Revista Colombiana de Estadística
January 2020, Volume 43, Issue 1, pp. 21 to 48
DOI: http://dx.doi.org/10.15446/rce.v43n1.77052

Two Useful Discrete Distributions to Model Overdispersed Count Data

Dos distribuciones discretas útiles para modelar datos de recuento sobredispersos

Josmar Mazucheli[1,a], Wesley Bertoli[2,b], Ricardo Oliveira[3,c]

[1]Department of Statistics, State University of Maringá, Maringá, Brazil

[2]Department of Statistics, Federal University of Technology - Paraná, Curitiba, Brazil

[3]Medical School of Ribeirão Preto, University of São Paulo, Ribeirão Preto, Brazil

Abstract

The methods to obtain discrete analogs of continuous distributions have been widely considered in recent years. In general, the discretization process provides probability mass functions that can be competitive with the traditional model used in the analysis of count data, the Poisson distribution. The discretization procedure also avoids the use of continuous distribution in the analysis of strictly discrete data. In this paper, we seek to introduce two discrete analogs for the Shanker distribution using the method of the infinite series and the method based on the survival function as alternatives to model overdispersed datasets. Despite the difference between discretization methods, the resulting distributions are interchangeable. However, the distribution generated by the method of the infinite series method has simpler mathematical expressions for the shape, the generating functions, and the central moments. The maximum likelihood theory is considered for estimation and asymptotic inference concerns. A simulation study is carried out in order to evaluate some frequentist properties of the developed methodology. The usefulness of the proposed models is evaluated using real datasets provided by the literature.

***Key words*:** Maximum likelihood estimation; Discrete distributions; Monte Carlo simulation; Overdispersion; Shanker distribution.

[a]PhD. E-mail: jmazucheli@gmail.com
[b]PhD. E-mail: wbsilva@utfpr.edu.br
[c]PhD. E-mail: rpuziol.oliveira@gmail.com

Resumen

Los métodos para obtener análogos discretos de distribuciones continuas han sido ampliamente considerados en los últimos años. En general, el proceso de discretización proporciona funciones de probabilidad en masa que pueden ser competitivas con el modelo tradicional utilizado en el análisis de datos de conteo, la distribución de Poisson. El procedimiento de discretización también evita el uso de la distribución continua en el análisis de datos estrictamente discretos. En este artículo, intentamos introducir dos análogos discretos para la distribución de Shanker utilizando el método de la serie infinita y el método basado en la función de supervivencia como alternativas para modelar conjuntos de datos sobre dispersados. A pesar de la diferencia entre los métodos de discretización, las distribuciones resultantes son intercambiables. Sin embargo, la distribución generada por el método de series infinitas tiene expresiones matemáticas más simples para la forma, las funciones de generación y los momentos centrales. La teoría de máxima verosimilitud se considera para la estimación y las preocupaciones de inferencia asintótica. Se lleva a cabo un estudio de simulación para evaluar algunas propiedades frecuentistas de la metodología desarrollada. La utilidad de los modelos propuestos se evalúa utilizando conjuntos de datos reales proporcionados por la literatura.

***Palabras clave*:** Estimación de máxima verosimilitud; Distribuciones discretas; Distribución de Shanker; Simulación del Monte Carlo; Sobredispersión.

1. Introduction

In recent decades, the building of a probabilistic distribution by discretization of a continuous random variable has been widely addressed in the literature. The main purpose of the discretization is to generate distributions that can be used for the analysis of strictly discrete data. For example, in survival analysis is common to use continuous distributions to model discrete data, so the discretization acts as a subterfuge to avoid this process. Several applications where continuous distributions were used to model discrete data can be found in Klein & Moeschberger (1997), Meeker & Escobar (1998), Kalbfleisch & Prentice (2002), Lee & Wang (2003), Lawless (2003), Collett (2003), Hamada, Wilson, Reese & Martz (2008), among others. A complete survey regarding all discretization methods and some discretized distributions can be found in Chakraborty (2015*a*).

One of the first proposed discretization methods is based on the definition of a probability mass function (pmf) that depends on an infinite series. The first traces of this method were presented by Good (1953), who proposed the discrete Good distribution to model population frequencies of species. Such an approach was considered by other authors to define discrete analogs, and we will point out a few. Haight (1957) proposed the discrete Pearson III distribution to model queues with baking and Siromoney (1964) introduced the Dirichlet's Series distribution as an alternative model to describe the frequency of wet days (rain-spells). After a long break, this method was revived by Kemp (1997) that formally introduced the discrete Normal distribution and derived its main statistical properties. The dis-

crete Exponential distribution was proposed by Sato, Ikota, Sugimoto & Masuda (1999) to describe the defect count frequencies on wafers or chips. Bi, Faloutsos & Korn (2001) introduced the discrete Lognormal distribution and with application to internet clickstream data, among others. Inusah & Kozubowski (2006) presented the discrete Laplace distribution discussing that, relative to the discrete Normal, the proposed model has closed forms for the pmf, for the generating functions and the central moments. The skewed version of the discrete Laplace distribution was proposed by Kozubowski & Inusah (2006). Further, Kemp (2008) introduced the discrete Half-Normal distribution studying its relation with other distributions and Nekoukhou, Alamatsaz & Bidram (2012) proposed the discrete Generalized Exponential distribution as an attempt to model ranking frequencies of graphemes in the Slovene language.

Another popular method to obtain discrete analogs of continuous random variables is the one based on the survival function (sf) of the original distribution. This method was proposed by Nakagawa & Osaki (1975) and has the interesting feature of preserving the original sf of its integer part for the generated pmf (Kemp 2004, Chakraborty 2015*a*). Several authors also considered the discretization method based on the sf, and we will point out a few. Nakagawa & Osaki (1975) proposed the discrete Weibull distribution and discussed its main properties. The Geometric-Weibull distribution considering a discrete analog for the Weibull component was introduced by Bracquemond & Gaudoin (2003). Roy (2004) proposed the discrete Rayleigh distribution and presented its usefulness in the stress-strength analysis. The discrete Burr and Pareto distributions were introduced by Krishna & Pundir (2009) for application in reliability estimation in series systems. Jazi, Lai & Alamatsaz (2010) proposed the discrete Inverse Weibull distribution and discussed different estimation methods for the model parameters. Gómez-Déniz & Calderín-Ojeda (2011) introduced the discrete Lindley distribution and illustrated its application using an automobile claim frequency data. The discrete Gamma distribution was proposed by Chakraborty & Chakravarty (2012), which derived several statistical properties of this model. Besides, Nekoukhou, Alamatsaz & Bidram (2013) presented the discrete Type II Generalized Exponential distribution, and Hussain & Ahmad (2014) introduced the discrete Inverse Rayleigh distribution as alternatives to model overdispersed count data.

The main aim of this paper is to use the methods of infinite series and of the sf to propose discrete analogs for the Shanker distribution, which is a one-parameter lifetime model proposed by Shanker (2015). We expect the proposed models to be suitable alternatives to model overdispersed count datasets. The Shanker distribution can be seen as a modification of the one-parameter Lindley distribution (Ghitany, Atieh & Nadarajah 2008).

A continuous random variable X is said to have Shanker distribution if its probability density function (pdf) can be written as

$$f_X(x;\theta) = \frac{\theta^2}{\theta^2+1}(\theta+x)\,e^{-\theta x}, \qquad x \in \mathbb{R}_+, \tag{1}$$

where $\theta \in \mathbb{R}_+$ is the shape parameter. The author has shown that this model is a two component mixture of an Exponential distribution with scale parameter θ

and a Gamma distribution having shape parameter 2 and scale parameter θ, with mixing proportions given, respectively, by $\theta^2\nu^{-1}$ and ν^{-1}, where $\nu = \theta^2 + 1$. For the one-parameter Lindley distribution the mixing proportions are, respectively, $\theta\nu^{-1}$ and ν^{-1}, where $\nu = \theta + 1$.

A comprehensive discussion about the statistical properties of the Shanker distribution, such as moments, hazard function, stochastic orderings, parameter estimation, among others is also presented on the mentioned paper. The corresponding sf is given by

$$S_x\left(x;\theta\right) = \left[1 + \frac{\theta x}{\theta^2 + 1}\right] e^{-\theta x}, \qquad x \in \mathbb{R}_+, \tag{2}$$

for $\theta \in \mathbb{R}_+$.

This paper is organized as follows. In Section 2, we briefly present the methods of infinite series and of the sf to define discrete analogs of continuous distributions. In Section 3, we introduce two types for the discrete Shanker distribution and derive the main statistical properties of each model. In Section 4, the problem of estimating the parameter of the proposed models is addressed, and inference procedures are discussed. In Section 5, a simulation study is conducted in order to evaluate the performance of the presented methodology. In Section 6, applications of the proposed models to real datasets are considered to illustrate its usefulness. Concluding remarks are addressed in Section 7.

2. Discretization Methods

In this section, we present two discretization methods that will be considered to obtain discrete analogs for the Shanker distribution. It is important to point out that the paper of Chakraborty (2015*a*) is possibly the only paper with exhaustive discussion on various methods of discretization.

2.1. Discretization by Infinite Series

The method of discretization by infinite series was firstly considered by Good (1953), which has proposed the discrete Good distribution to model population frequencies of species. A random variable Y is said to have a discrete Good distribution if its pmf can be written as

$$\mathrm{P}\left(Y = y;\alpha,\delta\right) = \frac{\delta^y y^\alpha}{\sum_{j=1}^{\infty}\delta^j j^\alpha}, \qquad y \in \mathbb{Z}_+,$$

for $\alpha \in \mathbb{R}$ and $\delta \in (0,1)$. The method of infinite series is characterized by the following definition.

Let X be a continuous random variable. If X has probability density function (pdf) $f_x\left(x;\boldsymbol{\theta}\right)$ with support on $\mathbb{R}$, then the corresponding discrete random variable Y has pmf given by

$$\mathrm{P}\left(Y=y;\boldsymbol{\theta}\right)=\frac{f_{X}\left(y;\boldsymbol{\theta}\right)}{\sum_{j=-\infty}^{\infty}f_{X}\left(j;\boldsymbol{\theta}\right)},\qquad y\in\mathbb{Z},$$

where $\boldsymbol{\theta}$ is the vector of parameters indexing the distribution of X.

