality we can restrict attention towards discrete infinite divisibility.

Distribution Function

Short for cdf shows the probability that the underlying variable is less than or equal to the specific point of the function. The cumulative distribution function of DMIR is,

$$F(x)=e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}, \ \alpha>0, \ \lambda>0, \ x=1,2,... \tag{5}$$

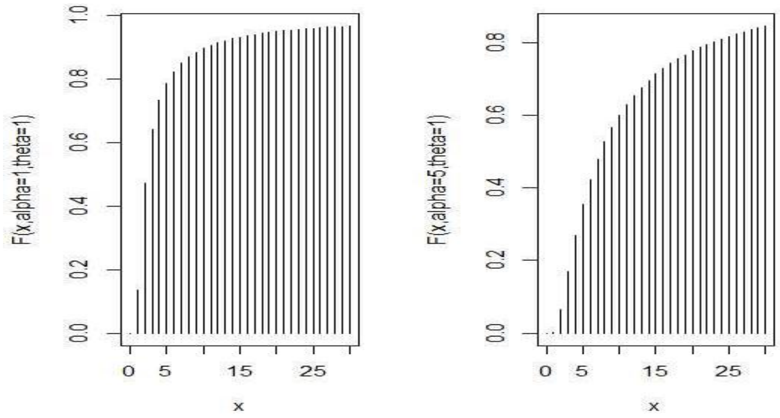

Figure 2: Distribution Function Plot of DMIR

The plot Fig.(2) shows the cumulative distribution function of Discrete Modified Inverse Distribution It shows the concave down increasing behavior for all values of both the parameters.

Survival Function

This shows the probability that any product will survive after specific time. The survival function of DMIR is given as,

$$S(x) = 1 - e^{-\frac{\alpha}{x} - \theta\left(\frac{1}{x}\right)^2}, \ \alpha > 0, \ \lambda > 0, \ x = 1, 2, \ ... \qquad (6)$$

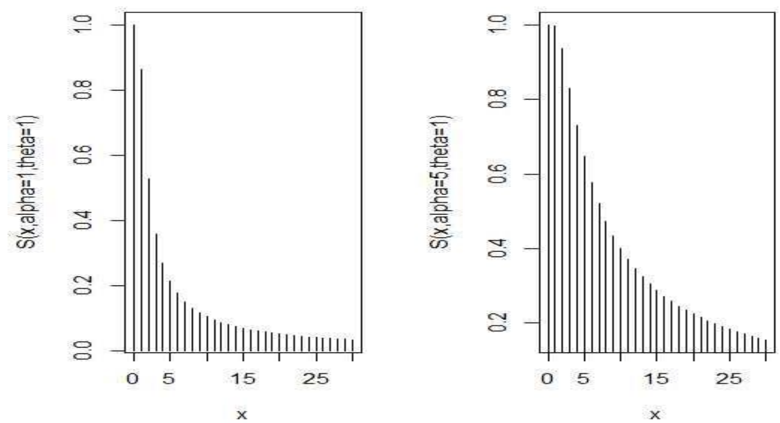

Figure 3: Survival Function Plot of DMIR

In Fig.(3), the survival function of DMIR shows monotone non-increasing function.

2.1. Statistical Properties

2.1.1. Residual Reliability Function

The $R(i \mid x)$ defines the probability that an object will survive at time i given it has survived till time X . The residual reliability function of DMIR is given as,

$$R(i \mid x) = \frac{S(x+i)}{S(x)}$$

$$R(i \mid x) = \frac{1-e^{-\frac{\alpha}{(x+i)}-\theta\left(\frac{1}{(x+i)}\right)^2}}{1-e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}}, \ \alpha > 0,\ \lambda > 0,\ x = 0,1,2,\ldots$$

2.1.2. Hazard Rate

The hazard rate is the ratio of probability mass function to the survival function. Hazard rate gives the probability that a failure of a product occur in specified interval.

$$h(x) = \frac{P(X = x)}{S(x)}$$

$$h(x) = \frac{S(x+1) - S(x)}{S(x)}$$

There are some problems with hazard rate given above. As $h(x) \le 1$, so failure rate function and failure probability create some confusion. Also the cumulative hazard rate function $\mathrm{H(x)} = \sum_{t=1}^{x} h(t) \neq -\ln S(x)$ (Xie et al. (2002)). Due to these problems a second rate of failure is defined. It is based on continuous case. In continuous case $h(x) = -\frac{d\ln(S(x))}{dx}$ and for the discrete case, it is replaced by $h(x) = -\left[\ln S(x+1) - \ln S(x)\right]$. Hence show the same monotonicity $H(x) = -\ln S(x)$.

