Methodology and Computing in Applied Probability
https://doi.org/10.1007/s11009-019-09731-7

On the Discrete Quasi Xgamma Distribution

Josmar Mazucheli[1] · **Wesley Bertoli**[2] · **Ricardo P. Oliveira**[3] · **André F. B. Menezes**[1]

Received: 15 October 2018 / Revised: 7 March 2019 / Accepted: 17 June 2019 /
Published online: 25 June 2019

Abstract

Methods to obtain discrete analogs of continuous distributions have been widely applied in recent years. In general, the discretization process provides probability mass functions that can be competitive with traditional models used in the analysis of count data. The discretization procedure also avoids the use of continuous distribution to model strictly discrete data. In this paper, we propose two discrete analogs for the quasi xgamma distribution as alternatives to model under- and overdispersed datasets. The methods of infinite series and survival function have been considered to derive the models and, despite the difference between the methods, the resulting distributions are interchangeable. Several statistical properties of the proposed models have been derived. The maximum likelihood theory has been considered for estimation and asymptotic inference concerns. An intensive simulation study has been carried out in order to evaluate the main properties of the maximum likelihood estimators. The usefulness of the proposed models has been assessed by using two real datasets provided by literature. A general comparison of the proposed models with some well-known discrete distributions has been provided.

Keywords Count data · Discretization methods · Quasi xgamma distribution · Data dispersion · Maximum likelihood estimation · Simulation study

Mathematics Subject Classification (2010) 62E15 · 62F10 · 62Q05

1 Introduction

In recent decades, the proposition of probabilistic models 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 in the analysis of strictly discrete data. For example, in survival analysis is common to use continuous distributions to

✉ Josmar Mazucheli
jmazucheli@gmail.com

1 Department of Statistics, State University of Maringá, Maringá, PR 87020-900, Brazil

2 Department of Statistics, Federal University of Technology - Paraná, Curitiba, PR 80230-901, Brazil

3 Medical School of Ribeirão Preto, University of São Paulo, Ribeirão Preto, SP 14049-900, Brazil

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 and Moeschberger (1997), Meeker and Escobar (1998), Kalbfleisch and Prentice (2002), Lee and Wang (2003), Lawless (2003), Collett (2003), Hamada et al. (2008), among many others. A complete survey regarding all discretization methods and some discretized distributions can be found in Chakraborty (2015b).

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), which has proposed the discrete Good distribution to model population frequencies of species. Such an approach was considered by other authors to define discrete analogs, and here 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 to model 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 discrete Exponential distribution was proposed by Sato et al. (1999) to describe the defect count frequencies on wafers or chips. Bi et al. (2001) introduced the discrete Log-normal distribution and showed that this model fits well to Internet click-stream data, among others. Inusah and Kozubowski (2006) presented the discrete Laplace distribution discussing that, relatively 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 and Inusah (2006). Besides, Kemp (2008) introduced the discrete Half-Normal distribution studying its relation with other distributions and Nekoukhou et al. (2012) proposed the discrete Generalized Exponential distribution as an attempt to model rank frequencies of graphemes in the Slovene language. Lisman and Van Zuylen (1972)

Another widespread method to obtain discrete analogs of continuous random variables is that one based on the survival function of the original distribution. This method was proposed by Nakagawa and Osaki (1975) and has the interesting feature of preserving the original survival function on its integer part for the generated pmf (Kemp 2004; Chakraborty 2015b). Several authors have considered the discretization method by survival function, and here we will also point out a few. Nakagawa and 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 and 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 and Pundir (2009) for application in reliability estimation in series systems. Jazi et al. (2010) proposed the discrete Inverse Weibull distribution and discussed different estimation methods for the model parameters. Gómez-Déniz and Calderín-ojeda (2011) introduced the discrete Lindley distribution and illustrated its application using automobile claim frequency data. The discrete Gamma distribution was proposed by Chakraborty and Chakravarty (2012), which derived several statistical properties of such a model. Also, Nekoukhou et al. (2013) presented the discrete Type II Generalized Exponential distribution, and Hussain and Ahmad (2014) introduced the discrete Inverse Rayleigh distribution as alternatives to model overdispersed count data.

The primary goal of this paper is to apply the methods of infinite series and survival function to derive discrete analogs for the quasi xgamma distribution, which is a 2-parameter lifetime model introduced and widely studied by Sen and Chandra (2017). We expect the proposed models to be suitable alternatives to model under- and overdispersed count datasets. A continuous random variable X is said to have xgamma distribution if its probability density function (pdf) can be written as

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

where $(\alpha,\theta) \in \mathbb{R}_+^2$ are the shape parameters. The authors has shown that this model can be derived as a 2-component mixture of an Exponential distribution with mean θ^{-1} and a Gamma distribution with shape parameter 3 and scale parameter θ, with mixing proportions given by $\alpha(1+\alpha)^{-1}$ and $(1+\alpha)^{-1}$, respectively.

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

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

for $(\alpha,\theta) \in \mathbb{R}_+^2$.

Remark 1 For specific values of parameter α, the quasi xgamma distribution has some particular cases.

i) If $\alpha = 0$, then X has a gamma distribution with shape parameter 3 and scale parameter θ, that is, $X \sim \text{Gamma}(3,\theta)$.

ii) If $\alpha = 1$, then Eq. 1 gives rise to a new class of distribution, with general pdf given by

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

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

iii) When $\alpha = \theta$, then Eq. 1 corresponds to the xgamma distribution, with pdf given by

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

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

This paper is organized as follow. In Section 2 we briefly present the methods of infinite series and survival function to derive discrete analogs of continuous distributions. In Section 3, we introduce two versions for the discrete quasi xgamma distribution, and we derive the main statistical properties of each model. In Section 4, the problem of estimating the model parameters is addressed, and classical inference procedures are discussed. In Section 5, the results of an intensive simulation study are presented in order to assess the main properties of the maximum likelihood estimators (MLEs). In Section 6, two real applications of the proposed models are exhibited as a way 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 quasi xgamma distribution. It is important to point out that the paper of Chakraborty (2015b) is possibly the only paper with an 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

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

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

Definition 1 Let X be a continuous random variable. If X has pdf $f_X\left(x;\boldsymbol{\theta}\right)$ with support on $\mathbb{R}$, then the corresponding discrete random variable Y has pmf given by

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

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

This method was exploited by several authors including Kulasekera and Tonkyn (1992), Doray and 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

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

One of the most recent examples of the use of this method is the discrete analog of the generalized Exponential distribution (Nekoukhou et al. 2012), whose pmf is given by

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

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 quasi xgamma distribution by this method.

2.2 Discretization by Survival Function

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

Definition 2 Let X be a continuous random variable. If X has survival function $S_X(x; \boldsymbol{\theta})$, then the discrete random variable $Y = \lfloor X \rfloor$ has pmf as follow

$$\mathrm{P}(Y = y; \boldsymbol{\theta}) = S_X(y; \boldsymbol{\theta}) - S_X(y + 1; \boldsymbol{\theta}), \quad y \in \mathbb{Z}_+, \tag{4}$$

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

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

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

and survival function

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

where $(\theta, \mu) \in \mathbb{R}_+^2$ are, respectively, the shape and the scale parameters, was one of the first discretized distributions by this method. Nakagawa and Osaki (1975) proposed the discrete Weibull distribution where the pmf for the random variable $Y = \lfloor X \rfloor$ is given by

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

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

3 The Discrete Quasi Xgamma Distribution

In this section, we will consider both methods previously presented to derive discrete analogs for the quasi xgamma distribution. For ease of notation, each probabilistic model provided by these methods will be denoted by DQX1 (type I discrete quasi xgamma) and DQX2 (type II discrete quasi xgamma) distributions, respectively. For each version of this model, the main statistical properties as the shape, the moments and the generating functions will be discussed.

3.1 Type I Discrete Quasi Xgamma Distribution

By considering Eq. 3, one can define the 2-parameter DQX1 distribution. We have the following definition.

Definition 3 Let X be a continuous random variable distributed accordingly to a xgamma distribution (1) with parameters $(\alpha, \theta) \in \mathbb{R}_+^2$. Let $h(z_1, z_2) = e^{z_1} - z_2$, $z_1 \in \mathbb{R}$ and $z_2 \in \mathbb{R}$. The pmf of Y having DQX1 distribution is given by

$$\mathrm{P}(Y = y; \alpha, \theta) = \frac{h(-\theta y, 0)}{q(\alpha, \theta)} \left[\alpha + \frac{\theta^2}{2} y^2\right], \quad y \in \mathbb{Z}_+, \tag{5}$$

for $(\alpha, \theta) \in \mathbb{R}_+^2$ and where

$$q(\alpha, \theta) = \frac{1}{2}\left\{\frac{h(\theta, 0)\left[2\alpha\, h^2(\theta, 1) + \theta^2 h(\theta, -1)\right]}{h^3(\theta, 1)}\right\}$$

is the inverse of the normalizing constant.

Proposition 1 *The Eq. 5 is a proper pmf.*

Proof Here we have to prove that $\sum_{y=0}^{\infty} \mathrm{P}(Y = y; \alpha, \theta) = 1$ for $(\alpha, \theta) \in \mathbb{R}_+^2$. Then

$$\begin{aligned}
\sum_{y=0}^{\infty} \mathrm{P}(Y = y; \theta) &= \sum_{y=0}^{\infty} \frac{h(-\theta y, 0)}{q(\alpha, \theta)}\left[\alpha + \frac{\theta^2}{2}\, y^2\right] \\
&= \frac{1}{q(\alpha, \theta)}\left\{\alpha \sum_{y=0}^{\infty} h(-\theta y, 0) + \frac{\theta^2}{2}\sum_{y=0}^{\infty} y^2 h(-\theta y, 0)\right\} \\
&= \frac{1}{q(\alpha, \theta)}\left\{\frac{\alpha\, h(\theta, 0)}{h(\theta, 1)} + \frac{\theta^2 h(\theta, 0)\, h(\theta, -1)}{2h^3(\theta, 1)}\right\} \\
&= 1,
\end{aligned}$$

which concludes the proof. □

For a discrete random variable Y having DQX1 distribution, we will adopt the notation $Y \sim \mathrm{DQX1}(\alpha, \theta)$. The pmf (5) does not involve complicated expressions, and therefore, the probabilities can be straightforwardly computed, as for example,

$$\mathrm{P}(Y = 0; \alpha, \theta) = \frac{\alpha}{q(\alpha, \theta)},$$

for $(\alpha, \theta) \in \mathbb{R}_+^2$. Figure 1 depicts the behavior of the pmf (5) for selected values of α and θ.

We have derived some theoretical properties of the DQX1 distribution. These properties are stated in the following propositions.

