STATISTICA, anno LXXXII, n. 4, 2022

DISCRETE NEW GENERALIZED PARETO DISTRIBUTION

Kuttan Pillai Jayakumar
Department of Statistics, University of Calicut, Kerala-673635, India.
Jiji Jose [1]
Department of Statistics, University of Calicut, Kerala-673635, India.

SUMMARY

In this paper we propose a discrete analogue of New Generalized Pareto distribution as a new discrete model using general approach of discretization of continuous distribution. The structural properties of the new distribution are discussed. The shape properties, moments, median, infinite divisibility and stress-strength properties are derived. Estimation of parameters are done using maximum likelihood method. An application of real data set shows the suitability of the proposed model.

Keywords: Discrete new generalized Pareto distribution; Hazard rate function; Maximum likelihood estimation; Stress-strength reliability.

1. INTRODUCTION

Pareto distribution is well known in the literature for modelling heavy tailed data and is used to model data from various fields such as Economics, Physics, Social Science, Medicine etc. (see Arnold, 2008; Jayakumar *et al.*, 2020, for details). Due to the application of Pareto distribution to set more flexibility in data modelling, a number of generalizations of Pareto distribution are developed in the literature. During the recent decade, several discrete versions of Pareto distributions have been developed owing to the development of new methodologies for generating new families of distributions. Examples include a new discrete Pareto type IV distribution developed by Ghosh (2020). Para and Jan (2016) introduced discrete three parameter Burr type XII and discrete Lomax to model count data of cysts of kidney using steriods, a data from medical science. Prieto *et al.* (2014) introduced and studied discrete generalized Pareto distribution, as a discrete analogue of continuous generalized Pareto distribution, used to model the number of road crashes on blackspots. Buddana and Kozubowski (2014) studied another version

[1] Corresponding Author. E-mail: jijijose2793@gmail.com

of discrete Pareto distribution. Krishna and Pundir (2009) used the discrete Burr and discrete Pareto distribution to model reliability estimation in series system.

The discrete distribution are useful when count data occurs. Discretizing a continuous model is an interesting technique began with the work of Nakagawa and Osaki (1975). Kemp (1997) studied discrete normal dstribution and Nakagawa and Osaki (1975) derived discrete Weibull distribution. A second type of discrete Weibull (DW) distributon was proposed by Stein and Dattero (1984) and third type by Padgett and Spurrier (1985). A discrete exponential distribution was examined by Sato *et al.* (1999). Roy (2003) studied another version of the discrete normal distribution. Discrete analogues of Laplace and skew Laplace was analysed by Inusah and Kozubowski (2006) and Kozubowski and Inusah (2006) respectively. Kemp (2008) examined discrete half normal distribution and Chakraborty and Chakravarty (2012, 2014, 2015) analysed the discrete versions of gamma, Gumbel and power distributions. Jayakumar *et al.* (2020) introduced and studied a new generalizations of Pareto distribution, using the concept of random minimum. In this paper, we introduce a discrete version of the New Generalized Pareto (NGP), distribution studied in Jayakumar *et al.* (2020). Note that the NGP distribution is a generalization of Pareto type III distribution and it is heavy tailed. Moreover, the NGP distribution is subexponential, has dominated variation property and is a member of class $\mathscr{L}$. Also if $X_1, X_2, ..., X_n$ are independent and identically distributed (iid) random variables having NGP distribution and if $S_n = X_1 + X_2 + ... + X_n$, then $P(S_n > x) \sim P(Y_n > x)$ where $Y_n = \max(X_1, X_2, ..., X_n)$. That is, the exceedances of high thresholds by the sum S_n is due to the exceedances of these thresholds by the largest value in the sample, which is of very use in insurance and risk modelling. For the application of subexponential distributions in various fields see Klüppelberg (1988).

In many real life situations where the variable under investigation is modelled by continuous distribution, it has been observed that often the variable is recorded as an integer valued one instead of real-valued, either because of its inherent nature or because of the limitation of measuring instruments which warrants the introduction of a discrete version of the existing continuous distributions. With this background, the primary goal is to provide a discrete analogue of new generalized Pareto distribuion, a competent model when compared with all other existing discrete Pareto models. From the study of newly proposed distribution, it is clear that the existing discrete Pareto models derived by Ghosh (2020), Para and Jan (2016), Prieto *et al.* (2014), Buddana and Kozubowski (2014) and Krishna and Pundir (2009) are special cases of Discrete New Generalized Pareto (DNGP) distribution. So our proposed model has a wider range of applications in various fields such as medical science, reliability, accident analysis etc. Also, from the Section of real data application, the proposed distribution showed better performance in terms of model selection and goodness of fit criterion.

The rest of the paper is organized as follows. In Section 2 we discussed the method of discretization used in the paper. DNGP distribution is introduced in Section 3. In Section 4 structural properties of DNGP distribution are studied. Here, the cumulative distribution, hazard rate function, moments, quantile function, and random variate generation of DNGP random variables are discussed. In Section 5 the parameters of DNGP are estimated using the method of maximum likelihood. Also, a simulation study is carried out to assess the performance of the estimates. In Section 6 we consider a real data set and it is showed that for modelling this data, the DNGP is more appropriate as compared to geometric, discrete Pareto, discrete generalized Pareto, discrete Pareto type IV, discrete Burr, exponential discrete Weibull and generalized discrete Weibull distributions. Conclusions are presented in Section 7.

2. Discretization of a continuous distribution

Deriving discrete analogues of continuous distribution by means of preserving one or more important traits of continuous distribution has received much attention in recent years. In literature, there are several ways proposed to derive discrete distributions from continuous distributions.These were discussed in detail by Chakraborty (2015).

In this paper, we discretize New Generalized Pareto (NGP) distribution introduced in Jayakumar *et al.* (2020) having probability density function (pdf)

$$g(x;\alpha,\beta,\gamma,\theta)=\frac{\alpha\beta^{\alpha}\theta(1-\gamma)\gamma^{\theta}}{1-\gamma^{\theta}}\frac{x^{\alpha\theta}-1}{(\gamma x^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta+1}},\ x>\beta \text{ and } \alpha,\beta,\gamma,\theta>0. \quad (1)$$

Let X be a continuous random variable. Then, the discrete analogue Y of X can be derived by using the survival function as follows. $S(\cdot)$ is the survival function of the random variable X, then

$$P(Y=y)=P(X\geq y)-P(X\geq y+1)=S(y)-S(y+1),\ y=0,1,2,3,\ldots, \quad (2)$$

where $Y=\lfloor X\rfloor$ is the largest integer less than or equal to X. The first and easiest using this approach is the geometric distribution with probability mass function (pmf)

$$p(x)=\theta^{x}-\theta^{x+1},\ x=0,1,2,\ldots,$$

which is derived by discretizing exponential distribution with survival function $S(x)=e^{-\lambda x}$ with $\lambda,x>0$, $\theta=e^{-\lambda}$ and $(0<\theta<1)$.