Using Eq.(4) and Eq.(6), the hazard rate function is,

$$h(x) = \frac{e^{-\frac{\alpha}{x+1}-\theta\left(\frac{1}{x+1}\right)^2} - e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}}{1-e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}}, \ \alpha > 0,\ \lambda > 0,\ x = 0,1,2,\ \ldots$$

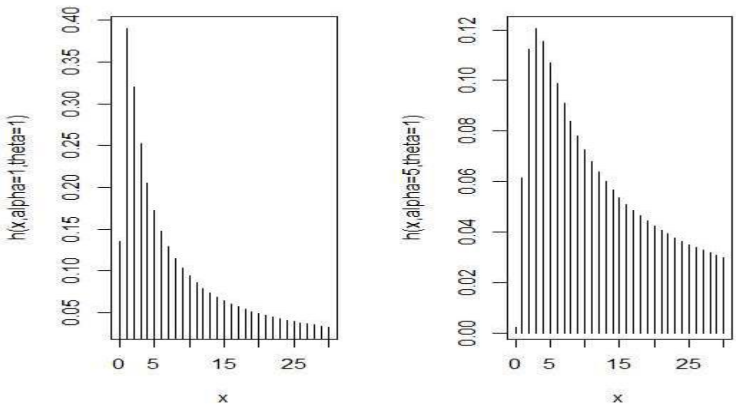

Figure 4: Hazard Rate of DMIR

In Fig.(4), the hazard rate of DMIR is shown by taking different values of α and θ. For small value of α the hazard rate function shows upside down bathtub shape and for large value of α hazard rate is increasing for early time period.

2.1.3. Second Rate of Failure

The second failure rate denoted by SRF is given as,

$$SRF = \ln \frac{S(x)}{S(x+1)}$$

By Using Eq.(6)

$$SRF = \ln \frac{1-e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}}{1-e^{-\frac{\alpha}{x+1}-\theta\left(\frac{1}{x+1}\right)^2}} \text{ , } \alpha > 0, \lambda > 0, x = 0, 1, 2, \ldots$$

2.1.4. Cumulative Hazard Rate

The cumulative hazard rate function is the probability of failure at any time x given that survives until time x. The cumulative hazard rate of DMIR is denoted by $H(x) = \sum_{t=1}^{x} h(t)$ and is given as,

$$H(x) = \sum_{t=1}^{x} \frac{e^{-\frac{\alpha}{t+1}-\theta\left(\frac{1}{t+1}\right)^2} - e^{-\frac{\alpha}{t}-\theta\left(\frac{1}{t}\right)^2}}{1-e^{-\frac{\alpha}{t}-\theta\left(\frac{1}{t}\right)^2}} \text{ , } \alpha > 0, \lambda > 0, x = 0, 1, 2, \ldots$$

2.1.5. Moment Properties.

Let $X=[Y]$ denotes a Discrete Modified Inverse random variable. Using the expression given by Hussain and Ahmad (2014).

$$E(X)=\sum_{x=0}^{\infty}\left(\phi(x)-\phi(x-1)\right)S(x)+\phi(0) \tag{7}$$

where $S(x)$ is the survival function and $\phi(x)$ is the function of X The mean and variance for the DMIR is obtained by using above expression and Eq.(6). Put $\phi(x)=X$ and $\phi(x)=X^2$.

$$E(X)=\sum_{x=1}^{\infty}1-e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2} \tag{8}$$

$$Var(X)=\sum_{x=1}^{\infty}(2x-1)\left[1-e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\right]-\left[\sum_{x=1}^{\infty}1-e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\right]^2 \tag{9}$$

The rth moment about origin of DMIR is obtained by substituting $\phi(x)=X^r$

$$E\left(X^r\right)=\sum_{x=1}^{\infty}\left(x^r-(x-1)^r\right)\left(1-e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\right)$$

By substituting $\phi(x)=e^{tx}$, we get the moment generating function of DMIR.

$$E\left(e^{tx}\right)=\sum_{x=1}^{\infty}\left(e^{tx}-e^{t(x-1)}\right)\left(1-e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\right)+1$$

The expression for probability generating function is obtained by $\phi(x)=t^x$

$$E\left(t^x\right)=\sum_{x=1}^{\infty}t^x\left(1-\frac{1}{t}\right)\left(1-e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\right)+1$$

Analytical expressions in closed form is not easy to calculate. Hence one can solve it numerically by using specialize software and by giving specific values to the parameters. If $E\left(X^k\right)$ gives finite answer, all the moments less than k will be finite. Numerically, it is proved that $E\left(X^k\right)$ gives finite result, resultantly all moments less than k are finite (see Rohatgi et al. (2003)).

2.1.6. Index of Dispersion

Index of dispersion (ID) is defined as,

$$ID = \frac{Variance}{Mean}$$

Index of dispersion is used to indicate that whether a distribution is appropriate for an under dispersed data or for an over dispersed data. ID < (>) 1 indicate that the distribution is under dispersed (over dispersed) (see chakraborty and chakraverty (2012). The ID of DMIR is given in the table below where the upper value indicate the mean and the middle value indicate the variance and the lower value indicates the ID factor for different values of α and θ.