This method was studied by several authors, including Kulasekera & Tonkyn (1992), Doray & Luong (1997), Kemp (1997) and Sato et al. (1999), which proposed a version of the method when the continuous random variable of interest is defined on $\mathbb{R}_{+}$. Thus, if the random variable X is defined on $\mathbb{R}_{+}$, the pmf of Y becomes

$$\mathrm{P}\left(Y=y;\boldsymbol{\theta}\right)=\frac{f_{X}\left(y;\boldsymbol{\theta}\right)}{\sum_{j=0}^{\infty}f_{X}\left(j;\boldsymbol{\theta}\right)},\qquad y\in\mathbb{Z}_{+}. \tag{3}$$

One of the most recent examples of the use of this method is that provided is the discrete analogue of the generalized Exponential distribution introduced by Nekoukhou et al. (2012) having pmf

$$\mathrm{P}\left(Y=y;\alpha,\lambda\right)=\lambda^{x-1}\left(1-\lambda^{x}\right)^{\alpha-1}\left[\sum_{i=1}^{\infty}\binom{\alpha-1}{j}\frac{\left(-1\right)^{j}\lambda^{j}}{1-\lambda^{1+j}}\right]^{-1},\qquad y\in\mathbb{Z}_{+},$$

for $\alpha\in\mathbb{R}_{+}$ and $\lambda\in(0,1)$.

A possible drawback of such method is the fact that, in some instances, the generated pmf may have no closed form, which is the case of the generalized Exponential model. However, it will be shown that this is not the case when obtaining the discrete analog for the Shanker distribution by this method.

2.2. Discretization by Survival Function

The method of discretization by sf was proposed by Nakagawa & Osaki (1975). This method allows us to discretize a continuous random variable from its sf. Several properties of the survival and of the risk functions were studied by Bracquemond & Gaudoin (2003), Roy (2003), Kemp (2004), Chakraborty (2015*a*), among others. The most important feature of this method is that it preserves the original sf on its integer part for the generated pmf (Chakraborty 2015*a*). Some other contributions in this area are given by Chakraborty & Chakravarty (2016), Chakraborty (2015*b*), Chakraborty & Gupta (2015) and Chakraborty & Chakravarty (2012). According to Roy (2003), we can define a discrete random variable from a continuous one as follows.

Let X be a continuous random variable. If X has sf $S_{X}\left(x;\boldsymbol{\theta}\right)$, then the discrete random variable $Y=\lfloor X\rfloor$ has pmf given by

$$\mathrm{P}\left(Y=y;\boldsymbol{\theta}\right)=S_{X}\left(y;\boldsymbol{\theta}\right)-S_{X}\left(y+1;\boldsymbol{\theta}\right),\qquad y\in\mathbb{Z}_{+}, \tag{4}$$

where $\lfloor\cdot\rfloor$ denotes the floor function, which returns the highest integer value smaller or equal to its argument.

It is noteworthy to mention that if the original sf has closed form, then the generated pmf will also have. For example, the Weibull distribution with pdf

$$f_X\left(x;\mu,\theta\right)=\frac{\theta}{\mu^{\theta}}\,x^{\theta-1}e^{-\left(\frac{x}{\mu}\right)^{\theta}},\qquad x\in\mathbb{R}_+,$$

and sf

$$S_X\left(x;\mu,\theta\right)=e^{-t(\frac{x}{\mu})^{\theta}},\qquad x\in\mathbb{R}_+,$$

where $\theta,\mu\in\mathbb{R}_+$ are, respectively, the shape and the scale parameters, was one of the first discretized distributions by this method. Nakagawa & Osaki (1975) proposed the discrete Weibull distribution which pmf for the random variable $Y=\lfloor X\rfloor$ is given by

$$\mathrm{P}\left(Y=y;\mu,\theta\right)=e^{-(\frac{y}{\mu})^{\theta}}-e^{-(\frac{y+1}{\mu})^{\theta}},\qquad y\in\mathbb{Z}_+,$$

for $(\theta,\mu)\in\mathbb{R}_+^2$. It is straightforward to prove that the above equation correspond to a proper pmf since it involves simple exponential terms.

3. The Discrete Shanker Distribution

In this section, we will consider both methods previously presented to define discrete analogs for the Shanker distribution. For ease of notation, each probabilistic model provided by these methods will be denoted by T1DS (Type I Discrete Shanker) and T2DS (Type II Discrete Shanker) distributions, respectively. For each version of this model, the main statistical properties as the shape, the generating functions and the central moments will be discussed.

3.1. Type I. Discrete Shanker Distribution

By considering equation (3), one can define the one-parameter T1DS distribution. We have the following definition.

Let X be a continuous random variable having Shanker distribution (1) with parameter $\theta\in\mathbb{R}_+$. Let $h\left(z\right)=e^z-1$, $z\in\mathbb{R}$. The pmf of Y having T1DS distribution is given by

$$\mathrm{P}\left(Y=y;\theta\right)=\frac{h^2\left(\theta\right)}{\theta h\left(\theta\right)+1}\left(\theta+y\right)e^{-\theta(y+1)},\qquad y\in\mathbb{Z}_+,\tag{5}$$

for $\theta\in\mathbb{R}_+$.

Proposition 1. *The equation* (5) *is a proper pmf.*

Proof. Here we have to prove that $\sum_{y=0}^{\infty} \mathrm{P}\left(Y=y;\theta\right)=1$ for $\theta \in \mathbb{R}_{+}$. Then,

$$\begin{aligned}\sum_{y=0}^{\infty} \mathrm{P}\left(Y=y;\theta\right) &= \sum_{y=0}^{\infty} \frac{h^{2}\left(\theta\right)}{\theta h\left(\theta\right)+1}\left(\theta+y\right)e^{-\theta(y+1)} \\ &= \frac{h^{2}\left(\theta\right)e^{-\theta}}{\theta h\left(\theta\right)+1}\left\{\theta\sum_{y=0}^{\infty}e^{-\theta y}+\sum_{y=0}^{\infty}ye^{-\theta y}\right\} \\ &= \frac{h^{2}\left(\theta\right)e^{-\theta}}{\theta h\left(\theta\right)+1}\left\{\frac{\theta e^{\theta}}{h\left(\theta\right)}+\frac{e^{\theta}}{h^{2}\left(\theta\right)}\right\} \\ &= 1,\end{aligned}$$

which concludes the proof. □

For a random variable Y behaving accordingly a T1DS distribution, we will adopt the notation $Y \sim \mathrm{T1DS}\left(\theta\right)$. The pmf (5) does not involves complicated expressions and therefore, the probabilities can be straightforwardly computed, as for example

$$\mathrm{P}\left(Y=0;\theta\right)=\frac{\theta h^{2}\left(\theta\right)e^{-\theta}}{\theta h\left(\theta\right)+1},$$

for $\theta \in \mathbb{R}_{+}$. Figure 1 depicts the behavior of the pmf (5) for selected values of θ.

We have derived some theoretical properties of the T1DS distribution. These properties are stated in the following propositions.

Proposition 2. *Let $Y \sim$ T1DS(θ). The sf of Y is given by*

$$S\left(y;\theta\right)=\frac{e^{-\theta y}\left[1-\left(\theta+y\right)h\left(-\theta\right)\right]}{\theta h\left(\theta\right)+1}, \qquad y \in \mathbb{Z}_{+},$$

for $\theta \in \mathbb{R}_{+}$.

Proof. By definition, $S\left(k;\theta\right)=\mathrm{P}\left(Y>k;\theta\right)=1-\mathrm{P}\left(Y\leqslant k;\theta\right)$. Then,

$$\begin{aligned}S\left(k;\theta\right) &= 1-\sum_{y=0}^{k}\frac{h^{2}\left(\theta\right)}{\theta h\left(\theta\right)+1}\left(\theta+y\right)e^{-\theta(y+1)} \\ &= 1-\frac{h^{2}\left(\theta\right)e^{-\theta}}{\theta h\left(\theta\right)+1}\left\{\theta\sum_{y=0}^{k}e^{-\theta y}+\sum_{y=0}^{k}ye^{-\theta y}\right\} \\ &= 1-\frac{h^{2}\left(\theta\right)e^{-\theta}}{\theta h\left(\theta\right)+1}\left\{\frac{\theta\left(e^{\theta}-e^{-\theta k}\right)}{h\left(\theta\right)}-\frac{ke^{-\theta k}}{h\left(\theta\right)}+\frac{e^{\theta}\left(1-e^{-\theta k}\right)}{h^{2}\left(\theta\right)}\right\} \\ &= \frac{e^{-\theta y}\left[1-\left(\theta+y\right)h\left(-\theta\right)\right]}{\theta h\left(\theta\right)+1},\end{aligned}$$

which concludes the proof. □

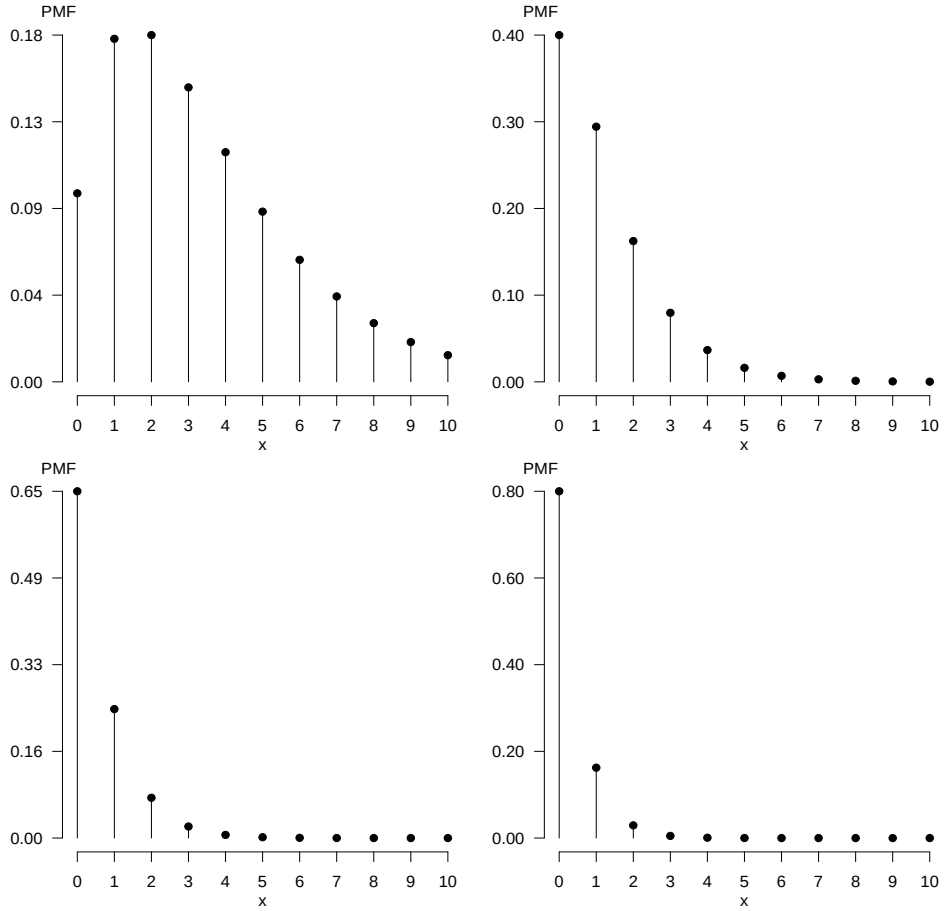

FIGURE 1: Behavior of the pmf of the T1DS distribution (upper-left-panel: $\theta = 0.5$, upper-right-panel: $\theta = 1.0$, lower-left-panel: $\theta = 1.5$ and lower-right-panel: $\theta = 2.0$).

