Artigo Original

DOI:10.5902/2179460X25186

Ciência e Natura, Santa Maria v.39 n.3, 2017, Set - Dez, p. 539 – 552
Revista do Centro de Ciências Naturais e Exatas - UFSM
ISSN impressa: 0100-8307 ISSN on-line: 2179-460X

CIÊNCIA e NATURA

A comparative study between two discrete Lindley distributions

Um estudo comparativo entre duas distribuições Lindley discretas

Ricardo Puziol Oliveira[1], Josmar Mazucheli[1] and Jorge Alberto Achcar[2]

[1]Department of Statistics, Maringá State University, PR, Brazil
rpuziol.oliveira@gmail.com; jmazucheli@gmail.com
[2]Department of Social Medicine, University of São Paulo, SP, Brazil
achcar@fmrp.usp.br

Abstract

The methods to generate a probability function from a probability density function has long been used in recent years. In general, the discretization process produces probability functions that could be alternatives to traditional distributions used in the analysis of count data as geometric, Poisson and negative binomial distributions. The discretization also avoids the use of a continuous distribution in the analysis of strictly discrete data. In this paper, using the method based on an infinite series, proposed by Good (1953), we studied an alternative discrete Lindley distribution to those studies in Gómez-Déniz and Calderín-Ojeda (2011) and Bakouch et al. (2014). For both distributions, a simulation study is carried out to examine the bias and mean squared error for the maximum likelihood estimators of the parameters as well the coverage probability and the length of coverage probability. For the discrete Lindley distribution obtained by infinite series method, we present the analytical expression for bias reduction of the maximum likelihood estimator. Some examples using real data from the literature show the potential of these distributions. Despite the discretization methods are quite different, the resulting distributions are interchangeable, however the distribution generated by an infinite series has simple mathematical expressions and can be used directly to count data in the presence of covariates.

Keywords: *Discretization methods, Lindley distribution, likelihood, series, survival analysis, Monte Carlo simulation.*

Recebido: 14/03/2017 Aceito: 11/05/2017

1 Introduction

In recent years, the generation of a discrete observation from a continuous random variable has been considered by several authors (see, for example, Chakraborty (2015)). Basically, the main purpose to discretize a continuous probability density function is to generate a distribution for the analysis of strictly discrete data. For example, in survival data analysis it is common to use continuous distributions for discrete data, so the discretization acts with a subterfuge to avoid this process. A lot of applications considering continuous distributions in the analysis of discrete data are presented in many lifetime books as for example: Hamada et al. (2008); Collett (2003); Lee and Wang (2003); Lawless (2003); Kalbfleisch and Prentice (2002); Meeker and Escobar (1998); Klein and Moeschberger (1997) and others.

One of the first discretized distributions introduced in the literature was the Weibull distribution. From the Weibull distribution with probability density function:

$$f(x \mid \mu,\beta) = \frac{\beta}{\mu^{\beta}} x^{\beta-1} \exp\left[-\left(\frac{x}{\mu}\right)^{\beta}\right] \tag{1}$$

and survival function:

$$S(x \mid \mu,\beta) = \exp\left[-\left(\frac{x}{\mu}\right)^{\beta}\right]. \tag{2}$$

Nakagawa and Osaki (1975) proposed the discrete Weibull distribution whose probability function can be written as:

$$P(X = x \mid \mu,\beta) = \exp\left[-\left(\frac{x}{\mu}\right)^{\beta}\right] - \exp\left[-\left(\frac{x+1}{\mu}\right)^{\beta}\right] \tag{3}$$

where $x \in \mathbb{N}$ and $\mu,\beta > 0$ are, respectively, the scale and shape parameters. It is easy to verify that (3) is, in fact, a probability function.

Recently Nekoukhou et al. (2012), using the method based on an infinite series, have introduced the Generalized Exponential distribution whose probability function is written as:

$$P(X = x \mid \alpha,\lambda) = \lambda^{x-1}(1-\lambda^{x})^{\alpha-1}\left[\sum_{j=1}^{\infty}\binom{\alpha-1}{j}\frac{(-1)^{j}\lambda^{j}}{1-\lambda^{1+j}}\right]^{-1} \tag{4}$$

where $x \in \mathbb{N}$, $\binom{\alpha-1}{j} = \frac{1}{j!}(\alpha-1)\cdots(\alpha-j)$, $0 < \lambda < 1$ and $\alpha > 0$.

In this paper, also considering the method based on an infinite series, we introduce an alternative discrete Lindley distribution and a comparison of this model with the version presented in Gómez-Déniz and Calderín-Ojeda (2011) and Bakouch et al. (2014). In Section 2, two discretization methods are presented and expressions resulting from its application in Lindley distribution are displayed in Section 3. In section 4, the biases and mean squared error of the maximum likelihood estimates are studied. Some applications are presented in Section 5 and in Section 6 we present some concluding remarks.

2 Discretization methods

2.1 Discretization by survival function

Proposed by Nakagawa and Osaki (1975), this method discretize a continuous random variable from its survival function. Some properties for a discrete analogue to continuous distributions obtained by this method were studied by Kemp (2004), Bracquemond and Gaudoin (2003), Roy (2003), Chakraborty (2015), among others.

Following Kemp (2004), we can define an discrete analogue to continuous random variable as follows:

Definition 1: *Let X a continuous random variable. If X has survival function $S_X(x)$, then the discrete random variable $Y = \lfloor X \rfloor$, where $\lfloor X \rfloor$ indicates the smallest integer part or equal to X, has PMF (probability mass function) written as:*

$$P(Y = k) = \sum_{j=0}^{1}(-1)^{j} S_X(k+j). \tag{5}$$

It is easily verified that (5) is, in fact, a probability function for $x \in \mathbb{N}$. If the survival function of X has compact form, then the PMF (5) will have compact form.

Some distributions discretized by this method introduced in the literature are: Inverse Rayleigh distribution (Hussain and Ahmad, 2014), Lindley distribution (Gómez-Déniz and Calderín-Ojeda, 2011; Bakouch et al., 2014), Type II generalized Exponential distribution (Nekoukhou et al., 2013), Gamma distribution (Chakraborty and Chakravarty, 2012), Inverse Weibull distribution (Aghababaei Jazi et al., 2010), Burr XII and Pareto distributions (Krishna and Pundir, 2009), Rayleigh distribution (Roy, 2004), geometric Weibull distribution (Bracquemond and Gaudoin, 2003), among others.

2.2 Discretization by an infinite series

The first traces of this method were presented in Good (1953) in a modeling study of population frequency of species. Later, other authors such as Kulasekera and Tonkyn (1992), Doray and Luong (1997), Kemp (1997), Sato et al. (1999) studied this method and showed a version of it when the support of continuous random variable is defined in $(-\infty, \infty)$ or $(0, \infty)$.