3. Discrete New Generalized Pareto distribution

Jayakumar and Sankaran (2016) defined Generalized Uniform distribution with pdf,

$$g(y;\gamma,\theta)=\frac{(1-\gamma)\theta\gamma^{\theta}}{(1-\gamma^{\theta})(y(1-\gamma)+\gamma)^{\theta+1}};0<y<1,\gamma>0,\theta>0 \quad (3)$$

and using the transformation $X=\frac{\beta}{Y^{1/\alpha}}$, Jayakumar *et al.* (2020) obtained New Generalized Pareto (NGP) distribution with four parameters and pdf given in Eq.(1).

The survival and hazard rate functions of NGP are:

$$S(x;\alpha,\beta,\gamma,\theta)=\frac{1}{1-\gamma^{\theta}}-\frac{\gamma^{\theta}}{1-\gamma^{\theta}}\left[\frac{x^{\alpha\theta}}{[\gamma x^{\alpha}+(1-\gamma)\beta^{\alpha}]^{\theta}}\right], x\geq\beta \tag{4}$$

and

$$h(x;\alpha,\beta,\gamma,\theta)=\frac{(\frac{1-\gamma}{\gamma})\frac{\theta\alpha\beta^{\alpha}}{x^{\alpha+1}}}{\left[(\frac{1-\gamma}{\gamma})(\frac{\beta}{x})^{\alpha}+1\right]\left[\left[(\frac{1-\gamma}{\gamma})(\frac{\beta}{x})^{\alpha}+1\right]^{\theta}-1\right]}, x>\beta. \tag{5}$$

The pmf of the discrete version of Y of the New Generalized Pareto distribution using the method in Eq.(2) is derived as,

$$\begin{aligned}P_Y(y;\alpha,\beta,\gamma,\theta)=&P(Y=y)=S_x(y)-S_x(y+1)\\=&\frac{\gamma^{\theta}}{1-\gamma^{\theta}}\left[\frac{(y+1)^{\alpha\theta}}{(\gamma(y+1)^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta}}-\frac{y^{\alpha\theta}}{(\gamma y^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta}}\right],\end{aligned} \tag{6}$$

where $y=[\beta],[[\beta+1],[[\beta+2],\ldots,\alpha,\beta,\theta>0$ and $0<\gamma<1$. We call this distribution as Discrete New Generalized Pareto distribution with parameters α,β,γ and θ and is denoted by DNGP$(\alpha,\beta,\gamma,\theta)$.

When $\theta=1$ and $\gamma\rightarrow 1$,the pmf becomes

$$P_Y(y;\alpha,\beta)=\left(\frac{\beta}{y}\right)^{\alpha}-\left(\frac{\beta}{y+1}\right)^{\alpha}, \tag{7}$$

which is the pmf of discrete Pareto distribution.

The pmf of DNGP is plotted in Figure 1. In particular, some representative plots of DNGP pmf for various parametric values are shown.

By fixing the parameters β and γ, the pmf of DNGP is plotted for different values of α and θ. From Figure 1, it appears that the distribution is rightly skewed. As the values of the parameter α increases, the distribution becomes more leptokurtic and the tail stretches more to the right. As the value of θ increases, the peakedness and skewness decreases as well. The set of plots shows that DNGP is flexible for modelling life time data as it has decreasing and increasing followed by decreasing shapes.

4. STRUCTURAL PROPERTIES OF DNGP$(\alpha,\beta,\gamma,\theta)$

In this Section we study some distribution properties of DNGP$(\alpha,\beta,\gamma,\theta)$. Here we obatin the expression for cumulative distribution function, survival and hazard rate function and then plot the hazard rates to identify their behaviour for different parametric values. The moments are numerically obtained and checked whether DNGP is better for both over and under dispersed data through the measure of dispersion index. A new method for generating random samples from DNGP distribution is presented. Finally we study the infinite divisibility and stress-strength reliability of DNGP distribution.

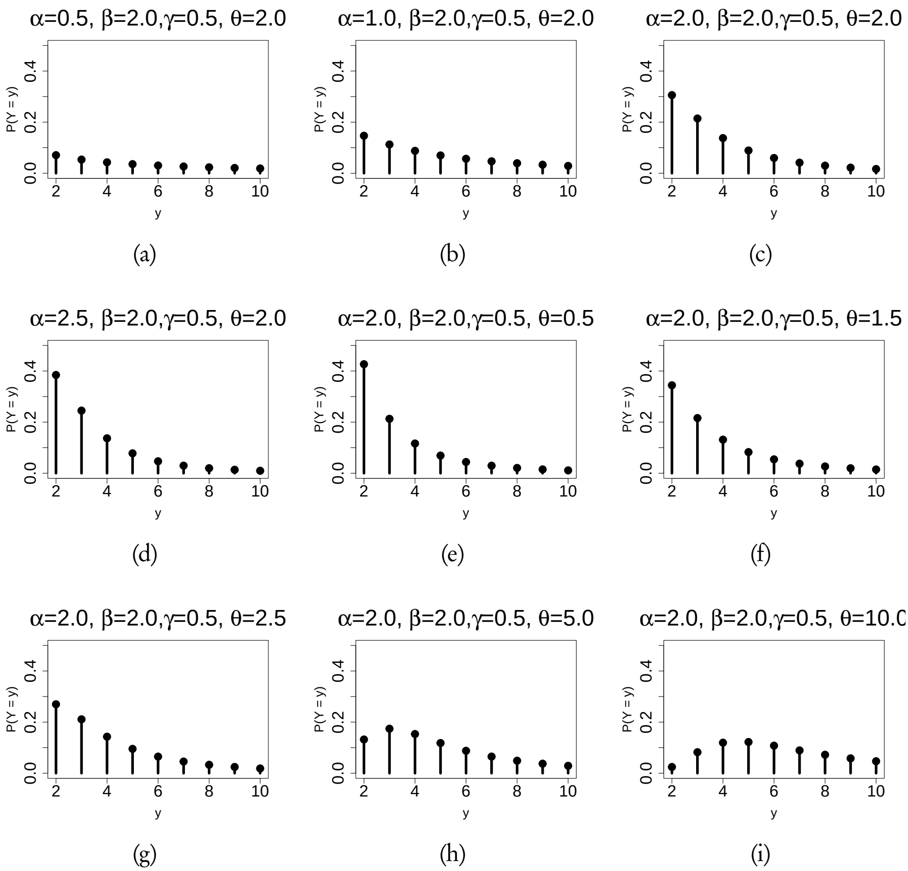

Figure 1 – PMF plot of DNGP($\alpha, \beta, \gamma, \theta$) for different values of $\alpha, \beta, \gamma, \theta$.