Proposition 2 *Let $Y \sim DQX1(\alpha, \theta)$. The survival function of Y is given by*

$$S(y; \alpha, \theta) = 1 - \frac{2\alpha\left[h(\theta, 0) - h(-\theta y, 0)\right] - \theta^2 t(y, \theta)}{2\, q(\alpha, \theta)\, h(\theta, 1)}, \quad y \in \mathbb{Z}_+,$$

for $(\alpha, \theta) \in \mathbb{R}_+^2$ and where

$$t(y, \theta) = \frac{y^2 h(-\theta y, 0)\left[h(\theta, 0)\, h(\theta, 2) + 1\right] + 2y h(-\theta y, 0)\, h(\theta, 0)\, h(\theta, 1) + h(\theta, 0)\, h(\theta, -1)\left[h(-\theta y, 0) - 1\right]}{h^2(\theta, 1)}.$$

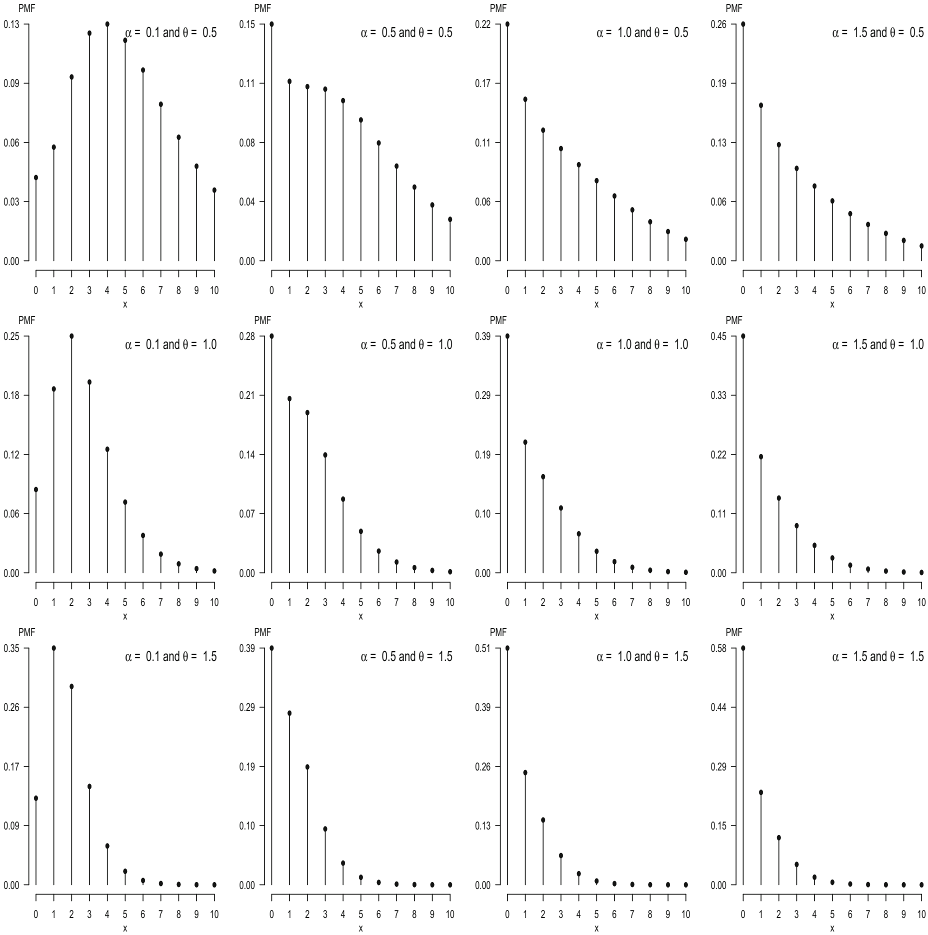

Fig. 1 Behavior of the DQX1 distribution for different values of α and θ

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

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

which concludes the proof. □

Proposition 3 *Let $Y \sim DQX1(\alpha, \theta)$. The probability generating function (pgf) of Y is given by*

$$G(s) = \frac{h(\theta, 0)}{2q(\alpha, \theta)} \left\{ \frac{2\alpha \left[h(\theta, 0)\, h(\theta, 2s) + s^2 \right] + \theta^2 s h(\theta, -s)}{h^3(\theta, s)} \right\},$$

for $(\alpha, \theta) \in \mathbb{R}_+^2$ and $s \neq e^{\theta}$.

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

$$\begin{aligned} G(s) &= \sum_{y=0}^{\infty} s^y \frac{h(-\theta y, 0)}{q(\alpha, \theta)} \left[\alpha + \frac{\theta^2}{2} y^2 \right] \\ &= \frac{1}{q(\alpha, \theta)} \left\{ \alpha \sum_{y=0}^{\infty} s^y h(-\theta y, 0) + \frac{\theta^2}{2} \sum_{y=0}^{\infty} y^2 s^y h(-\theta y, 0) \right\} \\ &= \frac{1}{q(\alpha, \theta)} \left\{ \frac{\alpha h(\theta, 0)}{h(\theta, s)} + \frac{s h(\theta, 0)\, h(\theta, -s)}{h^3(\theta, s)} \right\} \\ &= \frac{h(\theta, 0)}{2q(\alpha, \theta)} \left\{ \frac{2\alpha \left[h(\theta, 0)\, h(\theta, 2s) + s^2 \right] + \theta^2 s h(\theta, -s)}{h^3(\theta, s)} \right\}, \end{aligned}$$

which concludes the proof. □

Proposition 4 *Let $Y \sim DQX1(\alpha, \theta)$. The moment generating function (mgf) of Y is given by*

$$M(t) = \frac{h(\theta, 0)}{2q(\alpha, \theta)} \left\{ \frac{2\alpha \left[h(\theta, 0)\, h\left(\theta, 2e^t\right) + h(2t, 0) \right] + \theta^2 h(t, 0)\, h\left(\theta, -e^t\right)}{h^3(\theta, e^t)} \right\},$$

for $(\alpha, \theta) \in \mathbb{R}_+^2$ and $t \neq \theta$.

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

$$\begin{aligned} M(t) &= \sum_{y=0}^{\infty} e^{ty} \frac{h(-\theta y, 0)}{q(\alpha, \theta)} \left[\alpha + \frac{\theta^2}{2} y^2 \right] \\ &= \frac{1}{q(\alpha, \theta)} \left\{ \alpha \sum_{y=0}^{\infty} h[-y(\theta - t), 0] + \frac{\theta^2}{2} \sum_{y=0}^{\infty} y^2 h[-y(\theta - t), 0] \right\} \\ &= \frac{1}{q(\alpha, \theta)} \left\{ \frac{\alpha h(\theta, 0)}{h(\theta, e^t)} + \frac{h(t, 0)\, h(\theta, 0)\, h\left(\theta, -e^t\right)}{h^3(\theta, e^t)} \right\} \\ &= \frac{h(\theta, 0)}{2q(\alpha, \theta)} \left\{ \frac{2\alpha \left[h(\theta, 0)\, h\left(\theta, 2e^t\right) + h(2t, 0) \right] + \theta^2 h(t, 0)\, h\left(\theta, -e^t\right)}{h^3(\theta, e^t)} \right\}, \end{aligned}$$

which concludes the proof. □

Proposition 5 *Let $Y \sim DQX1(\alpha,\theta)$. The cumulant generating function (cgf) of Y is given by*

$$\begin{aligned} C(t) = {} & \log[h(\theta,0)] + \log\left\{2\alpha\left[h(\theta,0)\,h\left(\theta,2e^{t}\right) + h(2t,0)\right] + \theta^2 h(t,0)\,h\left(\theta,-e^{t}\right)\right\} \\ & - \log[2q(\alpha,\theta)] - 3\log\left[h\left(\theta,e^{t}\right)\right], \end{aligned}$$

for $(\alpha,\theta) \in \mathbb{R}_+^2$ and $t \neq \theta$.

Proof Straightforward. The result is obtained by noticing that $C(t) = \log[M(t)]$. □

Proposition 6 *Let $Y \sim DQX1(\alpha,\theta)$. The characteristic function (cf) of Y is given by*

$$\phi(t) = \frac{h(\theta,0)}{2q(\alpha,\theta)}\left\{\frac{2\alpha\left[h(\theta,0)\,h\left(\theta,2e^{it}\right) + e^{2it}\right] + \theta^2 e^{it} h\left(\theta,-e^{it}\right)}{h^3\left(\theta,e^{it}\right)}\right\},$$

for $(\alpha,\theta) \in \mathbb{R}_+^2$, $t \in \mathbb{R}$ and $i = \sqrt{-1}$ is the imaginary number.

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

Proposition 7 *Let $Y \sim DQX1(\alpha,\theta)$. The k^{th} moment of Y about the origin is given by*

$$\mu_k' = \frac{\left[3h(\theta,1) + h(-\theta,0) - h^2(\theta,0)\right]\left\{2\alpha\mathcal{P}_{-k}[h(-\theta,0)] + \theta^2\mathcal{P}_{-(k+2)}[h(-\theta,0)]\right\}}{2\alpha[1 - h(\theta,0)\,h(\theta,2)] - \theta^2 h(\theta,-1)}, \quad k \geqslant 1, \tag{6}$$

for $(\alpha,\theta) \in \mathbb{R}_+^2$ and where $\mathcal{P}$ is the general polylogarithm function defined as $\mathcal{P}_a(z) = \sum_{k=1}^{\infty} z^n n^{-a}$ for $|z| < 1$ and by an analytic continuation otherwise.

Proof By definition, $\mu_k' = E\left(Y^k\right)$. Then

$$\begin{aligned} \mu_k' &= \sum_{y=0}^{\infty} y^k \frac{h(-\theta y,0)}{q(\alpha,\theta)}\left[\alpha + \frac{\theta^2}{2}y^2\right] \\ &= \frac{1}{q(\alpha,\theta)}\left\{\alpha\sum_{y=0}^{\infty} y^k h(-\theta y,0) + \frac{\theta^2}{2}\sum_{y=0}^{\infty} y^{k+2} h(-\theta y,0)\right\} \\ &= \frac{1}{q(\alpha,\theta)}\left\{\alpha\mathcal{P}_{-k}[h(-\theta,0)] + \frac{\theta^2}{2}\mathcal{P}_{-(k+2)}[h(-\theta,0)]\right\} \\ &= \frac{\left[3h(\theta,1) + h(-\theta,0) - h^2(\theta,0)\right]\left\{2\alpha\mathcal{P}_{-k}[h(-\theta,0)] + \theta^2\mathcal{P}_{-(k+2)}[h(-\theta,0)]\right\}}{2\alpha[1 - h(\theta,0)\,h(\theta,2)] - \theta^2 h(\theta,-1)}. \end{aligned}$$

□

Now, from Eq. 6, the mean (μ) and the variance (σ^2) of Y are given, respectively, by

$$\mu = \mu_1' = \frac{2\alpha[h(\theta,0)\,h(\theta,2) + 1] + \theta^2[h(\theta,0)\,h(\theta,-4) + 1]}{h(\theta,1)\left[2\alpha[h(\theta,0)\,h(\theta,2) + 1] + \theta^2 h(\theta,-1)\right]},$$

and

$$\sigma^2 = \mu_2' - \left(\mu_1'\right)^2 = \frac{b\left(\alpha, \theta\right)}{h^2\left(\theta, 1\right)\left[2\alpha\left[h\left(\theta, 0\right)h\left(\theta, 2\right) + 1\right] + \theta^2 h\left(\theta, -1\right)\right]^2},$$

where

$$\begin{aligned} b\left(\alpha, \theta\right) &= 2h\left(\theta, 0\right)\left\{2\theta^4\left[h\left(\theta, 0\right)h\left(\theta, -1\right) + 1\right] + \alpha\theta^2\left[h^2\left(\theta, 0\right)\left[h\left(\theta, 0\right)h\left(\theta, -8\right) - 14\right] + 5\right]\right\} \\ &\quad + 4\alpha^2 h\left(\theta, 0\right)\left[h\left(\theta, 0\right)\left[h\left(\theta, 0\right)\left[h\left(\theta, 0\right)h\left(\theta, 4\right) + 6\right] - 4\right] + 1\right]. \end{aligned}$$