Definition 2: *Let X be a continuous random variable. If X has pdf $f(x)$ with support $-\infty < x < \infty$, then the discrete random variable corresponding Y has PMF as follows:*

$$P(Y = k) = \frac{f(k)}{\sum_{j=-\infty}^{\infty} f(j)}. \tag{6}$$

In the case where the support of X is $(0, \infty)$, according to Sato et al. (1999), the PMF of Y is:

$$P(Y = k) = \frac{f(k)}{\sum_{j=0}^{\infty} f(j)}. \tag{7}$$

Some distributions discretized by this method introduced in the literature are: Pearson III distribution (Haight, 1957), Dirichlet's series distribution (Siromoney, 1964), Gaussian distribution (Kemp, 1997), Gamma and exponential distributions (Sato et al., 1999), Log-Gaussian distribution (Bi et al., 2001), Laplace distribution (Inusah and J. Kozubowski, 2006), Skew-Laplace distribution (Kozubowski and Inusah, 2006), Half-Gaussian distribution (Kemp, 2008), Beta-exponential distribution (Nekoukhou et al., 2012), among others.

3 The discrete Lindley distribution

3.1 Discretization by survival function

Let X be a continuous random variable with Lindley distribution. Using the survival function of X, Gómez-Déniz and Calderín-Ojeda (2011) and Bakouch et al. (2014) presented the discrete Lindley distribution with PMF written in the form:

$$P(X = x \mid \beta) = \frac{e^{-\beta x}}{1+\beta}[\beta(1 - 2e^{-\beta}) + (1 - e^{-\beta})(1 + \beta x)] \tag{8}$$

where $x \in \mathbb{N}$ and $\beta > 0$.

The behavior of (8) for some values of β is showed in Figure 1. Note that the PMF is unimodal and when $\beta > 1$, the mode is centered at the value zero (Bakouch et al. (2014)).

From (8) we have:

$$\mathbb{E}(X) = \frac{e^{-\beta}(2\beta - \beta e^{-\beta} + 1 - e^{-\beta})}{(1+\beta)(1 - e^{-\beta})^2}, \tag{9}$$

and:

$$\mathbb{V}(X) = \frac{e^{-\beta}\left[(-3\beta^2 - 4\beta - 2)e^{-\beta} + (2\beta + e^{-2\beta} + 1)(\beta + 1)\right]}{(1+\beta)^2(1 - e^{-\beta})^4}. \tag{10}$$

Analyzing the ratio between $\mathbb{E}(X)$ and $\mathbb{V}(X)$ we can see that $\mathbb{E}(X) < \mathbb{V}(X)$ for all $\beta > 0$. So this discrete version should only be used in data analysis with overdispersion. For more details, see Bakouch et al. (2014).

3.2 Estimation

Let $x_1, \ldots, x_n$ be a random sample from a distribution with PMF (8); the log-likelihood function of the discrete Lindley distribution is given by:

$$l(\beta \mid \mathbf{x}) = -n\beta\overline{x} - n\log(1+\beta) + \sum_{i=1}^{n} \log\left[1 + (1+x_i)\beta - (1+2\beta+\beta x_i)\exp(-\beta)\right]. \tag{11}$$

The maximum likelihood estimator $\widehat{\beta}$ of β is obtained by solving numerically, for β, the equation $\frac{d}{d\beta} l(\beta \mid \mathbf{x}) = 0$, where:

$$\frac{d}{d\beta} l(\beta \mid \mathbf{x}) = -n\overline{x} - \frac{n}{1+\beta} + \sum_{i=1}^{n} \frac{1 + x_i - (1 - 2\beta + x_i + \beta x_i)e^{-\beta}}{1 + (1+x_i)\beta - (1+2\beta+\beta x_i)e^{-\beta}}. \tag{12}$$

Note that this expression is non-linear in β and it must be solved numerically. However $\widehat{\beta} \approx -0.5\left(1 - \sqrt{1+4\overline{x}}\right)$ if $e^{-\beta} \approx 1$ (Bakouch et al., 2014).

The confidence intervals for β as well hypothesis tests of interest can be constructed from the asymptotic normality of the maximum likelihood estimates considering large sample sizes.

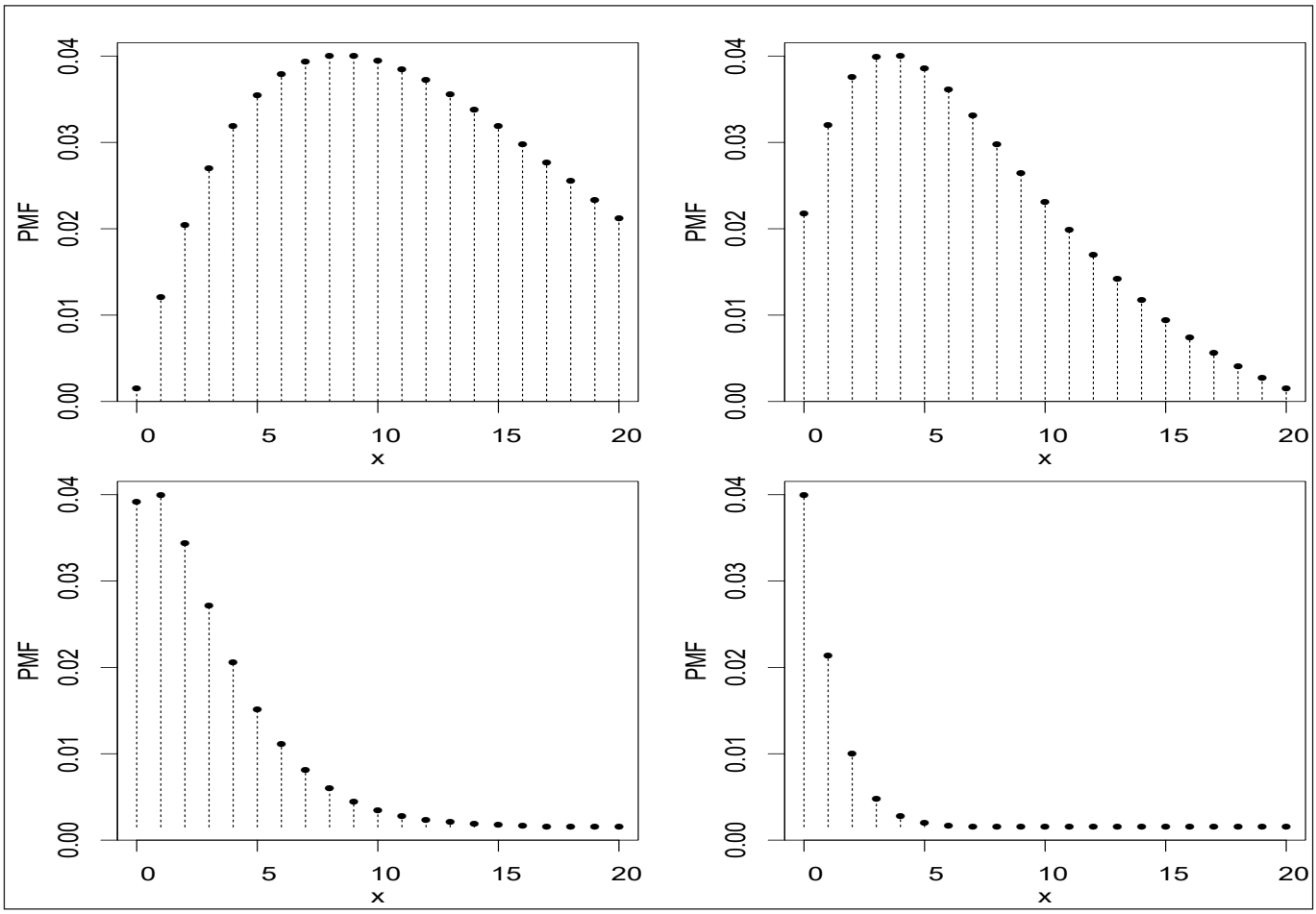

Figure 1: Behavior of the probability function of the discrete Lindley distribution, obtained by survival function, considering different values for β (upper-left panel: $\beta = 0.1$, upper-right panel: $\beta = 0.2$, lower-left panel: $\beta = 0.5$ and lower-right panel: $\beta = 1.2$).

