

REVIEW

Open Access



Generating discrete analogues of continuous probability distributions-A survey of methods and constructions

Subrata Chakraborty

Correspondence:
subrata_arya@yahoo.co.in
Department of Statistics, Dibrugarh
University, Dibrugarh 786004,
Assam, India

Abstract

In this paper a comprehensive survey of the different methods of generating discrete probability distributions as analogues of continuous probability distributions is presented along with their applications in construction of new discrete distributions. The methods are classified based on different criterion of discretization.

Keywords: Discrete analogue; Reliability function; Hazard rate function; Competing risk; Exponentiated distribution; Maximum entropy; Discrete pearson; T-X method

1. Introduction

Sometimes in real life it is difficult or inconvenient to get samples from a continuous distribution. Almost always the observed values are actually discrete because they are measured to only a finite number of decimal places and cannot really constitute all points in a continuum. Even if the measurements are taken on a continuous scale the observations may be recorded in a way making discrete model more appropriate.

In some other situation because of precision of measuring instrument or to save space, the continuous variables are measured by the frequencies of non-overlapping class interval, whose union constitutes the whole range of random variable, and multinomial law is used to model the situation.

In categorical data analysis with econometric approach existence of a continuous unobserved or latent variable underlying an observed categorical variable is presumed. Categorical variable is the observed as different discrete values when the unobserved continuous variable crosses a threshold value. Therefore, the inference is based on observed discrete values which are only indicative of the intervals to which unobserved continuous variable belongs but not its true values. Hence this is a case where one makes use of a discretization of the underlying continuous variable.

In survival analysis the survival function may be a function of count random variable that is a discrete version of underlying continuous random variable. For example the length of stay in an observation ward is counted by number of days or survival time of leukemia patients counted by number of weeks. From these examples it is clear that the continuous life time may not necessarily always be measured on a continuous scale but may often be counted as discrete random variables.

More over often the continuous failure time data generated from a complex system poses more derivational problem than that of a discrete version of the underlying continuous one. Despite these discrete life time distributions played only a marginal role in reliability analysis. Therefore, there is a need to focus on more realistic discrete life time distributions (Rezaei Roknabadi *et al.* 2009). That is discretization of a continuous lifetime model is an interesting and intuitively appealing approach to derive a discrete lifetime model corresponding to the continuous one (Lai 2013).

From the above discussion it can be inferred that many a times in real world the original variables may be continuous in nature but discrete by observation and hence it is reasonable and convenient to model the situation by an appropriate discrete distribution generated from the underlying continuous models preserving one or more important traits of the continuous distribution.

Deriving discrete analogues (Discretization) of continuous distribution has drawn attention of researchers. In recent decades a large number of research papers dealing with discrete distribution derived by discretizing a continuous one have appeared in a scattered manner in existing statistical literatures.

There are several ways to derive discrete distribution from continuous ones. In the current published literature we could find only two papers that dealt with surveys of discrete analogues of continuous distributions though in a limited manner. These are Bracquemond and Gaudoin (2003) who devoted a section on discrete life time distributions derived from continuous one in their survey on discrete life time distributions and Lai (2013) who presented construction of discrete lifetime distributions from continuous one in his paper concerning issues of construction of discrete life time distribution

With above background the main motivation of this article is to present a comprehensive method-wise survey of the different techniques of discretization of continuous distributions, with examples of their applications in construction of discrete analogues.

In the section 2 of this article discretization of continuous distributions are discussed method wise including composite methods, which comprise two stages using two different methods in separate subsections. In section 3 a discussion on the discretization highlighting its need, limitations and also a final conclusion is presented. Throughout the paper continuous random variable to be discretized is denoted by X while its discrete analogue by Y and with respect to discrete life time characteristics Kemp's (2004) convention is followed.

2. Discrete analogues

A continuous random variable may be characterized either by its probability density function (pdf), moment generating function (mgf), moments, hazard rate function etc. Basically construction of a discrete analogue from a continuous distribution is based on the principle of preserving one or more characteristic property of the continuous one.

The various methods by which discrete analogue Y of a continuous random variable X may be constructed can be classified as follows:

- I. Difference equation analogues of Pearsonian differential equation.
- II. Probability mass function (pmf) of Y retains the form of the pdf of X and support of Y is determined from full range of X .
- III. Pmf of Y retains the form of the pdf of X and support of Y is determined from a subset of the range of X .
- IV. Survival function (sf) of Y retains the form of the survival function of X and support of Y is determined from full range of X .
- V. Cumulative distribution function (cdf) of Y retains the form of the cdf of X and support of Y is determined from a subset of the range of X .
- VI. Hazard (failure) rate function of Y retains the form of the hazard (failure) rate function of X .
- VII. Moments of Y and X up to a certain order coincides.
- VIII. Any interval domain, any theoretically possible mean–variance pair for Y
- IX. Two stage composite methods

2.1 Discrete analogue of pearsonian system

Pearson (1895) starting with the difference equation

$$\frac{p_k - p_{k-1}}{p_k} = \frac{k - a}{b_0 + b_1k + b_2k^2} \tag{1}$$

defined the celebrated Pearsonian system of continuous distributions with pdf satisfying the differential equation

$$\frac{1}{f} \frac{df}{dx} = \frac{a + x}{c_0 + c_1x + c_2x^2}$$

Though Pearson himself did not pursue the development of a discrete analogue of his continuous system, the difference Eq. in (1) was used by Carver (1919, 1923). But he too did not attempt a thorough examination of the discrete distributions arising from Eq. (1).

Katz (1945, 1946, 1948, 1965) developed a discrete analogue of the Pearsonian system of continuous distributions by using the relationship

$$p_{k+1} = \frac{a + bk}{1 + k} p_k, \quad k = 0, 1, 2, \dots$$

Where $p_k = P(Y = k)$. The main motivation of Katz was to discriminate between binomial, Poisson and negative binomial distributions. Some notable related developments following Katz are as follows:

Ord (1967a, b, c, 1968) discussed discrete analogue of the Pearson continuous system by using the following difference equation:

$$\frac{p_k - p_{k-1}}{p_k} = \frac{a - k}{(a + b_0) + (b_1 - 1)k + b_2k(k - 1)}$$

Discrete t distribution: Ord (1968) also derived discrete analogue of various types of Pearsonian distributions. In particular, proposed discrete t distribution as a particular case of type VII distribution. The pmf of his discrete t was

$$p_k = \alpha_m / \prod_{j=1}^m \{(j+k+a)^2 + b^2\}, k \in \mathbf{Z}$$

Where $0 \leq a \leq 1$, $0 < b < \infty$, m is a non negative integer, and α_m is a normalizing constant. (see Johnson *et al.* 2005 for detail).

Gurland and Tripathi (1975) and Tripathi and Gurland (1977), studied the extended Katz family that satisfies the probability recurrence relation

$$p_{k+1} = \frac{a+bk}{c+k} p_k, k = 0, 1, 2, \dots$$

Sundt and Jewell (1981) investigated a family of distributions satisfying probability recurrence relation

$$p_{k+1} = \frac{a+b+ak}{1+k} p_k, k = 0, 1, 2, \dots \quad (\text{See also Willmot, 1988})$$

2.2 Methodology-II

In this method the pmf of the discrete random variable Y is derived as an analogue of the continuous random variable X with pdf $f(x)$, $-\infty < x < \infty$ as

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

The distribution generated using this technique may not always have a compact form due to the normalizing constant.

2.2.1 Good distribution

The first trace of this type of construction is seen in Good distribution (Good 1953) having pmf

$$P(Y = k) = q^k k^a / \sum_{j=1}^{\infty} q^j j^a, k = 1, 2, \dots; \alpha \in \mathbf{R} \quad (3)$$

When $a > -1$, this distribution can be derived as a discrete analogue of gamma distribution by considering

$$f(x) = \frac{1}{\theta^\beta \Gamma(\beta)} x^{\beta-1} e^{-x/\theta} \text{ in Eq. (2) and replacing } e^{-1/\theta} = q \text{ and } \beta - 1 = a.$$

This distribution was applied to model the population frequencies of species and the estimation of population parameters.

This distribution was extensively studied by Kulasekara and Tonkyn (1992) and Doray and Luong (1997).

The distribution in Eq. (3) is a special case of **Hurwitz-Lerch Zeta** Distribution (Zornig and Altmann 1995; Doray and Luong 1997; Gupta *et al.* 2008). For Hurwitz-Lerch Zeta functions see Gradshteyn and Ryzhik (2000). (see also section 11.2.20 of Johnson *et al.* 2005).

Another related distribution is the **discrete Pareto** distribution, also known as the Riemann zeta distribution (see page 527, Johnson *et al.* 2005).

Jamjoom (2013) investigated order statistics of the above distribution (also investigated by Alhazzani 2012) both in the “*i.i.d.*” and “*identical but not independent*” cases.

2.2.2 General Dirichlet distribution

Using general Dirichlet series, Siromoney (1964) studied the general Dirichlet Series distribution with pmf

$$P(Y = k) = a_k e^{-\lambda_k \theta} / \sum_{j=1}^{\infty} a_j e^{-\lambda_j \theta}, \quad k = 1, 2, \dots,$$

Various distributions were seen as particular cases as follows:

- For $\lambda_k = \log(k)$, the distribution reduces to Dirichlet series distribution with pmf

$$P(Y = k) = a_k k^{-\theta} / \sum_{j=1}^{\infty} a_j j^{-\theta}, \quad k = 1, 2, \dots$$

- For $a_k = a$ and $\lambda_k = \log(k)$, the distribution reduces to Zeta distribution with pmf

$$P(Y = k) = k^{-\theta} / \zeta(\theta); \quad k = 1, 2, \dots \quad \text{where } \zeta(\theta) = \sum_{j=1}^{\infty} j^{-\theta} \text{ is the Riemann Zeta}$$

function.

- Putting $\lambda_k = k$, $e^{-\theta} = \alpha$, gives power series distribution with pmf

$$P(Y = k) = a_k \alpha^k / \sum_{j=0}^{\infty} a_j \alpha^j, \quad k = 0, 1, 2, \dots$$

- For $a_k = k^a$, $e^{-\theta} = q$ and $\lambda_k = -k$ reduces to Good distribution

- For $a_k = k$, $\lambda_k = -k^2$ and $\theta = 1/2$ discrete Pearson distribution mentioned in Byers and Shenton (1994) having pmf

$$P(Y = k) \propto k e^{-k^2/2}, \quad k = 1, 2, \dots$$

- The discrete Pearson III distribution of Haight (1957) is a special

case with $a_k = (k + \nu)^a$ and $\lambda_k = k$ (page 532, Johnson *et al.* 2005) with pmf

$$P(Y = k) = e^{-k\theta} / [(k + \nu)^a \sum_{j=0}^{\infty} \{e^{-j\theta} / (j + \nu)^a\}], \quad k = 0, 1, 2, \dots$$

Siromoney (1964) applied this distribution to model frequency distribution of the length of wet spells during the period 1932-62 in a place called Tambaram in southern India.

2.2.3 Discrete normal distribution

A discrete normal distribution was investigated by many authors including Lisman and Van Zuylen (1972); Kemp (1997); Liang (1999); and Szablowski (2001).

The discrete normal distribution was derived as a discrete analogue of the normal distribution (Kemp 1997) by considering $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$, in Eq. (2) and substituting, $e^{(1-2\mu)/2\sigma^2} = \lambda$ and $e^{-1/\sigma^2} = q$. The resulting pmf is given by

$$P(Y = k) = \lambda^k q^{k(k-1)/2} / \sum_{j=-\infty}^{\infty} \lambda^j q^{j(j-1)/2}, \quad k = 0, \pm 1, \pm 2, \dots \tag{4}$$

This distribution is characterized by maximum entropy for specified mean and variance, and integer support on $(-\infty, +\infty)$. It can be derived as the distribution of the difference of two related **Heine** distribution (Benkherouf and Bather 1988; see also section 4.12.6 of Johnson *et al.* 2005 and references therein)

Weighted distribution of discrete normal with parameter (λ, q) with weight function of the form π^x is again discrete normal $(\pi\lambda, q)$.

For $\lambda = q^{1/2}$ and $q = e^{-2\beta}$ the pmf in Eq. (4) reduces to that of Das Gupta (1993) version of **discrete normal** distribution.

$$P(Y = k) = q^{k/2} e^{-\beta k(k-1)} / \sum_{j=-\infty}^{\infty} q^{j/2} e^{-\beta j(j-1)}, \quad k = 0, \pm 1, \pm 2, \dots$$

The distribution is log concave and unimodal like normal distribution.

