Original Research Article

A Discrete Analogue of Complementary Exponentiated-G Poisson Family of Distributions: Properties and Estimation

Abstract

This paper presented a two discrete family of life distributions called discrete complementary exponentiated-G Poisson family and discrete Zubair-G family as a special case from discrete complementary exponentiated-G Poisson family. Some basic distributional properties are derived. Such as hazard rate, moments, quantiles, order statistics and Rényi Entropy. A special sub-model of the discrete Zubair-G family, called the discrete Zubair Weibull distribution is considered in detail. Method of maximum likelihood are used under Type-II censored samples for estimating the parameters, survival, hazard rate and alternative hazard rate functions. Confidence intervals for the parameters are obtained. A simulation study is carried out to illustrate the theoretical results of the maximum likelihood estimation. Finally, the performance of the new distribution is compared with existing distributions using applications of three real data sets to show the suitability and flexibility of the proposed model.

Keywords*:* *Complementary exponentiated-G Poisson; Zubair-G distribution; Weibull distribution; Quantile function; Moments; Discrete order statistic; Type-II censored samples; Maximum likelihood method; Markov Chain Monte Carlo method.*

1. Introduction

Tahir and Cordeiro (2016) proposed *complementary exponentiated-G Poisson* (CEGP) family of distributions. The *cumulative distribution function* (cdf) of CEGP is given by

$$F(x;\alpha,\lambda,\xi)=\frac{e^{\alpha G^{\lambda}(x;\xi)}-1}{e^{\alpha}-1}. \qquad x\epsilon\mathbb{R},\quad \alpha,\lambda,\xi>0, \qquad (1)$$

where ξ is a vector of parameters and $G(x,\xi)$ is the cdf of the baseline model.

The *probability density function* (pdf), *survival function* (sf) and *hazard rate function* (hrf) of CEGP family are given, as follows:

$$f(x;\alpha,\lambda,\xi)=\frac{\lambda\alpha g(x;\xi)G^{\lambda-1}(x;\xi)e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}, \qquad x\epsilon\mathbb{R},\quad \alpha,\lambda,\xi>0, \qquad (2)$$

$$S(x;\alpha,\lambda,\xi)=\frac{e^{\alpha}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}. \qquad x\epsilon\mathbb{R},\quad \alpha,\lambda,\xi>0, \qquad (3)$$

and

$$h(x;\alpha,\lambda,\xi)=\frac{\lambda\alpha g(x;\xi)G^{\lambda-1}(x;\xi)e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-e^{\alpha G^{\lambda}(x;\xi)}},\qquad x\epsilon\mathbb{R},\quad \alpha,\lambda,\xi>0, \qquad (4)$$

where $g(x,\xi)$ is the pdf of the baseline model.

The discretization phenomenon generally arises when it becomes difficult to measure the life length of a device or product on a continuous scale. Such situations may arise when the observed lifetimes need to be recorded on a discrete scale instead of on a continuous analogue. In many practical situations, the reliability data are measured in terms of the numbers of runs, cycles or shocks the device sustains before it fails. For example, the number of times the devices are switched on/off, the lifetime of the switch is a *discrete random variable* (drv). Also, the number of voltages fluctuations, which an electronic or electrical item can endure before its failure, the life of weapon is measured by the number of rounds fired prior to failure, or the life of equipment is measured by the number of completed cycles or the number of times it operated before failure. Similarly, in survival analysis the sf may be a function of drv that is considered as a discrete version of the analogue *continuous random variable* (crv). Such as the length of stay in observation ward; when it is measured by the number of days, or the survival time that the leukemia patients survived since therapy may be counted by number of days or weeks.

Many researchers studied the general approach of discretization of some known continuous distributions for use as a discrete lifetime distribution. For example, Nakagawa and Osaki (1975) proposed a discrete Weibull distribution. Khan *et al.* (1989) introduced two discrete Weibull distributions and they presented a simple method to estimate the parameters for one of them. Mudholkar and Srivastava (1993) presented Exponentiated Weibull family for analyzing bathtub failure-rate data. Kemp (1997) introduced a discrete normal that is characterized by maximum entropy specified mean and variance. Roy (2003) introduced a discrete normal distribution. Inusah and Kozubowski (2006) obtained a discrete version of the Laplace distribution. Krishna and Pundir (2009) obtained the discrete Burr Type XII and Pareto distribution. Jazi *et al.* (2010) introduced discrete inverse Weibull distribution. Gomez-Deniz and Calderin-Ojeda (2011) proposed discrete Lindley. Al-Huniti and AL-Dayian (2012) introduced the discrete Burr Type III distribution. Nekoukhou and *et al*. (2012) proposed a discrete analogue of the generalized exponential distribution. Para and Jan (2014) introduced discretization of Burr-type III distribution. Hussain *et al.* (2016) introduced a two-parameter discrete Lindley distribution. Alamatsaz *et al*. (2016) proposed the discrete generalized Rayleigh distribution. Jayakumar and Sankaran (2017) introduced a new generalization of discrete Weibull distribution. Nurudeen and Abayomi (2017) presented a new discrete family of reduced modified Weibull distribution. Hegazy *et al*. (2018) introduced discrete Gompertz distribution. Helmy (2018) introduced discrete Burr Type II distribution. Para and Jan (2018) presented three parameter discrete generalized Inverse Weibull distribution. Maiti *et al*. (2018) proposed discrete X-Gamma distribution. Hegazy *et al*. (2019) introduced discrete inverted Kumaraswamy distribution. Elmorshedy and Eliwa (2019) presented a new two-parameter exponentiated discrete Lindley distribution. Almetwally (2020) proposed the discrete alpha power inverse Lomax distribution with application of COVID-19. Eliwa *et al*. (2020) introduced discrete Gompertz-G family of distributions. Almetwally *et al.* (2020) presented Managing Risk of Spreading COVID-19 in Egypt: modelling using a discrete Marshall–Olkin generalized Exponential distribution. Mable and Nurrohmah (2021) proposed the discrete Weibull-Geometric distribution. Recently Opone *et al.* (2021) introduced a discrete analogue of the continuous Marshall-Olkin Weibull distribution.

Although there are several discrete distributions in the statistical literature to model the above-mentioned situations, there is still a need to develop new discretized distribution that is suitable under different conditions. A flexible discrete generator of distributions is introduced, it's called *discrete* CEGP (DCEGP) family of distributions. Some reasons for introducing the DCEGP family are given below:

- **a.** To provide special models with all types of hazard rate functions.
- **b.** To propose more suitable models than other generated models under the same baseline distribution and other well-known models in the statistical literature.
- **c.** To improve the characteristics and flexibility of the present distributions.
- **d.** To introduce the extended version of the baseline distribution having closed form of the cdf, sf and hrf.

Discretization of Continuous Distribution

The general approach of discretizing a continuous variable can be used to construct a discrete model by introducing a grouping on the time axis. If the crv X has the sf, $S(x) = P(X > x)$, and times are grouped into unit intervals so that the discrete observed variable is *discrete x* $(dX) = X$, the largest integer part of X, the *probability mass function* (pmf) of dX can be written as

$$P(x) = P(dX = x) = P(x \leq dX < x + 1)$$
$$= S(x) - S(x + 1), \qquad x = 0,1,2, \dots . \qquad (5)$$

The pmf of the drv, dX can be viewed as discrete concentration of pdf of X. So, given any continuous distribution it is possible to construct corresponding discrete distribution using (5). One of the advantages of using this approach of discretizing is that the sf for discrete distributions has the same functional form of the sf for the continuous distributions; as a result, many reliability characteristics and properties remain unchanged. Thus, discretization of a continuous lifetime model according to this approach is an interesting and simple approach to derive a discrete lifetime model corresponding to the continuous one.

The rest of the paper is organized as follows: Section 2, the DCEGP family of distributions is introduced and some basic distributional properties. Estimation of the family parameter by methods of moments and ML in Section 3. The *discrete Zubair Weibull* (DZW) distribution is defined and some of its basic distributional properties and estimation of the model parameters in Section 4. Section 5 offers Monte Carlo simulation study to study the behavior of the maximum likelihood estimates and some concluding remarks. Also, three applications are analyzed to illustrate the suitability of the proposed model in Section 6.

2. Discrete Complementary Exponentiated-G Poisson Family

Using (5) dX can be viewed as the discrete analogue of the continuous CEGP family variable X, and it commonly said to follow DCEGP family of distributions with parameters α, λ and ξ, denoted by DCEGP (α, λ, ξ) family distribution, where the corresponding pmf of dX can be written as

$$P(x; \alpha, \lambda, \xi) = \frac{e^{\alpha G^{\lambda}(x+1;\xi)} - e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha} - 1}, \quad x = 0,1,2, \ldots, \quad \alpha, \lambda, \xi > 0. \quad (6)$$

The cdf, sf, hrf, *alternative* hrf (ahrf) and *reversed* hrf (rhrf), can be formulated as:

$$F(x; \alpha, \lambda, \xi) = \frac{e^{\alpha G^{\lambda}(x+1;\xi)} - 1}{e^{\alpha} - 1}, \quad x = 0,1,2, \ldots, \quad (7)$$

$$S(x; \alpha, \lambda, \xi) = \frac{e^{\alpha} - e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha} - 1}, \quad x = 0,1,2, \ldots, \quad (8)$$

$$h(x; \alpha, \lambda, \xi) = \frac{e^{\alpha G^{\lambda}(x+1;\xi)} - e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha} - e^{\alpha G^{\lambda}(x;\xi)}}, \quad x = 0,1,2, \ldots. \quad (9)$$

There are some problems associated with the definition of, three of the more notable ones are given below:

a. $h(x)$ is not additive for a competing risk model.

b. $h(x) \leq 1$ and it has the interpretation of a probability.

c. The cumulative hrf, $H(x) = \sum h(x) \neq -lnS(x)$. [see Lai (2013) and (2014)].

Therefore, it was necessary to find an alternative definition that is consistent with its continuous counterpart. Roy and Gupta (1992) give an excellent alternative definition of a discrete hrf denoted by $ah(x)$:

$$ah(x; \alpha, \lambda, \xi) = ln\left[\frac{s(x)}{s(x+1)}\right] = ln\left[\frac{e^{\alpha} - e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha} - e^{\alpha G^{\lambda}(x+1;\xi)}}\right], \quad x = 0,1,2, \ldots, \quad \alpha, \lambda, \xi > 0. \quad (10)$$

The two concepts $h(x)$ and $ah(x)$ have the same monotonic property, i.e., $ah(x)$ is increasing (decreasing) if $h(x)$ is increasing (decreasing).

and

$$rh(x; \alpha, \lambda, \xi) = \frac{e^{\alpha G^{\lambda}(x+1;\xi)} - e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha G^{\lambda}(x+1;\xi)} - 1}, \quad x = 0,1,2, \ldots, \quad \alpha, \lambda, \xi > 0. \quad (11)$$

2.1 Structural properties of discrete complementary exponentiated-G Poisson family

This subsection is devoted to obtaining some important distributional properties of DCEGP (α, λ, ξ) distribution, such as *mean residual lifetime function* (MRL), quantiles, moments, order statistics and Rényi Entropy.

2.1.1 Quantiles

The u^{th} quantile of dX, x_u, satisfies $p(X \leq x_u) \geq u$ and $p(X \geq x_u) \geq 1 - u$,i.e., $F(x_u - 1) < u \leq F(x_u)$. [see Rohatgi and Saleh (2001)].

proof

$p(X \le x_u) \ge u$, from (5)

$\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-1}{e^{\alpha}-1} \ge u$, hence

$$G(x_u+1;\xi) \ge \left(\frac{ln[u(e^{\alpha}-1)+1]}{\alpha}\right)^{\frac{1}{\lambda}}. \qquad (12)$$

Similarly, if $p(X \ge x_u) \ge 1-u$, one obtains

$$G(x_u;\xi) \ge \left(\frac{ln[u(e^{\alpha}-1)+1]}{\alpha}\right)^{\frac{1}{\lambda}}. \qquad (13)$$

2.1.2 The moments

a. The non-central moments

The r^{th} non-central moments of DCEGP distribution is given by

$$\acute{\mu_r} = \sum_x x^r \left[\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}\right], \qquad r = 1,2,.... \qquad (14)$$

The mean of DCEGP distribution is given by:

$$\acute{\mu_1} \equiv \mu = \sum_x x \left[\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}\right]. \qquad (15)$$

The second non-central moments of DCEGP distribution is given by

$$\acute{\mu_2} = \sum_x x^2 \left[\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}\right]. \qquad (16)$$

b. The central moments

The central moments of DCEGP family distribution can be obtained by using the relation between the central and non-central moments as given below

$$\mu_r = \sum_{j=0}^{r} \binom{r}{j} (-1)^j \mu^j \mu'_{r-j}, \qquad r = 1,2,... \; . \qquad (17)$$

The variance of DCEGP distribution can be obtained by using equations (16) and (17) as follows:

$$\mu_2 = \sum_x x^2 \left[\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}\right] - \left[\sum_x x \left[\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}\right]\right]^2. \qquad (18)$$

2.1.3 The order statistics

Let $F_i(x;\alpha,\lambda,\xi)$; the cdf of the i^{th} order statistics for random sample $X_1, X_2, \dots, X_n$, from the DCEGP (α,λ,ξ) [see Arnold *et al.* (2008)], is given by

$$F_{i:n}(x;\alpha,\lambda,\xi) = \sum_{r=i}^{n} \binom{n}{r} [\mathrm{F}(x;\alpha,\lambda,\xi)]^r \, [1-\mathrm{F}(x;\alpha,\lambda,\xi)]^{n-r}. \qquad (19)$$

Using the binomial expansion for $[1-\mathrm{F}(x;\alpha,\lambda,\xi)]^{n-r}$ and substituting (7) in (19).

Hence

$$F_{i:n}(x;\alpha,\lambda,\xi)=\sum_{r=i}^{n}\binom{n}{r}\sum_{j=0}^{n-r}\binom{n-r}{j}(-1)^{j}\left[\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-1}{e^{\alpha}-1}\right]^{r+j}. \tag{20}$$

The pmf of DCEGP can be obtained using (20), (see Arnold *et al*. (2008)).