Proposition 3. *Let $Y \sim T1DS(\theta)$. The probability generating function (pgf) of Y is given by*

$$G(s) = \frac{h^2(\theta)\left[\theta\left(e^{\theta} - s\right) + s\right]}{\left[\theta h(\theta) + 1\right]\left(e^{\theta} - s\right)^2},$$

for $s \neq e^{\theta}$.

Proof. By definition, $G(s) = \mathbb{E}\left(s^Y\right)$. For $s \neq e^{\theta}$, we have that

$$\begin{aligned} G(s) &= \sum_{y=0}^{\infty} s^y \frac{h^2(\theta)}{\theta h(\theta)+1}(\theta+y)e^{-\theta(y+1)} \\ &= \frac{h^2(\theta)e^{-\theta}}{\theta h(\theta)+1}\left\{\theta\sum_{y=0}^{\infty} s^y e^{-\theta y} + \sum_{y=0}^{\infty} y s^y e^{-\theta y}\right\} \\ &= \frac{h^2(\theta)e^{-\theta}}{\theta h(\theta)+1}\left\{\frac{\theta}{1-se^{-\theta}} + \frac{se^{-\theta}}{\left(1-se^{-\theta}\right)^2}\right\} \\ &= \frac{h^2(\theta)\left[\theta\left(e^{\theta}-s\right)+s\right]}{\left[\theta h(\theta)+1\right]\left(e^{\theta}-s\right)^2}, \end{aligned}$$

which concludes the proof. □

Proposition 4. *Let $Y \sim$ T1DS(θ). The moment generating function (mgf) of Y is given by*

$$M(t) = \frac{h^2(\theta)\left[\theta\left(e^{\theta}-e^{t}\right)+e^{t}\right]}{\left[\theta h(\theta)+1\right]\left(e^{\theta}-e^{t}\right)^2}, \tag{6}$$

for $t \neq \theta$.

Proof. By definition, $M(t) = \mathbb{E}\left(e^{tY}\right)$. For $t \neq \theta$, we have that

$$\begin{aligned} M(t) &= \sum_{y=0}^{\infty} e^{ty} \frac{h^2(\theta)}{\theta h(\theta)+1}(\theta+y)e^{-\theta(y+1)} \\ &= \frac{h^2(\theta)e^{-\theta}}{\theta h(\theta)+1}\left\{\theta\sum_{y=0}^{\infty} e^{-y(\theta-t)} + \sum_{y=0}^{\infty} y e^{-y(\theta-t)}\right\} \\ &= \frac{h^2(\theta)e^{-\theta}}{\theta h(\theta)+1}\left\{\frac{\theta}{1-e^{-(\theta-t)}} + \frac{e^{-(\theta-t)}}{\left(1-e^{-(\theta-t)}\right)^2}\right\} \\ &= \frac{h^2(\theta)\left[\theta\left(e^{\theta}-e^{t}\right)+e^{t}\right]}{\left[\theta h(\theta)+1\right]\left(e^{\theta}-e^{t}\right)^2}, \end{aligned}$$

which concludes the proof. □

Proposition 5. *Let $Y \sim$ T1DS(θ). The cumulant generating function (cgf) of Y is given by*

$$C(t) = 2\log\left[h(\theta)\right] + \log\left[\theta\left(e^{\theta}-e^{t}\right)+e^{t}\right] - \log\left[\theta h(\theta)+1\right] - 2\log\left(e^{\theta}-e^{t}\right),$$

for $t \neq \theta$.

Proof. Straightforward. Since $C(t) = \log\left[M(t)\right]$, the result follows. □

Proposition 6. *Let $Y \sim T1DS(\theta)$. The characteristic function (cf) of Y is given by*

$$\phi(t) = \frac{h^2(\theta)\left[\theta\left(e^{\theta} - e^{it}\right) + e^{it}\right]}{\left[\theta h(\theta) + 1\right]\left(e^{\theta} - e^{it}\right)^2},$$

for $t \in \mathbb{R}$ and $i = \sqrt{-1}$ is the imaginary number.

Proof. Straightforward. The result is obtained by noticing that $\phi(t) = M(it)$. □

It can be easily noticed that equation (6) is infinitely differentiable on t, since it involves exponential terms of its argument. Thus, from Proposition 4, the ordinary moments of Y can be derived by

$$\mu'_r = \frac{h^2(\theta)}{\left[\theta h(\theta) + 1\right]}\left\{\theta\frac{\mathrm{d}^r}{\mathrm{d}t^r}\frac{1}{(e^{\theta} - e^{t})} + \frac{\mathrm{d}^{r+1}}{\mathrm{d}t^{r+1}}\frac{1}{(e^{\theta} - e^{t})}\right\}_{t=0}, \qquad r \geqslant 1,$$

for $\theta \in \mathbb{R}_+$. Hence, the mean (μ) and the variance (σ^2) of Y are given, respectively, by

$$\mu = \mu'_1 = \frac{\theta h(\theta) + \left(e^{\theta} + 1\right)}{\left[\theta h(\theta) + 1\right] h(\theta)},$$

and

$$\sigma^2 = \mu'_2 - \left(\mu'_1\right)^2 = \frac{\left[\theta^2 h^2(\theta) + \left(e^{\theta} + 3\right)\theta h(\theta) + 2\right] e^{\theta}}{\left[\theta h(\theta) + 1\right]^2 h^2(\theta)}.$$

A normalized measure of dispersion can be obtained by using the variance-to-mean relationship. This measure is the well-known index of dispersion (ID) which, in this case, is given by

$$\mathrm{ID} = \frac{\sigma^2}{\mu} = \frac{\left[\theta^2 h^2(\theta) + \left(e^{\theta} + 3\right)\theta h(\theta) + 2\right] e^{\theta}}{\left[\theta h(\theta) + 1\right]\left[\theta h(\theta) + \left(e^{\theta} + 1\right)\right] h(\theta)}. \tag{7}$$

Analogously, the coefficient of variation (CV) of Y has the form

$$\mathrm{CV} = \frac{\sigma}{\mu} = \frac{e^{\theta/2}\sqrt{\left[\theta^2 h^2(\theta) + \left(e^{\theta} + 3\right)\theta h(\theta) + 2\right]}}{\theta h(\theta) + \left(e^{\theta} + 1\right)}.$$

Another useful measure is the zero-modification (ZM) index

$$\mathrm{ZM} = 1 + \mu^{-1}\log\left[\mathrm{P}(Y = 0)\right],$$

which is defined based on the Poisson distribution. This index can be easily interpreted since $\mathrm{ZM} > 0$ indicates zero-inflation, $\mathrm{ZM} < 0$ indicates zero-deflation and $\mathrm{ZM} = 0$ indicates no zero-modification. For the T1DS distribution, we have that the ZM index is given by

$$\mathrm{ZM} = 1 + \frac{\left[\log(\theta) + 2\log\left[h(\theta)\right] - \log\left[\theta h(\theta) + 1\right] - \theta\right]\left[\theta h(\theta) + 1\right] h(\theta)}{\theta h(\theta) + \left(e^{\theta} + 1\right)}. \tag{8}$$

The asymmetry degree and the flatness of a probabilistic model are usually measured by its coefficients of skewness and kurtosis, respectively. The first one can be computed by the third central moment, normalized by the variance raised to the power $3/2$ and the latter is given by the fourth central moment divided by the square of the variance. These coefficients are essential to characterize the shape of any distribution but, for the T1DS model, extensive and very complicated expressions were obtained for such measures. For this reason, the expressions of these coefficients are omitted here. However, Table 1 summarizes, for selected values of θ, the nature and the behavior of these coefficients along with the measures previously stated in the propositions.

Table 1: Theoretical descriptive statistics under T1DS distribution.

θ	Measures						
	Mean	Variance	ID	CV	ZM	Skewness	Kurtosis
0.5	3.4604	8.0707	2.3322	0.8208	0.3239	1.3979	5.9460
1.0	1.1639	1.8413	1.5820	1.1658	0.2119	1.5946	6.5432
1.5	0.4938	0.6527	1.3209	1.6351	0.1344	2.0171	8.2771
2.0	0.2405	0.2842	1.1819	2.2170	0.0819	2.5495	10.9373
2.5	0.1271	0.1404	1.1048	2.9487	0.0492	3.2249	15.0018
3.0	0.0703	0.0748	1.0616	3.8808	0.0295	4.1028	21.5466

When assessing equation (8) more deeply, we have obtained that ZM $\to 0$ as $\theta \to \infty$ and ZM $\to 1$ as $\theta \to 0$. This implies that, besides the usual case (ZM $= 0$), the T1DS distribution is suitable to deal with zero-inflation, but is not indicated to modeling zero-deflated datasets. Further, the coefficient of variation, the coefficient of skewness, and the coefficient of kurtosis are increasing as θ increases. On the other hand, the larger values for the mean, variance and index of dispersion are obtained for small values of θ. The T1DS distribution may be considered an interesting alternative to model overdispersed datasets ($\sigma^2 > \mu$). This can be seen by noticing that the index of dispersion is a decreasing function on θ and also ID $\to 1$ as $\theta \to \infty$ and ID $\to \infty$ as $\theta \to 0^+$. In fact, $\lim_{\theta \to 0}$ ID is undefined since the lateral limits are not equal, but the results presented in Table 1 allow us to conclude that ID increases as θ approaches zero. Therefore, as ID is always greater than 1, we conclude that $\sigma^2 > \mu$ for all $\theta \in \mathbb{R}_+$.

