Revista Colombiana de Estadística
January 2016, Volume 39, Issue 1, pp. 45 to 61
DOI: http://dx.doi.org/10.15446/rce.v39n1.55138

A Two Parameter Discrete Lindley Distribution

Distribución Lindley de dos parámetros

Tassaddaq Hussain[1,a], Muhammad Aslam[2,b], Munir Ahmad[3,c]

[1]Department of Statistics, Government Postgraduate College, Rawalakot, Pakistan

[2]Department of Statistics, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

[3]National College of Business Administration and Economics, Lahore, Pakistan

Abstract

In this article we have proposed and discussed a two parameter discrete Lindley distribution. The derivation of this new model is based on a two step methodology i.e. mixing then discretizing, and can be viewed as a new generalization of geometric distribution. The proposed model has proved itself as the least loss of information model when applied to a number of data sets (in an over and under dispersed structure). The competing models such as Poisson, Negative binomial, Generalized Poisson and discrete gamma distributions are the well known standard discrete distributions. Its Lifetime classification, kurtosis, skewness, ascending and descending factorial moments as well as its recurrence relations, negative moments, parameters estimation via maximum likelihood method, characterization and discretized bi-variate case are presented.

***Key words*:** Characterization, Discretized version, Estimation, Geometric distribution, Mean residual life, Mixture, Negative moments.

Resumen

En este artículo propusimos y discutimos la distribución Lindley de dos parámetros. La obtención de este Nuevo modelo está basada en una metodología en dos etapas: mezclar y luego discretizar, y puede ser vista como una generalización de una distribución geométrica. El modelo propuesto demostró tener la menor pérdida de información al ser aplicado a un cierto número de bases de datos (con estructuras de supra y sobredispersión). Los

[a]Professor. E-mail: taskho2000@yahoo.com
[b]Professor. E-mail: aslam_ravian@hotmail.com
[c]Professor. E-mail: drmunir@ncbae.edu.pk

modelos estándar con los que se puede comparar son las distribuciones Poisson, Binomial Negativa, Poisson Generalizado y Gamma discrete.Su clasificación de tiempo de vida, kurtosis, sesgamiento, momentos factorials ascendientes y descendientes, al igual que sus relaciones de recurrencia, momentos negativos, estimación de parámetros via máxima verosimilitud, caracterización y discretización del caso bivariado son presentados.

***Palabras clave*:** caracterización, estimación, distribución Geométrica, momentos negativos, mixtura, versión discretizada, vida media residual.

1. Introduction

For the last few decades discretized distributions have been studied extensively to model the discrete failure time data in statistical literature. Generally, discretized versions are obtained from any continuous distribution defined on the real line $\mathbb{R}$ with probability density function (pdf) $f(x)$, and are based on the support; the set of integers $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$ have a probability mass function that takes either of the two forms:

$$P(Y = x) = S(x) - S(x + 1), \tag{1}$$

or

$$P(Y = x) = f(x) \left(\sum_{k=-\infty}^{\infty} f(k) \right)^{-1}. \tag{2}$$

The former in statistical literature is known as the discrete concentration approach in which $S(x)$ is the preserved survival function of continuous distribution at integers, and the latter is called the time discretization approach in which $f(x)$ is the preserved probability density function (pdf) of the continuous distribution at integers.

Following the approach as given in equation (1) Nakagawa & Osaki (1975), Roy (2003, 2004), Krishna & Pundir (2009, 2007), Jazi, Lai & Alamatsaz (2010), Dĕniz & Ojeda (2011), Chakraborty & Chakravarty (2012), Al-Huniti & Al-Dayian (2012) and Hussain & Ahmad (2012, 2014) discretized continuous Weibull, Normal, Rayleigh, Burr and Pareto, Maxwell, Inverse Weibull, Lindley, Gamma, Inverse Gamma and Burr type -III, Inverse Gamma and Inverse Rayleigh distributions respectively. Similarly Kemp (1997), Szablowski (2001), Inusah & Kozubowski (2006), Kozubowski & Inusah (2006), Kemp (2006) and Nekoukhou, Alamatsaz & Bidram (2012) adopted the latter approach as is presented in equation (2) to discretize the Normal, Laplace, skew Laplace, half Normal and Generalized Exponential distributions respectively. Such discretized versions are being applied in the field of actuaries, engineering and biostatistics in which the lifetimes of persons, organisms or products are measured in months, weeks or days.

The discretization phenomenon generally arises when it becomes impossible or inconvenient to measure the life length of a product or device on a continuous scale. Such situations may arise, when the lifetimes need to be recorded on a discrete scale rather than on a continuous analogue. For examples the number of to and for motions of a pendulum or spring device before resting, the number of times devices are switched on/off, the number of days a patient stays in an observation ward, the length of successful stay of a pig (in terms of number of days/weeks, in the laboratory) and the number of weeks/months/years a cancer patient survives after treatment etc. Although there are a number of discrete distributions in the literature to model the above mentioned situations there is still a lot of space left to develop new discretized distribution that is suitable under different conditions.

In this article, a two parameter discrete Lindley distribution is proposed. It does not only have a simple structure, positively skewed and leptokurtic, but it also has more flexibility than the Děniz & Ojeda (2011) single parameter discrete Lindley distribution it also has less loss of information compared with standard discrete distributions. Derivation of two parameter discrete Lindley distribution, along with some properties, are given in section two, section three deals with estimation of parameters. In section four characterization issue is addressed, section five addresses the application of the proposed model and in the sixth section we study a discretized bivariate version of the two parameter Lindley distribution.

