

**DISCRETE GENERALIZED WEIBULL DISTRIBUTION:
PROPERTIES AND APPLICATIONS IN MEDICAL SCIENCES**

B.A. Para and T.R. Jan

Department of statistics, University of Kashmir, Srinagar, J&K (India)

Email: parabilal@gmail.com

drtrjan@gmail.com

ABSTRACT

In this paper, we introduce a discrete analogue of generalized Weibull distribution (DGWD) as a new discrete model. Initially we study some fundamental distributional properties of this new discrete model and discuss unimodality and failure rate functions. Finally, we discuss the application of the model with a data set studied by Morgan et al. (2007) and McElduff (2012).

KEYWORDS

Simulation, Hazard Rate, Index of Dispersion.

1. INTRODUCTION

Researchers obtain new probability models by using different techniques such as compounding, discretization, transmutation etc. in order to illustrate a phenomenon in the form of mathematical expressions. Recently discretization of continuous probability distributions has also received attention from researchers from past decade. Discretized Statistical models form a basic field of study to handle discrete lifetime data and also count data in a wide variety of disciplines such as biological and medical sciences, physical sciences, engineering, agriculture and so on. As plethora of the lifetimes in reliability are of continuous in nature and therefore many continuous lifetime models have been discussed in statistical modelling literature [See for example Kapur and Lamberson (1997), Lawless (1982) and Sinha (1986)]. However, there are many situations when it is impractical in lifetime experiments to quantify the life span of a component or a device on a continuous scale. There are many components which operate in cycles and the experimenter counts number of completed cycles before failure. We can quote a well-known example of copier whose lifetime is simply the total number of copies it produces. Another referred example is the life length of an off/on switching component or machine whose life length depends on the number of times we switch the device on/off, is a random variable with discrete nature, or lifetime of a component which receives a number of shocks prior to failure. In reliability theory, we sometimes record the number of hours/days a cancer patient survives since the time of therapy, or the times from remission to relapse are also most of the times recorded in number of hours/days. In the current era discrete distributions play special role in the field of reliability theory and other applied fields. In this background, the well-known distributions like GD (Geometric distribution) and NBD (Negative Binomial Distribution) are the well-known discrete versions for the ED (Exponential distribution) and GD (Gamma distribution),

respectively. In the recent past, researchers discretized many continuous distributions for modeling count data and discrete lifetime data from various applied fields. As mentioned before only, the discrete parallels of gamma and exponential distributions are NBD and GD. Standard continuous Weibull model [Khan et al. (1989), 33_34] was discretized by many authors with different functionality. Roy (2003 and 2004) studied discrete analogue of the RD (Rayleigh distribution) and ND (Normal distribution) as a new discrete alternatives for RD (Rayleigh distribution) and ND (Normal distribution). Krishna and Pundir (2009) proposed Discrete versions of two parameter BXII (Burr type XII) and PD (Pareto distribution). Recently two parameter DIWD (discrete inverse Weibull distribution) was proposed Jazi et al. (2010), along with the applications in series systems in reliability. Para and Jan (2014) discussed a two parameter DBIII (discrete Burr type III distribution) as a reliability model to fit a various range of discrete lifetime data. DLD (discrete Lindley distribution) was discussed by Gómez-Déniz and Calderín-Ojeda (2011) by discretizing the continuous failure model of the Lindley distribution. Nekoukhou et al. (2012) studied a discrete version of the GED (generalized exponential distribution), which is another generalization of the GD (geometric distribution), and some of its distributional and structural properties were discussed. Para and Jan (2016a) recently introduced a discrete version of three parameter Burr type XII and Lomax distribution as a new discrete distributions to model counts of cysts of kidneys using steroids. Para and Jan (2016b) also introduced a discrete version of Loglogistic distribution to model a data from medical genetics.

We propose a discrete generalized Weibull distribution (DGWD) as there is a need to find more accurate discrete survival models or probability models in biological sciences, medical science and other applied fields, to fit to various discrete count data sets. It is well known in general that a generalized model is having more flexibility than the base model and it is favored by data analysts in analyzing data.

There are many generalized versions of Weibull distribution. Gusmão et al. (2011) studied a generalized inverse Weibull distribution by adding another parameter γ to the standard inverse Weibull distribution. Inverting this model gives another generalized version of Weibull distribution as given by

$$f(x) = \begin{cases} \gamma\alpha\beta^{-\alpha}x^{\alpha-1}e^{-\gamma(x/\beta)^\alpha} & x > 0; \alpha > 0; \beta > 0; \gamma > 0 \\ 0 & \text{elsewhere} \end{cases}$$

2. DISCRETE GENERALIZED WEIBULL MODEL

We discuss two methods of discretizing a continuous random variable as

(i) Method First:

For a continuous random variable X on R with *pdf* $f(x)$, a discrete random variable Y can be defined that has integer support on $(-\infty, \infty)$ as:

$$p(Y = k) = f(k) / \sum_{j=-\infty}^{\infty} f(j); \quad k = 0, \pm 1, \pm 2, \pm 3, \dots$$

This method was used by Kemp (1997) to obtain a discrete version of the normal distribution [See also Inusha and Kozubowski (2006)].

(ii) Method Second:

Roy (1993) introduced a discretization method using reliability function of the model concerned.

$$p(X = x) = s(x) - s(x + 1) \quad \text{when } x = 0, 1, 2, \dots$$

Roy (1993) applied this method for discretizing ED (Exponential distribution), $s(x)$ being the survival function of the exponential random variate.

A DGW (discrete generalized Weibull) variable, dX can be observed as the discrete concentration of the continuous generalized Weibull variable X using the second method of discretization.

$$p(dX = x) = s(x) - s(x + 1)$$

The probability mass function of the proposed model takes the form

$$p(x) = \theta^{(x/\beta)^\alpha} - \theta^{((x+1)/\beta)^\alpha} \quad x = 0, 1, 2, \dots \tag{2.1}$$

where $\theta = \exp(-\gamma)$, $0 < \theta < 1$, $\alpha > 0$, $\beta > 0$, $\gamma > 0$, and the cumulative distribution function is given by

$$F(x) = 1 - \theta^{((x+1)/\beta)^\alpha} \quad \text{where } \theta = \exp(-\gamma), 0 < \theta < 1, \alpha > 0, \beta > 0, \gamma > 0. \tag{2.2}$$

The quantile function for three parameter discrete generalized Weibull distribution is obtained by inverting (2.2)

$$X_\phi = \left\lceil \beta \left(\frac{\log(1-\phi)}{\log \theta} \right)^{1/\alpha} - 1 \right\rceil, \text{ where } \theta = \exp(-\gamma), 0 < \theta < 1, \alpha > 0, \beta > 0,$$

where $\lceil . \rceil$ denotes the GIF (floor function/greatest integer function) (the largest integer that is less than or equal). In particular, median is

$$X_{0.5} = \left\lceil \beta \left(\frac{\log(2)}{\log \theta} \right)^{1/\alpha} - 1 \right\rceil.$$

where $\lfloor . \rfloor$ denotes the ceiling function (the smallest integer that is greater than or equal).