Proposition 7. *Let $Y \sim$ T1DS(θ). The mode (y_0) of Y is given by*

$$y_0 = \begin{cases} 0, & \text{if } \ r(\theta) < 0 \\ \lceil r(\theta) \rceil, & \text{if } \ r(\theta) > 0, \end{cases} \tag{9}$$

where $r(\theta) = h^{-1}(\theta) - \theta$ is a real-valued threshold and $\lceil \cdot \rceil$ is the ceiling function, which returns the lowest integer value greater or equal to its argument. If $r(\theta) = k$, $k \in \mathbb{Z}_+$, then the T1DS distribution is bimodal with modes k and $k+1$.

Proof. The ratio of consecutive probabilities is given by

$$\frac{\mathrm{P}(Y = y+1; \theta)}{\mathrm{P}(Y = y; \theta)} = \left[1 + \frac{1}{(\theta + y)}\right] e^{-\theta}, \qquad y \in \mathbb{Z}_+, \tag{10}$$

for $\theta \in \mathbb{R}_+$. From (10), it is clear that $\mathrm{P}(Y = y+1;\theta) = \mathrm{P}(Y = y;\theta)$ if $y = r(\theta) = h^{-1}(\theta) - \theta$. More generally, we have the following relations

i) $\mathrm{P}(Y = y+1;\theta) < \mathrm{P}(Y = y;\theta)$ if $y > r(\theta)$;

ii) $\mathrm{P}(Y = y+1;\theta) = \mathrm{P}(Y = y;\theta)$ if $y = r(\theta)$;

iii) $\mathrm{P}(Y = y+1;\theta) > \mathrm{P}(Y = y;\theta)$ if $y < r(\theta)$.

Now, let $k \in \mathbb{Z}_+$. By (i), if $r(\theta) < 0$ then $y_0 = 0$ since $y \in \mathbb{Z}_+$. If $r(\theta) > 0$ and $r(\theta) \neq k$ then $\mathrm{P}(Y = r(\theta);\theta) = 0$ and therefore, by (i) and (iii), $y_0 = \lceil r(\theta) \rceil$, that is, $\mathrm{P}(Y = \lceil r(\theta) \rceil;\theta) > \mathrm{P}(Y = k;\theta)$ for all $k \neq \lceil r(\theta) \rceil$. Finally, if $r(\theta) > 0$ and $r(\theta) = k$ then

$$\mathrm{P}(Y = k-1;\theta) \overset{\text{(iii)}}{<} \mathrm{P}(Y = k;\theta) \overset{\text{(ii)}}{=} \mathrm{P}(Y = k+1;\theta) \overset{\text{(i)}}{>} \mathrm{P}(Y = k+2;\theta),$$

and therefore, both k and $k+1$ are modes, implying bimodality. This concludes the proof. □

Proposition 8. *The T1DS distribution has an increasing hazard rate.*

Proof. Since (10) is a decreasing function on y, $\mathrm{P}(Y = k;\theta)$ is log-concave and therefore, the T1DS distribution has an increasing hazard rate. Hence, the proof. □

For instance, if $\theta = 0.5$ then $r(\theta) \approx 1.04$ and hence, $y_0 = 2$, as can be seen in the upper-left-panel of the Figure 1. In addition, it can also be proved that equation (5) satisfies $\mathrm{P}^2(Y = y;\theta) \geqslant \mathrm{P}(Y = y-1;\theta)\,\mathrm{P}(Y = y+1;\theta)$ for $r(\theta) \neq k$, which implies unimodality (see Theorem 3 by Keilson & Gerber (1971)). The relationship between log-concavity, unimodality and increasing hazard rate of discrete distributions has been discussed by Grandell (1997).

Proposition 9. *The T1DS distribution has heavy tails as θ approaches zero.*

Proof. The heavy-tail (HT) index is defined by the ratio (10) when $y \to \infty$. A discrete distribution is said to have heavy tails if $\mathrm{HT} \to 1$ when $y \to \infty$. Hence,

$$\lim_{\theta\to 0} \mathrm{HT} = \lim_{\theta\to 0} \left\{ e^{-\theta} \lim_{y\to\infty} \left[1 + \frac{1}{(\theta + y)} \right] \right\} = \lim_{\theta\to 0} e^{-\theta} = 1,$$

which concludes the proof. □

3.2. Type II. Discrete Shanker Distribution

By considering equations (2) and (4), one can define the one-parameter T2DS distribution. We have the following definition.

Let X be a continuous random variable having Shanker distribution (1) with parameter $\theta \in \mathbb{R}_+$. Let $h(z_1, z_2) = 1 + z_1(z_1 + z_2)$, $z_1 \in \mathbb{R}_+$ and $z_2 \in \mathbb{Z}_+$. The pmf of $Y = \lfloor X \rfloor$ having T2DS distribution is given by

$$\mathrm{P}(Y = y; \theta) = \frac{e^{-\theta y}}{\theta^2 + 1}\left[h(\theta, y) - h(\theta, y+1)\, e^{-\theta}\right], \qquad y \in \mathbb{Z}_+, \tag{11}$$

for $\theta \in \mathbb{R}_+$.

Proposition 10. *The equation* (11) *is a proper pmf.*

Proof. The result comes analogous to the proof of Proposition 1. □

For a random variable Y distributed accordingly a T2DS distribution, we will adopt the notation $X \sim \mathrm{T2DS}(\theta)$. For this version, the probabilities can be easily computed as noticed for the T1DS distribution. Then,

$$\mathrm{P}(Y = 0; \theta) = \frac{1}{\theta^2 + 1}\left[h(\theta, 0) - h(\theta, 1)\, e^{-\theta}\right],$$

for $\theta \in \mathbb{R}_+$. Figure 2 depicts the behavior of the pmf (11) of Y, using selected values for θ.

We have also derived some theoretical properties of the T2DS distribution. These properties are stated in the following propositions.

Proposition 11. *Let $Y \sim T2DS(\theta)$. The sf of Y is given by*

$$S(y; \theta) = \frac{h(\theta, y+1)\, e^{-\theta(y+1)}}{\theta^2 + 1}, \qquad y \in \mathbb{Z}_+,$$

for $\theta \in \mathbb{R}_+$.

Proof. The result comes analogous to the proof of Proposition 2. □

Proposition 12. *Let $Y \sim T2DS(\theta)$. The pgf of Y is given by*

$$G(s) = \frac{\left(\theta^2 + 1\right)\left(e^{2\theta} + s\right) - \left[\left(\theta^2 + 1\right)(s+1) - \theta(s-1)\right] e^{\theta}}{\left(\theta^2 + 1\right)\left(s - e^{\theta}\right)^2},$$

for $s \neq e^{\theta}$.

Proof. The result comes analogous to the proof of Proposition 3. □

Proposition 13. *Let $Y \sim T2DS(\theta)$. The mgf of Y is given by*

$$M(t) = \frac{\left(\theta^2 + 1\right)\left[\left(e^{\theta} - e^{t} - 1\right) e^{\theta} + e^{t}\right] + \theta e^{\theta}\left(e^{t} - 1\right)}{\left(\theta^2 + 1\right)\left(e^{t} - e^{\theta}\right)^2}, \tag{12}$$

for $t \neq \theta$.

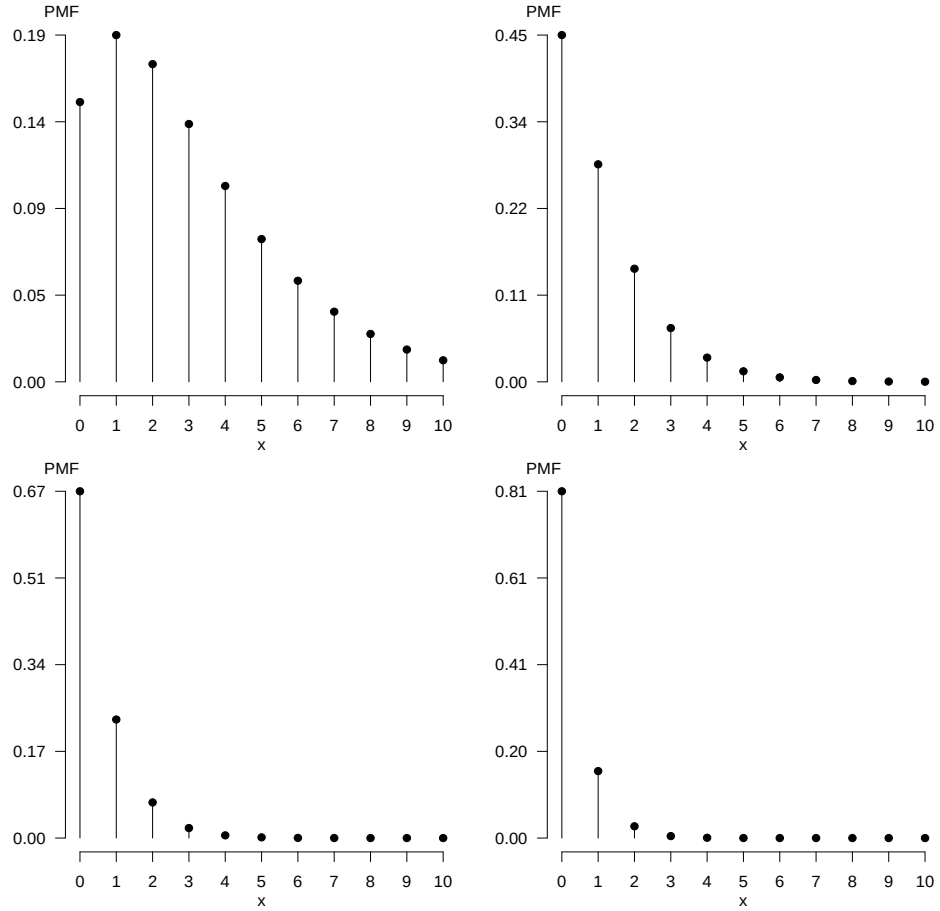

FIGURE 2: Behavior of the pmf of the T2DS distribution (upper-left-panel: $\theta = 0.5$, upper-right-panel: $\theta = 1.0$, lower-left-panel: $\theta = 1.5$ and lower-right-panel: $\theta = 2.0$).

