A New Discrete Raleigh Distribution and its Application in Immunogold Assay Data

Krishnakumari.K[1*], Dais George[2].

[1]Research Scholar, St.Thomas College, Palai (SAS SNDP YOGAM College, Konni), Kerala, India; E-mail: krishnavidyadharan@gmail.com

[2]Catholicate College, Pathanamthitta, Kerala, India.

Abstracts: Recently, discretization of continuous distribution was enchanted by many researchers especially in the field of reliability engineering and life testing. This is because the existing models are unsuited for many practical situations. For discretizing continuous distributions, various methods are available in literature. Among them a newly developed method is the two-stage composite discretization method which includes three different methodologies. Using these methodologies, we propose some new discretized distributions obtained through the discretization of Rayleigh distribution. We compare the hazard rate functions of Rayleigh, Discrete Rayleigh with the newly proposed discrete distributions. The flexibility of the model is illustrated using immunogold assay data.

Keywords: Discretization, Estimation, Hazard rate function, Rayleigh Distribution.

1. INTRODUCTION

Recently discrete distributions played a key role in modeling survival data. In survival analysis, usually the variables are measured are in continuous scale. But there exist many practical situations which demands discrete measurements. For example, the time spent by a patient in a hospital is counted as number of days or weeks. Here instead of measuring the lifetime in a continuous scale we treat it in discrete manner. Also, the discrete measurements will make the models more appropriate. Usually, in reliability analysis continuous distributions play a vital role in modeling. But sometimes it produces a lot of derivational difficulties, and the approximation procedures results in discretization. Since too many discrete distributions are available in literature, they are not able to model many practical situations and hence many continuous distributions must be discretized. Now a days there are various methods available for discretizing continuous distributions. Alzaatrech et al. (2012) and Chakraborty (2015) carried out a survey of these methods and using them many continuous distributions were discretized. For details see, Lisman and van Zuylen (1972), Stein and Dattero (1984), Kemp (1997), Das Gupta (1993), Gupta and Kundu (1999), Szablowski (2001), Roy (2003,2004), Inusah and Kozubowski (2006), Kozubowski and Inusah (2006), Krishna and Pundir (2009), Jazi et al. (2010), Nekoukhou et al. (2012), Lekshmi and Sebastian (2014), Nekoukhou Hamed Bidram (2015), Chakraborty (2015), Chakraborty and Chakravarthy (2016), Josmar et al.(2017), Berhane Abebe and Rama Shanker (2018), Ganji and Gharari (2018), Jayakmar and Sankaran (2018), Yari and Tondpour (2018), Abebe and Shukla(2019), Krishnakumari and Dais (2020), Krishnakumari and Dais(2021) and Morshedy et.al (2020). So, we concentrate our study in the development of some new discretized distributions from their continuous counterpart.

The remaining part of the paper is organized as follows. In section 2 the discretization method developed by Yari and Tondpour (2018) is discussed. Some new discrete lifetime distributions are proposed and studied in section 3. In section 4, using hazard rate functions, we compare the newly proposed distributions with the existing distributions. Parameter estimation is done in section 5. An applications of the newly proposed New discrete Rayleigh (Type1) distribution is discussed in section 6. We concluded the article by section 7.

2. DISCRETIZATION METHODS

In recent years discretization of continuous distribution has got great recognition in many real-world scenarios. A lot of discrete distributions are so far proposed and studied by many researchers. For details, see the books by Balakrishnan and Nevzorov (2003), Jonson et al. (2005) and Consul and Famoye (2006). In this section we are concentrating on the discretization of Rayleigh distributions using the methods proposed by Yari and Tondpour (2018). He proposed three methodologies and each of which consists of two stages. In the first stage a new continuous random variable is constructed from the underlying continuous random variable and in the second stage, this new random variable is discretized by maintaining the same hazard rate function. The following are the three methodologies used.