If $i=1$ in (20) one can obtain the cdf of the first order statistics, as given below

$$F_{1}(x;\alpha,\lambda,\xi)=1-[1-\mathrm{F}(x;\alpha,\lambda,\xi)]^{n}$$

$$=1-\left[\frac{e^{\alpha}-e^{\alpha G^{\lambda}(x+1;\xi)}}{e^{\alpha}-1}\right]^{n}. \tag{21}$$

If $i=n$ in (20) one can obtain the cdf of the largest order statistics, as given below

$$F_{n}(x;\alpha,\lambda,\xi)=[\mathrm{F}(x;\alpha,\lambda,\xi)]^{n}$$

$$=\left[\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-1}{e^{\alpha}-1}\right]^{n}. \tag{22}$$

The pmf of i^{th} order statistics of DCEGP, is defined by

$$P_{i:n}(x;\alpha,\lambda,\xi)=\frac{n!}{(r-1)!(n-r)!}\int_{F(x-;\xi)}^{F(x;\xi)}v^{r-1}(1-v)^{n-r}dv. \tag{23}$$

Using the binomial expansion for $(1-v)^{n-r}$

$$P_{i:n}(x;\alpha,\lambda,\xi)=\frac{n!}{(r-1)!(n-r)!}\sum_{j=0}^{n-r}\binom{n-r}{j}\frac{(-1)^{j}}{s+j}\left\{[F(x;\alpha,\lambda,\xi)]^{s+j}-[F(x-;\alpha,\lambda,\xi)]^{s+j}\right\}. \tag{24}$$

The pmf of the smallest order statistics is obtained by substituting i=1 in (23) as follows:

$$P_{1}(x;\alpha,\lambda,\xi)=\left[\frac{e^{\alpha}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}\right]^{n}-\left[\frac{e^{\alpha}-e^{\alpha G^{\lambda}(x+1;\xi)}}{e^{\alpha}-1}\right]^{n}. \tag{25}$$

The pmf of the largest order statistics is obtained by substituting i=n in (23) as follows:

$$P_{n}(x;\alpha,\lambda,\xi)=\left[\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-1}{e^{\alpha}-1}\right]^{n}-\left[\frac{e^{\alpha G^{\lambda}(x;\xi)}-1}{e^{\alpha}-1}\right]^{n}. \tag{26}$$

2.1.4 Rényi Entropy

Entropy refers to the amount of uncertainty associated with a random variable X. It has many applications in several fields such as quantum information, econometrics, survival analysis, information theory, and computer science [see Rényi (1961)]. can be expressed as

$$I_{\eta}(x)=\frac{1}{1-\eta}log\sum_{x=0}^{\infty}f^{\eta}(x;\alpha,\lambda,\xi),$$

when $X\sim$DCEGP(α,λ,ξ), Rényi entropy can be derived as

$$I_{\eta}(x)=\frac{1}{1-\eta}log\sum_{x=0}^{\infty}\left(\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}\right)^{\eta},\quad \eta>0,(\eta\neq 1). \tag{27}$$

The Shannon entropy can be defined by

$$H(f) = E[-logf(x;\alpha,\lambda,\xi)] = -logf(x;\alpha,\lambda,\xi)\sum_{x=0}^{\infty} f(x;\alpha,\lambda,\xi). \quad (28)$$

When $X\sim$DCEGP(α,λ,ξ), Shannon entropy can be derived as

$$H(f) = -log\left(\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}\right)\sum_{x=0}^{\infty}\left(\frac{e^{\alpha G^{\lambda}(x+1;\xi)}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}\right). \quad (29)$$

It is observed that the Shannon entropy can be calculated as a special case of the Rényi entropy when $\eta \to 1$.

2.1.5 Mean residual lifetime function, mean time to failure, mean time between failure, and Availability

The MRL of DCEGP is defined as [see Lawless (1982) and Kemp (2004)]

$$MRL(x) = \frac{1}{s(x_0)}\sum_{x=x_0+1}^{\infty} s(x),$$

$$= \frac{e^{\alpha}-\sum_{x=x_0+1}^{\infty} e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-e^{\alpha G^{\lambda}(x_0;\xi)}}, \qquad x_0 = 0,1,2,\dots, \quad \alpha,\lambda,\xi > 0. \quad (30)$$

Mean Time to Failure (MTTF), *Mean Time between Failure* (MTBF) *and Availability* (AV) are reliability terms based on methods and procedures for lifecycle predictions for a product. MTTF, MTBF and Av are ways of providing a numeric value based on a compilation of data to quantify a failure rate and the resulting time of expected performance. In addition, in request to design and manufacture a maintainable system, it is necessary to predict the MTTF, MTBF, and AV. [see Eliwa *et al.* (2020)].

The MTBF and MTTF is given as

$$MTBF = \frac{-x}{\log\left[\frac{e^{\alpha}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}\right]}, \qquad x = 1,2,\dots, \quad (31)$$

$$MTTF = \sum_{x=1}^{\infty}\frac{e^{\alpha}-e^{\alpha G^{\lambda}(x;\xi)}}{e^{\alpha}-1}, \qquad x = 1,2,\dots. \quad (32)$$

The AV is considered as being the probability that the component is successful at time t, i.e.

$$AV = \frac{MTTF}{MTBF}. \quad (33)$$

3. Maximum Likelihood Estimation for Discrete Complementary Exponentiated-G Poisson Family

In this section, the ML estimators of the parameters, sf, hrf and ahrf based on Type-II censored samples are derived.

Let $(x_1, x_2, \dots, x_n)$ be a random sample from DCEGP (α, λ, ξ) family distribution with density function as $f(x; \alpha, \lambda, \xi)$. The likelihood function of DCEGP (α, λ, ξ) family based on Type-II censored sample corresponding (6) and (8) is:

$$L\left(\underline{\psi}; \underline{x}\right) \propto \left\{\prod_{i=1}^{r} P\left(x_{(i)}, \underline{\psi}\right)\right\}\left[S\left(x_{(r)}, \underline{\psi}\right)\right]^{n-r},$$

$$\propto \left\{\prod_{i=1}^{r} \frac{e^{\alpha G^{\lambda}(x_i+1;\xi)} - e^{\alpha G^{\lambda}(x_i;\xi)}}{e^{\alpha}-1}\right\}\left[\frac{e^{\alpha} - e^{\alpha G^{\lambda}(x_r;\xi)}}{e^{\alpha}-1}\right]^{n-r}, \tag{34}$$

where $\underline{\psi} = \alpha, \lambda, \xi$.

The natural logarithm of the likelihood function is given by

$$l \equiv lnL\left(\underline{\psi}; \underline{x}\right) \propto ln \prod_{i=1}^{r}\left[\frac{e^{\alpha G^{\lambda}(x_i+1;\xi)} - e^{\alpha G^{\lambda}(x_i;\xi)}}{e^{\alpha}-1}\right] + (n-r) ln\left[\frac{e^{\alpha} - e^{\alpha G^{\lambda}(x_r;\xi)}}{e^{\alpha}-1}\right],$$

$$\propto \sum_{i=1}^{r} ln\left[e^{\alpha G^{\lambda}(x_i+1;\xi)} - e^{\alpha G^{\lambda}(x_i;\xi)}\right] + (n-r) ln\left[e^{\alpha} - e^{\alpha G^{\lambda}(x_r;\xi)}\right]$$
$$-n \ln(e^{\alpha} - 1). \tag{35}$$

Depending on the invariance property, the ML estimators of sf, hrf and ahrf can be obtained by replacing α, λ and ξ with their corresponding ML estimators $\hat{\alpha}, \hat{\lambda}$ and $\hat{\xi}$, respectively, in (8), (9) and (10), as given below.

$$\hat{S}_{ML}(x_0) = \frac{e^{\hat{\alpha}} - e^{\hat{\alpha} G^{\hat{\lambda}}(x_0;\hat{\xi})}}{e^{\hat{\alpha}}-1}, \tag{36}$$

$$\hat{h}_{ML}(x_0) = \frac{e^{\hat{\alpha} G^{\hat{\lambda}}(x+1;\hat{\xi})} - e^{\hat{\alpha} G^{\hat{\lambda}}(x;\hat{\xi})}}{e^{\hat{\alpha}} - e^{\hat{\alpha} G^{\hat{\lambda}}(x_0;\hat{\xi})}}, \tag{37}$$

and

$$\widehat{ah}_{ML}(x_0) = ln\left(e^{\hat{\alpha}} - e^{\hat{\alpha} G^{\hat{\lambda}}(x_0;\hat{\xi})}\right) - l\,n\left(e^{\hat{\alpha}} - e^{\hat{\alpha} G^{\hat{\lambda}}(x_0+1;\hat{\xi})}\right). \tag{38}$$

4. The Discrete Zubair -G Family

In this section, we are interested with the *Zubair-G* (Z-G) family introduced by [Ahmed (2018)], which is considered as a special case from CEGP family when $\lambda = 2$.

Hence, a discrete analogue of Z-G family, called *discrete* Z-G (DZ-G) family of distributions, can be optioned when $\lambda = 2$ in DCEGP family.

The pmf, cdf, sf, hrf, ahrf and rhrf, can be written as:

$$P(x; \alpha, \xi) = \frac{e^{\alpha G^2(x+1;\xi)} - e^{\alpha G^2(x;\xi)}}{e^{\alpha}-1}, \qquad x = 0,1,2, \dots, \quad \alpha, \xi > 0, \tag{39}$$

$$F(x; \alpha, \xi) = \frac{e^{\alpha G^2(x+1;\xi)} - 1}{e^{\alpha}-1}, \qquad x = 0,1,2, \dots, \tag{40}$$

$$S(x;\alpha,\xi)=\frac{e^{\alpha}-e^{\alpha G^2(x;\xi)}}{e^{\alpha}-1}, \qquad x=0,1,2,\dots, \quad (41)$$

$$h(x;\alpha,\xi)=\frac{e^{\alpha G^2(x+1;\xi)}-e^{\alpha G^2(x;\xi)}}{e^{\alpha}-e^{\alpha G^2(x;\xi)}}, \qquad x=0,1,2,\dots, \quad (42)$$

$$ah(x;\alpha,\xi)=ln\left[\frac{e^{\alpha}-e^{\alpha G^2(x;\xi)}}{e^{\alpha}-e^{\alpha G^2(x+1;\xi)}}\right], \qquad x=0,1,2,\dots\ ,\ \alpha,\xi>0, \quad (43)$$

and

$$rh(x;\alpha,\lambda,\xi)=\frac{e^{\alpha G^2(x+1;\xi)}-e^{\alpha G^2(x;\xi)}}{e^{\alpha G^2(x+1;\xi)}-1}, \qquad x=0,1,2,\dots\ , \quad \alpha,\xi>0. \quad (44)$$

4.1 Discrete Zubair Weibull distribution

Weibull distribution is one of the best-known lifetime distributions. It is commonly used for analyzing biological, engineering, medical, and hydrological data sets. This model has been exhaustively used for describing hazard rates an important quantity of survival analysis. Weibull distribution is a reasonable choice due to its negatively and positively skewed density shapes. However, this distribution is not a good model for explanation phenomenon with non-monotone failure rates, which can be found on data from applications in reliability and biological studies. Thus, developed forms of the Weibull model have been sought in many applied areas. Adding parameters to a well-defined distribution has been indicated as a good methodology for providing more flexible new classes of distributions. The cdf and pdf of two parameter Weibull distribution is given by

$$G(x;\xi)=1-e^{-\beta x^{\theta}}\ , \qquad x,\xi>0. \quad (45)$$

$$g(x;\xi)=\theta\lambda x^{\theta-1}e^{-\beta x^{\theta}}. \qquad x,\xi>0. \quad (46)$$

With reparameterization $\gamma=e^{-\beta}$, $(0<\gamma<1)$ hence

$$G(x;\xi)=1-\gamma^{x^{\theta}}\ , \qquad x,\xi>0. \quad (47)$$

where $\xi=(\theta,\gamma)$. [see Murthy *et al*. (2004)]

Based on the cdf of the two parameter Weibull distribution. Then, the pmf and cdf of DZW distribution can be expressed as follows:

$$P(x;\alpha,\theta,\gamma)=\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2}-e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1}, \quad x=0,1,2,\dots, \quad \alpha,\theta>0,\ \ 0<\gamma<1, \quad (48)$$

and

$$F(x;\alpha,\theta,\gamma)=\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2}-1}{e^{\alpha}-1}, \qquad x=0,1,2,\dots, \quad \alpha,\theta>0,\ \ 0<\gamma<1. \quad (49)$$

The sf, hrf and ahrf of the DZW distribution are given by

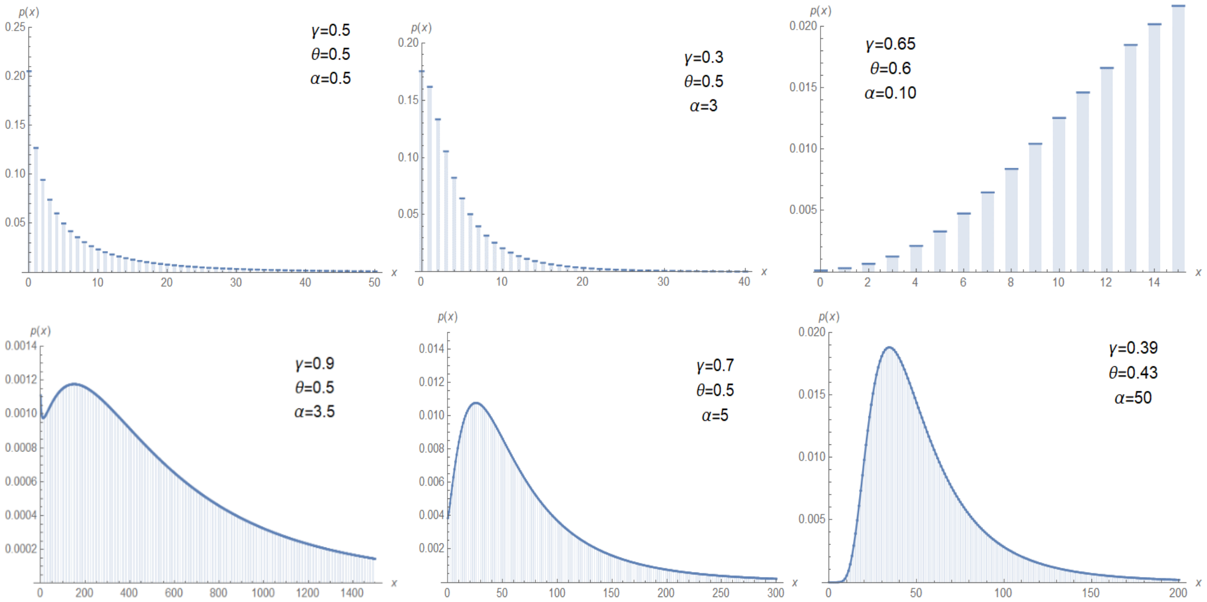
$$S(x;\alpha,\theta,\gamma)=\frac{e^{\alpha}-e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1}, \qquad x=0,1,2,\dots, \quad \alpha,\theta>0,\ \ 0<\gamma<1, \quad (50)$$

$$h(x,\alpha,\theta,\gamma)=\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^{2}}-e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^{2}}}{e^{\alpha}-e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^{2}}},\qquad x=0,1,2,\ldots,\quad \alpha,\theta>0,\ \ 0<\gamma<1,\qquad (51)$$