Harris et al. (2001) *applied this distribution in dynamic analysis of rural retail establishment count data.*

2.2.4 Discrete exponential distribution

Sato *et al.* (1999) proposed discrete exponential distribution having similar looking structure starting with the continuous exponential distribution having pdf

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0, \alpha > 0, \lambda > 0.$$

The pmf of their discrete exponential distribution was

$$P(Y = k) = (1 - e^{-\lambda}) e^{-\lambda k}, \quad k = 0, 1, 2, \dots \tag{5}$$

This is the geometric distribution with pmf

$$P(Y = k) = (1 - p)p^k, \quad k = 0, 1, 2, \dots, \text{ where } p = e^{-\lambda}.$$

Sato et al. (1999) *applied this distribution to model defect count distribution in semiconductor deposition equipment and defect count distribution per chips.*

It can be easily checked that the pmf in Eq. (5) can be derived as a discrete analogue of exponential distribution by considering $f(x) = \lambda e^{-\lambda x}, x > 0$, in Eq. (2).

2.2.5 Discrete Gamma distribution

Sato *et al.* (1999) also briefly discussed the convolution of their discrete exponential distribution to present a discrete Gamma distribution having pmf

$$P(Y = k) = \frac{\prod_{i=1}^{m-1} (k + i)}{(m-1)!} (1 - e^{-\lambda})^m e^{-\lambda k}, \quad k = 0, 1, 2, \dots$$

This can be easily seen as the negative binomial with pmf

$$P(Y = k) = \binom{m + k - 1}{k} p^k (1 - p)^m, \quad k = 0, 1, 2, \dots \text{ when } p = e^{-\lambda}.$$

Sato et al. (1999) applied this distribution to model defect count distribution in semiconductor deposition equipment and defect count distribution per chips.

2.2.6 Discrete log normal distribution

Considering $f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right]$ in Eq. (2), Bi et al. (2001) proposed a discrete distribution with pmf

$$P(Y = k) = e^{-\frac{(\ln k - \mu)}{2\sigma^2}} / k \sum_{j=1}^{\infty} e^{-\frac{(\ln j - \mu)}{2\sigma^2}}, \quad k = 1, 2, \dots \tag{6}$$

and called it **discrete Gaussian exponential** (DGX) distribution. It is easy to see that Eq. (6) can be derived as a discrete analogue of log normal distribution.

This distribution reduces to a discrete generalised Zipf distribution in limit as $\mu \rightarrow -\infty$ (see Bi et al. 2001) with pmf

$$P(Y = k) \propto \frac{1}{k} \exp\left[-\frac{(\ln k - \mu)}{\sigma^2}\right] \propto k^{-(1 - \mu/\sigma^2)}$$

This distribution was applied to model four extremely skewed count data sets namely Text data from the English Bible, Sales data from a large retailer chain, Telecommunications data customer data from an AT&T service of monthly usage volumes, and Click stream data and browsing behavior of internet users.

2.2.7 Discrete half normal distribution

Kemp (2006) presented a discrete half normal distribution as a maximum entropy distribution for given mean and variance with support 0, 1, 2, ... The pmf is given by

$$P(Y = k) = \theta^k q^{k(k-1)/2} / \sum_{j=0}^{\infty} \theta^j q^{j(j-1)/2}, \quad k = 0, 1, 2, \dots$$

This can be seen as discretization of continuous half normal in the same way as in section 2.2.3. It can arise as a limiting **q-hyper-Poisson-I** (Kemp 2002) distribution and also as a mixture of **Heine** distributions (Benkherouf and Bather 1988).

Khorashadizadeh et al. (2012) referred this distribution as **discrete truncated normal**. For an approximation result on this distribution see Byers and Shenton (1994).

2.2.8 Discrete Laplace (double exponential)

Inusah and Kozubowski (2006) proposed a discrete analogue of Laplace (Double exponential) distribution having pmf

$$P(Y = k) = \frac{1-p}{1+p} p^{|k|}, \quad k = 0, \pm 1, \pm 2, \dots; \quad 0 < p < 1, \tag{7}$$

This distribution inherits many properties of its continuous counterpart namely unimodality, infinite divisibility, maximum entropy distribution for given absolute moment. Also arises as the difference of two *i.i.d.* geometric random variables.

This distribution can be derived as a discrete analogue of Laplace distribution by considering

$$f(x) = \frac{1}{2\sigma} \exp(-|x|/\sigma), \text{ in Eq. (2) and substituting, } e^{-1/\sigma} = p.$$

Inusah and Kozubowski (2006) *applied this distribution in modelling different currency exchange rate data.* Meyer et al. (2013) *applied it for estimating Y-STR haplotype frequencies.*

2.2.9 Discrete Skew Laplace

Considering skew Laplace distribution with pdf

$$f(x) = \frac{1}{\sigma} \frac{m}{1+m^2} \begin{cases} \exp(-xm/\sigma), & x \geq 0 \\ \exp(-x/m\sigma), & x < 0 \end{cases}$$

as the base distribution, a discrete analogue was first proposed by Kozubowski and Inusah (2006). It's pmf is given by

$$P(Y = k) = \frac{(1-p)(1-q)}{1-pq} \begin{cases} p^k, & k = 0, 1, 2, 3, \dots \\ q^{|k|}, & k = 0, -1, -2, -3, \dots \end{cases} \tag{8}$$

Where $e^{-1/\sigma} = p$ and $e^{-1/k\sigma} = q$, $p \in (0, 1)$ and $q \in (0, 1)$.

For $p = q$ Eq. (8) reduces to Eq. (7). Arises as the difference of two independently but not identically distributed geometric random variables.

This distribution was also applied for modeling currency exchange rates.

Another discrete distribution that generalizes the discrete skew Laplace distribution was proposed by Lekshmi and Sebastian (2014). This new **Generalized Discrete Laplace** distribution can be derived as the difference of two independently distributed **negative binomial** (NB) random variables with same dispersion parameter.

2.2.10 Discrete generalized exponential distribution

The generalized exponential distribution of Gupta and Kundu (1999) has pdf

$$f(x) = \alpha\lambda(1-e^{-\lambda x})^{\alpha-1}e^{-\lambda x}, \quad x > 0, \alpha > 0, \lambda > 0.$$

A discrete analogue of this distribution was proposed by Nekoukhou *et al.* (2012) with pmf

$$P(Y = k) = Cp^{k-1}(1-p^k)^{\alpha-1} / \sum_{j=0}^{\infty} \binom{\alpha-1}{j} \frac{(-1)^j}{1-p^{1+j}}, \quad k = 1, 2, 3, \dots$$

Where $C = \sum_{j=0}^{\infty} \binom{\alpha-1}{j} \frac{(-1)^j p^j}{1-p^{1+j}}, \quad e^{-\lambda} = p$

Nekoukhou et al. (2012) *applied this distribution to model rank frequencies of graphemes in a Slavic language called 'Slovene'.*

Among the various distribution described in section 2.2 above, discrete normal in section 2.2.3, discrete half normal in section 2.2.7 and discrete Laplace distribution in section 2.2.8 can also be classified as generated to **preserve the maximum entropy** property of their continuous counterpart.

2.3 Methodology-III

This is a modification of the method-II (Barbiero 2010). Here the discrete analogue is derived to have a finite support.

Suppose X is a continuous random variable with pdf $f_X(x)$, $-\infty < x < \infty$. Y is the discrete analogue with the support consisting of k points to be derived from the range of X . Let $g = (1 - k)/2$, k odd positive integer and $y_i = g - 1 + i$, $i = 1, 2, \dots, k$.

For an example consider the case of discretizing X . Let $c_i = \Phi(y_i)$, $y_i = F_X^{-1}(c_i)$, where $\Phi(y_i)$ is the cdf of $N(0, 1)$ and $F_X()$ is the cdf of the X .

Then the pmf of Y with support $\{y_1, y_2, \dots, y_k\}$ is given by

$$P(Y = y_i) = f_X(y_i) / \sum_{i=1}^k f_X(y_i), \quad i = 1, 2, \dots, k$$

Barbiero (2010) gave examples of discrete gamma with 5 points and Weibull with 9 points support.

This method generates discrete analogue of continuous distribution with limited support like beta distribution. Here if X is symmetrical then Y retains expected value of X and pmf of Y retains the structure of the pdf of X .

For this method to be implemented the continuous cdf must be invertible, the support of the resulting discrete distribution may not be set of integers.

Barbiero (2010) *has applied this method to estimate the reliability of systems for which stress and strength are defined as complex functions, and whose reliability is not derivable through analytic techniques.*

2.3.1 Discrete power function distribution

The pdf of the continuous finite range power-function distribution having pdf

$$f(x) = px^{p-1}/\theta^p, \quad 0 < x < \theta, p > 0, \theta > 0$$

was introduced by Mukherjee and Islam (1983). Lai and Wang (1995) discretized the above distribution to derive a finite range discrete distribution with pmf

$$P(Y = k) = k^\alpha / c(N, \alpha), \quad k = 0, 1, \dots, N; \alpha > 0; \alpha \in \mathbf{R}$$

Where $c(n, j) = \sum_{k=0}^n k^\alpha = \frac{B_{j+1}(n+1) - B_{j+1}}{j+1}$, $B_m(x)$ is the m^{th} Bernoulli Polynomial defined

$$\text{as } B_m(x) = \sum_{k=0}^m B_k(x) x^{m-k}.$$

This distribution can model bathtub-shaped hazard rate as well as upside-down bathtub-shaped mean residual life. They studied various other reliability properties and applied this model to fit a mortality data.

2.4 Methodology-IV

Following Kemp's (2004) convention here we consider the definition of the discrete sf defined as $S_Y(k) = P(Y \geq k)$ and accordingly the cdf $F_Y(k) = P(Y \leq k)$ is related to the sf as $S_Y(k) = 1 - F_Y(k - 1)$.

If the underlying continuous random variable X has the survival function (sf) $S_X(x)$, then the random variable $Y = \lfloor X \rfloor$ = largest integer less or equal to X will have the pmf

$$\begin{aligned}
 P(Y = k) &= P(k \leq X < k + 1) \\
 &= F_X(k + 1) - F_X(k) = S_X(k) - S_X(k + 1), \quad k = 0, 1, 2, \dots
 \end{aligned}
 \tag{9}$$

[Since for continuous random variable X , $P(X = x) = 0$ and $F_X(k) = 1 - S_X(k)$]

The method can be viewed deriving a discrete concentration (Roy 2003) of the random variable X and also as a process of time discretization (Bracquemond and Gaudoin 2003) in the context of X representing life. It is possibly the easiest method of construction.

The resulting pmf will be in a compact form if the continuous sf is in compact form. This method preserves the sf that is $S_Y(k) = S_X(k)$.

One limitation of this technique is the concentration on the left limit of the equal intervals in which the support of the continuous random variable is partitioned.

Alternatively, by considering $Y = \lceil X \rceil$ smallest integer greater than or equal to X one can get a discrete version of X with following pmf that will preserve the cdf.

$$\begin{aligned}
 P(Y = k) &= P(X = k) = P(k - 1 < X \leq k) \\
 &= P(X \leq k) - P(X \leq k - 1) = F_X(k) - F_X(k - 1), \quad k = 0, 1, 2, \dots
 \end{aligned}$$

(see Lai 2012; Bracquemond and Gaudoin 2003).

It may be noted here that $\lceil X \rceil = \lfloor X \rfloor + 1$.

2.4.1 Discrete exponential distribution

If the underlying distribution is exponential with sf

$$S_X(x) = P(X \geq x) = \exp(-\theta x),$$

then the pmf of its discrete version is given by

$$P(Y = k) = \exp(-\theta k) - \exp(-\theta(k + 1)) = q^k - q^{k+1} = (1 - q)q^k, \quad k = 0, 1, 2, \dots$$

where $q = \exp(-\theta)$. This is the geometric distribution (Bracquemond and Gaudoin 2003).

2.4.2 Discrete Weibull distribution

Weibull distribution is widely accepted failure model but in practice, the failure data are often measured in discrete time such as cycles, blows, shocks, or revolutions. Discrete Weibull was proposed to find a discrete distribution corresponding to the Weibull.