2.1 Methodology I

Here, in the first stage, a continuous random variable X with cumulative distribution function F(x) and support [0, ∞) is used to construct a new continuous random variable X_1 having hazard rate function

$$h_{x1}(x) = e^{-F(x)},\ \ x \geq 0 \tag{1}$$

In the second stage, a discrete analogue Y of X_1 is derived by using the following methodology where hazard rate function of Y retains the form of hazard rate function of X_1. If the continuous random variable X_1 has survival function S_{x1} (x) and hazard rate function h_{x1} (x) then the survival function of the discrete analogue Y is given by

$$s_y(k) = \left(1 - h_{x1}(1)\right)\left(1 - h_{x1}(2)\right) \dots \left(1 - h_{x1}(k-1)\right); k = 1,2, \dots, m$$

and the corresponding probability mass function is

$$P(Y=k) = \begin{cases} h_{x1}(0); k = 0 \\ ((1 - h_{x1}(1))(1 - h_{x1}(2) \dots \left(\left(1 - h_{x1}(k-1)\right)\right) h_{x1}(2k); \\ \qquad\qquad\qquad\qquad k = 1,2, \dots\dots, m \\ 0; otherwise \end{cases} \tag{2}$$

If P(Y=k) is such that the total probability is not equal to one, then we shall multiply every P(y) by the positive constant w that will ensure the total probability equals to one. Hence the probability mass function takes the form

$$P(Y=y) = \begin{cases} w; \qquad y = 0 \\ w h_{x1}(y) \prod_{i=1}^{y-1} (1 - h_{x1}(i)); \qquad y = 1,2, \dots m \\ 0; otherwise \end{cases} \tag{3}$$

where m can be infinite or infinite since $h_{x1}(x)$is always between zero and one.

Now, by using (3) the resulting pmf of Y in new methodology is

$$P(Y=y) = \begin{cases} w; \quad y = 0 \\ we^{-Fx(x)} \prod_{i=1}^{y-1} (1 - e^{-Fx(i)}); \quad y = 1,2,\ldots m \\ 0; otherwise \end{cases} \tag{4}$$

2.2 Methodology II

In this method, in the first stage a new continuous random variable X_1 having hazard rate function

$$h_{x1}(x) = \frac{2F_X(x)}{1 + F_X(x)}$$

is constructed using continuous random variable X with cumulative distribution function $F_X(x)$ and support [0, ∞). Then, in the second stage, a discrete analogue Y of X_1 is derived by using (3). Also note that discrete distributions obtained in this methodology has increasing hazard rate function.

2.3 Methodology III

Here, in the first stage a continuous random variable X with cumulative distribution function $F_X(x)$ and Support [0, ∞). is used to construct a new continuous random variable X_1 having hazard rate function

$$h_{x1}(x) = \frac{1}{f_x(x) + 1} x \geq 0$$

Then in the second stage, a discrete analogue Y of X_1 is derived by using (3). Here the hazard rate function of Y is increasing (decreasing) on (a, b) where a, b € R^+ if and only if $F_X(x)$ is decreasing (increasing) on the same interval.

In the first two methods, hazard rate functions of Y are decreasing and increasing respectively and in the third method it can be increasing, U shaped or modified unimodal. An important advantage of the method is that discrete analogues obtained have monotonic and non-monotonic hazard rate functions.

3. SOME NEW DISCRETE LIFETIME DISTRIBUTIONS

Rayleigh distribution is the commonly used probability distribution for modeling data relating to reliability and life testing. It has emerged as a special case of Weibull distribution. It is widely used in communication theory, physical sciences, and reliability engineering. A discrete analogue of continuous distribution is derived by maintaining one or more discriminative property of the continuous distribution. A discrete Rayleigh distribution is used in the field of reliability approximation. For details see Roy (2004). In this section, we introduce some new discrete distributions derived from the continuous Rayleigh distributions namely, New discrete Rayleigh (Type I, II and III) using the methodologies explained in section 2.