Table 1
Table of Index of Dispersion

α / θ	**0.5**	**1**	**3**	**5**	**7**	**9**
0.5	2.9143 90.6891 31.1183	2.3079 92.9130 40.2587	4.3744 98.5080 22.5192	5.1246 104.3146 20.3485	5.7534 109.9115 19.1037	6.3037 115.2855 18.2885
1	4.9082 170.2043 34.7058	5.2090 170.5494 32.7413	6.0913 173.5818 28.4967	6.7525 177.4267 26.2757	7.31580 181.3887 24.7941	7.8162 185.3110 23.7086
3	11.5023 421.9706 36.6858	11.6455 421.4094 36.1864	12.1431 420.3283 34.6146	12.5727 420.1510 33.4178	12.9606 420.4411 32.4400	13.3179 421.0392 31.6145
5	16.8346 608.5575 36.1492	16.9245 607.8779 35.9170	17.2595 605.7173 35.0947	17.5664 604.1452 34.3921	17.8526 602.9698 33.7749	18.1227 602.0883 33.2229
7	21.4162 752.0231 35.1147	21.4807 751.3294 34.9769	21.7283 748.8639 34.4649	21.9617 746.7805 34.0037	22.1841 744.9861 33.5820	22.3971 743.4216 33.1280
9	25.4690 863.0311 33.9248	25.5188 863.3554 33.8321	25.7124 860.8435 33.4797	25.8983 858.5859 33.1522	26.0775 856.5341 32.8457	26.2511 854.6536 32.5569

As shown in Table 1 that the ID for different values of α and θ is greater than 1 so it is observed that DMIR is appropriate for over dispersed data.

2.2. Further Properties

2.2.1. Quantile Function

The quantile function is derive by using the Eq.(5).

$$q = F\left(x\right)$$

$$q = e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}$$

Let $t = 1/x$ and hence the equation become

$$\ln\left(q\right) = -\left(\alpha\right)t - \left(\theta\right)\left(t\right)2$$

By making use of quadratic formula and Hence $x = \left(1/t\right)$

$$x = \frac{2\theta}{-\alpha + \sqrt{(a)^2 - 4\theta\left(\ln(q)\right)}} \tag{10}$$

Put $q = 1/2$ the median of DMIR is as,

$$x_m = \frac{2\theta}{-\alpha + \sqrt{(a)^2 - 4\theta\left(\ln(0.5)\right)}} \tag{11}$$

Similarly, the co-efficient of skewness is calculated by using the Bowley's co-efficient of Skewness.

$$Sk = \frac{Q\left(\frac{1}{4}\right) + Q\left(\frac{3}{4}\right) - 2x_m}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)}$$

where $Q\left(1/4\right)$ and $Q\left(3/4\right)$ are obtained by putting $q = 1/4$ and $q = 3/4$ in Eq.(10) and xm is the median in Eq.(11). The co-efficient of kurtosis is calculated by using the Percentile co-efficient of Kurtosis.

$$K = \frac{\frac{Q\left(\frac{1}{4}\right) + Q\left(\frac{3}{4}\right)}{2}}{P\left(\frac{90}{100}\right) - P\left(\frac{10}{100}\right)}$$

where $Q\left(1/4\right)$ and $Q\left(3/4\right)$ are obtained by putting $q = 1/4$ and $q = 3/4$ in Eq.(10). $P\left(90/100\right)$ and $P\left(10/100\right)$ is obtained by putting $q = 90/100$ and $q = 10/100$ in Eq.(10).

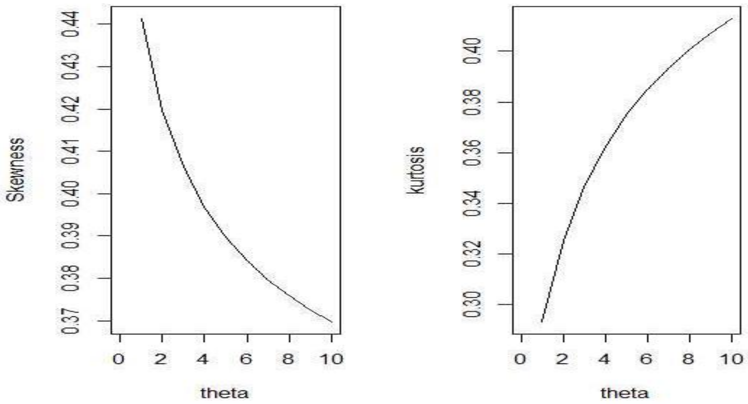

Figure 5: Skewness and Kurtosis of DMIR

Fig.(5) shows the skewness and kurtosis of DMIR for fixed $\alpha = 1$. It portrays that the co-efficient of skewness is decreasing for increasing values of θ. It can also be seen that the new distribution is positively skewed. The graph of kurtosis shows that as the value of parameter θ increases, the percentile co-efficient of kurtosis increases. Here, the co-efficient of Kurtosis by fixing α and increasing θ, the value is less than 0.256 which means it shows the shape is platykurtic.