2. Definition and Properties of a Two Parameter Discrete Lindley Distribution

2.1. Derivation

The phenomenon of mixing and then discretizing the continuous distributions with the help of proper weights and a set of parameters is so far a new one. In this manuscript, after adopting the mixing and discretization technique, we have proposed a two parameter discrete Lindley distribution, which can be viewed as a new generalization of geometric distribution. Suppose $W_1 \sim \text{Gamma}(1, \theta)$ and $W_2 \sim \text{Gamma}(2, \theta)$, on mixing these densities with probabilities $p_1 = \frac{\theta}{\theta+\beta}$ and $p_2 = \frac{\beta}{\theta+\beta}$ so that $p_1 + p_2 = 1$ and $\beta \geq 0$, the resulting distribution of the random variable X will be a two parameter Lindley distribution, i.e. $X \sim p_1W_1 + p_2W_2$. This implies that

$$f(x) = \frac{\theta^2}{\theta+\beta}(1+\beta x)\exp(-\theta x), \quad x \geq 0, \beta \geq 0 \quad \text{and} \quad \theta > 0. \tag{3}$$

In order to model the discrete actuarial failure data Děniz & Ojeda (2011) proposed asingle parameter discrete Lindley distribution which is not considered a flexible model for analyzing different life time and actuarial data. It is used to model the over dispersed data pattern (see Děniz & Ojeda 2011), which is mathematically a complicated one. Therefore, to increase the flexibility for modeling purposes, we developed a two parameter discrete Lindley distribution by inserting equation (3) into equation (2).

Figure 1 depicts the curve behavior under different parameter combinations; the curve changes to a reverse J shape as $p \to 0$ and $\beta \to 0$. Moreover, peakedness of the curve decreases with a long right tail as p and β increases.

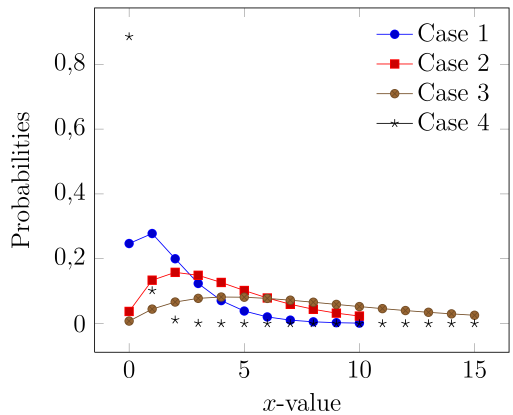

FIGURE 1: Probability graph for Two parameter Discrete Lindley Distribution Case 1. $\beta = 1.5$; $p = 0.45$; Case 2. $\beta = 4.5$; $p = 0.65$; Case 3. $\beta = 6.25$; $p = 0.80$; Case 4. $\beta = 0.15$; $p = 0.10$.

Definition 1. A random variable Y has a two parameter discrete Lindley distribution with parameters $0 < p < 1$ and $\beta \geq 0$, denoted by TDL (p, β), is defined as

$$P(Y = x) = p_x = \frac{(1-p)^2(1+\beta x)p^x}{(1+p(\beta-1))},\ x = 0, 1, 2, 3, \ldots, 0 < p < 1 \,\text{and}\, \beta \geq 0 \quad (4)$$

where $\exp(-\theta) = p$. For $\beta = 0$ the distribution geometrically reduces, and for $\beta = 1$ it becomes one a parameter discrete Lindley distribution. The recursive relation between TDL^s probabilities is given by

$$(1+\beta x)P(Y = x+1) = p(1+\beta(x+1))P(Y = x), \quad (5)$$

for $0 < p < 1$,$\beta \geq 0$ and $x = 0, 1, 2, 3, \ldots$

2.2. Reliability Characteristics of TDL

Although, most of the statistical research is based on continuous lifetime probability distributions to model the real lifetime phenomena, reliability engineers/anthropologists are looking for solutions, for which time can be interpreted as a discrete variable such as: the number of times a piece of equipment is operated; the life of a switch being measured by the number of strokes; the life of equipment being measured by the number of cycles it completes or the number of times it is operated prior to failure; the life of a weapon is measured by the number of rounds fired until failure; and the number of years of a married couple successfully completed; or the number of miles that a plane is flown before failures etc. In order to do this they started to discretize the continuous lifetime distributions. In reliability theory, the survivor, the hazard, the cumulative hazard,the accumulated hazard and mean residual life functions are important characteristics upon which the classification of discrete lifetime probability distribution is made. These classifications, in turn, point out the nature of the product for the reliability analyst, which could be for example, the increasing (decreasing) failure rate IFR (DFR) class, increasing (decreasing) failure rate average IFRA (DFRA) class, the new better (worse) than used NBU (NWU) class, new better (worse) than used in expectation NBUE (NWUE) class and increasing (decreasing) mean residual lifetime IMRL (DMRL) class etc. (see Kemp 2004). These classes are generally based on reliability/survival functions which give the probability that a component will survive beyond a specified time. The basic definition and formulae of the above mentioned characteristics for TDL are given below. Reliability function of TDL is defined and expressed as $S_x = P(Y \geq x) = \sum_{k=x}^{\infty} P(Y = k)$,

$$S_x = \frac{p^x((1-p)(1+\beta x) + p\beta)}{1 + p(\beta - 1)},$$

$0 < p < 1, \beta \geq 0$ and $x = 0, 1, 2, 3, \ldots$

Its failure rate function which gives the probability of failure given that it has not occurred before a specific time, is defined as $h_x = \frac{P(Y=x)}{P(Y \geq x)}$

$$h_x = \frac{(1-p)^2(1+\beta x)}{((1-p)(1+\beta x) + p\beta)},$$

$0 < p < 1,\beta \geq 0$ and $x = 0, 1, 2, 3, \ldots$

Its Mean Residual Life (MRL) function is defined and expressed as $MRL(Y) = E(Y - x|Y \geq x)$,

$$MRL(Y) = \frac{ph_x}{(1-p)^2} + \frac{p\beta(1+p)}{((1-p)(1+\beta x) + p\beta)(1-p)},$$

$0 < p < 1, \beta \geq 0$ and $x = 0, 1, 2, 3, \ldots$

Theorem 1. *If $Y \sim TDL(p, \beta)$ then the probability mass function (pmf)of the random variable Y is log-concave for all choices of β and independent of p.*

Proof. In order to show that the two parameter discrete Lindley distribution, as defined in equation (4), is log-concave, it is sufficient to show that for $\beta \geq 0$, $(p_x)^2 \geq p_{x-1}p_{x+1} \forall x = 1, 2, \ldots$ This implies that $(1 + \beta x)^2 \geq (1 + \beta x)^2 - \beta^2$. Generally, it is seen that the log-concave probability mass functions are strongly unimodal (see Kielson & Gerber 1971, Nekoukhou et al. 2012) and have an increasing failure rate (IFR) which suggest an intuitive concept caused by product wearing out. □

Therefore, we have the following Corollary.