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

$$\mathrm{ID} = \frac{\sigma^2}{\mu} = \frac{b\left(\alpha, \theta\right)\left\{h\left(\theta, 1\right)\left[2\alpha\left[h\left(\theta, 0\right)h\left(\theta, 2\right) + 1\right] + \theta^2 h\left(\theta, -1\right)\right]\right\}^{-1}}{2\alpha\left[h\left(\theta, 0\right)h\left(\theta, 2\right) + 1\right] + \theta^2\left[h\left(\theta, 0\right)h\left(\theta, -4\right) + 1\right]}. \tag{7}$$

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

$$\mathrm{CV} = \frac{\sigma}{\mu} = \frac{\sqrt{b\left(\alpha, \theta\right)}}{2\alpha\left[h\left(\theta, 0\right)h\left(\theta, 2\right) + 1\right] + \theta^2\left[h\left(\theta, 0\right)h\left(\theta, -4\right) + 1\right]}.$$

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

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

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 DQX1 distribution, we have that the ZM index is given by

$$\mathrm{ZM} = 1 + \frac{h\left(\theta, 1\right)\left[2\alpha\left[h\left(\theta, 0\right)h\left(\theta, 2\right) + 1\right] + \theta^2 h\left(\theta, -1\right)\right]\left\{\log\left(\alpha\right) - \log\left[q\left(\alpha, \theta\right)\right]\right\}}{2\alpha\left[h\left(\theta, 0\right)h\left(\theta, 2\right) + 1\right] + \theta^2\left[h\left(\theta, 0\right)h\left(\theta, -4\right) + 1\right]}. \tag{8}$$

The asymmetry degree and the flatness of a distribution can be 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 DQX1 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 α and θ, the nature and the behavior of these coefficients along with the measures previously presented.

When assessing Eq. 8 more deeply, we have obtained that $\mathrm{ZM} \to 0$ as $\theta \to \infty$ and $\mathrm{ZM} \to 1$ as $\theta \to 0$. This implies that, besides the usual case ($\mathrm{ZM} = 0$), the DQX1 distribution is suitable to deal with zero-inflation but is not indicated to model zero-deflated datasets. Further, it is clear that the coefficient of skewness and the coefficient of kurtosis are increasing as α and θ increases. On the other hand, the higher values of the mean, of the variance and the index of dispersion are obtained when α and θ are simultaneously small.

Figure 2 depicts the behavior of Eq. 7. The feature that deserves to be highlighted is that the DQX1 distribution is suitable to deal with overdispersion and underdispersion as the index of dispersion can be either higher or smaller than 1 for certain values of parameters α and θ.

Table 1 Theoretical descriptive statistics under DQX1 distribution

α	θ	Measures						
		Mean	Variance	ID	CV	ZM	Skewness	Kurtosis
0.50	0.25	8.9386	53.0013	5.9295	0.8145	0.7173	1.1423	4.7333
1.00	0.50	3.5056	12.3745	3.5300	1.0035	0.5684	1.3630	5.2937
1.50	0.75	1.8878	4.9517	2.6231	1.1788	0.4578	1.5857	6.0867
2.00	1.00	1.1645	2.4866	2.1353	1.3541	0.3711	1.7963	6.9856
2.50	1.50	0.5843	0.9536	1.6321	1.6713	0.2508	2.1017	8.4004
1.00	3.00	0.2905	0.3049	1.0496	1.9008	0.0273	1.9586	6.9706
1.50	2.00	0.4646	0.6278	1.3513	1.7055	0.1657	1.9673	7.4146
2.00	1.50	0.6435	1.0489	1.6300	1.5916	0.2554	1.9694	7.6505
2.50	1.00	1.0715	2.2820	2.1298	1.4099	0.3623	1.9105	7.5983
3.00	0.75	1.4857	3.8704	2.6052	1.3242	0.4280	1.9178	7.7776

Proposition 8 *The DQX1 distribution has an increasing hazard rate.*

Proof The ratio of consecutive probabilities is given by

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

One can notice that the previous equation is a decreasing function on y. In this case, it follows that $\mathrm{P}\left(Y=y;\alpha,\theta\right)$ is log-concave and therefore, the DQX1 distribution has an increasing hazard rate. Hence the proof. $\square$

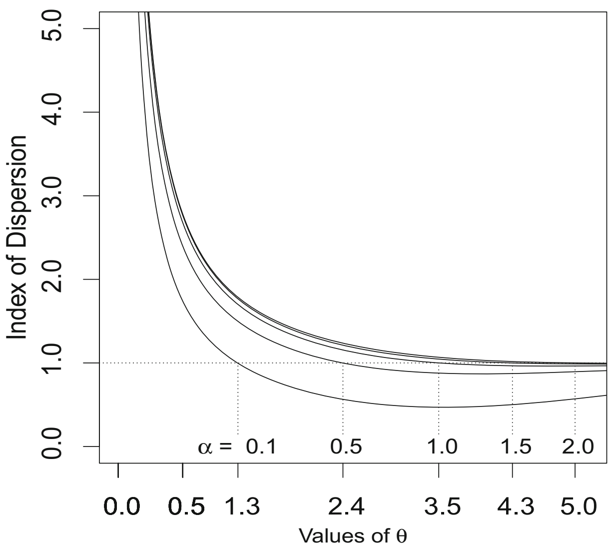

Fig. 2 Behavior of the index of dispersion of DQX1 distribution

In such context, it can also be proved that Eq. 5 satisfies $\mathrm{P}^2(Y = y; \alpha, \theta) \geqslant \mathrm{P}(Y = y - 1; \alpha, \theta) \times \mathrm{P}(Y = y + 1; \alpha, \theta)$, which implies unimodality (see Theorem 3 by Keilson and Gerber 1971). The relationship between log-concavity, unimodality and increasing hazard rate of discrete distributions has been discussed by Grandell (1997).

Proposition 9 *The DQX1 distribution has heavy tails as θ approaches zero.*

Proof The heavy-tail (HT) index is defined by

$$\mathrm{HT} = \lim_{y\to\infty} \frac{\mathrm{P}(Y = y + 1; \alpha, \theta)}{\mathrm{P}(Y = y; \alpha, \theta)},$$

for $(\alpha, \theta) \in \mathbb{R}_+^2$. For the DQX1 distribution, one can easily obtain $\mathrm{HT} = h(-\theta, 0)$. A discrete distribution is said to have heavy tails if $\mathrm{HT} \to 1$ when $y \to \infty$. Hence,

$$\lim_{\theta\to 0} \mathrm{HT} = \lim_{\theta\to 0} h(-\theta, 0) = 1,$$

which concludes the proof. □

3.2 Type II Discrete Quasi Xgamma Distribution

By considering Eqs. 2 and 4, one can define the 2-parameter DQX2 distribution. We have the following definition.

Definition 4 Let X be a continuous random variable distributed accordingly to a xgamma distribution (1) with parameters $\alpha \neq -1 \in \mathbb{R}, \theta \in \mathbb{R}_+$. Let $h(z_1, z_2, z_3) = z_2^2/2\,(z_1 + 1/z_2)^2 + z_3 + 1/2;\ z_1, z_2, z_3 \in \mathbb{R}$. The pmf of $Y = \lfloor X \rfloor$ having DQX2 distribution is give by

$$\mathrm{P}(Y = y; \alpha, \theta) = \frac{e^{-\theta y}}{\alpha + 1}\left[h(y, \theta, \alpha) - h(y + 1, \theta, \alpha)e^{-\theta}\right], \quad y \in \mathbb{Z}_+, \tag{9}$$

for $\alpha \neq -1 \in \mathbb{R}$ and $\theta \in \mathbb{R}_+$.

Proposition 10 *The* Eq. 9 *is a proper pmf.*

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

For a random variable Y distributed accordingly to a DQX2 distribution, we will adopt the notation $X \sim \mathrm{DQX2}(\alpha, \theta)$. For this version, the probabilities can be easily computed as noticed for the DQX1 distribution. Then

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

for $\alpha \neq -1 \in \mathbb{R}$ and $\theta \in \mathbb{R}_+$. Figure 3 shows the behavior of the pmf (9) of Y, using selected values for α and θ.

We have also derived some theoretical properties of the DQX2 distribution. These properties are stated in the following propositions.

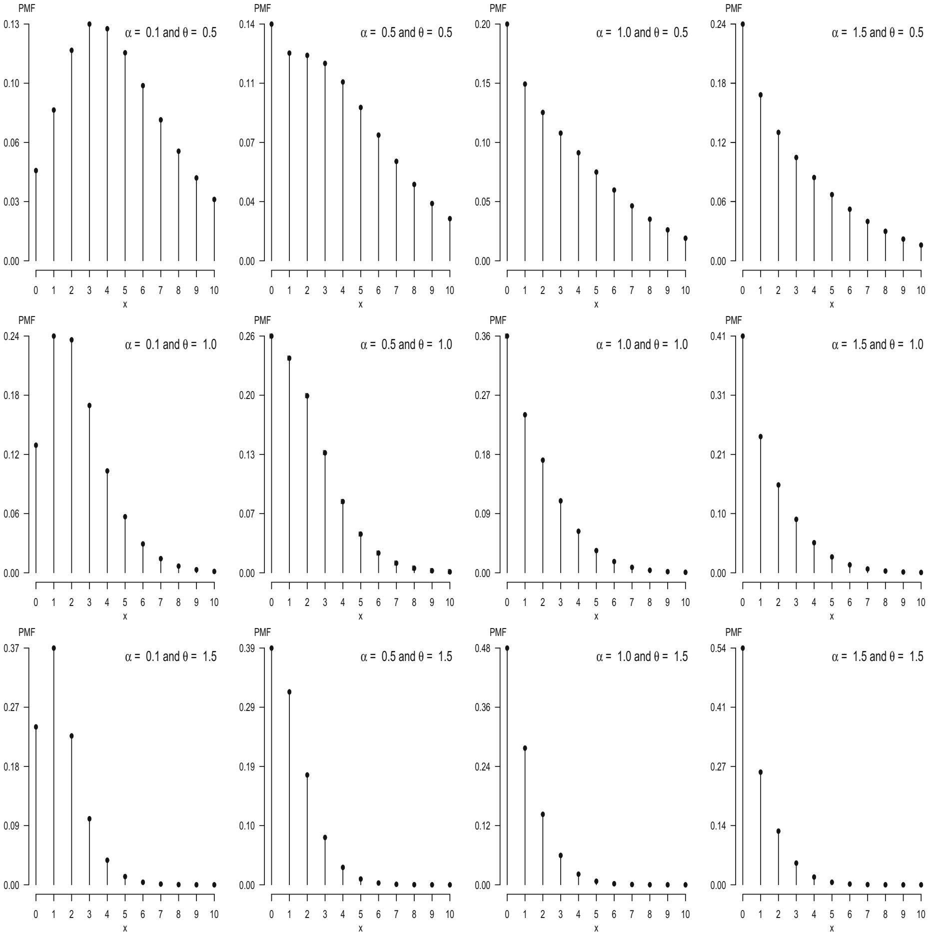

Fig. 3 Behavior of the DQX2 distribution for different values of α and θ

Proposition 11 *Let $Y \sim DQX2\,(\alpha, \theta)$. The survival function of Y is given by*

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

for $\alpha \neq -1 \in \mathbb{R}$ and $\theta \in \mathbb{R}_{+}$.