3.3 Discretization by infinite series

By the discretization method presented in Section 2.2 the discrete Lindley has PMF written in the form:

$$P(X = x \mid \beta) = (1 + x)e^{-\beta(x+2)}(e^{\beta} - 1)^2 \tag{13}$$

where $x \in \mathbb{N}$ and $\beta > 0$.

According to the equation (13), it is observed that the PMF is unimodal with mode given as follows:

$$x_0 = \begin{cases} \left\lfloor \dfrac{e^{-\beta}}{1 - e^{-\beta}} \right\rfloor, & \text{if} \quad \dfrac{e^{-\beta}}{1 - e^{-\beta}} \notin \mathbb{Z}, \\ \dfrac{e^{-\beta}}{1 - e^{-\beta}}, & \text{if} \quad \dfrac{e^{-\beta}}{1 - e^{-\beta}} \in \mathbb{Z} \end{cases} \tag{14}$$

In fact, note that:

$$\begin{aligned} [P(X = x)]^2 &= (1 + x)^2(e^{-\beta(x+2)})^2(e^{\beta} - 1)^4 \\ &= (1 + x)^2 e^{-\beta(x+1)} e^{-\beta(x+3)}(e^{\beta} - 1)^4 \\ &\geq x e^{-\beta(x+1)}(e^{\beta} - 1)^2(x + 2)e^{-\beta(x+3)}(e^{\beta} - 1)^2. \end{aligned}$$

The right side of the above inequality is the same as $P(X = x - 1)P(X = x + 1)$. So, the equation (13) satisfies the log-concavity inequality $P^2(X = x) \geq P(X = x - 1)P(X = x + 1)$ for $x = 1, 2, \ldots$ and, therefore, by Theorem 3 from

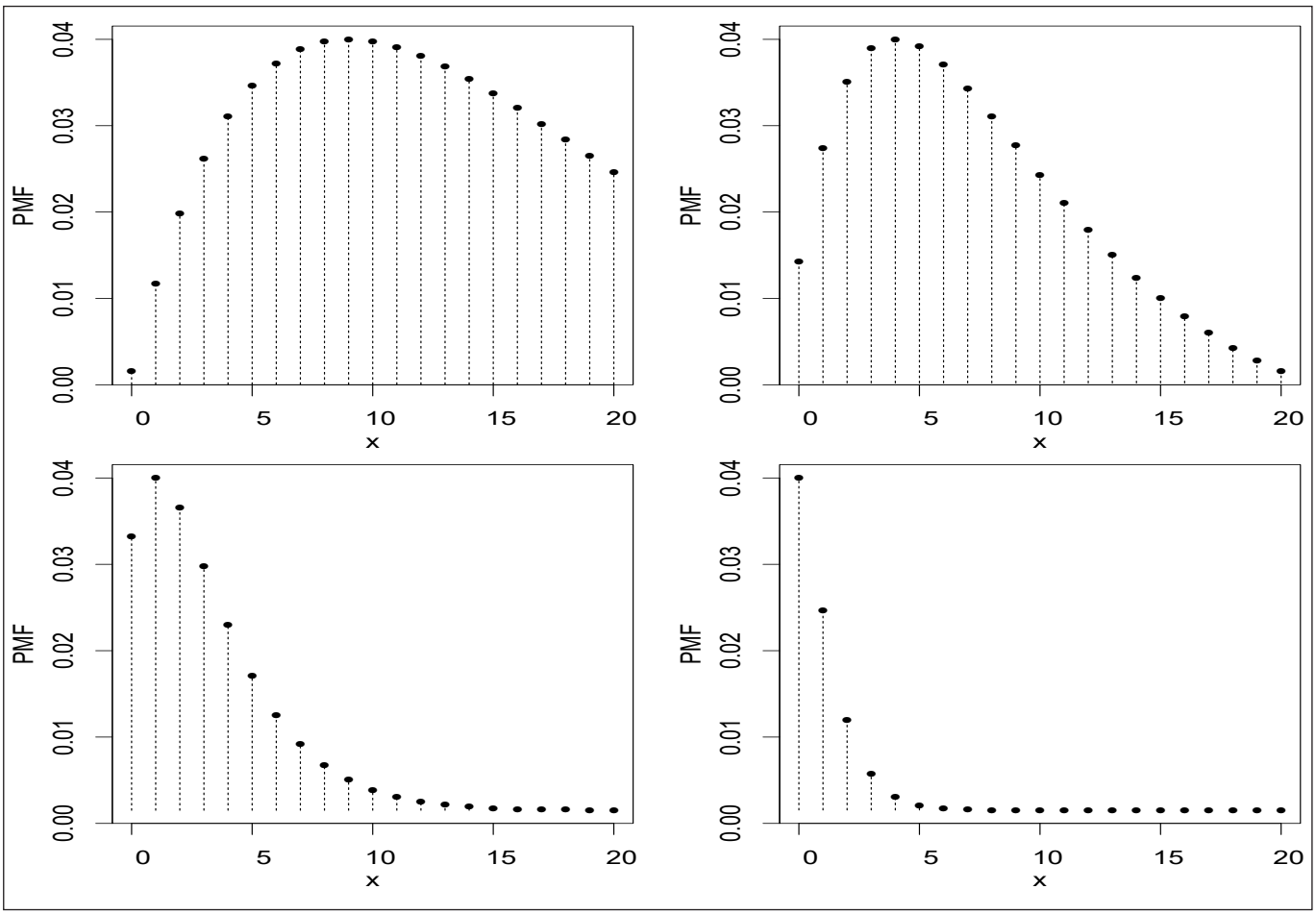

Figure 2: Behavior of the probability function of the discrete Lindley distribution, obtained by infinite series, considering different values for β (upper-left panel: $\beta = 0.1$, upper-right panel: $\beta = 0.2$, lower-left panel: $\beta = 0.5$ and lower-right panel: $\beta = 1.2$).

Keilson and Gerber (1971), is unimodal. In Figure 2 it is illustrated the behavior of (13) for some values of β.