The value of α has an important role in defining the shape of the *pmf*. As the value of θ increases the skewness decreases. The *pmf* (2.1) at $x = 0$ and $\beta = 1$ is $p(x) = 1 - \theta$.

In this section, we discuss that DGWD (discrete generalized Weibull distribution) can be nested to different distributions under the specific parameter settings.

- (i) For $\alpha = 1$ and $\beta = 1$, discrete generalized Weibull distribution becomes Geometric distribution.
- (ii) For $\alpha = 2$ and $\beta = 1$, discrete generalized Weibull distribution becomes discrete Rayleigh distribution.

2.1 Random Data Generation from Discrete Generalized Weibull Distribution

For the purpose of simulating a sequence of random numbers t_1, t_2, \dots, t_n of the discrete generalized Weibull random variable X with pmf $p(X = t_i) = p_i, \sum_{i=0}^k p_i = 1$ and

$$\text{a cdf } F(X = t_i) = \sum_{i=0}^{t_i} p_i, \quad i = 0, 1, 2, \dots$$

Step 1: Generating random number m using uniform distribution $U(0, 1)$.

Step 2: Generate random number t_i based on

$$\begin{aligned} &\text{if } m \leq p_0 = F(t_0) \text{ then } X = t_0, \\ &\text{if } p_0 < m \leq p_0 + p_1 = F(t_1) \text{ then } X = t_1, \\ &\quad \vdots \\ &\text{if } \sum_{j=0}^{k-1} p_j < m \leq \sum_{j=0}^k p_j = F(t_k) \text{ then } X = t_k. \end{aligned}$$

In order to generate n random numbers from discrete generalized Weibull distribution, t_1, t_2, \dots, t_n , repeat step 1 to step 2 n times.

3. RELIABILITY MEASURES OF DISCRETE GENERALIZED WEIBULL RANDOM VARIABLE dX

(a) Survival function

$$s(x) = p(dX \geq x) = \theta^{(x/\beta)^\alpha} \quad x = 0, 1, 2, \dots \text{ where } \theta = e^{-\gamma}, \quad 0 < \theta < 1, \quad \alpha > 0, \quad \beta > 0,$$

$s(x)$ is same as for continuous generalized Weibull distribution, and discrete generalized Weibull distribution at the integer values of x .

(b) Hazard Rate, $r(x)$, is given by

$$r(x) = \frac{p(x)}{s(x)} = \frac{\theta^{(x/\beta)^\alpha} - \theta^{((x+1)/\beta)^\alpha}}{\theta^{(x/\beta)^\alpha}}, \quad x = 0, 1, 2, \dots$$

$$\text{where } \theta = e^{-\gamma}, \quad 0 < \theta < 1, \quad \alpha > 0, \quad \beta > 0$$

(c) Second Hazard Rate is given by

$$h(x) = \log \left(\theta^{(x/\beta)^\alpha - ((x+1)/\beta)^\alpha} \right), \quad x = 0, 1, 2, \dots \text{ where } \theta = e^{-\gamma}, \quad 0 < \theta < 1, \quad \alpha > 0, \quad \beta > 0.$$

Fig. 3.1 to Fig. 3.3 illustrates the hazard rate plot for discrete generalized Weibull model for different values of parameters. It could be seen that hazard functions, $r(x)$ and $h(x)$ are always monotonic decreasing functions if

$$\alpha < \frac{\log(\log(\theta^2)/\log(\theta))}{\log 2} = \omega \text{ (say), where } \theta = e^{-\gamma}, 0 < \theta < 1, \alpha > 0, \beta > 0$$

For $\alpha = \omega$, discrete generalized Weibull model has a constant failure rate and for $\alpha > \omega$, discrete generalized Weibull model has an increasing failure rate.

3.1 Moments of Discrete Generalized Weibull Distribution

$$E(X^r) = \sum_{x=0}^{\infty} x^r p(x) = \sum_{x=0}^{\infty} [x^r - (x-1)^r] s(x)$$

Now,

$$E(X) = \sum_{x=1}^{\infty} s(x) = \sum_{x=1}^{\infty} \theta^{(x/\beta)^\alpha}, \quad E(X^2) = \sum_{x=1}^{\infty} (2x-1)s(x) = \sum_{x=1}^{\infty} (2x-1)\theta^{(x/\beta)^\alpha}$$

$$\text{and } V(X) = \sum_{x=1}^{\infty} (2x-1)\theta^{(x/\beta)^\alpha} - \left(\sum_{x=1}^{\infty} \theta^{(x/\beta)^\alpha} \right)^2.$$

The *mgf* of the discretized generalized Weibull distribution is

$$\begin{aligned} M_x(t) &= \sum_{x=0}^{\infty} e^{tx} p(x) = \sum_{x=0}^{\infty} e^{tx} \left(\theta^{(x/\beta)^\alpha} - \theta^{((x+1)/\beta)^\alpha} \right) \\ &= e^0 (\theta^{\psi(0;\beta,\alpha)}) + e^t (\theta^{\psi(1;\beta,\alpha)}) + e^{2t} (\theta^{\psi(2;\beta,\alpha)}) + e^{3t} (\theta^{\psi(3;\beta,\alpha)}) + \dots \\ &\quad - \left\{ e^0 (\theta^{\psi(1;\beta,\alpha)}) + e^t (\theta^{\psi(2;\beta,\alpha)}) + e^{2t} (\theta^{\psi(3;\beta,\alpha)}) + e^{3t} (\theta^{\psi(4;\beta,\alpha)}) + \dots \right\}, \\ &= 1 + (e^t - 1)(\theta^{\psi(1;\beta,\alpha)}) + (e^{2t} - e^t)(\theta^{\psi(2;\beta,\alpha)}) + (e^{3t} - e^{2t})(\theta^{\psi(3;\beta,\alpha)}) + \dots \\ &= 1 + \sum_{x=1}^{\infty} (e^{xt} - e^{(x-1)t}) \theta^{\psi(x;\beta,\alpha)}, \end{aligned}$$

where $\psi(x;\beta,\alpha) = (x/\beta)^\alpha, \beta > 0$ and $\alpha > 0$.