Corollary 1. *If $Y \sim TDL(p, \beta)$ then the mode of the random variable Y is located at m and m satisfies $\frac{p\beta-(1-p)}{(1-p)\beta} \leq m \leq \frac{\beta-(1-p)}{(1-p)\beta}$. This implies that $p_x \geq p_{x-1}$ for $x \leq m$ and $p_{x+1} \leq p_x$ for $x \geq m$ as stated by Kielson & Gerber (1971), Abouammoh & Mashhour (1981) and Nekoukhou et al. (2012). From this, the following chain of implication for a two parameter discrete Lindley (TDL) distribution (see Kemp 2004) is suggested $IFR \Rightarrow IFRA \Rightarrow NBU \Rightarrow NBUE \Rightarrow DMRL$.*

Theorem 2. *If $Y \sim TDL(p, \beta)$, the probability generating function (pgf) of Y is given by*

$$G_Y(t) = \frac{(1-p)^2(1-pt(1-\beta))}{(1-pt)^2(1-p(1-\beta))},$$

$\beta \geq 0, 0 < p < 1$ *for* $0 < pt < 1$.

Proof. By definition the pgf can be expressed as $G_Y(t) = E(t^Y) = \sum_{x=0}^{\infty} t^x P(Y = x)$, this implies that

$$G_Y(t) = \frac{(1-p)^2}{1+p(\beta-1)} \sum_{x=0}^{\infty} (pt)^x (1+\beta x),$$

which on simplification yields

$$G_Y(t) = \frac{(1-p)^2(1-pt(1-\beta))}{(1-pt)^2(1-p(1-\beta))},$$

$\beta \geq 0, 0 < p < 1$ for $0 < pt < 1$.

On Replacing t by e^t we get a moment generating function (mgf) such as

$$M_Y(t) = \frac{(1-p)^2(1-pe^t(1-\beta))}{(1-pe^t)^2(1-p(1-\beta))}, \tag{6}$$

$\beta \geq 0, 0 < p < 1$ for $0 < pe^t < 1$. □

The first four derivatives of equation (6), with respect to t at $t = 0$, yield the first four moments about origin i.e. $\mu_r' = \frac{d^r M_Y(t)}{dt^r}|_{t=0}$ which after simplification are:

$$\mu_1' = \frac{(1-p+\beta(1+p))p}{(1+p(\beta-1))(1-p)},$$

$$\mu_2' = \frac{(1-p^2+\beta(1+4p+p^2))p}{(1+p(\beta-1))(1-p)^2},$$

$$\mu_3' = \frac{(1+3p-3p^2-p^3+\beta(1+11p+11p^2+p^3))p}{(1+p(\beta-1))(1-p)^3},$$

$$\mu_4' = \frac{(1+10p-10p^3-p^4+\beta(1+26p+66p^2+26p^3+p^4))p}{(1+p(\beta-1))(1-p)^4},$$

$$Var(Y) = \frac{(1-p)^2+(1-3p^2+2p)\beta+2(p\beta)^2}{(1+p(\beta-1))^2(1-p)^2},$$

$$Index of Dispersion = \frac{(1-p)^2+(1-3p^2+2p)\beta+2(p\beta)^2}{(1+p(\beta-1))(1+p(\beta-1)+\beta)(1-p)}.$$

2.3. Index of Dispersion

Index of dispersion (ID) for discrete distributions is defined as variance to mean ratio, it indicates whether a certain distribution is suitable for under or over dispersed data sets, and is used widely in ecology as a standard measure for measuring clustering (over dispersion) or repulsion (under dispersion) (see Johnson, Kotz & Kemp 1992). If ID$\geq 1 (\leq 1)$ the distribution is over dispersed (under dispersed). It is observed that if either $p \to 1$ or $\beta \to 0$, the distribution will always follow over dispersion and if $p \to 0$, $\beta \to \infty$ it will follow the under dispersion phenomenon. Moreover, the distribution is positively skewed with a longer tail compared to a one parameter discrete Lindley and leptokurtic in nature, which is evident from the ratio of the square of the third mean moment to the cube of the second mean moment and the ratio of the fourth mean moment to the square of variance. TDL is positively skewed for $\beta \to 0$ and $p \to 0$. Also, it approaches 2 and zero (0) as $p \to 1$ and $\beta \to \infty$ respectively, which is evident in Table 1. Moreover, it is leptokurtic in nature, and has high peakedness as $p \to 0$ and $\beta \to 0$ (see Table 2. Its peakedness approaches six as $p \to 1$ and for smaller values of p and higher values of β it becomes equal to 3.

Table 1: Skewness.

$\beta \downarrow$ p$\rightarrow$	0.05	0.10	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.05	21.09	11.61	6.95	5.45	4.75	4.34	4.08	3.86	3.59	3.11
0.50	14.89	8.24	4.90	3.78	3.19	2.82	2.55	2.34	2.17	2.05
3.50	4.31	2.34	1.68	1.68	1.78	1.87	1.93	1.97	1.99	1.99
6.50	2.18	1.26	1.28	1.58	1.80	1.92	1.97	1.99	2.00	2.00
10.50	1.12	0.77	1.22	1.67	1.90	1.99	2.02	2.02	2.01	2.00

Table 2: Kurtosis.