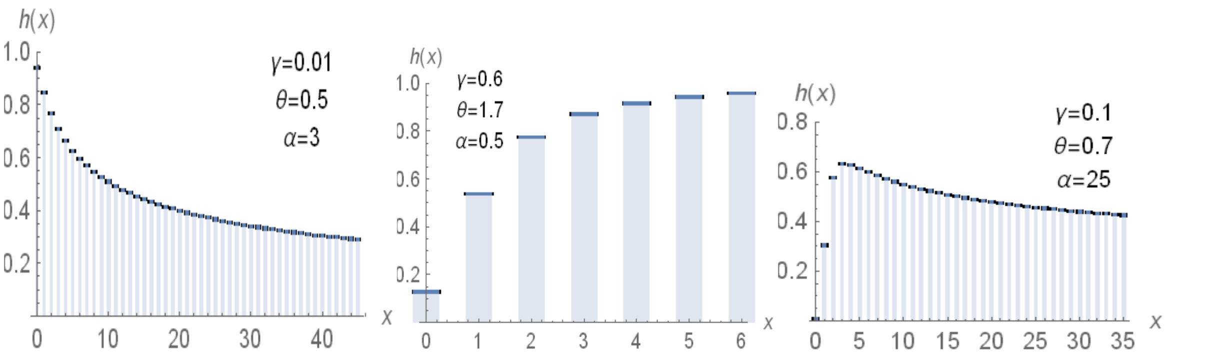
and

$$ah(x;\alpha,\theta,\gamma)=ln\left[\frac{e^{\alpha}-e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^{2}}}{e^{\alpha}-e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^{2}}}\right],\qquad x=0,1,2,\ldots,\quad \alpha,\theta>0,\ \ 0<\gamma<1.\qquad (52)$$

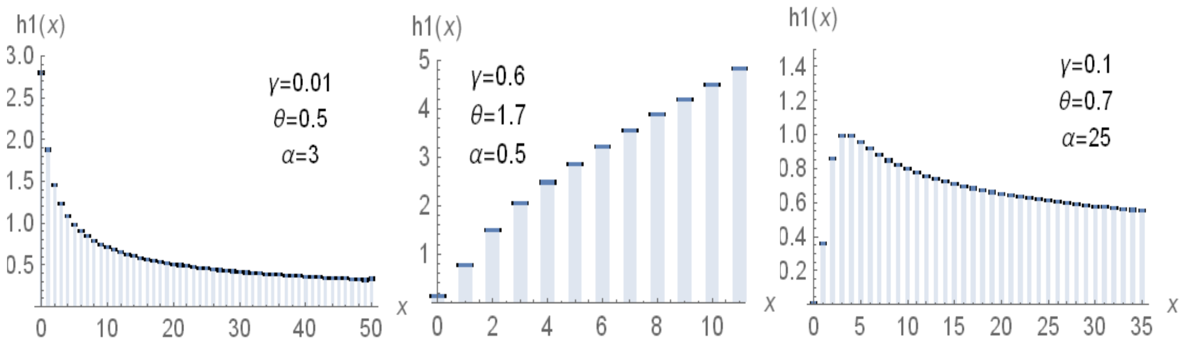
The plots of the pmf, hrf and ahrf of the DZW (α,θ,γ) for different values of the parameters are displayed, in Figures 1 - 3 respectively. Figures 1 displays some plots of the pmf for various parameter values. The pmf can take various shapes including decreasing, increasing, decreasing followed by unimodal, unimodal, left and right skewed, which gives it flexibility in handling most real data sets. Figures 2 and Figures 3 show some plots of the hrf and ahrf for various values of the parameters which are decreasing, increasing and unimodal shapes.



Figures 1. Plots of the pmf of DZW (α,θ,γ) for different values of α,θ and γ



Figures 2. Plots of the hrf of DZW (α,θ,γ) for different values of α,θ and γ



Figures 3. Plots of the ahrf of DZW (α, θ, γ) for different values of α, θ and γ

4.2 Special sub-model

In this subsection, two sub-models can be derived from DZW distribution given in (48); *discrete Zubair exponential* (DZEx) and *discrete Zubair Rayleigh* (DZR).

4.2.1 The discrete Zubair exponential distribution

The DZEx distribution is a special case of DZW, when $\theta = 1$ in (48), (49), (50), (51) and (52) with pmf, cdf, sf, hrf and ahrf are as follows:

$$P(x;\alpha,\gamma) = \frac{e^{\alpha\left(1-\gamma^{(x+1)}\right)^2} - e^{\alpha\left(1-\gamma^{(x)}\right)^2}}{e^{\alpha}-1}, \quad x = 0,1,2,\dots, \ \alpha > 0 \quad 0 < \gamma < 1. \quad (53)$$

$$F(x;\alpha,\gamma) = \frac{e^{\alpha\left(1-\gamma^{(x+1)}\right)^2} - 1}{e^{\alpha}-1}, \quad x = 0,1,2,\dots, \ \alpha > 0 \quad 0 < \gamma < 1, \quad (54)$$

$$S(x;\alpha,\gamma) = \frac{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x)}\right)^2}}{e^{\alpha}-1}, \quad x = 0,1,2,\dots, \ \alpha > 0 \quad 0 < \gamma < 1, \quad (55)$$

$$h(x;\alpha,\gamma) = \frac{e^{\alpha\left(1-\gamma^{(x+1)}\right)^2} - e^{\alpha\left(1-\gamma^{(x)}\right)^2}}{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x)}\right)^2}}, \quad x = 0,1,2,\dots, \ \alpha > 0 \quad 0 < \gamma < 1, \quad (56)$$

and

$$ah(x;\alpha,\gamma) = ln\left[\frac{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x+1)}\right)^2}}{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x)}\right)^2}}\right]. \quad x = 0,1,2,\dots, \ \alpha > 0 \quad 0 < \gamma < 1. \quad (57)$$

The plots of the pmf, hrf and ahrf of the DZEx (α, γ) for different values of the parameters are displayed, in Figures 4 – 6, respectively. Figure 4 shows some plots of the pmf for various parameter values. The plots shown that the pmf can be decreasing, increasing, unimodal, left and right skewed. Figure 5 and Figure 6 present some plots of the hrf and ahrf for various values of the parameters which are increasing and constant shape.

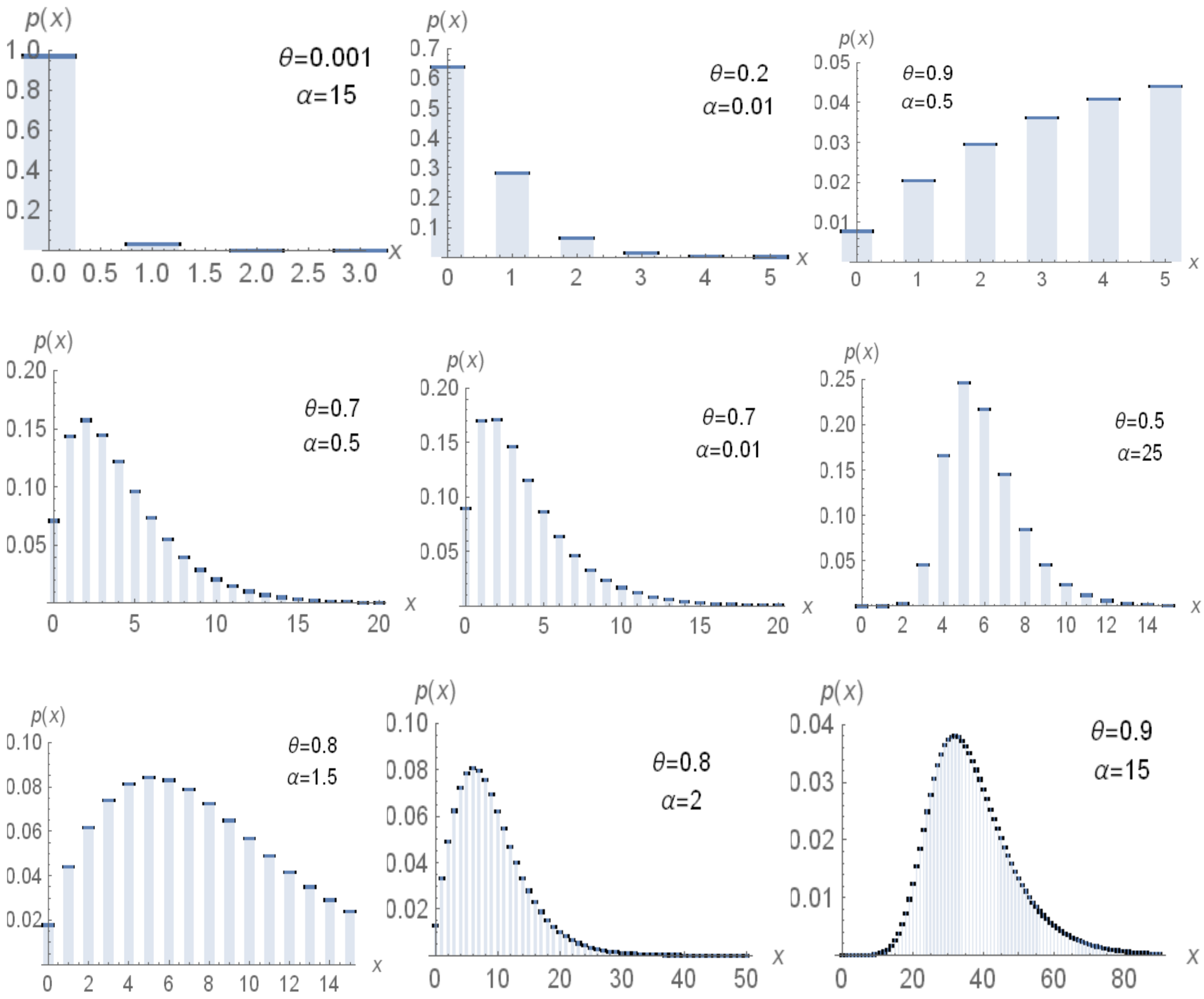


Figure 4. Plots of the pmf of DZEx (α, γ) for different values of α and γ

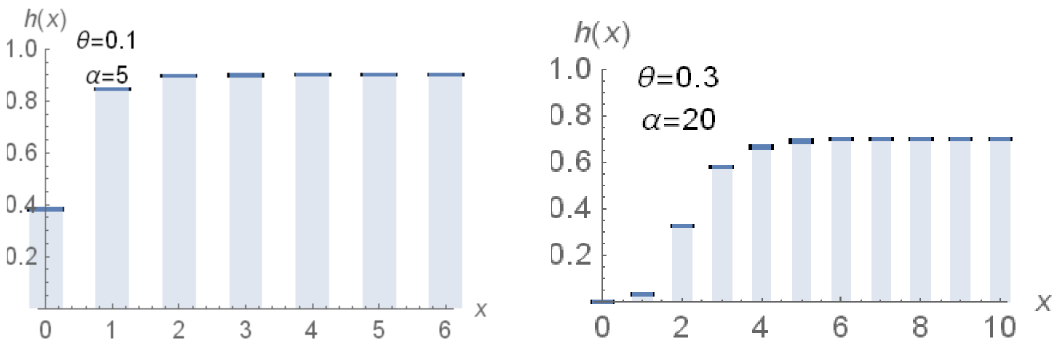


Figure 5. Plots of the hrf of DZEx (α, γ) for different values of α and γ

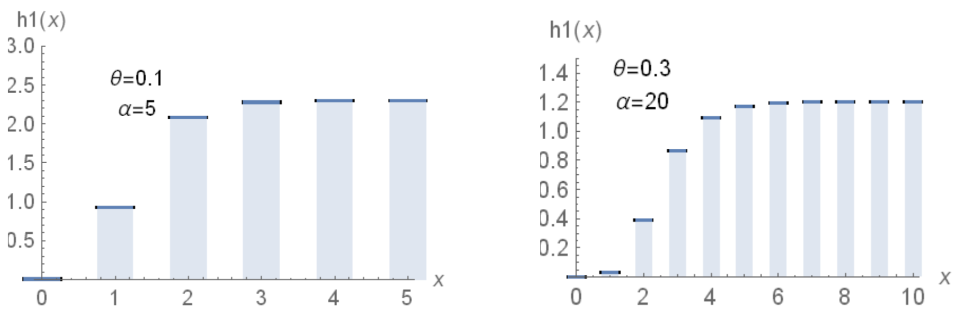


Figure 6. Plots of the ahrf of DZEx (α, γ) for different values of α and γ

4.2.2 The discrete Zubair Rayleigh distribution

The DZR distribution is a special case of DZW, when $\theta = 2$ (48), (49), (50), (51) and (52) with pmf, cdf, sf, hrf and ahrf are as follows:

$$P(x;\alpha,\gamma) = \frac{e^{\alpha\left(1-\gamma^{(x+1)^2}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^2}\right)^2}}{e^{\alpha}-1}, \quad x = 0,1,2,\dots, \quad \alpha > 0, \ 0 < \gamma < 1, \tag{58}$$

$$F(x;\alpha,\gamma) = \frac{e^{\alpha\left(1-\gamma^{(x+1)^2}\right)^2} - 1}{e^{\alpha}-1}, \quad x = 0,1,2,\dots, \quad \alpha > 0, \ 0 < \gamma < 1, \tag{59}$$

$$S(x;\alpha,\gamma) = \frac{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x)^2}\right)^2}}{e^{\alpha}-1}, \quad x = 0,1,2,\dots, \quad \alpha > 0, \ 0 < \gamma < 1, \tag{60}$$

$$h(x,\alpha,\gamma) = \frac{e^{\alpha\left(1-\gamma^{(x+1)^2}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^2}\right)^2}}{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x)^2}\right)^2}}, \quad x = 0,1,2,\dots, \quad \alpha > 0, \ 0 < \gamma < 1, \tag{61}$$

and

$$ah(x;\alpha,\gamma) = ln\left[\frac{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x)^2}\right)^2}}{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x+1)^2}\right)^2}}\right], \quad x = 0,1,2,\dots, \quad \alpha > 0, \ 0 < \gamma < 1. \tag{62}$$

The plots of the pmf, hrf and ahrf of the DZR (α,γ) for different values of the parameters are displayed, in Figures 7-9, respectively. Figure 7 presents some plots of the pmf for various parameter values. The plots of the pmf can be decreasing, increasing, unimodal, right skewed, and left skewed. Figures 8 and Figures 9 show some plots of the hrf and ahrf for various values of the parameters which are increasing shape.