For a random variable X with PMF (13) the corresponding probability generating function and the moment generating function can be expressed, respectively as:

$$G(k) = \mathbb{E}\left(k^X\right) = \left(\frac{e^\beta - 1}{e^\beta - k}\right)^2 \quad \text{and} \quad M(k) = \mathbb{E}\left(e^{kX}\right) = \frac{2e^{-2k}\left(e^\beta - 1\right)^2}{\left(e^{\beta-k} - 1\right)^3}.$$

Different from the discretized version obtained by survival function, the version proposed here has simple expressions for the mean and variance:

$$\mathbb{E}(X) = \frac{2}{e^\beta - 1} \quad \text{and} \quad \mathbb{V}(X) = \frac{2e^\beta}{(e^\beta - 1)^2}\ .$$

For all $\beta > 0$ it is easily to see that $\mathbb{E}(X) < \mathbb{V}(X)$. In this way, this distribution can be used in the count data analysis with overdispersion. The dispersion index is written as $\frac{e^\beta}{e^\beta - 1}$.

3.4 Estimation

Let $x_1, \ldots, x_n$ be a random sample from (13); the log-likelihood function is given by:

$$l(\beta \mid \mathbf{x}) \propto 2n \log(e^\beta - 1) - \beta(2n + n\overline{x}). \tag{15}$$

The maximum likelihood estimator of β is obtained by solving $\frac{d}{d\beta} l(\beta \mid \mathbf{x}) = 0$ in β. That is:

$$\frac{d}{d\beta} l(\beta \mid \mathbf{x}) = \frac{2ne^\beta}{e^\beta - 1} - 2n - n\overline{x} \quad \text{that is,} \quad \widehat{\beta} = \log\left(1 + \frac{2}{\overline{x}}\right). \tag{16}$$

The second derivative of the log-likelihood function is given by:

$$\frac{d^2}{d\beta^2} l(\beta \mid \mathbf{x}) = -\frac{2ne^\beta}{\left(e^\beta - 1\right)^2} \tag{17}$$

therefore:

$$\mathbb{E}\left[-\frac{d^2}{d\beta^2} l(\beta \mid \mathbf{x})\right] = \frac{2ne^\beta}{\left(e^\beta - 1\right)^2}. \tag{18}$$

Solving (18) locally in $\widehat{\beta}$ we have:

$$\widehat{Var}(\widehat{\beta}) = \frac{2}{n\overline{x}\,(2 + \overline{x})}. \tag{19}$$

Theorem 1: *The estimator $\widehat{\beta}$ of β is positively biased, that is, $E(\widehat{\beta}) - \beta > 0$.*
Proof: Let

$$\widehat{\beta} = g(\overline{X})$$

and

$$g(t) = \log\left(1 + \frac{2}{t}\right),$$

for $t > 0$. Since

$$g''(t) = \frac{4(t+1)}{t^2(t+1)^2} > 0,$$

$g(t)$ is strictly convex. Thus, by Jensen's inequality, we have $\mathbb{E}\left(g(\overline{X})\right) > g(\mathbb{E}(\overline{X}))$. Finally, since:

$$g(\mathbb{E}(\overline{X})) = g\left(\frac{2}{e^{\beta}+1}\right) = \log(1+e^{\beta}-1) = \beta.$$

we obtain $\mathbb{E}(\widehat{\beta}) > \beta$. Therefore, the estimator $\widehat{\beta}$ of β is positively biased.

Cox and Snell (1968) provided a framework for estimating the bias, to $O(n^{-1})$ for the maximum likelihood estimators of the parameters of regular densities. Then, subtracting the estimated bias from the original maximum likelihood estimator produces a bias-corrected estimator that is unbiased to $O(n^{-2})$. This type of bias adjustment can be applied successfully in the discrete Lindley distribution given in (13). Following Cox and Snell (1968) we have:

$$BIAS(\widehat{\beta}) = \left(\kappa^{11}\right)^2 [0.5\kappa_{111} + \kappa_{11,1}] + O\left(n^{-2}\right) \tag{20}$$

where $\kappa^{11} = \mathbb{E}\left[-\frac{\partial^2}{\partial\beta^2} l(\beta \mid \mathbf{x})\right]^{-1} = \frac{1}{2ne^{\beta}}\left(e^{\beta}-1\right)^2$, $\kappa_{11,1} = \mathbb{E}\left[\frac{\partial^2}{\partial\beta^2} l(\beta \mid \mathbf{x}) \times \frac{\partial}{\partial\beta} l(\beta \mid \mathbf{x})\right] = 0$ and $\kappa_{111} = \mathbb{E}\left[\frac{\partial^3}{\partial\beta^3} l(\beta \mid \mathbf{x})\right] = \frac{1}{\left(e^{\beta}-1\right)^3} 2ne^{\beta}\left(e^{\beta}+1\right)$.

In this way, the bias-corrected maximum likelihood estimator $\widehat{\beta}_{CMLE}$ can be written as:

$$\widehat{\beta}_{CMLE} = \widehat{\beta} - \frac{1}{4n}\left(e^{\widehat{\beta}} - e^{-\widehat{\beta}}\right). \tag{21}$$

Re-parameterizing (13) in terms of the mean $\theta = \dfrac{2}{e^{\beta}-1}$ we have $\beta = \log(\frac{2+\theta}{\theta})$ such that $\widehat{\theta} = \overline{x}$. The bias-corrected maximum likelihood estimator for θ is given by $\widehat{\theta} - \frac{1+\widehat{\theta}}{n}$. It is important to point out that in terms of θ we have

$$P(X = x \mid \theta) = 4\,(1+x)\,\theta^x\,(2+\theta)^{-(2+x)}, \tag{22}$$

such that $\mathbb{E}\,(X) = \theta$ and $\mathbb{V}(X) = \theta + \frac{\theta^2}{2}$.

4 Simulation study

In this section we estimated, by Monte Carlo simulation, the biases, the mean squared errors, the coverage probabilities and the coverage lengths for the maximum likelihood estimator, $\widehat{\beta}$, for discrete Lindley distributions obtained by survival function and infinity series. For computational stability, we assumed the values $\beta = 0.2, 0.5, 0.8, 1.0, 1.2$ and sample sizes $n = 10, 20, \ldots, 90, 100$. For each scenario, we calculated:

$$BIAS(\widehat{\beta}) = \frac{1}{N}\sum_{i=1}^{N}(\widehat{\beta}_i - \beta), \quad MSE(\widehat{\beta}) = \frac{1}{N}\sum_{i=1}^{N}(\widehat{\beta}_i - \beta)^2,$$

$$CP_{\beta}(n) = \frac{1}{N}\sum_{i=1}^{N} I\{\widehat{\beta}_i - 1.96\widehat{s}_{\widehat{\beta}_i} < \beta < \widehat{\beta}_i + 1.96\widehat{s}_{\widehat{\beta}_i}\} \text{ and } CL_{\beta}(n) = \frac{3.92}{N}\sum_{i=1}^{N} s_{\widehat{\beta}_i}$$

where $I\{\cdot\}$ denotes the indicator function and the number of simulations, $N = 10.000$. The simulation study are performed using R version 3.3.0 (R Core Team, 2015).

In Table 1, it is presented the simulation results for discrete Lindley distribution obtained by survival function. In Tables 2 and 3, are presented the simulation results for maximum likelihood estimator and bias-corrected maximum likelihood estimator for the discrete Lindley distribution obtained by infinite series.