2.2.2. Mean Deviation.

The Mean deviation about mean is,

$$\delta_1(x) = \sum_{x=0}^{\infty} | x - \mu | \, p(x)$$

Using the result in Eq.(4) and for the solution of above equation (see Bakouch, Jazi and Nadarajah (2014)), we get

$$= 2\mu F(\mu) - 2\sum_{x=0}^{\mu} x p(x)$$

$$\delta_1(x) = 2\mu\left(e^{-\frac{\alpha}{\mu}-\theta\left(\frac{1}{\mu}\right)^2} \right) + 2\sum_{x=1}^{\mu} e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2} - 2\mu\left(e^{-\frac{\alpha}{\mu+1}-\theta\left(\frac{1}{\mu+1}\right)^2} \right)$$

The mean deviation from median is,

$$\delta_2(x) = \sum_{x=0}^{\infty} | x - M | \, p(x)$$

Using the result in Eq.(4) and for the solution of above equation (see Bakouch, Jazi and Nadarajah (2014)), we get

$$= M - 2\sum_{x=0}^{M} xp(x)$$

$$\delta_2(x) = M + 2\sum_{x=1}^{M} e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2} - 2M\left(e^{-\frac{\alpha}{M+1}-\theta\left(\frac{1}{M+1}\right)^2}\right)$$

2.2.3. Distribution of ith Order Statistics

Let $X_{(1)}, X_{(2)}, \ldots X_{(n)}$ be the order sample of size n from DMIR. The pmf of $X_{(i)}$ is,

$$P\left(X_{(i)} = x\right) = \left(X_{(i)} \le x\right) - P\left(X_{(i)} \le x-1\right)$$

$$P(X_i = x) = \int_0^{F(x)} \frac{1}{B(i, n-i+1)} u^{i-1}(1-u)^{n-i}\,du - \int_0^{F(x-1)} \frac{1}{B(i, n-i+1)} u^{i-1}(1-u)^{n-i}\,du$$

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

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

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

$$= i\binom{n}{i} \sum_{j=0}^{n-i} \binom{n-i}{j} (-1)^j \frac{u^{i+j}}{(i+j)}$$

$$= \binom{n}{i} \sum_{j=0}^{n-i} \binom{n-i}{j} \frac{(-1)^j (i)_j}{(i+1)_j} \left(e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2 (i+j)} - e^{-\frac{\alpha}{x-1}-\theta\left(\frac{1}{x-1}\right)^2 (i+j)} \right)$$

$$= \binom{n}{i} \sum_{j=0}^{n-i} \binom{n-i}{j} \frac{(-1)^j (i)_j}{(i+1)_j} e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2 (i+j)} - \binom{n}{i} \sum_{j=0}^{n-i} \binom{n-i}{j} \frac{(-1)^j (i)_j}{(i+1)_j} e^{-\frac{\alpha}{x-1}-\theta(\frac{1}{x-1})^2 (i+j)} \qquad (12)$$

Results given by Hussain and Ahmad (2014)

$$\sum_{j=0}^{n-i}\binom{n-i}{j}\frac{(-1)^j (i)_j}{(i+1)_j} e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2 (i+j)} = \left(e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2 (i)}\right) {}_2F_1\left(-n+i, i, i+1; e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\right) \tag{13}$$

and

$$\sum_{j=0}^{n-i}\binom{n-i}{j}\frac{(-1)^j (i)_j}{(i+1)_j} e^{-\frac{\alpha}{x-1}-\theta(\frac{1}{x-1})^2 (i+j)}$$

$$= \left(e^{-\frac{\alpha}{x-1}-\theta\left(\frac{1}{x-1}\right)^2 (i)}\right) {}_2F_1\left(-n+i, i, i+1; e^{-\frac{\alpha}{x-1}-\theta\left(\frac{1}{x-1}\right)^2}\right) \tag{14}$$

where

$$F_2(-m, b, c; z) = \sum_{j=0}^{m}\binom{m}{j}(-1)^j \frac{(b)_j}{(c)_j} z^j$$

and

$(b)_j = 1; j = 0$ and $j = 0$ *and* $b(b+1)....(b+n-1); j > 0$

Hence substituting Eq.(13) and Eq. (14) in Eq. (12)

$$P(X_{(i)} = x) = (i)K_i \left[\left(e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2 (i)}\right) {}_2F_1\left(-n+i, i, i+1; e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\right)\right.$$

$$\left. -\left(e^{-\frac{\alpha}{x-1}-\theta\left(\frac{1}{x-1}\right)^2 (i)}\right) {}_2F_1\left(-n+i, i, i+1; e^{-\frac{\alpha}{x-1}-\theta\left(\frac{1}{x-1}\right)^2}\right)\right] \tag{15}$$