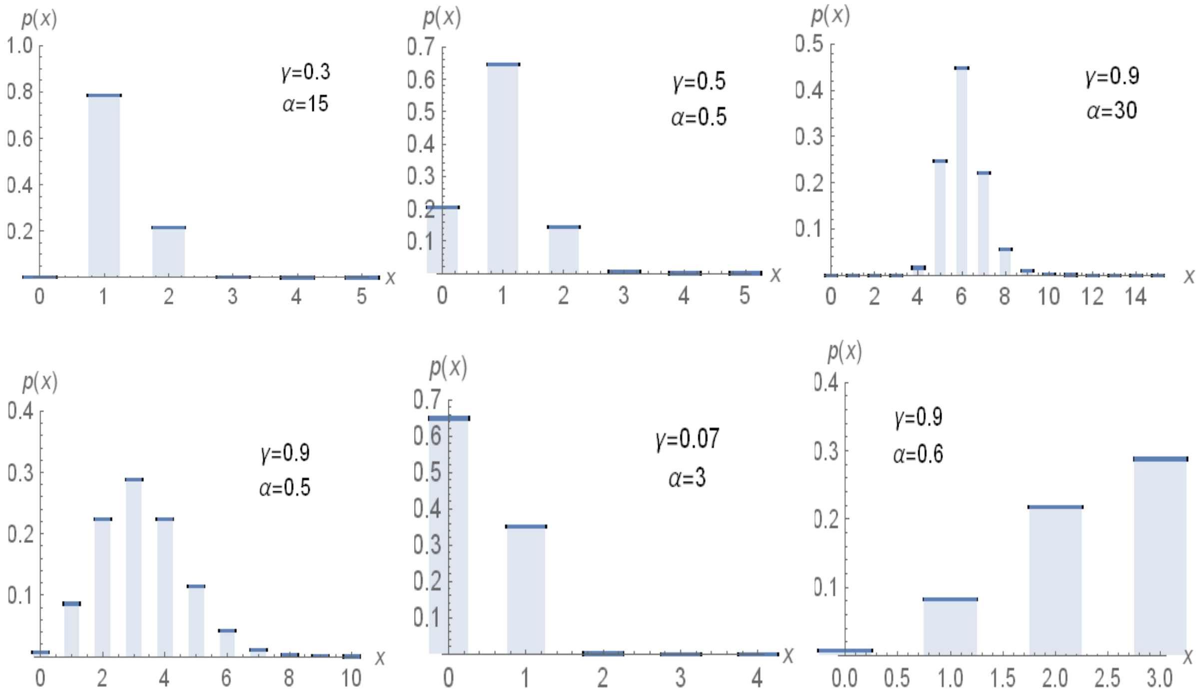
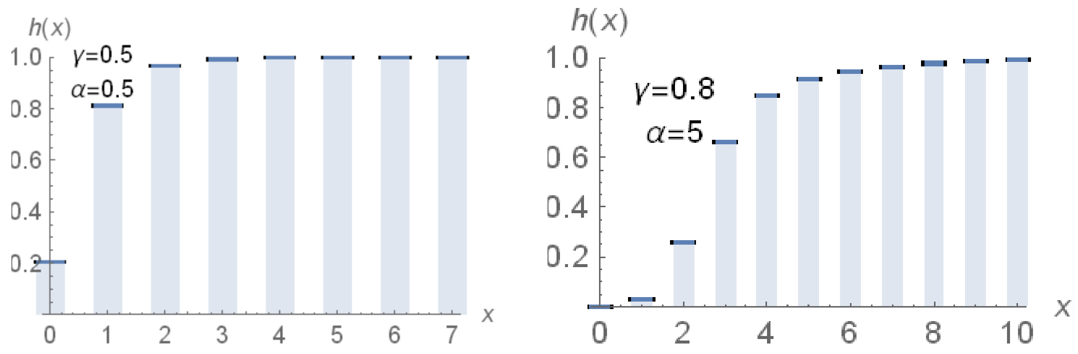
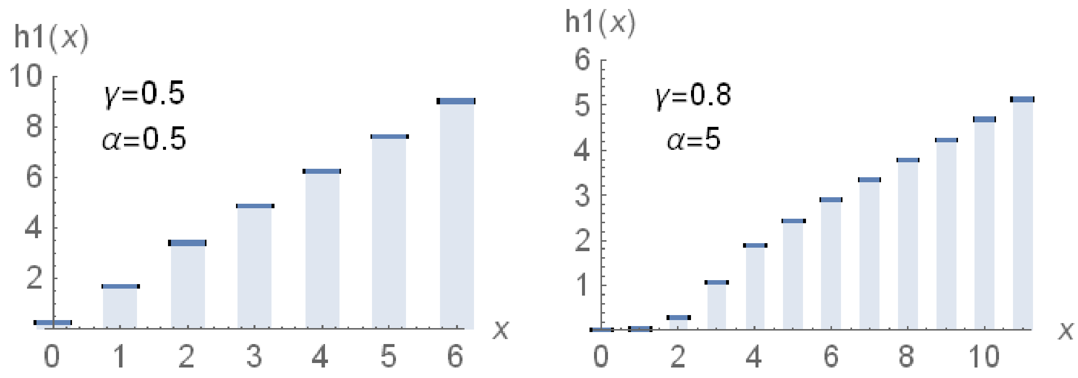


Figure 7. Plots of the pmf of DZR (α,γ) for different values of α and γ



Figures 8. Plots of the hrf of DZR (α, γ) for different values of α and γ



Figures 9. Plots of the ahrf of DZR (α, γ) for different values of α and γ

4.3 Some statistical properties of discrete Zubair Weibull distribution

In this subsection, some basic properties of the DZW distribution such as quantile function, order statistics, Rényi entropy, MRL, MTBF, MTTF and AV are derived.

4.3.1 Quantile function

The u^{th} quantile x_u, of DZW (α, θ, γ) is given by

$$x_u = \left\lceil \left(\frac{ln\left[1-\left(\frac{ln[u(e^{\alpha}-1)+1]}{\alpha}\right)^{\frac{1}{2}}\right]}{ln\gamma} \right)^{\frac{1}{\theta}} - 1 \right\rceil, \qquad 0 < u < 1. \qquad (63)$$

Where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x.

4.3.2 The moments and related concepts of discrete Zubair Weibull distribution

Using (47) and substituting in (14), The first four non-central moments of DZW distribution are

$$\acute{\mu}_1 = \mu = \sum_{x=0}^{\infty} x \left[\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1} \right], \quad (64)$$

$$\acute{\mu}_2 = \sum_{x=0}^{\infty} x^2 \left[\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1} \right], \quad (65)$$

$$\acute{\mu}_3 = \sum_{x=0}^{\infty} x^3 \left[\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1} \right], \quad (66)$$

$$\acute{\mu}_4 = \sum_{x=0}^{\infty} x^4 \left[\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1} \right]. \quad (67)$$

The variance of DZW distribution is

$$\mu_2 = \sum_{x=0}^{\infty} x^2 \left[\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1} \right] - \left[\sum_{x=0}^{\infty} x \left(\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1} \right) \right]^2. \quad (68)$$

Using the first four non-central moments in (64) - (67), can obtain the *skewness* (Sk) and *kurtosis* (Kur) from the following relations, respectively, as

$$SK = \frac{\acute{\mu}_3 - 3\acute{\mu}_2\acute{\mu}_1 + 2(\acute{\mu}_1)^3}{\mu_2^{3/2}}, \quad (69)$$

and

$$Kur = \frac{\acute{\mu}_4 - 4\acute{\mu}_2\acute{\mu}_1 + 6\acute{\mu}_2(\acute{\mu}_1)^2 - 3(\acute{\mu}_1)^4}{\mu_2^2}, \quad (70)$$

The *index of dispersion* (ID) and *coefficient of variation* (CV) for DZW distribution can be obtained as

$$\text{ID} = \frac{\mu_2}{\mu}, \qquad \text{and} \qquad CV = \frac{(\mu_2)^{1/2}}{\mu}. \quad (71)$$

Mathematica 11 is used to obtain these characteristics numerically. Some numerical values of mean, median, variance, SK, Kur and ID of DZW (α, θ, γ), DZEx (α, γ) and DZR (α, γ) for different values of the parameters are given in Tables 1 - 3, respectively.

Table 1
Mean, Median, Variance, Sk, Kur, ID and CV of DZW distribution for some values of the parameters (α, θ, γ)

Parameter			Descriptive Measures						
α	**θ**	**γ**	**Mean**	**Median**	**Variance**	**Sk**	**Kur**	**ID**	**CV**
2	**3**	**0.4**	0.6063	0.5241	0.1521	-1.6688	4.7073	0.2509	0.6432
10			0.9729	0.9952	0.1262	-0.3912	3.1761	0.1297	0.3651
15			1.1834	1.1882	0.0625	-0.1458	3.1187	0.0528	0.2113
0.5	**0.6**	**0.5**	3.6786	2.4736	16.3265	1.6756	5.1786	4.4382	1.0984
	0.8		3.0991	2.1581	6.4038	1.3080	4.8277	2.0663	0.8166
	1.2		1.8430	1.4712	2.7721	1.2580	3.9858	1.5041	0.9034
	1.3		1.4472	1.2847	1.4435	0.7478	2.7551	0.9974	0.8302
15	**2**	**0.8**	2.9206	2.8669	0.2317	0.6857	2.9190	0.0793	0.1648
		0.1	0.7575	0.6771	0.1335	0.4230	2.3084	0.1762	0.4823
		0.01	0.5777	0.6364	0.0579	-0.1010	1.6438	0.1002	0.4165

Table 2
Mean, Median, Variance, Sk, Kur, ID and CV of DZE distribution for some values of the parameters α when $\gamma = 0.3$

Parameter	Descriptive Measures						
α	**Mean**	**Median**	**Variance**	**Sk**	**Kur**	**IOD**	**CV**
0.5	1.0394	0.6653	1.0629	2.1233	8.0620	1.0226	0.9919
3	1.5567	0.9855	1.6784	1.7108	5.4039	1.0782	0.8322
5	1.7735	1.4192	1.0500	1.0041	3.4686	0.5920	0.5778
10	2.1822	2.0049	0.8670	0.6420	2.8952	0.3973	0.4267
20	3.0144	3.0357	1.1097	-0.1177	2.7618	0.3681	0.3494
30	3.1570	3.2251	0.8967	-0.1189	2.0882	0.2840	0.2999

Table 3
Mean, Median, Variance, Sk, Kur, ID and CV of DZR distribution for some values of the parameters α when $\gamma = 0.5$

Parameter	Descriptive Measures						
α	Mean	Median	Variance	Sk	Kur	IOD	CV
0.05	0.8376	0.6466	0.3532	0.6777	3.0209	0.4217	0.7095
0.5	0.8057	0.7278	0.3327	0.7751	3.1266	0.4129	0.7159
2	1.0844	0.9885	0.3133	0.9087	4.2529	0.2889	0.5162
5	1.3920	1.3562	0.3034	-0.0062	2.3149	0.2180	0.3957
10	1.6078	1.6752	0.2721	-0.0166	2.2827	0.1692	0.3244
25	1.9737	1.9841	0.2008	0.1330	2.0964	0.1017	0.2270

From Tables 1, 2 and 3 can conclude that:

- Table 1 shows that the mean and median increase, the variance, Sk and Kur decrease for fixed values of θ and γ when the parameter α increases. The mean, median, variance, Sk and Kur decrease for fixed values of θ and α when the parameter γ decreses. The mean, median, variance, Sk and Kur decrease for fixed values of γ and α when the parameter θ increases.
- One can deduce that the DZW distribution can be used to analyze under-dispersed, over-dispersed and equi-dispersed data sets.
- From Table 2, one can observe that the mean, median and variance increase for fixed values of γ when the parameter α is increases.
- Table 3 indicates that the mean and median increase, the variance decreases for fixed values of γ when the parameter α increases.

One can conclude that the DZW, DZEx and DZR distribution are flexible distributions and can be used in modeling different types of datasets which are suitable for modeling positively or negatively skewed and either platykurtic (Kur < 3) or leptokurtic (Kur > 3) data.

4.3.3 The order statistics of discrete Zubair Weibull distribution

The cdf of the i^{th} order statistics for a random sample $X_1, X_2, \dots, X_n$, from the DZW (α, θ, γ), is given by

$$F_{i:n}(x; \alpha, \xi) = \sum_{r=i}^{n} \binom{n}{r} \sum_{j=0}^{n-r} \binom{n-r}{j} (-1)^j \left[\frac{e^{\alpha\left(1-\gamma^{(x+1)^\theta}\right)^2} - 1}{e^{\alpha} - 1} \right]^{r+j}. \quad (72)$$

If $i = 1$ in (72) one can obtain the cdf of the first order statistics, as given below

$$F_1(x;\alpha,\theta,\gamma) = 1 - \left[\frac{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2}}{e^{\alpha}-1}\right]^n. \quad (73)$$

If $i = n$ in (72) one can obtain the cdf of the largest order statistics, as given below

$$F_n(x;\alpha,\theta,\gamma) = \left[\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2} - 1}{e^{\alpha}-1}\right]^n. \quad (74)$$

Then, the pmf of the i^{th} order statistics, is

$$P_{i:n}(x;\alpha,\theta,\gamma) = \frac{n!}{(r-1)!(n-r)!}\sum_{j=0}^{n-r}\binom{n-r}{j}\frac{(-1)^j}{s+j}\left\{\left[\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2}-1}{e^{\alpha}-1}\right]^{s+j} - \left[\frac{e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}-1}{e^{\alpha}-1}\right]^{s+j}\right\}, \quad (75)$$

The pmf of the smallest order statistics is obtained by substituting $i = 1$ in (75) as follows:

$$P_1(x;\alpha,\theta,\gamma) = \left[\frac{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1}\right]^n - \left[\frac{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2}}{e^{\alpha}-1}\right]^n. \quad (76)$$

The pmf of the largest order statistics is obtained by substituting $i = n$ in (75) as follows:

$$P_n(x;\alpha,\theta,\gamma) = \left[\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2}-1}{e^{\alpha}-1}\right]^n - \left[\frac{e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}-1}{e^{\alpha}-1}\right]^n. \quad (77)$$

4.3.4 Rényi Entropy

The Rényi entropy is

$$I_\eta(x) = \frac{1}{1-\eta}\log\sum_{x=0}^{\infty}\left(\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1}\right)^{\eta}. \quad \eta > 0, (\eta \neq 1). \quad (78)$$

The Shannon entropy can be defined by

$$H(f) = -\log\left(\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1}\right)\sum_{x=0}^{\infty}\left(\frac{e^{\alpha\left(1-\gamma^{(x+1)^{\theta}}\right)^2} - e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha}-1}\right). \quad (79)$$

The Shannon entropy can be calculated as a special case of the Rényi entropy when $\eta \to 1$.