3.1 New Discrete Rayleigh Distribution (Type I)

The probability density function and distribution function of a Rayleigh distribution with parameter σ > 0 is given respectively by

$$f(x) = \frac{x}{\sigma^2} e^{-\left(\frac{x^2}{2\sigma^2}\right)}; \quad x \geq 0, \ \sigma > 0 \tag{5}$$

and

$$F(x) = 1 - e^{-\left(\frac{x^2}{2\sigma^2}\right)}; x \geq 0, \ \sigma > 0$$

Using methodology 1 in section 2, the pmf of New Discrete Rayleigh distribution of Type I is obtained as

$$P_Y(y) = e^{e^{-\left(\frac{y^2}{2\sigma^2}\right)}-1} \prod_{i=1}^{y-1} \left(1 - e^{e^{-\left(\frac{i^2}{2\sigma^2}\right)}-1}\right); \quad y = 1,2,3, \dots, m \tag{6}$$

The plot of New Discrete Rayleigh Type I distribution is given in Figure. 1.

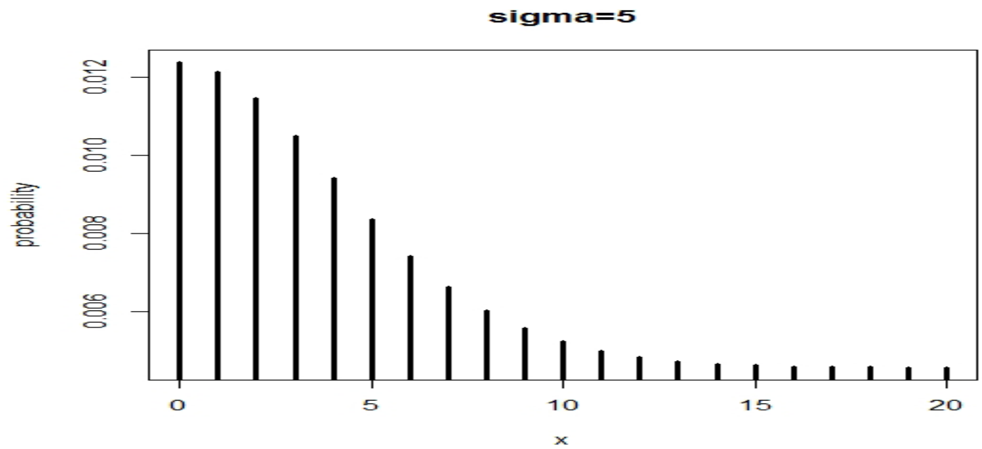

Figure. 1.New Discrete Rayleigh Type I Distribution

Its hazard rate function is given by

$$h_y(y) = e^{e^{-\left(\frac{y^2}{2\sigma^2}\right)}-1} \tag{7}$$

and is decreasing.

3.2 New Discrete Rayleigh Distribution (Type II)

The pmf of New Discrete Rayleigh distribution of Type II is obtained by using methodology 2 in section 2 as

$$P_Y(y) = w \frac{2\left(1-e^{-\left(\frac{y^2}{2\sigma^2}\right)}\right)}{2-e^{-\left(\frac{y^2}{2\sigma^2}\right)}} \prod_{i=1}^{y-1} \left(1 - \frac{2\left(1-e^{-\left(\frac{i^2}{2\sigma^2}\right)}\right)}{2-e^{-\left(\frac{i^2}{2\sigma^2}\right)}}\right); \quad y = 1,2,3, \dots \dots, m \tag{8}$$

Figure. 2 gives the plot of New Discrete Rayleigh Type II distribution.