If $X \sim$ Weibull distribution with pdf and sf

$$f_X(x) = \frac{\beta}{\lambda} \left(\frac{x}{\lambda}\right)^{\beta-1} \exp\left\{-\left(\frac{x}{\lambda}\right)^\beta\right\} \text{ and } S_X(x) = \exp\left\{-\left(\frac{x}{\lambda}\right)^\beta\right\}$$

Considering the sf of the Weibull in the Eq. (9), and substituting $q = \exp\left[-\left(\frac{1}{\lambda}\right)^\beta\right]$, Nakagawa and Osaki (1975) first proposed discrete Weibull distribution with pmf

$$P(Y = k) = q^{k^\beta} - q^{(k+1)^\beta}, \quad k = 0, 1, 2, \dots; \beta > 0; 0 < q < 1 \tag{10}$$

If X_1, X_2, \dots, X_n are i.i.d. discrete Weibull in Eq. (10) then $\min(X_1, X_2, \dots, X_n)$ is also a discrete Weibull. (see also Almalki (2014).

Khan *et al.* (1989) and Kulasekara (1994) considered estimation of this distribution. Englehardt and Li (2011) applied this distribution in modeling microbial counts. See also Bakouch *et al.* (2012) and Khorashadizadeh *et al.* (2012) for applications.

2.4.3 Discrete geometric Weibull distribution

Often we see systems possessing two phase life. First the stable phase having a constant failure rate until the change point time τ followed by next step which is the wear out phase with a larger increasing failure rate. Zacks (1984) considered the failure distribution in the wear out phase as Weibull to obtain the sf of the exponential Weibull distribution as

$$S_X(x) = \exp\left[-\lambda x - \{\lambda(x-\tau)^+\}^\beta\right], \text{ where } X^+ = \max(0, X).$$

The corresponding discrete version referred to as the discrete geometric Weibull was proposed by Bracquemond and Gaudoin (2003) with pmf

$$P(Y = k) = \exp\left[-\lambda(k-1) - \{\lambda(k-\tau-1)^+\}^\beta\right] - \exp\left[-\lambda k - \{\lambda(k-\tau)^+\}^\beta\right]$$

2.4.4 Discrete normal distribution

Roy (2003) considered discrete normal distribution with pmf

$$P(Y = k) = \Phi((k+1-\mu)/\sigma) - \Phi((k-\mu)/\sigma), \quad k = 0, \pm 1, \pm 2, \dots; \sigma > 0; -\infty < \mu < +\infty$$

where $\Phi(\cdot)$ is the cumulative distribution function (cdf) of standard normal distribution.

An application of the distributions for evaluating the reliability of complex systems was elaborated as an alternative to simulation methods Roy (2003).

2.4.5 Discrete Rayleigh distribution

If $X \sim$ Rayleigh distribution then its pdf and sf are respectively given by

$$f_X(x) = (x/\sigma^2) \exp[-x^2/2\sigma^2] \text{ and sf } S_X(x) = \exp[-x^2/2\sigma^2], \quad x > 0.$$

Discrete Rayleigh distribution (Roy 2004) has pmf

$$P(Y = k) = \theta^{k^2} - \theta^{(k+1)^2}, \quad k = 0, 1, 2, \dots; \quad 0 < \theta < 1.$$

This is a particular case of the discrete Weibull distribution of Nakagawa and Osaki (1975) stated in section 2.4.2.

Roy (2004) *applied this distribution in reliability modeling and in approximating probability integrals arising out of a reliability analysis in continuous setting.*

2.4.6 Discrete Maxwell distribution

If $X \sim$ Maxwell distribution then its pdf and sf are respectively given by

$$f_X(x) = \frac{4}{\sqrt{\pi}} \frac{1}{\theta^{3/2}} x^2 e^{-x^2/\theta} \text{ and sf } S_X(x) = 1 - \frac{\Gamma(3/2, x^2/\theta)}{\Gamma(3/2)}, \quad x > 0.$$

Krishna and Pundir (2007) studied discrete Maxwell distribution having pmf

$$P(Y = k) = \frac{4}{\sqrt{\pi}} \frac{1}{\theta^{3/2}} Q(k, 2, \theta), \quad k = 0, 1, 2, \dots; \quad \theta > 0$$

$$\text{where } Q(k, 2, \theta) = \int_k^{k+1} u^2 e^{-u^2/\theta} du.$$

2.4.7 Discrete extended exponential distribution (Telescopic)

If $X \sim$ extended exponential distribution then its pdf and sf are respectively given by

$$f_X(x) = \alpha g'_\theta(x) e^{-\alpha g_\theta(x)}, \text{ and sf } S_X(x) = e^{-\alpha g_\theta(x)}, \alpha, x > 0.$$

Where $g_\theta(x)$ is a strictly increasing function of x with $g_\theta(0) = 0$ and $g_\theta(x) \rightarrow \infty$ as $x \rightarrow \infty$ (Rezaei Roknabadi 2000, 2006).

Rezaei Roknabadi *et al.* (2009) obtained the pmf of their telescopic distribution by discretizing the extended exponential distribution as

$$P(Y = k) = q^{g_\theta(k)} - q^{g_\theta(k+1)}, k = 0, 1, 2, \dots, \text{ where } q = e^{-\alpha}, 0 < q < 1$$

Rezaei Roknabadi *et al.* (2009) have shown that this family of distribution belongs to IFR (increasing Failure Rate) class if any one of the following is true:

- i. $g_\theta^*(y) = g_\theta(y + 1) - g_\theta(y)$ is an increasing function of y .
- ii. For every sequence $\{q^{g_\theta(i+y)} - q^{g_\theta(y)}\}, i = 0, 1, 2, \dots$ is decreasing
- iii. For all $j_1, j_2, k_1, k_2 \in \{0, 1, \dots\}$ such that $j_1 < j_2$ and $k_1 < k_2$

$g_\theta(j_1 - k_1) - g_\theta(j_2 - k_2) \leq g_\theta(j_2 - k_1) - g_\theta(j_1 - k_2)$. That is satisfying the Polya sequence of order two for reliability function.

- iv. $\{g_\theta(y)\}, y = 0, 1, \dots$ is convex.

Further by taking $T_\theta(y) = \frac{1}{2} \{2g_\theta(y + 1) - g_\theta(y) - g_\theta(y + 2)\}$ it was proved by that the family is IFR (DFR) iff $T_\theta(y) > (<) 0$ and CFR iff $T_\theta(y) = 0$.

Following are some important distributions that belong to this family:

- i. Discrete exponential
- ii. Discrete Rayleigh
- iii. Discrete Weibull
- iv. Discrete Linear Exponential
- v. Discrete Gompertz

This class of distribution was reinvestigated under the name **discretized general class of continuous distribution** in the chapter IV of a Masters Thesis by Al-Masoud (2013).

They obtained the following distributions as particular cases:

- i. **Discrete Modified Weibull Extension** Distribution: By taking $g_\theta(x) = \exp(x/\theta)^\beta - 1$.
The pmf is of the form

$$P(Y = k) = q^{-1} \left[q^{\exp(k/\theta)^\beta} - q^{\exp((k+1)/\theta)^\beta} \right], k = 0, 1, 2, \dots$$

from which the discretized model of Chen (2000) is derived by putting $\theta = 1$.

This can be seen as a discretized version of the Modified Weibull Extension of Xie *et al.* (2002) having sf $S_X(x) = \exp[\lambda\alpha\{1 - \exp(x/\theta)^\beta\}], x > 0, \lambda > 0, \theta > 0, \beta > 0$ after appropriate re-parameterization.

ii. **Discrete Modified Weibull Type I Distribution:** By taking $g_{\theta}(x) = (\delta/\alpha)x + x^{\beta}$. The pmf is given by

$$P(Y = k) = q^{\alpha k + k^{\beta}} - q^{\alpha(k+1) + (k+1)^{\beta}}, \quad k = 0, 1, 2, \dots$$

This distribution is discretized version of the Modified Weibull Type I Distribution Sarhan and Zaindin (2009) having sf $S_X(x) = \exp[-\alpha x - \lambda x^{\beta}]$, $x > 0, \lambda > 0, \alpha, \beta > 0$ after appropriate re-parameterization. Al-Masoud (2013) derived and studied it in detail the discretized linear failure rate distribution as a special case by putting $\beta = 2$.

iii. **Discrete Modified Weibull Type II Distribution:** By taking $g_{\theta}(x) = e^{\alpha x} x^{\beta}$. The pmf is given by

$$P(Y = k) = q^{k^{\beta} e^{\alpha k}} - q^{(k+1)^{\beta} e^{\alpha(k+1)}}, \quad k = 0, 1, 2, \dots$$

This is a discretized version of the Modified Weibull Type II Distribution Lai *et al.* (2003) having sf $S_X(x) = \exp[-\lambda x^{\beta} e^{\alpha x}]$, $x > 0, \lambda > 0, \alpha, \beta > 0$ after appropriate re-parameterization. Reliability characteristics and parameter estimation of the above particular cases are also discussed in detail by Al-Masoud (2013).

The discrete modified Weibull distribution of Nooghabi *et al.* (2011) having pmf $P(Y = k) = q^{k^{\beta} c^k} - q^{(k+1)^{\beta} c^{k+1}}$, $k = 0, 1, \dots, 0 < q < 1, c \geq 0, \beta > 0$ is a particular case when $\alpha = 1$. The hazard rate function is increasing as well as bathtub shaped. (see also Almalki 2014)

iv. **Discrete Reduced Modified Weibull:** By taking $g_{\theta}(x) = \sqrt{x}(1 + bc^x)$. Almalki (2014) derived this distribution starting with continuous modified Weibull (Almalki 2014) having respective pdf and sf

$$f_X(x) = \frac{1}{2\sqrt{x}} \{ \alpha + \beta(1 + 2\lambda x) e^{\lambda x} \} e^{-\alpha\sqrt{x} - \beta\sqrt{x}e^{\lambda x}}, \quad x > 0, \alpha, \beta, \lambda > 0$$

and
$$S_X(x) = \frac{\exp[-\alpha\sqrt{x} - \beta\sqrt{x}e^{\lambda x}]}{q^{\sqrt{x}(1+bc^x)}}, \quad x > 0, \alpha, \beta, \lambda > 0$$

where $q = e^{-\alpha}, b = \beta/\alpha$ and $c = e^{\lambda}$ and $0 < q < 1, b > 0$ and $c \geq 1$. The corresponding pmf is given by

$$P(Y = k) = q^{\sqrt{k}(1+bc^k)} - q^{\sqrt{k+1}(1+bc^{k+1})}, \quad k = 0, 1, \dots \tag{11}$$

For $b = 0$ the distribution in Eq. (11) reduces to Discrete Weibull of Nakagawa and Osaki (1975) (see section 2.4.2 of this paper). Almalki (2014) applied this distribution to fit four data sets and compared the results with discrete Weibull, discrete additive Weibull and discrete modified Weibull distributions (see also Almalki and Nadarajah (2014).

2.4.8 Discrete Burr distribution

Krishna and Pundir (2009) studied discrete Burr distribution by considering $X \sim$ Burr distribution with pdf and sf

$$f_X(x) = \alpha\beta x^{\alpha-1} / (1 + x^{\alpha})^{\beta+1}, \quad x > 0, \alpha, \beta > 0 \text{ and } S_X(x) = (1 + x^{\alpha})^{-\beta} \text{ respectively.}$$

The pmf of their discrete Burr distribution is given by

$$P(Y = k) = \theta^{\log(1+k^\alpha)} - \theta^{\log(1+(k+1)^\alpha)}, \quad k = 0, 1, 2, \dots; \quad 0 < \theta < 1 \quad (12)$$

Where $\theta = e^{-\beta}$. See also Khorashadizadeh et al. (2012).

2.4.9 Discrete Pareto distribution

Krishna and Pundir (2009) derived the discrete Pareto distribution as a particular case of their discrete Burr distribution putting $\alpha = 1$ in the pmf in Eq. (12).

An application in reliability estimation in series system and a real data example on dentistry using this distribution is also discussed.

2.4.10 Discrete inverse Weibull distribution

If X follows Weibull, then the distribution X^{-1} is said to follow the inverse Weibull distribution. Jazi *et al.* (2010) proposed discrete inverse Weibull distribution by considering $X \sim$ Inverse Weibull distribution with sf $S_X(x) = 1 - \exp[-ax^{-\beta}]$. The pmf of inverse Weibull distribution is given by

$$P(Y = k) = \begin{cases} q, & k = 1 \\ q^{k^{-\beta}} - q^{(k-1)^{-\beta}}, & k = 2, 3, \dots; \beta > 0; 0 < q < 1 \end{cases}$$

Where $q = e^{-a}$. They studied its distributional and reliability properties and parameter estimation.