4.1. Cumulative distribution function

The cumulative distribution function (cdf) of DNGP($y;\alpha,\beta,\gamma,\theta$) is obtained as

$$\begin{aligned} F(y) =& P(Y \leq y) = 1 - S_x(y) + P(Y = y) \\ =& 1 - \frac{1}{1-\gamma^\theta} + \frac{\gamma^\theta}{1-\gamma^\theta} \frac{(y+1)^{\alpha\theta}}{(\gamma(y+1)^\alpha + (1-\gamma)\beta^\alpha)^\theta} \end{aligned} \tag{8}$$

where $y = [\beta], [\beta+1], [\beta+2], \ldots$. Here, note that

$$F(0) = 1 - \frac{1}{1-\gamma^\theta} + \frac{\gamma^\theta}{(1-\gamma^\theta)(\gamma + (1-\gamma)\beta^\alpha)^\theta}.$$

The proportion of positive values is

$$1 - F(0) = \frac{1}{1-\gamma^\theta} - \frac{\gamma^\theta}{(1-\gamma^\theta)(\gamma + (1-\gamma)\beta^\alpha)^\theta}.$$

Also,

$$P(a < Y \leq b) = \frac{\gamma^\theta}{1-\gamma^\theta} \left[\frac{(b+1)^{\alpha\theta}}{(\gamma(b+1)^\alpha + (1-\gamma)\beta^\alpha)^\theta} - \frac{(a+1)^{\alpha\theta}}{(\gamma(a+1)^\alpha + (1-\gamma)\beta^\alpha)^\theta} \right].$$

4.2. Survival and hazard rate functions

The survival function of DNGP(α,β,γ,θ) is given by,

$$\begin{aligned} S(y) =& P(Y > y) = 1 - P(Y \leq y) \\ =& \frac{1}{1-\gamma^\theta} - \frac{\gamma^\theta}{1-\gamma^\theta} \left[\frac{(y+1)^{\alpha\theta}}{(\gamma(y+1)^\alpha + (1-\gamma)\beta^\alpha)^\theta} \right]. \end{aligned} \tag{9}$$

The hazard rate function of DNGP(α,β,γ,θ) is given by

$$\begin{aligned} h(y) =& P(Y = y / Y \geq y) = \frac{P(Y = y)}{P(Y \geq y)} \\ =& \frac{\gamma^\theta}{1-\gamma^\theta} \left[1 - \frac{y^{\alpha\theta}(\gamma(y+1)^\alpha + (1-\gamma)\beta^\alpha)^\theta}{(y+1)^{\alpha\theta}(\gamma(y^\alpha + (1-\gamma)\beta^\alpha)^\theta} \right]. \end{aligned} \tag{10}$$

The reverse hazard rate and second rate of failure are respectively

$$\begin{aligned} h^*(x;\alpha,\beta,\gamma,\theta) =& P(Y = y / Y \leq y) = \frac{P(Y = y)}{P(Y \leq y)} \\ =& \frac{(y+1)^{\alpha\theta}(\gamma(y^\alpha + (1-\gamma)\beta^\alpha)^\theta - y^{\alpha\theta}(\gamma(y+1)^\alpha + (1-\gamma)\beta^\alpha)^\theta}{[(y+1)^{\alpha\theta} - (\gamma(y+1)^\alpha + (1-\gamma)\beta^\alpha)^\theta][\gamma(y^\alpha + (1-\gamma)\beta^\alpha]^\theta} \end{aligned} \tag{11}$$

and

$$h^{**}(x;\alpha,\beta,\gamma,\theta)=\log\left[\frac{S(y)}{S(y+1)}\right]$$
$$=\frac{\left[(\gamma(y+2)^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta}\right]\left[(\gamma(y+1)^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta}-\gamma^{\theta}(y+1)^{\alpha\theta}\right]}{\left[(\gamma(y+1)^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta}\right]\left[(\gamma(y+2)^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta}-\gamma^{\theta}(y+2)^{\alpha\theta}\right]}. \tag{12}$$

The accumulated hazard function, $H(y)$ is given by,

$$H(y)=\sum_{k=0}^{y}\frac{\gamma^{\theta}}{1-\gamma^{\theta}}\left[1-\frac{k^{\alpha\theta}(\gamma(k+1)^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta}}{(k+1)^{\alpha\theta}(\gamma(k^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta}}\right]. \tag{13}$$

Figure 2 represents hazard plots of DNGP for various values of α,β,γ and θ. From the plots, it is clear that hazard rate can be decreasing and increasing followed by decreasing shapes.

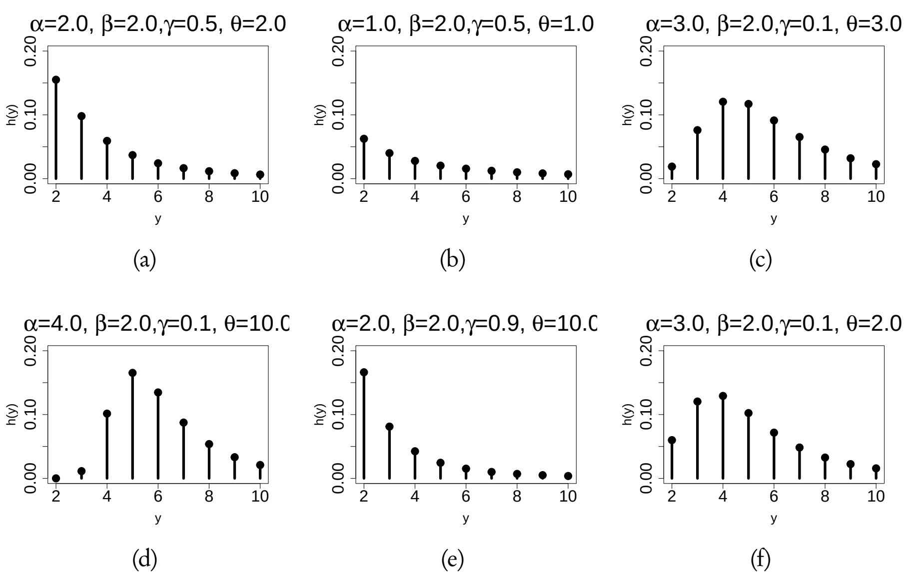

Figure 2 – Hazard rate of DNGP$(\alpha,\beta,\gamma,\theta)$ for different values of $\alpha,\beta,\gamma,\theta$.