$\beta \downarrow p \rightarrow$	0.05	0.10	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.05	26.08	16.59	11.93	10.42	9.69	9.26	8.95	8.65	8.26	7.50
0.50	19.41	12.71	9.31	8.08	7.43	6.98	6.66	6.39	6.19	6.05
3.50	7.28	5.63	5.40	5.59	5.76	5.87	5.94	5.97	5.99	6.00
6.50	4.85	4.52	5.18	5.64	5.85	5.95	5.98	6.00	6.00	6.00
10.50	3.74	4.23	5.37	5.83	5.99	6.02	6.03	6.02	6.01	6.00

Theorem 3. *If $Y\sim$ TDL(p,β) and r^{th}, descending factorial moment of Y is given by*

$$\mu'_{(r)} = \frac{r! p^r (1-p)^{-r} (1 + \beta r + (\beta - 1)p)}{(1 + p(\beta - 1))}, \tag{7}$$

where $\beta \geq 0, 0 < p < 1, r = 0, 1, \ldots, (a)_n = a(a+1)(a+2)\cdots(a+n-1)$ *and* $\mu'_{(0)} = 1$.

Proof. The r^{th} descending factorial moment for r.v. Y can be defined as $\mu'_{(r)} = E(Y^{(r)}) = \sum_{x=0}^{\infty} x^{(r)} P(Y = x)$, which will from now an use the relation $x^{(r)} =$

$x(x-1)\dots(x-r+1)=\frac{x!}{(x-r)!}$. Then, we have

$$\mu'_{(r)}=\frac{(1-p)^2}{(1+p(\beta-1))}\sum_{x=r}^{\infty}\frac{x!}{(x-r)!}(1+\beta x)p^x,$$

by using the binomial series $(1-z)^{-a}=\sum_{x=0}^{\infty}\frac{(a)_n z^n}{n!}$, (see Rainville 1965). The followings reached:

$$\mu'_{(r)}=\frac{(1-p)^2 r!p^r}{(1+p(\beta-1))} \times\left\{(1-p)^{-(r+1)}+\beta r+(r+1)(r+1)p+(r+1)_2(r+2)\frac{p^2}{2!}\dots\right\},$$

$$\mu'_{(r)}=\frac{(1-p)^{-r}r!p^r}{(1+p(\beta-1))}\left\{(1-p)^{-(r+1)}+\beta r(1-p)^{-(r+1)}+\beta p(r+1)(1-p)^{-(r+2)}\right\},$$

After some algebraic manipulation, equation (7) is attained, which generates a recursive relation between r and $r-1$, descending factorial moments such as

$$\mu'_{(r)}\left\{(1-p)(1+\beta(r-1)+(\beta-1)p)\right\}=\left\{rp(1+\beta r+(\beta-1)p)\right\}\mu'_{(r-1)}$$

$\beta\geq 0, 0<p<1, r=1,2,\dots$ □

Theorem 4. *If $Y\sim TDL(p,\beta)$, the r^{th} ascending factorial moment of Y is given by*

$$\mu'_{[r]}=\frac{r!p(1-p)^{-r}(1+\beta-p+\beta pr)}{(1+p(\beta-1))}, \tag{8}$$

where $\beta\geq 0, 0<p<1, r=0,1,2,\dots,\mu'_{[0]}=\frac{p(1+\beta-p)}{(1+p(\beta-1))}$ and $(a)_n=a(a+1)(a+2)\dots(a+n-1)$.

Proof. The r^{th}, absolute ascending factorial moment can be defined and expressed as $\mu'_{[r]}=E((Y)_r)=\sum_{x=0}^{\infty}(x)_r P(Y=x)$. This implies that

$$\mu'_{[r]}=\frac{(1-p)^2}{(1+p(\beta-1))}\left\{\sum_{x=0}^{\infty}(x)_r p^x+\beta\sum_{x=0}^{\infty}x(x)_r p^x\right\},$$

by using the series $(1-z)^{-a}=\sum_{x=0}^{\infty}\frac{(a)_n z^n}{n!}$ (see Rainville 1965), we can get equation (8), where $\sum_{x=0}^{\infty}(x)_r p^x=r!p(1-p)^{-r-1}$ and $\beta\sum_{x=0}^{\infty}x(x)_r p^x=\beta r!p(1-p)^{-r-2}(1+pr)$. Equation (8) yields a recursive relation between r and $r-1$ ascending factorial moments such as

$$\mu'_{[r]}\left\{(1-p)(1+\beta-p+\beta p(r-1))\right\}=\left\{r(1+\beta-p+\beta pr)\right\}\mu'_{[r-1]},$$

$\beta\geq 0, 0<p<1, r=1,2,\dots$ □

Thus the theorem.

Theorem 5. *If $Y \sim TDL(p,\beta)$ the first order negative moment of Y is given by*

$$E(Y+a)^{-1} = \frac{(1-p)^2}{(1+p(\beta-1))} \times \left\{a^{-1}{}_2F_1(a,1;\, a+1;\, p) + \beta p(a+1)^{-1}{}_2F_1(a+1,2;\, a+2;\, p)\right\}, \quad (9)$$

where $a > 0, \beta \geq 0, 0 < p < 1$ and ${}_2F_1(a,b;\, c;\, z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n z^n}{(c)_n n!}$.

***Proof*.** By definition $E(\frac{1}{Y+a}) = \sum_{x=0}^{\infty}(x+a)^{-1}P(Y=x)$, so we have

$$E(\frac{1}{Y+a}) = \frac{(1-p)^2}{(1+p(\beta-1))}\sum_{x=0}^{\infty}(x+a)^{-1}p^x(1+\beta x).$$

After simplification we get equation (9) where $\sum_{x=0}^{\infty}(x+a)^{-1}p^x = a^{-1}{}_2\mathrm{F}_1(a,1;\, a+1;\, p)$ and $\sum_{x=0}^{\infty}(x+a)^{-1}\beta x p^x = \beta p(a+1)^{-1}{}_2\mathrm{F}_1(a+1,2;\, a+2;\, p)$. □

Corollary 2. *The s^{th} order negative moment of $Y \sim TDL(p,\beta)$ can be expressed as*

$$E(Y+a)^{-s} = \frac{(1-p)^2}{(1+p(\beta-1))}a^{-s}{}_{s+1}F_s(a,\ldots,a,1;a+1,\ldots,a+1;p) + \frac{(1-p)^2}{(1+p(\beta-1))}\beta p(a+1)^{-s}{}_{s+1}F_s(a+1,\cdots,a+1,2;a+2,\cdots,a+2;p),$$

where

$${}_sF_u(a_1,\ldots,a_s;b_1,\ldots,b_u;\, z) = \sum_{n=0}^{\infty}\frac{(a_1)_n,\ldots,(a_s)_n z^n}{(b_1)_n,\ldots,(b_u)_n n!},$$

the series converges for $s = u+1$ and $|z| < 1$.