where

$$K_{i=} \frac{1}{(i)}\binom{n}{i}$$

The cumulative distribution of $X_{(i)}$ is

$$P(X_i < x) = \sum_{t=i}^{n}\binom{n}{t} F(x)^t (1-F(x))^{n-t}$$

$$= \frac{n!}{(i-1)!(n-1)!}\int_0^{F(x)} u^{i-1}(1-u)^{n-i} du$$

$$= i\binom{n}{i}\int_0^{F(x)} u^{i-1}\sum_{j=0}^{n-i}\binom{n-i}{j}(-1)^j u^j du$$

$$= i\binom{n}{i}\sum_{j=0}^{n-i}\binom{n-i}{j}(-1)^j \int_0^{F(x)} u^{i+j-1} du$$

$$= i\binom{n}{i}\sum_{j=0}^{n-i}\binom{n-i}{j}\frac{(-1)^j}{(i+1)_j}e^{-\frac{\alpha}{x}-\theta(\frac{1}{x})^2(i+j)}$$

$$P(X_i < x) = F_1\left(-n+i, i, i+1; e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\right). \tag{16}$$

3. ESTIMATION

The parameters of DMIR is obtained by using the method of moments and method of proportion (see Khan (1989) and Jazi (2010)) and the likelihood method of estimation.

3.1. Method of Maximum Likelihood Estimation

Let $X_1, X_2, ..., X_n$ be a random sample of lifetime of n items. The sample is independently identically distributed from DMIR distribution, the likelihood function is given as,

$$l(\alpha,\theta) = \prod_{i=1}^{n} P(X = x) = \prod_{i=1}^{n}\left(e^{-\frac{\alpha}{x+1}-\theta\left(\frac{1}{x+1}\right)^2} - e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\right)$$

$$\ln l(\alpha,\theta) = \sum_{i=1}^{n}\ln P(X = x) = \sum_{i=1}^{n}\ln\left(e^{-\frac{\alpha}{x+1}-\theta\left(\frac{1}{x+1}\right)^2} - e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\right)$$

Partially differentiate above equation with respect to α and θ.

$$\frac{\partial \ln l}{\partial \alpha}(\alpha,\theta) = \sum_{i=1}^{n}\left(\frac{e^{-\frac{\alpha}{x+1}-\theta\left(\frac{1}{x+1}\right)^2}\left(-\frac{1}{x+1}\right) - e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\left(-\frac{1}{x}\right)}{e^{-\frac{\alpha}{x+1}-\theta\left(\frac{1}{x+1}\right)^2} - e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}}\right) \tag{17}$$

$$\frac{\partial \ln l}{\partial \theta}(\alpha,\theta) = \sum_{i=1}^{n}\left(\frac{e^{-\frac{\alpha}{x+1}-\theta\left(\frac{1}{x+1}\right)^2}\left(-\left(\frac{1}{x+1}\right)^2\right) - e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}\left(-\left(\frac{1}{x}\right)^2\right)}{e^{-\frac{\alpha}{x+1}-\theta\left(\frac{1}{x+1}\right)^2} - e^{-\frac{\alpha}{x}-\theta\left(\frac{1}{x}\right)^2}}\right) \tag{18}$$

Equate Eq.(17) and Eq.(18) to zero. Then mle's of α and θ can be obtained by using numerical methods.

3.2. Proportion Method

A new method of proportion is proposed by Khan et al. (1989) and later on Jazi (2010) to estimate the parameters. Here we used this method to estimate the parameters of Discrete Modified Inverse Rayleigh Distribution. Let $X_1, X_2, ..., X_n$ be a random sample taken from DMIR distribution, then define an indicator variable as,

$$\begin{cases} 1 \text{ for } x_i = 0 \\ 0 \text{ for otherwise} \end{cases} \tag{19}$$

Now $W = \sum_{i=i}^{n} I(x_i)$ be the number of 0's in the sample. So the probability $P_0(\alpha, \theta)$ is estimates by proportion W/n. Now define another indicator variable as, Here it is clear that $P_1(\alpha, \theta)$ can be obtained by using the proportion $Y/n = \sum_{i=1I}^{n}(x_i)/n$, where Y shows the proportion of 1's in the sample. Solve the two equations simultaneously to get the estimates.

$$\hat{\alpha} = -\log\left(\frac{W}{n}\right) - \hat{\theta}$$

$$\hat{\theta} = -4\log\left(\frac{Y+W}{n}\right) - 2\hat{\alpha}$$

Substitute the value of $\hat{\theta}$ in $\hat{\alpha}$, the value of $\hat{\alpha}$ becomes,

$$\hat{\alpha} = \log\left(\frac{W/n}{((W+Y)/n)^4}\right)$$

Thus, the above estimates are consistent estimates of the parameters. W/n and $(W+Y)/n$ are the unbiased and empirical consistent estimates of the $P(X \le 0)$ and $P(X \le 1)$.