4.3. Moments

The r^{th} moment about origin is given by

$$\mu_r' = \sum_{y=[\beta]}^{\infty} y^r \frac{\gamma^\theta}{1-\gamma^\theta} \left[\frac{(y+1)^{\alpha\theta}}{(\gamma(y+1)^\alpha + (1-\gamma)\beta^\alpha)^\theta} - \frac{y^{\alpha\theta}}{(\gamma y^\alpha + (1-\gamma)\beta^\alpha)^\theta} \right]. \tag{14}$$

The moments are numerically obtained by using R programming for given values of $\alpha, \beta, \gamma, \theta$. The index of dispersion of DNGP for different values of the parameters are given in Table 1. From Table 1, it can be seen that for two sets of values for beta and gamma (β =0.5 and 1.0, γ = 0.1) and for fixed θ, the index of dispersion $(E(X^2) - [E(X)]^2)/E(X)$ decreases as α increases. Also for fixed α, the index of dispersion increases as well as decreases as θ increases. Depending upon the parametric values, it can be found that index of dispersion can be greater than 1 as well as less than 1. Therefore DNGP models are better for both under and over dispersed data.

TABLE 1
Index of dispersion for DNGP$(\alpha, \beta, \gamma, \theta)$ for different values of α and θ for $\beta = 0.5, \gamma = 0.1$ and $\beta = 1.0, \gamma = 0.1$.

	$\alpha\downarrow$ $\theta\rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
	0.5	4.200	4.294	4.391	4.491	4.593	4.697	4.803	4.909	5.015	5.121
β - 0.5	1.0	3.164	3.183	3.204	3.228	3.254	3.284	3.318	3.355	3.396	3.442
γ - 0.1	2.0	2.058	2.046	2.032	2.017	2.000	1.983	1.964	1.945	1.925	1.906
	3.0	1.293	1.276	1.258	1.240	1.221	1.202	1.182	1.163	1.144	1.125
	5.0	0.975	0.963	0.951	0.938	0.925	0.912	0.898	0.883	0.869	0.854
	0.5	4.173	4.309	4.446	4.582	4.717	4.851	4.983	5.111	5.237	5.359
β - 1.0	1.0	2.900	3.004	3.112	3.223	3.338	3.455	3.576	3.700	3.825	3.953
γ - 0.1	2.0	1.571	1.609	1.641	1.670	1.695	1.717	1.736	1.752	1.766	1.779
	3.0	0.797	0.831	0.862	0.890	0.915	0.937	0.956	0.973	0.987	0.998
	5.0	0.168	0.181	0.195	0.208	0.222	0.235	0.247	0.259	0.271	0.282

In Table 2, we present raw moments about origin, central moments, skewness and kurtosis based on moments of DNGP$(\alpha, \beta, \gamma, \theta)$ for given values of $\alpha, \beta, \gamma, \theta$. From Table 2, it is clear that DNGP is positively skewed since the sign of μ_3 is positive for all the parametric values and DNGP is leptokurtic since the measure of kurtosis is greater than 3 for all given values of the parameters.

TABLE 2
Moments, Skewness and Kurtosis for $\beta = 0.9$, $\gamma = 0.9$ and various values of α and θ.

Parameter	Raw moments	Central Moments	Skewness	Kurtosis
$\alpha = 0.5$	$\mu_1{}' = 1.908$			
$\theta = 0.5$	$\mu_2{}' = 9.416$	$\mu_2 = 5.776$		
	$\mu_3{}' = 62.221$	$\mu_3 = 22.218$	2.559	4.932
	$\mu_4{}' = 473.521$	$\mu_4 = 164.623$		
$\alpha = 2.0$	$\mu_1{}' = 1.227$			
$\theta = 0.5$	$\mu_2{}' = 2.980$	$\mu_2 = 1.475$		
	$\mu_3{}' = 12.596$	$\mu_3 = 5.320$	8.814	15.960
	$\mu_4{}' = 76.432$	$\mu_4 = 34.740$		
$\alpha = 3.0$	$\mu_1{}' = 0.902$			
$\theta = 1.0$	$\mu_2{}' = 1.415$	$\mu_2 = 0.602$		
	$\mu_3{}' = 3.642$	$\mu_3 = 1.280$	7.524	21.616
	$\mu_4{}' = 16.043$	$\mu_4 = 7.824$		
$\alpha = 5.0$	$\mu_1{}' = 0.653$			
$\theta = 2.0$	$\mu_2{}' = 0.715$	$\mu_2 = 0.288$		
	$\mu_3{}' = 0.888$	$\mu_3 = 0.044$	0.081	5.413
	$\mu_4{}' = 1.486$	$\mu_4 = 0.450$		
$\alpha = 2.0$	$\mu_1{}' = 1.345$			
$\theta = 5.0$	$\mu_2{}' = 3.495$	$\mu_2 = 1.685$		
	$\mu_3{}' = 15.379$	$\mu_3 = 6.146$	7.890	14.086
	$\mu_4{}' = 94.639$	$\mu_4 = 40.007$		

4.4. Quantile function and random number generation

The quantile function of the random variable Y having DNGP$(\alpha,\beta,\gamma,\theta)$ distribution is,

$$\phi(m)=F^{-1}(m)=[y_m]=\beta\left[\frac{1-\gamma}{\gamma}\frac{1}{((m-1)(1-\gamma^{\theta})+1)^{-1/\theta}-1}\right]^{1/\alpha}-1,\ \ 0\leq m\leq 1, \tag{15}$$

where $[y_m]$ is the smallest integer greater than or equal to y_m. In particular, the median is given by

$$\text{Median}=\phi(0.5)=\beta\left[\frac{1-\gamma}{\gamma}\frac{1}{\left[1-\left(\frac{1-\gamma^{\theta}}{2}\right)^{-1/\theta}\right]-1}\right]^{1/\alpha}-1,\ \ 0\leq m\leq 1. \tag{16}$$

Here we use usual inverse transformation method for generating samples from proposed model. Let M be a random number taken from $U(0,1)$. Then, a random number Y, following DNGP, is sampled using the Eq.(15).