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

Proposition 12 *Let $Y \sim DQX2\,(\alpha, \theta)$. The pgf of Y is given by*

$$G(s) = \frac{2[h(\theta, 1, \alpha) - (\alpha + 1)e^{\theta}]q(s, \alpha, \theta)}{(\alpha + 1)(s - e^{\theta})^{3} r(\alpha, \theta)},$$

where

$$\begin{aligned} & q(s,\alpha,\theta) \\ & = \left[(\alpha+1)(e^{3\theta}-s^2)+(h(\theta,1,-2\alpha)-3)(e^{\theta}-1)se^{\theta}+h(\theta,1,\alpha)(s^2-e^{\theta})e^{\theta}\right](e^{\theta}-1), \end{aligned}$$

and

$$r(\alpha,\theta)=(e^{\theta}-\theta-1)^2-[e^{\theta}\theta^2-(e^{\theta}-1)^2(1+2\alpha)],$$

for $\alpha \neq -1 \in \mathbb{R}$, $\theta \in \mathbb{R}_+$ *and* $s \neq e^{\theta}$.

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

Proposition 13 *Let* $Y \sim DQX2\,(\alpha,\theta)$. *The mgf of* Y *is given by*

$$M\,(t)=\frac{2[h(\theta,1,\alpha)-(\alpha+1)e^{\theta}]q(e^{t},\alpha,\theta)}{(\alpha+1)(e^{t}-e^{\theta})^3 r(\alpha,\theta)},$$

for $\alpha \neq -1 \in \mathbb{R}$, $\theta \in \mathbb{R}_+$ *and* $t \neq \theta$.

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

Proposition 14 *Let* $Y \sim DQX2\,(\alpha,\theta)$. *The cgf of* Y *is given by*

$$\begin{aligned} C\,(t) \;=\; & \log(2)+\log([h(\theta,1,\alpha)-(\alpha+1)e^{\theta}])+\log(q(e^{t},\alpha,\theta))-\log(\alpha+1) \\ & -3\log(e^{t}-e^{\theta})-\log(r(\alpha,\theta)), \end{aligned}$$

for $\alpha \neq -1 \in \mathbb{R}$, $\theta \in \mathbb{R}_+$ *and* $t \neq \theta$.

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

Proposition 15 *Let* $Y \sim DQX2\,(\alpha,\theta)$. *The cf of* Y *is given by*

$$\phi\,(t)=\frac{2[h(\theta,1,\alpha)-(\alpha+1)e^{\theta}]q(e^{it},\alpha,\theta)}{(\alpha+1)(e^{it}-e^{\theta})^3 r(\alpha,\theta)},$$

for $\alpha \neq -1 \in \mathbb{R}$, $\theta \in \mathbb{R}_+$ *and* $t \in \mathbb{R}$.

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

Proposition 16 *Let* $Y \sim DQX2\,(\alpha,\theta)$. *The* k^{th} *moment of* Y *about the origin is given by*

$$\mu'_k=\frac{[2\theta(1-(1+\theta)e^{-\theta})]\mathcal{P}_{-(k+1)}\left[e^{-\theta}\right]-\theta^2(e^{-\theta}-1)\mathcal{P}_{-(k+2)}\left[e^{-\theta}\right]-2[h(\theta,1,\alpha)e^{-\theta}-\alpha-1]\mathcal{P}_{-k}\left[e^{-\theta}\right]}{2(\alpha+1)},$$
$$k \geqslant 1,$$

for $(\alpha,\theta) \in \mathbb{R}^2_+$ *and where* $\mathcal{P}$ *is the general polylogarithm function defined in Proposition 7.*

Proof The result comes analogously to the proof of Proposition 7. □

Now, from Eq. 10, the mean (μ) and the variance (σ^2) of Y are given, respectively, by

$$\mu = \frac{b(\alpha, \theta)}{(\alpha + 1)(e^{\theta} - 1)^3},$$

where $b(\alpha, \theta) = e^{\theta}\left[h(\theta, 1, \alpha)(e^{\theta} + 1) - 2(\alpha + \theta + 1)\right] - (\alpha + 1)(e^{\theta} - 1)$ and

$$\sigma^2 = \frac{c(\alpha, \theta) - [b(\alpha, \theta)]^2}{(\alpha + 1)^2(e^{\theta} - 1)^6},$$

where $c\,(\alpha, \theta) = (\alpha + 1)(e^{\theta} - 1)^2\{e^{\theta}\left[8\,h\,(\theta, -1/2, -\alpha/4) - 10 + e^{\theta}\,(1/7\,h(\theta, 7, -14\alpha)\right.$ $\left.-15/7) + h(\theta, 1, \alpha)(1 + e^{\theta} + e^{2\theta})\right] + \alpha + 1\}$.

Analogously to the DQX1 distribution, the index of dispersion (ID) of DQX2 distribution, whose behavior is depicted in Fig. 4, is given by

$$\mathrm{ID} = \frac{c(\alpha, \theta) - [b(\alpha, \theta)]^2}{b(\alpha, \theta)(\alpha + 1)(e^{\theta} - 1)^3}, \tag{10}$$

and the CV has the form

$$\mathrm{CV} = \frac{\sqrt{c(\alpha, \theta) - [b(\alpha, \theta)]^2}}{b(\alpha, \theta)}.$$

As for the DQX1 distribution, the coefficients of skewness and kurtosis of the DQX2 distribution have extensive and very complicated expression. These expressions will also be omitted, but Table 2 summarizes, for selected values of α and θ, the nature and the behavior of these coefficients along with the measures previously presented in this subsection.

For this version, the ZM index is given by

$$\mathrm{ZM} = 1 + \frac{\left[\log\left(h(0, \theta, \alpha) - h(1, \theta, \alpha)e^{-\theta}\right) - \log(\alpha + 1)\right](\alpha + 1)(e^{\theta} - 1)^3}{b(\alpha, \theta)}. \tag{11}$$

The limit properties of Eq. 11 are equal to those obtained for Eq. 8, i.e., ZM $\to 0$ as $\theta \to \infty$ and ZM $\to 1$ as $\theta \to 0$. This implies that the DQX2 distribution is also suitable to deal

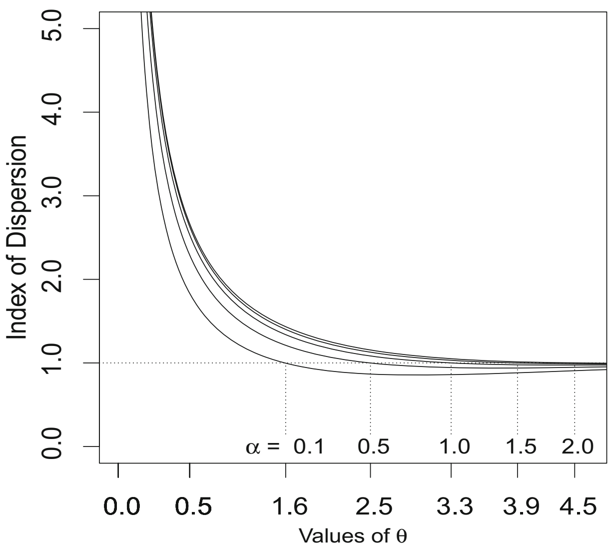

Fig. 4 Behavior of the index of dispersion of DQX2 distribution

Table 2 Theoretical descriptive statistics under DQX2 distribution

α	θ	Measures						
		Mean	Variance	ID	CV	ZM	Skewness	Kurtosis
0.50	0.25	8.8403	51.5098	5.8267	0.8119	0.7072	1.1689	4.8389
1.00	0.50	3.5207	11.9186	3.3853	0.9806	0.5484	1.3668	5.3793
1.50	0.75	1.9369	4.8140	2.4853	1.1328	0.4320	1.5484	6.0302
2.00	1.00	1.2209	2.4602	2.0150	1.2847	0.3427	1.7149	6.7145
2.50	1.50	0.6328	0.9701	1.5330	1.5566	0.2176	1.9473	7.7009
1.00	3.00	0.2722	0.2770	1.0178	1.9338	0.0086	1.9699	7.0331
1.50	2.00	0.4915	0.6135	1.2483	1.5937	0.1176	1.8258	6.9011
2.00	1.50	0.6903	1.0469	1.5165	1.4821	0.2150	1.8328	7.1188
2.50	1.00	1.1296	2.2902	2.0274	1.3397	0.3400	1.8163	7.2175
3.00	0.75	1.5463	3.8993	2.5217	1.2770	0.4168	1.8460	7.4645

with zero-inflation but is not indicated to model zero-deflated datasets. On the other hand, since $\theta \in \mathbb{R}_+$, the central moments of the DQX2 distribution present the same behavior respect to those derived for the DQX1 distribution. Moreover, since equation (10) presents the same limit properties of Eq. 7, we conclude that the DQX2 distribution may also be considered as an alternative to model under- and overdispersed datasets.

Proposition 17 *The DQX2 distribution has an increasing hazard rate.*

Proof The ratio of consecutive probabilities is given by

$$\frac{\mathrm{P}\left(Y=y+1;\alpha,\theta\right)}{\mathrm{P}\left(Y=y;\alpha,\theta\right)}=\frac{e^{-\theta(y+1)}[h(y+1,\theta,\alpha)-h(y+2,\theta,\alpha)e^{-\theta}]}{e^{-\theta y}[h(y,\theta,\alpha)-h(y+1,\theta,\alpha)e^{-\theta}]}. \tag{12}$$

One can notice that Eq. 12 is also a decreasing function on y. In this case, it follows that $\mathrm{P}\left(Y=y;\alpha,\theta\right)$ is log-concave and therefore, the DQX2 distribution has an increasing hazard rate. Hence the proof. □

For the DQX2 distribution, it also holds that Eq. 9 satisfies the inequality

$$\mathrm{P}^2\left(Y=y;\alpha,\theta\right)\geqslant\mathrm{P}\left(Y=y;\alpha,\theta\right)\mathrm{P}\left(Y=y+1;\alpha,\theta\right).$$

Proposition 18 *The DQX2 distribution has heavy tails as θ approaches zero.*

Proof The result comes analogously to the proof of Proposition 9. □

4 Maximum Likelihood Estimation

In this section, we will address the issue of estimating the parameter θ of both versions of the discrete quasi xgamma distribution. We have adopted the classical approach, and here we will derive the maximum likelihood function for the DQX1 and DQX2 models. Using these functions, one can obtain point estimates for parameter θ in each case. Moreover, suitable estimates for the confidence intervals (CIs) can be obtained using large sample approximations, that are based on the asymptotic properties of the MLEs.