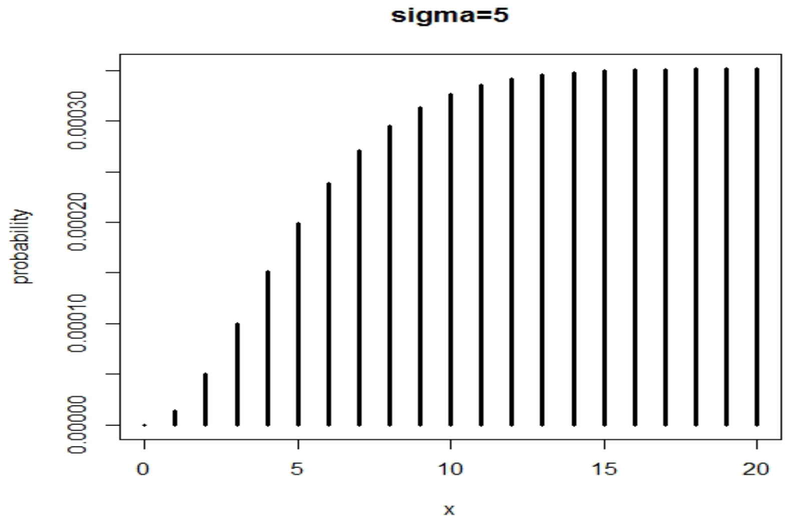

Figure. 2. New Discrete Rayleigh Type II Distribution

Its hazard rate function is increasing, and it is

$$h_y(y) = \frac{2\left(1-e^{-\left(\frac{y^2}{2\sigma^2}\right)}\right)}{2-e^{-\left(\frac{y^2}{2\sigma^2}\right)}} \tag{9}$$

3.3 New Discrete Rayleigh Distribution (Type III)

By applying methodology 3 in section 2, the pmf and hazard rate function of a discrete Rayleigh distribution of Type III is obtained respectively as

$$P_Y(y) = \frac{1}{1+\frac{y}{\sigma^2}e^{-\left(\frac{y^2}{2\sigma^2}\right)}} \prod_{i=1}^{y-1}\left(\frac{1}{1+\frac{y}{\sigma^2}e^{-\left(\frac{y^2}{2\sigma^2}\right)}}\right); \quad y = 1,2,3,\ldots\ldots,m \tag{10}$$

and

$$h_y(y) = \frac{1}{1+\frac{y}{\sigma^2}e^{-\left(\frac{y^2}{2\sigma^2}\right)}} \tag{11}$$

The hazard rate function is U shaped if $\sigma > 2$. The plot of New Discrete Rayleigh

Type III distribution is given in Figure. 3.

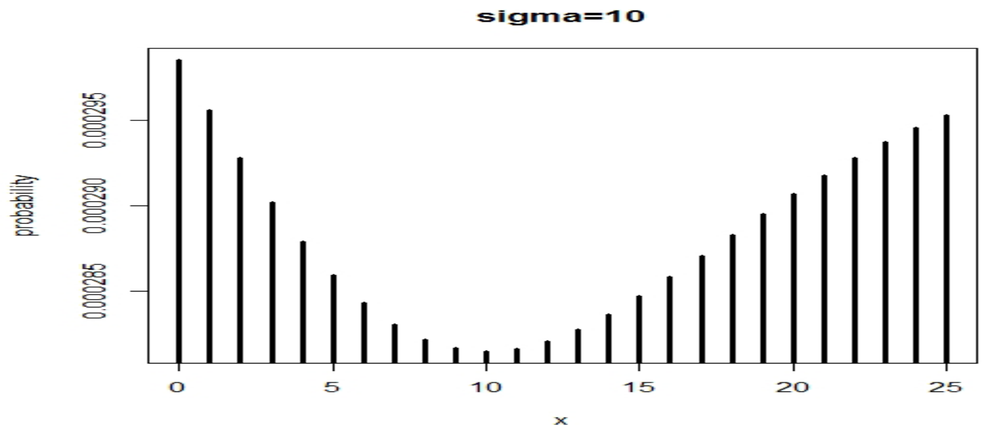

Figure. 3. New Discrete Rayleigh Type III Distribution

4. COMPARISON

Lifetime data are usually measured in a continuous manner. But there exists a lot of situations where it can be treated in a discrete way. Hazard rate is an unavoidable characteristic in reliability and life data analysis. Different shapes of hazard rate functions can give different distributions. Based on the conditions imposed on the parameters, the hazard rate functions may have increased, decreasing and bathtub shapes. Discrete analogues derived from the continuous distributions have the same shape as that of the continuous one for the same parametric values. In this section we discuss about the acquiescence of this fact.

Here, we consider the comparison of Rayleigh distribution, discrete Rayleigh distribution introduced by Roy (2004) and the newly proposed New Discrete Rayleigh distributions (Type I, Type II and Type III) using hazard rate function.