Differentiating the moment generating function r times with respect to t gives the r th moment of the DGWD(θ, β, α) in the form

$$M_x^{(r)}(t) |_{t=0} = \sum_{x=1}^{\infty} [x^r e^{xt} - (x-1)^r e^{(x-1)t}] \theta^{\psi(x;\beta,\alpha)}$$

The probability generating function of the discretized generalized Weibull distribution is

$$\begin{aligned} G_{[x]}(t) &= \sum_{x=0}^{\infty} t^x p(x) = \sum_{x=0}^{\infty} t^x \left(\theta^{(x/\beta)^\alpha} - \theta^{((x+1)/\beta)^\alpha} \right) = \sum_{x=0}^{\infty} t^x (s(x) - s(x+1)) \\ &= t^0 s(0) + t^1 s(1) + t^2 s(2) + \dots - \left\{ t^0 s(1) + t^1 s(2) + t^2 s(3) + \dots \right\} \\ &= 1 + (t-1)s(1) + (t^2 - t)s(2) + (t^3 - t^2)s(3) + \dots = 1 + (t-1) \sum_{x=1}^{\infty} t^{x-1} s(x). \end{aligned}$$

The first and the second derivative of the probability generating function with respect to t of the $DGWD(x; \theta, \beta, \alpha)$ are

$$G'_{[x]}(t) = \sum_{x=1}^{\infty} (xt - x + 1)t^{x-2}s(x),$$

$$G''_{[x]}(t) = \sum_{x=1}^{\infty} (x-1)t^{x-3} \{(t-1)(x-2) + 2t\} s(x).$$

At $t = 1$ these equations yield the first and second factorial moments.

$$E(x) = G'_{[x]}(t)|_{t=1} \text{ and } E(x^2) = G'_{[x]}(t)|_{t=1} + G''_{[x]}(t)|_{t=1}.$$

Table 3.1 exhibits the index of dispersion, for three parameter discrete generalized Weibull distribution using different values of the parameters θ , β and α . It can be realized that the variance to mean ratio shows that discrete generalized Weibull model is over-dispersed as well as under-dispersed.