4.1 Inference Under DQX1 Distribution

Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ a random sample of size n from the DQX1 distribution and $\boldsymbol{y} = (y_1, \ldots, y_n)$ its observed values. The log-likelihood function of the vector $\boldsymbol{\theta} = (\alpha, \theta)$ can be expressed as

$$\ell_n(\boldsymbol{\theta}; \boldsymbol{y}) = -n\left\{\theta\overline{y} + \log\left[q(\alpha, \theta)\right]\right\} + \sum_{i=1}^{n} \log\left[\alpha + \frac{\theta^2}{2} y_i^2\right], \tag{13}$$

where $\overline{y}$ is the sample mean. The MLE $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ can be obtained by direct maximization of the log-likelihood function (13). Hence, the components of the *score* vector, $U_{\boldsymbol{\theta}} = (U_\alpha, U_\theta)^\intercal$, are given by

$$U_\alpha = \frac{\partial \ell_n(\alpha, \theta; \boldsymbol{y})}{\partial \alpha} = -\frac{n}{q(\alpha, \theta)} \frac{\partial q(\alpha, \theta)}{\partial \alpha} + 2\sum_{i=1}^{n} \frac{1}{2\alpha + \theta^2 y_i^2},$$

and

$$U_\theta = \frac{\partial \ell_n(\alpha, \theta; \boldsymbol{y})}{\partial \theta} = -n\overline{y} - \frac{n}{q(\alpha, \theta)} \frac{\partial q(\alpha, \theta)}{\partial \theta} + 2\theta \sum_{i=1}^{n} \frac{y_i^2}{2\alpha + \theta^2 y_i^2}.$$

There is no closed-form solution for the MLE of $\boldsymbol{\theta}$, and therefore, standard optimization algorithms such as Newton-Raphson based methods may be used to obtain numerical estimates. Now, the Hessian matrix can be obtained as

$$\mathcal{H}(\boldsymbol{\theta}) = \begin{bmatrix} U_{\alpha\alpha} & U_{\alpha\theta} \\ U_{\theta\alpha} & U_{\theta\theta} \end{bmatrix},$$

where

$$U_{\alpha\alpha} = \frac{n}{q(\alpha, \theta)} \left\{ \frac{1}{q(\alpha, \theta)} \left[\frac{\partial q(\alpha, \theta)}{\partial \alpha}\right]^2 - \frac{\partial^2 q(\alpha, \theta)}{\partial \alpha^2} \right\} - 4\sum_{i=1}^{n} \frac{1}{\left[2\alpha + \theta^2 y_i^2\right]^2},$$

$$\begin{aligned} U_{\theta\theta} &= \frac{n}{q(\alpha, \theta)} \left\{ \frac{1}{q(\alpha, \theta)} \left[\frac{\partial q(\alpha, \theta)}{\partial \theta}\right]^2 - \frac{\partial^2 q(\alpha, \theta)}{\partial \theta^2} \right\} \\ &\quad -2\left\{ \sum_{i=1}^{n} \frac{y_i^2}{2\alpha + \theta^2 y_i^2} + 2\theta^2 \sum_{i=1}^{n} \frac{y_i^4}{\left(2\alpha + \theta^2 y_i^2\right)^2} \right\}, \end{aligned}$$

and

$$\begin{aligned} U_{\alpha\theta} = U_{\theta\alpha} &= \frac{\partial^2 \ell_n(\alpha, \theta; \boldsymbol{y})}{\partial \alpha \partial \theta} = \frac{n}{q(\alpha, \theta)} \left\{ \frac{1}{q(\alpha, \theta)} \frac{\partial q(\alpha, \theta)}{\partial \alpha} \frac{\partial q(\alpha, \theta)}{\partial \theta} - \frac{\partial q(\alpha, \theta)}{\partial \alpha \partial \theta} \right\} \\ &\quad -4\theta \sum_{i=1}^{n} \frac{y_i^2}{\left[2\alpha + \theta^2 y_i^2\right]^2}. \end{aligned}$$

Now, the Fisher information of $\boldsymbol{\theta}$ is given by

$$\mathcal{I}(\boldsymbol{\theta}) = -E_{\mathbf{Y}}\left[\mathcal{H}(\boldsymbol{\theta})\right] = \begin{bmatrix} S_{\alpha\alpha} & S_{\alpha\theta} \\ S_{\theta\alpha} & S_{\theta\theta} \end{bmatrix},$$

where

$$S_{\alpha\alpha} = -\frac{n}{q(\alpha,\theta)}\left\{\frac{1}{q(\alpha,\theta)}\left[\frac{\partial q(\alpha,\theta)}{\partial\alpha}\right]^2 - \frac{\partial^2 q(\alpha,\theta)}{\partial\alpha^2}\right\}$$
$$+\frac{2nh^3(\theta,1)\,{}_3\mathcal{F}_2\left[1,a_1,-a_1;1-a_1,1+a_1;h(-\theta,0)\right]}{\alpha h(\theta,0)\left[2\alpha h^2(\theta,1)+\theta^2 h(\theta,-1)\right]},$$

where ${}_p\mathcal{F}_q$ is the generalized hypergeometric function (Slater 1966) and $a_1 = \theta^{-1}\sqrt{-2\alpha}$. Also,

$$S_{\theta\theta} = -\frac{n}{q(\alpha,\theta)}\left\{\frac{1}{q(\alpha,\theta)}\left[\frac{\partial q(\alpha,\theta)}{\partial\theta}\right]^2 - \frac{\partial^2 q(\alpha,\theta)}{\partial\theta^2}\right\} + \frac{2nh(\theta,-1)}{2\alpha h^2(\theta,1)+\theta^2 h(\theta,-1)}$$
$$+\frac{4n\theta^2 h^3(\theta,1)\,{}_6\mathcal{F}_5\left[2,2,2,2,1-a_1,1+a_1;1,1,1,2-a_1,2+a_1;h(-\theta,0)\right]}{h(2\theta,0)\left(\theta^2+2\alpha\right)\left[2\alpha h^2(\theta,1)+\theta^2 h(\theta,-1)\right]},$$

and

$$S_{\alpha\theta} = S_{\theta\alpha} = -\frac{n}{q(\alpha,\theta)}\left\{\frac{1}{q(\alpha,\theta)}\frac{\partial q(\alpha,\theta)}{\partial\alpha}\frac{\partial q(\alpha,\theta)}{\partial\theta} - \frac{\partial q(\alpha,\theta)}{\partial\alpha\partial\theta}\right\}$$
$$+\frac{4n\theta h^3(\theta,1)\,{}_4\mathcal{F}_3\left[2,2,1-a_1,1+a_1;1,2-a_1,2+a_1;h(-\theta,0)\right]}{h(2\theta,0)\left(\theta^2+2\alpha\right)\left[2\alpha h^2(\theta,1)+\theta^2 h(\theta,-1)\right]}.$$

By the maximum likelihood theory, a consistent estimator for the covariance matrix of $\hat{\boldsymbol{\theta}}$ is obtained by the inverse of the Fisher information of $\boldsymbol{\theta}$, evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, i.e.,

$$\hat{\Sigma}_{\boldsymbol{\theta}} = \begin{bmatrix} S^{-1}_{\hat{\alpha}\hat{\alpha}} & S^{-1}_{\hat{\alpha}\hat{\theta}} \\ S^{-1}_{\hat{\theta}\hat{\alpha}} & S^{-1}_{\hat{\theta}\hat{\theta}} \end{bmatrix}.$$

Finally, in order to obtain intervallic estimates for parameters α and θ, one can use large sample approximations for the $100 \times (1-\alpha)\,\%$ two-sided CIs as

$$\hat{\alpha} \pm z_{1-\alpha/2}\sqrt{S^{-1}_{\hat{\alpha}\hat{\alpha}}} \quad \text{and} \quad \hat{\theta} \pm z_{1-\alpha/2}\sqrt{S^{-1}_{\hat{\theta}\hat{\theta}}},$$

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

4.2 Inference Under DQX2 Distribution

Let $\mathbf{Y} = (Y_1,\ldots,Y_n)$ a random sample of size n from the DQX2 distribution and $\boldsymbol{y} = (y_1,\ldots,y_n)$ its observed values. The log-likelihood function of the vector $\boldsymbol{\theta} = (\alpha,\theta)$ can be expressed as

$$\ell_n(\boldsymbol{\theta};\boldsymbol{y}) = -n\{\theta\overline{y} + \log(\alpha+1)\} + \sum_{i=1}^{n}\log\left[h(y_i,\theta,\alpha) - h(y_i+1,\theta,\alpha)e^{-\theta}\right], \quad (14)$$

where $\overline{y}$ is the sample mean. The MLE $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ can be obtained by direct maximization of the log-likelihood function (14). Hence, the components of the *score* vector, $U_{\boldsymbol{\theta}} = (U_\alpha, U_\theta)^\intercal$, are given by

$$U_\alpha = \frac{\partial\ell_n(\alpha,\theta;\boldsymbol{y})}{\partial\alpha} = -\frac{n}{\alpha+1} + \sum_{i=1}^{n}\frac{1-e^{-\theta}}{h(y_i,\theta,\alpha) - h(y_i+1,\theta,\alpha)e^{-\theta}},$$

and

$$
\begin{aligned}
U_\theta &= \frac{\partial \ell_n(\alpha, \theta; \boldsymbol{y})}{\partial \theta} \\
&= -n\overline{y} + \sum_{i=1}^{n} \frac{y_i(\theta y_i + 1) + h(y_i + 1, \theta, \alpha)e^{-\theta} - (y_i + 1)(\theta y_i + \theta + 1)e^{-\theta}}{h(y_i, \theta, \alpha) - h(y_i + 1, \theta, \alpha)e^{-\theta}}.
\end{aligned}
$$

There is no closed-form solution for the MLE of $\boldsymbol{\theta}$, and therefore, standard optimization algorithms such as Newton-Raphson based methods may be used to obtain numerical estimates. Now, the Hessian matrix can be obtained as

$$
\mathcal{H}(\boldsymbol{\theta}) = \begin{bmatrix} U_{\alpha\alpha} & U_{\alpha\theta} \\ U_{\theta\alpha} & U_{\theta\theta} \end{bmatrix},
$$

where

$$
U_{\alpha\alpha} = \frac{n}{(\alpha+1)^2} - \sum_{i=1}^{n} \frac{(1 - e^{-\theta})^2}{\left[h(y_i, \theta, \alpha) - h(y_i + 1, \theta, \alpha)e^{-\theta}\right]^2},
$$

$$
\begin{aligned}
U_{\theta\theta} = \sum_{i=1}^{n} \Bigg\{ & \frac{y_i^2 + e^{-\theta}(y_i + 1)[2(\theta y_i + \theta + 1) - (y_i + 1)]}{h(y_i, \theta, \alpha) - h(y_i + 1, \theta, \alpha)e^{-\theta}} - \frac{h(y_i + 1, \theta, \alpha)e^{-\theta}}{h(y_i, \theta, \alpha) - h(y_i + 1, \theta, \alpha)e^{-\theta}} \\
& - \left[\frac{-h(y_i + 1, \theta, \alpha)e^{-\theta}}{h(y_i, \theta, \alpha) - h(y_i + 1, \theta, \alpha)e^{-\theta}} + \frac{y_i(\theta y_i + 1) - (y_i + 1)(\theta y_i + \theta + 1)e^{-\theta}}{h(y_i, \theta, \alpha) - h(y_i + 1, \theta, \alpha)e^{-\theta}}\right]^2 \Bigg\},
\end{aligned}
$$

and

$$
\begin{aligned}
U_{\alpha\theta} = U_{\theta\alpha} = \sum_{i=1}^{n} \Bigg\{ & \frac{e^{-\theta}}{h(y_i, \theta, \alpha) - h(y_i + 1, \theta, \alpha)e^{-\theta}} \\
& - \frac{[y_i(\theta y_i + 1) - (y_i + 1)(\theta y_i + \theta + 1)e^{-\theta}](1 - e^{-\theta})}{[h(y_i, \theta, \alpha) - h(y_i + 1, \theta, \alpha)e^{-\theta}]^2} \\
& - \frac{h(y_i + 1, \theta, \alpha)(e^{-\theta} - e^{-2\theta})}{[h(y_i, \theta, \alpha) - h(y_i + 1, \theta, \alpha)e^{-\theta}]^2} \Bigg\}.
\end{aligned}
$$