Application of this model in lifetimes of certain electronic devices was also considered by Jazi et al. (2010).

2.4.11 Discrete Inverse Rayleigh distribution

Inverse Rayleigh distribution is a particular case of inverse Weibull distribution when $\beta = 2$ with sf $S_X(x) = 1 - \exp[-a/x^2]$. Hussain and Ahmad (2014) proposed discrete inverse Rayleigh distribution with pmf

$$P(Y = k) = q^{1/(k+1)^2} - q^{1/k^2}, \quad 0 < q < 1, \quad k = 0, 1, 2, \dots \quad \text{where } \theta = e^{-a}.$$

Hussain and Ahmad (2014) *applied this distribution to model two real life count data.*

2.4.12 Discrete Lindley distribution

If $X \sim$ Lindley distribution then its pdf and sf are respectively given by

$$f_X(x) = \frac{\theta^2}{1 + \theta} (1 + x)e^{-x\theta} \quad \text{and} \quad S_X(x) = \frac{e^{-x\theta}(1 + \theta + \theta x)}{1 + \theta}, \quad x > 0.$$

Gómez-Déniz and Calderin-Ojeda (2011) proposed a discrete Lindley distribution having pmf

$$P(Y = k) = \frac{\lambda^k}{1 - \log \lambda} [\lambda \log \lambda + (1 - \lambda)(1 - \log \lambda^{k+1})], \quad 0 < \lambda < 1, \quad k = 0, 1, 2, \dots$$

Bakouch *et al.* (2012) again re-investigated this distribution and studied many additional properties of extensively.

This distribution was applied to model the collective risk model when both number of claims and size of a single claim are included in the model.

2.4.13 Discrete generalized exponential distribution

The generalized exponential distribution of Gupta and Kundu (1999) has pdf

$$f_X(x) = \alpha\lambda(1-e^{-\lambda x})^{\alpha-1}e^{-\lambda x} \text{ and sf } S_X(x) = 1-(1-e^{-\lambda x})^\alpha, x > 0, \alpha > 0, \lambda > 0$$

Nekoukhou *et al.* (2011) proposed a discrete analogue of this distribution with pmf given by

$$P(Y = k) = (1-p^{k+1})^\alpha - (1-p^k)^\alpha \tag{13}$$

They applied this distribution to model a discrete data se related to accidents of 647 women working on Shells for 5 weeks.

This distribution was first mentioned in Jiang (2010) and later independently derived as exponentiated-exponential-geometric distribution using T-X method in Alzaatreh *et al.* (2012), as an exponentiated geometric in Chakraborty and Gupta (2015).

2.4.14 Discrete gamma distribution

The Gamma distribution with parameters n and θ having pdf

$$f_X(x) = \frac{1}{\theta^n \Gamma(n)} x^{n-1} e^{-x/\theta} \text{ and sf } S_X(x) = \frac{1}{\theta^n \Gamma(n)} \int_x^\infty u^{n-1} e^{-u/\theta} du = \frac{1}{\Gamma(n)} \Gamma(n, x/\theta)$$

$$\text{Where } \Gamma(n, x/\theta) = \frac{1}{\theta^n} \int_x^\infty u^{n-1} e^{-u/\theta} du = \int_{x/\theta}^\infty u^{n-1} e^{-u} du$$

Chakraborty and Chakravarty (2012) defined a discrete gamma distribution with the pmf

$$P(Y = k) = (1/\Gamma(n))\Gamma(n, k/\theta, (k + 1)/\theta) ; k = 0, 1, \dots; n > 0, \theta > 0$$

Where $\Gamma(n, k/\theta, (k + 1)/\theta) = \Gamma(n, k/\theta) - \Gamma(n, (k + 1)/\theta)$.

The authors studied many properties including classification of failure rate and applied this distribution in empirical modelling of two discrete failure time data related to computer break down and time to death of leukemia patients.

2.4.15 Discrete Burr-III distribution

Al-Huniti and Al-Dayian (2012) discussed Discrete Burr III Distribution starting with the continuous one having the pdf and sf

$$f_X(x) = c k x^{-c-1} (1 + x^{-c})^{-d-1}, x > 0, c, d > 0 \text{ and}$$

$$S_X(x) = 1 - \exp[-d \log(1 + x^{-c})] \text{ respectively.}$$

The pmf of is given by

$$P(Y = k) = \theta^{\log\{1+(1+k)^{-c}\}} - \theta^{\log\{1+k^{-c}\}}, k = 0, 1, \dots, \text{ where } \theta = e^{-d}.$$

They have established the characterization property that distribution of the minimum order statistic from a sample of size n is Discrete Burr III distribution (c, θ^n) iff the sample is from Discrete Burr III distribution (c, θ) .

Para and Jan (2014) reinvestigated exactly the same distribution.

2.4.16 Discrete log-logistic distribution

It is a special case of discrete Burr distribution obtained by putting $\theta = e^{-1}$ in the pmf in Eq. (13). Khorashadizadeh et al. (2012).

2.4.17 Discrete generalized gamma distribution

The generalized gamma distribution with parameters $k, \theta,$ and c has pdf

$$f_X(x) = (c/(\theta^{cn}\Gamma k)) x^{cn-1} e^{-(x/\theta)^c}, t \geq 0; n, \theta, c > 0$$

and sf $S_X(x) = (1/\Gamma n)\Gamma_n((x/\theta)^c)$ respectively.

Where $\Gamma_n((t/\theta)^c) = \int_{(t/\theta)^c}^{\infty} v^{n-1} e^{-v} dv = (c/\theta^{cn}) \int_t^{\infty} u^{cn-1} e^{-(u/\theta)^c} du$

and $\Gamma_n(a) = \int_a^{\infty} v^{n-1} e^{-v} dv$ being the upper incomplete gamma function.

Starting with a statistical mechanical set up Chakraborty (2015a) defined a discrete generalized gamma distribution with the pmf

$$P(Y = k) = (1/\Gamma n)\Gamma_n((k/\theta)^c, ((k + 1)/\theta)^c); k = 0, 1, \dots; n > 0, \theta > 0, c > 0$$

Where $\Gamma_n((k/\theta)^c, ((k + 1)/\theta)^c) = (c/(\theta^{cn})) \int_k^{k+1} u^{cn-1} e^{-(u/\theta)^c} du.$

A number of existing and new distributions are seen as particular cases the discrete generalized gamma distribution $dy(n, \theta, c)$ for various values of the parameters n, θ and $c.$

For

- i. $c = 1,$ **discrete gamma** distribution $dy(n, \theta)$ (Chakraborty and Chakravarty 2012).
- ii. $n = 1,$ **discrete Weibull** distribution (Nakagawa and Osaki 1975).
- iii. $c = 1$ and $\theta = 1,$ One parameter **discrete gamma** distribution $dy(n)$ with pmf $P(Y = k) = (1/\Gamma n)\Gamma(n, k, (k + 1))$ (Chakraborty and Chakravarty 2012).
- iv. $c = 1$ and $n = 1,$ **geometric distribution** with pmf $P(Y = k) = q^k - q^{k+1} = (1 - q)q^k, k = 0, 1, 2, \dots,$ where $q = e^{-1/\theta}.$
- v. $c = 2,$ a **discrete hydrograph** distribution with pmf $P(Y = k) = 2/\theta^{2n}\Gamma k t^{cn-1} e^{-(t/\theta)^c}.$
- vi. $c = 2$ and $n \leftarrow n/2,$ **discrete generalized Rayleigh** distribution

$$P[Y = k] = (1/\Gamma(n/2))\Gamma_{n/2}((k/\theta)^2, (k + 1/\theta)^2); k = 0, 1, \dots; n > 0, \theta > 0$$

- vii. $c = 2, k = 1,$ **discrete Rayleigh** distribution (Roy 2004).
- viii. $c = 2, n = 3/2$ and $\theta \leftarrow \sqrt{\theta},$ **discrete Maxwell-Boltzmann** Krishna and Pundir (2007) distribution with pmf

$$P[Y = k] = (2/\sqrt{\pi})\Gamma_{3/2}(k^2/\theta, (k + 1)^2/\theta); k = 0, 1, \dots; \theta > 0.$$

ix. $c = 2$ and $n = 1/2$, **discrete half-Normal** distribution

$$P(Y = k) = \text{Erf} \left[\left(\frac{k}{\theta}, \frac{k+1}{\theta} \right) \right] = \text{Erf}[a, b] = 2 \left[\Phi(\sqrt{2}b) - \Phi(\sqrt{2}a) \right],$$

$b > a > 0, \theta > 0$, where $\Phi(\cdot)$ is the cdf of standard normal distribution.

x. Large $n, \mu = \log \theta + (1/c)\log n$ and $\sigma = 1/c\sqrt{n}$, **discrete lognormal** distribution with pmf

$$P(Y = k) = \Phi\{(\log(k+1)-\mu)/\sigma\} - \Phi\{(\log(k)-\mu)/\sigma\}; k = 0, 1, 2, \dots$$

Chakraborty (2015a) has shown that this distribution is IFR if $c > 1$, DFR if $k \leq 1, c < 1$ and CFR if $k = 1, c = 1$. *Application of the distribution in modelling two real life count data sets was also demonstrated by the author.*

2.4.18 Discrete Logistic distribution

The logistic distribution with parameters $\mu(-\infty < \mu < \infty)$ and $p (0 < p < 1)$ has pdf

$$f_X(x) = \frac{\exp\{-(x-\mu)/\beta\}}{\beta[1 + \exp\{-(x-\mu)/\beta\}]^2} \text{ and sf}$$

$$S_X(x; p, \mu) = p^{x-\mu} / (1 + p^{x-\mu}); x \in \mathbf{R}, 0 < p < 1.$$

A random variable Y is said to have a discrete logistic distribution Chakraborty and Chakravarty (2013) with parameter $p (0 < p < 1)$ and $-\infty < \mu < \infty$, if its pmf has the form

$$P(Y = k) = \frac{(1-p)p^{k-\mu}}{(1 + p^{k-\mu})(1 + p^{k-\mu+1})}, k \in \mathbf{Z}.$$

Chakraborty and Chakravarty (2013) *applied this distribution to model a real life count data in Z.*

Khorashadizadeh et al. (2012) considered the monotonic behavior of log odd ratio for standard discrete logistic distribution and **discrete truncated logistic** distribution and their relation with IFR class. They have also considered several other discrete lifetime distributions such as discrete Burr XII, Discrete log logistic (Krishna and Pundir 2009), Discrete Weibull (Nakagawa and Osaki 1975), discrete half normal Kemp et al. (2006). Discrete truncated logistic distribution was also considered in Bracquemond and Gaudoin (2003).

2.4.19 Another Discrete Skew Laplace distribution

Barbiero (2014) proposed an alternative discrete skew Laplace distribution by discretizing alternative parameterized skew Laplace distribution having respective pdf and sf

$$f_X(x) = \frac{\log p \log q}{\log(pq)} \begin{cases} p^x, & x \geq 0 \\ q^{-x}, & x < 0 \end{cases} \text{ and}$$

$$S_X(x) = \begin{cases} \frac{\log q}{\log(pq)} p^x, & x \geq 0 \\ 1 - \frac{\log q}{\log(pq)} q^{-x}, & x < 0, 0 < p < 1, 0 < q < 1. \end{cases}$$

The resulting pmf is given by

$$P(Y = k) = \frac{1}{\log(pq)} \begin{cases} \log p [q^{-(k+1)}(1-q)], & k = \dots, -2, -1 \\ \log q [p^k(1-p)], & k = 0, 1, 2, \dots \end{cases}$$

This distribution was applied to model two real life count data.

2.4.20 Discrete Gumbel distribution

The pdf and sf of the Gumbel (Type I) extreme value distribution is given by

$$f_X(x) = \sigma^{-1} e^{-(x-\mu)/\sigma} \exp[-e^{-(x-\mu)/\sigma}] \quad x \in \mathbf{R}, -\infty < \mu < \infty, \sigma > 0$$

and $S(x) = 1 - \exp[-e^{-(x-\mu)/\sigma}]$ respectively.

Chakraborty and Chakravarty (2014) proposed a discrete Gumbel distribution by discretizing the Gumbel distribution with pmf

$$P(Y = k) = e^{-\alpha p^{k+1}} - e^{-\alpha p^k}, k \in \mathbf{Z}, 0 < p < 1, \alpha > 0$$

After the re-parameterization $p = e^{-1/\sigma}$ and $\alpha = p^{-\mu}$.