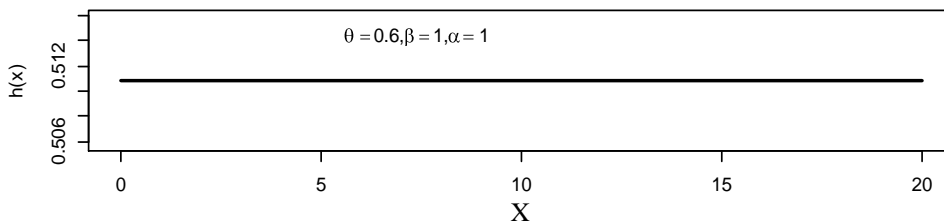


Fig. 3.1: $h(x)$ Plot for $DGWD(\theta, \beta, \alpha)$

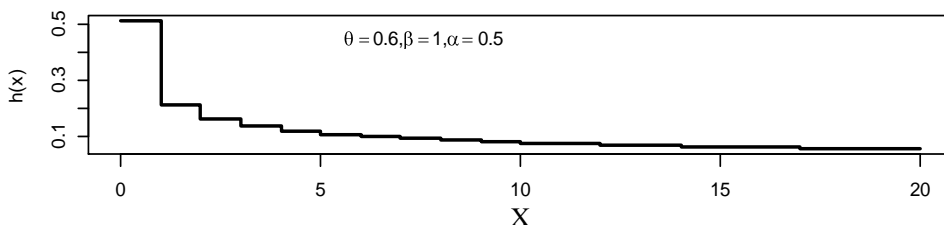


Fig. 3.2: $h(x)$ Plot for $DGWD(\theta, \beta, \alpha)$

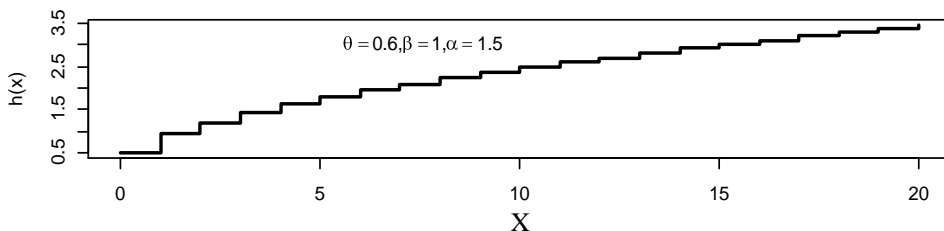


Fig. 3.3: $h(x)$ Plot for $DGWD(\theta, \beta, \alpha)$

Table 3.1
Index of Dispersion for DGWD for Different Combinations of Parameters

		Different Values of θ								
		$\beta = 1$	0.05	0.10	0.22	0.30	0.35	0.42	0.50	0.65
α	0.40	3.702	5.945	14.008	23.297	31.895	49.876	85.331	273.026	
	0.50	2.209	3.089	5.771	8.428	10.675	14.987	22.623	55.899	
	0.60	1.638	2.090	3.348	4.485	5.392	7.042	9.765	20.220	
	0.80	1.205	1.366	1.793	2.149	2.417	2.876	3.577	5.901	
	0.90	1.111	1.211	1.481	1.706	1.874	2.159	2.585	3.941	
	1.00	1.053	1.111	1.282	1.429	1.538	1.724	2.000	2.857	
	1.10	1.015	1.044	1.147	1.242	1.315	1.440	1.625	2.197	
	1.20	0.990	0.998	1.051	1.111	1.159	1.243	1.370	1.765	
	1.40	0.964	0.943	0.930	0.942	0.959	0.993	1.052	1.251	
	1.50	0.958	0.928	0.890	0.887	0.893	0.911	0.949	1.090	
	1.80	0.951	0.906	0.821	0.783	0.766	0.753	0.751	0.791	
	1.90	0.951	0.904	0.808	0.761	0.740	0.718	0.707	0.727	
	2.00	0.950	0.902	0.799	0.745	0.718	0.690	0.671	0.674	
	2.50	0.950	0.900	0.782	0.706	0.662	0.607	0.557	0.505	
3.00	0.950	0.900	0.780	0.700	0.651	0.584	0.512	0.412		
α	$\beta = 1.5$	0.05	0.10	0.22	0.30	0.35	0.42	0.50	0.65	
	0.40	4.923	8.174	20.030	33.813	46.616	73.458	126.491	407.748	
	0.5	2.848	4.128	8.059	11.984	15.315	21.731	33.121	82.915	
	0.6	2.049	2.717	4.569	6.248	7.593	10.043	14.099	29.724	
	0.8	1.414	1.674	2.330	2.867	3.269	3.958	5.008	8.488	
	0.9	1.263	1.437	1.873	2.222	2.480	2.913	3.558	5.600	
	1.0	1.157	1.275	1.573	1.812	1.987	2.277	2.702	4.006	
	1.1	1.079	1.156	1.364	1.532	1.655	1.858	2.153	3.036	
	1.2	1.019	1.066	1.209	1.329	1.418	1.564	1.776	2.402	
	1.4	0.929	0.936	0.996	1.057	1.105	1.185	1.302	1.646	
	1.5	0.895	0.886	0.918	0.961	0.996	1.056	1.146	1.408	
	1.8	0.812	0.772	0.750	0.759	0.771	0.796	0.836	0.963	
	1.9	0.789	0.742	0.708	0.710	0.717	0.735	0.766	0.867	
	2.3	0.709	0.640	0.575	0.561	0.558	0.560	0.570	0.612	
2.5	0.673	0.598	0.523	0.505	0.501	0.499	0.503	0.531		
α	$\beta = 2.0$	0.05	0.10	0.22	0.30	0.35	0.42	0.50	0.65	
	0.40	6.107	10.354	25.990	44.265	61.273	96.977	167.591	542.422	
	0.50	3.470	5.144	10.315	15.506	19.923	28.441	43.589	109.905	
	0.60	2.455	3.334	5.774	7.994	9.776	13.026	18.415	39.213	
	0.80	1.636	1.990	2.869	3.585	4.122	5.039	6.437	11.073	
	0.90	1.434	1.679	2.274	2.746	3.092	3.672	4.534	7.260	
	1.00	1.288	1.462	1.883	2.211	2.449	2.842	3.414	5.161	
	1.10	1.177	1.302	1.608	1.845	2.015	2.294	2.695	3.886	
	1.20	1.087	1.177	1.403	1.579	1.705	1.911	2.203	3.053	
	1.40	0.949	0.992	1.119	1.221	1.295	1.415	1.584	2.062	
	1.50	0.892	0.920	1.015	1.095	1.153	1.247	1.379	1.751	
	1.80	0.753	0.752	0.790	0.829	0.859	0.908	0.977	1.171	
	1.90	0.713	0.707	0.734	0.765	0.789	0.829	0.887	1.046	
	2.00	0.676	0.666	0.684	0.709	0.729	0.762	0.810	0.943	
2.50	0.519	0.502	0.502	0.510	0.518	0.532	0.553	0.613		

4. ESTIMATION OF PARAMETERS OF DISCRETE GENERALIZED WEIBULL DISTRIBUTION

i) Parameter Estimation Based on Maximum Likelihood Method

Let $x_1, x_2, x_3, \dots, x_n$ be a random sample of size n . If x_i 's are assumed to be IID random variables following generalized discrete Weibull distribution with likelihood function is given by

$$L(\theta, \beta, \alpha; x) = \prod_{i=1}^n p(x_i) = \prod_{i=1}^n \left(\theta^{(x_i/\beta)^\alpha} - \theta^{((x_i+1)/\beta)^\alpha} \right) \quad (4.1)$$

$$\log L(\theta, \beta, \alpha; x) = \sum_{i=1}^n \log \left(\theta^{(x_i/\beta)^\alpha} - \theta^{((x_i+1)/\beta)^\alpha} \right) \quad (4.2)$$

Taking partial derivatives of $\log L(\theta, \beta, \alpha; x)$ with respect to θ, β and α and equating to zero, we can get the normal equations to find the ML estimates.

$$\frac{\partial \log L(\theta, \beta, \alpha; x)}{\partial \theta} = \sum_{i=1}^n \frac{(x_i/\beta)^\alpha \hat{\theta}^{(x_i/\beta)^\alpha - 1} - ((x_i+1)/\beta)^\alpha \hat{\theta}^{((x_i+1)/\beta)^\alpha - 1}}{\hat{\theta}^{(x_i/\beta)^\alpha} - \hat{\theta}^{((x_i+1)/\beta)^\alpha}} \quad (4.3)$$

$$\begin{aligned} & \frac{\partial \log L(\theta, \beta, \alpha; x)}{\partial \beta} \\ &= \sum_{i=1}^n \frac{\frac{\alpha}{\hat{\beta}^2} \theta^{(x_i/\hat{\beta})^\alpha} x_i \log(\theta) \left(x_i / \hat{\beta} \right)^{\alpha-1} - \frac{\alpha}{\hat{\beta}^2} \theta^{((x_i+1)/\hat{\beta})^\alpha} (x_i+1) \log(\theta) \left((x_i+1) / \hat{\beta} \right)^{\alpha-1}}{\theta^{(x_i/\hat{\beta})^\alpha} - \theta^{((x_i+1)/\hat{\beta})^\alpha}} \end{aligned} \quad (4.4)$$

$$\begin{aligned} & \frac{\partial \log L(\theta, \beta, \alpha; x)}{\partial \alpha} \\ &= \sum_{i=1}^n \frac{\theta^{(x_i/\beta)^\alpha} (x_i/\beta)^{\hat{\alpha}} \log(\theta) \log(x_i/\beta) - \theta^{((x_i+1)/\beta)^\alpha} ((x_i+1)/\beta)^{\hat{\alpha}} \log(\theta) \log((x_i+1)/\beta)^{\hat{\alpha}}}{\theta^{(x_i/\beta)^\alpha} - \theta^{((x_i+1)/\beta)^\alpha}} \end{aligned} \quad (4.5)$$

The solution of above normal equations cannot be obtained in closed form, so by using numerical methods like Newton Raphson Method, the MLE of (θ, β, α) are the solution of above log-likelihood equations (4.3), (4.4) and (4.5).

Now we discuss four potential cases for estimation of parameters.

Case I: Considering known α and β , and unknown θ .

$$\left. \frac{\partial \log L(\theta, \beta, \alpha; x)}{\partial \theta} \right|_{\alpha=\hat{\alpha}, \beta=\hat{\beta}, \theta=\hat{\theta}} = 0 \Rightarrow \sum_{i=1}^n \frac{(x_i/\beta)^\alpha \hat{\theta}^{(x_i/\beta)^\alpha - 1} - ((x_i+1)/\beta)^\alpha \hat{\theta}^{((x_i+1)/\beta)^\alpha - 1}}{\hat{\theta}^{(x_i/\beta)^\alpha} - \hat{\theta}^{((x_i+1)/\beta)^\alpha}} = 0 \quad (4.6)$$

By solving equation (4.6) numerically using Newton Raphson Method gives the maximum likelihood estimate $\hat{\theta}$ of θ .