Proof. The result comes analogous to the proof of Proposition 4. □

Proposition 14. *Let $Y \sim T2DS(\theta)$. The cgf of Y is given by*

$$C(t) = \log\left\{\left(\theta^2+1\right)\left[\left(e^{\theta}-e^{t}-1\right)e^{\theta}+e^{t}\right]+\theta e^{\theta}\left(e^{t}-1\right)\right\} - \log\left(\theta^2+1\right)-2\log\left(e^{t}-e^{\theta}\right),$$

for $t \neq \theta$.

Proof. The result comes analogous to the proof of Proposition 5. □

Proposition 15. *Let $Y \sim T2DS(\theta)$. The cf of Y is given by*

$$\phi(t) = \frac{\left(\theta^2+1\right)\left[\left(e^{\theta}-e^{it}-1\right)e^{\theta}+e^{it}\right]+\theta e^{\theta}\left(e^{it}-1\right)}{\left(\theta^2+1\right)\left(e^{it}-e^{\theta}\right)^2},$$

for $t \in \mathbb{R}$ and $i = \sqrt{-1}$ is the imaginary number.

Proof. The result comes analogous to the proof of Proposition 6. □

For this version, it is also clear that equation (12) is infinitely differentiable on t. Therefore, from Proposition 13, the ordinary moments of Y can be derived by

$$\mu'_r = \left\{ e^{\theta} \left[\frac{\mathrm{d}^r}{\mathrm{d}t^r} \frac{\left(e^{\theta} - e^t - 1\right)}{\left(e^t - e^{\theta}\right)^2} + \frac{\theta}{\theta^2 + 1} \frac{\mathrm{d}^r}{\mathrm{d}t^r} \frac{\left(e^t - 1\right)}{\left(e^t - e^{\theta}\right)^2} \right] + \frac{\mathrm{d}^r}{\mathrm{d}t^r} \frac{e^t}{\left(e^t - e^{\theta}\right)^2} \right\}_{t=0}, \qquad r \geqslant 1,$$

for $\theta \in \mathbb{R}_+$. Hence, the mean and the variance of Y are given, respectively, by

$$\mu = \frac{\left(\theta^2 + 1\right)\left(e^{\theta} - 1\right) + \theta e^{\theta}}{\left(\theta^2 + 1\right)\left(e^{\theta} - 1\right)^2},$$

and

$$\sigma^2 = \frac{\left(e^{\theta} - 1\right)\theta^4 + \left(e^{2\theta} - 1\right)\theta^3 + \left[2\left(e^{2\theta} + 1\right) - 5e^{\theta}\right]\theta^2}{(\theta^2 + 1)^2(e^{\theta} - 1)^4} + \frac{\left(e^{2\theta} - 1\right)\theta + \left(e^{\theta} - 2\right)e^{\theta} + 1}{(\theta^2 + 1)^2(e^{\theta} - 1)^4}.$$

Now, the ID of Y is given by

$$\mathrm{ID} = \frac{\left(e^{\theta} - 1\right)\theta^4 + \left(e^{2\theta} - 1\right)\theta^3 + \left[2\left(e^{2\theta} + 1\right) - 5e^{\theta}\right]\theta^2}{\left[\left(\theta^2 + 1\right)\left(e^{\theta} - 1\right) + \theta e^{\theta}\right]\left(\theta^2 + 1\right)\left(e^{\theta} - 1\right)^2} + \frac{\left(e^{2\theta} - 1\right)\theta + \left(e^{\theta} - 2\right)e^{\theta} + 1}{\left[\left(\theta^2 + 1\right)\left(e^{\theta} - 1\right) + \theta e^{\theta}\right]\left(\theta^2 + 1\right)\left(e^{\theta} - 1\right)^2},$$

and the CV has the form

$$\mathrm{CV} = \frac{\sqrt{g\left(\theta\right)}}{\left(\theta^2 + 1\right)\left(e^{\theta} - 1\right) + \theta e^{\theta}},$$

where

$$g\left(\theta\right) = \left(e^{\theta} - 1\right)\theta^4 + \left(e^{2\theta} - 1\right)\theta^3 + \left[2\left(e^{2\theta} + 1\right) - 5e^{\theta}\right]\theta^2 + \left(e^{2\theta} - 1\right)\theta + \left(e^{\theta} - 2\right)e^{\theta} + 1.$$

As for the T1DS distribution, the coefficients of skewness and kurtosis of the T2DS distribution are represented by an extensive and very complicated expression. These expressions will also be omitted, but Table 2 summarizes, for selected values of θ, the nature and the behavior of these coefficients along with the measures previously presented in this subsection.

Table 2: Theoretical descriptive statistics under T2DS distribution.

θ	Measures						
	Mean	Variance	ID	CV	ZM	Skewness	Kurtosis
0.5	3.1088	7.8607	2.5292	0.9009	0.3915	1.4522	6.1036
1.0	1.0423	1.7050	1.6356	1.2528	0.2300	1.7173	7.0468
1.5	0.4578	0.6090	1.3307	1.7047	0.1379	2.1104	8.7843
2.0	0.2290	0.2709	1.1830	2.2732	0.0824	2.6168	11.3721
2.5	0.1231	0.1361	1.1046	2.9967	0.0492	3.2770	15.3967
3.0	0.0688	0.0727	1.0609	3.9229	0.0294	4.1470	21.9449

For this version, the ZM index is given by

$$\mathrm{ZM} = 1 + \frac{\left[\log\left[h\left(\theta,0\right) - h\left(\theta,1\right)e^{-\theta}\right] - \log\left(\theta^2+1\right)\right]\left(\theta^2+1\right)\left(e^{\theta}-1\right)^2}{\left(\theta^2+1\right)\left(e^{\theta}-1\right)+\theta e^{\theta}}. \quad (13)$$

The limit properties of (13) are equal to those obtained for (8), that is, $\mathrm{ZM} \to 0$ as $\theta \to \infty$ and $\mathrm{ZM} \to 1$ as $\theta \to 0$. This implies that the T2DS distribution is also suitable to deal with zero-inflation, but is not indicated to modeling zero-deflated datasets. On the other hand, since $\theta \in \mathbb{R}_+$, the central moments of the T2DS distribution present the same behavior concerning those derived from the T1DS distribution. Moreover, since equation (13) presents the same limit properties of equation (7), we conclude that the T2DS distribution may also be considered as an alternative to model overdispersed datasets.

Proposition 16. *Let $Y \sim T2DS(\theta)$. The mode (y_0) of Y is given by* (9)*, where*

$$r\left(\theta\right) = \frac{2}{\left(e^{\theta}-1\right)} - \frac{\theta^2+1}{\theta},$$

is a real-valued threshold. If $r\left(\theta\right) = k$, $k \in \mathbb{Z}_+$, then the T2DS distribution is bimodal with modes k and $k+1$.

Proof. The ratio of consecutive probabilities is given by

$$\frac{\mathrm{P}\left(Y=y+1;\theta\right)}{\mathrm{P}\left(Y=y;\theta\right)} = \frac{\left[h\left(\theta,y+1\right) - h\left(\theta,y+2\right)\right]e^{-\theta}}{\left[h\left(\theta,y\right) - h\left(\theta,y+1\right)\right]}, \quad (14)$$

for $\theta \in \mathbb{R}_+$. Thus, the result comes analogous to the proof of Proposition 7. □

Proposition 17. *The T2DS distribution has an increasing hazard rate.*

Proof. One can notice that equation (14) is also a decreasing function on y. In this case, if follows that $\mathrm{P}\left(Y=k;\theta\right)$ is log-concave and therefore, the T2DS distribution has an increasing hazard rate. Hence, the proof. □

For the T2DS distribution, it can also be proved that equation (11) satisfies $\mathrm{P}^2\left(Y=y;\theta\right) \geqslant \mathrm{P}\left(Y=y-1;\theta\right)\mathrm{P}\left(Y=y+1;\theta\right)$ for $r\left(\theta\right) \neq k$, which implies unimodality. In addition, one can notice that the form of the mode of the T2DS

distribution is exactly equal to the T1DS model but, in this case, if $\theta = 0.5$ then $r(\theta) \approx 0.58$ and hence, the mode is 1, as can be seen in the upper-left-panel of the Figure 2.

Proposition 18. *The T2DS distribution has heavy tails as θ approaches zero.*

Proof. By considering the HT index previously defined, we have that

$$\lim_{\theta\to 0} \mathrm{HT} = \lim_{\theta\to 0}\left\{e^{-\theta}\lim_{y\to\infty}\frac{[h(\theta,y+1)-h(\theta,y+2)]}{[h(\theta,y)-h(\theta,y+1)]}\right\} = \lim_{\theta\to 0} e^{-\theta} = 1,$$

which concludes the proof. □

4. Maximum Likelihood Estimation

In this section, we will address the issue of estimating the parameter θ of both versions of the discrete Shanker distribution. We have adopted the frequentist approach, and here we will derive the maximum likelihood function for the T1DS and T2DS models. Using these functions, one can obtain point estimates for parameter θ in each case. Moreover, suitable estimates for the confidence intervals can be obtained using large-sample approximations, that is based on the asymptotic properties of the maximum likelihood estimators.