Now, the Fisher information of $\boldsymbol{\theta}$ is given by

$$
\mathcal{I}(\boldsymbol{\theta}) = -E_{\mathbf{Y}}\left[\mathcal{H}(\boldsymbol{\theta})\right] = \begin{bmatrix} S_{\alpha\alpha} & S_{\alpha\theta} \\ S_{\theta\alpha} & S_{\theta\theta} \end{bmatrix},
$$

where

$$
\begin{aligned}
S_{\alpha\alpha} = {}_3\mathcal{F}_2 \Bigg[& 1, k_2(1,1,\theta) + \frac{\sqrt{k_1(\alpha,\theta)}}{(e^\theta - 1)\theta}, k_2(1,1,\theta) - \frac{\sqrt{k_1(\alpha,\theta)}}{(e^\theta - 1)\theta}; \\
& k_2(\theta+1,2,\theta) - \frac{\sqrt{k_1(\alpha,\theta)}}{(e^\theta - 1)\theta}, k_2(\theta+1,2,\theta) + \frac{\sqrt{k_1(\alpha,\theta)}}{(e^\theta - 1)\theta}; \frac{1}{e^\theta} \Bigg] - \frac{n}{(\alpha+1)^2} + \\
& \frac{2e^\theta(1 - e^{-\theta})^2}{2e^\theta(\alpha+1)^2 - (\theta^2 + 2(\alpha+\theta+1))(\alpha+1)},
\end{aligned}
$$

where ${}_p\mathcal{F}_q$ is the generalized hypergeometric function, with k_1 and k_2 defined by

$$
k_1(\alpha, \theta) = (e^{2\theta} + 1)(-2\alpha - 1) + (\theta^2 + 4\alpha + 2)e^\theta \quad \text{and} \quad k_2(a_1, a_2, \theta) = \frac{a_1 e^\theta - a_2\theta - 1}{(e^\theta - 1)\theta}.
$$

Due to the complex form of the terms $U_{\alpha\theta} = U_{\theta\alpha}$ and $U_{\theta\theta}$ from Hessian matrix, the components $S_{\alpha\theta} = S_{\theta\alpha}$ and $S_{\theta\theta}$ from Fisher information matrix demands massive calculations and also depends on the hypergeometric function stated for $S_{\alpha\alpha}$. These calculations are not illustrated here. By the maximum likelihood theory, a consistent estimator for the covariance matrix of $\hat{\boldsymbol{\theta}}$ is obtained by the inverse of the Fisher information of $\boldsymbol{\theta}$, evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, i.e.,

$$\hat{\Sigma}_{\boldsymbol{\theta}} = \begin{bmatrix} S^{-1}_{\hat{\alpha}\hat{\alpha}} & S^{-1}_{\hat{\alpha}\hat{\theta}} \\ S^{-1}_{\hat{\theta}\hat{\alpha}} & S^{-1}_{\hat{\theta}\hat{\theta}} \end{bmatrix}.$$

Finally, in order to obtain intervallic estimates for parameters α and θ, one can use large sample approximations for the $100 \times (1 - \alpha)\,\%$ two-sided CIs as

$$\hat{\alpha} \pm z_{1-\alpha/2}\sqrt{S^{-1}_{\hat{\alpha}\hat{\alpha}}} \quad \text{and} \quad \hat{\theta} \pm z_{1-\alpha/2}\sqrt{S^{-1}_{\hat{\theta}\hat{\theta}}},$$

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

5 Simulation Study

In this section we have estimated, using $B = 10,000$ Monte Carlo simulation, the bias (B), the root mean squared error (RMSE), the coverage probabilities (CP) and the coverage lengths (LCP) of the MLE of $\hat{\boldsymbol{\theta}}$ of both versions of the discrete quasi xgamma distribution. To run the simulation, we have considered $\alpha = 0.5$, 1.0 and 1.5 and $\theta = 0.5$, 1.0, 1.5 and 2.0 for sample sizes ranging from 20 to 200 by 30. The inverse-transform method for discrete distributions (Rubinstein and Kroese 2008) was implemented to generate the pseudo-random samples. The simulation process was performed using `Ox Console` (Doornik 2007). Let $\beta = \alpha$ or θ. The quantities of interest were estimated by the following expressions.

- $\mathrm{B}(\hat{\beta}) = B^{-1}\sum_{i=1}^{B}(\hat{\beta}_i - \beta)$;
- $\mathrm{RMSE}(\hat{\beta}) = B^{-1/2}\sqrt{\sum_{i=1}^{B}(\hat{\beta}_i - \beta)^2}$;
- $\mathrm{CP}_{\beta}(n) = B^{-1}\sum_{i=1}^{B}\mathcal{I}\left\{\hat{\beta}_i - 1.96\,\hat{\mathrm{SE}}\left(\hat{\beta}_i\right) < \theta < \hat{\beta}_i + 1.96\,\hat{\mathrm{SE}}\left(\hat{\beta}_i\right)\right\}$, where $\mathcal{I}\{\cdot\}$ denotes the indicator function and $\hat{\mathrm{SE}}\left(\hat{\beta}_i\right)$ stands for the estimated asymptotic standard error of $\hat{\beta}_i$;
- $\mathrm{LCP}_{\beta}(n) = 3.92\,B^{-1}\sum_{i=1}^{B}\hat{\mathrm{SE}}\left(\hat{\beta}_i\right)$.

For both versions, the behavior of the average bias is illustrated in Figs. 5 and 6. The average root mean squared error is showed in Figs. 7 and 8. The results obtained for the coverage probabilities and the coverage lengths are reported in Tables 3 and 4.

From the obtained results, in each scenario and for both versions of the discrete xgamma distribution, we have that the bias of $\hat{\boldsymbol{\theta}}$ is positive and tends to zero when the sample size increases. For α the bias is negative in all scenarios. Also, the mean squared error of $\hat{\boldsymbol{\theta}}$ tends

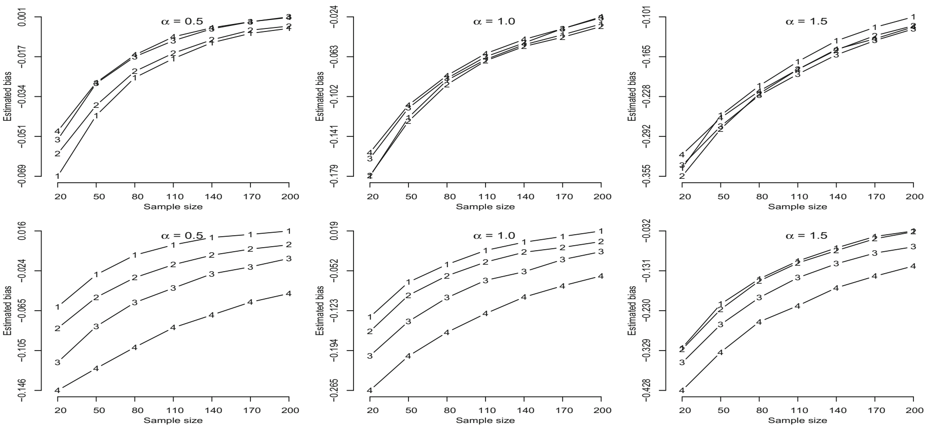

Fig. 5 Estimated bias for α ($1 : \theta = 0.5$, $2 : \theta = 1.0$, $3 : \theta = 1.5$ and $4 : \theta = 2.0$). Upper-panel: DQX1 distribution. Lower-panel: DQX2 distribution

to zero in each case. Besides, one can notice that the coverage probabilities for parameter θ are always higher than 94% for both discretizations and the coverage length tends to zero when the sample size increases.

6 Application to Real-Life Data

In this section, we will present two applications using real datasets as a way to show that the proposed models may be attractive alternatives to some standard discrete distributions. All computations were performed using the `R` environment (R Development Core Team 2017).

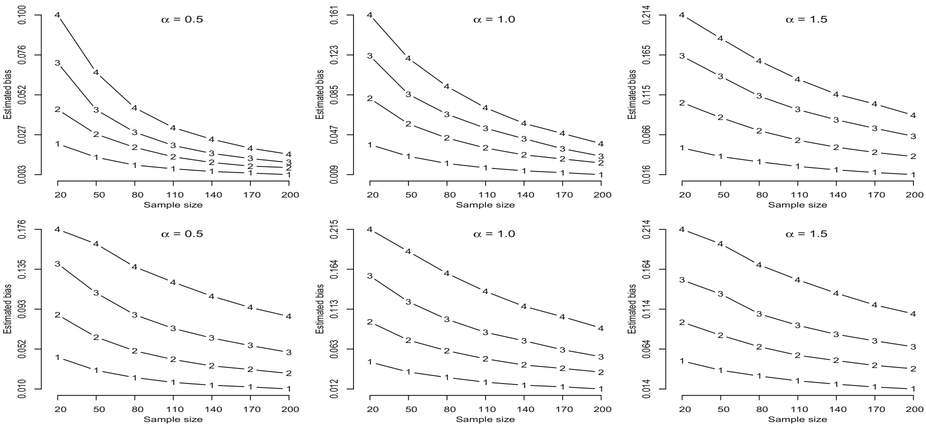

Fig. 6 Estimated bias for θ ($1 : \theta = 0.5$, $2 : \theta = 1.0$, $3 : \theta = 1.5$ and $4 : \theta = 2.0$). Upper-panel: DQX1 distribution. Lower-panel: DQX2 distribution

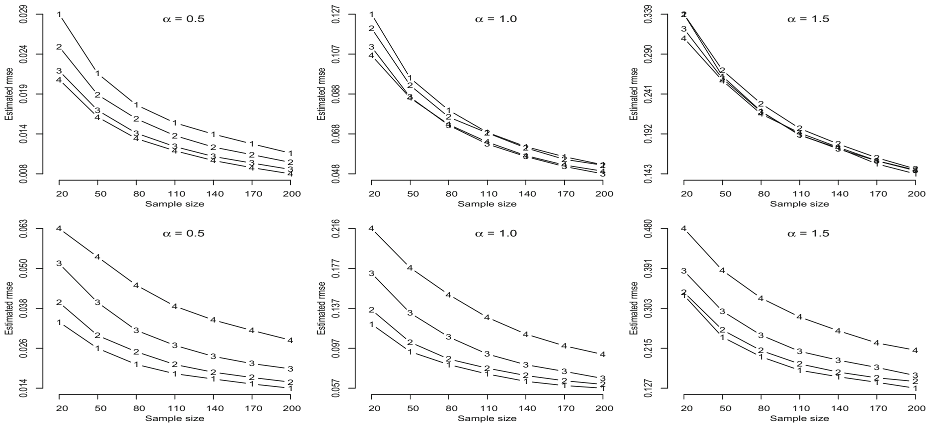

Fig. 7 Estimated RMSE for α ($1 : \theta = 0.5$, $2 : \theta = 1.0$, $3 : \theta = 1.5$ and $4 : \theta = 2.0$). Upper-panel: DQX1 distribution. Lower-panel: DQX2 distribution