4.5. Infinite Divisibility

The structural property of infinite divisibility of the distribution is a characteristic that has close relation to the Central Limit Theorem and waiting time distributions. In modeling, it is desirable to know whether a given distribution is infinitely divisible or not. Therefore, according to Steutel and van Harn (2003), Proposition 9.2, page 56, if $p_x, x\in\mathbb{N}_0$ is infinitely divisible, then $p_x\leq e^{-1}$ for all $x\in\mathbb{N}$. Also from Theoreom 3.2 of Steutel and van Harn (2003), if for atleast one case for which p_x is greater than $1/e$, then pmf cannot be compound Poisson and hence it cannot be infinitely divisible. Here, DNGP$(2,2,0.9,5)$ distribution, we can see $p_2=0.4770173>e^{-1}=0.367$. So, in general, DNGP$(\alpha,\beta,\gamma,\theta)$ distributions are not infinitely divisible. In addition, since the class of self decomposable and stable distributions, in their discrete concept, are sub-class of infinitely divisible distributions, we can conclude that DNGP distribution can be neither self decomposable nor stable, in general.

4.6. Stress-strength reliability

Stress-strength reliability analysis is widely used in reliability engineering. Suppose Y denotes the strength of a component subject to a component Z. Then $R=P(Y>Z)$ is a measure of system performance called stress-strength parameter. In literature, the estimation of R has been considered when Y and Z are independently and identically distributed. The stress-strength model is defined in discrete case by

$$P(Y>Z)=\sum_{y=0}^{\infty}P_Y(y)F_Z(y),$$

where $P_Y(y)$ and $F_Z(y)$ denotes pmf and cdf of the discrete random variable Y and Z respectively.

Let $Y \sim \text{DNGP}(\alpha_1,\beta_1,\gamma_1,\theta_1)$ and $Z \sim \text{DNGP}(\alpha_2,\beta_2,\gamma_2,\theta_2)$. Denote $\lambda_1=(\alpha_1,\beta_1,\gamma_1,\theta_1)^T$ and $\lambda_2=(\alpha_2,\beta_2,\gamma_2,\theta_2)^T$. Let $(y_1,y_2,\ldots,y_n)$ and $(z_1,z_2,\ldots,z_m)$ be n and m independent observations taken from DNGP(λ_1) and DNGP(λ_2), respectively. Then,

$$\begin{aligned}R=P(Y>Z)=\sum_{y=[\beta]}^{\infty}\frac{\gamma_1^{\theta_1}\gamma_2^{\theta_2}}{(1-\gamma_1^{\theta_1})(1-\gamma_2^{\theta_2})}\Bigg[&\frac{(y+1)^{\alpha_1\theta_1}}{(\gamma_1(y+1)^{\alpha_1}+(1-\gamma_1)\beta_1^{\alpha_1})^{\theta_1}}+\\&-\frac{y^{\alpha_1\theta_1}}{(\gamma_1 y^{\alpha_1}+(1-\gamma_1)\beta_1^{\alpha_1})^{\theta_1}}\Bigg]\Bigg[\frac{(y+1)^{\alpha_2\theta_2}}{(\gamma_2(y+1)^{\alpha_2}+(1-\gamma_2)\beta_2^{\alpha_2})^{\theta_2}}-1\Bigg].\end{aligned}\tag{17}$$

The total likelihood function of R is given by $L_R(\lambda^*)=L_n(\lambda_1)L_m(\lambda_2)$, where $\lambda^*=(\lambda_1,\lambda_2)$. The score vector of R is given by

$$T_R(\lambda^*)=\left(\frac{\partial L_R}{\partial\alpha_1},\frac{\partial L_R}{\partial\beta_1},\frac{\partial L_R}{\partial\gamma_1},\frac{\partial L_R}{\partial\theta_1},\frac{\partial L_R}{\partial\alpha_2},\frac{\partial L_R}{\partial\beta_2},\frac{\partial L_R}{\partial\gamma_2},\frac{\partial L_R}{\partial\theta_2}\right).$$

The maximum likelihood estimate (mle) $\hat{\lambda}$ of λ^* can be obtained by solving $T_R(\lambda^*)=0$. Then, we substitute the mle $\hat{\lambda}$ in Eq.(17). Thus, we can obtain the stress-strength parameter R.

5. Estimation and Simulation Study

Here, we consider the method of maximum likelihood for estimating the parameters of DNGP. Consider a random sample $(y_1,y_2,\ldots,y_n)$ of size n, drawn from DNGP$(\alpha,\beta,\gamma,\theta)$ with unknown parameter vector $\lambda=(\alpha,\beta,\gamma,\theta)^T$. Then, the likelihood function is given by

$$L(y;\alpha,\beta,\gamma,\theta)=\frac{\gamma^{\theta n}}{(1-\gamma^{\theta})^n}\prod_{j=1}^{n}\left[\frac{(y_j+1)^{\alpha\theta}}{(\gamma(y_j+1)^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta}}-\frac{y_j^{\alpha\theta}}{(\gamma y_j^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta}}\right]\tag{18}$$

and log-likelihood becomes

$$\begin{aligned}\log L(y;\alpha,\beta,\gamma,\theta)=&n\theta\log\gamma-n\log(1-\gamma^{\theta})+\sum_{j=1}^{n}\log\Big[(y_j+1)^{\alpha\theta}(\gamma(y_j+1)^{\alpha}+\\&+(1-\gamma)\beta^{\alpha})^{-\theta}-y_j^{\alpha\theta}(\gamma y_j^{\alpha}+(1-\gamma)\beta^{\alpha})^{-\theta}\Big].\end{aligned}\tag{19}$$

The corresponding likelihood equations are (by taking partial derivatives of LogL w.r.t α, β, γ and θ respectively)

$$\frac{\partial \log L}{\partial \alpha} = \sum_{j=1}^{n} \theta \Bigg\{ \left[\frac{\log(1+y_j)(1+y_j)^{\alpha\theta} A_2^{-\theta} - \log(y_j) y_j^{\alpha\theta} A_1^{-\theta} + (1+y_j)^{\alpha\theta} A_2^{-1-\theta}(-\beta^{\alpha}\theta \log(\beta)\beta^{\alpha}\gamma \log(\beta) - \gamma \log(1+y_j)(1+y_j)^{\alpha})}{(1+y_j)^{\alpha\theta} A_2^{-\theta} - y_j^{\alpha\theta} A_1^{-\theta})} \right] - \left[\frac{y_j^{\alpha\theta} A_1^{-1-\theta}(-\beta^{\alpha}\theta \log(\beta)\beta^{\alpha}\gamma \log(\beta) - \gamma \log(y_j) y_j^{\alpha})}{(1+y_j)^{\alpha\theta} A_2^{-\theta} - y_j^{\alpha\theta} A_1^{-\theta})} \right] \Bigg\}, \tag{20}$$