Case II: Considering known α and unknown β and θ

$$\left. \frac{\partial \log L(\theta, \beta, \alpha; x)}{\partial \beta} \right|_{\alpha=\hat{\alpha}, \beta=\hat{\beta}, \theta=\hat{\theta}} = 0 \text{ yields}$$

$$\frac{\sum_{i=1}^n \frac{\alpha}{\hat{\beta}^2} \theta^{(x_i/\hat{\beta})^\alpha} x_i \log(\theta) \left(x_i / \hat{\beta}\right)^{\alpha-1} - \frac{\alpha}{\hat{\beta}^2} \theta^{((x_i+1)/\hat{\beta})^\alpha} (x_i + 1) \log(\theta) \left((x_i + 1) / \hat{\beta}\right)^{\alpha-1}}{\theta^{(x_i/\hat{\beta})^\alpha} - \theta^{((x_i+1)/\hat{\beta})^\alpha}} = 0 \tag{4.7}$$

In order to get the ML estimates of the unknown parameters, we solve Equations (4.6) and (4.7) analytically using some numerical method like Newton Raphson Method to get the maximum likelihood estimates $\hat{\beta}$ and $\hat{\theta}$ of the parameters β and θ .

Case III: Considering known β and unknown θ and α .

$$\left. \frac{\partial \log L(\theta, \beta, \alpha; x)}{\partial \alpha} \right|_{\alpha=\hat{\alpha}, \beta=\hat{\beta}, \theta=\hat{\theta}} = 0 \text{ yields}$$

$$\sum_{i=1}^n \frac{\theta^{(x_i/\beta)^{\hat{\alpha}}} (x_i / \beta)^{\hat{\alpha}} \log(\theta) \log(x_i / \beta) - \theta^{((x_i+1)/\beta)^{\hat{\alpha}}} ((x_i + 1) / \beta)^{\hat{\alpha}} \log(\theta) \log((x_i + 1) / \beta)}{\theta^{(x_i/\beta)^{\hat{\alpha}}} - \theta^{((x_i+1)/\beta)^{\hat{\alpha}}}} = 0 \tag{4.8}$$

In a similar way as in case II, we solve the equations (4.6) and (4.8) analytically using some numerical method like Newton Raphson Method to get the maximum likelihood estimates $\hat{\theta}$ and $\hat{\alpha}$ of the parameters θ and α .

Case IV: Considering known θ , β and α .

In case of all unknown parameters, we solve the Equations (4.6), (4.7) and (4.8) analytically using some numerical method like Newton Raphson Method to get the maximum likelihood estimates $\hat{\theta}$, $\hat{\beta}$ and $\hat{\alpha}$ of the parameters θ , β and α respectively.

ii) Parameter Estimation using Proportion Method

Khan et al. (1989) considered and provided motivation of using method of proportions to estimate the unknown parameters for standard discrete Weibull distribution. We are going to provide a similar method for DGWD (discrete generalized Weibull distribution) for the same motives as sketched by Khan et al. (1989). Let $x_1, x_2, x_3, \dots, x_n$ be a random sample drawn from pmf (2.1). Put the indicator function

$$I_u(x_i) = \begin{cases} 1 & \text{if } x_i = u \\ 0 & \text{if } x_i \neq u \end{cases}$$

Define $f_u = \sum I_u(x_i)$ is the count of u in the observed sample. So, the probability $p(u; \theta, \beta, \alpha)$ can be estimated by proportion (relative frequency) $R_u = \frac{f_u}{n}$. Here the following four cases of parameter estimation using proportion method are considered.

Case I: Considering unknown θ and known α and β .

This case involves only one equation to be solved for estimation of parameter θ . The unknown parameter θ has a proportion estimate in exact solution, where

$$p(0; \theta, \beta, \alpha) = 1 - \theta^{\beta - \alpha} = f_0 / n, \\ \text{where } \theta = \exp(-\gamma), 0 < \theta < 1, \alpha > 0, \beta > 0, \gamma > 0, \theta^* = (1 - f_0 / n)^{\beta \alpha} \quad (4.9)$$

f_0 denotes the counts of zero's in a sample of size n .

Case II: Considering unknown β and θ and known α .

Let f_1 denote the zero counts in sample of size n .

$$p(1; \theta, \beta, \alpha) = \theta^{\beta - \alpha} - \theta^{(2/\beta)\alpha} = f_1 / n, \quad \text{where } \theta = \exp(-\gamma), 0 < \theta < 1, \alpha > 0, \beta > 0 \quad (4.10)$$

By solving together equations (4.9) and (4.10) numerically using some numerical method like Newton Raphson Method to get the proportion estimates β^* and θ^* of the parameters β and θ .

Case III: Considering unknown θ and α and known β .

Proceeding in a similar way as in case II, the solution of equations (4.9) and (4.10) numerically gives the proportion estimates θ^* and α^* of the parameters θ and α .

Case IV: Considering unknown θ , β and α .

Let f_2 denote the counts of two's in the sample of size n .

$$p(2; \theta, \beta, \alpha) = \theta^{(2/\beta)\alpha} - \theta^{(3/\beta)\alpha} = f_2 / n \quad \text{where } \theta = \exp(-\gamma), 0 < \theta < 1, \alpha > 0, \beta > 0 \quad (4.11)$$

The solution of equations (4.9), (4.10) and (4.11) analytically using a numerical method like Newton Raphson Method gives the proportion estimates θ^* , β^* and α^* of the parameters θ , β and α respectively.

5. SIMULATION STUDY

In this section, we investigate the performance of the ML estimators for a finite sample size n . Simulation study based on different $DGWD(x, \alpha, \beta, \theta)$ distribution is carried out. The random observations are generated by using the inverse cdf method presented in section 3.1 from $DGWD(\alpha, \beta, \theta)$. A simulation study was carried out for each triplet (α, β, θ) taking four parameter combinations as $(\alpha = 0.8, \beta = 0.5, \theta = 0.4)$, $(\alpha = 1.5, \beta = 1.0, \theta = 0.6)$, $(\alpha = 1.7, \beta = 1.2, \theta = 0.7)$, $(\alpha = 2.0, \beta = 1.5, \theta = 0.8)$ and the process was repeated 500 times by taking different sample sizes $n = (25, 50, 100, 200, 300, 500)$. The simulated results are given in Table 5.1. We observe in Table 5.1 that the agreement between theory and practice improves as the sample size n increases. MSE and Variance of the estimators suggest us that the estimators are consistent and the maximum likelihood method performs quite well in estimating the model parameters of the proposed distribution.