3.3. Method of Moments

Method of moments to estimate the parameter is stated as to equate the first two population moments with the sample moments and solve the equations simultaneously. Here the moments are not obtained in closed form so these equations cannot be solved with this technique. So we used the method of pseudo moments by minimizing the $S(\alpha, \theta)$ with respect to the α and θ as proposed by Khan (1989).

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

where $M_1 = \frac{\sum_{i=1}^{n} x_i}{n}$ and $M_2 = \frac{\sum_{i=1}^{n} x_i^2}{n}$

$$E(X) = \sum_{i=1}^{n} xp(x) \text{ and } E\left(X^2\right) = \sum_{i=1}^{n} x^2 p(x)$$

Simulation: A simulation study was performed by generating 500 samples of various sizes from DMIR. Applying bivariate N-R method with each of these samples, we obtained mle's of both the parameters with Fisher Information matrix. The method of moments is also used. This method yields smaller variance. Simulation results of all the mle method are given below in the table. In the table all the entries, above are bias and below are mse for 500 samples of size n = 60, 80, 100.

Table 2
Simulation Results for Parameters above (Below) Bias (MSE)

		Mle's
α=0.5, θ=2	n=60	0.3323 -0.2300 (0.1992) (0.0895)
	n=80	0.4137 -0.2955 (0.2782) (0.1361)
	n=100	0.4641 -0.3275 (0.3830) (0.1827)
α=1, θ=2	n=60	-0.7754 0.5342 (1.3641) (0.6105)
	n=80	-0.4295 0.3129 (0.3632) (0.2024)
	n=100	-0.4349 0.3054 (0.2742) (0.1287)
α=2, θ=2	n=60	-0.779 0.4431 (0.9107) (0.2790)
	n=80	-0.9597 0.5658 (1.4042) (0.4588)
	n=100	-0.9886 0.0517 (1.6333) (0.6719)

From the above result, it is evident that the method of maximum likelihood can be taken up as a consistent method. The estimates obtained by method of mle method are also asymptotically unbiased and consistent.

4. GOODNESS OF FIT

In this section, we use different goodness of fit test on real data sets to check whether the proposed distribution fits the data set better or not. Also the comparison is made with Discrete Inverse Rayleigh distribution. The parameters are estimated by using the likelihood method of estimation in this section. Proposed distribution is also fitted on the over-dispersed data set to check its appropriateness on the over-dispersed data. Different goodness of fit test are given below that we have used.

H_o : The data follow the proposed distribution
H_A : The data do not follow the proposed distribution.

4.1. Kolmogorov-Smirnov Test

This goodness of fit test encounter an empirical cumulative function and cumulative function of the continuous/discrete distribution. Let a random sample from a distribution with cdf $F(x)$ is consider. The empirical cumulative distribution function is given as,

$$F_n(x) = \frac{No.\ of\ \ obs. < x}{n}$$

The K.S test statistics is defines as,

$$D = \sup |F_n(x) - F(x)|$$

4.2. Chi-Square Test

The chi-square test is another goodness of fit test to check whether the data come from a distribution or not. This test statistic applies to the group data. There are number of ways to calculate the number of classes ad form group frequency distribution. The test statistics is defines as,

$$X^2 = \sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i}$$

where o_i and e_i are the observed and expected frequencies. The null hypothesis is rejected if the calculated value is greater than the $\chi^2{}_{k-m-1}$, where k is the number of classes and m is the number of the parameters.

4.3. Cramer-Von Mises Test

Like other non-parameter goodness of fit test, cramer-von mises test is another test that have been shown a most powerful test. The original test was developed by Cramer and Richard von Mises. The test statistic is given as,

$$T = \frac{1}{12n} + \sum_{i=1}^{n} \left[\frac{2i-1}{2n} - F(x_i) \right]$$

where $F(x)$ is the cdf of the proposed distribution. The null hypothesis will be rejected if the p-value is less than the significance level.

4.4. Anderson Darling Test

This test is the extension of cramer-von mises test and also used to check the goodness of the fit. This test was developed by Anderson and Darling, The test statistic is given as,

$$AD = -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \left[\ln F(x_i) + \ln\left(1 - F(x_{n+1-i})\right) \right]$$

The null hypothesis will be rejected if the p-value is less than the significance level.

5. APPLICATION

In this section, discrete Modified Inverse Rayleigh distribution is fitted on real data sets and compared with Discrete Inverse Rayleigh Distribution (DIR) and Generalized Poisson Distribution (GPD). For convenience, Parameters are estimated on R computational package by using the method of maximum likelihood distribution.

Example 1

Data is given in time to death (in weeks) of AG positive leukemia patients (chakraborty and chakravarty (2012) also see Hand el al. (1994)).

65, 156, 100, 134, 16, 108, 121, 4, 39, 143, 56, 26, 22, 1, 1, 5, 65.

