



The Beta-Rayleigh Distribution on the Lattice of Integers

Vahid Nekoukhou

University of Isfahan, Khansar Unit

Received: 9/14/2015 Approved: 5/28/2016

Abstract. In this paper, a discrete analog of the beta-Rayleigh distribution is studied. This new distribution contains the generalized discrete Rayleigh and discrete Rayleigh distributions as special sub-models. Some distributional and moment properties of the new discrete distribution as well as its order statistics are discussed. We will see that the hazard rate function of the new model can be increasing, bathtub-shaped and upside-down bathtub. Estimation of the parameters is illustrated and, finally, the model with a real data set is examined.

Keywords. Discrete Rayleigh distribution; generalized discrete Rayleigh distribution; exponentiated discrete Weibull distribution; hazard rate function.

MSC 2010: 60E05, 62E10.

1 Introduction

Eugene et al. (2002) introduced a general class of distributions generated from the logit of a beta random variable with cumulative distribution function (cdf)

$$F(x; a, b, \theta) = I_{G(x)}(a, b) = \frac{1}{B(a, b)} \int_0^{G(x)} t^{a-1} (1-t)^{b-1} dt, \quad x \in \mathbb{R}, \quad (1)$$

where $a > 0$ and $b > 0$ are two parameters whose role is to control the skewness and tail weight, $\boldsymbol{\theta}$ is the parameter vector of the absolutely continuous cdf G , $I_y(a, b) = \frac{B_y(a, b)}{B(a, b)}$ is the incomplete beta function ratio and $B_y(a, b) = \int_0^y t^{a-1}(1-t)^{b-1} dt$ denotes the incomplete beta function. In recent years, the class of beta- G (BG) distributions has received much attention in the literature. For instance, Cordeiro et al. (2013a) and Singla et al. (2012) considered the beta exponentiated (generalized) Weibull distribution and investigated its properties and application. Cordeiro and Lemonte (2011a) and (2011b) studied the beta Laplace and beta-half-Cauchy distributions, respectively. Morais et al. (2013) considered the beta generalized logistic distribution and investigated some of its important properties and features. Cordeiro et al. (2013b) introduced the beta generalized Rayleigh distribution and considered its applications to lifetime data. Cordeiro et al. (2013c) and Bidram et al. (2013) introduced the beta Weibull-geometric distribution. As alternative studies in this connection, we can address the beta generalized Pareto distribution proposed by Mahmoudi (2011), beta Lomax distribution of Lemonte and Cordeiro (2013), beta modified Weibull distribution of Silva et al. (2010) and beta log-normal distribution given by Castellares et al. (2013).

Recently, constructing discrete counterparts of known continuous distributions in order to study new discrete distributions has received much attention in the literature. Indeed, discrete distributions play key roles in survival analysis. The lifetimes of many components are being measured by the number of completed cycles of operation or strokes. Even for a continuous operation, involving a continuous measurement of lifetime, observations made at periodic time points give rise to a discrete situation, and a discrete model may be more appropriate (cf. Roy, 2004). In this way, Nekoukhoo and Bidram (2015a) introduced a class of discrete BG distributions on the set of integers $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ emerges as

$$f(y; a, b, \boldsymbol{\theta}) = P(Y = y) = I_{G(y+1)}(a, b) - I_{G(y)}(a, b).$$

They investigated that the above probability mass function (pmf) can also be expressed as

$$f(y; a, b, \boldsymbol{\theta}) = \sum_{i=0}^{\infty} \omega_i(a, b) [\{G(y+1)\}^{a+i} - \{G(y)\}^{a+i}], \quad (2)$$

where

$$\omega_i(a, b) = \frac{\Gamma(a+b)(-1)^i}{\Gamma(a)\Gamma(b-i)(a+i)!}.$$

The cdf of a discrete BG distribution is also given by

$$F(y; a, b, \boldsymbol{\theta}) = \sum_{i=0}^{\infty} \omega_i(a, b) \{G([y] + 1)\}^{a+i}, \quad y \in \mathbb{R}, \quad (3)$$

where $[.]$ denotes the integer function.

The last authors introduced the discrete beta generalized exponential (DBGE) distribution by inserting the cdf of the generalized exponential (GE) distribution of Gupta and Kundu (1999) into Equation (2). Indeed, they studied a discrete analog of the beta generalized exponential distribution of Barreto-Souza et al. (2010).

Here, we attempt to introduce the *discrete beta-Rayleigh* (DBR) distribution as another member of the class of discrete BG distributions. This is so because discrete random variables can also be encountered frequently for many different practical reasons. For example, in lifetime modeling, field failures are often collected and reported daily, weekly, and so forth. Systems often operate in cycles and the experimenter observes the number of cycles successfully completed prior to failure (cf. Lee and Cha, 2015).

We will see that the hazard rate function of the new discrete distribution can be increasing, bathtub-shaped and upside-down bathtub. Hence, this discrete analog of the beta-Rayleigh distribution is more flexible than the classic discrete distributions.

The paper is organized as follows: Section 2, introduces the DBR distribution. Some important features and properties of the new distribution such as the probability mass and cumulative distribution functions are studied. In addition, the mean and variance of the DBR distribution will be obtained. The researcher will consider the behavior of the hazard rate function and show that the new distribution exhibits bathtub, upside-down bathtub and monotonically increasing hazard rates. The infinite divisibility of the distribution in question is discussed. The cdf and pmf of the order statistics of DBR distributions will be attained. The estimation process of the parameters is discussed and a kind of simulated example is provided. Additionally, the stress-strength parameter is illustrated and the new model with a real data set is also examined. Some concluding remarks are given in Section 3.