In every scenario, for both discretizations, we have that the bias of $\widehat{\beta}$ is positive and tends to zero when the sample size increases. It is also observed that the mean square error of $\widehat{\beta}$ tends to zero in every scenario. Related to the coverage probabilities, we have $\mathrm{CP}_{\beta}(n)$ ranging from 0.94 to 0.96 and the coverage length tends to zero when the sample size increases.

Table 1: Estimated bias, mean-squared error, coverage probability and length of coverage probability for β (by survival function).

		Values of β						Values of β				
	n	0.2	0.5	0.8	1.0	1.2	n	0.2	0.5	0.8	1.0	1.2
BIAS	10	0.0141	0.0263	0.0640	0.0674	0.0902	10	0.9537	0.9358	0.9620	0.965	0.9610
	20	0.0080	0.0106	0.0340	0.0332	0.0435	20	0.9459	0.9480	0.9620	0.954	0.9590
	30	0.0056	0.0074	0.0200	0.0188	0.0339	30	0.9448	0.9418	0.9510	0.967	0.9540
	40	0.0039	0.0041	0.0129	0.0132	0.0250	40	0.9488	0.9500	0.9390	0.959	0.9410
	50	0.0030	0.0027	0.0080	0.0114	0.0201	50	0.9357	0.9458	0.9499	0.964	0.9550
	60	0.0025	0.0015	0.0055	0.0083	0.0170	60	0.9409	0.9510	0.9510	0.964	0.9570
	70	0.0022	0.0001	0.0050	0.0078	0.0153	70	0.9436	0.9519	0.9530	0.967	0.9560
	80	0.0018	0.0007	0.0037	0.0081	0.0125	80	0.9488	0.9469	0.9520	0.957	0.9500
	90	0.0018	0.0003	0.0039	0.0085	0.0102	90	0.9498	0.9469	0.9500	0.956	0.9389
	100	0.0016	0.0007	0.0013	0.0084	0.0070	100	0.9409	0.9480	0.9510	0.961	0.9459
MSE	10	0.0029	0.0177	0.0559	0.0826	0.1487	10	0.1895	0.4805	0.8229	1.0454	1.3118
	20	0.0014	0.0074	0.0216	0.0383	0.0564	20	0.1300	0.3284	0.5567	0.7079	0.8769
	30	0.0009	0.0050	0.0137	0.0217	0.0364	30	0.1049	0.2662	0.4457	0.5676	0.7076
	40	0.0006	0.0036	0.0100	0.0163	0.0262	40	0.0901	0.2289	0.3820	0.4881	0.6069
	50	0.0005	0.0028	0.0077	0.0124	0.0203	50	0.0802	0.2041	0.3393	0.4355	0.5398
	60	0.0004	0.0022	0.0064	0.0102	0.0163	60	0.0730	0.1858	0.3087	0.3960	0.4911
	70	0.0003	0.0018	0.0055	0.0089	0.0136	70	0.0675	0.1715	0.2855	0.3664	0.4538
	80	0.0003	0.0017	0.0049	0.0078	0.0124	80	0.0630	0.1606	0.2666	0.3428	0.4233
	90	0.0002	0.0015	0.0043	0.0069	0.0108	90	0.0594	0.1513	0.2514	0.3232	0.3981
	100	0.0002	0.0014	0.0038	0.0062	0.0096	100	0.0563	0.1436	0.2376	0.3066	0.3765

Table 2: Estimated bias, mean-squared error, coverage probability and length of coverage probability for β (by infinite series).

		Values of β						Values of β				
	n	0.2	0.5	0.8	1.0	1.2	n	0.2	0.5	0.8	1.0	1.2
BIAS	10	0.0137	0.0262	0.0537	0.0570	0.0748	10	0.949	0.950	0.964	0.972	0.957
	20	0.0068	0.0097	0.0286	0.0279	0.0351	20	0.947	0.955	0.953	0.956	0.960
	30	0.0049	0.0065	0.0173	0.0161	0.0283	30	0.948	0.939	0.957	0.959	0.953
	40	0.0035	0.0038	0.0111	0.0113	0.0210	40	0.940	0.959	0.948	0.965	0.952
	50	0.0027	0.0024	0.0072	0.0105	0.0169	50	0.940	0.944	0.951	0.967	0.952
	60	0.0025	0.0019	0.0048	0.0075	0.0135	60	0.940	0.953	0.959	0.955	0.953
	70	0.0023	0.0010	0.0043	0.0070	0.0123	70	0.936	0.947	0.956	0.959	0.947
	80	0.0019	0.0011	0.0032	0.0076	0.0098	80	0.942	0.949	0.954	0.950	0.956
	90	0.0020	0.0006	0.0036	0.0076	0.0080	90	0.948	0.947	0.953	0.959	0.951
	100	0.0017	0.0008	0.0013	0.0077	0.0056	100	0.945	0.956	0.944	0.958	0.944
MSE	10	0.0029	0.0171	0.0479	0.0675	0.1256	10	0.1877	0.4676	0.7761	0.9787	1.2153
	20	0.0012	0.0067	0.0187	0.0314	0.0472	20	0.1284	0.3196	0.5296	0.6682	0.8200
	30	0.0008	0.0045	0.0121	0.0185	0.0312	30	0.1039	0.2592	0.4258	0.5379	0.6640
	40	0.0006	0.0034	0.0090	0.0143	0.0229	40	0.0893	0.2232	0.3657	0.4632	0.5706
	50	0.0004	0.0027	0.0070	0.0110	0.0176	50	0.0796	0.1991	0.3254	0.4137	0.5081
	60	0.0004	0.0021	0.0059	0.0090	0.0140	60	0.0726	0.1816	0.2961	0.3764	0.4622
	70	0.0003	0.0018	0.0051	0.0079	0.0118	70	0.0671	0.1677	0.2739	0.3482	0.4273
	80	0.0003	0.0016	0.0045	0.0069	0.0105	80	0.0627	0.1569	0.2558	0.3259	0.3987
	90	0.0003	0.0014	0.0039	0.0061	0.0093	90	0.0591	0.1478	0.2413	0.3072	0.3752
	100	0.0002	0.0013	0.0034	0.0055	0.0083	100	0.0560	0.1403	0.2282	0.2915	0.3551

Table 3: Estimated bias, mean-squared error, coverage probability and length of coverage probability for β_{CMLE} (by infinite series).