Theorem 6. *If $Y \sim TDL(p,\beta)$, the s^{th} order negative factorial moment of Y is given by*

$$\mu'_{-[s]} = \frac{(1-p)^2}{(1+p(\beta-1))} \times \left\{((a)_s)^{-1}{}_2F_1(a,1;\, a+s;\, p) + \beta p((a+1)_s)^{-1}{}_2F_1(a+1,2;\, a+s+1;\, p)\right\}, \quad (10)$$

where $a > 0, s = 0,1,2,\ldots,\beta \geq 0, 0 < p < 1$ and ${}_2F_1(a,b;\, c;\, z) = \sum_{n=0}^{\infty}\frac{(a)_n(b)_n z^n}{(c)_n n!}$.

***Proof*.** The s^{th} order negative factorial moment for Y is defined as

$$\mu'_{-[s]} = E((\frac{1}{Y+a})^{[s]}) = \sum_{x=0}^{\infty}\prod_{i=0}^{s-1}(x+a+i)^{-1}P(Y=x),$$

on incorporating equation (4) we get

$$\mu'_{-[s]} = \frac{(1-p)^2}{(1+p(\beta-1))} \sum_{x=0}^{\infty} \prod_{i=0}^{s-1} (x+a+i)^{-1} p^x (1+\beta x),$$

which after simplification yields equation (10).

Where $\sum_{x=0}^{\infty} \prod_{i=0}^{s-1} (x+a+i)^{-1} p^x = ((a)_s)^{-1} {}_2\mathrm{F}_1(a, 1;\, a+s;\, p)$ and $\sum_{x=0}^{\infty} \prod_{i=0}^{s-1} (x+a+i)^{-1} \beta x p^x = \beta p((a+1)_s)^{-1} {}_2\mathrm{F}_1(a+1, 2;\, a+s+1;\, p)$. □

***Note* 1.**

$${}_2\mathrm{F}_1(a, b;\, c;\, z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n z^n}{(c)_n n!},$$

and $(a)_n = a(a+1)(a+2)\cdots(a+n-1)$, are the hypergeometric series function and the Pochhammer's symbol respectively, the series converges for $a, b, c \geq 0$ and $|z| \leq 1$.

3. Parameter Estimation

3.1. Maximum Likelihood Method

If $Y_1, Y_2, \ldots, Y_n$ be a random sample drawn identically independently from the two parameters discrete Lindley (TDL) distribution with observed values $x_1, x_2, \ldots, x_n$ then the joint probability function for TDL distribution can be expressed as

$$f(x_1, x_2, \ldots, x_n; p, \beta) = \frac{(1-p)^{2n}}{(1+p(\beta-1))^n} p^{\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} (1+\beta x_i),$$

$$\ln(L(p;\beta)) = 2n\ln(1-p) - n\ln((1+p(\beta-1))) + \sum_{i=1}^{n} x_i \ln p + \sum_{i=1}^{n} \ln(1+\beta x_i). \quad (11)$$

By partially differentiating both sides of equation (11) with respect to p and β, equating them to zero, we get MLE^s of p and β respectively, which can be shown as

$$\frac{\partial \ln(L(p,\beta))}{\partial p} = -\frac{2n}{1-p} - \frac{n(\beta-1)}{1+p(\beta-1)} + \frac{\sum_{i=1}^{n} x_i}{p} = 0,$$

$$\frac{2p}{1-p} + \frac{p(\beta-1)}{1+p(\beta-1)} = \bar{x}, \quad (12)$$

$$\frac{\partial \ln(L(p,\beta))}{\partial \beta} = -\frac{np}{1+(\beta-1)p} + \sum_{i=1}^{n} \frac{x_i}{1+\beta x_i} = 0,$$

this implies that

$$\sum_{i=1}^{n} \frac{x_i}{1+\beta x_i} = \frac{np}{1+p(\beta-1)}. \quad (13)$$

The MLE^s are computed using a computational package such as Mathematica [7.0]. In view of the regularity conditions as stated by Rohatgi and Saleh on page 419, the MLE^s i.e. $(\hat{p}, \hat{\beta})$ of TDL has a bivariate normal distribution with mean (p,β) and a variance-covariance matrix $(\mathrm{I}(p,\beta))^{-1}$. Thus $(\hat{p}, \hat{\beta}) \sim$ BVN((p,β), $(\mathrm{I}(p,\beta))^{-1}$) where I($p$,$\beta$) denotes the information matrix and is given below

$$I((p,\beta)|_{p=\hat{p},\beta=\hat{\beta}}) = \begin{bmatrix} E(\frac{-\partial^2 \ln L(p,\beta)}{\partial p^2}) & E(\frac{-\partial^2 \ln L(p,\beta)}{\partial p \partial \beta}) \\ & \\ E(\frac{-\partial^2 \ln L(p,\beta)}{\partial p \partial \beta}) & E(\frac{-\partial^2 \ln L(p,\beta)}{\partial \beta^2}) \end{bmatrix}$$

as entropy of a random variable is considered as the measure of the uncertainty of the random variable. It is used to measure the amount of information required to describe the random variable. In this regard, the entropy of the two parameter discrete Lindley (TDL) distribution is defined as:

$$H(Y) = \ln(\frac{1+p(\beta-1)}{(1-p)^2}) - \ln p(\frac{(1+\beta+p(\beta-1))p}{(1-p)(1+p(\beta-1))}) - \beta E\left\{x_2\mathrm{F}_1(1,1;\,2;\,-\beta x)\right\}$$

where $\ln(1+x) = x_2\mathrm{F}_1(1,1;\,2;\,-x)$ and ${}_2\mathrm{F}_1(a,b;\,c;\,z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n z^n}{(c)_n n!}$.