They investigated the distributional, reliability and monotonic properties, different parameter estimation methods.

Chakraborty and Chakravarty (2014) applied this distribution to model three real life count data related to maximum flood discharges and annual maximum wind speeds from literature.

2.4.21 Discrete Additive Weibull distribution

If X_1 and X_2 are independent Weibull with sf $\exp[-\lambda_1 x_1^\theta]$ and $\exp[-\lambda_2 x_2^\gamma]$ respectively, then the distribution of $X = \min\{X_1, X_2\}$ is referred to as the additive Weibull distribution having sf

$$S_X(x) = \exp[-\lambda_1 x^\theta - \lambda_2 x^\gamma], \quad x > 0, \theta, \gamma, \lambda_1, \lambda_2 \in (0, \infty). \\ = q_1^{x^\theta} q_2^{x^\gamma}, \quad x = 0, 1, \dots \text{ where } q_1 = e^{-\lambda_1}, q_2 = e^{-\lambda_2}.$$

Bebbington *et al.* (2012) introduced the discrete additive Weibull distribution with four parameters. The sf and the pmf of this distribution are respectively given by

$$S_Y(k) = q_1^{k^\theta} q_2^{k^\gamma}, \quad x = 0, 1, \dots \text{ and}$$

$$P_Y(k) = q_1^{k^\theta} q_2^{k^\gamma} - q_1^{(k+1)^\theta} q_2^{(k+1)^\gamma}, \quad k = 0, 1, \dots, 0 < q_1, q_2 < 1, \theta, \gamma > 0$$

This distribution is IFR if $\theta \geq 1$ and $\gamma > 1$ ($\theta > 1$ and $\gamma \geq 1$), DFR if $\theta \leq 1$ and $\gamma < 1$ ($\theta < 1$ and $\gamma \leq 1$) and is bathtub shaped if $\theta < 1 < \gamma$ ($\gamma < 1 < \theta$) (see also Almalki 2014).

2.4.22 Discrete power distribution

Chakraborty and Chakravarty (2015) proposed a versatile new discrete distribution as a discrete analogue of the two sided power distribution of Van Drop and Kotz (2002a, b). The pmf of the discrete power distribution is given by

$$P(Y = k) = \begin{cases} \frac{(k-a+1)^n - (k-a)^n}{(b-a)(m-a)^{n-1}}, & k = a, a+1, \dots, m-1 \\ \frac{(b-k)^n - (b-k-1)^n}{(b-a)(b-m)^{n-1}}, & k = m, m+1, \dots, b-1 \end{cases}$$

Where a, b and $a \leq m \leq b$ are integers, and n is any positive real number. Some of its important distributional and reliability properties were investigated. Estimation methods of parameters were presented.

For more on general continuous triangular and two-sided power distributions see Zocchi and Kokonendji (2013) and for application of discrete triangular distribution in kernel estimation for discrete functions see Kokonendji and Zocchi (2010).

2.5 Methodology-V

If the underlying continuous random variable X has the cdf $F_X(x) = \Pr(X \leq x)$ then the pmf of the discrete analogue Y is given by

$$P(Y = k) = F_X(k + \delta) - F_X(k - [1 - \delta]), \quad 0 < \delta < 1 \tag{14}$$

Where the parameter $0 < \delta < 1$ is so chosen that the first two raw moments of X and Y remains close (Roy and Dasgupta 2001). Except for a shift in the location by δ the pmf in Eq. (14) preserves the form of the original cdf.

For example if X follows a normal and some other symmetrical unimodal distribution the optimal choice of δ is 0.5 so that the pmf in Eq. (14) reduces to

$$P(Y = k) = F_X(k + 0.5) - F_X(k - [1 - 0.5])$$

The choice of number of point of discretization is derived from a compromise between the accuracy and computational load of the results. Hence for reducing computational overload number of points should be small say 3 and for increasing accuracy the number of points should be large say 9.

Applied in approximating system reliability of complex systems under stress-strength model.

Note that for $\delta = 0$ and $\delta = 1$ the Eq. (14) reduces to the discrete analogues of X simple defined by $Y = \lceil X \rceil$ and $Y = \lceil X \rceil - 1$ with respective pmfs

$$P(Y = k) = F_X(k) - F_X(k-1) \quad \text{and} \quad P(Y = k) = F_X(k+1) - F_X(k).$$

2.5.1 Discrete Ade's distribution

Suppose that W has a gamma distribution with parameters n , and θ , has pdf $f_W(w) = (\theta^n / \Gamma(n)) w^{n-1} e^{-\theta w}$, $w \geq 0$; $n, \theta > 0$. Then

$$X = \begin{cases} 0, & \text{if } 0 \leq w \leq 1 \\ (\log w)^b, & \text{if } w \geq 1 \end{cases}$$

follows Ade's distribution with parameters n, θ, b .

The discrete Ade's distribution of Perry and Taylor (1985) is defined as

$$P(Y = k) = \begin{cases} P(0 \leq X < 0.5), & \text{if } k = 0 \\ P(i - 0.5 \leq X < i + 0.5), & \text{if } k = i, i = 1, 2, \dots \end{cases}$$

Perry and Taylor (1985) *fitted this distribution to 22 entomological data sets with encouraging results.*

2.6 Methodology-VI

This method preserves the hazard rate function. If the underlying continuous random variable X has the sf $S_X(x) = P(X \geq x)$ and hazard rate function $\lambda_X(x) = f_X(x)/S_X(x)$ then the sf of the discrete analogue Y is given by

$$P(Y \geq k) = (1 - \lambda_X(1))(1 - \lambda_X(2)) \cdots (1 - \lambda_X(k-1)), \quad k = 1, 2, \dots, m$$

The corresponding pmf is then given by

$$P(Y = k) = (1 - \lambda_X(1))(1 - \lambda_X(2)) \cdots (1 - \lambda_X(k-1)) [1 - (1 - \lambda_X(k))] \\ = (1 - \lambda_X(1))(1 - \lambda_X(2)) \cdots (1 - \lambda_X(k-1)) \lambda_X(k).$$

$$P(Y = k) = \begin{cases} \lambda_X(0), & k = 0 \\ (1 - \lambda_X(1))(1 - \lambda_X(2)) \cdots (1 - \lambda_X(k-1)) \lambda_X(k), & k = 1, 2, \dots, m \\ 0, & \text{else} \end{cases}$$

Note that here the range of Y that is value of m is determined so as to satisfy the condition that $0 \leq \lambda_X(x) < 1$ and multiply every $P(Y = k)$ by a positive normalizing constant to ensure the total probability equals to 1. Such a choice of is not going to affect the functional form of the failure rate. This approach though was highlighted by Roy and Ghosh (2009) was in fact used by Stein and Dattero way back in 1984 and preserves failure (hazard) rate function.

Bracquemond and Gaudoin (2003) though maintained that failure distribution with bounded support appears unrealistic from the point of view of applications since one cannot sure to ascertain that a system will necessarily fail in less than m counts.

2.6.1 Discrete Weibull

Hazard rate function of $X \sim$ Weibull distribution is given by

$$\lambda_X(x) = cx^{\beta-1}, \quad x > 0$$

Stein and Dattero (1984) presented a discretization of Weibull distribution with pmf

$$P(Y = k) = ck^{\beta-1} \prod_{j=1}^{k-1} (1 - cj^{\beta-1}), \quad k = 1, 2, \dots, m; \beta > 0; 0 < c \leq m \tag{15}$$

where the parameter m is determined in such a way that $0 \leq \lambda_X(x) < 1$.

$$m = \begin{cases} c^{-[1/(\beta-1)]}, & \text{if } \beta > 1 \\ +\infty, & \text{if } \beta \leq 1 \end{cases}$$

where $\lfloor X \rfloor$ = largest integer less or equal to X . For this distribution the hazard and sf rate function are respectively given by

$$\lambda_Y(k) = \begin{cases} ck^{\beta-1}, & k = 1, 2, \dots, m \\ 0, & k = 0 \text{ or } k > m \end{cases}$$

$$\text{and } S_Y(k) = \prod_{j=1}^{k-1} [1 - cj^{\beta-1}], \quad k = 1, 2, \dots, m.$$

Note that the distribution in Eq. (15) and the discrete Weibull defined in Eq. (10) coincides and reduces to geometric distribution when $c = 1 - q$ and $\beta = 1$. Khan *et al.* (1989) dealt with the estimation of the parameters of this distribution.

A connection is shown to the famous Birthday Problem and to the lifetime of a series system of components.

2.6.2 Discrete Rayleigh

The continuous Rayleigh distribution has

$$S_X(x) = \exp(-x^2/2\sigma^2) \text{ and } \lambda_X(x) = x/\sigma^2, \quad x > 0.$$

So the effective support of the discrete Rayleigh will have to be determined from the condition that $0 \leq \lambda_X(x) < 1$ which in this case implies $0 \leq x < \sigma^2$. Thus if we take $\sigma^2 = 2$, the range of X will be $0 \leq X < 2$.

2.6.3 Discrete Lomax

The continuous Lomax distribution has

$$S_X(x) = (1 + x/\beta)^{-\alpha} \text{ and } \lambda_X(x) = \alpha/(\beta + x), \quad x > 0.$$

So the effective support of the discrete Lomax (Roy and Ghosh 2009) will have to be determined from the condition that $0 \leq \lambda_X(y) < 1$ which in this case implies $y \geq \alpha - \beta$.

For details regarding above method of construction see Roy and Ghosh (2009) who have applied the above two distributions to approximate the reliability of complex systems approximating reliability under a stress strength model where exact determination of survival probability is analytically intractable.

2.6.4 Another Discrete Weibull

This method ensures that the alternative discrete hazard rate function of the discrete analogue is exactly same the hazard rate of the underlying continuous one. Alternative discrete hazard rate was defined by Roy and Gupta (1992) as $\lambda_Y^*(k) = \log [S_Y(k)/S_Y(k + 1)]$. This definition overcomes some of the problems classical definition of discrete hazard rate (see also Lai 2013). Consequently, the discrete alternative cumulative hazard rate defined as

$$H_Y^*(k) = \sum_{i=1}^k \lambda_i^*(k) \text{ obeys } H_Y^*(k) = -\log[1-F_Y(k)].$$

It can be easily checked that

$$\begin{aligned} \lambda_Y^*(k) &= \log[S_Y(k)/S_Y(k + 1)] = -\log[S_Y(k + 1)/S_Y(k)] \\ &= -\log[1-P(Y = k)/S_Y(k)] = -\log[1-\lambda_Y(k)] \end{aligned}$$

Hence $\lambda_Y(k) = 1 - \exp[-\lambda_Y^*(k)]$.

In this method of discretization if the underlying continuous random variable X has hazard rate function $\lambda_X(x)$, then the hazard rate function of the discrete analogue Y is given by $\lambda_Y(k) = 1 - \exp[-\lambda_X(k)]$ that is by taking $\lambda_Y^*(k) = \lambda_X(k)$. The pmf is the obtained by equation

$$P(Y = k) = (1-\lambda_Y(1))(1-\lambda_Y(2)) \cdots (1-\lambda_Y(k-1))\lambda_Y(k).$$

For example, if $X \sim$ Weibull distribution with hazard rate function $\lambda_X(x) = c x^{\beta-1}$, $x > 0$ and a discrete analogue is obtained by Padgett and Spurrier (1985) with pmf of the discrete Weibull is given by

$$\begin{aligned}
 P(Y = k) &= \left(1 - e^{-c k^{\beta-1}}\right) \prod_{j=1}^{k-1} \left(e^{-c j^{\beta-1}}\right) = \left(1 - e^{-c k^{\beta-1}}\right) e^{-c \sum_{j=1}^{k-1} j^{\beta-1}}, k = 1, 2, \dots; \beta \in \mathbb{R}; c \in \mathbb{R}^+ \\
 &= \left(1 - e^{-c(k+1)^{\beta-1}}\right) e^{-c \sum_{j=1}^k j^{\beta-1}}, k = 0, 1, 2, \dots; \beta \in \mathbb{R}; c \in \mathbb{R}^+
 \end{aligned}$$

For this distribution $\lambda_Y(k) = 1 - e^{-c k^{\beta-1}}$ and $\lambda_Y^*(k) = c k^{\beta-1}, k = 1, 2, \dots; \beta \in \mathbb{R}; c \in \mathbb{R}^+$. Lai (2013) also derived a discrete inverse Weibull using this method. See also Almalki (2014); Lai (2013) and Bracquemond and Gaudoin (2003).