4.1. Inference Under T1DS Distribution

Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ a random sample of size n from the T1DS distribution and $\boldsymbol{y} = (y_1, \ldots, y_n)$ its observed values. The log-likelihood function for parameter θ can be expressed as

$$\ell_n(\theta;\boldsymbol{y}) = -n\left\{\theta(\overline{y}+1) - 2\log[h(\theta)] + \log[\theta h(\theta)+1]\right\} + \sum_{i=1}^{n}\log(\theta+y_i), \quad (15)$$

where $\overline{y}$ is the sample mean. The maximum likelihood estimator (MLE) of θ can be obtained by direct maximization of the log-likelihood function. Thus, the first order derivative of (15) respect to θ (score function) is given by

$$U_n(\theta;\boldsymbol{y}) = \frac{\mathrm{d}}{\mathrm{d}\theta}\ell_n(\theta;\boldsymbol{y}) = -n(\overline{y}+1) + \frac{2ne^{\theta}}{h(\theta)} - \frac{n\left[\theta e^{\theta}+h(\theta)\right]}{\theta h(\theta)+1} + \sum_{i=1}^{n}\frac{1}{\theta+y_i}.$$

There is no closed-form solution for the MLE of θ and therefore, standard optimization algorithms such Newton-Raphson based methods may be used to obtain numerical estimates. By the maximum likelihood theory, a consistent estimator for the variance of $\widehat{\theta}$ is obtained by the inverse of the Fisher information, that is,

$$\begin{aligned} I_n(\theta) =& \mathbb{E}\left[-\frac{\mathrm{d}}{\mathrm{d}\theta}U_n(\theta;\mathbf{Y})\right] \\ =& n\left[\frac{2e^{\theta}}{h^2(\theta)} - \frac{1+e^{\theta}\left[\theta(\theta-1)+\left(e^{\theta}-4\right)\right]}{[\theta h(\theta)+1]^2} + \frac{h^2(\theta)e^{-\theta}}{\theta h(\theta)+1}\zeta\left(e^{-\theta},1,\theta\right)\right], \end{aligned}$$

where ζ is the Lerch-Phi function (Bateman & Erdélyi 1953) defined as $\zeta(z,a,v) = \sum_{j=0}^{\infty} z^j (j+v)^{-a}$ for $|z| < 1$.

Finally, in order to obtain intervallic estimates for θ, one can use large-sample approximations for the $100 \times (1-\alpha)\,\%$ two-sided confidence interval (CI), that is,

$$\widehat{\theta} \pm z_{_{1-\alpha/2}} \sqrt{I_n^{-1}(\hat{\theta})},$$

where $z_{_{1-\alpha/2}}$ is the upper $(\alpha/2)^{th}$ percentile of the standard Normal distribution.

4.2. Inference Under T2DS Distribution

Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ a random sample of size n from the T2DS distribution and $\boldsymbol{y} = (y_1, \ldots, y_n)$ its observed values. The log-likelihood function for parameter θ can be expressed as

$$\ell_n(\theta; \boldsymbol{y}) = -n\left[\theta\,\overline{y} + \log\left(\theta^2+1\right)\right] + \sum_{i=1}^{n} \log\left[h(\theta, y_i) - h(\theta, y_i+1)\, e^{-\theta}\right], \quad (16)$$

where $\overline{y}$ is the sample mean. The MLE of θ can be obtained by direct maximization of the log-likelihood function. The first order derivative of (16) respect to θ is given by

$$U_n(\theta; \boldsymbol{y}) = -\,n\left[\overline{y} + \frac{2\theta}{\theta^2+1}\right] + \sum_{i=1}^{n} \frac{(2\theta + y_i) + [\theta(\theta + y_i + 1) - (2\theta + y_i + 1) + 1]\, e^{-\theta}}{[(\theta + y_i)\,\theta + 1] - [(\theta + y_i + 1)\,\theta + 1]\, e^{-\theta}}.$$

Now, in order to estimate the variance of $\widehat{\theta}$, one have to obtain the Fisher information of θ. In this case, this quantity has the form

$$I_n(\theta) = \frac{n\theta^2\left[\theta\left(\theta^2+\theta+3\right)+1\right]^2\, {}_6\mathrm{F}_5\left(1, a_1, a_1, a_2, a_2, b_1; a_3, a_3, b_2, a_4, a_4; e^{-\theta}\right)}{\left(\theta^2+1\right)^3\left[\left(\theta^2+1\right)\left(e^{\theta}-1\right)-\theta\right]e^{\theta}},$$

where ${}_p\mathrm{F}_q$ is the generalized hypergeometric function (Slater 1966), whose arguments are given by

$$a_1 = \frac{[\theta(\theta+2\theta+3)+2]\,e^{\theta} - \left[\theta\left(\theta^2+4\theta+3\right)+4\right] - \sqrt{c_1}}{2\left(\theta^2+1\right)\left(e^{\theta}-1\right)},$$

$$a_2 = \frac{[\theta(\theta+2\theta+3)+2]\,e^{\theta} - \left[\theta\left(\theta^2+4\theta+3\right)+4\right] + \sqrt{c_1}}{2\left(\theta^2+1\right)\left(e^{\theta}-1\right)},$$

$$a_3 = \frac{\theta e^{\theta}\left(\theta^2+3\right) - \left[\theta\left(\theta^2+2\theta+3\right)+2\right] - \sqrt{c_1}}{2\left(\theta^2+1\right)\left(e^{\theta}-1\right)},$$

and

$$a_4 = \frac{\theta e^{\theta}\left(\theta^2+3\right) - \left[\theta\left(\theta^2+2\theta+3\right)+2\right] + \sqrt{c_1}}{2\left(\theta^2+1\right)\left(e^{\theta}-1\right)},$$

where $c_1 = \theta^2\left(\theta^2+3\right)^2\left(e^{2\theta}+1\right) - \left[\theta^2\left(2\theta^4-8\theta^2-10\right)+4\right]e^{\theta}$,

$$b_1 = \frac{\left(\theta^2+1\right)\left(e^{\theta}-1\right)-\theta}{\theta\left(e^{\theta}-1\right)} \quad \text{and} \quad b_2 = \frac{\left[\left(\theta^2+1\right)+\theta\right]e^{\theta}-(\theta+1)^2}{\theta\left(e^{\theta}-1\right)}.$$

Again, there is no closed-form solution for the MLE of θ. In this case, we can adopt the same procedure presented in the previous subsection to obtain point and intervallic estimates for parameter θ.

5. Simulation Study

In this section, we have estimated, using $B = 10,000$ Monte Carlo simulation, the biases, the mean squared error, the coverage probabilities and the coverage lengths of the MLE $\widehat{\theta}$ of both versions of the discrete Shanker distribution. To run the simulation, we have considered $\theta = 0.3, 0.6, 0.8, 1.0, 1.5, 1.8$ and 2.0 and sample sizes ranging from 20 to 200 by 30. The inverse-transform method for discrete distributions (Rubinstein & Kroese 2008) was implemented to generate the pseudo-random samples. The simulation process was performed using *Ox Console* (Doornik 2007). The quantities of interest were estimated by the following expressions.

- $BIAS(\widehat{\theta}) = \dfrac{1}{B}\sum_{i=1}^{B}(\widehat{\theta}_i - \theta)$.
- $MSE(\widehat{\theta}) = \dfrac{1}{B}\sum_{i=1}^{B}(\widehat{\theta}_i - \theta)^2$.
- $CL_{\theta}(n) = \dfrac{3.92}{B}\sum_{i=1}^{B} s_{\widehat{\theta}_i}$.
- $CP_{\theta}(n) = \dfrac{1}{B}\sum_{i=1}^{B}\mathcal{I}\{\widehat{\theta}_i - 1.96\widehat{s}_{\widehat{\theta}_i} < \theta < \widehat{\theta}_i + 1.96\widehat{s}_{\widehat{\theta}_i}\}$, where $\mathcal{I}\{\cdot\}$ denotes the indicator function.

For both versions, the behavior of the average bias and average mean squared error are shown in Figures 3 and 4. The results for the coverage probabilities and the coverage lengths are reported in Tables 3 and 4.

From Figures 3 and 4, in each scenario and for T1DS and T2DS distributions, we have that the bias of $\widehat{\theta}$ is positive and tends to zero when the sample size increases. Also, the mean squared error of $\widehat{\theta}$ tends to zero in each case. For the coverage probabilities, we have $CP_{\theta}(n)$ around 0.94 and 0.96 for both discretizations, and the coverage length tends to zero when the sample size increases. Table 5 reports the percentage of times, out of $10,000$ Monte Carlo simulations, that the Voung's test (Vuong 1989) judges that the generated data is coming from the same distribution.

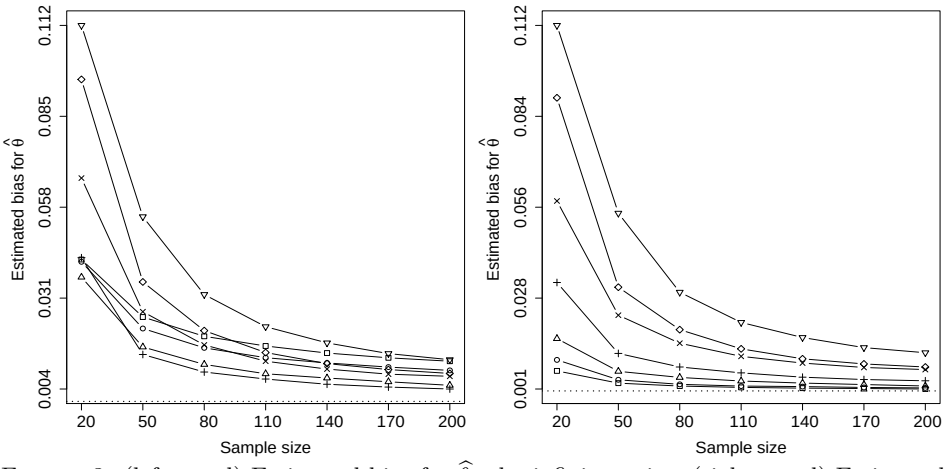

FIGURE 3: (left-panel) Estimated bias for $\widehat{\theta}$ – by infinite series. (right-panel) Estimated bias for $\widehat{\theta}$ – by survival function ($\square : \theta = 0.3$, $\bigcirc : \theta = 0.6$, $\triangle : \theta = 0.9$, $+ : \theta = 1.0$, $\times : \theta = 1.5$, $\diamond : \theta = 1.8$ and $\triangledown : 2.0$).