$$\frac{\partial \log L}{\partial \beta} = \sum_{j=1}^{n} \frac{\alpha \beta^{\alpha-1}\theta(\gamma-1)((1+y_j)^{\alpha\theta} A_2^{-1-\theta} - y_j^{\alpha\theta} A_1^{-1-\theta})}{(1+y_j)^{\alpha\theta} A_2^{-\theta} - y_j^{\alpha\theta} A_1^{-\theta})}, \tag{21}$$

$$\frac{\partial \log L}{\partial \gamma} = \frac{n\theta}{\gamma} + \frac{n\theta\gamma^{\theta-1}}{1-\gamma^{\theta}} + \sum_{j=1}^{n} \frac{\left(1+y_j\right)^{\alpha\theta}\left(\beta\alpha\theta A_2^{-1-\theta} - \theta(1+y_j)^{\alpha} A_2^{-1-\theta}\right) - y_j^{\alpha\theta}\left(\beta\alpha\theta A_1^{-1-\theta} - \theta y_j^{\alpha} A_1^{-1-\theta}\right)}{(1+y_j)^{\alpha\theta} A_2^{-\theta} - y_j^{\alpha\theta} A_1^{-\theta})}, \tag{22}$$

$$\frac{\partial \log L}{\partial \theta} = n \log \gamma + \frac{n\gamma^{\theta}\log\gamma}{1-\gamma^{\theta}} + \sum_{j=1}^{n} \frac{A_1^{-\theta} y_j^{\alpha\theta}\left(\log A_1 - \alpha \log(y_j)\right) - A_2^{-\theta}(1+y)_j^{\alpha\theta}\left(\log A_2 - \alpha \log(1+y_j)\right)}{(1+y_j)^{\alpha\theta} A_2^{-\theta} - y_j^{\alpha\theta} A_1^{-\theta}}, \tag{23}$$

where $A_1 = \beta^{\alpha}(1-\gamma) + \gamma y_j^{\alpha}$ and $A_2 = \beta^{\alpha}(1-\gamma) + \gamma(1+y_j)^{\alpha}$. The mle(s) of α, β, γ and θ can be obtained by setting Eq. (20) to Eq. (23) equal to zero and solving simultaneously with the help of statistical packages like `optim` or `nlm` in `R` programming.

Here, we perform simulation studies to find out the performance of mle of α, β, γ, and θ of the proposed DNGP model for different sample sizes. The different sample sizes taken here are $n = 100, 250, 500$ and 1000. We repeat the process 2000 times and report the average estimates of the parameters using `nlm` function of `stats` package in `R` programming (R Core Team (2013)). In Table 3, we present the mle of α, β, γ, and θ of DNGP$(\alpha, \beta, \gamma, \theta)$ distribution together with their mean square errors for different values of n. From Table 3, it is clear that as sample size increases, the estimates are close to the parameter values and mean squared errors of the estimates are also decreases.

TABLE 3
mle of α, β , γ and θ.

Parameter	n	$\hat{\alpha}$ (MSE($\hat{\alpha}$))	$\hat{\beta}$ (MSE($\hat{\beta}$))	$\hat{\gamma}$ (MSE($\hat{\gamma}$))	$\hat{\theta}$ (MSE($\hat{\theta}$))
$\alpha = 1.0$	100	0.91(0.424)	1.10(0.522)	0.20(0.508)	0.99(0.003)
$\beta = 1.0$	250	0.98(0.034)	1.07(0.215)	0.17(0.250)	0.99(0.003)
$\gamma = 0.1$	500	1.03(0.029)	1.02(0.031)	0.16(0.184)	0.99(0.000)
$\theta = 1.0$	1000	1.02(0.024)	1.03(0.025)	0.15(0.133)	1.02(0.000)
α =2.0	100	2.10(0.473)	1.19(1.771)	0.25(0.140)	3.20(2.043)
$\beta = 1.0$	250	2.09(0.733)	1.18(1.567)	0.243(0.092)	3.18(1.603)
$\gamma = 0.2$	500	2.03(0.043)	1.13(0.064)	0.233(0.054)	3.09(0.404)
θ =3.0	1000	2.01(0.000)	1.01(0.008)	0.203(0.000)	3.09(0.386)
$\alpha = 2.0$	100	2.02(0.025)	3.12(1.317)	0.590(0.472)	1.02(0.046)
$\beta = 3.0$	250	2.02(0.023)	3.11(1.132)	0.493(0.257)	1.02(0.042)
$\gamma = 0.5$	500	2.01(0.019)	3.10(1.321)	0.543(0.257)	1.02(0.055)
$\theta = 1.0$	1000	2.02(0.012)	3.01(0.658)	0.508(0.072)	1.02(0.022)
$\alpha = 2.0$	100	2.03(0.030)	2.70(2.073)	0.594(0.418)	2.08(0.344)
$\beta = 2.5$	250	2.04(0.099)	2.59(0.358)	0.574(0.274)	1.94(0.195)
$\gamma = 0.5$	500	1.92(0.343)	2.46(0.083)	0.554(0.146)	2.06(0.175)
$\theta = 2.0$	1000	2.02(0.015)	2.46(0.081)	0.514(0.007)	2.05(0.175)

6. Application

In this Section, we consider a real data set for illustrating the method we discussed in Section 5. The data consist of the 72 exceedances for the years 1958-1984 (rounded to one decimal place) of flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada (Choulakian and Stephens, 2001). The data are reported in Table 4.

TABLE 4
Data on the 72 exceedances for the years 1958-1984 of flood peaks (in m^3/s) of the Wheaton River.

1.7	1.4	0.6	9.0	5.6	1.5	2.2	18.7	2.2	1.7	30.8	2.5
14.4	8.5	39.0	7.0	13.3	27.4	1.1	25.0	0.32	0.1	4.2	1.0
0.4	11.6	15.0	0.4	25.0	27.1	20.6	14.1	11.0	2.8	3.4	20.2
5.3	22.1	7.3	14.1	11.9	16.8	0.7	1.1	22.9	9.9	21.5	5.3
1.9	2.5	1.7	10.4	27.6	9.7	13.0	14.4	0.1	10.7	36.4	27.5
12.0	1.7	1.1	30.0	2.7	2.5	9.3	37.6	0.6	3.6	64.0	27.0