2 The DBR distribution

Here, we construct a discrete analog of the beta-Rayleigh (BR) distribution by inserting the cdf of the Rayleigh distribution,

$$F(y; \theta) = 1 - e^{-\theta y^2}, \quad y > 0, \quad (4)$$

where $\theta > 0$ is the scale parameter, into Equation (2). Consequently, the pmf of a DBR distribution for $y \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$ is given by

$$\begin{aligned} p_y &= \sum_{i=0}^{\infty} \omega_i(a, b) \left\{ \left[1 - e^{-\theta(y+1)^2} \right]^{a+i} - \left[1 - e^{-\theta y^2} \right]^{a+i} \right\} \\ &= \sum_{i=0}^{\infty} \omega_i(a, b) \left\{ \left(1 - p^{(y+1)^2} \right)^{a+i} - \left(1 - p^{y^2} \right)^{a+i} \right\}, \end{aligned} \quad (5)$$

where $a > 0$, $b > 0$ and $0 < p = e^{-\theta} < 1$ are the model parameters. For integer values of b , the sum in Equation (5) is finite and stops at $b - 1$; see Nekoukhou and Bidram (2015a).

The pmf given by Equation (5) defines the DBR distribution and we will denote it by $DBR(a, b, p)$ in the rest of the paper.

It is interesting to note that the pmf of a DBR distribution is a linear combination of the two-parameter generalized discrete Rayleigh (GDR) probability mass functions, introduced by Nekoukhou and Bidram (2015b). A GDR distribution, with parameters $0 < p < 1$ and $\gamma > 0$, is a special sub-model of the three-parameter exponentiated discrete Weibull distribution of Nekoukhou and Bidram (2015b) and has the following pmf

$$f_{GDR}(y; p, \gamma) = \{1 - p^{(y+1)^2}\}^\gamma - \{1 - p^{y^2}\}^\gamma \quad (6)$$

$$= \sum_{j=1}^{\infty} (-1)^{j+1} \binom{\gamma}{j} \{(p^j)^{y^2} - (p^j)^{(y+1)^2}\}, \quad (7)$$

where $\binom{\gamma}{j} = \frac{\Gamma(\gamma+1)}{\Gamma(\gamma+1-j)j!}$ and for integer $\gamma > 0$, the sum in Eq. (7) stops at γ . Others view of point, a GDR distribution can be considered as the exponentiated discrete Rayleigh (DR) distribution of Roy (2004). Therefore,

we have the following representation for the pmf of a DBR distribution

$$f_{DBR}(y; a, b, p) = \sum_{i=0}^{\infty} \omega_i(a, b) f_{GDR}(y; p, a + i).$$

For $b = 1$, the $DBR(a, b, p)$ distribution reduces to the $GDR(p, a)$ distribution and if in addition $a = 1$, then the DR distribution is obtained.

The pmf of the DBR distribution can also be written as a linear combination of the DR pmfs as follows

$$f_{DBR}(y; a, b, p) = \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \omega_i(a, b) (-1)^{j+1} \binom{a+i}{j} \{(p^j)^{y^2} - (p^j)^{(y+1)^2}\}.$$

For integer values of a , $\sum_{j=1}^{\infty}$ should be replaced by $\sum_{j=1}^{a+i}$ in the above equation. Figure 1 indicates the behavior of DBR pmfs for some values of their parameters.

The cdf of a DBR distribution, using Equation (3), is of the form

$$F_{DBR}(y; a, b, p) = \sum_{i=0}^{\infty} \omega_i(a, b) F_{GDR}(y; p, a + i), \quad y \geq 0,$$

where F_{GDR} denotes the cdf of $GDR(p, a + i)$ distributions. Hence, we have the following result

$$F_{DBR}(y; a, b, p) = \sum_{i=0}^{\infty} \omega_i(a, b) \left(1 - p^{([y]+1)^2}\right)^{a+i}, \quad y \geq 0. \quad (8)$$

2.1 Properties of DBR Distributions

The survival and hazard rate functions of a random variable Y following a $DBR(a, b, p)$ distribution are given, respectively, by

$$\bar{F}_{DBR}(y; a, b, p) = P(Y > y) = 1 - \sum_{i=0}^{\infty} \omega_i(a, b) \left(1 - p^{([y]+1)^2}\right)^{a+i}, \quad y \geq 0$$

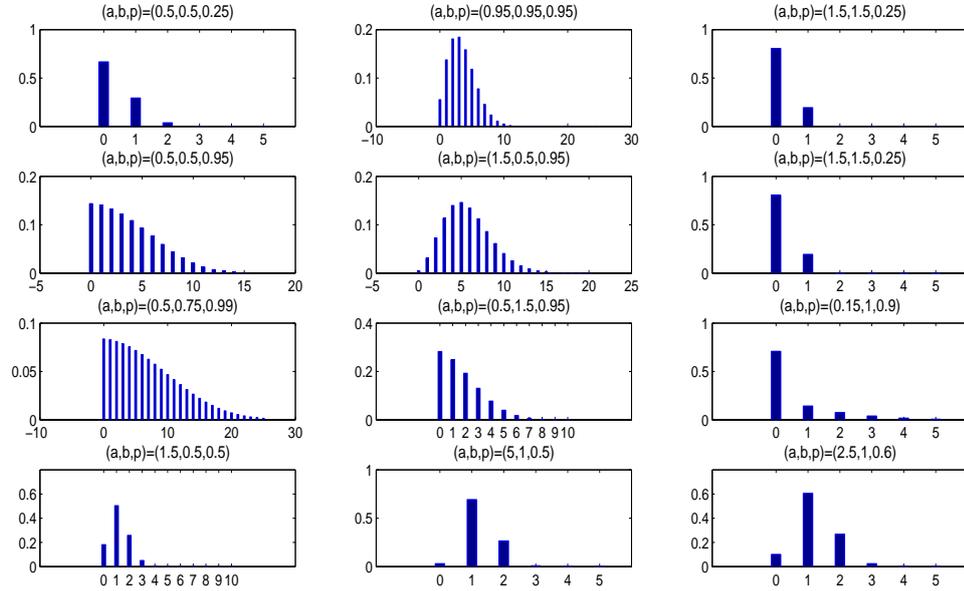


Figure 1. Pmfs of DBR distributions for some selected parameters values.

and

$$h_{DBR}(y; a, b, p) = \frac{\sum_{i=0}^{\infty} \omega_i(a, b) \left\{ \left(1 - p^{(y+1)^2}\right)^{a+i} - \left(1 - p^{y^2}\right)^{a+i} \right\}}{1 - \sum_{i=0}^{\infty} \omega_i(a, b) \left\{1 - p^{(y+1)^2}\right\}^{a+i}}, y \in \mathbb{N}_0.$$

As mentioned before, for an integer $b > 0$, $\sum_{i=0}^{\infty}$ should be replaced by $\sum_{i=0}^{b-1}$ in the above relations. Figure 2 illustrates the hazard rate function behavior of the DBR distribution. As we see from Figure 2, the hazard rate function of a DBR distribution can be increasing, bathtub-shaped and upside-down bathtub depending on the parameters values.