Table 3
Data on Time to Death of Leukemia Patients

Model	AIC	BIC	K.S test	p-value	cvm test	p-value	AD	p-value
DMIR	194.0	196.0	0.3445	0.0354	0.6348	0.0172	3.8736	0.0103
DIR	266.7552	267.5884	0.7215	0.0014	2.7184	0.000	39.506	0.0238
GPD	333.8409	332.6741	0.8742	0.0190	3.7964	0.000		

The summary of the fitted model is given in Table 3 and this is compare with Discrete Inverse Rayleigh distribution and Generalized Poisson Distribution on the basis of Akaike information criteria(AIC) and Bayesian information criteria (BIC) and Kolmogrov-Smirnov(K-S) test and Cramer-Von Mises (cvm) test and Andersen Darling(AD) Test. As the AIC and BIC is small as compared to the other distributions so the new model provides better fit. The p-value of K.S test shows that the Discrete Modified Inverse Rayleigh distribution better fit at significance level $\alpha = 1$. Similarly p-value for cramer-von mises test and Anderson Darling Test shows that at significance level $\alpha = 1$ the Discrete Inverse Rayleigh distribution and Generalized Poisson distribution do not good fit at all, where Discrete Modified Inverse Rayleigh distribution provide good fit. We are unable to reject null hypothesis and conclude that the data follow the proposed distribution.

Example 2

The given data shows the number of stillbirths for 402 New Zealand white rabbit litters (by Para and Jaan (2016)). This data set is an over-dispersed data as the ratio of its variance to its mean is greater than 1. So to check the appropriateness of the new model on the over-dispersed data set, model is fitted on the given data.

Table 4
Data on Number of Stillbirths for White Rabbit Litters

No. of Stillbirths	0	1	2	3	4	5	6	7	8	Total
Frequency	314	48	20	7	5	2	2	1	3	402

Let X be the number of stillbirths in 402 New Zealand white rabbit litters. In this example Pearson'Chi-square test is used to check the goodness of the fitted new proposed model. The value of Chi-square test statistics along with its p-value and the value of Akaike Information Criteria and Bayesian Information Criteria is shown in the table.

Table 5
Data on Number of Stillbirths for White Rabbit Litters

No. of Stillbirths	0	1	2	3	4	5	6	7	8	Total
Frequency	314	48	20	7	5	2	2	1	3	402
Exp. fre DMIR	313.48	55.32	13.99	5.91	3.26	2.01	1.33	0.97	0.72	402
Exp. fre DIR	312.20	65.18	13.49	4.83	2.26	1.23	0.74	0.48	0.33	402
Exp. fre GPD	313.64	49.52	18.16	14.14	4.63	2.68	1.63	1.03	0.67	402

Table 6
Summary Results for the Data of Number of Stillbirths for White Rabbit Litters

Model	Estimates	Chi-square	p-value
DMIR	$\alpha\hat{}=0.0961$, $\hat{}\theta=0.1526$	6.447	0.0918
DIR	$q\hat{}=0.7766$	12.5187	0.0055
GPD	$a\hat{}=0.2482$, $\hat{}z=0.4524$	4.5212	0.2104

Table 6 shows that the Discrete Modified Inverse Rayleigh distribution and GPD fits the data better than Discrete Inverse Rayleigh Distribution. The p-value of Chi-square test statistic shows that DIR do not provide the good fit at all, whereas Discrete Modified Inverse Rayleigh Distribution and Generalized Poisson Distribution gives good fit. Based on the value of Chi-square test statistics along with its p-value, we are able to accept the null hypothesis and conclude that data follows the proposed distribution.

Example 3

In this example, Discrete Modified Inverse Rayleigh distribution is fitted to over-dispersed data set. Here the mean=0.6825 and variance of the data set is 0.8137 so the index of dispersion is 1.1922 which shows that the data is over dispersed. The given data set is the distribution of yeast cells in 400 squares of haemacytometer observed by Students (1907) (see Hussain and Ahmad (2014)).

Table 7
Data on Yeast Cells in 400 Squares of Haemacytometer

No. of Cells	0	1	2	3	4	5
Freq.	213	128	37	18	3	1
DMIR(0.00017,0.6203)	215.076	127.43	30.85555	11.45	5.42	9.8
DIR(0.5335)	213.5913	128.7474	31.1424	11.5464	5.4641	9.96
GPD(0.6245,0.0848)	214.2116	126.6415	44.0050	11.7725	2.6886	0.5530

The value of the chi-square along with its p-value for Discrete Modified Inverse Rayleigh distribution, DIR distribution and GPD is given in Table 8.