Table 5.1
Average Bias, MSE and Variance for Simulated Results of ML Estimates

Sample Size		$(\alpha = 0.8, \beta = 0.5, \theta = 0.4)$			$(\alpha = 1.5, \beta = 1.0, \theta = 0.6)$		
		Bias	Variance	MSE	Bias	Variance	MSE
25	α	0.354254	0.686322	0.811818	0.156408	0.200135	0.224598
	β	0.042313	0.017507	0.019297	0.074247	0.045086	0.050598
	θ	0.016466	0.021298	0.02157	-0.01996	0.005871	0.00627
50	α	0.170404	0.279043	0.308081	0.034085	0.041747	0.042909
	β	0.035725	0.011474	0.01275	0.069805	0.027959	0.032832
	θ	-0.00563	0.007489	0.007521	-0.02659	0.002853	0.00356
100	α	0.071532	0.02656	0.031677	0.022412	0.019644	0.020146
	β	0.037706	0.007903	0.009325	0.032372	0.005728	0.006776
	θ	-0.0073	0.001301	0.001354	-0.01022	0.000715	0.000819
200	α	0.00144	0.012302	0.012305	0.019303	0.012486	0.012859
	β	0.015188	0.003514	0.003744	0.020094	0.003019	0.003422
	θ	-0.01264	0.000699	0.000858	-0.00781	0.000457	0.000518
300	α	0.005938	0.008122	0.008158	0.008124	0.007138	0.007204
	β	0.014712	0.002924	0.003141	0.019941	0.002399	0.002796
	θ	-0.00554	0.000422	0.000452	-0.00818	0.000271	0.000338
500	α	0.003817	0.00306	0.003074	0.005239	0.005958	0.005985
	β	0.01347	0.00163	0.001811	0.00882	0.00118	0.001258
	θ	-0.00631	0.000168	0.000208	-0.00505	0.000191	0.000217

Table 5.1 (contd...)

Sample Size		$(\alpha = 0.8, \beta = 0.5, \theta = 0.4)$			$(\alpha = 1.5, \beta = 1.0, \theta = 0.6)$		
		Bias	Variance	MSE	Bias	Variance	MSE
25	α	0.10132	0.177988	0.188254	0.098607	0.173873	0.183597
	β	0.06196	0.062467	0.066306	0.124706	0.038745	0.054296
	θ	-0.01652	0.009179	0.009452	-0.02162	0.004743	0.00521
50	α	0.060455	0.051417	0.055071	0.088946	0.0754	0.083311
	β	0.033915	0.016241	0.017391	0.041369	0.031117	0.032828
	θ	-0.01455	0.002574	0.002785	-0.00988	0.003387	0.003484
100	α	0.053763	0.029788	0.032678	-0.00549	0.01881	0.01884
	β	0.041619	0.012894	0.014627	0.058699	0.018328	0.021774
	θ	-0.01015	0.001266	0.001369	-0.0128	0.002241	0.002405
200	α	-0.00314	0.012977	0.012987	-0.00184	0.016614	0.016618
	β	0.037516	0.005292	0.0067	0.030852	0.008759	0.009711
	θ	-0.01366	0.00065	0.000837	-0.00742	0.00137	0.001425
300	α	0.008932	0.008095	0.008175	0.014248	0.008586	0.008789
	β	0.024452	0.002749	0.003347	0.011125	0.003344	0.003468
	θ	-0.00728	0.000273	0.000326	-0.00471	0.00064	0.000662
500	α	0.000131	0.003587	0.003587	0.033694	0.005929	0.007064
	β	0.015715	0.000951	0.001198	0.011538	0.002167	0.0023
	θ	-0.0059	0.000178	0.000213	0.002684	0.000302	0.000309

6. SOME THEOREMS ASSOCIATED WITH DISCRETE GENERALIZED WEIBULL DISTRIBUTION

In this section, we provide few important theorem.

Theorem 1:

If $X \sim GWD(x; \gamma, \beta, \alpha)$ then $Y = \left[\left(e^{\left(\frac{x}{\beta} \right)^\alpha} - 1 \right)^{\frac{1}{\alpha}} \right] \beta$ follows discrete Burr-Type XII

distribution i.e., $DBD-XII(x; \theta, \beta, \alpha)$, where $\theta = e^{-\gamma}$, $0 < \theta < 1$, $\beta > 0$, $\alpha > 0$, $\gamma > 0$.

Proof:

$$p[Y \geq y] = p \left[\left[\left(e^{\left(\frac{x}{\beta} \right)^\alpha} - 1 \right)^{\frac{1}{\alpha}} \right] \beta \geq y \right]$$

$$\begin{aligned}
 &= p \left[\left(\frac{X}{\beta} \right)^\alpha \geq \log \left[\left(\frac{y}{\beta} \right)^\alpha + 1 \right] \right] \\
 &= p \left[X \geq \left(\log \left[\left(\frac{y}{\beta} \right)^\alpha + 1 \right] \right)^{\frac{1}{\alpha}} \beta \right] = \left[\left(\frac{y}{\beta} \right)^\alpha + 1 \right]^{-\gamma},
 \end{aligned}$$

where $\beta > 0, \alpha > 0, \gamma > 0$, which is a three parameter DBXIID (Discrete Burr type XII distribution).

Theorem 2:

If $X \sim GWD(x; \gamma, \beta, \alpha)$ then $Y = \left[\left(e^{\left(\frac{x}{\beta} \right)^\alpha} - 1 \right) \beta \right]$ follows discrete Pareto distribution

where i.e., $DP(y; \theta, \beta)$ where $\theta = e^{-\gamma}, 0 < \theta < 1, \beta > 0, \gamma > 0$.

Proof:

$$\begin{aligned}
 p[Y \geq y] &= p \left[\left(e^{\left(\frac{X}{\beta} \right)^\alpha} - 1 \right) \beta \geq y \right] = p \left[\left(\frac{X}{\beta} \right)^\alpha \geq \log \left[1 + \left(\frac{y}{\beta} \right) \right] \right] \\
 &= p \left[X \geq \left(\log \left[1 + \left(\frac{y}{\beta} \right) \right] \right)^{1/\alpha} \beta \right] = \theta^{\log \left[1 + \left(\frac{y}{\beta} \right) \right]} = \left(1 + \left(\frac{y}{\beta} \right) \right)^{-\gamma},
 \end{aligned}$$

where $\beta > 0, \gamma > 0$, which is the reliability function of a DPD (discrete Pareto distribution).