The mean and variance of a random variable $Y \sim DBR(a, b, p)$ are given, respectively, by

$$E(Y) = \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \sum_{y=0}^{\infty} \omega_i(a, b) (-1)^{j+1} \binom{a+i}{j} p^{j(y+1)^2}$$

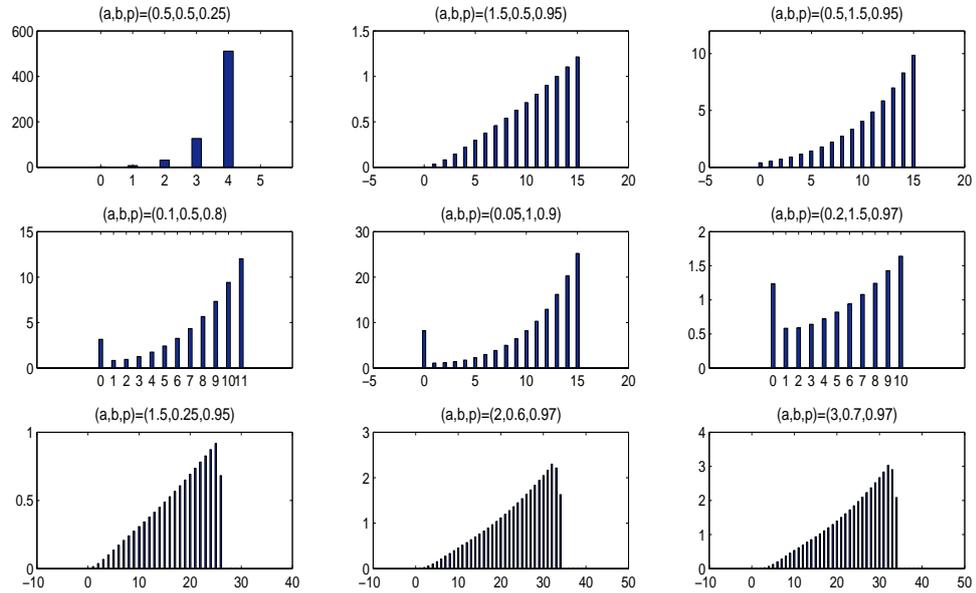


Figure 2. Hazard rate function plots of DBR distributions for selected parameters values.

and

$$Var(Y) = 2 \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \sum_{y=0}^{\infty} \omega_i(a, b) (-1)^{j+1} \binom{a+i}{j} y p^{j(y+1)^2} + E(Y) - \{E(Y)\}^2.$$

The mean and variance of a DBR distribution for different values of its parameters are calculated in Table 1. It appears that the mean increases with p for fixed values of a and b . In addition, DBR models are appropriate for modeling both over- and under-dispersed data since, in these models, the variance can be larger or smaller than the mean which is not the case with some standard classical discrete distributions. Hence, the parameters of the distribution can be adjusted to suit most data sets.

2.2 Infinite Divisibility

The researchers here make the following note in regards to the famous structural property of infinite divisibility of the distribution in question. Such a characteristic has a close relation to the Central Limit Theorem and

Table 1. Mean (Variance) of $DBR(a, b, p)$ for different values of a , b and p .

$p = 0.25$				
a/b	0.50	1.00	1.50	2.00
0.50	0.3744 (0.3189)	0.1359 (0.1214)	0.0578 (0.0546)	0.0257 (0.0251)
1.00	0.5645 (0.3787)	0.2539 (0.1973)	0.1252 (0.1100)	0.0625 (0.0586)
1.50	0.6910 (0.3826)	0.3563 (0.2411)	0.1959 (0.1584)	0.1069 (0.0956)
2.00	0.7841 (0.3684)	0.4453 (0.2626)	0.2662 (0.1966)	0.1563 (0.1320)
$p = 0.75$				
a/b	0.50	1.00	1.50	2.00
0.50	1.3092 (1.6289)	0.7169 (0.7358)	0.4846 (0.04566)	0.3571 (0.3233)
1.00	1.8367 (1.5728)	1.1523 (0.8241)	0.8491 (0.5726)	0.6684 (0.4450)
1.50	2.1338 (1.4413)	1.4365 (0.7841)	1.1120 (0.5657)	0.9114 (0.4562)
2.00	2.3305 (1.3414)	1.6362 (0.7348)	1.3052 (0.5344)	1.0971 (0.4335)
$p = 0.90$				
a/b	0.25	0.50	0.75	1.00
0.10	1.6692 (7.3656)	0.7927 (2.5034)	0.5180 (1.3191)	0.3854 (0.8434)
0.25	2.9635 (9.5886)	1.6100 (3.9834)	1.1186 (2.3173)	0.8630 (1.5742)
0.50	4.0250 (9.4214)	2.4597 (4.4814)	1.8208 (2.8446)	1.4635 (2.0527)
0.75	4.5943 (8.7414)	2.9917 (4.3671)	2.3015 (2.8851)	1.8998 (2.1485)

waiting time distributions. Thus, it is a desirable question in modeling to know whether a given distribution is infinitely divisible or not. To settle this question, we recall that according to Steutel and van Harn (2004, pp. 56), if p_x , $x \in \mathbb{N}_0$, is infinitely divisible, then $p_x \leq e^{-1}$ for all $x \in \mathbb{N}$. However, e.g., in a $DBR(2.25, 1.5, 0.8)$ distribution we see that $p_2 = 0.413 > e^{-1} = 0.367$. Therefore, in general, $DBR(a, b, p)$ distributions are not infinitely divisible. In addition, since the classes of self-decomposable and stable distributions, in their discrete concepts, are subclasses of infinitely divisible distributions, we conclude that a DBR distribution can be neither self-decomposable nor stable in general.