Barbiero et al. (2013) discussed parameter estimation by different methods for this distribution in details with applications of real data fitting showing how the type III discrete Weibull distribution can fit real data.

2.7 Methodology-VII

This is a process proposed by Luceno (1999) of approximating a continuous random variable X having pdf $f(x), a \leq x \leq b$ by a discrete random variable Y taking values y_1, y_2, \dots, y_{MN} having pmf $P(Y = y_j) = p_j; j = 1, 2, \dots, MN$ such that both X and Y have same finite r^{th} moment for $r = 0, 1, \dots, 2N - 1$ and their cdf coincides at least at $M + 1$ points. Here the support of random variable Y i.e., $\{y_1, y_2, \dots, y_{MN}\}$ is roots of polynomial equation of N^{th} degree and not necessarily be the integers. As such derived distribution is not discrete in the sense of having integer support. So this is rather a way of approximating a pdf $f_X(x)$ by a pmf $\{p_j\}, j = 1, 2, \dots, MN$ which retain common moments and cdf value at the points of discretization. A list of approximation of some classical probability distribution proposed by Luceno (1999) is given in Table 1 below:

The gamma (t, α) distribution has mean t/α and variance t/α^2 ; the superscript Gauss-Hermite (GH), Gauss-Laguerre (GLa), Gauss-Jacobi (GJ) and Gauss-Legendre (GLe) refer to the polynomial names; the subscript j varies in $\{1, 2, \dots, N\}$.

This method may require solution of system of non-linear equations in addition to the requirement of the existence of moments of the continuous distribution.

2.8 Methodology-VIII

Hagmark (2008) presented a method for constructing nonnegative integer-valued random variables with any interval domain, any theoretically possible mean–variance pair, and different shapes using basic tool of a mean preserving discretization method in which the discretization of a nonnegative initial random variable X with cdf $F_X(x)$ is defined as the count variable Y with cdf

Table 1 Approximation of some classical probability distribution

Distribution	Abscissa	Probability	Weight	Polynomial
$N(\mu, \sigma^2)$	$\mu + \sigma X_j^{(GH)}$	$\frac{1}{\sqrt{\pi}} W_j^{(GH)}$	$\exp(-x^2), -\infty < x < \infty$	GH
gamma (t, a)	$\frac{1}{a} X_j^{(GLa)}$	$\frac{1}{\Gamma t} W_j^{(GLa)}$	$x^{t-1} \exp(-x), 0 < x < \infty$	GLa
beta (a, β)	$\frac{(1+x_j^{(GJ)})}{2}$	$\frac{t^{1-a-\beta} W_j^{(GJ)}}{\text{beta}(a, \beta)}$	$\frac{(1-x)^{a-1}}{(1+x)^{1-\beta}}, -1 < x < 1$	GJ
uniform (a, b)	$a + \frac{b-a}{2(1+x_j^{(GLE)})}$	$\frac{1}{2} W_j^{(GLE)}$	$1, -1 < x < 1$	GLE

$$F_Y(y) = P(Y \leq y) = \int_y^{y+1} F_X(x) dx$$

He has shown that under this construction

- (i) $E(X) = E(Y)$ and (ii) $Var(X) \leq Var(Y) \leq Var(X) + \min\{E(X), 1/4\}$.

Note that $F_Y(y)$ is actually the average of $F_X(\cdot)$ in the interval $(n, n + 1)$ under assumption of uniform distribution in that interval. Hagmark (2008) asserted that every count variable is a discretization of an initial continuous random variable which is seldom unique. He gave example initial continuous distribution of which Poisson distribution is a discretized version and algorithms to generate discrete distributions using this method.

2.9 Two stage composite methods

2.9.1 Discretized Exponentiated models

In this method the discrete analogue of the continuous random variable X having cdf $F_X(x)$ and sf $S_X(x)$ is derived as a discrete random variable Y having pmf

$$P(Y = k) = S_X(k)^\alpha - S_X(k + 1)^\alpha, y = 0, 1, 2, \dots; \alpha > 0$$

Thus basically first the continuous distribution function is exponentiated and the resulting exponentiated continuous distribution is then discretized by using the methodology-IV.

For example, by exponentiating the cdf of the continuous exponential distribution Gupta and Kundu (1999) derived generalized exponential distribution having pdf, cdf and sf

$$f_x(x) = \alpha \lambda (1 - e^{-\lambda x})^{\alpha-1} e^{-\lambda x}, x > 0, \alpha > 0, \lambda > 0.$$

$$F_X(x) = (1 - e^{-\lambda x})^\alpha \text{ and } S_X(x) = 1 - (1 - e^{-\lambda x})^\alpha \text{ respectively.}$$

Writing $q = e^{-\lambda}$, a discrete analogue of this distribution can be obtained with pmf

$$P(Y = k) = (1 - q^{k+1})^\alpha - (1 - q^k)^\alpha, k = 0, 1, 2, \dots \tag{16}$$

Which is the distribution mentioned in Eq. (13) and again later in Eq. (20). (see Mudholkar *et al.* (1995) for exponentiated Weibull).

Remark 1. One can use the exponentiation of sf and then discretize to get different analogues. Also one can use other methodologies instead of method III to generate different discrete analogues of the exponentiated continuous distributions.

2.9.2 Two-fold competing risk models

In this method (Jiang 2010) first two continuous random variables X_1 and X_2 having sfs $S_{X_1}(x)$ and $S_{X_2}(x)$ are combined to produce a new random variable X having sf $S_X(x) = S_{X_1}(x)S_{X_2}(x)$.

Then a discrete analogue Y of X is derived from $S_X(x)$ by using methodology-IV. The resulting pmf is

$$P(Y = k) = S_X(k) - S_X(k + 1) = S_{X_1}(k)S_{X_2}(k) - S_{X_1}(k + 1)S_{X_2}(k + 1) \\ = S_{X_1}(k)P_{X_2}(Y = k) + S_{X_2}(k)P_{X_1}(Y = k) - P_{X_1}(Y = k)P_{X_2}(Y = k), k = 0, 1, 2, \dots; \alpha > 0$$

Where $P_{X_i}(Y = k) = S_{X_i}(k) - S_{X_i}(k + 1)$ is the discrete analogue of the continuous random variable X_1 . Clearly, the random variable X is equal to minimum $\{X_1, X_2\}$.

Discrete additive Weibull distribution discussed in the section 2.4.21 can be seen as an example of this construction.

Remark 2.

- i. Obviously, one can generalize this to more than two i.e. manifold competing risk models model.
- ii. Discretized exponentiated method can be seen as a particular case of this method when the X 's are identical.

2.9.3 Marshall and Olkin followed by method-III

In this method first the sf $S_X(x)$ of a continuous random variable X is generalized by adding an extra parameter α using Marshall and Olkin (1997) scheme then discretize by using the methodology-IV. The generalized sf is then

$$S_X(x, \alpha) = \frac{\alpha S_X(x)}{1 - (1 - \alpha) S_X(x)} \tag{17}$$

and the corresponding pmf of the discrete analogue by method-IV is

$$P(Y = k) = S_X(k, \alpha) - S_X(k + 1, \alpha) = \frac{\alpha \{S_X(k) - S_X(k + 1)\}}{\{1 - (1 - \alpha) S_X(k)\} \{1 - (1 - \alpha) S_X(k + 1)\}} \tag{18}$$

2.9.3.1 Generalization of the geometric distribution Gómez-Déniz (2010) proposed and studied a new generalization of the geometric distribution by using this scheme of discretization. They started with X following exponential distribution with sf $S_X(x) = \exp(-\theta x) = q^x$, where $q = e^{-\lambda}$ and used the construction in Eq. (18) generalize the geometric distribution with pmf

$$P(Y = k) = \frac{\alpha q^k (1 - q)}{\{1 - (1 - \alpha) q^{k+1}\} \{1 - (1 - \alpha) q^k\}}, \quad k = 0, 1, 2, \dots$$

2.9.3.2 Discrete half normal Gómez-Déniz *et al.* (2014) proposed a discrete version of the half-normal distribution by using this scheme of discretization and investigated its generalization with applications.

First taking $S_X(x) = \Phi_X(x)$ where $\Phi_X(x)$ is the cdf of $N(0, \sigma)$ in Eq. (17) a generalization of the normal distribution is obtained with sf

$$S_X(x) = \frac{\alpha \{1 - \Phi_X(x)\}}{1 - (1 - \alpha) \{1 - \Phi_X(x)\}}, \quad -\infty < x < \infty$$

Then the sf for the corresponding distribution in \mathbf{R}^+ which can be considered as a generalization of the **half-normal** distribution is given by

$$\frac{S_X(x)}{S_X(0)} = \frac{(1 + \alpha) \{1 - S_X(x)\}}{1 - (1 - \alpha) \{1 - S_X(x)\}}, \quad x > 0$$

Now employing the methodology-IV the pmf of discrete generalized half normal distribution is obtained as

$$P(Y = k) = \frac{(1 + \alpha)\{\Phi_X(k + 1) - \Phi_X(k)\}}{\{1 - (1 - \alpha)(1 - \Phi_X(k))\}\{1 - (1 - \alpha)(1 - \Phi_X(k + 1))\}}$$

In particular for $\alpha = 1$, we get the discrete half normal distribution Chakraborty (2015a) with pmf

$$P(Y = k) = 2[\Phi_X(k + 1) - \Phi_X(k)]$$

2.9.4 T-X method

Suppose $F_X(x)$, $h_X(x)$ and $H_X(x) = -\log(1 - F_X(x))$ be respectively the cdf, the hazard rate function and cumulative hazard rate function of any random variable X . $f_T(t)$ and $F_T(t)$ be the pdf and cdf of another continuous random variable T defined on $(0, \infty)$. The cdf of the random variable Y having T - X family of distributions defined by Alzaatreh *et al.* (2012) is then given by

$$F_Y(y) = \int_0^{-\log(1-F_X(y))} f_T(t) dt = F_T\{-\log(1-F_X(y))\},$$

when X is a continuous random variable the corresponding pdf of the T - X family can be obtained as

$$f_Y(y) = \frac{f_X(y)}{1-F_X(y)} f_T(-\log(1-F_X(y))) = h_X(y) f_T(H_X(y)).$$

If X is a discrete random variable, the T - X family is a family of discrete distribution transformed from the **non-negative** continuous random variable T . The pmf of the T - X family of discrete distribution can be found as

$$P(Y = k) = F_Y(k) - F_Y(k-1) = F_T\{-\log(1-F_X(k))\} - F_T\{-\log(1-F_X(k-1))\}; k = 0, 1, 2, \dots \tag{19}$$

As such we can see that this method is essentially employing discretization method on the T - X pdf to generate new discrete distribution.

If X is a geometric random variable with cdf $F_X(x) = 1 - p^{x+1}$, $x = 0, 1, 2, \dots$, then T - X family in Eq. (19) is referred to as the **T-geometric** family with pmf

$$\begin{aligned} &= F_T\{-\log(p^{k+1})\} - F_T\{-\log(p^k)\} \\ &= F_T\{-(k+1)\log p\} - F_T\{-k\log p\}; k = 0, 1, 2, \dots \end{aligned} \tag{20}$$

In particular if X is a geometric random variable with parameter $p = e^{-1} = 0.3679$, then pmf of the **T-geometric** family reduces to $P(Y = k) = F_T(k + 1) - F_T(k)$, $k = 0, 1, 2, \dots$ (see section 2.5).

Alzaatreh *et al.* (2012) proved many properties of this family including the unimodality of the **T-geometric** family given that the non-negative continuous random variable T is unimodal with a unique mode.

For example if the random variable T follows the exponentiated-exponential distribution (Gupta and Kundu 1999) with cdf

$$F_T(t) = (1 - e^{-\lambda t})^\alpha, t > 0, \alpha > 0, \lambda > 0$$

then *T-geometric* family in Eq. (20) leads to the **exponentiated exponential–geometric** distribution (EEGD) with pmf

$$P(Y = k) = (1 - p^{\lambda(k+1)})^\alpha - (1 - p^{\lambda k})^\alpha, k = 0, 1, 2, \dots \quad (21)$$

On replacing p^λ by θ , we will have

$$P(Y = k) = (1 - \theta^{k+1})^\alpha - (1 - \theta^k)^\alpha, 0 < \theta < 1, k = 0, 1, 2, \dots$$

Note that, if $\alpha = 1$, i.e. the random variable T has exponential distribution, and then the EEGD reduces to the geometric distribution. Also observe the similarity of Eq. (21) with Eqs. in (13) and (16).