		Values of β						Values of β				
	n	0.2	0.5	0.8	1.0	1.2	n	0.2	0.5	0.8	1.0	1.2
BIAS	10	0.0029	-0.0016	0.0044	-0.0086	-0.0139	10	0.941	0.933	0.946	0.943	0.939
	20	0.0015	-0.0037	0.0053	-0.0031	-0.0052	20	0.942	0.947	0.953	0.946	0.948
	30	0.0015	-0.0023	0.0020	-0.0041	0.0019	30	0.942	0.939	0.957	0.959	0.945
	40	0.0009	-0.0028	-0.0003	-0.0037	0.0014	40	0.934	0.952	0.943	0.965	0.952
	50	0.0006	-0.0028	-0.0018	-0.0015	0.0014	50	0.937	0.944	0.944	0.959	0.952
	60	0.0008	-0.0024	-0.0026	-0.0024	0.0006	60	0.936	0.953	0.950	0.955	0.948
	70	0.0008	-0.0028	-0.0021	-0.0015	0.0013	70	0.935	0.942	0.953	0.959	0.947
	80	0.0007	-0.0022	-0.0024	0.0002	0.0002	80	0.940	0.949	0.954	0.950	0.956
	90	0.0009	-0.0023	-0.0014	0.0010	-0.0005	90	0.946	0.947	0.949	0.955	0.944
	100	0.0007	-0.0018	-0.0031	0.0017	-0.0021	100	0.944	0.952	0.944	0.956	0.944
MSE	10	0.0024	0.0145	0.0380	0.0530	0.0911	10	0.1782	0.4423	0.7280	0.9115	1.1156
	20	0.0011	0.0063	0.0167	0.0281	0.0415	20	0.1251	0.3110	0.5138	0.6462	0.7896
	30	0.0007	0.0043	0.0113	0.0173	0.0284	30	0.1021	0.2546	0.4174	0.5262	0.6479
	40	0.0005	0.0033	0.0085	0.0136	0.0214	40	0.0882	0.2202	0.3603	0.4557	0.5603
	50	0.0004	0.0026	0.0068	0.0105	0.0167	50	0.0788	0.1970	0.3215	0.4084	0.5008
	60	0.0004	0.0021	0.0058	0.0087	0.0134	60	0.0720	0.1799	0.2932	0.3724	0.4567
	70	0.0003	0.0018	0.0050	0.0076	0.0113	70	0.0666	0.1665	0.2716	0.3450	0.4230
	80	0.0003	0.0016	0.0044	0.0067	0.0102	80	0.0623	0.1559	0.2539	0.3233	0.3952
	90	0.0002	0.0014	0.0039	0.0059	0.0090	90	0.0588	0.1469	0.2397	0.3051	0.3723
	100	0.0002	0.0012	0.0034	0.0054	0.0082	100	0.0557	0.1395	0.2269	0.2896	0.3526

5 Aplications

5.1 Application 1 (without covariates)

Consider a dataset related to the number of times that a computer break-down in each of 128 consecutive weeks of operation (Chakraborty and Chakravarty, 2012). The mean and variance are given respectively by, $\bar{x} = 4.023$ times and $s^2 = 14.464$ times^2, which evidences overdispersion. The fit of a discrete Lindley distribution obtained by infinite series (DLIS) was compared to the fit of another discrete Lindley distribution obtained by survival function (DLS), $P(X = x \mid \beta) = e^{-\beta x}(1+\beta)^{-1}[\beta(1-e^{-\beta}) + (1-e^{-\beta})(1+\beta x)]$ (Bakouch et al., 2014), a discrete Rayleigh (DR), $P(X = x \mid \theta) = \theta^{x^2} - \theta^{(x+1)^2}$ (Roy, 2004), a geometric (G), $P(X = x \mid \theta) = \theta^x - \theta^{(x+1)}$, and a Poisson (P), $P(X = x \mid \beta) = \frac{e^{-\beta}\beta^x}{x!}$.

The parameters were estimated by maximum likelihood method (MLE) and to compare the fits we considered the values of $-logL$, AIC, BIC and the χ^2 goodness-of-fit (see, Table 5). We conclude that, between DLIS and DLS, the results are almost the same. But, in terms of equations and computational stability, the DLIS distribution has a better fit when compared to the others distributions considered in this application.

Table 4: Observed and expected number of times that computer break-down considering the DLIS, DLS, DR, G, P and NB distributions.

Number of Break-Down	Observed	Expected					
		DLIS	DLS	DR	G	P	NB
0	15	14.11	16.48	3.63	25.48	2.29	16.09
1	18	18.85	18.91	10.29	20.40	9.21	19.38
2	24	18.88	18.14	15.29	16.34	18.53	18.46
3	14	16.82	15.95	18.03	13.09	24.85	16.03
4	15	14.04	13.32	18.43	10.48	25.00	13.25
5	10	11.25	10.76	16.91	8.39	20.12	10.62
6	8	8.77	8.48	14.17	6.72	13.49	8.33
+6	24	25.24	25.93	31.22	27.06	14.48	25.79
Total	128	128	128	128	128	128	128

Table 5: Parameter estimates and goodness-of-fit measures.

Distribution	MLE	S.E	$-\log L$	χ^2	p-value	D.F	AIC	BIC
DLIS	$\widehat{\beta} = 0.403$	0.025	316.795	2.286	0.891	6	635.59	638.44
DLS	$\widehat{\beta} = 0.381$	0.024	316.679	2.744	0.840	6	635.35	638.21
DR	$\widehat{\theta} = 0.971$	0.002	346.751	55.01	0.001	6	695.50	698.35
G	$\widehat{\theta} = 0.800$	0.015	320.925	11.07	0.085	6	643.85	646.70
P	$\widehat{\lambda} = 4.023$	0.177	384.276	102.8	0.001	6	770.55	773.40
NB	$\widehat{p} = 1.718$	0.320	316.471	2.496	0.777	5	636.94	642.64
	$\widehat{\mu} = 4.024$	0.324						

5.2 Application 2 (with covariates)

In this application, we considered a dataset introduced by Long (1990) related to the number of publications produced by Ph.D. biochemists to illustrate the application of a discrete Lindley distributions (DLIS and DLS) in presence of covariates. Its fit is compared to the negative binomial distribution.

Table 6: Dataset: Number of publications produced by Ph.D. biochemists.

Variable	Description	n	Mean	Variance	Min	Max
art	# articles produced in last three years of Ph.D	915	1.69	3.71	0	19
x_1	1 for females (two levels)	915	-	-	0	1
x_2	1 for married (two levels)	915	-	-	0	1
x_3	# of children under age six	915	0.50	0.59	0	3
x_4	prestige of Ph.D. program	915	3.10	0.97	0.73	3
x_5	# articles by mentor in last three years	915	8.77	89.94	0	77

This dataset have also been analyzed by Long et al. (2001) and is available from the Stata website `http://www.stata-press.com/data/lf2/couart2.dta`. The mean number of articles is 1.69 and the variance is 3.71, a little more than twice the mean (see Table 6). The data are over-dispersed. Results are showed in Tables 7 and 8. For both distributions we consider: $\log(\beta) = \beta_0 + \sum_{i=1}^{5} \beta_i x_i$ where x_i are described in Table 6.

Table 7: Parameter estimates and standard errors for Negative Binomial and Discrete Lindley models.