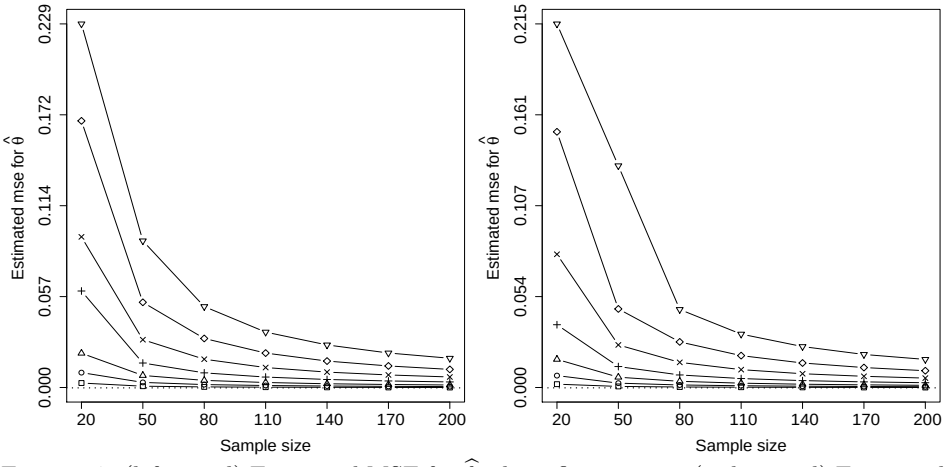

FIGURE 4: (left-panel) Estimated MSE for $\widehat{\theta}$ – by infinite series. (right-panel) Estimated MSE for $\widehat{\theta}$ – by survival function ($\square : \theta = 0.3$, $\bigcirc : \theta = 0.6$, $\triangle : \theta = 0.9$, $+ : \theta = 1.0$, $\times : \theta = 1.5$, $\diamond : \theta = 1.8$ and $\triangledown : 2.0$).

Using the same simulation scenarios as previously described, we have estimated the probability of correct selection (PCS) using the difference between the maximized log-likelihood functions as the discrimination criterion. Let ℓ_k be the log-likelihood function of the TkDS distribution. We choose T1DS or T2DS as the preferred model if the statistic $T_n = \ell_1(\hat{\theta}; \boldsymbol{y}) - \ell_2(\hat{\theta}; \boldsymbol{y})$ is greater than or less than zero, respectively. Estimates of the PCS's are shown in the Figure 5.

TABLE 3: Estimated coverage probability and length of coverage probability for $\widehat{\theta}$ (T1DS distribution).

Quantity	n	Values for θ						
		0.3	0.6	0.9	1.0	1.5	1.8	2.0
$CP_\theta(n)$	20	0.9528	0.9510	0.9522	0.9552	0.9628	0.9590	0.9599
	50	0.9476	0.9503	0.9496	0.9537	0.9533	0.9588	0.9571
	80	0.9516	0.9503	0.9515	0.9517	0.9517	0.9492	0.9588
	110	0.9502	0.9490	0.9518	0.9528	0.9506	0.9533	0.9586
	140	0.9494	0.9501	0.9500	0.9522	0.9490	0.9513	0.9547
	170	0.9508	0.9486	0.9503	0.9504	0.9505	0.9504	0.9559
	200	0.9500	0.9512	0.9497	0.9486	0.9507	0.9515	0.9502
$CL_\theta(n)$	20	0.1790	0.3264	0.4897	0.7087	0.9965	1.3622	1.7805
	50	0.1121	0.2041	0.3041	0.4333	0.6001	0.8011	1.1449
	80	0.0884	0.1610	0.2395	0.3403	0.4690	0.6233	0.8103
	110	0.0753	0.1371	0.2038	0.2894	0.3979	0.5279	0.6840
	140	0.0667	0.1215	0.1805	0.2560	0.3517	0.4662	0.6033
	170	0.0605	0.1102	0.1636	0.2320	0.3186	0.4222	0.5456
	200	0.0558	0.1016	0.1508	0.2138	0.2934	0.3887	0.5021

TABLE 4: Estimated coverage probability and length of coverage probability for $\widehat{\theta}$ (T2DS distribution).

Quantity	n	Values for θ						
		0.3	0.6	0.9	1.0	1.5	1.8	2.0
$CP_\theta(n)$	20	0.9685	0.9658	0.9615	0.9574	0.9661	0.9687	0.9667
	50	0.9614	0.9498	0.9483	0.9531	0.9568	0.9534	0.9558
	80	0.9584	0.9554	0.9531	0.9503	0.9536	0.9490	0.9529
	110	0.9541	0.9680	0.9572	0.9496	0.9494	0.9513	0.9563
	140	0.9573	0.9608	0.9631	0.9516	0.9497	0.9506	0.9578
	170	0.9567	0.9548	0.9587	0.9504	0.9506	0.9500	0.9556
	200	0.9542	0.9456	0.9480	0.9482	0.9517	0.9514	0.9542
$CL_\theta(n)$	20	0.2035	0.3714	0.5576	1.2078	1.0897	1.4578	1.8686
	50	0.1226	0.2272	0.3420	0.4805	0.6547	0.8588	1.1078
	80	0.0953	0.1779	0.2683	0.3770	0.5114	0.6678	0.8545
	110	0.0806	0.1510	0.2279	0.3205	0.4336	0.5654	0.7214
	140	0.0710	0.1335	0.2017	0.2835	0.3832	0.4993	0.6362
	170	0.0642	0.1209	0.1827	0.2570	0.3471	0.4522	0.5755
	200	0.0590	0.1113	0.1682	0.2367	0.3198	0.4163	0.5295

TABLE 5: Percentage of times out of 10,000 that the null hypothesis is not rejected.

θ	Data generated from T1DS					Data generated from T2DS				
	Sample size					Sample size				
	20	50	100	200	500	20	50	100	200	500
0.3	75.93	74.32	68.55	60.03	53.95	92.87	92.59	88.54	80.74	71.12
0.6	88.66	88.02	87.93	86.14	75.83	94.54	94.45	93.04	91.52	84.05
0.9	94.01	93.45	93.08	91.97	89.05	96.42	95.89	95.13	94.70	92.87
1.0	95.65	94.33	94.22	93.28	91.06	96.70	96.17	95.79	95.69	93.78
1.5	99.49	98.75	96.75	95.38	94.79	99.59	99.17	97.52	96.30	95.78
1.8	99.86	99.75	98.29	95.98	94.71	99.88	99.82	98.82	96.55	95.62
2.0	99.95	99.93	99.30	96.97	95.28	99.97	99.92	99.36	97.17	95.70

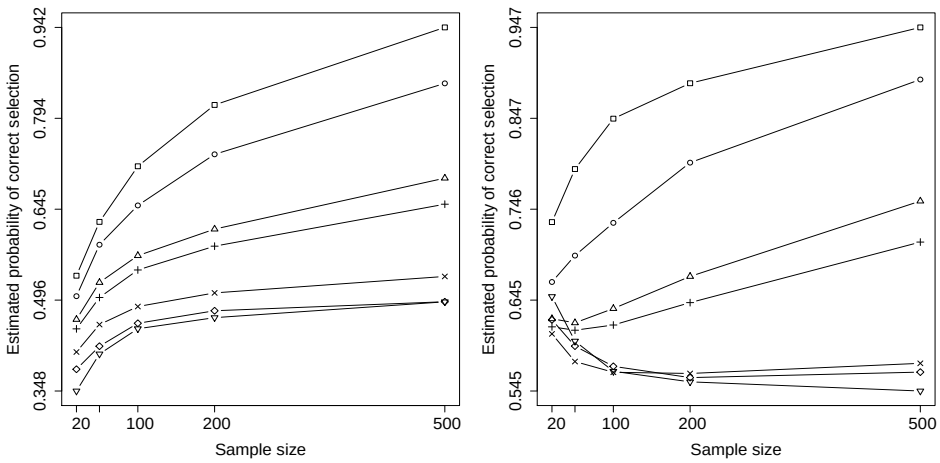

FIGURE 5: Left-panel: Estimated PCS when data are generated from T1DS distribution. Right-panel: Estimated PCS when data are generated from T2DS ($\square : \theta = 0.3$, $\bigcirc : \theta = 0.6$, $\triangle : \theta = 0.9$, $+ : \theta = 1.0$, $\times : \theta = 1.5$, $\diamond : \theta = 1.8$, and $\triangledown : \theta = 2.0$).

6. Application to real data

In this section, both versions of the discrete Shanker distribution are considered to model two real datasets from different areas. The goodness-of-fit of the proposed models is compared with those accessed by the Poisson (P) and Negative Binomial (NB) distributions. The parameterizations considered to fit the NB model is the same implemented in the `R` software. For the first application, we have considered the total number of borers per hill in each plot for a control group and three treatment groups. This dataset was firstly analyzed by Bliss & Fisher (1953). In a field experiment of insect pests on the corn borer, four treatments were arranged in 15 randomized blocks. At the end of the season, eight hills of corn were selected randomly in each plot, and the borers were recorded. Here, we are considering the data from the second treatment (Saha (2008), Table 9). The second one relates to the number of contract strikes in US manufacturing beginning each month between January 1968 and December 1976 (Kennan 1985). All computations to obtain the results presented in this section were performed using the `R` environment (R Development Core Team 2017). The executable *scripts* are available from the authors upon justified request.

Table 6 presents some descriptive statistics for each dataset. The raw data used in this section can be found in Appendix A. The initial analysis highlights the presence of overdispersion (see the index of dispersion), justifying the choice of the discrete Shanker distribution to describe such data. Moreover, the sample mode of the second dataset is greater than 0 and so, if one of the versions of our model fit these data, then we expect a value smaller than 0.50 for the MLE of θ in this case.

TABLE 6: Variables and descriptive statistics for each dataset.

Dataset	Variable	n	Mean	Median	Var.	ID (%)	CV (%)
1	Number of borers per hill	120	1.48	1.00	3.19	215.26	120.46
2	Number of contract strikes	108	5.24	5.00	14.07	268.52	71.58

In Table 7 we present the frequency distribution of each sample. The expected frequencies were obtained through the estimated probabilities, that were computed using the MLEs. Frequencies in bold relate to those one closer to the observed ones. The results show that the proposed models provide reliable fit in both cases.

TABLE 7: Observed and expected frequencies from the fitted models.