Hazard rate function of Rayleigh distribution is given by the equation

$$h_x(x)=x/\sigma^2$$

Its graph is given in figure 4.

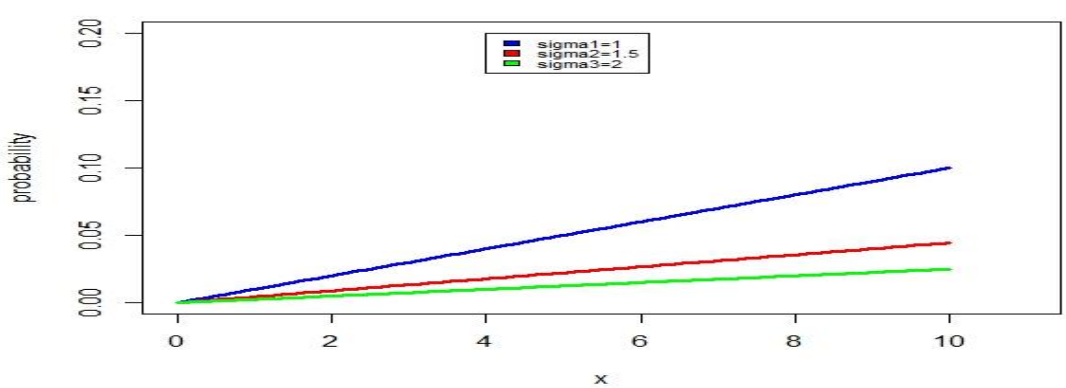

Figure 4. Hazard plot- Rayleigh distribution

The hazard rate function of discrete Rayleigh distribution of Roy (2004) is given by

$$h_x(x) = 1 - \theta^{(x+1)^2 - x^2}$$

Its hazard plot is given in Figure 5.

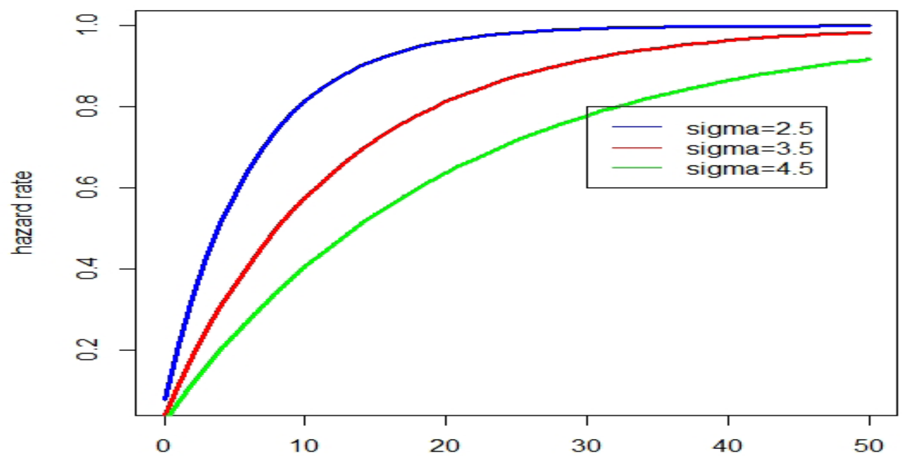

Figure. 5. Hazard plot-Discrete Rayleigh Distribution

The hazard rate function of New Discrete Rayleigh distribution (Type I, Type II and Type III) is obtained, and it is given in section 3 by (7), (9) and (11).

The corresponding plots are given in Figure 6.