4. Characterization

Theorem 7. *Let Y be a nonnegative discrete random variable with probability mass function $P(Y = x)$ and $x \in \mathbb{Z}^+$; it then will follow the two parameter discrete Lindley distribution with parameters p and β iff*

$$MRL(Y) = \frac{ph_x}{(1-p)^2} + \frac{p\beta(1+p)}{((1-p)(1+\beta x)+p\beta)(1-p)}, \forall x \in \mathbb{Z}^+ \tag{14}$$

where $h_x = \frac{P(Y=x)}{S_x}$, $S_x = \frac{p^x((1-p)(1+\beta x)+p\beta)}{1+p(\beta-1)}$,$0 < p < 1$ and $\beta \geq 0$.

***Proof*.** Necessity: According to Kemp (2004) the MRL function is defined as $MRL(Y) = \frac{\sum_{k=x+1}^{\infty} S_k}{S_x}$, this implies that

$$MRL(Y) = \frac{((1-p)+\beta p)\sum_{k=x+1}^{\infty} p^k + \beta(1-p)\sum_{k=x+1}^{\infty} kp^k}{(1+p(\beta-1))S_x},$$

$$MRL(Y) = \frac{p^x p(1-p)\left\{(1-p)(1+\beta x)+(1+p)\beta\right\}}{(1-p)^2(1+p(\beta-1))S_x},$$

After simplification we get

$$MRL(Y) = \frac{ph_x}{(1-p)^2} + \frac{p\beta(1+p)}{\left\{(1-p)(1+\beta x)+p\beta\right\}(1-p)},$$

Sufficency: Suppose equation (14) holds then it can be written as

$$\sum_{k=x+1}^{\infty} S_k = \frac{pP(Y=x)}{1-p} + \frac{p\beta(1+p)p^x}{1+p(\beta-1)}, \tag{15}$$

Also

$$\sum_{k=x+1}^{\infty} S_k = \frac{P(Y=x+1)}{1-p} + \frac{2p^2\beta p^x}{1+p(\beta-1)}, \tag{16}$$

on comparing equation (15) with equation (16) we get

$$P(Y=x+1) - pP(Y=x) = \frac{p\beta p^x(1-p)^2(1+\beta x)}{(1+p(\beta-1))(1+\beta x)},$$

$$(1+\beta x)P(Y=x+1) = (1+(x+1)\beta)P(Y=x+1)p,$$

which gives $P(Y=0)=p_0=\frac{(1-p)^2}{1+p(\beta-1)}$ and $P(Y=x)=p_0p^x(1+\beta x)$. □

Theorem 8. *The random variable $Y \sim TDL(p,1)$ iff it can be written as $Y \equiv X_1+X_2$ where $X_i \sim G_0(q)$ i.e. $P(X=x_i)=p^{x_i}q, x_i=0,1,\dots$ for $i=1,2$ are independent random variables.*

***Proof*.** From equation (6) we get

$$M_Y(t) = \frac{(1-p)^2(1-pe^t(1-\beta))}{(1-pe^t)^2(1-p(1-\beta))},$$

$\beta \geq 0$, $0<p<1$ for $0<pe^t<1$.

For $\beta=1$ it simplifies as

$$M_Y(t) = \frac{(1-p)^2}{(1-pe^t)^2},$$

$0<p<1$ for $0<pe^t<1$ therefore

$$M_Y(t) = M_{X_1}(t)M_{X_2}(t),$$

$0<p<1$ for $0<pe^t<1$ which implies the result. □

5. Real Data Examples

We used two data sets reported by Chakraborty & Chakravarty (2012), to investigate the competence of the proposed model. Suitsbility of the proposed model is tested via the p-value and the *Akaike*^{s} information criteria (AIC) proposed by Hirotsugu Akaike in 1971. It was then compared with Poisson, Negative binomial, Generalized Poisson and discrete gamma distributions; as stated in Chakraborty & Chakravarty (2012).

Data set 1: The first data set contains observations on a number of European red mites on apple leaves and is presented in Table 3. Clearly this data set belongs to on over dispersed structure that has $ID = 1.9828$. We present the MLE, observed and expected frequencies, Log-Likelihood (LL), Chi-square values, p-values and AIC values for two parameter discrete Lindley distribution: $\hat{\beta} = 0.146$; $\hat{p} = 0.479$; $LL = -222.3$; $d.f. = 5$; $\chi^2 = 2.514$; p-value $= 0.774$; $AIC = 448.76$.

TABLE 3: Example 1.

Red mites	0	1	2	3	4	5	6	7	8	Total
Frequency	70	38	17	10	9	3	2	1	0	150
TDL(p,β)	68.90	37.81	20.42	10.89	5.75	3.00	1.56	0.81	0.41	150

Data set 2: The second data set in table 4 is about the observations collected from number of strikes in UK coal mining industries in four successive week periods during 1948-1959. The clearly data mentioned in Table 4 is an under dispersed data set with $ID = 0.7467$. The MLE, observed and expected frequencies, Log-Likelihood (LL), Chi-square values, p-values and AIC values for the two parameter discrete Lindley distribution are given below: $\hat{\beta} = 8.3855$; $\hat{p} = 0.176$; $LL = -187.44$; $d.f. = 3$; $\chi^2 = 0.5108$; p-value $= 0.92$; $AIC = 378$.

TABLE 4: Example 2.

Number of outbreaks	0	1	2	3	4	Total
Frequency	46	76	24	9	1	156
TDL(p,β)	45.95	76.08	25.39	6.59	1.53	156

It can be observed that in these examples the proposed model not only gives high p-values but also a minimum AIC compared to the distributions mentioned by Chakraborty & Chakravarty (2012). It therefore depicts the situation that the proposed model has the least loss of information in comparison with to the standard distributions.