4.3.5 Mean residual lifetime function, mean time to failure, mean time between failure, and Availability

The MRL, MTBF, MTTF and AV of the DZW are as follows:

$$MRL(x) = \frac{e^{\alpha} - \sum_{x=x_0+1}^{\infty} e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^2}}{e^{\alpha} - e^{\alpha\left(1-\gamma^{(x_0)^{\theta}}\right)^2}}, \qquad x_0 = 1,2,\ldots. \quad (80)$$

$$MTBF = \frac{-x}{\log\left(e^{\alpha}-e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^{2}}\right)-\log(e^{\alpha}-1)}, \qquad x = 1,2,\dots . \quad (81)$$

$$MTTF = \sum_{x=1}^{\infty}\frac{e^{\alpha}-e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^{2}}}{e^{\alpha}-1}, \qquad x = 1,2,\dots . \quad (82)$$

and

$$AV = \frac{\left[\sum_{x=1}^{\infty}\frac{e^{\alpha}-e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^{2}}}{e^{\alpha}-1}\right]\left[\log\left(e^{\alpha}-e^{\alpha\left(1-\gamma^{(x)^{\theta}}\right)^{2}}\right)-\log(e^{\alpha}-1)\right]}{-x}. \quad (83)$$

4.4 Maximum Likelihood Estimation of discrete Zubair Weibull distribution

In this subsection, the unknown parameters, sf, hrf and ahrf of the DZW distribution can be estimated using ML method.

The likelihood function of the DZW (α,θ,γ) distribution based on Type-II censored sample is given by

$$L\left(\underline{\varphi};\underline{x}\right) \propto \left\{\prod_{i=1}^{r}\frac{e^{\alpha\left(1-\gamma^{(x_i+1)^{\theta}}\right)^{2}}-e^{\alpha\left(1-\gamma^{(x_i)^{\theta}}\right)^{2}}}{e^{\alpha}-1}\right\}\left[\frac{e^{\alpha}-e^{\alpha\left(1-\gamma^{(x_r)^{\theta}}\right)^{2}}}{e^{\alpha}-1}\right]^{n-r}, \quad (84)$$

The natural logarithm of the likelihood function is given by

$$l \equiv lnL\left(\underline{\varphi};\underline{x}\right) = \sum_{i=1}^{r} ln\left[e^{\alpha\left(1-\gamma^{(x_i+1)^{\theta}}\right)^{2}} - e^{\alpha\left(1-\gamma^{(x_i)^{\theta}}\right)^{2}}\right]$$
$$+(n-r)ln\left[e^{\alpha} - e^{\alpha\left(1-\gamma^{(x_r)^{\theta}}\right)^{2}}\right] - n\ln(e^{\alpha}-1), \quad (85)$$

where $\underline{\varphi}$ is a vector of parameters α,θ and γ.

By differentiating the log likelihood function with respect to the parameters α, θ and γ as follows:

$$\frac{\partial l}{\partial \alpha} = \sum_{i=1}^{r}\frac{\left[\left(1-\gamma^{(x_i+1)^{\theta}}\right)^{2}e^{\alpha\left(1-\gamma^{(x_i+1)^{\theta}}\right)^{2}}\right]-\left[\left(1-\gamma^{(x_i)^{\theta}}\right)^{2}e^{\alpha\left(1-\gamma^{(x_i)^{\theta}}\right)^{2}}\right]}{e^{\alpha\left(1-\gamma^{(x_i+1)^{\theta}}\right)^{2}}-e^{\alpha\left(1-\gamma^{(x_i)^{\theta}}\right)^{2}}}$$
$$+\frac{(n-r)\left[e^{\alpha}-\left(1-\gamma^{(x_r)^{\theta}}\right)^{2}e^{\alpha\left(1-\gamma^{(x_r)^{\theta}}\right)^{2}}\right]}{e^{\alpha}-e^{\alpha\left(1-\gamma^{(x_r)^{\theta}}\right)^{2}}}-\frac{ne^{\alpha}}{e^{\alpha}-1}, \quad (86)$$

$$\frac{\partial l}{\partial \gamma} = \sum_{i=1}^{r} \frac{2\alpha\left\{\left[(x_i)^{\theta}\left(\gamma^{(x_i)^{\theta}-1}\right)\left(1-\gamma^{(x_i)^{\theta}}\right)e^{\alpha\left(1-\gamma^{(x_i)^{\theta}}\right)^2}\right]-\left[(x_i+1)^{\theta}\left(\gamma^{(x_i+1)^{\theta}-1}\right)\left(1-\gamma^{(x_i+1)^{\theta}}\right)e^{\alpha\left(1-\gamma^{(x_i+1)^{\theta}}\right)^2}\right]\right\}}{e^{\alpha\left(1-\gamma^{(x_i+1)^{\theta}}\right)^2}-e^{\alpha\left(1-\gamma^{(x_i)^{\theta}}\right)^2}} + \frac{2\alpha(n-r)\left(1-\gamma^{(x_r)^{\theta}}\right)(x_r)^{\theta}\left(\gamma^{(x_r)^{\theta}-1}\right)e^{\alpha\left(1-\gamma^{(x_r)^{\theta}}\right)^2}}{e^{\alpha}-e^{\alpha\left(1-\gamma^{(x_r)^{\theta}}\right)^2}}, \quad (87)$$

and

$$\frac{\partial l}{\partial \theta} = \sum_{i=1}^{r} \frac{2\alpha\left\{\left[\left(\gamma^{(x_i)^{\theta}}\right)\left(ln\left[\gamma^{(x_i)}\right]\right)\left(1-\gamma^{(x_i)^{\theta}}\right)e^{\alpha\left(1-\gamma^{(x_i)^{\theta}}\right)^2}\right]-\left[\left(\gamma^{(x_i+1)^{\theta}}\right)\left(ln\left[\gamma^{(x_i+1)}\right]\right)\left(1-\gamma^{(x_i+1)^{\theta}}\right)e^{\alpha\left(1-\gamma^{(x_i+1)^{\theta}}\right)^2}\right]\right\}}{e^{\alpha\left(1-\gamma^{(x_i+1)^{\theta}}\right)^2}-e^{\alpha\left(1-\gamma^{(x_i)^{\theta}}\right)^2}} + \frac{2\alpha(n-r)\left(1-\gamma^{(x_r)^{\theta}}\right)\left(\gamma^{(x_r)}\right)\left(ln\left[\gamma^{(x_r)}\right]\right)e^{\alpha\left(1-\gamma^{(x_r)^{\theta}}\right)^2}}{e^{\alpha}-e^{\alpha\left(1-\gamma^{(x_r)^{\theta}}\right)^2}}. \quad (88)$$

The ML estimates of the parameters α, θ and γ can be obtained by equating (86)-(88) to zeros and solving numerically.

The ML estimators of sf, hrf and ahrf can be derived using the invariance property by replacing α, θ and γ by their corresponding ML estimators $\hat{\alpha}$, $\hat{\theta}$and $\hat{\gamma}$, respectively, in (8), (9) and (10) as given below

$$\hat{S}_{ML}(x_0) = \frac{e^{\hat{\alpha}}-e^{\hat{\alpha}\left(1-\hat{\gamma}^{(x_0)^{\hat{\theta}}}\right)^2}}{e^{\hat{\alpha}}-1}, \quad (89)$$

$$\hat{h}_{ML}(x_0) = \frac{e^{\hat{\alpha}\left(1-\hat{\gamma}^{(x+1)^{\hat{\theta}}}\right)^2}-e^{\hat{\alpha}\left(1-\hat{\gamma}^{(x)^{\hat{\theta}}}\right)^2}}{e^{\hat{\alpha}}-e^{\hat{\alpha}\left(1-\hat{\gamma}^{(x_0)^{\hat{\theta}}}\right)^2}}, \quad (90)$$

and

$$\widehat{ah}_{ML}(x_0) = ln\left(e^{\hat{\alpha}} - e^{\hat{\alpha}\left(1-\hat{\gamma}^{(x_0)^{\hat{\theta}}}\right)^2}\right) - l\,n\left(e^{\hat{\alpha}} - e^{\hat{\alpha}\left(1-\hat{\gamma}^{(x_0+1)^{\hat{\theta}}}\right)^2}\right). \quad (91)$$

5. Numerical Results

This section aims to investigate the precision of the theoretical results of estimation based on the simulated and real data.

5.1 Simulation study

In this subsection, a simulation study is presented to illustrate the application of the various theoretical results developed in the previous section based on generated data from DZW (α, θ, γ) distribution, for different sample sizes (n=30, 60, 100, 200 and 500) using *number of replications* (NR)=1000. The computations are performed using Mathematica 11.

The following steps are used to generate Type-II censored sample from DZW (α, θ, γ) distribution as follows:

Step 1: Two different combinations of population parameter values are selected

I. $(\alpha = 3, \theta = 0.5, \gamma = 0.3)$ and
II. $(\alpha = 2, \theta = 0.5, \gamma = 0.9)$

based on two levels of $\frac{r}{n} \times 100$ percentage of uncensored observations Type-II censoring 70% and 100%. from the DZW distribution for different samples of size n.

Step 2: Generate 1000 random samples of size, n =30, 60, 100, 200 and 500 from DZW (α, θ, γ) distribution using the following transformation:

$$x_i = \left\lceil \left(\frac{ln\left[1 - \left(\frac{ln[u(e^{\alpha}-1)+1]}{\alpha}\right)^{\frac{1}{2}}\right]}{ln\gamma} \right)^{\frac{1}{\theta}} - 1 \right\rceil, \qquad i = 1,2,..,n$$

where u_i are random samples from uniform distribution and then taking the ceiling.

Step 3: The averages, *estimated risks* (ERs), *relative errors* (REs), variances of ML estimates of the parameters, sf, hrf and ahrf are computed as follows:

1. Average $= \frac{\sum_{i=1}^{NR} estimated\ value}{NR}$
2. Estimated risk $= \frac{\sum_{i=1}^{NR} (estimated\ value - true\ value)^2}{NR}$
3. Relative error $= \frac{\sqrt{ER(estimated\ value)}}{true\ value}$
4. Variance = $ER\ (estimated\ value) - bais^2\ (estimated\ value)$.

Step 4: The averages, ERs, REs, variances of ML estimates of the parameters, sf, hrf and ahrf are calculated for each model parameters and for each sample size.

Table 4
Average, ERs, REs, Variances of the ML estimates and 95% confidence intervals of the parameters α, θ, and γ based on Type-II censoring of DZW
$(NR = 1000,\ \alpha = 3, \theta = 0.5\ and\ \gamma = 0.3)$

n	r	Parameters	Average	ER	RE	Variance	UL	LL	Length
30	21	α	4.6297	2.7264	0.5504	0.0704	5.1499	4.1095	1.0404
		θ	0.6780	0.0400	0.4001	0.0084	0.8571	0.4988	0.3583
		γ	0.3726	0.0064	0.2674	0.0012	0.4397	0.3054	0.1342
	30	α	4.6025	2.6129	0.5388	0.0448	5.0174	4.1876	0.8298
		θ	0.5899	0.0125	0.2240	0.0045	0.7209	0.4589	0.2620
		γ	0.3454	0.0034	0.1943	0.0013	0.4171	0.2736	0.1436
60	42	α	4.5661	2.4794	0.5249	0.0267	4.8864	4.2458	0.6406
		θ	0.6708	0.0336	0.3668	0.0044	0.8015	0.5402	0.2613
		γ	0.3747	0.0061	0.2610	0.0006	0.4208	0.3285	0.0923
	60	α	4.5607	2.4525	0.5220	0.0168	4.8148	4.3066	0.5082
		θ	0.5867	0.0097	0.1965	0.0021	0.6774	0.4960	0.1814
		γ	0.3454	0.0027	0.1729	0.0006	0.3946	0.2961	0.0985
100	70	α	4.5318	2.3592	0.5120	0.0129	4.7539	4.3096	0.4444
		θ	0.6652	0.0301	0.3472	0.0028	0.7696	0.5608	0.2088
		γ	0.3750	0.0059	0.2561	0.0003	0.4076	0.3425	0.0651
	100	α	4.5437	2.3932	0.5157	0.0101	4.7403	4.3472	0.3931
		θ	0.5809	0.0079	0.1774	0.0013	0.6522	0.5096	0.1426
		γ	0.3447	0.0024	0.1620	0.0003	0.3822	0.3072	0.0750
200	140	α	4.5077	2.2786	0.5032	0.0055	4.6535	4.3618	0.2918
		θ	0.6622	0.0279	0.3340	0.0016	0.7399	0.5846	0.1553
		γ	0.3750	0.0058	0.2533	0.0002	0.3992	0.3508	0.0484
	200	α	4.5344	2.3596	0.5120	0.0051	4.6749	4.3940	0.2809
		θ	0.5795	0.0070	0.1670	0.0007	0.6297	0.5292	0.1005
		γ	0.3444	0.0022	0.1550	0.0002	0.3713	0.3176	0.0537
500	350	α	4.4910	2.2254	0.4973	0.0023	4.5840	4.3980	0.1860
		θ	0.6613	0.0266	0.3264	0.0006	0.7101	0.6125	0.0975
		γ	0.3750	0.0057	0.2517	0.0001	0.3915	0.3586	0.0330
	500	α	4.5277	2.3357	0.5094	0.0020	4.6149	4.4404	0.1745
		θ	0.5768	0.0062	0.1573	0.0003	0.6102	0.5433	0.0670
		γ	0.3435	0.0020	0.1477	0.0001	0.3602	0.3267	0.0336