Parameter	Negative Binomial	DLIS	DLS	95% Conf. Int. N. Binomial	95% Conf. Int. DLIS	95% Conf. Int. DLS
β_0	0.2561 (0.1386)	0.2529 (0.1425)	0.5038 (0.1446)	(-0.015,0.527)	(-0.026,0.532)	(0.220, 0.787)
β_1	-0.2164 (0.0727)	-0.2159 (0.0747)	0.2092 (0.0758)	(-0.358,-0.074)	(-0.362,-0.069)	(0.060, 0.357)
β_2	0.1505 (0.0821)	0.1504 (0.0844)	-0.1471 (0.0857)	(-0.010,0.311)	(-0.015,0.316)	(-0.315, 0.020)
β_3	-0.1764 (0.0531)	-0.1761 (0.0545)	0.1708 (0.0549)	(-0.280,-0.072)	(-0.283,-0.069)	(0.063, 0.278)
β_4	0.01527 (0.0360)	0.0156 (0.0371)	-0.0154 (0.0376)	(-0.055,0.085)	(-0.057,0.088)	(-0.089, 0.058)
β_5	0.02908 (0.0035)	0.02926 (0.0036)	-0.0292 (0.0037)	(0.022,0.035)	(0.022,0.036)	(-0.036,-0.021)
α	0.4416 (0.0530)			(0.337,0.545)		

It is observed from the results in Table 7, that the DLIS distribution estimates are not very different from those obtained assuming the negative binomial model, and both sets would led to the same conclusions and looking at the standard errors, we see that both approaches to overdispersion lead to very similar estimated standard errors. However, the LDS estimates, except for the sign, are basically the same of the others models. Now, looking regression coefficients, we conclude that, in DLS distribution, β_2, β_4 are not significant; in DLIS distribution, $\beta_0, \beta_2, \beta_4$ are not significant; and, in negative binomial distribution, $\beta_0, \beta_2, \beta_4$ are not significant (see confidence intervals in Table 7). Also, looking the AIC (Akaike, 1974), AICc (Cavanaugh, 1997) and BIC (Bhat and Kumar, 2010) criterion introduced in Table 8, they are, basically, the same, but the DLS model is better in terms of parsimony and goodness of fit.

Table 8: Goodness-of-fit measures.

Goodness-of-fit Criteria	Negative Binomial	DLIS	DLS
-2 Log Likelihood	3121.9	3123.0	3133.0
AIC (smaller is better)	3135.9	3135.0	3141.0
AICC (smaller is better)	3136.0	3135.1	3141.2
BIC (smaller is better)	3169.6	3164.0	3160.2

5.3 Application 3 (with covariates)

In this application, we considered the dataset analyzed by Deb and Trivedi (1997) and Liu and Cela (2008) to illustrate just the application of discrete Lindley (DLIS), zero-inflated discrete Lindley (ZIDLIS) and Hurdle discrete Lindley (HDLIS) models in the presence of covariates (see, Remark 1). Its fit is compared to the Poisson, negative binomial, zero-inflated Poisson and Hurdle Poisson models. For all distributions we consider: $\log(\beta) = \beta_0 + \sum_{i=1}^{7} \beta_i x_i$ and $logit\,(p) = \alpha_0 + \sum_{i=1}^{7} \alpha_i x_i$ where x_i are describe in Table 9.

Table 9: Dataset: The number of hospital stays of 4406 respondents who were aged 66 or older and covered by Medicare program.

Variable	Description	n	Mean	Variance	Min	Max
hosp	# of hospital stays	4406	0.30	0.56	0	8
x_1	1 if self-perceived health is excellent	4406	0.08	0.7	0	1
x_2	1 if self-perceived health is poor	4406	0.13	0.11	0	1
x_3	# of chronic conditions	4406	1.54	1.82	0	8
x_4	age in years (divided by 10)	4406	7.4	0.40	6.6	10.9
x_5	1 if the person is male	4406	0.40	0.24	0	1
x_6	# of years of education	4406	10.29	13.98	0	18
x_7	1 if the person is covered by private insurance	4406	0.78	0.17	0	1

Remark 1: *In this application, we only used DLIS distribution since using DLS distribution we had computational instability for the parameter estimations.*

This dataset was originally obtained from National Medical Expenditure Survey (NMES) conducted in 1987 including 4406 respondents who were aged 66 or older and covered by Medicare program. The dataset description and summary statistics are given in Table 9 and we can show that the variance of hosp is about two times of the mean, implying the possibility of overdispersion.

Estimated coefficients of all models together with related statistics are listed in Tables 10 and 11. While Poisson regression provides a baseline model for count data, the other models demonstrate the better fit when compared to the basic Poisson regression model. The zero-inflated discrete Lindley model has the best fit when compared to the others models.

Looking at the standard errors of all models, we see that both approaches to overdispersion lead to very similar estimated standard errors and looking the AIC, AICc and BIC criterion, we conclude that the zero-inflated discrete Lindley model is the best fitted model in terms of goodness of fit.

Table 10: Parameter estimates and standard errors for all models.

Parameter	P	NB	DL	HP	ZIP	HDL	ZIDL
β_0	-3.3290	-3.7526	-3.5168	4.2294	4.2660	4.2294	5.6076
	(0.3397)	(0.4468)	(0.3761)	(0.4889)	(0.9712)	(0.4889)	(1.3472)
β_1	-0.7234	-0.6979	-0.7134	0.5826	-0.3699	0.5826	-0.4987
	(0.1756)	(0.1933)	(0.1809)	(0.1991)	(0.7174)	(0.1991)	(0.8369)
β_2	0.6262	0.6139	0.6202	-0.6953	-0.5897	-0.6953	-0.8732
	(0.0679)	(0.0954)	(0.0770)	(0.1073)	(0.1952)	(0.1073)	(0.3857)
β_3	0.2645	0.2894	0.2739	-0.3078	-0.2801	-0.3078	-0.4955
	(0.0183)	(0.0265)	(0.0211)	(0.0289)	(0.0624)	(0.0289)	(0.0994)
β_4	0.1864	0.2384	0.2101	-0.2750	-0.4060	-0.2750	-0.6517
	(0.0420)	(0.0553)	(0.0464)	(0.0606)	(0.1198)	(0.0606)	(0.1734)
β_5	0.1032	0.1539	0.1244	-0.1947	-0.3348	-0.1947	-0.6514
	(0.0563)	(0.0730)	(0.0620)	(0.0801)	(0.1624)	(0.0801)	(0.2317)
β_6	-0.0002	-0.0023	-0.0013	-0.0059	-0.0194	-0.0059	0.0135
	(0.0079)	(0.0102)	(0.0087)	(0.0113)	(0.0221)	(0.0113)	(0.0310)
β_7	0.1087	0.0939	0.1026	-0.0192	0.2249	-0.0192	0.1797
	(0.0693)	(0.0905)	(0.0766)	(0.0994)	(0.1961)	(0.0994)	(0.2834)
α		1.7667					
		(0.1605)					
α_0				-0.4818	-0.3665	-0.9664	-0.6733
				(0.5626)	(0.5720)	(0.6871)	(0.5709)
α_1				-0.9435	-0.9200	-1.0049	-0.8980
				(0.4953)	(0.4585)	(0.5254)	(0.4147)
α_2				0.3374	0.3249	0.3568	0.3738
				(0.1008)	(0.1012)	(0.1267)	(0.1162)
α_3				0.1427	0.1277	0.1529	0.1105
				(0.0297)	(0.0339)	(0.0375)	(0.0335)
α_4				-0.0108	-0.0244	0.0059	-0.0375
				(0.0683)	(0.0688)	(0.0836)	(0.0687)
α_5				-0.0382	-0.0596	-0.0411	-0.1203
				(0.0923)	(0.0991)	(0.1124)	(0.1020)
α_6				-0.0181	-0.0125	-0.0161	0.0042
				(0.0129)	(0.0135)	(0.0156)	(0.0136)
α_7				0.2592	0.2292	0.2595	0.1690
				(0.1140)	(0.1140)	(0.1375)	(0.1198)

P: Poisson, NB: Negative Binomial, HP: Hurdle Poisson, ZIP: Zero-Inflated Poisson, DL: Discrete Lindley. HDL: Hurdle Discrete Lindley and ZIDL: Zero-Inflated Discrete Lindley

Table 11: Goodness-of-fit measures.