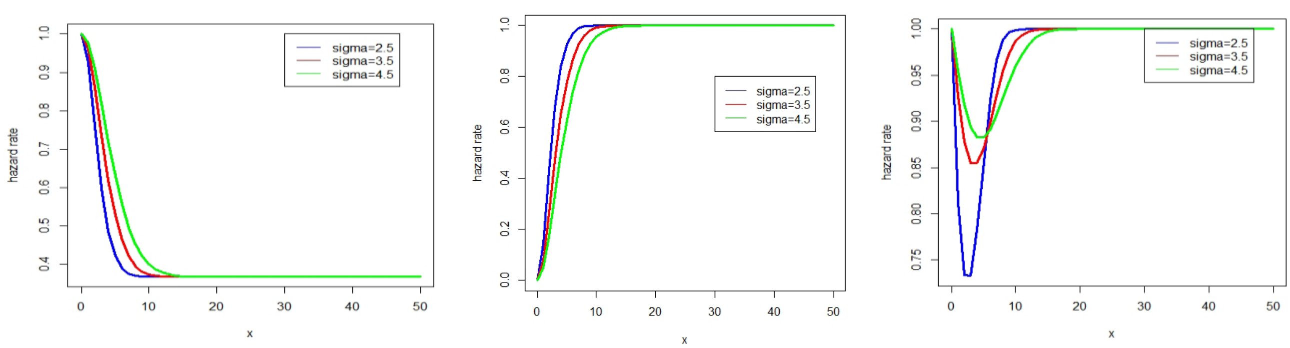

Figure. 6. Plots of hazard rate (i) New Discrete Rayleigh Type I (ii) New Discrete

Rayleigh Type II (iii) New Discrete Rayleigh Type III

Figure 4 and 5 shows that hazard rate function of Rayleigh and discrete Rayleigh distributions are always increasing. But in the case of the newly proposed distributions for all values of σ, it is decreasing for New Discrete Rayleigh Type I and increasing for New Discrete Rayleigh Type II distribution. The hazard rate is a U-shaped curve for New Discrete Rayleigh Type III distribution for $\sigma > 2$.

5. ESTIMATION

We estimate the parameters of New Discrete Rayleigh distribution Type I distribution using the method of proportion.

6. REAL DATA ANALYSIS

As an application of New Discrete Rayleigh distribution Type I distribution, we consider the data se used by Cullen, Walsh, Nicholson, and Harris (1990). Here we give the counts of sites with 1, 2, 3, 4 and 5 particles from immunogold assay data and 122, 50, 18, 4 and 4 were the counts. Figure 7 shows the observed data.

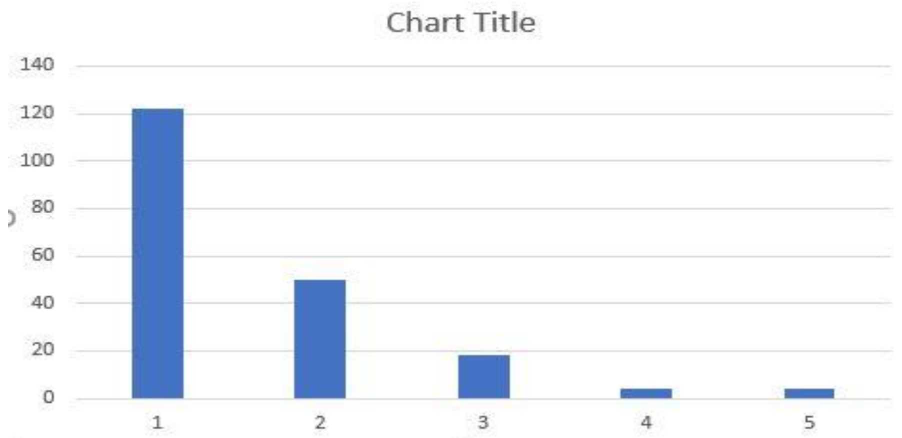

Figure 7. Observed data

Since it has the same shape as that of New Discrete Rayleigh Type I distribution, the empirical data is fitted for New Discrete Rayleigh Type I distribution. The estimated value of σ is 2.2. The embedded graph is shown in figure 8.