Theorem 3:

If $X \sim BD-III(x; \gamma, \beta, \alpha)$, with survival function, $s(x) = 1 - (1 + \beta x^{-\alpha})^{-\gamma}$, then $Y = \left[\left(\log \left(1 + \beta x^{-\alpha} \right) \right)^{\frac{1}{\alpha}} \beta \right]$ follows discrete generalized Weibull distribution i.e.,

$DGWD(y; \theta, \beta, \alpha)$, where $\theta = e^{-\gamma}, 0 < \theta < 1, \beta > 0, \alpha > 0, \gamma > 0$.

Proof:

Consider

$$\begin{aligned}
 e^{-\gamma \ln \left[\left(\frac{y}{\beta} \right)^\alpha + 1 \right]} &= \left(\left(\frac{y}{\beta} \right)^\alpha + 1 \right)^{-\gamma} \\
 p[Y \geq y] &= p \left[\left[\left(\log(1 + \beta x^{-\alpha}) \right)^{\frac{1}{\alpha}} \beta \geq y \right] \right] \\
 p[Y \geq y] &= 1 - p \left[X \geq \frac{\left(\left(\frac{y}{\beta} \right)^\alpha - 1 \right)^{-\frac{1}{\alpha}}}{\beta} \right] \\
 &= 1 - \left[1 - \left[1 + \beta \left(\left(\frac{y}{\beta} \right)^\alpha - 1 \right)^{-\frac{1}{\alpha}} \right]^{-\gamma} \right] = e^{-\gamma \left(\frac{y}{\beta} \right)^\alpha}
 \end{aligned}$$

where $\beta > 0$, $\alpha > 0$, $\gamma > 0$, which is the survival function of a discrete generalized Weibull distribution.

7. APPLICATION OF DISCRETE GENERALIZED WEIBULL DISTRIBUTION

Here we consider a data set related to counts of stillbirths for 402 New Zealand white rabbit litters initially discussed by Morgan, Palmer and Ridout (2007) in the situation of Score Tests as given in the Table 7.1(a). The distribution is apparently zero inflated with 78.1% of the litters having no stillbirths and over dispersion is clearly present as the variance (1.37) is much larger than the mean (0.45). Table 7.1(b) exhibits some descriptive measures of stillbirths in 402 New Zealand white rabbit litters based on 1000 bootstrap samples.

Table 7.1(a)
Counts of stillbirths for 402 New Zealand White Rabbit Litters

Number of Stillbirths	0	1	2	3	4	5	6	7	≥8	Total
Frequency	314	48	20	7	5	2	2	1	3	402

Table 7.1(b)
Descriptive Statistics of Counts of Stillbirths
in 402 New Zealand White Rabbit Litters

Measure	Statistic Calculated	Standard Error (SE)	Bootstrap ^a			
			Bias	Standard Error (SE)	95% Confidence Interval	
					Lower	Upper
Mean	.4527	.05850	.0037	.0580	.3458	.5721
Standard Deviation	1.17285		-.00253	.13316	.89890	1.42456
Variance	1.376		.012	.312	.808	2.029
Skewness	3.870	.122	-.069	.346	3.146	4.531
Kurtosis	17.765	.243	-.476	3.626	11.236	25.802
N	402		0	0	402	402

a. Bootstrap results are based on 1000 bootstrap samples

Parameter estimation was done by using the R package “FAdist” to obtain ML estimates of unknown parameters. The ML estimates obtained using fitdistr procedure as given in the following Table 7.2.

Table 7.2
Parameter Estimates by ML Method for Fitted Distributions

Distribution	Parameter Estimates	Model Function
Discrete generalized inverse Weibull distribution	$\theta = 0.17,$ $\alpha = 1.47,$ $\beta = 0.26$	$p(x) = \theta \left(\frac{\beta}{x+1}\right)^\alpha - \theta \left(\frac{\beta}{x}\right)^\alpha \quad x = 0, 1, 2, \dots$ $0 < \theta < 1, \alpha > 0, \beta > 0$
Discrete generalized Weibull distribution	$\theta = 0.15,$ $\alpha = 0.60,$ $\beta = 1.45$	$p(x) = \theta \left(\frac{x}{\beta}\right)^\alpha - \theta \left(\frac{x+1}{\beta}\right)^\alpha \quad x = 0, 1, 2, \dots$ $0 < \theta < 1, \alpha > 0, \beta > 0$
Poisson Lindley	$\theta = 2.83$	$p(x) = \frac{\theta^2(x + \theta + 2)}{(\theta + 1)^{x+3}} \quad x = 0, 1, 2, \dots \quad \theta > 0$
Poisson	$\lambda = 0.45$	$p(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad \lambda > 0; \quad x = 0, 1, 2, \dots$
DRayleigh	$q = 0.64$	$p(x) = q^{x^2} - q^{(x+1)^2} \quad 0 < q < 1; \quad x = 0, 1, 2, \dots$
Geometric	$q = 0.69$	$p(x) = q^x p \quad x = 0, 1, 2, \dots \quad q = 1 - p, \quad 0 < q < 1$
Zero Inflated Poisson	$\alpha = 0.73,$ $\lambda = 1.68$	$p(x) = \begin{cases} \alpha + (1 - \alpha) \frac{e^{-\lambda} \lambda^x}{x!}, & \lambda > 0; x = 0 \\ (1 - \alpha) \frac{e^{-\lambda} \lambda^x}{x!}, & \lambda > 0; x = 1, 2, 3, \dots \end{cases}$ $0 < \alpha < 1; \quad \lambda > 0$

Computation of the expected frequencies for fitting DGWD (discrete generalized Weibull), DGIWD (discrete generalized inverse Weibull distribution), PLD (Poisson Lindley distribution), PD (Poisson distribution), GD (Geometric distribution), ZIPD (Zero Inflated Poisson distribution) and DRD (DRayleigh distribution) using R and Pearson's Chi-square test is employed to assess the goodness of fit of considered models. The calculated expected counts for each discussed model are given in the Table 7.3.