2.3 Order Statistics

Order statistics are among the most fundamental tools in non-parametric statistics and inference. They enter the problems of estimation and hypoth-

esis testing in a variety of ways.

Nekoukhou and Bidram (2015a) obtained some general relations for the probability mass and distribution functions of order statistics of the class of discrete BG distributions. In the present section, we establish such relations for order statistics of DBR distributions. More precisely, let $F_i(y; a, b, p)$ and $f_i(y; a, b, p)$ be the cdf and pmf of the i -th order statistic of a random sample of size n drawn from a $DBR(a, b, p)$ distribution. One can show that the cdf of the i -th order statistic is of the form

$$F_i(y; a, b, p) = \sum_{k=i}^n \sum_{j=0}^{n-k} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \dots \sum_{m_{k+j}=0}^{\infty} \delta_{k,j} F_{GDR}(y; p, a(k+j) + \sum_{r=1}^{k+j} m_r), \quad (9)$$

where

$$\delta_{k,j} = \frac{\binom{n}{k} \binom{n-k}{j} (-1)^{j+\sum_{r=1}^{k+j} m_r} \{\Gamma(a+b)\}^{k+j}}{\{\Gamma(a)\}^{k+j} \prod_{r=1}^{k+j} \Gamma(b-m_r) m_r! (a+m_r)}.$$

It should be noted that for integer $b > 0$ the infinite sums in Equation (9) stop at $b - 1$; see Nekoukhou and Bidram (2015a). As we see, the cdf of the i th order statistic of a DBR distribution is an infinite linear combination of $GDR(p, a(k+j) + \sum_{r=1}^{k+j} m_r)$ distribution functions. In particular, when $b = 1$, one can show that Equation (9) reduces to the cdf of the i th order statistic of a $GDR(p, a)$ distribution given by Nekoukhou and Bidram (2015b).

The corresponding pmf of the i -th order statistic for a non-negative integer y , $f_i(y) = F_i(y) - F_i(y - 1)$, is also given by

$$f_i(y; a, b, p) = \sum_{k=i}^n \sum_{j=0}^{n-k} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \dots \sum_{m_{k+j}=0}^{\infty} \delta_{k,j} f_{GDR} \left(y; p, a(k+j) + \sum_{r=1}^{k+j} m_r \right),$$

where f_{GDR} denotes the pmf of a GDR distribution given by (6)-(7). In view of the fact that $f_i(y; a, b, p)$ is an infinite linear combination of $GDR(p, a(k+j) + \sum_{r=1}^{k+j} m_r)$ pmfs, we may obtain some properties of order statistics, such as their moments, from the corresponding GDR distributions.

2.4 Estimation

Let Y be a random variable from the DBR distribution and $\boldsymbol{\theta} = (a, b, p)^T$ be the parameters vector. The likelihood function for a single observation y is given by

$$\ell(\boldsymbol{\theta}) = \sum_{i=0}^{\infty} \omega_i(a, b) \left\{ \left(1 - p^{(y+1)^2}\right)^{a+i} - \left(1 - p^{y^2}\right)^{a+i} \right\}. \quad (10)$$

The first derivatives of the likelihood function with respect to the components of $\boldsymbol{\theta}$ are given by

$$\begin{aligned} \frac{\partial \ell}{\partial a} &= \sum_{i=0}^{\infty} \omega_i(a, b) \left\{ \psi(a+b) - \psi(a) - \frac{1}{a+i} \right\} \\ &\quad \times \left\{ \left(1 - p^{(y+1)^2}\right)^{a+i} - \left(1 - p^{y^2}\right)^{a+i} \right\} \\ &\quad + \omega_i(a, b) \left\{ \left(1 - p^{(y+1)^2}\right)^{a+i} \ln \left(1 - p^{(y+1)^2}\right) - \left(1 - p^{y^2}\right)^{a+i} \ln \left(1 - p^{y^2}\right) \right\}, \end{aligned}$$

$$\frac{\partial \ell}{\partial b} = \sum_{i=0}^{\infty} \omega_i(a, b) (\psi(a+b) - \psi(b-i)) \left\{ \left(1 - p^{(y+1)^2}\right)^{a+i} - \left(1 - p^{y^2}\right)^{a+i} \right\}$$

and

$$\begin{aligned} \frac{\partial \ell}{\partial p} &= \sum_{i=0}^{\infty} \omega_i(a, b)(a+i) \left\{ \left(1 - p^{y^2}\right)^{a+i-1} p^{y^2-1} y^2 \right. \\ &\quad \left. - \left(1 - p^{(y+1)^2}\right)^{a+i-1} p^{(y+1)^2-1} (y+1)^2 \right\}, \end{aligned}$$

where $\psi(\cdot)$ is the well-known digamma function.

Now, let y_1, y_2, \dots, y_n be n observations of a random sample drawn from a $DBR(a, b, p)$ distribution. The total likelihood function is obtained by $\ell_n(\boldsymbol{\theta}) = \prod_{k=1}^n \ell_k(\boldsymbol{\theta})$, where $\ell_k(\boldsymbol{\theta})$; $k = 1, 2, \dots, n$, is given by Equation (10). The maximum likelihood estimate (MLE) of $\boldsymbol{\theta}$, say $\hat{\boldsymbol{\theta}}$, is obtained by solving the nonlinear equation

$$\mathbf{U}_n = (\partial \ell_n / \partial a, \partial \ell_n / \partial b, \partial \ell_n / \partial p)^T = \mathbf{0},$$

using a numerical method.