6.1 Corn Borers

For the first application, we will consider a dataset on the total number of borers per hill in each plot for a control group and three treatment groups, initially analyzed by Bliss and 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 at random in each plot and the borers recorded from each hill. Here we will use the data from the second treatment (Saha 2008, Table 9). This dataset will be denoted by DS1. For the sake of comparison we have considered the Poisson (P), the Negative Binomial (NB) and COM-Poisson (COM-P) models. The pmf of the COM-P distribution is given by

$$\mathrm{P}\left(Y = y; \lambda, \nu\right) = \frac{\lambda^y}{(y!)^\nu}\left[\sum_{j=0}^{\infty}\frac{\lambda^j}{(j!)^\nu}\right]^{-1}, \quad y \in \mathbb{Z}_+,$$

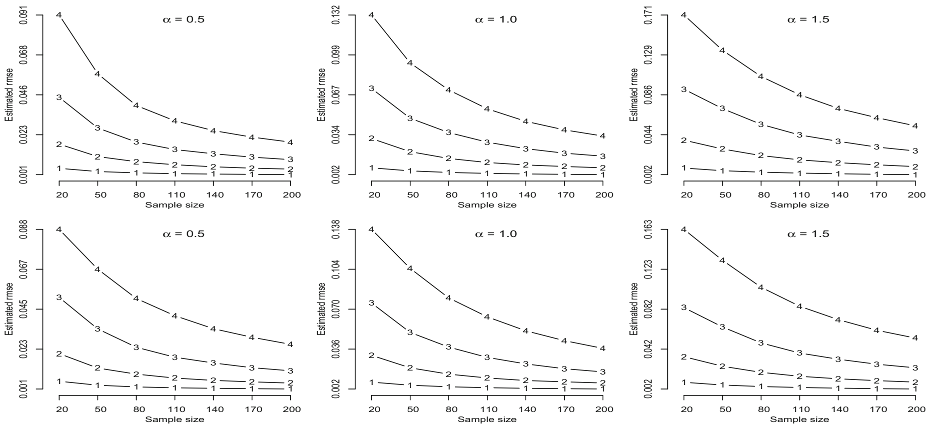

Fig. 8 Estimated RMSE for θ ($1 : \theta = 0.5$, $2 : \theta = 1.0$, $3 : \theta = 1.5$ and $4 : \theta = 2.0$). Upper-panel: DQX1 distribution. Lower-panel: DQX2 distribution

Table 3 Estimated CP and LCP under DQX1 distribution

n	CP_α	CP_θ	LCP_α	LCP_θ	CP_α	CP_θ	LCP_α	LCP_θ
	$\theta = 0.5$ and $\alpha = 0.5$				$\theta = 0.5$ and $\alpha = 1.0$			
20	0.900	0.952	0.879	0.261	0.866	0.956	1.919	0.342
50	0.917	0.960	0.738	0.208	0.896	0.952	1.657	0.275
80	0.933	0.956	0.652	0.178	0.914	0.955	1.491	0.237
110	0.937	0.960	0.588	0.158	0.918	0.955	1.360	0.211
140	0.943	0.958	0.541	0.144	0.925	0.958	1.259	0.192
170	0.948	0.955	0.503	0.133	0.929	0.956	1.177	0.178
200	0.952	0.956	0.471	0.124	0.937	0.957	1.113	0.166
	$\theta = 0.5$ and $\alpha = 1.5$				$\theta = 1.0$ and $\alpha = 0.5$			
20	0.828	0.963	3.170	0.405	0.896	0.959	0.824	0.538
50	0.866	0.950	2.817	0.332	0.918	0.960	0.686	0.427
80	0.890	0.943	2.546	0.288	0.931	0.959	0.606	0.367
110	0.904	0.943	2.365	0.259	0.939	0.955	0.547	0.326
140	0.909	0.947	2.248	0.238	0.940	0.958	0.504	0.296
170	0.915	0.953	2.117	0.221	0.946	0.958	0.469	0.274
200	0.920	0.952	2.010	0.207	0.953	0.958	0.438	0.255
	$\theta = 1.0$ and $\alpha = 1.0$				$\theta = 1.0$ and $\alpha = 1.5$			
20	0.865	0.955	1.884	0.720	0.815	0.967	3.169	0.857
50	0.887	0.956	1.649	0.583	0.852	0.948	2.832	0.709
80	0.910	0.954	1.492	0.504	0.880	0.943	2.609	0.617
110	0.913	0.956	1.368	0.449	0.890	0.944	2.441	0.557
140	0.926	0.956	1.269	0.410	0.903	0.946	2.312	0.511
170	0.931	0.955	1.186	0.379	0.907	0.948	2.189	0.475
200	0.932	0.958	1.124	0.355	0.918	0.951	2.081	0.446
	$\theta = 1.5$ and $\alpha = 0.5$				$\theta = 1.5$ and $\alpha = 1.0$			
20	0.896	0.965	0.767	0.833	0.854	0.975	1.888	1.132
50	0.916	0.962	0.647	0.660	0.889	0.961	1.640	0.910
80	0.932	0.965	0.568	0.564	0.907	0.956	1.471	0.784
110	0.935	0.963	0.512	0.500	0.921	0.955	1.351	0.701
140	0.942	0.960	0.470	0.455	0.928	0.956	1.253	0.638
170	0.947	0.961	0.436	0.420	0.935	0.963	1.178	0.590
200	0.947	0.961	0.409	0.391	0.938	0.960	1.119	0.552
	$\theta = 1.5$ and $\alpha = 1.5$				$\theta = 2.0$ and $\alpha = 0.5$			
20	0.802	0.986	3.242	1.359	0.898	0.977	0.754	1.197
50	0.848	0.955	2.862	1.118	0.920	0.961	0.626	0.940
80	0.877	0.947	2.630	0.974	0.936	0.964	0.549	0.798
110	0.891	0.949	2.451	0.877	0.939	0.963	0.496	0.706
140	0.902	0.943	2.331	0.807	0.946	0.960	0.453	0.640
170	0.910	0.946	2.212	0.750	0.950	0.957	0.420	0.589
200	0.911	0.950	2.116	0.704	0.952	0.956	0.393	0.549

Table 3 (continued)

n	CP_α	CP_θ	LCP_α	LCP_θ	CP_α	CP_θ	LCP_α	LCP_θ
	$\theta = 2.0$ and $\alpha = 1.0$				$\theta = 2.0$ and $\alpha = 1.5$			
20	0.851	0.995	1.919	1.624	0.817	0.998	3.486	1.989
50	0.884	0.979	1.647	1.294	0.836	0.984	2.986	1.608
80	0.905	0.971	1.470	1.108	0.863	0.970	2.715	1.392
110	0.912	0.966	1.348	0.984	0.879	0.962	2.517	1.243
140	0.919	0.969	1.246	0.893	0.893	0.955	2.378	1.136
170	0.926	0.970	1.167	0.824	0.899	0.954	2.236	1.052
200	0.931	0.966	1.109	0.770	0.906	0.954	2.139	0.986

Table 4 Estimated CP and LCP under DQX2 distribution

n	CP_α	CP_θ	LCP_α	LCP_θ	CP_α	CP_θ	LCP_α	LCP_θ
	$\theta = 0.5$ and $\alpha = 0.5$				$\theta = 0.5$ and $\alpha = 1.0$			
20	0.921	0.942	1.032	0.268	0.919	0.945	2.157	0.339
50	0.938	0.945	0.868	0.211	0.942	0.938	1.853	0.272
80	0.952	0.947	0.768	0.180	0.946	0.941	1.647	0.233
110	0.956	0.946	0.693	0.159	0.953	0.940	1.506	0.208
140	0.961	0.949	0.637	0.145	0.958	0.942	1.391	0.189
170	0.964	0.943	0.590	0.133	0.963	0.942	1.298	0.174
200	0.964	0.948	0.553	0.124	0.964	0.942	1.224	0.163
	$\theta = 0.5$ and $\alpha = 1.5$				$\theta = 1.0$ and $\alpha = 0.5$			
20	0.873	0.952	3.287	0.388	0.912	0.961	1.084	0.555
50	0.917	0.941	2.958	0.320	0.930	0.956	0.900	0.436
80	0.941	0.939	2.702	0.278	0.939	0.953	0.792	0.371
110	0.955	0.938	2.507	0.249	0.949	0.956	0.717	0.329
140	0.959	0.940	2.358	0.228	0.949	0.959	0.660	0.298
170	0.960	0.945	2.237	0.212	0.951	0.955	0.615	0.275
200	0.967	0.948	2.107	0.198	0.962	0.955	0.577	0.256
	$\theta = 1.0$ and $\alpha = 1.0$				$\theta = 1.0$ and $\alpha = 1.5$			
20	0.908	0.959	2.203	0.696	0.871	0.968	3.433	0.800
50	0.928	0.950	1.886	0.557	0.914	0.948	3.026	0.655
80	0.943	0.948	1.675	0.478	0.938	0.943	2.800	0.572
110	0.947	0.953	1.528	0.425	0.947	0.945	2.585	0.512
140	0.958	0.952	1.417	0.387	0.953	0.946	2.418	0.468
170	0.952	0.951	1.318	0.357	0.957	0.948	2.287	0.434
200	0.960	0.951	1.245	0.333	0.959	0.950	2.171	0.407
	$\theta = 1.5$ and $\alpha = 0.5$				$\theta = 1.5$ and $\alpha = 1.0$			
20	0.923	0.984	1.252	0.918	0.891	0.982	2.408	1.128
50	0.926	0.965	1.036	0.721	0.915	0.963	2.026	0.894
80	0.940	0.964	0.909	0.611	0.933	0.957	1.799	0.763
110	0.943	0.963	0.819	0.539	0.941	0.957	1.642	0.677
140	0.951	0.960	0.757	0.489	0.951	0.955	1.510	0.614
170	0.953	0.962	0.703	0.449	0.957	0.959	1.424	0.566
200	0.957	0.964	0.662	0.418	0.959	0.959	1.344	0.528

Table 4 (continued)

n	CP_α	CP_θ	LCP_α	LCP_θ	CP_α	CP_θ	LCP_α	LCP_θ
	$\theta = 1.5$ and $\alpha = 1.5$				$\theta = 2.0$ and $\alpha = 0.5$			
20	0.863	0.991	3.725	1.286	0.985	1.000	1.600	1.416
50	0.910	0.964	3.225	1.045	0.950	0.995	1.292	1.115
80	0.928	0.954	2.923	0.899	0.944	0.985	1.120	0.941
110	0.942	0.955	2.706	0.804	0.947	0.977	1.013	0.832
140	0.944	0.951	2.538	0.735	0.951	0.967	0.929	0.750
170	0.946	0.951	2.385	0.679	0.958	0.963	0.866	0.689
200	0.954	0.955	2.251	0.634	0.960	0.963	0.814	0.641
	$\theta = 2.0$ and $\alpha = 1.0$				$\theta = 2.0$ and $\alpha = 1.5$			
20	0.906	1.000	2.825	1.717	0.868	0.999	4.264	1.942
50	0.908	0.992	2.364	1.360	0.891	0.992	3.667	1.568
80	0.925	0.975	2.086	1.154	0.910	0.974	3.325	1.347
110	0.931	0.969	1.897	1.020	0.919	0.965	3.026	1.194
140	0.943	0.966	1.760	0.922	0.927	0.960	2.842	1.087
170	0.946	0.965	1.647	0.849	0.937	0.955	2.667	1.001
200	0.951	0.964	1.555	0.788	0.943	0.955	2.533	0.934

where $\lambda \in \mathbb{R}_+$ and $\nu \in \mathbb{R}_+ \cup \{0\}$.