Since the data set is continuous, here first we discretize the data. The parameters are estimated by using the method of maximum likelihood (using R software). We compare the fit of the DNGP distribution with the following discrete lifetime distributions:

a Geometric (G) distribution having pmf

$$P(Y=y)=(1-p)^{y}p\ ;0<p<1,\ y=0,1,2,\ldots.$$

b Discrete Pareto (DP) (see Krishna and Pundir, 2009) distribution having pmf

$$P(Y=y)=\theta^{\log(1+y)}-\theta^{\log(2+y)};\ 0<\theta<1,\ y=0,1,2\ldots.$$

c Discrete Generalized Pareto (DGP) (see Prieto *et al.*, 2014) distribution having pmf

$$P(Y=y)=(1+\lambda(y-\mu))^{-\alpha}-(1+\lambda(y-\mu+1))^{-\alpha};\lambda>0,\mu>0,\alpha>0,\\ y=\mu,\mu+1,\ldots.$$

d Exponentiated Discrete Weibull (EDW) distribution (see Nekoukhou and Bidram, 2015) having pmf

$$P(Y=y)=(1-p^{(y+1)^{\alpha}})^{\gamma}-(1-p^{y^{\alpha}})^{\gamma};\quad \alpha>0,\quad \gamma>0,\ \ 0<p<1,\\ y=0,1,2,\ldots.$$

e Generalized Discrete Weibull (GDW) distribution (see Jayakumar and Sankaran, 2018) having pmf

$$P(y=y)=\frac{\theta\left[(1-p^{(y+1)^{\beta}})^{\alpha}-(1-p^{y^{\beta}})^{\alpha}\right]}{\left\{1-\bar{\theta}[1-(1-p^{y^{\beta}})^{\alpha}]\right\}\left\{1-\bar{\theta}[1-(1-p^{(y+1)^{\beta}})^{\alpha}]\right\}}\ ;\quad 0<p<1,\quad \alpha>0,\beta>0,\theta>0,\\ y=0,1,2,\ldots.$$

f Discrete Burr XII (DBD XII) distribution (see Para and Jan, 2016) having pmf

$$P(Y=y)=\beta^{\log(1+(\frac{y}{\gamma})^{c})}-\beta^{\log(1+(\frac{y+1}{\gamma})^{c})}\ ;0<\beta<1,\ \gamma>0,\ c>0,\ \text{y=0,1,2,}\ldots.$$

g Discrete Burr (DB) (see Krishna and Pundir, 2009) distribution having pmf

$$P(Y=y)=\theta^{\log(1+y^{\alpha})}-\theta^{\log(1+(1+y^{\alpha}))}\ ;0<\theta<1,\ \alpha>0\ ,\ y=0,1,2,\ldots.$$

h Discrete Pareto type IV (DP IV) (see Ghosh, 2020) distribution having pmf

$$P(Y=y)=\theta^{\log(1+(\frac{y}{\sigma})^{(1/\gamma)})}-\theta^{\log(1+(\frac{(y+1)}{\sigma})^{(1/\gamma)})};0<\theta<1,\sigma>0,\gamma>0,\\ y=0,1,2,3,\ldots.$$

TABLE 5
Parameter estimates and goodness of fit for various models fitted for the dataset.

Model	MLEs(SE)	-LogL	AIC	BIC	AICc	HQIC	K-S	p-value
DP	$\hat{p} = 0.643(0.03)$	294.06	590.12	592.39	590.17	591.02	0.38	0.00
DB	$\hat{\alpha} = 170.666(0.00)$	270.24	544.47	549.02	544.64	546.28	0.30	0.00
	$\hat{\theta} = 0.997(0.00)$							
G	$\hat{p} = 0.073(0.01)$	257.50	517.00	519.28	517.06	517.91	0.14	0.11
DGP	$\hat{\lambda} = 4.459(1.22)$	257.14	520.27	527.10	520.68	522.99	0.38	0.00
	$\hat{\mu} = 1.218(1.44)$							
	$\hat{\alpha} = 0.304(3.06)$							
DP IV	$\hat{\theta} = 1.109 \times 10^{-5}(0.00)$	257.25	520.48	527.31	520.84	523.20	0.15	0.09
	$\hat{\sigma} = 108.728(5.95)$							
	$\hat{\gamma} = 0.884(0.08)$							
DB XII	$\hat{\beta} = 6.106 \times 10^{-5}(0.00)$	257.27	520.53	527.36	520.89	523.25	0.11	0.30
	$\hat{\gamma} = 92.001(0.23)$							
	$\hat{c} = 1.143(0.07)$							
EDW	$\hat{p} = 0.695(0.39)$	256.65	519.30	526.13	519.65	522.01	0.11	0.35
	$\hat{\alpha} = 0.648(0.36)$							
	$\hat{\gamma} = 2.73(3.29)$							
GDW	$\hat{\alpha} = 3.381(1.15)$	256.64	521.29	530.39	521.88	524.91	0.11	0.35
	$\hat{\beta} = 0.588(0.78)$							
	$\hat{\theta} = 1.160(0.02)$							
	$\hat{p} = 0.605(1.93)$							
DNGP	$\hat{\alpha} = 2.028(0.25)$	242.33	492.65	501.76	493.25	496.28	0.09	0.54
	$\hat{\beta} = 1.808(0.03)$							
	$\hat{\gamma} = 0.002(0.00)$							
	$\hat{\theta} = 6.627 \times 10^{-6}(0.09)$							

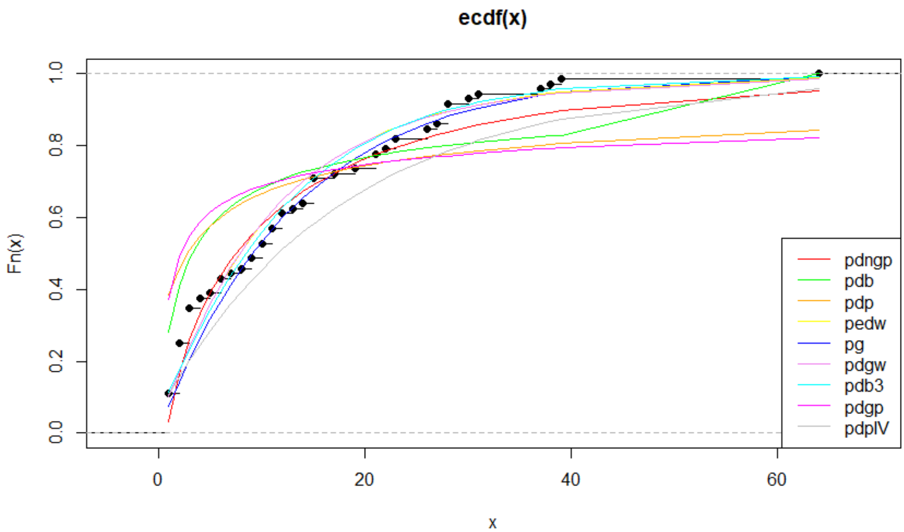

Figure 3 – Empirical and fitted cdfs for the dataset.