Table 5

Average, ERs, REs, Variances of the ML estimates and 95% confidence intervals of the $S(x_0), h(x_0)$ and $ah(x_0)$ based on Type-II censoring of DZW ($NR = 1000, x_0 = 1,\ \alpha = 3, \theta = 0.5\ and\ \gamma = 0.3$)

n	r	sf, hrf and ahrf	Average	ER	RE	Variance	UL	LL	Length
30	21	$S(x_0)$	0.9461	0.0152	0.1493	0.0004	0.9837	0.9086	0.0751
		$h(x_0)$	0.1305	0.0055	0.3789	0.0012	0.1986	0.0623	0.1362
		$ah\ (x_0)$	0.1406	0.0077	0.4025	0.0017	0.2209	0.0603	0.1605
	30	$S(x_0)$	0.9344	0.0125	0.1354	0.0004	0.9733	0.8955	0.0778
		$h(x_0)$	0.1269	0.0056	0.3810	0.0008	0.1818	0.0720	0.1099
		$ah\ (x_0)$	0.1362	0.0078	0.4046	0.0001	0.1999	0.0726	0.1273
60	42	$S(x_0)$	0.9466	0.0151	0.1488	0.0001	0.9705	0.9227	0.0479
		$h(x_0)$	0.1273	0.0053	0.3714	0.0006	0.1737	0.0809	0.0928
		$ah\ (x_0)$	0.1365	0.0075	0.3955	0.0008	0.1904	0.0826	0.1078
	60	$S(x_0)$	0.9347	0.0123	0.1345	0.0001	0.9600	0.9094	0.0506
		$h(x_0)$	0.1263	0.0052	0.3693	0.0004	0.1637	0.0889	0.0747
		$ah\ (x_0)$	0.1353	0.0073	0.3937	0.0005	0.1783	0.0922	0.0861
100	70	$S(x_0)$	0.9464	0.0149	0.1482	0.0001	0.9633	0.9294	0.0339
		$h(x_0)$	0.1257	0.0052	0.3591	0.0002	0.1566	0.0949	0.0617
		$ah\ (x_0)$	0.1345	0.0073	0.3928	0.0003	0.1700	0.0991	0.0708
	100	$S(x_0)$	0.9343	0.0122	0.1337	0.0001	0.9542	0.9145	0.0397
		$h(x_0)$	0.1251	0.0052	0.3688	0.0002	0.1520	0.0983	0.0538
		$ah\ (x_0)$	0.1338	0.0074	0.3939	0.0002	0.1646	0.1030	0.0615
200	140	$S(x_0)$	0.9460	0.0148	0.1475	0.0001	0.9584	0.9335	0.0249
		$h(x_0)$	0.1255	0.0051	0.3649	0.0001	0.1475	0.1035	0.0440
		$ah\ (x_0)$	0.1342	0.0073	0.3901	0.0002	0.1594	0.1089	0.0505
	200	$S(x_0)$	0.9343	0.0121	0.1334	0.0001	0.9487	0.9199	0.0288
		$h(x_0)$	0.1250	0.0052	0.3665	0.0001	0.1451	0.1050	0.0401
		$ah\ (x_0)$	0.1336	0.0073	0.3918	0.0001	0.1567	0.1107	0.0459
500	350	$S(x_0)$	0.9457	0.0147	0.1470	1.7654×10^{-5}	0.9539	0.9374	0.0165
		$h(x_0)$	0.1257	0.0050	0.3613	0.0001	0.1396	0.1118	0.0278
		$ah\ (x_0)$	0.1343	0.0071	0.3868	0.0001	0.1502	0.1184	0.0318
	500	$S(x_0)$	0.9339	0.0120	0.1327	2.0876×10^{-5}	0.9428	0.9249	0.0179
		$h(x_0)$	0.1251	0.0051	0.3638	3.7328×10^{-5}	0.1371	0.1131	0.0239
		$ah\ (x_0)$	0.1336	0.0072	0.3893	4.8802×10^{-5}	0.1473	0.1200	0.0274

Table 6

Average, ERs, REs, Variances of the ML estimates and 95% confidence intervals of the parameters α, θ and γ based on Type-II censoring of DZW

($NR = 1000,\ \alpha = 2, \theta = 0.5\ and\ \gamma = 0.9$)

n	r	parameters	Average	ER	RE	Variance	UL	LL	Length
30	21	α	2.6655	1.1243	0.5302	0.6813	4.2834	1.0477	3.2357
		θ	0.4820	0.0005	0.0468	0.0002	0.5113	0.4527	0.0586
		γ	0.8781	0.0006	0.0271	0.0001	0.8992	0.8570	0.0422
	30	α	2.5837	0.9255	0.4810	0.5848	4.0825	1.0849	2.9976
		θ	0.4810	0.0005	0.0467	0.0002	0.5075	0.4544	0.0531
		γ	0.8785	0.0005	0.0256	6.8749×10^{-5}	0.8948	0.8623	0.0325
60	42	α	2.5921	0.6888	0.4150	0.3382	3.7320	1.4523	2.2797
		θ	0.4814	0.00047	0.0432	0.0001	0.5031	0.4598	0.04322
		γ	0.8791	0.00046	0.0240	2.9596×10^{-5}	0.8898	0.8685	0.0213
	60	α	2.5627	0.6402	0.4001	0.3235	3.6775	1.4479	2.2296
		θ	0.4804	0.00046	0.0431	7.9796×10^{-5}	0.4979	0.4629	0.0350
		γ	0.8791	0.00045	0.0242	3.6383×10^{-5}	0.8902	0.8673	0.0236
100	70	α	2.5212	0.5010	0.3378	0.1849	3.3641	1.6783	1.6858
		θ	0.4806	0.00044	0.0420	6.2843×10^{-5}	0.4961	0.4650	0.0311
		γ	0.8798	0.00042	0.0229	1.1530×10^{-5}	0.8875	0.8721	0.0153
	100	α	2.5310	0.4633	0.3403	0.1813	3.3656	1.6963	1.6693
		θ	0.4803	0.00043	0.0415	4.3745×10^{-5}	0.4933	0.4674	0.0259
		γ	0.8797	0.00042	0.0229	1.4123×10^{-5}	0.8871	0.8723	0.0147
200	140	α	2.5094	0.3565	0.2985	0.0970	3.1199	1.8989	1.2211
		θ	0.4804	0.00041	0.0406	2.8335×10^{-5}	0.4909	0.4700	0.0209
		γ	0.8801	0.00040	0.0223	6.8408×10^{-6}	0.8852	0.8749	0.0103
	200	α	2.4973	0.3406	0.2918	0.0933	3.0960	1.8986	1.1974
		θ	0.4803	0.0004	0.0405	2.1318×10^{-5}	0.4893	0.4712	0.0181
		γ	0.8801	0.0004	0.0222	6.5930×10^{-6}	0.8852	0.8751	0.0100
500	350	α	2.4817	0.2699	0.2597	0.0379	2.8632	2.1003	0.7628
		θ	0.4801	0.0004	0.0404	9.6757×10^{-6}	0.4862	0.4740	0.0122
		γ	0.8804	0.00039	0.0219	2.3514×10^{-6}	0.8834	0.8774	0.0060
	500	α	2.4767	0.2623	0.2561	0.0351	2.8438	2.1096	0.7342
		θ	0.4801	0.0004	0.0402	8.6952×10^{-6}	0.4859	0.4743	0.0116
		γ	0.8804	0.00038	0.0218	2.2952×10^{-6}	0.8834	0.8775	0.0059

Table 7

Average, ERs, REs, Variances of the ML estimates and 95% confidence intervals of the $S(x_0), h(x_0)$ and $ah(x_0)$ based on Type-II censoring of DZW

$(NR = 1000, x_0 = 1,\ \alpha = 2, \theta = 0.5\ and\ \gamma = 0.9)$

n	r	sf, hrf and ahrf	Average	ER	RE	Variance	UL	LL	Length
30	21	$S(x_0)$	0.9967	2.4992×10^{-6}	0.0016	2.4869×10^{-6}	0.9998	0.9936	0.0062
		$h(x_0)$	0.0029	1.8138×10^{-6}	0.4545	1.8096×10^{-6}	0.0055	0.0003	0.0053
		$ah\ (x_0)$	0.0029	1.8270×10^{-6}	0.4556	1.8236×10^{-6}	0.0056	0.0003	0.0053
	30	$S(x_0)$	0.9966	2.4712×10^{-6}	0.0016	2.4058×10^{-6}	0.9996	0.9935	0.0061
		$h(x_0)$	0.0030	1.8129×10^{-6}	0.4544	1.8091×10^{-6}	0.0057	0.0004	0.0052
		$ah\ (x_0)$	0.0030	1.8270×10^{-6}	0.4555	1.8231×10^{-6}	0.0057	0.0004	0.0052
60	42	$S(x_0)$	0.9967	1.2893×10^{-6}	0.0011	1.2789×10^{-6}	0.9990	0.9945	0.0044
		$h(x_0)$	0.0029	9.4929×10^{-7}	0.3288	9.4392×10^{-7}	0.0048	0.0010	0.0038
		$ah\ (x_0)$	0.0029	9.5595×10^{-7}	0.3295	9.5062×10^{-7}	0.0048	0.0010	0.0038
	60	$S(x_0)$	0.9967	1.2970×10^{-6}	0.0011	1.2701×10^{-6}	0.9989	0.9945	0.0044
		$h(x_0)$	0.0029	9.3891×10^{-7}	0.3270	9.3830×10^{-7}	0.0048	0.0010	0.0038
		$ah\ (x_0)$	0.0029	9.4539×10^{-7}	0.3277	9.4479×10^{-7}	0.0048	0.0010	0.0038
100	70	$S(x_0)$	0.9967	7.2310×10^{-7}	0.0009	6.9827×10^{-7}	0.9983	0.9950	0.0033
		$h(x_0)$	0.0029	5.1639×10^{-7}	0.2425	5.1557×10^{-7}	0.0043	0.0015	0.0028
		$ah\ (x_0)$	0.0029	5.1973×10^{-7}	0.2430	5.1892×10^{-7}	0.0044	0.0015	0.00282
	100	$S(x_0)$	0.9967	7.0701×10^{-7}	0.0008	6.8655×10^{-7}	0.9983	0.9951	0.0032
		$h(x_0)$	0.0029	5.1195×10^{-7}	0.2415	5.1017×10^{-7}	0.0043	0.0015	0.0027
		$ah\ (x_0)$	0.0029	5.1518×10^{-7}	0.2419	5.1342×10^{-7}	0.0043	0.0015	0.00280
200	140	$S(x_0)$	0.9967	3.8034×10^{-7}	0.0006	3.6268×10^{-7}	0.9979	0.9955	0.0024
		$h(x_0)$	0.0029	2.7290×10^{-7}	0.1763	2.7048×10^{-7}	0.0039	0.0019	0.0020
		$ah\ (x_0)$	0.0029	2.7459×10^{-7}	0.1766	2.7218×10^{-7}	0.0039	0.0019	0.0020
	200	$S(x_0)$	0.9967	3.8255×10^{-7}	0.0006	3.5804×10^{-7}	0.9979	0.9955	0.0023
		$h(x_0)$	0.0029	2.6756×10^{-7}	0.1746	2.6668×10^{-7}	0.0039	0.0019	0.0020
		$ah\ (x_0)$	0.0029	2.6921×10^{-7}	0.1749	2.6834×10^{-7}	0.0039	0.0019	0.0020
500	350	$S(x_0)$	0.9967	1.6061×10^{-7}	0.0004	1.3750×10^{-7}	0.9974	0.9959	0.0015
		$h(x_0)$	0.0029	1.0475×10^{-7}	0.1092	1.0356×10^{-7}	0.0036	0.0023	0.0013
		$ah\ (x_0)$	0.0029	1.0537×10^{-7}	0.1094	1.0417×10^{-7}	0.0036	0.0023	0.0013
	500	$S(x_0)$	0.9967	1.5455×10^{-7}	0.0003	1.2888×10^{-7}	0.9974	0.9960	0.0014
		$h(x_0)$	0.0029	9.8354×10^{-8}	0.1058	9.7620×10^{-8}	0.0035	0.0023	0.0012
		$ah\ (x_0)$	0.0029	9.8938×10^{-8}	0.1060	9.8202×10^{-8}	0.0035	0.0023	0.0012

5.2 Concluding Remarks

From Tables 4, 5, 6 and 7, one can deduce that:

1. The REs, ERs, and variances of the MLEs of the parameters $\alpha, \theta, and\, \gamma$ decrease when the sample size n increases. also, the REs, ERs, and and variances of the MLEs of the sf, hrf and the ahrf decrease when the sample size increases. Also, the lengths of the confidence intervals get shorter when the sample size increases.
2. The REs, ERs, and variances of the ML estimates of the parameters, sf, hrf, and ahrf estimates decrease when the level of censoring decreases. The lengths of the confidence intervals become narrower when the sample size increases.
3. In general, all the results of REs, ERs and variances obtained for complete sample sizes, less than the corresponding results for censored samples. Also, results perform better when *n* and *r* larger.

6. Applications

This Section is devoted to illustrate the flexibility and applicability of the proposed distribution using three real data sets. The first two real data sets are discrete real data sets whereas the third is count data set.

6.1 Discrete data

For each data set, the DZW distribution is compared with some existing distributions such as *discrete Weibull* (DW) introduced by [Nakagawa and Osaki (1975)], *generalization of discrete Weibull* (GDW) presented by [Jayakumar and Sankran (2017)], *discrete Weibull-geometric* (DW-G) presented by [Mabel *et al.* (2021)], *discrete Marshall-Olkin Weibull* (DMOW) proposed by [Opone *et al.* (2021)] and *discrete Marshall-Olkin generalized exponential* (DMOGE) presented by [Almetwally *et al.* (2020)]. The comparison was presented based on some criteria. These criteria are *Kolmogorov-Smirnov* (K-S) statistic and its p-value, the -2*log-likelihood* ($-2lnL$), *Akaike information criterion* (*AIC*), *Bayesian information criterion* (*BIC*) and *Akaike information criterion with correction* (*AICc*):

$$AIC = 2k - 2ln(L)$$

$$AICc = AIC + 2\frac{k(k+1)}{n-k-1}$$

$$BIC = kln(n) - 2ln(L)$$

where k denotes the number of the estimated parameters, lnL is the log-likelihood function evaluated at the maximum likelihood estimates, and n is the sample size. The distribution with the smallest values of these statistics is the best for fitting the data.