	P	NB	HP	ZIP	DL	HDL	ZIDL
-2 Log Likelihood	6091.9	5713.1	5758.4	5755.9	5832.8	5699.7	5685.1
AIC (smaller is better)	6107.9	5731.1	5790.4	5787.9	5848.8	5731.7	5717.1
AICC (smaller is better)	6108.0	5731.2	5790.6	5788.0	5848.8	5731.8	5717.2
BIC (smaller is better)	6159.0	5788.6	5892.7	5890.1	5899.9	5833.9	5819.3

P: Poisson, NB: Negative Binomial, HP: Hurdle Poisson, ZIP: Zero-Inflated Poisson, DL: Discrete Lindley HDL: Hurdle Discrete Lindley and ZIDL: Zero-Inflated Discrete Lindley

6 Conclusion

In this paper, considering a discretization method based on an infinite series, we introduce an alternative discrete Lindley distribution. Some characteristics and properties of this distribution were presented and studied which it was found that it can be used in the analysis of data with overdispersion. Monte Carlo studies showed that the biases and mean squared errors of this distribution are asymptotically non-biased and has small values compared to discrete Lindley distribution obtained by survival function considered in Bakouch et al. (2014) and has great coverage probabilities ranging from 0.94 to 0.96 and the coverage length goes to zero when the sample size increases. In the considered applications, the DLIS distribution had a better or equivalent fit compared to other distributions considered in the applications leading to the conclusion that this distribution could be a good alternative for overdispersed count data in presence or not of covariates, especially, it is better than DLS distribution in computational aspects (simulation and estimation), equations and goodness-of-fit.

References

Aghababaei Jazi, M., Lai, C. D., Hossein Alamatsaz, M. (2010). A discrete inverse Weibull distribution and estimation of its parameters. *Statistical Methodology*, *7*, 121–132.

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, *19*(6), 716–723.

Bakouch, H. S., Jazi, M. A., Nadarajah, S. (2014). A new discrete distribution. *Statistics*, *48*(1), 200–240.

Bhat, H. S., Kumar, N. (2010). On the derivation of the bayesian information criterion. *School of Natural Sciences, University of California*.

Bi, Z., Faloutsos, C., Korn, F. (2001). The DGX distribution for mining massive, skewed data. Em: *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, pp. 17–26.

Bracquemond, C., Gaudoin, O. (2003). A survey on discrete lifetime distributions. *International Journal of Reliability, Quality and Safety Engineering*, *10*(01), 69–98.

Cavanaugh, J. E. (1997). Unifying the derivations for the akaike and corrected akaike information criteria. *Statistics & Probability Letters*, *33*(2), 201–208.

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

Chakraborty, S., Chakravarty, D. (2012). Discrete gamma distributions: properties and parameter estimations. *Communications in Statistics-Theory and Methods*, *41*(18), 3301–3324.

Collett, D. (2003). *Modelling Survival Data in Medical Research*, 2º edn. Chapaman and Hall, New York.

Cox, D. R., Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society Series B (Methodological)*, *30*(2), 248–275.

Deb, P., Trivedi, P. K. (1997). Demand for medical care by the elderly: a finite mixture approach. *Journal of applied Econometrics*, *12*(3), 313–336.

Doray, L. G., Luong, A. (1997). Efficient estimators for the good family. *Communications in Statistics-Simulation and Computation*, *26*(3), 1075–1088.

Gómez-Déniz, E., Calderín-Ojeda, E. (2011). The discrete Lindley distribution: properties and applications. *Journal of Statistical Computation and Simulation*, *81*(11), 1405–1416.

Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, *40*(3-4), 237–264.

Haight, F. A. (1957). Queueing with balking. *Biometrika*, *44*(3/4), 360–369.

Hamada, M. S., Wilson, A. G., Reese, C. S., Martz, H. F. (2008). *Bayesian reliability*. Springer Series in Statistics, Springer, New York.

Hussain, T., Ahmad, M. (2014). Discrete inverse Rayleigh distribution. *Pak J Statist*, *30*(2), 203–222.

Inusah, S., J. Kozubowski, T. (2006). A discrete analogue of the Laplace distribution. *Journal of Statistical Planning and Inference*, *136*.

Kalbfleisch, J. D., Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*, 2º edn. Wiley, New York, NY.

Keilson, J., Gerber, H. (1971). Some results for discrete unimodality. *Journal of the American Statistical Association*, *66*.

Kemp, A. W. (1997). Characterizations of a discrete normal distribution. *Journal of Statistical Planning and Inference*, *63*(2), 223 – 229, in Honor of C.R. Rao.

Kemp, A. W. (2004). Classes of discrete lifetime distributions. *Taylor & Francis*.

Kemp, A. W. (2008). The discrete half-normal distribution. Em: *Advances in mathematical and statistical modeling*, Springer, pp. 353–360.

Klein, J. P., Moeschberger, M. L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag, New York.

Kozubowski, T. J., Inusah, S. (2006). A skew Laplace distribution on integers. *Annals of the Institute of Statistical Mathematics*, *58*(3), 555–571.

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

Kulasekera, K., Tonkyn, D. W. (1992). A new discrete distribution, with applications to survival, dispersal and dispersion. *Communications in Statistics-Simulation and Computation*, *21*(2), 499–518.

Lawless, J. F. (2003). *Statistical models and methods for lifetime data*, 2º edn. Wiley Series in Probability and Statistics, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ.

Lee, E. T., Wang, J. W. (2003). *Statistical methods for survival data analysis*, 3º edn. Wiley Series in Probability and Statistics, Hoboken, NJ.

Liu, W., Cela, J. (2008). Count data models in SAS. Em: *SAS Global Forum*, Citeseer.

Long, J. S. (1990). The origins of sex differences in science. *Social forces*.

Long, J. S., Freese, J., et al. (2001). Predicted probabilities for count models. *Stata Journal*, *1*(1), 51–7.

Meeker, W. Q., Escobar, L. A. (1998). *Statistical Methods for Reliability Data*. John Wiley & Sons, New York.

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

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

Nekoukhou, V., Alamatsaz, M. H., Bidram, H. (2013). Discrete generalized exponential distribution of a second type. *Statistics*, *47*, 876–887.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Roy, D. (2003). The discrete normal distribution. *Communication in Statistics- Theory and Methods*, *32*, 1871–1883.

Roy, D. (2004). Discrete Rayleigh distribution. *Reliability, IEEE Transactions on*, *53*(2), 255–260.

Sato, H., Ikota, M., Sugimoto, A., Masuda, H. (1999). A new defect distribution metrology with a consistent discrete exponential formula and its applications. *Semiconductor Manufacturing, IEEE Transactions on*, *12*(4), 409–418.

Siromoney, G. (1964). The general Dirichlet's series distribution. *Journal of the Indian Statistical Association*, *2*.