6. Discretized Bivariate Case

Bivariate discrete random variables defined on integers or on non-negative values are used to model the paired count data that arise in a number of situations such as: in the analysis of accidents, e.g. the number of accidents in a site before and after infrastructure changes; in epidemiological analysis, e.g. incidents of different diseases in a series of districts; in medical research, e.g. the number of seizures before and after treatment etc. In this regard, literature on bivariate discrete distribution is sparse and worth mentioning particularly in terms of bivariate discretized distributions. Here, we give a bivariate discretized Lindley distribution by discretizing the bivariate continuous distribution, using the following equation:

$$P(Y_1 = x_1, Y_2 = x_2) = f(x_1, x_2) \left(\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} f(k_1, k_2) \right)^{-1} \tag{17}$$

6.1. Derivation

Suppose $\underline{W}' = (W_1, W_2)$ and $\underline{Z}' = (Z_1, Z_2)$ are two random vectors of independent random variables, each of which is distributed $Gamma(1, \theta)$ and $Gamma(2, \theta)$ respectively. Now, consider $p_1 = \frac{\theta}{\theta+\beta}$ and $p_2 = \frac{\beta}{\theta+\beta}$ are the respective weights of $\underline{W}'$ and $\underline{Z}'$ such that $p_1 + p_2 = 1$, $\theta > 0$ and $\beta \geq 0$. On mixing these densities, the resulting distribution of the random vector $\underline{X}' = (X_1, X_2)$ will be a bi variate two parameter Lindley distribution i.e. $\underline{X}' \sim p_1\underline{W}' + p_2\underline{Z}'$, and can be expressed as

$$f_{X_1,X_2}(x_1, x_2) = \frac{\theta^3}{\theta+\beta}(1 + \theta\beta x_1 x_2)(\exp(-\theta(x_1 + x_2))), \tag{18}$$

$x_1, x_2 \geq 0$, $\beta \geq 0$ and $\theta > 0$.

On substituting equation (18) into equation (17) we get a discretized version of bivariate Lindley distribution, expressed as:

$$P(Y_1 = x_1, Y_2 = x_2) = \frac{(1 - e^{-\theta})^4(1 + \theta\beta x_1 x_2)\exp(-\theta(x_1 + x_2))}{(1 - e^{-\theta})^2 + \beta\theta e^{-2\theta}}, \tag{19}$$

$x_1, x_2 \in \mathbb{Z}^+$, $\beta \geq 0$ and $\theta > 0$.

Its moment generating function is defined as:

$$M_{Y_1,Y_2}(t_1, t_2) = \frac{(1 - e^{-\theta})^4\left\{(1 - e^{-(\theta-t_1)})(1 - e^{-(\theta-t_2)}) + \beta\theta e^{-2\theta+t_1+t_2}\right\}}{((1 - e^{-\theta})^2 + \beta\theta e^{-2\theta})(1 - e^{-(\theta-t_1)})^2(1 - e^{-(\theta-t_2)})^2},$$

$|t_1, t_2| < 1$, $\theta \geq (t_1, t_2)$, $\beta \geq 0$ and $\theta > 0$.

Similarly, marginal probability mass functions of Y_i for $i = 1, 2$ are given by:

$$g_i(Y_i = x_i) = \frac{(1 - e^{-\theta})^2}{(1 - e^{-\theta})^2 + \beta\theta e^{-2\theta}}(1 + (\theta\beta x_i - 1)e^{-\theta})\exp(-\theta(x_i)), \tag{20}$$

$x_i \in \mathbb{Z}^+$, $\beta \geq 0$ and $\theta > 0$, and conditional probability mass functions of $Y_j \mid Y_i = x_i$, denoted by $g_j(Y_j \mid Y_i = x_i)$ for $i \neq j = 1, 2$, is expressed as:

$$g_j(Y_j \mid Y_i = x_i) = \frac{(1 - e^{-\theta})^2}{(1 + (\beta\theta x_i - 1)e^{-\theta}}(1 + \theta\beta x_i x_j)\exp(-\theta(x_j)), \tag{21}$$

$x_i, x_j \in \mathbb{Z}^+$, $\beta \geq 0$ and $\theta > 0$.

From equation (19) and (20) we get:

$$E(Y_1, Y_2) = \frac{e^{-2\theta}\left\{(1 - e^{-\theta})^2 + \beta\theta(1 + e^{-\theta})^2\right\}}{(1 - e^{-\theta})^2 + \beta\theta e^{-2\theta}(1 - e^{-\theta})^2}, \tag{22}$$

$Y_i, Y_j \in \mathbb{Z}^+$, $\beta \geq 0$ and $\theta > 0$.

$$E(Y_i) = \frac{e^{-\theta}\left\{(1 - e^{-\theta})^2 + \beta\theta e^{-\theta}(1 + e^{-\theta})\right\}}{(1 - e^{-\theta})^2 + \beta\theta e^{-2\theta}(1 - e^{-\theta})}, \tag{23}$$

$Y_i \in \mathbb{Z}^+$ for $i = 1, 2$, $\beta \geq 0$ and $\theta > 0$.

This implies that

$$Cov(Y_1, Y_2) = \frac{e^{-2\theta}\left\{(1-e^{-\theta})^2 + \beta\theta(1+e^{-\theta})^2\right\}}{(1-e^{-\theta})^2 + \beta\theta e^{-2\theta}(1-e^{-\theta})^2} - \frac{e^{-2\theta}\left\{(1-e^{-\theta})^2 + \beta\theta e^{-\theta}(1+e^{-\theta})\right\}^2}{\left\{(1-e^{-\theta})^2 + \beta\theta e^{-2\theta}\right\}^2 (1-e^{-\theta})^2}, \quad (24)$$

$Y_i \in \mathbb{Z}^+$ for $i = 1, 2$,$\beta \geq 0$ and $\theta > 0$. Its is obvious from equation (24) that for $\beta = 0$ the $Cov(Y_1, Y_2) = 0$.

Theorem 9. *Let $\underline{Y}' = (Y_1, Y_2)$ be a discrete random vector distributed according to equation (19), then the probability mass functions of $U = Y_1 + Y_2$ are expressed respectively as:*

$$g_U(U = u) = \frac{(1-e^{-\theta})^4(1+u)e^{-\theta u}}{(1-e^{-\theta})^2 + \beta\theta e^{-2\theta}}(1 + \frac{\beta\theta u(u-1)}{6}), \quad (25)$$

$u \in \mathbb{Z}^+$, $\beta \geq 0$ *and* $\theta > 0$.