The values of estimates, standard error (SE), the log-likelihood function (-LogL), the Kolmogrov-Smirnov (K-S) statistic, Akaike Information Criterion (AIC), Akaike Information Criterion with correction (AICc), Bayesian Information Criterion (BIC) and Hannon-Quinn Information Criterion (HQIC) are calculated for the nine distributions in order to verify which distribution fits better to the data and presented in Table 5. The better distribution corresponds to smaller -LogL, K-S, AIC, AICc, BIC, and HQIC values and larger p value. Here, AIC = -2LogL + $2k$, AICc = -2LogL + $(\frac{2kn}{n-k-1})$, BIC = -2LogL + $k\log(n)$, HQIC = -2LogL + $2k\log(\log(n))$, where L is the likelihood function evaluated at the mle(s), k is the number of parameters, and n is the sample size. The K-S distance, $D_n = \sup_y |F(y) - F_n(y)|$, where $F_n(y)$ is the empirical distribution function. From Table 5, we can see that -LogL, K-S, AIC, AICc, BIC, and HQIC values are smallest for DNGP with -LogL = 242.33, AIC = 492.65, AICc = 493.25, BIC = 501.76, HQIC = 496.28 and K-S statistic value is 0.09. Also DNGP has highest $p-$ value (i.e., 0.54). Therefore, the DNGP distribution gives better fit to the data compared to the other eight models.

Figure 3 shows the plot of the cdf of nine models in comparison with empirical distribution function of the given data. This figure indicates DNGP is superior to other models in terms of model fitting.

7. Conclusion

In this paper, we have introduced and studied Discrete New Generalized Pareto distribution (DNGP) as a discrete analogue of New Generalized Pareto (NGP) and derived different structural properties of DNGP. Estimation of the parameters of DNGP distribution are carried out using the method of maximum likelihood. The proposed distribution DNGP has wide range of applications in various fields and is a competitor of existing discrete Pareto models. Analysis of a real data set is presented to show that the flexibility and application of the distribution.

Acknowledgements

Authors wish to thank the Editor in Chief and the two reviewers for their constructive suggestions, which greatly improved the presentation of the paper.

References

B. C. Arnold (2008). *Pareto and generalized Pareto distributions*. In D. Chotikapanich (ed.), *Modeling Income Distributions and Lorenz Curves*, Springer, New York, pp. 119–145.

A. Buddana, T. J. Kozubowski (2014). *Discrete Pareto distributions*. Economic Quality Control, 29, no. 2, pp. 143–156.

S. Chakraborty (2015). *Generating discrete analogues of continuous probability distributions-a survey of methods and constructions*. Journal of Statistical Distributions and Applications, 2, no. 1, pp. 1–30.

S. Chakraborty, D. Chakravarty (2012). *Discrete gamma distributions: properties and parameter estimation*. Communications in Statistics-Theory and Methods, 41, no. 18, pp. 3301–3324.

S. Chakraborty, D. Chakravarty (2014). *A discrete Gumbel distribution*. arXiv preprint arXiv:1410.7568.

S. Chakraborty, D. Chakravarty (2015). *A discrete power distribution*. arXiv preprint arXiv:1501.06299.

V. Choulakian, M. A. Stephens (2001). *Goodness-of-fit tests for the generalized Pareto distribution*. Technometrics, 43, no. 4, pp. 478–484.

I. Ghosh (2020). *A new discrete Pareto type (IV) model: theory, properties and applications*. Journal of Statistical Distributions and Applications, 7, no. 1, pp. 1–17.

S. Inusah, T. J. Kozubowski (2006). *A discrete analogue of the Laplace distribution*. Journal of Statistical Planning and Inference, 136, no. 3, pp. 1090–1102.

K. Jayakumar, B. Krishnan, G. Hamedani (2020). *On a new generalization of Pareto distribution and its applications*. Communications in Statistics-Simulation and Computation, 49, no. 5, pp. 1264–1284.

K. Jayakumar, K. K. Sankaran (2016). *On a generalisation of uniform distribution and its properties*. Statistica, 76, no. 1, pp. 83–91.

K. Jayakumar, K. K. Sankaran (2018). *A generalization of discrete Weibull distribution*. Communications in Statistics-Theory and Methods, 47, no. 24, pp. 6064–6078.

A. W. Kemp (1997). *Characterizations of a discrete normal distribution*. Journal of Statistical Planning and Inference, 63, no. 2, pp. 223–229.

A. W. Kemp (2008). *The discrete half-normal distribution*. In *Advances in Mathematical and Statistical Modeling*, Springer, Birkhäuser Boston., Boston, pp. 353–360.

C. Klüppelberg (1988). *Subexponential distributions and integrated tails*. Journal of Applied Probability, 25, no. 1, pp. 132–141.

T. J. Kozubowski, S. Inusah (2006). *A skew Laplace distribution on integers*. Annals of the Institute of Statistical Mathematics, 58, no. 3, pp. 555–571.

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

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

V. Nekoukhou, H. Bidram (2015). *The exponentiated discrete Weibull distribution*. Sort, 39, pp. 127–146.

W. Padgett, J. D. Spurrier (1985). *On discrete failure models*. IEEE Transactions on Reliability, 34, no. 3, pp. 253–256.

B. Para, T. Jan (2016). *On discrete three parameter Burr type XII and discrete Lomax distributions and their applications to model count data from medical science*. Biometrics & Biostatistics International Journal, 4, no. 2, pp. 1–15.

F. Prieto, E. Gómez-Déniz, J. M. Sarabia (2014). *Modelling road accident blackspots data with the discrete generalized Pareto distribution*. Accident Analysis & Prevention, 71, pp. 38–49.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. ISBN 3-900051-07-0.

D. Roy (2003). *The discrete normal distribution*. Communications in Statistics-Theory and Methods, 32, no. 10, pp. 1871–1883.

H. SATO, M. IKOTA, A. SUGIMOTO, H. MASUDA (1999). *A new defect distribution metrology with a consistent discrete exponential formula and its applications*. IEEE Transactions on Semiconductor Manufacturing, 12, no. 4, pp. 409–418.

W. E. STEIN, R. DATTERO (1984). *A new discrete Weibull distribution*. IEEE Transactions on Reliability, 33, no. 2, pp. 196–197.

F. W. STEUTEL, K. VAN HARN (2003). *Infinite Divisibility of Probability Distributions on the Real Line*. CRC Press, Boca Raton.