Data set I

The first data set consists of the 2003 final examination marks of 48 slow space students in mathematics in the Indian Institute of Technology at Kanpur. This data set is taken from Gupta and Kundu (2009). The data are**: 29, 25, 50, 15, 13, 27, 15, 18, 7, 7, 8, 19, 12, 18, 5, 21, 15, 86, 21, 15, 14, 39, 15, 14, 70, 44, 6, 23, 58, 19, 50, 23, 11, 6, 34, 18, 28, 34, 12, 37, 4, 60, 20, 23, 40, 65, 19 and 31.**

Table 8 presents the ML estimates corresponding *standard errors* (SEs) as well as $-2lnL$, AIC, BIC, CAIc, K-S statistic and its p-value. The result in this table indicates that the compared distributions fit for this data, the DZW distribution has the lowest *K–S* value and highest p-value. Also, the smallest values of $-2lnL$, AIC, BIC, CAIc. Consequently, the DZW distribution is the best model compared with other distributions used based on the given criteria.

Table 8
Goodness-of-fit measures for fitted models of real data Set I

Model	parameter	Estimate	SEs	K-S	P-value	-2*ln*L	AIC	BIC	AICc
DZW	α θ γ	3.4358 0.9448 0.8873	0.7819 0.8080 0.8086	0.125	0.8375	399.279	405.279	410.893	405.825
DW	q β	0.6443 0.3330	0.8111 0.8144	0.2708	0.0536	498.679	502.679	506.422	502.946
GDW	α θ p β	4.7121 2.9576 0.1225 0.1961	0.7687 0.7869 0.8166 0.8158	0.2708	0.0544	481.543	489.543	497.028	490.473
DW-G	p λ α	0.3555 0.8047 0.6428	0.8142 0.8094 0.8111	0.2708	0.0539	517.043	523.043	528.657	523.589
DMOW	α β γ	8.0239 0.5973 0.5230	0.7350 0.8116 0.8124	0.2083	0.2338	416.184	422.184	427.797	422.729
DMOGE	α θ γ	2.9844 0.9121 0.8507	0.7866 0.8083 0.8090	0.1875	0.3525	407.133	413.133	418.746	413.678

The *total time test* (TTT) plot can be used to get information about the shape of the hrf of a give data set, which helps in selecting a particular model to fit a given data set. Fitted pmf, P-P and Q-Q plots indicate that the proposed distribution fit for the datasets used.

Figure 10 shows TTT plot of this data set which indicates that this data has an increasing-shaped hazard rate. The fitted pmf, P-P and Q-Q plots indicate that the DZW distribution provides the best fit for this data.

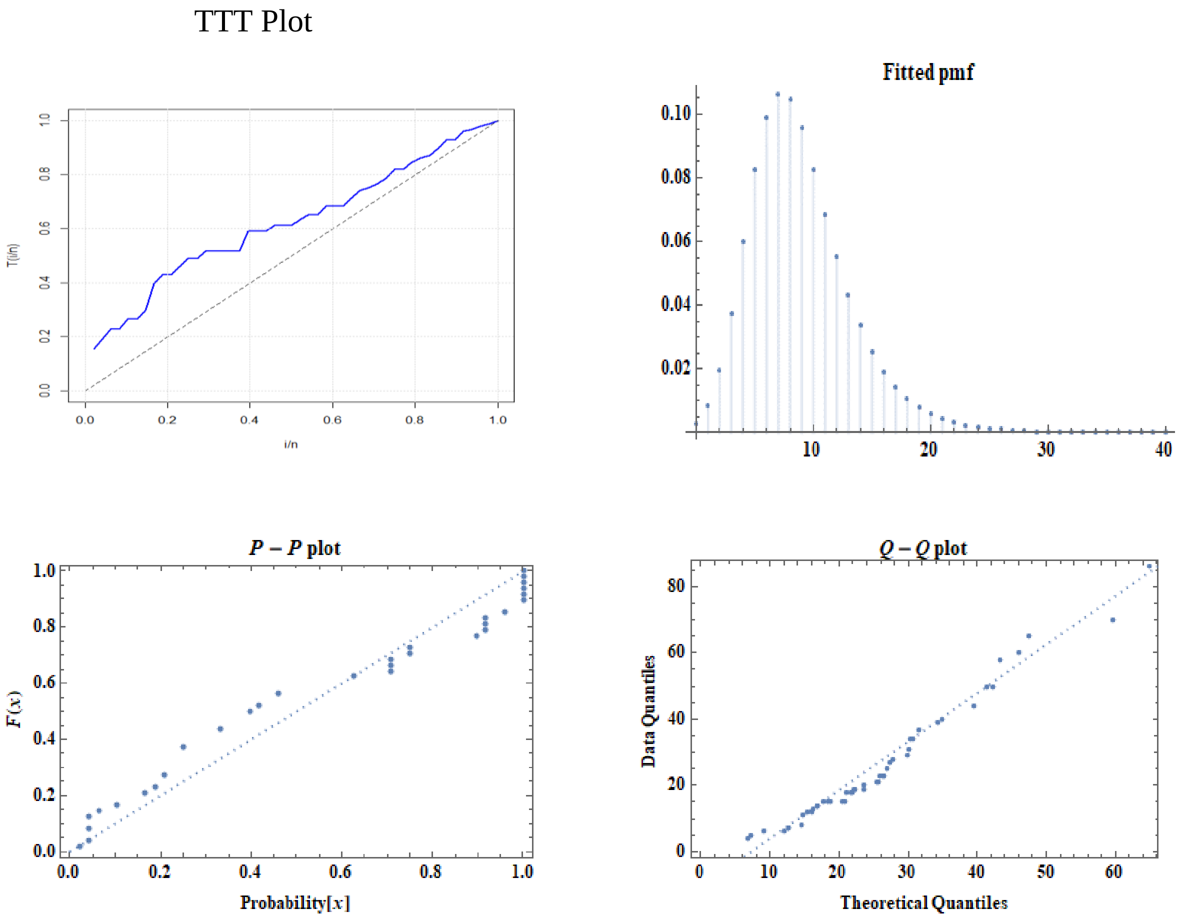


Figure 10.

TTT, fitted pmf, P-P and Q-Q plots of the DZW distribution for Data Set I

Data set II

The second data set refers to survival times of 44 patients suffering from head and neck cancer who retreated using a combination of radio therapy. This data is given by Afify *et al.* (2021). The data are: **12, 32, 37, 24, 24, 74, 81, 26, 41, 58, 63, 68, 78, 47, 55, 84, 155, 159, 92, 94, 110, 127, 130, 133, 140, 112, 119, 146, 173, 179, 194, 195,339, 432, 209, 249, 281, 319, 469, 725, 817, 519, 633 and 1776**.

Table 9 presents the ML estimates and corresponding SEs, *K-S* statistic with its corresponding p-value, $-2lnL$, AIC, BIC and CAIc. From Table 10, it is observed that all models fit the data set. However, the proposed distribution has smallest values of $-2lnL$, AIC, BIC, CAIc, lowest *K–S* value and highest p-value. Hence, proposed distribution is the best fit for this data compared with other distributions considered here.

Figure 11 displays TTT plot of data set II which indicates that this data has a unimodal-shaped hazard rate. The fitted pmf, P-P and Q-Q plots indicate that the DZW distribution provides the best fit for this data.

Table 9
Goodness-of-fit measures for fitted models of real data Set II

Model	parameter	Estimate	SEs	K-S	P-value	-2*ln*L	AIC	BIC	AICc
DZW	α θ γ	0.4665 0.9071 0.8965	1.1064 1.0982 1.0984	0.1136	0.9430	349.885	355.885	360.282	356.742
DW	q β	0.9900 1.7624	1.0967 1.0824	0.25	0.1282	591.126	595.126	598.694	595.419
GDW	α θ p β	0.6563 3.8529 0.9201 1.0662	1.1029 1.0441 1.0980 1.0953	0.2727	0.0755	660.309	668.309	675.446	669.335
DW-G	p λ α	0.3319 0.5499 0.0785	1.1089 1.1049 1.1136	0.2273	0.2072	600.657	606.657	612.009	607.257
DMOW	α β γ	1.2866 0.9404 1.0016	1.0912 1.0976 1.0965	0.2045	0.3188	575.518	581.518	586.87	582.118
DMOGE	α θ γ	1.9829 0.9095 0.4929	1.0784 1.0982 1.1059	0.1818	0.4655	557.31	563.31	568.31	563.91

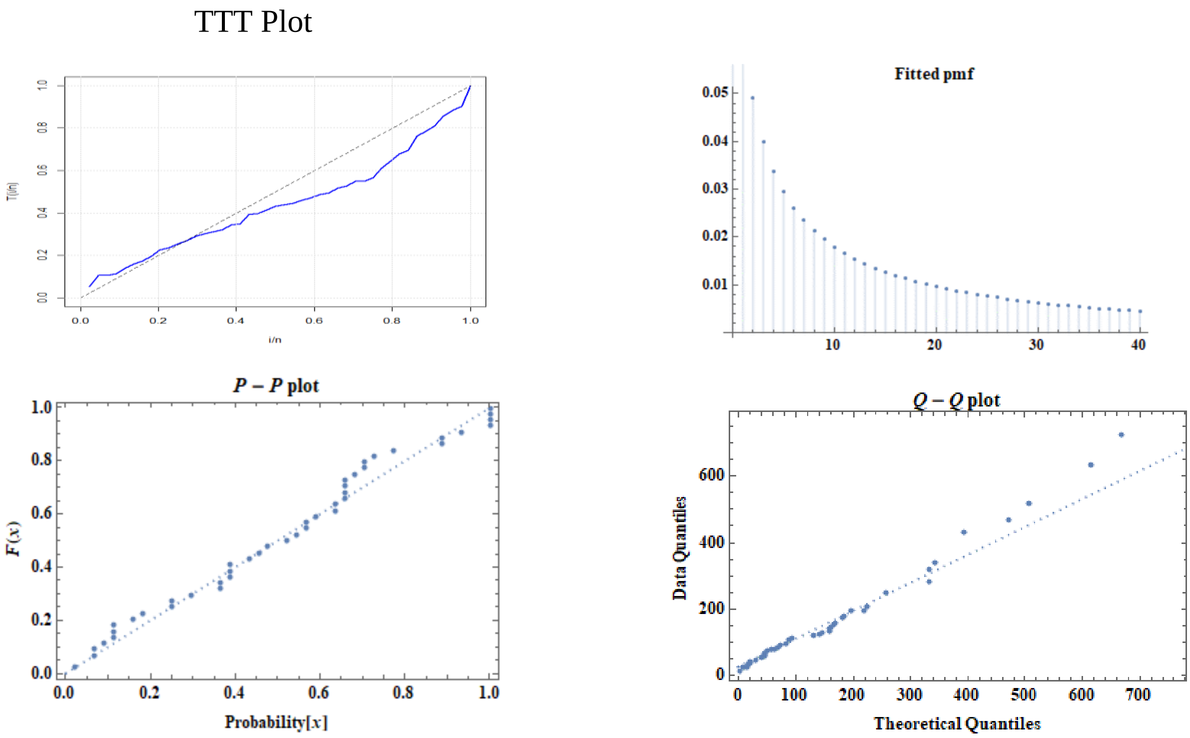


Figure 11.

TTT, fitted pmf, P-P and Q-Q plots of the DZW distribution for Data Set II

6.2 Count data

In the third data set the proposed distribution is compared with DW, GDW, DW-G, DMOW and DMOGE distributions. It represents 40 observations of time-to-failure (10^3h) of turbocharger of one type of engine. This data is provided by Xu *et al.* (2003).

X	0	1	2	3	4	5	6	7	8	9	**Total**
Freq	0	0	2	2	2	7	6	8	9	4	**40**

The third data set is displayed in Table 10. The ML estimates of the parameters of the considered distributions and corresponding SEs, χ^2 statistic, corresponding p-value, *degrees of freedom* (df), $-2lnL$, AIC, BIC and AICc criteria are computed.

Table 10
Parameter estimates and goodness of fit for various models fitted for the 40 observations of time-to-failure (10^3h) of turbocharger of one type of engine

Model	Parameter	Estimate	SEs	χ^2	Df	P-value	-2*ln*L	AIC	BIC	AICc
DZW	α θ γ	7.0820 1.5447 0.7996	0.2263 0.3606 0.3844	8.736	2	0.3650	233.365	239.365	244.718	239.965
DW	q β	0.1133 0.4792	0.4053 0.3943	12.265	3	0.1398	331.563	335.563	338.941	335.887
GDW	α θ p β	9.2455 5.5544 0.7003 1.5163	0.2965 0.2266 0.3875 0.3615	11.393	1	0.1804	249.337	257.337	264.092	258.48
DW-G	p λ α	1.5122 0.7944 0.0180	0.3616 0.3845 0.4082	15.330	2	0.0530	353.259	359.259	364.325	359.925
DMOW	α β γ	2.8416 0.8357 1.5419	0.3165 0.3832 0.3607	10.385	2	0.2390	241.609	247.609	252.676	248.276
DMOGE	α θ γ	3.8874 0.4949 3.0833	0.2791 0.3938 0.3079	10.264	2	0.2470	238.3	244.3	249.366	244.966

One can observe from Table 10, based on the p-value, that all the distributions fit for this data. The DZW distribution has the smallest χ^2 value and the highest p-value. Based on $-2lnL$, AIC, BIC and AICc criteria, the proposed distribution has the smallest values followed by the DMOGE distribution.

Figure 12 shows that TTT plot of this data set indicates that this data has an increasing-shaped hazard rate. The fitted pmf, P-P and Q-Q plots indicate that the DZW distribution fits the data very well.