The Fisher information matrix is given by

$$\mathbf{I}(\boldsymbol{\theta}) = [I_{\theta_i, \theta_j}]_{3 \times 3}; \quad i, j = 1, 2, 3,$$

whose components can be calculated, numerically, by the relation

$$I_{\theta_i, \theta_j} = E\left(-\frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j}\right); \quad i, j = 1, 2, 3.$$

The total Fisher information matrix is given by $\mathbf{I}_n(\boldsymbol{\theta}) = n\mathbf{I}(\boldsymbol{\theta})$ which can be approximated by

$$\mathbf{I}_n(\hat{\boldsymbol{\theta}}) \approx \left[-\frac{\partial^2 \ell_n}{\partial \theta_i \partial \theta_j} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}\right]_{3 \times 3} \quad i, j = 1, 2, 3,$$

and, therefore, under conditions given in Ferguson (1996), $\hat{\boldsymbol{\theta}}$ has an asymptotic normal distribution as $N_3(\boldsymbol{\theta}, \mathbf{I}_n(\hat{\boldsymbol{\theta}})^{-1})$, or equivalently, $\mathbf{I}_n(\hat{\boldsymbol{\theta}})^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T$ has a multivariate standard normal distribution. Asymptotic normal distributions are usually used for constructing approximate confidence intervals, confidence regions, and testing hypotheses of the parameters. For example, an asymptotic confidence interval with confidence coefficient $1 - \gamma$ for parameters θ_j ; $j = 1, 2, 3$, is computed by $(\hat{\theta}_j \mp z_{1-\gamma/2} \sqrt{I_{\hat{\theta}_j, \hat{\theta}_j}})$, where $I_{\hat{\theta}_j, \hat{\theta}_j}$ is the j -th diagonal element of $\mathbf{I}_n(\hat{\boldsymbol{\theta}})^{-1}$ and $z_{1-\gamma/2}$ is the $(1 - \gamma/2)$ -th quantile of the standard normal distribution.

2.5 A Simulated Example

It is evident that if V is a beta random variable with parameters a and b , then the cdf of the random variable $X = G^{-1}(V)$ coincides with cdf (1). Hence, for a Rayleigh random variable with cdf (4), we have

$$X = \left(-\frac{1}{\theta} \log(1 - V)\right)^{\frac{1}{2}},$$

which helps to generate data from a BR distribution. Thus, we can consider $[X]$ in order to generate data from a $DBR(a, b, p)$ distribution. Table 2 presents the maximum likelihood estimates of $\boldsymbol{\theta} = (a, b, p)^T$ and also contains their standard errors for the sample sizes $n = 100, 200, 500$ and

Table 2. MLEs of the DBR parameters and related standard errors.

n	\hat{a}	\hat{b}	\hat{p}	$\hat{SE}(\hat{a})$	$\hat{SE}(\hat{b})$	$\hat{SE}(\hat{p})$
$(a, b, p) = (1, 2, 0.25)$						
100	0.9032	1.7134	0.2112	0.6214	0.5453	0.2214
200	0.8641	1.9121	0.2398	0.4547	0.4161	0.1241
500	0.9736	1.8931	0.2478	0.2121	0.2365	0.0712
1000	0.9897	1.9721	0.2514	0.1015	0.1151	0.0301
$(a, b, p) = (2, 3, 0.5)$						
100	1.8878	2.8658	0.4546	0.6119	0.5447	0.2120
200	1.9621	2.8979	0.4889	0.3326	0.2882	0.1523
500	2.0368	2.9813	0.5107	0.2559	0.1812	0.0653
1000	1.9978	3.0151	0.5097	0.1116	0.1110	0.0209
$(a, b, p) = (2, 2, 0.75)$						
100	1.9297	2.2524	0.7314	0.5337	0.4879	0.2147
200	1.9633	1.8925	0.7442	0.3965	0.2295	0.1221
500	1.9845	2.1013	0.7506	0.1412	0.1511	0.0748
1000	2.1412	2.0412	0.7521	0.1034	0.1101	0.0311

1000 as a simulated example. Standard errors are attained by means of asymptotic covariance matrix of the MLEs of DBR parameters when the Newton-Raphson procedure converges in, e.g., MATLAB software.

Figure 3 is plotted for more values of sample sizes and for $(a, b, p) = (1, 2, 0.25)$ to show how the standard errors vary with respect to n . As we see, the standard errors of the estimated parameters decrease to zero as $n \rightarrow \infty$.

2.6 Stress-strength Parameter

$R = P(X > Y)$ is a measure of component reliability and its estimation problem when X and Y are independent and follow a specified common distribution has been discussed widely in the literature. Suppose that the random variable X is the strength of a component which is subjected to a random stress Y . Estimation of R when X and Y are independent and identically distributed following a well-known distribution has been considered in the literature. However, data may be discrete by nature, for example, the

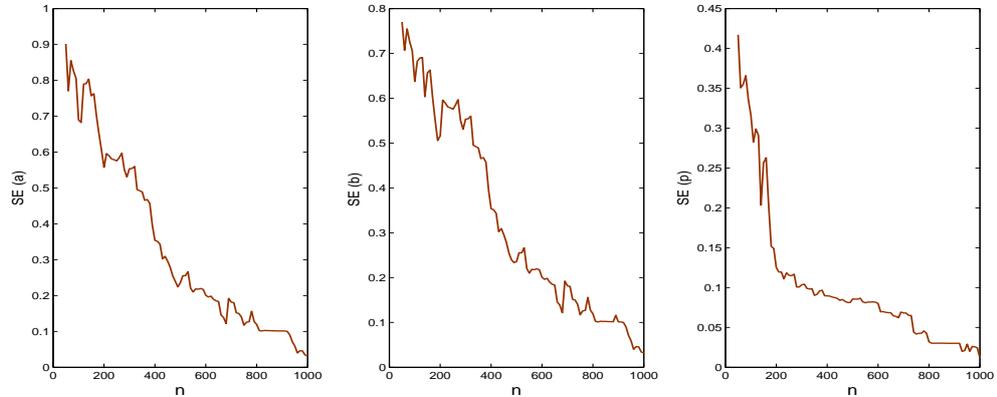


Figure 3. Standard errors of the MLEs of $(a, b, p) = (1, 2, 0.25)$ for $n=50, 100, 150, \dots, 1000$.

number of events occurring in a certain spatial or temporal interval. When discrete data are derived from discretization, instead of numerical measurements on X and Y , they are presented in a form of ordered categories.