Table 5 reports the MLEs, the asymptotic standard errors and some goodness-of-fit (gof) measures for each fitted model. One can notice that both versions of the discrete quasi xgamma distributions are presenting the smallest values of AIC and BIC criteria. Also, by the expected frequencies presented in Table 6, one we can conclude that the DQX1 distribution provides the best fit among all models considered here.

6.2 Outbreaks Strikes

For the second application, we will consider a dataset from the literature related to the number of outbreaks of strikes in the UK coal mining industries in four successive week

Table 5 Parameter estimates and gof measures for each model (DS1)

Model	Parameter	MLE (SE)	AIC	BIC	χ^2 (p-value)	d.f.
DQX1	α	0.4824 (0.1521)	543.61	549.18	1.23 (0.942)	5
	θ	0.6601 (0.0565)				
DQX2	α	0.5684 (0.2026)	544.34	549.91	1.80 (0.876)	5
	θ	0.6237 (0.0535)				
COM-P	λ	1.7604 (0.4023)	548.27	553.84	4.89 (0.429)	5
	ν	3.1668 (0.2718)				
NB	μ	1.0412 (0.1200)	545.98	551.56	3.41 (0.637)	5
	ϕ	0.1962 (0.0732)				
P	λ	3.1667 (0.1624)	615.01	617.80	164.55 (< 0.001)	6

Table 6 Observed and expected frequencies from each fitted model (DS1)

Counts	Observed	DQX1	DQX2	COM-P	NB	P
0	24	**23.0443**	22.1321	19.6019	20.0322	5.0573
1	16	**17.2879**	18.8819	22.1785	20.8576	16.0147
2	16	**17.2735**	17.8497	19.6742	18.9557	25.3565
3	18	**16.1094**	15.8359	15.8500	15.9099	26.7652
4	15	**13.5218**	12.9872	12.1237	12.6207	21.1892
6	15	18.0172	17.2811	**15.4781**	16.6029	20.5025
9	12	**11.0304**	10.9533	10.1648	10.7823	4.9186
$\geqslant 10$	4	3.7154	**4.0789**	4.9287	4.2388	0.1961

Estimated expected frequencies in bold relate to those closer to the observed ones

periods during the years 1948-1959 (Kendall 1961). This dataset presents underdispersion characteristics since ID ≈ 0.75. Other authors including Castillo and Pérez-Casany (1998), Ridout and Besbeas (2004) and Chakraborty and Chakravarty (2012) have also used this data for illustration purposes. This dataset will be denoted by DS2. We have also fitted the NB distribution, but it is worthwhile to mention that their maximum likelihood estimates are unique only in overdispersed case.

Table 7 reports the MLEs, the asymptotic standard errors and some gof measures for each fitted model. One can notice that both versions of the discrete quasi xgamma distributions are presenting the smallest values of AIC and BIC criteria. On the other hand, the COM-P distribution also has presented a reasonably fit. Besides, by Table 8, one we can notice that the DQX2 distribution better fits lower counts when compared with the other fitted models. Although the AIC and BIC values of the COM-P distribution are very close to those provided by the DQX1 and DQX2 distributions, from the chi-square statistic, we can affirm that the DQX2 fits better than DQX1 and COM-P models. In addition, our proposals are not defined in terms of an infinite sum as the case of the COM-P distribution.

Table 7 Parameter estimates and gof measures for each model (DS2)

Model	Parameter	MLE (SE)	AIC	BIC	χ^2 (p-value)	d.f.
DQX1	α	0.1758 (0.0368)	379.12	385.22	1.92 (0.383)	2
	θ	2.3016 (0.1507)				
DQX2	α	-0.1097 (0.0880)	378.79	384.89	1.51 (0.470)	2
	θ	2.1335 (0.1507)				
COM-P	λ	1.4830 (0.2521)	380.02	386.12	2.89 (0.236)	2
	ν	1.7686 (0.2945)				
NB	μ	0.9936 (0.0801)	388.23	394.33	10.74 (0.005)	2
	ϕ	112.8508 (187.3175)				
P	λ	0.9936 (0.0798)	388.92	387.11	10.41 (0.015)	1

Table 8 Observed and expected frequencies from each fitted model (DS2)

Counts	Observed	DQX1	DQX2	COM-P	NB	P
0	46	46.0551	**46.0280**	47.4834	58.0018	57.7583
1	76	74.0515	**74.9294**	70.4349	57.1351	57.3880
2	24	28.2635	**26.9609**	30.6592	28.3900	28.5101
3	9	6.3074	6.5022	6.5141	9.4871	**9.4424**
$\geqslant 4$	1	1.3225	1.5795	**0.9084**	2.9859	2.9012

Estimated expected frequencies in bold relate to those closer to the observed ones

7 Concluding Remarks

In this paper, two versions of the discrete quasi xgamma distribution were introduced as alternatives to model count datasets presenting overdispersion or underdispersion. To derive the proposed models, we have considered the methods of infinite series and survival function. The main statistical properties as the mean, the variance, the generating moments and the coefficients of variation, skewness, and kurtosis for each version were derived. Also, it was shown that both versions of the discrete quasi xgamma 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 the parameters of both versions. Also, we have performed an intensive 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 have indicated the suitability of the considered methodology. The usefulness of the proposed models was assessed by fitting them to real datasets provided by literature. The model selection was performed by using the AIC and the BIC criteria. The goodness-of-fit was evaluated through the χ^2 statistic. The obtained results have demonstrated that the DQX1 and DQX2 distributions are competitive with standard discrete models as the Poisson, the Negative Binomial and the COM-Poisson distributions. As a final note, we would like to mention that we are currently developing an `R` package containing a complete toolkit to fit the proposed models. In the current stage, the executable *scripts* used to fit both versions of the discrete xgamma distribution can be made available by the authors upon justified request.

Acknowledgements Josmar Mazucheli gratefully acknowledge the partial financial support from Fundação Araucária (Grant 064/2019 - UEM/Fundação Araucá). The research of Wesley Bertoli is partially supported by the Federal University of Technology - Paraná and by a doctoral grant from Fundação Araucária (CP 18/2015). The Brazilian organization CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) supports the research of Ricardo P. Oliveira (GC 001).

References

Bakouch HS, Jazi MA, Nadarajah S (2014) A new discrete distribution. Statistics 48(1):200–240
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 CI, Fisher RA (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. Int J Reliab Qual Saf Eng 10(1):69–98

Castillo JD, Pérez-Casany M (1998) Weighted poisson distributions for overdispersion and underdispersion situations. Ann Inst Stat Math 50(3):567–585

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

Chakraborty S (2015b) 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 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 RD (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, 2nd. Chapaman and Hall, New York

Doornik JA (2007) Object–oriented matrix programming using Ox, 3rd. Timberlake Consultants Press and Oxford, London

Doray LG, Luong A (1997) Efficient estimators for the Good family. Commun Stat Simul Comput 26(3):1075–1088

Gómez-Déniz E, Calderín-ojeda E (2011) The discrete Lindley distribution: properties and applications. J Stat Comput Simul 81(11):1405–1416

Good IJ (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 FA (1957) Queueing with balking. Biometrika 44(3/4):360–369

Hamada MS, Wilson AG, Reese CS, Martz HF (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 TJ (2006) A discrete analogue of the Laplace distribution. Journal of Statistical Planning and Inference 136(3):1090–1102

Jazi MA, Lai CD, Alamatsaz MH (2010) A discrete inverse Weibull distribution and estimation of its parameters. Stat Methodology 7(2):121–132

Kalbfleisch JD, Prentice RL (2002) The statistical analysis of failure time data, 2nd. Wiley, New York

Keilson J, Gerber H (1971) Some results for discrete unimodality. J Am Stat Assoc 66(334):386–389

Kemp AW (1997) Characterizations of a discrete normal distribution. Journal of Statistical Planning and Inference 63(2):223–229

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

Kemp AW (2008) The discrete half–normal distribution. In: Advances in mathematical and statistical modeling. Birkhäuser, Boston, pp 353–360

Kendall MG (1961) Natural law in social sciences. J R Stat Soc Ser A 124:1–19

Klein JP, Moeschberger ML (1997) Survival analysis: techniques for censored and truncated data. Springer, New York

Kozubowski TJ, Inusah S (2006) A skew Laplace distribution on integers. Ann Inst Stat Math 58(3):555–571

Krishna H, Pundir PS (2009) Discrete Burr and discrete Pareto distributions. Stat Methodology 6(2):177–188

Kulasekera KB, Tonkyn DW (1992) A new discrete distribution, with applications to survival, dispersal and dispersion. Commun Stat Simul Comput 21(2):499–518

Lawless JF (2003) Statistical models and methods for lifetime data, 2nd. Wiley, Hoboken

Lee ET, Wang JW (2003) Statistical methods for survival data analysis, 3rd. Wiley, Hoboken

Lisman JHC, Van Zuylen MCA (1972) Note on the generation of most probable frequency distributions. Statistica Neerlandica 26(1):19–23

Meeker WQ, Escobar LA (1998) Statistical methods for reliability data. Wiley, New York

Nakagawa T, Osaki S (1975) The discrete Weibull distribution. IEEE Trans Reliab R-24(5):300–301

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

Nekoukhou V, Alamatsaz MH, 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

Ridout MS, Besbeas P (2004) An empirical model for underdispersed count data. Stat Model 4(1):77–89
Roy D (2003) The discrete normal distribution. Communication in Statistics - Theory and Methods 32(10):1871–1883
Roy D (2004) Discrete Rayleigh distribution. IEEE Trans Reliab 53(2):255–260
Rubinstein RY, Kroese DP (2008) Simulation and the Monte Carlo method. Wiley series in probability and statistics, 2nd. Wiley, Hoboken
Saha KK (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 Trans Semicond Manuf 12(4):409–418
Sen S, Chandra N (2017) The quasi xgamma distribution with application in bladder cancer data. Journal of Data Science 15:61–76
Siromoney G (1964) The general Dirichlet's series distribution. Journal of the Indian Statistical Association 2-3(2):1–7
Slater LJ (1966) Generalized hypergeometric functions. Cambridge University Press, Cambridge