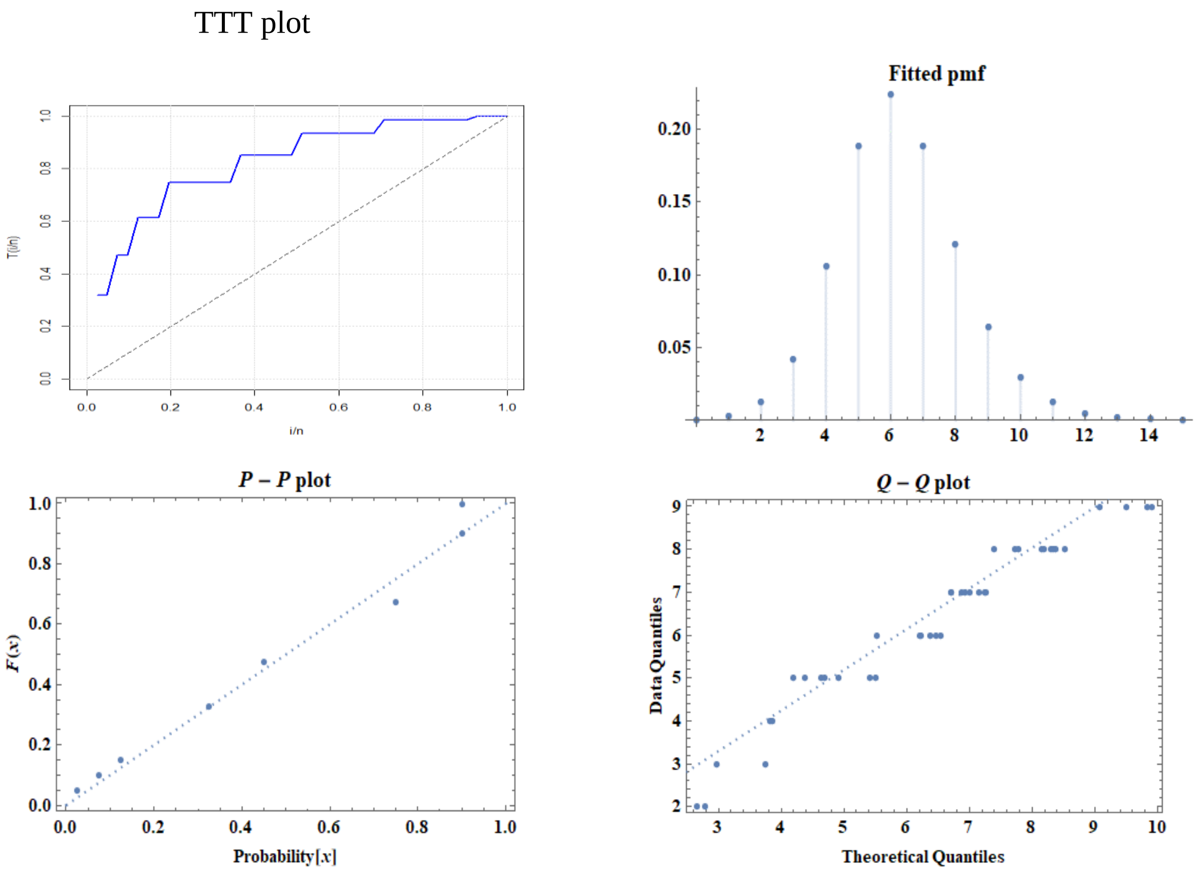


Figure 12.

TTT, Fitted pmf, P-P and Q-Q plots of the DZW distribution for Data Set III

7. Conclusion

The objective of this paper is to attract wider applications in engineering, medicine, biological, and other fields of research. Hence, in this paper, a discrete family of distributions is proposed, it is called DCEGP family. Some of its distributional properties including hazard rate, moments, quantiles, order statistics and Rényi Entropy are studied. Three special models of the family are discussed. The proposed family can be used for modeling count and lifetime data since the hrf has different shapes. The method of the ML is used to estimate the unknown parameters, sf, hrf and ahrf. Simulation study is carried out to illustrate the theoretical results and three real datasets with unimodal and increasing hazard rate are analyzed to demonstrate the suitability and flexibility of DZW distribution. TTT, Fitted pmf, P-P and Q-Q plots indicate more support for DZW distribution. These real datasets are compared with DW, GDW, DW-G, DMOW and DMOGE distributions. Finally, through these comparisons, one can conclude that the DZW distribution is the best distribution for fitting these data sets.

References

Afify, A. Z., El-Morshedy, M., Eliwa, M. S. (2021). A New skewed discrete model: properties, inference, and applications. *Pakistan Journal of statistics and operation research*, 17(4): 799-816. https://pjsor.com/index.php/index.

Ahmed, Z. (2018). The Zubair-G family of distributions. *Annals of Data Science*, 7(2): 195–208. https://doi.org/10.1007/s40745-018-0169-9.

Alamatsaz, M. H., Dey, S., Dey, T. and Shams Harandi, S. (2016). Discrete generalized Rayleigh distribution. *Pakistan Journal of Statistics*, 32(1): 1-20.

AL-Huniti, A. A. and AL-Dayian, G. R. (2012). Discrete Burr Type III distribution. *American Journal of Mathematics and Statistics*, 2(5): 145-152.

Almetwally, E. M. (2020). Discrete alpha power inverse Lomax distribution with application of COVID-19 Data. *International Journal of Applied Mathematics and Statistical Sciences (IJAMSS)*, 9(6): 2319-3972.

Almetwally, E. M., Almongy, H. M. and Saleh, H. A. (2020). Managing risk of spreading "COVID-19" in Egypt: Modelling using a discrete Marshall–Olkin generalized Exponential distribution. *International Journal of Probability and Statistics*, 9(2): 33-41.

Arnold, B., Balakrishnan, N. and Najaraja, H. N. (2008). *A first course in order statistics*. John-Wiley and Sons, New York.

Chakraborty, S. (2015). Generating discrete analogues of continuous probability distributions – A survey of methods and constructions. *Journal of Statistical Distributions and Applications*, 2(6):1-30.

Chakraborty, S. and Chakravorty, D. (2016). A new discrete probability distribution with integer support on $(-\infty, \infty)$. *Communication in Statistics – Theory and Methods*, 45(2):492-505. http://www.tandfonline.com/loi/lsta20.

Eliwa, M. S., Alhussain, Z. A. and El-Morshedy, M. (2020). Discrete Gompertz-G family of distributions for over- and under-dispersed data with properties, estimation, and applications. *Mathematics.* 8(3): 358. http://creativecommons.org/licenses/by/4.0/.

Elmorshedy, M., Eliwa, S. M. and Nagy, H. (2019). A new two-parameter exponentiated discrete Lindley distribution, estimation and application. *Journal of Applied Statistics*, 47(2): 1-22. https://www.researchgate.net/publication/334326226.

Gomez-Deniz, E. and Calderin-Ojeda, E. (2011). The discrete Lindley distribution: properties and application. *Journal of Statistics Computation and Simulation*, 81(11): 1405-1416.

Gupta, R. D. and Kundu., D. (2009). A new class of weighted exponential distributions. *Statistics*, 43(6): 621-634.

Hegazy, M. A., EL-Helbawy, A. A., Abd EL-Kader, R. E. and AL-Dayian, G.R. (2018). Discrete Gompertz distribution: properties and estimation. *Journal of Biostatistics and Biometric Applications*, 3(3): 307.

Hegazy, M. A., EL-Helbawy, A. A., Abd EL-Kader, R. E. and AL-Dayian, G.R. (2019). Discrete Inverted Kumaraswamy distribution: properties and estimation. https://www.researchgate.net/publication/333173090.

Helmy, N. (2018). Discrete Burr Type II distribution. *Al-Azhar Scientific Journal of the Commercial Faculties*, 20(2): 1-21.

Hinkley, D. (1977). On quick choice of power transformation. *Journal of the Royal Statistical Society,* Series C, 26, 67-69. https://doi.org/10.2307/2346869

Hussain, T., Aslam, M. and Ahmad, M. (2016). A Two parameter discrete Lindley distribution. *Revista Colombiana de Estadistica*, 39(1): 45-61.

Inusah, S. and Kozubowski, J. T. (2006). A discrete analogue of the Laplace distribution. *Journal of Statistics Planning Inference*, 136(3): 1090-1102.

Jayakumar, K. and Sankaran, K. K. (2017). A generalization of discrete Weibull distribution. *Communications in Statistics-Theory and Methods*, 47(24): 6064-6078. https://www.researchgate.net/publication/321953467.

Jazi, M. A., Lai, C. D. and Alamatsaz, M. H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. *Elsevier Statistical Methodology,* 7(2): 121-132.

Kemp, A. (2004). Classes of discrete lifetime distributions, *Communications in Statistics Theory and Methods*, 33(12): 3069– 3093.

Kemp, A.W. (1997). Characterization of a discrete normal distribution**.** *Journal of Statistical Planning and Inference*, 63(2): 223-229.

Khan, M. S. A., Khalique, A. and Abouammoh, A. M. (1989). On estimating parameters in a discrete Weibull distribution. *IEEE Transactions on Reliability,* 383: 348-350.

Krishna, H. and Pundir, P. S. (2009). Discrete Burr and discrete Pareto distributions. *Statistical Methodology,* 6(2): 177-188.

Lai, C. D. (2013). Issues concerning constructions of discrete lifetime models. *Quality Technology and Quantitative Management*, 10(2): 251-262. https://www.researchgate.net/publication/237843383.

Lai, C. D. (2014). *Generalized Weibull distribution*. Springer Heidelberg, New York, Dordrecht, London.

Lai, C. D., Murthy, D. N. P. and Xie, M. (2006*).* *Weibull distributions and their applications. In book: Springer handbook of engineering statistics.* Springer, New York. https://www.researchgate.net/publication/37628953.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. New York: Wiley.

Lee, E.T. and Wang, J. W. (2003). *Statistical methods for survival data analysis,* 3rd edition, Wiley, New York.

Mabel, Y. O., Novita, M. and Nurrohmah, S. (2021). Discrete Weibull-geometric distribution. *Journal of Physics Conference Series,* 1725.

Maiti, S. S., Dey, M. and Seema, S. (2018). discrete X-Gamma distribution: properties, estimation and an application to the collective risk model. *Journal of Reliability and Statistical Studies,* 11(1):117-132.

Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *Reliability, IEEE Transactions on*, 42(2): 299-302.

Murthy, D. N. P., Xie, M. and Jiang, R. (2004). *Weibull models*. John Wiley and Sons, New York. http://dx.doi.org/10.1002/047147326X.

Nakagawa, T. and Osaki, S. (1975). The discrete Weibull distribution. *IEEE Transactions on Reliability*, 24(5): 300-301.

Nekoukhou, V., Alamatsaz, M. H. and Bidram, H. (2012). A discrete analog of generalized Exponential distribution. *Communication in Statistics- Theory and Methods,* 41(11): 2000-2013.

Nurudeen, A. (2017). A new discrete family of reduced modified Weibull distribution. *International Journal of Statistical Distributions and Applications***,** 3(3): 25-31.

Opone, F. and Akata, I. (2021). A discrete analogue of the continuous Marshall-Olkin Weibull distribution with application to count data. *Earthline Journal of Mathematical Sciences,* 5(2): 415-428. https://www.researchgate.net/publication/346084901.

Para, B. A. and Jan, T. R. (2018). On three parameter discrete generalized inverse Weibull distribution: properties and applications. *Annals of Data Science,* 6(3): 549–570.

Para, B. A. and Jan, T.R. (2014). Discrete generalized Burr-Type XII distribution. *Journal of Modern Applied Statistical Method,* 13(2): 244-258.

Rényi, A. (1961). On measures of entropy and information. *Mathematical Statistics and Probability,* 1: 547–561.

Rohatgi, V. K. and Saleh, E. A. K. (2001). *An introduction to probability and statistics*. 2nd Edition, John- Wiley and Sons, New York.

Roy, D. (2003). Discrete normal distribution. *Communication in Statistics- Theory and Methods*, 32(10): 1871-1883.

Roy, D. and Gupta, R. P. (1992). Classifications of discrete lives. *Microelectronics and Reliability,* 34(3): 253-255.

Tahir, M. H. and Cordeiro, G. M. (2016). Compounding of distributions: a survey and new generalized classes. *Journal of Statistical Distributions and Applications*, 3(13): 1-35.

Tahir, M. H. and Cordeiro, G. M. (2019). Letter to the editor concerning "The Zubair-G family of distributions: properties and applications" *Annals of Data Science,* https://www.researchgate.net/publication/332404752.

Xie, M., Gaudoin, O. and Bracquemond, C. (2002). Redefining failure rate function for discrete distribution. *International Journal of Reliability, Quality and Safety Engineering,* 9(3): 275-286.

Xu, K., Xie, M., Tang, L. C. and Ho, S. L. (2003) Application of neural networks in forecasting engine systems reliability, *Applied Soft Computing*, 2(4): 255-268.

8. Appendix

a. $h(x)$ is not additive for a competing risk model.

Proof

The hrf in discrete case is defined as

$$h(x) = \frac{S(x)-S(x+1)}{S(x)}. \quad \text{(AI)}$$

are not additive for series system. That is, if we have *n* discrete components in series,

$$h_n(x) = \frac{\prod_{i=1}^{n} S_i(x) - \prod_{i=1}^{n} S_i(x+1)}{\prod_{i=1}^{n} S_i(x)},$$

$$= 1 - \prod_{i=1}^{n} \left\{\frac{S_i(x+1)}{S_i(x)}\right\},$$

$$= 1 - \prod_{i=1}^{n}\{1 - h_i(x)\} \neq \sum_{i=1}^{n} h_i(x),$$

and hence,

$$ah_n(x) = \sum_{i=1}^{n} ah_i(x). \quad \text{(AII)}$$

c. The cumulative hrf, $H(x) = \sum h(x) \neq -ln[S(x)]$.

proof

In continuous case the cumulative hrf is:

$$\Lambda(x) = -\ln\,[S(x)], \quad \text{(AIII)}$$

In discrete case, if the cumulative hrf gives

$$H(x) = \sum_{i=0}^{n} h_i(x), \quad \text{(AV)}$$

from (AIII)

$$\Lambda(x) = -\ln\,[S(x)] = -\ln\,[\prod_{i=1}^{n}\{1 - h_i(x)\}],$$

$$= -\sum_{i=0}^{n}\{1 - h_i(x)\}. \quad \text{(AVI)}$$

from (AV)

$$H(x) = \sum_{i=1}^{n} h_i(x) = \sum_{i=0}^{n} \frac{S_i(x)-S_i(x+1)}{S_i(x)},$$

$$= \sum_{i=0}^{n} \left\{1 - \frac{S_i(x+1)}{S_i(x)}\right\}. \quad \text{(AVI)}$$

In general, $H(x) \neq \Lambda(x)$ in discrete case (if (AI) is used as the failure rate function).

[see Xie *et al.* (2002) and Kemp (2004)]