Here, we consider the estimation of R in a DBR model, in view of the fact that a relatively small amount of work is devoted to discrete or categorical data.

The stress-strength parameter, in discrete case, is defined as

$$R = P(X > Y) = \sum_{x=0}^{\infty} f_X(x)F_Y(x),$$

where f_X and F_Y denote the pmf and cdf of the independent discrete random variables X and Y , respectively. Now, let $X \sim DBR(\boldsymbol{\theta}_1)$ and $Y \sim DBR(\boldsymbol{\theta}_2)$, where $\boldsymbol{\theta}_1 = (a_1, b_1, p)^T$ and $\boldsymbol{\theta}_2 = (a_2, b_2, p)^T$. Using Equations (5) and (8), we obtain

$$R = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} w_i(a_1, b_1)w_j(a_2, b_2) \times \sum_{x=0}^{\infty} \left\{ (1 - p^{(x+1)^2})^{(a_1+i)} - (1 - p^{x^2})^{(a_1+i)} \right\} (1 - p^{(x+1)^2})^{(a_2+j)}. \quad (11)$$

Now, assume that x_1, x_2, \dots, x_n and y_1, y_2, \dots, y_m are independent observations from $X \sim DBR(\boldsymbol{\theta}_1)$ and $Y \sim DBR(\boldsymbol{\theta}_2)$, respectively. The total

Table 3. Results of ten shots fired.

Numbers	1	2	3	4	5	6	7	8	9
Frequency	2	4	10	22	26	18	12	4	2

likelihood function is $\ell_R(\boldsymbol{\theta}^*) = \ell_n(\boldsymbol{\theta}_1)\ell_m(\boldsymbol{\theta}_2)$, where $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$. The score vector is given by

$$U_R(\boldsymbol{\theta}^*) = (\partial\ell_R/\partial a_1, \partial\ell_R/\partial a_2, \partial\ell_R/\partial b_1, \partial\ell_R/\partial b_2, \partial\ell_R/\partial p),$$

and the MLE of $\boldsymbol{\theta}^*$, say $\hat{\boldsymbol{\theta}}^*$, may be attained from the nonlinear equation $U_R(\hat{\boldsymbol{\theta}}^*) = \mathbf{0}$. Thus, by inserting the MLEs in Equation (11) the stress-strength parameter R will be estimated.

2.7 Application

In this section, the DBR distribution will be examined with a real data set, given by Consul and Jain (1973). This data set in which considers the results of ten shots fired from a rifle at each of 100 targets has been displayed in Table 3. Gómez-Déniz (2010) used these data in order to study its generalization of the geometric distribution.

Here, we compare the capacity of the DBR distribution with some of its rival models in discrete data modeling. The maximum likelihood method is used to obtain the estimates of the DBR parameters. Comparing the DBR distribution with its rival models is performed by using the Akaike information criterion (AIC) and Kolmogrov-Smirnov (K-S) test statistic. Table 4 indicates the fitting computations which consists of the MLEs, AICs and the values of K-S test statistic determined by the fitting models.

The discrete double generalized exponential (DDGE) and discrete beta-exponential (DBE) distributions, which are special sub-models of the DBGE distribution, and also the discrete Weibull (DW) distribution of Nakagawa and Osaki (1975) are compared with the DBR distribution. In addition, the two-parameter discrete generalized exponential distribution of a second type (DGE_2) of Nekoukhou et al. (2013) which is a generalization of the geometric distribution, $p_y = (1-p)p^y$, is also fitted to these data. The EDW distribution and its special case, GDR distribution, and also generalized geometric (GG) distribution of Gómez-Déniz (2010) are other rival models. The distributions in Table 4 are considered, briefly, in the Appendix.

According to the AICs and the values of K-S test statistics in Table 4, it

Table 4. Summary of computations.

Models	MLEs	AIC	K-S statistic
DBR	$(\hat{a}, \hat{b}, \hat{p}) = (1.52, 1.01, 0.97)$	48.6272	0.1807
EDW	$(\hat{\alpha}, \hat{\gamma}, \hat{p}) = (3.72, 1.05, 0.99)$	54.1788	0.2814
GDR	$(\hat{p}, \hat{\gamma}) = (0.94, 3.29)$	52.3959	0.4318
DBGE	$(\hat{a}, \hat{b}, \hat{\alpha}, \hat{p}) = (1.37, 1.53, 1.37, 0.89)$	55.4358	0.3192
DBE	$(\hat{a}, \hat{b}, \hat{p}) = (1.22, 0.99, 0.90)$	57.6651	0.4253
DDGE	$(\hat{b}, \hat{\alpha}, \hat{p}) = (1.32, 1.11, 0.94)$	59.7067	0.3450
DGE_2	$(\hat{a}, \hat{p}) = (18.51, 0.53)$	53.9190	0.2424
DW	$(\hat{\alpha}, \hat{p}) = (3.80, 0.99)$	51.9968	0.2819
GG	$(\hat{\alpha}, \hat{\theta}) = (394.75, 0.33)$	52.3626	0.3004
Geometric	$\hat{p} = 0.83$	50.6606	0.3103

seems that the DBR model gives a satisfactory fit to these real data. Figures 4 and 5 indicate the fitted models and corresponding probability-probability (p-p) plots of the fitted models, respectively. As we see, the figures also confirm the goodness of fit of the DBR distribution.