Proof. Suppose $\underline{Y}' = (Y_1, Y_2)$ follows a two parameter discrete bivariate Lindley distribution as defined in equation (19). Let us consider $U = Y_1 + Y_2$ and $V = Y_1$. Now by using the change of variable technique we can write the joint probability mass function of U and V as:

$$P(U = u, V = v) = \frac{(1-e^{-\theta})^4}{(1-e^{-\theta})^2 + \beta\theta e^{-2\theta}}(1 + \theta\beta v(u-v))(\exp(-\theta u)), \quad (26)$$

$0 \leq v \leq u \in \mathbb{Z}^+$, $\beta \geq 0$ and $\theta > 0$.

Now, on summing over v we get the probability mass function of U as expressed in equation (25). □

Theorem 10. *If $\underline{Y}' = (Y_1, Y_2)$ is a discrete random vector distributed according to equation (19) then the probability mass functions of $V = Y_1 - Y_2$ are expressed respectively as:*

$$g_V(V = v) = \frac{(1-e^{-\theta})^4 e^{-\theta v}}{((1-e^{-\theta})^2 + \beta\theta e^{-2\theta})(1-e^{-2\theta})^3} \times ((1-e^{-2\theta})^3 + \beta\theta e^{-2\theta}(v - 1 - (v+1)e^{-2\theta})), \quad (27)$$

$v \in \mathbb{Z}$,$\beta \geq 0$ *and* $\theta > 0$.

Proof. Suppose $\underline{Y}' = (Y_1, Y_2)$ is defined according to equation (19). Now consider $V = Y_1 - Y_2$ and $M = Y_1$. According to change of variable technique, the joint probability mass function of V and M can be expressed as:

$$P(V = v, M = m) = \frac{(1-e^{-\theta})^4}{(1-e^{-\theta})^2 + \beta\theta e^{-2\theta}}(1 + \theta\beta m(v+m))(\exp(-\theta(v+2m))), \quad (28)$$

$v \in \mathbb{Z}$, $m \in \mathbb{Z}^+$, $\beta \geq 0$ and $\theta > 0$.

Now, on summing over m we get the probability mass function of V, as expressed in equation (27). □

7. Conclusion

A two parameter discrete Lindley distribution has been proposed. Its various distributional properties, reliability characteristics and characterization have been studied. It was found that this distribution has a simple structure, is more mathematically amenable, more flexible and has a longer tail than the one parameter discrete Lindley and other models in modeling actuarial and other count data from various fields such as ecology, health, psychology, sociology and engineering. It also has a less loss of information compared to the standard discrete distributions. Further issues such as characterization and mixtures are currently being researched and may be discussed in the further papers.

Acknowledgements

The authors are deeply thankful to the editor and the reviewer for their valuable suggestions to improve the quality of the paper.

[Received: September 2014 — Accepted: March 2015]

References

Abouammoh, A. & Mashhour, A. (1981), 'A note on the unimodality of discrete distributions', *Communication in Statistics Theory and Methods* **10**(13), 1345–1354.

Al-Huniti, A. & Al-Dayian, G. (2012), 'Discrete burr type-iii distribution', *American Journal of Mathematics and Statistic* **2**(5), 145–152.

Chakraborty, S. & Chakravarty, D. (2012), 'Discrete gamma distributions: Properties and parameters estimations', *Communication in Statistics Theory and Methods* **41**(18), 3301–3324.

Dĕniz, E. & Ojeda, E. (2011), 'The discrete lindley distribution: Properties and application', *Journal of Statistical Computation and Simulation* **81**(11), 1405–1416.

Hussain, T. & Ahmad, M. (2012), Discrete inverse gamma distribution, *in* M. Ahmad, ed., 'Official Statistics and its Impact on Good Governance and Economy of Pakistan', 9**th** International Conference on Statistical Sciences, Islamic Countries Society of Statistical Sciences, Lahore, Pakistan, pp. 381–394.

Hussain, T. & Ahmad, M. (2014), 'Discrete inverse rayleigh distribution', *Pakistan Journal of Statistics* **3**(2), 203–222.

Inusah, S. & Kozubowski, J. (2006), 'A discrete analogue of the laplace distribution', *Journal of Statistical Planning and Inference* **136**, 1090–1102.

Jazi, M., Lai, C. & Alamatsaz, M. (2010), 'A discrete inverse weibull distribution and estimation of its parameters', *Statistical Methodology* **7**, 121–132.

Johnson, N., Kotz, S. & Kemp, A. (1992), *Univariate Discrete Distributions*, 2 edn, John Wiley and Sons, New York.

Kemp, A. (1997), 'Statistical methodology', *Journal of Statistical Planning and Inference* **63**, 223–229.

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

Kemp, A. (2006), 'The discrete half-normal distribution', *Advances in mathematical and statistical modeling* **9**, 353–360.

Kielson, J. & Gerber, H. (1971), 'Some results for discrete unimodality', *Journal of American Statistical Association* **66**, 386–389.

Kozubowski, J. & Inusah, S. (2006), 'A skew laplace distribution on integers', *AISM* **58**, 555–571.

Krishna, H. & Pundir, P. (2007), 'Discrete maxwell distribution'. http://interstat.statjournals.net/YEAR/2007/articles/0711003.pdf (accessed in July 2010).

Krishna, H. & Pundir, P. (2009), 'Discrete burr and discrete pareto distributions', *Statistical Methodology* **6**, 177–188.

Nakagawa, T. & Osaki, S. (1975), 'The discrete weibull distribution', *IEEE Transactions on Reliability* **24**(5), 300–301.

Nekoukhou, V., Alamatsaz, M. & Bidram, H. (2012), 'A discrete analog of the generalized exponential distribution', *Communication in Statistics Theory and Methods* **41**(11), 2000–2013.

Rainville, E. (1965), *Special Functions*, The Macmillan Company, New York.

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

Roy, D. (2004), 'Discrete rayleigh distribution', *IEEE Transactions on Reliability* **52**(2), 255–260.

Szablowski, P. (2001), 'Discrete normal distribution and its relationship with jacobi theta functions', *Statistics and Probability Letters* **52**, 289–299.