Table 7.3
Table for Goodness-of-Fit

X	Observed	DGIWD	Poisson	DRayleigh	Geometric	ZIP	PLD	DGWD
0	314.000	313.503	255.626	143.452	276.719	314.000	276.850	313.876
1	48.000	54.170	115.731	189.764	86.238	33.770	87.207	48.111
2	20.000	15.096	26.198	61.214	26.876	28.442	26.662	18.927
3	7.000	6.555	3.954	7.225	8.376	15.969	7.977	9.038
4	5.000	3.520	0.447	0.338	2.610	6.725	2.348	4.787
5	2.000	2.144	0.041	0.006	0.813	2.265	0.682	2.712
6	2.000	1.419	0.003	0.000	0.254	0.636	0.196	1.612
7	1.000	0.996	0.000	0.000	0.079	0.153	0.056	0.995
8	3.000	4.598	0.000	0.000	0.036	0.039	0.022	1.943
χ^2 P-Values --		0.1260	0.000	0.0000	0.0000	0.0007	0.00001	0.4402

X: Number of stillbirths in 402 New Zealand white rabbit litters

Observed: Observed frequency of stillbirths in 402 New Zealand white rabbit litters

The p-values based on Pearson's Chi-square statistic are 0.1260, 0.000, 0.000, 0.000, 0.0007, 0.00001 and 0.4402 for DGIWD (discrete generalized inverse Weibull distribution), PD (Poisson distribution), DRD (Discrete Rayleigh distribution), GD (Geometric distribution), ZIPD (Zero Inflated Poisson distribution), PLD (Poisson Lindley distribution) and DGWD (discrete generalized Weibull distribution), respectively (see Table 7.3). This reveals that DGIWD (discrete generalized inverse Weibull distribution), PD (Poisson distribution), DRD (Discrete Rayleigh distribution), GD (Geometric distribution), ZIPD (Zero Inflated Poisson distribution) and PLD (Poisson Lindley distribution) are not good fit at all, whereas DGWD (discrete generalized Weibull distribution) provides a good fit. The null hypothesis that data come from DGWD (discrete generalized Weibull distribution) is accepted.

We have compared DGWD (discrete generalized Weibull distribution) with DGIWD (discrete generalized inverse Weibull distribution), PD (Poisson distribution), DRD (Discrete Rayleigh distribution), GD (Geometric distribution), ZIPD (Zero Inflated Poisson distribution) and PLD (Poisson Lindley distribution) using the AIC [Akaike (1974)] and BIC [Schwarz (1978)] criterion.

Table 7.4
AIC, BIC and NLL (Negative Loglikelihood)

Criterion	DGIWD	Poisson	DRayleigh	Geometric	ZIP	PLD	DGWD
NLL	339.116	431.593	621.015	351.895	362.319	365.203	335.841
AIC	684.233	865.186	1244.030	707.791	726.638	732.405	677.682
BIC	696.222	869.183	1248.027	715.784	730.635	736.402	689.672

From Table 7.4, Comparing the fits using AIC and BIC criterion, it is noticeable that AIC and BIC criterion favors DGWD (discrete generalized Weibull distribution) in comparison with the DGIWD (discrete generalized inverse Weibull distribution), PD (Poisson distribution), DRD (Discrete Rayleigh distribution), GD (Geometric distribution), ZIPD (Zero Inflated Poisson distribution) and PLD (Poisson Lindley distribution), in the case of number of stillbirths in 402 New Zealand white rabbit litters.

REFERENCES

1. Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723. doi:10.1109/tac.1974.1100705.
2. De Gusmão, F.R., Ortega, E.M. and Cordeiro, G.M. (2011). The generalized inverse Weibull distribution. *Statistical Papers*, 52(3), 591-619.
3. Gómez-Déniz, E. and Calderín-Ojeda, E. (2011). The discrete Lindley distribution: properties and applications. *Journal of Statistical Computation and Simulation*, 81(11), 1405-1416. doi:10.1080/00949655.2010.487825.
4. Inusah, S. and Kozubowski, T. (2006). A discrete analogue of the Laplace distribution. *Journal of Statistical Planning and Inference*, 136(3), 1090-1102.
5. Jazi, M.A., Lai, C. and Alamatsaz, M.H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. *Statistical Methodology*, 7(2), 121-132. doi:10.1016/j.stamet.2009.11.001.
6. Kapur, K.C. and Lamberson, L.R. (1997). *Reliability in Engineering Design*, John Wiley & Sons, New York.
7. Kemp, A. (1997). Characterizations of a discrete normal distribution. *Journal of Statistical Planning and Inference*, 63(2), 223-229.
8. Khan, M., Khalique, A. and Abouammoh, A. (1989). On estimating parameters in a discrete Weibull distribution. *IEEE Transactions on Reliability*, 38(3), 348-350. doi:10.1109/24.44179.
9. Krishna, H. and Pundir, P.S. (2009). Discrete Burr and discrete Pareto distributions. *Statistical Methodology*, 6(2), 177-188. doi:10.1016/j.stamet.2008.07.001.
10. Lawless, J.F. (1982). *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, New York.
11. McElduff, F.C. (2012). *Models for Discrete Epidemiological and Clinical data*. A thesis presented for the degree of Doctor of Philosophy University College London.
12. Morgan, B.J.T., Palmer, K.J. and Ridout, M.S. (2007). Score Test Oddities. *The American Statistician*, 61, 285-288.

13. Nekoukhou, V., Alamatsaz, M.H. and Bidram, H. (2012). A Discrete Analog of the Generalized Exponential Distribution. *Communications in Statistics - Theory and Methods*, 41(11), 2000-2013. doi:10.1080/03610926.2011.555044.
14. Para, B.A. and Jan, T.R. (2014). Discretization of Burr-Type III Distribution, *Journal of Reliability and Statistical Studies*, 7(2), 87-94.
15. Para, B.A. and Jan, T.R. (2016a). On Discrete Three Parameter Burr Type XII and Discrete Lomax Distributions and Their Applications to Model Count Data from Medical Science. *Biom Biostat. Int. J.*, 4(2), 00092. DOI: 10.15406/bbij.2016.04.00092.
16. Para, B.A. and Jan, T.R. (2016b). Discrete Version of Log-Logistic Distribution and Its Applications in Genetics. *International Journal of Modern Mathematical Sciences*, 14(4), 407-422.
17. Roy, D. (1993). Reliability Measures in the Discrete Bivariate Set-Up and Related Characterization Results for a Bivariate Geometric Distribution. *Journal of Multivariate Analysis*, 46(2), 362-373. doi:10.1006/jmva.1993.1065.
18. Roy, D. (2003). The Discrete Normal Distribution. *Communications in Statistics - Theory and Methods*, 32(10), 1871-1883. doi:10.1081/sta-120023256.
19. Roy, D. (2004). Discrete Rayleigh Distribution. *IEEE Transactions on Reliability*, 53(2), 255-260. doi:10.1109/tr.2004.829161
20. Schwarz, G. (1978). Estimating the Dimension of a Model. *Ann. Statist.*, 6(2), 461-464. doi:10.1214/aos/1176344136.
21. Sinha, S.K. (1986). *Reliability and Life Testing*, Wiley Eastern Ltd, New Delhi.