One can construct approximate confidence intervals for the parameters of the DBR model. Indeed, such confidence intervals are attained by means of asymptotic covariance matrix of the MLEs of DBR parameters when the Newton-Raphson procedure converges. For instance, 95% asymptotic confidence intervals for DBR parameters are calculated as $a \in (1.5214 \mp 0.8449)$, $b \in (1.0115 \mp 0.7497)$ and $p \in (0.9662 \mp 0.2103)$.

3 Conclusions

We introduced a new discrete distribution, called *discrete beta-Rayleigh* (DBR) distribution, motivated by the fact that it provides greater flexibility in order to analyze various discrete data. Indeed, the DBR distribution is appropriate for modeling both over- and under-dispersed data and its hazard rate function can be increasing, bathtub-shaped and unimodal. That is, DBR distributions can be used as improved models for analyzing failure data in discrete case.

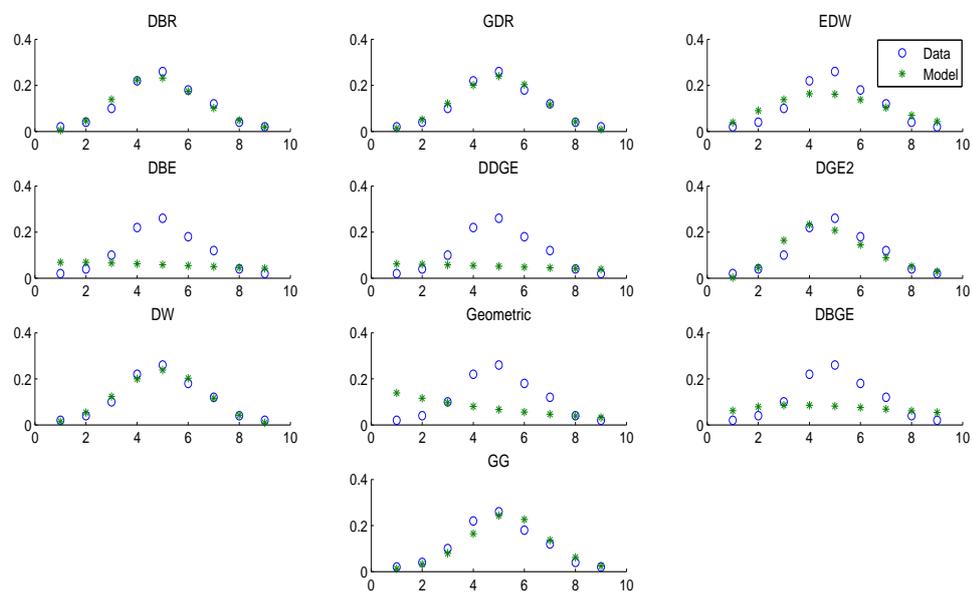


Figure 4. The fitted models.

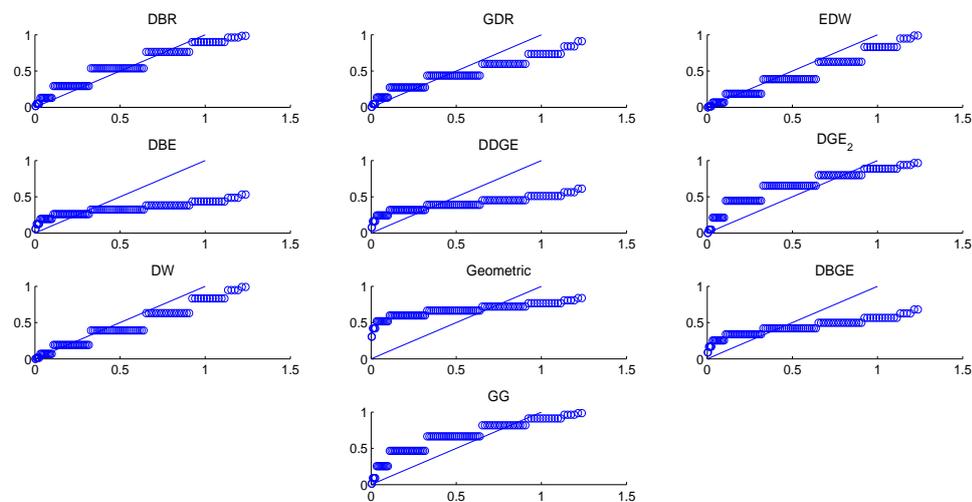


Figure 5. The p-p plots of the fitted models.

Acknowledgement

The author sincerely thanks the anonymous referees for their valuable comments that led to improvement of the paper.

References

- Barreto-Souza, W., Santos, A.H. and Cordeiro, G.M. (2010). The Beta Generalized Exponential Distribution. *J. Stat. Comput. Simul.*, **80**, 159-172.
- Bidram, H., Behboodian, J. and Towhidi, M. (2013). The Beta Weibull-geometric Distribution. *J. Stat. Comput. Simul.*, **83**, 52-67.
- Castellares, F., Montenegro, L.C. and Cordeiro, G.M. (2013). The Beta Log-normal Distribution. *J. Stat. Comput. Simul.*, **83**, 203-228.
- Consul, P.C and Jain, G.C. (1973). A Generalization of the Poisson Distribution. *Technometrics*, **15**, 791-799.
- Cordeiro, G.M., Gomes, A.E., Silva, C.Q. and Ortega, E.M.M. (2013a). The Beta Exponentiated Weibull Distribution. *J. Stat. Comput. Simul.*, **83**, 114-138.
- Cordeiro, G.M., Cristino, C.T., Hashimoto, E.M. and Ortega, E.M.M. (2013b). The Beta Generalized Rayleigh Distribution with Applications to Lifetime Data. *Stat. Papers*, **54**, 133-161.
- Cordeiro, G.M., Silva, G.O. and Ortega, E.M.M. (2013c). The Beta-Weibull Geometric Distribution. *Statistics*, **47**, 817-834.
- Cordeiro, G.M. and Lemonte, A.J. (2011a). The Beta Laplace Distribution. *Stat. & Prob. Lett.*, **81**, 973-982.
- Cordeiro, G.M. and Lemonte, A.J. (2011b). The Beta-half-Cauchy Distribution. *J. Prob. Stat.*, **2011**, 1-18.
- Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal Distribution and Its Applications. *Comm. Stat. Theo. Meth.*, **31**, 497-512.
- Ferguson, T.S. (1996). *A Course in Large Sample Theory*. Chapman and Hall, London.
- Gómez-Déniz E. (2010). Another Generalization of the Geometric Distribution, *Test*, **19**, 399-415.
- Gupta, R.D. and Kundu, D. (1999). Generalized Exponential Distributions. *Aust. NZ. J. Stat.*, **41**, 173-188.