2.9.5 Generalization of the *T-X* method

Let $f_T(t)$ be the pdf of a continuous random variable T defined on $[a, b]$ and $W(\cdot)$ be a absolutely continuous and monotonically non-decreasing function with $W(0) \rightarrow a$ and $W(1) \rightarrow b$. Then the cdf of the **generalized *T-X* family** of distributions defined by Alzaatreh *et al.* (2013) is given by

$$f_Y(y) = \int_0^{W(F(y))} f_T(t) dt = F_T[W(F_X(y))],$$

If X is a non-negative discrete random variable, then the pmf of this generalized *T-X* family of discrete distribution can be found as

$$P(Y = k) = F_T\{W(F_X(k))\} - F_T\{W(F_X(k-1))\}; k = 0, 1, 2, \dots \quad (22)$$

Obviously Eq. (22) reduces to Eq. (19) when $W(x) = -\log(1 - x)$.

Akinsete *et al.* (2014) considered T as the Kumaraswamy (1980) distribution with cdf $F_T(t) = 1 - (1 - t^\alpha)^\beta, 0 < t < 1, \alpha > 0, \beta > 0$, X as the geometric random variable with cdf $F_X(x) = 1 - p^{x+1}, x = 0, 1, 2, \dots$ and $W(x) = x$ in Eq. (21) to propose the *Kumaraswamy-geometric* distribution with pmf

$$P(Y = k) = [1 - (1 - p^k)^\alpha]^\beta - [1 - (1 - p^{k+1})^\alpha]^\beta, k = 0, 1, 2, \dots \quad (23)$$

Note that for $\beta = 1$, Eq. (23) reduces to Eq. (21). Akinsete *et al.* (2014) also proved that this distribution can also be derived by considering log-Kumaraswamy distribution instead of Kumaraswamy distribution and taking $W(x) = -\log(1 - x)$.

2.9.6 Method of discretization after transmutation

Chakraborty (2015b) recently introduced the idea of discretization of transmuted continuous distributions. A random variable Z is said to be constructed by the quadratic rank transmutation map method of Shaw and Buckley (2007) by transmuting another random variable X with cdf $F_X(\cdot)$ if the cumulative distribution function (cdf) of Z is given by

$$F_Z(z) = (1 + \alpha)F_X(z) - \alpha(F_X(z))^2, -1 < \alpha < 1$$

So given a cdf $F_X(x)$ of a continuous random variable X , it is first transmuted to $F_Z(z)$ by adding an extra parameter α using Shaw and Buckley (2007) scheme then discretized by using the methodology-IV. The corresponding pmf of the new discrete distribution is then given by

$$P(Y = k) = (1 + \alpha)[F_X(k + 1) - F_X(k)] - \alpha[(F_X(k + 1))^2 - (F_X(k))^2]$$

For example considering $F_X(x) = 1 - \beta e^{-\beta x}$, the pdf and cdf of the transmuted exponential distribution derived using the quadratic rank transmutation by Shaw and Buckley (2007) are respectively given by

$$f_Z(z) = \beta e^{-\beta z}(1 - \alpha) + 2\alpha\beta e^{-2\beta z}, z > 0, \beta > 0, -1 < \alpha < 1$$

$$F_Z(z) = (1 + \alpha)(1 - e^{-\beta z}) - \alpha(1 - e^{-\beta z})^2, z > 0, \beta > 0, -1 < \alpha < 1 \text{ (Shaw and Buckley, 2007).}$$

Now using the methodology-IV the pmf of the discrete analogue Y of transmuted exponential, is obtained as

$$\begin{aligned} P(Y = k) &= F_Z(k + 1) - F_Z(k) \\ &= (1 - \alpha)q^k(1 - q) + \alpha q^{2k}(1 - q^2), k = 0, 1, 2, \dots; 0 < q < 1, -1 < \alpha < 1 \end{aligned}$$

with $e^{-\beta} = q$. This is the transmuted geometric distribution proposed recently by Chakraborty (2015b) and studied in detail by Chakraborty and Bhati (2015).

3. Discussion and conclusions

3.1. Benefits of discretization of continuous probability distribution

When only an approximating discrete random variable is observable, estimation procedures employing the hypothetical continuous random variables are sometime biased and hence a discrete distribution is more appropriate for an observed data (Holland 1975).

Discretization of continuous distribution may be looked upon as a filtering process which may help in reducing of noise present in the data. Especially data sets having a high amount of background noise can gain from this process.

Discretization may bring in computational easiness.

3.2 Limitation of discretization of continuous probability distribution

When a continuous probability density function is discretized to a probability mass function there will always be some loss of information. As such one should try to strike a balance between the need for discretization and resulting loss of information or accuracy. Also attention should be paid to select the best one among the available techniques of discretization. Some of the criteria for selection mentioned in Bracquemond and Gaudoin (2003) are simple and flexible expressions, physical basis for the distribution, interpretation of model parameter, and efficiency of estimators.

3.3 Concluding remarks

The discretization of a continuous distribution using different methods has attracted renewed attention of researchers in last few years. Though a large number of such distributions are now available in the literature, still new discrete analogues are being added to the existing collection. There is still enough scope to contribute new discrete versions using different methods since not all methods received same attention of the researchers. This article is aimed at providing up to date information on this

vibrant research topic. Future research in this area may be to search for different constructions that might ensure preservation of multiple characteristics of the continuous distribution in its discretized version, developing inferential procedures for these discrete analogues etc. among others. We have not discussed detail properties of the methods and discretized distributions presented in this survey as those can be accessed from the respective original references.

Competing interests

The author declares that he has no competing interests.

Authors' contributions

The author contribution is sole.

Acknowledgement

Author would like to acknowledge the remarks and suggestions made by various research scholars and scientists in the *International Conference on Statistical Data Mining for Bioinformatics, Health, Agriculture and Environment, at Rajshahi University, Bangladesh* during December 22 – 24, 2012 where an earlier version of this paper was first presented as an Invited talk.

Author would also like to acknowledge the anonymous referee and editor for their comments and suggestions on the first draft of this paper which lead to substantial improvements in presentation.

Received: 4 February 2015 Accepted: 1 July 2015

Published online: 05 August 2015

References

- Akinsete, A, Famoye, F, Lee, C: The Kumaraswamy-Geometric distribution. *J. Stat. Distributions. Appl.* **1**, 17 (2014)
- Alhazzani, NS: Modeling discrete life data in reliability and its applications. Master thesis. King Saud University, Riyadh (2012)
- Al-Huniti, AA, AL-Dayian, GR: Discrete Burr type III distribution. *Am. J. Math. Stat.* **2**(5), 145–152 (2012). doi:10.5923/jajms.20120205.07
- Almalki, SJ: Statistical analysis of lifetime data using new modified Weibull distributions, a thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Engineering and Physical Sciences. (2014)
- Almalki, SJ, Nadarajah, S: A new discrete modified Weibull distribution. *IEEE. Trans. Reliability.* **63**(1), 66–180 (2014). doi:10.1109/TR.2014.2299691
- Al-Masoud, TA: A discrete general class of continuous distributions. Master of Science (Statistics) thesis. Faculty of Science, King Abdul Aziz University, Jeddah-Saudi Arabia (2013)
- Alzaatreh, A, Lee, C, Famoye, F: On the discrete analogue of continuous distributions. *Stat. Methodol.* **9**, 589–603 (2012)
- Alzaatreh, A, Lee, C, Famoye, F: A new method for generating families of continuous distributions. *Metron* **71**, 63–79 (2013)
- Bakouch, HS, Jazi, MA, Nadarajah, S: A new discrete distribution. *Statistics, iFirst*, 1–41 (2012). doi/10.1080/02331888.2012.716677
- Barbiero, A: A discretizing method for reliability computation in complex stress-strength models. *World. Acad. Sci. Eng. Technol.* **47**, 75–81 (2010)
- Barbiero, A: An Alternative discrete skew Laplace distribution. *Stat. Methodol.* **16**, 47–67 (2014)
- Barbiero, A: Parameter estimation for type III discrete Weibull distribution: a comparative study. *J. Probab. Stat.*, Volume 2013, Article ID 946562, <http://dx.doi.org/10.1155/2013/946562>
- Bebbington, M, Lai, CD, Wellington, M, Zitikis, R: The discrete additive Weibull distribution: a bathtub shaped hazard for discontinuous failure data. *Reliability Engineering and System Safety* (2012)
- Benkherouf, L, Bather, JA: Oil exploration: sequential decisions in the face of uncertainty. *J. Appl. Probability.* **25**, 529–543 (1988)
- Bi, Z, Faloutsos, C, Korn, F: The “DGX” distribution for mining massive skewed data. 7th Conf. on Knowledge discovery and data Mining, San Francisco (2001)
- Bracquemond, C, Gaudoin, O: A survey on discrete life time distributions. *Int. J. Reliabil. Qual. Saf. Eng.* **10**, 69–98 (2003)
- Byers, Jr, RH, Shenton, LR: Surprising approximations to the half-normal. *Journal of Statistical Computation and Simulation* **49**(3), 215–216 (1994)
- Carver, HC: On the graduation of frequency distributions. *Proc. Casual. Actuarial. Soc. Am.* **6**, 52–72 (1919)
- Carver, HC: Frequency curves. *Handbook of Mathematical Statistics*. Rietz, HL (editor), Cambridge MA: Riverside, 92–119 (1923)
- Chakraborty, S: A new discrete distribution related to generalized gamma distribution and its properties. *Commun. Stat. Theory. Methods.* **44**(8), 1691–1705 (2015). doi:10.1080/03610926.2013.781635
- Chakraborty, S, Chakravarty, D: Discrete gamma distributions: properties and parameter estimation. *Commun. Stat. Theory. Methods.* **41**(18), 3301–3324 (2012)
- Chakraborty, S, Chakravarty, D: A discrete Gumbel distribution. arXiv:1410.7568 [math.ST], 28 Oct 2014
- Chakraborty, S, Chakravarty, D: A discrete two sided power distribution. arXiv:1501.06299[math.ST], 26 Jan 2015
- Chakraborty, S, Chakravarty, D: A new symmetric discrete probability distribution with integer support on $(-\infty, \infty)$. Published on line 12th August 2013, *Communications in Statistics-Theory and Methods* (2013)
- Chakraborty, S, Gupta, RD: Exponentiated geometric distribution: another generalization of geometric distribution. *Commun. Stat. Theory. Methods.* **44**(6), 1143–1157 (2015). doi:10.1080/03610926.2012.763090
- Chakraborty, S: Transmuted geometric distribution and its properties. arXiv:1502.04203 [math.ST], (2015b)