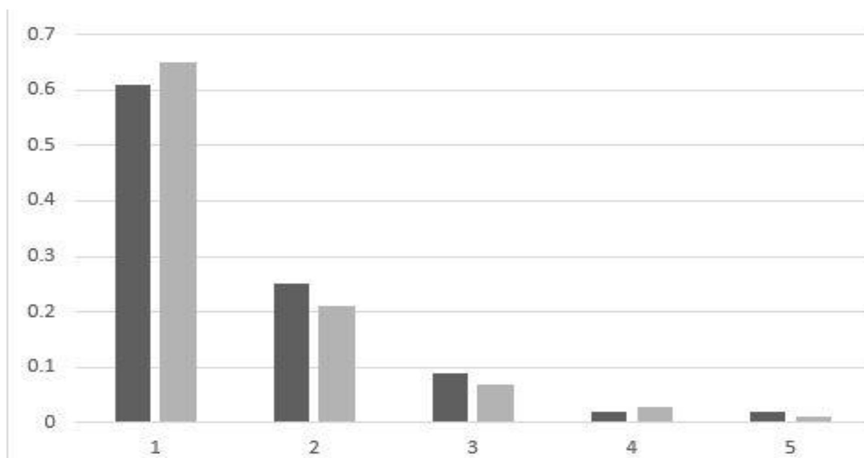

Figure 8. Embedded graph

The goodness of fit of the proposed model is checked using chi-square statistic and it is found that the New Discrete Rayleigh Type I distribution is a good fit for immunogold assay data. Also, for comparing its suitability with Discrete Burr XII (Type I) distribution (Yari and Tondpour (2018)), the loglikelihood, AIC and BIC values are tabulated and are shown in Table 1.

Table 1: Values of MLE's, Chi-square, LL, AIC, and BIC values

Distribution fitted	Estimated values	Chi-square	LL	AIC	BIC
Discrete Burr XII (Type I)	p=0.9554 c=0.3596	1.7031	-206.7	417.48	418.07
NewDiscrete Rayleigh (Type I)	σ =2.2	1.439	-174.281	350.562	350.8586

Table 1 shows that New Discrete Rayleigh (Type I) distribution is a better model than Discrete Burr XII (Type I) distribution.

Summary

We usually encounter many practical situations where lifetimes are appropriate for discrete measurement rather than continuous one. In the past decades discrete distributions like Geometric and Negative binomial are used to model discrete lifetime data. But these distributions are not able to model many practical situations. Now a days, there are various discretization methods available in literature. This leads to the development of discretizing continuous distributions. In this work we, derive some new lifetime distributions viz. Discrete Rayleigh (Type I, II and III). Since in reliability and lifetime analysis, hazard function is a chief characteristic, a comparative study based on hazard rate functions of these distributions with the existing distributions is also done. The study reveals that for all values of σ, the hazard rate function is increasing for Rayleigh and discrete Rayleigh distributions. But for New Discrete Rayleigh Type I distribution, it is decreasing and for New Discrete Rayleigh Type II distribution it is increasing. The hazard rate is a U-shaped curve for New Discrete Rayleigh Type III distribution for $\sigma > 2$. A real data analysis is carried out using immunogold assay data. The flexibility of the newly proposed model is compared with Discrete Burr Type I distribution and finds that the proposed model is a better model than the existing one.

REFERENCES

[1] Abebe, B and Shanker, R; A discrete Lindley distribution with applications in biological sciences, Biom. Biostat. Int. Journal, 7(1) 1-5 (2018).

[2] Abebe, B and Shukla, K.K.; A discrete Pranav distribution and its applications, Biometrics and Biostatistics International Journal, 8(1) 33-37(2019).

[3] Alzaatreh, A., Carl Lee; Felix Famoye; On the discrete analogue of continuous distributions, Statistical Methodology, 9(6)589-603 (2012).

[4] Balakrishnan, N and Nevzorov V.B.; A primer on statistical distributions, John Wiley and Sons, Inc, Hoboken, NJ. (2003).

[5] Chakraborty, S., Generating discrete analogues of continuous probability distributions-A survey of methods and constructions, Journal of Statistical Distributions and Applications, 2(6) (2015).

[6] Chakraborty, S. and Chakravarthy, D.; A new discrete probability distribution with integer support on (-∞,+∞), Communications in Statistics Theory and Methods, 45(2) 492-505 (2016)

[7] Consul, P.C. and Famoye, F.; Lagrangian Probability Distributions, Birkhauser, Boston, Massachusetts (2006).