Lee, H. and Cha, J.H. (2015). On Two General Classes of Discrete Bivariate Distributions. *The American Statistician*, DOI: 10.1080/00031305.2015.1044564.

Lemonte, A.J. and Cordeiro, G.M. (2013). An Extended Lomax Distribution. *Statistics*, **47**, 800-816.

Mahmoudi, E. (2011). The Beta Generalized Pareto Distribution with Application to Lifetime Data. *Math. Comput. Simul.*, **81**, 2414-2430.

Morais, A.L., Cordeiro, G.M. and Audrey, H.M.A. (2013). The Beta Generalized Logistic Distribution. *Brazilian J. Prob. Stat.*, **27**, 185-200.

Nakagawa, T. and Osaki, S. (1975). The Discrete Weibull Distribution. *IEEE Transact. Reliab.*, **24**, 300-301.

Nekoukhou, V., Alamatsaz, M.H. and Bidram, H. (2013). Discrete Generalized Exponential Distribution of a Second Type. *Statistics*, **47**, 876-887.

Nekoukhou, V. and Bidram, H. (2015a). A New Four-Parameter Discrete Distribution with Bathhtub and Unimodal Failur Rate. *J. Applied Statistics*, **42**, 2654-2670.

Nekoukhou, V. and Bidram, H. (2015b). The Exponentiated Discrete Weibull Distribution. *SORT*, **39**, 127-146.

Roy, D. (2004). Discrete Rayleigh Distribution. *IEEE Trans. Reliab*, **53**, 255-260.

Silva, G.O., Ortega, E.M.M. and Cordeiro, G.M. (2010). The Beta Modified Weibull Distribution. *Life. data Anal.*, **16**, 409-430.

Singla, N., Jain, K. and Sharma, S.K. (2012). The Beta Generalized Weibull Distribution: Properties and Applications. *Reli. Engin. Sys. Saf.*, **102**, 5-15.

Stutel, F.W. and van Harn, K. (2004). *Infinite Divisibility of Probability Distributions on the Real Line*. New York: Marcel Dekker.

Appendix

In this section we consider the rival models indicated in Table 4.

1) The exponentiated discrete Weibull (EDW) distribution of Nekoukhou and Bidram (2015b), for $y \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$, has the following pmf

$$\begin{aligned} f(y; p, \alpha, \gamma) &= \{1 - p^{(y+1)^\alpha}\}^\gamma - \{1 - p^{y^\alpha}\}^\gamma \\ &= \sum_{j=1}^{\infty} (-1)^{j+1} \binom{\gamma}{j} \{p^{jy^\alpha} - p^{j(y+1)^\alpha}\}, \end{aligned}$$

where $0 < p < 1$, $\alpha > 0$, $\gamma > 0$ and $\binom{\gamma}{j} = \frac{\Gamma(\gamma+1)}{\Gamma(\gamma+1-j)j!}$. For integer $\gamma > 0$, the above sum stops at γ .

2) Discrete Weibull distribution of Nakagawa and Osaki (1975), with pmf

$$f(y; p, \alpha) = (1 - p^{(y+1)^\alpha}) - (1 - p^{y^\alpha}), \quad (12)$$

is a special case of the EDW distribution, when $\gamma = 1$.

3) Discrete generalized exponential distribution of a second type, $DGE_2(\gamma, p)$, of Nekoukhou et al. (2013) with pmf

$$f(y; p, \gamma) = \{1 - p^{(y+1)}\}^\gamma - \{1 - p^y\}^\gamma,$$

can be considered as another special case of the EDW distribution, by choosing $\alpha = 1$.

4) The pmf of a discrete beta generalized exponential (DBGE) distribution proposed by Nekoukhou and Bidram (2015a), with parameters $a > 0$, $b > 0$, $\alpha > 0$ and $0 < p < 1$, is of the form

$$\begin{aligned} f(y; a, b, \alpha, p) &= \sum_{i=0}^{\infty} \omega_i(a, b) \{(1 - p^{y+1})^{\alpha(a+i)} - (1 - p^y)^{\alpha(a+i)}\} \\ &= \sum_{i=0}^{\infty} \omega_i(a, b) \sum_{j=1}^{\infty} (-1)^{j+1} \binom{\alpha(a+i)}{j} p^{jy} (1 - p^j), \end{aligned} \quad (13)$$

where $y \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$. The above pmf is, in fact, a linear combination of the probability mass functions of the DGE_2 distributions.

A discrete double generalized exponential (DDGE) distribution of Nekoukhou and Bidram (2015a) is a special case of the DBGE distribution, when $a = 1$. Moreover, a discrete beta-exponential (DBE) distribution of Nekoukhou and Bidram (2015a) is another sub-model of DBGE distribution, when $\alpha = 1$.

5) A generalization of the geometric distribution has been introduced by Gómez-Déniz (2010). The generalized geometric (GG) distribution of Gómez-Déniz (2010), for $y \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$, has the following pmf

$$f(y; \alpha, \theta) = \frac{\alpha \theta^y (1 - \theta)}{\{1 - (1 - \alpha) \theta^{y+1}\} \{1 - (1 - \alpha) \theta^y\}}$$

in which $\alpha > 0$ and $0 < \theta < 1$ are the model parameters.

Vahid Nekoukhou

Department of Statistics,

Khansar Faculty of Mathematics and Computer Science,

University of Isfahan, Khansar Unit,

Isfahan, Iran.

email: *v.nekoukhou@gmail.com*