Counts	Observed	Expected			
		P	NB	T1DS	T2DS
Dataset 1					
0	43	27.23	**44.28**	39.17	41.26
1	35	40.38	31.08	**34.51**	32.67
2	17	29.95	**19.10**	21.84	20.77
3	11	14.81	**11.17**	12.15	11.95
4	5	5.49	6.38	**6.32**	6.50
$\geqslant 5$	9	2.13	**7.27**	5.72	6.45
Dataset 2					
0	5	0.57	**5.42**	4.29	7.26
1	12	3.00	10.11	11.43	**11.72**
2	14	7.86	12.69	**13.90**	13.10
3	11	13.72	13.94	13.86	**12.77**
4	9	17.98	12.66	12.60	**11.59**
5	14	18.84	**11.25**	10.84	10.07
6	9	16.46	9.53	**9.01**	8.48
$\geqslant 7$	34	**28.98**	28.64	27.34	27.55

TABLE 8: Parameter estimates and gof measures for the fitted models.

Model	Par.	MLE (SE)	95% CI		χ^2 (p-value)	d.f.
			Lower	Upper		
Dataset 1						
P	μ	1.483 (0.111)	1.265	1.701	38.59 (< 0.001)	4
NB	μ	1.483 (0.162)	1.167	1.800	1.47 (0.689)	3
	ϕ	1.333 (0.644)	0.601	2.065		
T1DS	θ	0.884 (0.048)	0.789	0.978	3.71 (0.446)	4
T2DS	θ	0.820 (0.050)	0.723	0.918	2.36 (0.671)	4
Dataset 2						
P	μ	5.241 (0.220)	4.809	5.672	76.64 (< 0.001)	6
NB	μ	5.241 (0.369)	4.517	5.964	3.69 (0.594)	5
	ϕ	2.897 (0.644)	1.634	4.159		
T1DS	θ	0.357 (0.022)	0.312	0.401	4.31 (0.635)	6
T2DS	θ	0.330 (0.022)	0.288	0.373	4.67 (0.586)	6

The MLEs, SEs, and 95% asymptotic CIs for the parameters of each fitted model are presented in Table 8. The goodness-of-fit was assessed using the χ^2 statistic. For Dataset 1, the chi-squared value for the T2DS distributions is $\chi^2 = 2.36$, with a corresponding p-value ≈ 0.68, highlighting the adherence of the T2DS distribution. Also, for Dataset 2, we have obtained $\chi^2 = 4.67$ (p-value ≈ 0.59) for the T1DS model. The goodness-of-fit accessed by T1DS distribution was found to be quite similar to T2DS model. Model selection was performed using the Akaike information criterion with correction for finite samples (AICc), the Bayesian information criterion (BIC), and the Hannan-Quinn information criterion (HQC). These measures are presented in Table 9. One can notice that the smaller values of the given criteria are provided by one of the discrete Shanker versions. Therefore, we may conclude that exists evidence that the proposed models adhere well to the considered datasets and hence, they can be regarded as excellent alternatives for the modeling of count data in the presence of overdispersion.

Table 9: Comparison criteria for the fitted models.

Model	Dataset 1			Dataset 2		
	AICc	BIC	HQC	AICc	BIC	HQC
P	440.41	443.16	441.51	641.64	644.28	642.69
NB	404.71	410.18	406.87	570.69	575.94	572.75
T1DS	405.14	407.90	406.24	**569.42**	**572.06**	**570.47**
T2DS	**403.65**	**406.40**	**404.75**	570.55	573.19	571.60

7. Concluding Remarks

In this paper, two versions of the discrete Shanker distribution were introduced as alternatives to model overdispersed count datasets. To derive the proposed models, we have considered the methods of infinite series and survival function. Some statistical properties as the mean, variance, coefficients of variation, skewness, and kurtosis for each version were discussed. Also, it was shown that both versions of the discrete Shanker distribution are suitable options to deal with zero-inflated datasets. Moreover, we have derived the log-likelihood, the *score* function, and we have considered asymptotic intervallic estimation for parameter θ of both versions. Also, we have performed a Monte Carlo simulation study where the bias, the mean squared error, and the coverage lengths of the MLEs as well the coverage probability of the asymptotic CIs were computed. These measures indicate the suitability of the considered methodology. The usefulness of the proposed models was evaluated by fitting each one to two datasets with characteristics of overdispersion. The model selection was performed using the AICc, BIC, and HQC criteria. The goodness-of-fit was assessed by the χ^2 statistic. The obtained results demonstrate that the T1DS and T2DS distributions can be competitive with standard discrete models provided by literature.

Acknowledgements

The authors would like to thank the anonymous referees for its insightful suggestions that certainly contributed to improving this work. Josmar Mazucheli gratefully acknowledges the partial financial support from the Paraná Research Foundation (FA) - Grant 64/2019. The research of Wesley Bertoli is supported by the Federal University of Technology - Paraná and by the Paraná Research Foundation (FA) - Doctoral Grant: CP 18/2015.

Appendix A. Raw Datasets

The two real datasets used in the paper to illustrate the usefulness of the proposed models are provided in Table A1.

Table A1: Real datasets used in Section 6.

Dataset 1								
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	2	2	2
2	2	2	2	2	2	2	2	2
2	2	2	2	2	3	3	3	3
3	3	3	3	3	3	3	4	4
4	4	4	5	5	5	5	6	7
7	8	8						
Dataset 2								
5	4	6	16	5	8	8	9	10
10	7	1	6	5	6	5	13	6
10	13	4	8	5	0	2	2	2
8	4	11	4	8	9	9	4	0
9	8	5	5	10	3	5	4	6
6	5	1	2	2	2	2	4	3
2	3	1	2	0	1	1	1	1
5	7	2	9	3	6	9	3	3
5	9	10	9	15	18	13	10	9
7	7	0	3	3	4	2	1	2
2	3	0	5	5	1	1	1	1
8	5	9	6	3	4	6	2	3

[Received: December 2018 — Accepted: August 2019]

References

Bateman, H. & Erdélyi, A. (1953), *Higher transcendental functions*, Vol. 2, McGraw-Hill, New York.

Bi, Z., Faloutsos, C. & Korn, F. (2001), The DGX distribution for mining massive, skewed data, *in* 'Proceedings of the seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining', ACM, pp. 17–26.

Bliss, C. I. & Fisher, R. A. (1953), 'Fitting the negative binomial distribution to biological data', *Biometrics* **9**(2), 176–200.

Bracquemond, C. & Gaudoin, O. (2003), 'A survey on discrete lifetime distributions', *International Journal of Reliability, Quality and Safety Engineering* **10**(1), 69–98.

Chakraborty, S. (2015*a*), '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. (2015*b*), 'A new discrete distribution related to generalized Gamma distribution and its properties', *Communications in Statistics - Theory and Methods* **44**(8), 1691–1705.

Chakraborty, S. & Chakravarty, D. (2012), 'Discrete Gamma distributions: Properties and parameter estimation', *Communications in Statistics - Theory and Methods* **41**(18), 3301–3324.

Chakraborty, S. & Chakravarty, D. (2016), 'A new discrete probability distribution with integer support on $(-\infty, +\infty)$', *Communications in Statistics - Theory and Methods* **45**(2), 492–505.

Chakraborty, S. & Gupta, R. D. (2015), 'Exponentiated Geometric distribution: Another generalization of Geometric distribution', *Communications in Statistics - Theory and Methods* **44**(6), 1143–1157.

Collett, D. (2003), *Modelling survival data in medical research*, 2 edn, Chapman and Hall, New York.

Doornik, J. A. (2007), *Object-oriented matrix programming using Ox*, 3 edn, London: Timberlake Consultants Press and Oxford.

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

Ghitany, M. E., Atieh, B. & Nadarajah, S. (2008), 'Lindley distribution and its application', *Mathematics and Computers in Simulation* **78**(4), 493–506.

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.

Grandell, J. (1997), *Mixed Poisson processes*, Vol. 77, Chapman and Hall/CRC.

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', *Pakistan Journal of Statistics* **30**(2), 203–222.

Inusah, S. & Kozubowski, T. J. (2006), 'A discrete analogue of the Laplace distribution', *Journal of Statistical Planning and Inference* **136**(3), 1090–1102.

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

Kalbfleisch, J. D. & Prentice, R. L. (2002), *The statistical analysis of failure time data*, 2 edn, Wiley, New York.

Keilson, J. & Gerber, H. (1971), 'Some results for discrete unimodality', *Journal of the American Statistical Association* **66**(334), 386–389.

Kemp, A. W. (1997), 'Characterizations of a discrete Normal distribution', *Journal of Statistical Planning and Inference* **63**(2), 223–229.

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

Kemp, A. W. (2008), *The discrete Half-Normal distribution*, Birkhäuser Boston, Boston, pp. 353–360. In Advances in Mathematical and Statistical Modeling.

Kennan, J. (1985), 'The duration of contract strikes in U.S. manufacturing', *Journal of Econometrics* **28**(1), 5–28.

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. B. & 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*, 2 edn, John Wiley & Sons, Hoboken, New York.

Lee, E. T. & Wang, J. W. (2003), *Statistical methods for survival data analysis*, 3 edn, John Wiley & Sons, Hoboken, New York.

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* **R-24**(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**(11), 2000–2013.

Nekoukhou, V., Alamatsaz, M. H. & Bidram, H. (2013), 'Discrete generalized Exponential distribution of a second type', *Statistics - A Journal of Theoretical and Applied Statistics* **47**(4), 876–887.

R Development Core Team (2017), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. *http://www.R-project.org.

Roy, D. (2003), 'The discrete Normal distribution', *Communication in Statistics - Theory and Methods* **32**(10), 1871–1883.

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

Rubinstein, R. Y. & Kroese, D. P. (2008), *Simulation and the Monte Carlo method*, Wiley Series in Probability and Statistics, 2 edn, John Wiley & Sons, Hoboken, New York.

Saha, K. K. (2008), 'Analysis of one-way layout of count data in the presence of over or under dispersion', *Journal of Statistical Planning and Inference* **138**(7), 2067–2081.

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

Shanker, R. (2015), 'Shanker distribution and its applications', *International Journal of Statistics and Applications* **5**(6), 338–348.

Siromoney, G. (1964), 'The general Dirichlets Series distribution', *Journal of the Indian Statistical Association* **2-3**(2), 1–7.

Slater, L. J. (1966), *Generalized hypergeometric functions*, Cambridge University Press, London.

Vuong, Q. H. (1989), 'Likelihood ratio tests for model selection and non-nested hypotheses', *Econometrica* **57**(2), 307–333.