- Chakraborty, S, Bhati, D: Transmuted geometric distribution with applications in modeling and regression analysis of count data. Under review (2015)
- Chen, Z: A new two - parameter lifetime distribution with bathtub shape or increasing failure rate function. *Stat. Probabil. Lett.* **49**(2), 155–161 (2000)
- Das Gupta, R: Cauchy equation on discrete domain and some characterizations. *Theor. Probability. Appl.* **38**, 318–328 (1993)
- Doray, LG, Luong, A: Efficient estimators for the Good family. *Comm. Statist. Simul. Comput.* **26**, 1075–1088 (1997)
- Englehardt, JD, Li, R: The discrete Weibull distribution: an alternative for correlated counts with confirmation for microbial counts in water. *Risk Anal.* **31**(3), 370–381 (2011)
- Gómez-Déniz, E: Another generalization of the geometric distribution. *Test* **19**, 399–415 (2010)
- Gómez-Déniz, E, Calderin-Ojeda, E: The discrete Lindley distribution: properties and applications. *J. Stat. Comput. Simul.* **81**(11), 1405–1416 (2011)
- Gómez-Déniz, E, Vazquez-Polo, Garcia-Garcia, V: A discrete version of the half-normal distribution and its generalization with applications. *Stat Papers* **55**(2), 497–511 (2014)
- Good, IJ: The population frequencies of species and the estimation of population parameters. *Biometrika* **40**, 237–264 (1953)
- Gradshteyn, IS, Ryzhik, IM: *Tables of Integrals, Series, and Products*. Sixth Edition, Academic Press, London (2000)
- Gupta, RD, Kundu, D: Generalized exponential distributions. *Aust. N. Z. J. Stat.* **41**(2), 173–188 (1999)
- Gupta, P, Gupta, RC, Ong, SH, Srivastava, HM: A class of Hurwitz-Lerch Zeta distributions and their applications in reliability. *Appl. Math. Comput.* **196**, 521–531 (2008)
- Gurland, J, Tripathi, RC: Estimation of parameters on some extensions of the Katz family of discrete distributions involving hypergeometric functions. *Statistical Distributions in Scientific Work, Vol. 1: Models and Structures*, Patil, GP, Kotz, S, Ord, JK (editors), 59–82 (1975)
- Hagmark, PE: On construction and simulation of count data models. *Math. Comput. Simul.* **77**, 72–80 (2008)
- Haight, FA: Queueing with balking. *Biometrika* **44**, 360–369 (1957)
- Harris, TR, Shonkwiler, JS, Lin, Y: Application of discrete normal distribution for dynamic rural retail sector analysis: Selected paper at the annual AAEA meeting Chicago, Illinois, August 5–8. (2001)
- Holland, BS: Some Results on the discretization of continuous probability distributions. *Technometrics* **17**(3), 333–339 (1975)
- Hussain, T, Ahmad, M: Discrete inverse Rayleigh distribution. *Pak. J. Statist.* **30**(2), 203–222 (2014)
- Inusah, S, Kozubowski, TJ: A discrete analogue of the Laplace distribution. *J. Stat. Planning. Inference.* **136**, 1090–1102 (2006)
- Jamjoom, AA: Order statistics from discrete gamma distribution. *J. Am. Sci.* **9**(7), 487–498 (2013)
- Jazi, MA, Lai, CD, Alamatsaz, MH: A discrete inverse Weibull distribution and estimation of its parameters. *Stat. Methodol.* **7**(2), 121–132 (2010). doi:10.1016/j.stamet.11.001
- Jiang, R: Discrete competing risk model with application to modeling bus-motor failure data. *Reliability. Eng. Syst. Saf.* **95**, 981–988 (2010)
- Johnson, NL, Kemp, AW, Kotz, S: *Univariate discrete distributions*. Wiley, New York (2005)
- Katz, L: Characteristics of frequency functions defined by first order difference equations. PhD Thesis. University of Michigan, Ann Arbor, MI (1945)
- Katz, L: On the class of functions defined by the difference equation $(x + 1) f(x + 1) = (a + bx) f(x)$. *Ann. Math. Stat.* **17**, 501 (1946) (abstract)
- Katz, L: Frequency functions defined by the Pearson difference equation. *Ann. Math. Stat.* **19**, 120 (1948) (abstract)
- Katz, L: Unified treatment of a broad class of discrete probability distributions. *Classical and contagious discrete distributions*. Patil, G.P., (editor), Calcutta: Statistical Publishing Society; Oxford: Pergamon, 175–182 (1965)
- Kemp, AW: Characterization of a discrete normal distribution. *J. Stat. Planning. Inference.* **63**, 223–229 (1997)
- Kemp, CD: q -Analogues of the hyper-Poisson distribution. *J. Stat. Planning. Inference.* **101**, 179–183 (2002)
- Kemp, AW: Classes of discrete lifetime distributions. *Commun. Stat. Theor. Methods.* **33**(12), 3069–3093 (2004)
- Kemp, AW: The discrete half-normal distribution. *International Conference on Mathematical and Statistical Modeling in the Honor of Enrique Castillo*, June 28–30 (2006)
- Khan, MS, Khalique, A, Aboummoh, AM: On estimating parameters in a discrete Weibull distribution. *IEEE. Trans. Reliability.* **38**(3), 348–350 (1989)
- Khorashadizadeh, M, Rezaei Roknabadi, AH, Mohtashami Borzadaran, GR: Characterisation of life distributions using log-odds rate in discrete ageing. *Commun. Stat. Theor. Methods.* **42**(1), 76–87 (2012)
- Kokonendji, CC, Zocchi, SS: Extensions of discrete triangular distributions and boundary bias in kernel estimation for discrete functions. *Stat. Probabil. Lett.* **80**(21–22), 1655–1662 (2010)
- Kozubowski, TJ, Inusah, S: A skew Laplace distribution on Integers. *Ann. Inst. Stat. Math.* **58**, 555–571 (2006)
- Krishna, H, Pundir, PS: Discrete Burr and discrete Pareto distributions. *Stat. Methodol.* **6**, 177–188 (2009)
- Krishna, H, Pundir, PS: Discrete Maxwell distribution. *Interstat*, <http://interstat.statjournals.net/YEAR/2007/articles/0711003.pdf> (2007)
- Kulasekara, KB: Approximate mle of the parameters of a discrete Weibull distribution with type-I censored data. *Microelectron. Reliab.* **34**, 1185–1188 (1994)
- Kulasekara, KB, Tonkyn, DW: A new discrete distribution with application to survival, dispersal and dispersion. *Commun. Stat. Simul. Comput.* **21**, 499–518 (1992)
- Kumaraswamy, P: A generalized probability density function for double-bounded random processes. *Hydrology* **46**, 79–88 (1980)
- Lai, CD: Constructions and applications of lifetime distributions. *Appl. Stochastic. Models. Bus. Ind.* **29**, 127–140 (2012)
- Lai, CD: Issues concerning constructions of discrete lifetime models. *Qual. Technol. Quant. Manag.* **10**(2), 251–262 (2013)
- Lai, CD, Wang, DQ: A finite range discrete life distribution. *Int. J. Reliability. Qual. Saf. Eng.* **2**(2), 147–160 (1995)
- Lai, CD, Xie, M, Murthy, DNP: A modified Weibull distribution. *Reliability. IEEE. Trans.* **52**(1), 33–37 (2003)
- Lekshmi, S, Sebastian, S: A skewed generalized discrete Laplace distribution. *Int. J. Math. Stat. Invent.* **2**(3), 95–102 (2014)
- Liang, TC: Monotone empirical Bayes tests for a discrete normal distribution. *Stat. Probabil. Lett.* **44**, 241–249 (1999)
- Lisman, JHC, Van Zuylen, MCA: Note on the generation of the most probable frequency distribution. *Statistica Neerlandica* **26**, 19–23 (1972)
- Luceno, A: Discrete approximations to continuous univariate distributions – an alternative to simulation. *J. R. Statist. Soc. B* **61**(2), 345–352 (1999)

- Marshall, AW, Olkin, I: A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika* **84**(3), 641–652 (1997)
- Meyer, M, Poul Svante, A, Morling, EN: A gentle introduction to the discrete Laplace method for estimating Y-STR haplotype frequencies. arXiv:1304.2129v4 [stat.AP], 16 Oct 2013
- Mudholkar, GS, Srivastava, DK, Marshall, F: The exponentiated Weibull family: a reanalysis of the Bus-motor-failure data. *Technometrics* **37**(4), 436–445 (1995)
- Mukherjee, SP, Islam, A: A finite-range distribution of failure times. *Naval. Res. Logistics. Quart.* **30**(3), 487–491 (1983)
- Nakagawa, T, Osaki, S: The discrete Weibull distribution. *IEEE. Trans. Reliability.* **R-24**(5), 300–301 (1975)
- Nekoukhou, VM, Alamatsaz, MH, Bidram, H: Discrete generalized exponential distribution of the second type. *Stat. J. Theor. Appl. Stat.* **47**(4), 876–887 (2011). doi:10.1080/02331888.2011.633707
- Nekoukhou, VM, Alamatsaz, MH, Bidram, H: A discrete analog of the generalized exponential distribution. *Commun. Stat. Theor. Methods.* **41**, 2000–2013 (2012)
- Nooghabi, MS, Roknabady, AHR, Borzadaran, GRM: Discrete modified Weibull distribution. *Metron* **LXIX**, 207–222 (2011)
- Ord, JK: On a system of discrete distributions. *Biometrika* **54**, 649–656 (1967a)
- Ord, JK: On Families of discrete distributions. Ph. D. University of London, Thesis, London (1967b)
- Ord, JK: The discrete Student's t distribution. *Ann. Math. Stat.* **39**, 1513–1516 (1968)
- Ord, JK: On families of discrete distributions. Ph.D. thesis, Univ. of London (1967a)
- Padgett, WT, Spurrer, JD: On Discrete Failure Models. *IEEE. Transaction on Reliability,* R-34 (3), 253-256 (1985)
- Para, BA, Jan, TR: Discretization of Burr-Type III distribution. *J. Reliability. Stat. Stud.* **7**(2), 87–94 (2014)
- Pearson, K: Contributions to the mathematical theory of evolution I. Skew distribution in homogeneous material. *Philos. Trans. R. Soc. Lond. A.* **186**, 343–414 (1895)
- Perry, JN, Taylor, LR: Ades: New ecological families of species-specific frequency distributions that describe repeated spatial samples with an intrinsic power law variance-mean property. *J. Anim Ecol.* **54**, 931–953 (1985)
- Rezaei Roknabadi, AH: Some discrete life models. Ph. D. Thesis, Ferdowsi University of Mashad, Iran (2000)
- Rezaei Roknabadi, AH: Characterisation and model selections through reliability measures in the discrete case. *Stat. Probabil. Lett.* **43**, 197–206 (2006)
- Rezaei Roknabadi, AH, Mohtashami Borzadaran, GR, Khorashadzadeh, M: Some aspects of discrete telescopic hazard rate function in telescopic families. *Econ. Qual. Control.* **24**(1), 35–42 (2009)
- Roy, D: The discrete normal distribution. *Commun. Stat. Theor. Methods.* **32**(10), 1871–1883 (2003)
- Roy, D: Discrete Rayleigh distribution. *IEEE. Trans. Reliability.* **53**(2), 255–260 (2004)
- Roy, D, Dasgupta, T: A discretizing approach for evaluating reliability of complex systems under stress-strength model. *IEEE. Trans. Reliability.* **50**(2), 145–150 (2001)
- Roy, D, Ghosh, T: A new discretization approach with application in reliability estimation. *IEEE. Trans. Reliability.* **58**(3), 456–461 (2009)
- Roy, D, Gupta, PL: Classifications of discrete lives. *Microelectron. Reliab.* **32**(10), 1459–1473 (1992)
- Sarhan, A, Zaindin, M: Modified Weibull distribution. *Appl. Sci.* **11**, 123–136 (2009)
- Sato, H, Ikota, M, Aritoshi, S, Masuda, H: A new defect distribution meteorology with a consistent discrete exponential formula and its applications. *IEEE. Trans. Semicond. Manufactur.* **12**(4), 409–418 (1999)
- Shaw, W, Buckley, I: The alchemy of probability distributions: beyond Gram-Charlier expansions and a skew-kurtotic-normal distribution from a rank transmutation map. Research report. (2007)
- Siromoney, G: The general Dirichlet's series distribution. *J. Indian. Stat. Assoc.* **2 & 3**, 69–74 (1964)
- Stein, WE, Dattero, R: A new discrete Weibull distribution. *IEEE. Trans. Reliability.* **R-33**(2), 196–197 (1984)
- Sundt, B, Jewell, WS: Further results on recursive evaluation of compound distributions. *ASTIN. Bull.* **18**, 27–39 (1981)
- Szablowski, PJ: Discrete normal distribution and its relationship with Jacobi Theta functions. *Stat. Probabil. Lett.* **52**, 289–299 (2001)
- Tripathi, RC, Gurland, J: A general family of discrete distributions with hypergeometric probabilities. *J. R. Stat. Soc. B.* **39**, 349–356 (1977)
- Van Drop, JR, Kotz, S: A novel extension of the triangular distribution and its parameter estimation. *J. R. Stat. Soc. D.* **51**(1), 63–79 (2002a)
- Van Drop, JR, Kotz, S: The standard two sided power distribution and its properties: with application in financial engineering. *Am. Stat.* **56**(2), 90–99 (2002b)
- Willmot, GE: Sundt and Jewell's family of discrete distributions. *ASTIN. Bull.* **18**, 17–29 (1988)
- Xie, M, Tang, Y, Goh, TN: A modified Weibull extension with bathtub-shaped failure rate function. *Reliab. Eng. Syst. Saf.* **76**(3), 279–285 (2002)
- Zacks, S: Estimating the shift to wear-out of systems having exponential Weibull life. *Oper. Res.* **32**, 741–749 (1984)
- Zocchi, SS, Kokonendji, CC: On general continuous triangular and two-sided power distributions. *Communications in Statistics-Theory and Methods*, In Press (2013). doi:10.1080/03610926.2013.824102.
- Zornig, P, Altmann, G: Unified representation of Zipf distributions. *Comput. Statist. Data. Anal.* **19**, 461–473 (1995)