[8] Cullen, M.J., Walsh, J., Nicholson, I.V. and Harris, J.B.; Ultrastructural localization of dystrophin in human muscle by using gold immuno labelling, Proc.Royal Society London, Series B 20, 197-210, (1990)

[9] Dasgupta, R.; Cauchy equation on discrete domain and some characterization theorems, Theory of Probability and its Applications, 38(3) 520-524 (1993).

[10] Ganji, M and Gharari, F.; A new method for generating discrete analogues of continuous distributions, Journal of Statistical Theory and Applications, 17(1) 39-58 (2018)

[11] Gupta, R. D. and Kundu, D.; Generalized Exponential distributions, Austral.NZ.J.Statist., 41(2) 173-188 (1999).

[12] Inusah, S and Kozubowski, T. J.; A Discrete analogue of the Laplace distribution, Journal of Statistical Planning and Inference, 136(3) 1090-1102 (2006).

[13] Jayakumar, K and Sankaran, K. K.; A generalization of discrete Weibull distribution, Communications in Statistics-Theory and Methods, 47(24) 6064-6078 (2018).

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

[15] Johnson N. L., Kemp A.W. and Kotz, S.; Univariate Discrete Distributions, third edition, John Wiley and Sons, Inc, Hoboken, NJ (2005)

[16] Josmar, M., Wesley, BDS and Ricardo, P.O.; On the discrete Shanker distribution, Chilean Journal of Statistics, 136(3) 6-14 (2017).

[17] Kemp, A. W.; Characterizations of discrete Normal distribution, Journal of Statistical Planning and Inference, 63 223-229 (1997).

[18] Kozubowski, T. J. and Insuah, S.; A skew Laplace distribution on Integers, AISM, 58 555-571 (2006).

[19] Krishna, H. and Pundir,S.P.; Discrete Burr and discrete Pareto distributions, Statistical Methodology, 6 177-188 (2009).

[20] Krishnakumari, K and George, D.; Discrete Esscher transformed Laplace distribution and its application, Stochastic Modeling and Applications, 24(2) 149-159 (2020).

[21] Krishnakumari, K and George, D.; A New Discrete Weibull distribution and its application in immunogold assay data, Stochastic Modeling and Applications, 25(2) 23-34 (2021).

[22] Lekshmi, V.S and Sebastian, S.; A skewed generalized discrete Laplace distribution, International Journal of Mathematics and Statistics Invention, 2(3) 95-102 (2014).

[23] Lisman, J. H. C. and van Zuylen M. C. A.; Note on the generation of most probable frequency distributions, Statistica, Neerlan, 26(1) 19-23 (1972).

[24] M. Morshedy, M., Eliwa, M. S. and Altun E.; Discrete Burr-Hatke distribution with properties, Estimation methods and regression model, IEEE Access, 74359-74370 (2020).

[25] Nekoukhou, V., Alamatsaz, M. H. and Bidram, H.; A discrete analogue of the generalized Exponential distribution, Communications in Statistics-Theory and Methods, 41(11) 2000-2013 (2012).

[26] Nekoukhou, V and Bidram, H.; The exponentiated discrete Weibull distribution, SORT, 39(1) 127-146 (2015).

[27] Roy, D.; The discrete Normal distribution, Communications in Statistics Theory and Methods, 32, 1871-1883 (2003).

[28] Roy, D.; Discrete Rayleigh Distribution, IEEE Transactions on Reliability, 53(2) June (2004).

[29] Stein W. E. and Dattero W.; A new discrete Weibull distribution, IEEE Transactions on Reliability, R-33(2) June 196-197 (1984).

[30] Szablowski, P. J.; Discrete normal distribution and its relationship with Jacobi Theta functions, Statistics and Probability letters, 52 289-299 (2001).

[31] Yari, G and Tondpour, Z.; Some new discretization methods with application in reliability, Applications and Applied Mathematics, 13(2) 664-676(2018).

DOI: https://doi.org/10.15379/ijmst.v10i4.2156