Table 8
Value of Chi-Square Statistics along with p-value on Data on Yeast Cells in 400 Squares of Haemacytometer

No. of Cells	Chi-square	p-value
DMIR(0.00017,0.6203)	2.08	0.1492
DIR(0.5336)	2.0199	0.3660
GPD(0.6245,0.0848)	8.1609	0.0043

The p-value of chi-square test statistic shows that the GPD do not fit good on the data set, whereas DMIR and DIR provides good fit. The p-value for DMIR distribution leads to the acceptance of the null hypothesis and conclude that the data follows the DMIR distribution.

CONCLUSION

The main objective of this study is to propose a new discrete distribution which is suitable for modeling lifetime data. We derive the Discrete Modified Inverse rayleigh distribution. We derive the basic reliability and distributional properties. The graphical description is also presented. This distribution is suitable fit for lifetime data showing either decreasing or upside down bathtub hazard rate. The parameter estimation is undertaken by the three methods. Also the simulation study is performed by all three methods. The DMIR is fitted on real data sets of leukemia patients time to death, yeast cells in squares of haemacytometer and number of stillbirths. We conclude that DMIR is better fit than DIR distribution on the basis of AIC, BIC, Kolmogrov-Smirnov test, Cramer-von Mises test, Anderson Darling Test and Chi-square goodness of fit test.

ACKNOWLEDGEMENT

The authors are grateful to the references and editors and reviewers for their suggestions which made its improvement possible.

REFERENCES

1. Alamatsaz, M.H., Dey, S., Dey, T. and Harandi, S.S. (2016). Discrete Generalized Rayleigh Distribution. *Pakistan Journal of Statistics*, 32(1), 1-20.
2. AL-Huniti, A.A. and AL-Dayian, G.R. (2012). Discrete Burr Type 3 Distribution. *American Journal of Mathematics and Statistics*, 2(5), 145-152.
3. Al-Masoud, T.A., Elgayar, D.H. and Atallah, D.A. (2013). *A Discrete General Class of Continuous Distributions*. King Abdul Aziz University, Jeddah-Saudi Arabia.
4. Bakouch, H.S., Jazi, M.A. and Nadarajah, S. (2014). A new discrete distribution. *Statistics*, 48(1), 200-240.

5. Chakraborty, S. and Chakravarty, D. (2012). Discrete Gamma Distribution: Properties and Parameter Estimation. *Communications in Statistics-Theory and Methods*, 41(18), 3301-3324.
6. Gomez-Deniz, E. and Calderin-Ojeda, E. (2011). The discrete Lindley distribution: properties and applications. *Journal of Statistical Computation and Simulation*, 81(11), 1405-1416.
7. Hussain, T. and Ahmad, M. (2014). Discrete Inverse Rayleigh Distribution. *Pakistan Journal of Statistics*, 30(2), 203-222.
8. Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. and Ostrowski, E.O. (1994). *A Hand Book of Small Data Sets*. London: Chapman and Hall.
9. Inusah, S. and Kozubowski, T.J. (2006). A discrete analogue of the Laplace distribution. *Journal of Statistical Planning and Inference*, 136, 1090-1102.
10. Inusah, S. and Kozubowski, T.J. (2006). A Skew Laplace Distribution on Integers. The Institute of Statistical Mathematics. *ASIM*, 58, 555-571.
11. Jazi, M.A., Lai, C.D. and Alamatsaz, M.H. (2010). A discrete Inverse Weibull distribution and estimation of its parameters. Statistical Methodology, 7, 121-132.
12. Khan, A.M.S., Khalique, A. and Abouammoh, A.M. (1989). On estimating parameters in a discrete Weibull distribution. *IEEE Transactions on Reliability*, 38(3), 348-350.
13. Khan, M.S. (2014). Modified Inverse Rayleigh Distribution. *International Journal of Computer Applications*, 87(13), 2833.
14. Khan, S. and King, R. (2012). Modified Inverse Weibull Distribution. *Journal of Statistics Applications and Probability*, 1(2), 115-132.
15. Krishna, H. and Pundir, P.S. (2007). *Discrete Maxwell Distribution*. http://interstat.statjournals.net/Y EAR/2007/articles/0711003.pdf.
16. Krishna, H. and Pundir, P.S. (2009). Discrete Burr and Discrete Pareto distributions. *Statistical Methodology*, 6, 177-188.
17. Nakagawa, T. and Osaki, S. (1975). The discrete Weibull Distribution. *IEEE Transactions on Reliability*, 24(5), 300-301.
18. Para, B.A. and Jan, T.R. (2016). Discrete Version of Log-Logistic Distribution and its Application in Genetics. *International Journal of Modern Mathematical Sciences*, 14(4), 407-422.
19. Rohatgi, V.K. (2003). *Statistical Inference*. Courier Corporation.
20. Steutel, F.W. and Van Harn, K. (2004). *Infinite Divisibility of Probability Distributions on the Real Line*. New York: Marcel Dekker.
21. Xie, M., Gaudoin, O. and Bracwuemond, C. (2002). Redefining failure rate function for discrete distribution. *Int. J. Reliab. Quali. Safety Engg*., 9(3), 275-285.