A general stochastic model for bivariate episodes driven by a gamma sequence

Charles K. Amponsah, Tomasz J. Kozubowski[*] and Anna K. Panorska

*Correspondence:
tkozubow@unr.edu
Department of Mathematics &
Statistics, University of Nevada,
Reno, USA

Abstract

We propose a new stochastic model describing the joint distribution of (X, N), where N is a counting variable while X is the sum of N independent gamma random variables. We present the main properties of this general model, which include marginal and conditional distributions, integral transforms, moments and parameter estimation. We also discuss in more detail a special case where N has a heavy tailed discrete Pareto distribution. An example from finance illustrates the modeling potential of this new mixed bivariate distribution.

Keywords: BEG model, Distribution theory, Heavy tail, Financial data, Maximum likelihood estimation, Power law, Random summation

AMS Subject Classification: 60E05; 62E10; 62E15; 62F10; 62H05; 62H12; 62P05

1 Introduction

We propose a new stochastic model describing the joint distribution of

$$(X, N) \stackrel{d}{=} \left(\sum_{i=1}^{N} X_i, N \right), \tag{1}$$

where N is an integer-valued random variable on $\mathbb{N} = \{1, 2, \ldots\}$ while the $\{X_i\}$ are independent and identically distributed (IID) random variables on $\mathbb{R}_+$, independent of N. Our new model is a generalization of the BEG and the BGG models introduced in Kozubowski and Panorska (2005) and Barreto-Souza (2012), respectively, where N was geometrically distributed with the probability mass function (PMF)

$$P(N = n) = p(1 - p)^{n-1}, \quad n \in \mathbb{N}, \tag{2}$$

and the $\{X_i\}$ were either exponential (BEG), with the probability density function (PDF)

$$f(x) = \beta e^{-\beta x}, \quad x \in \mathbb{R}_+, \tag{3}$$

or gamma distributed (BGG), with the PDF

$$f(x) = \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta x}, \quad x \in \mathbb{R}_+. \tag{4}$$

Models of this kind were shown to be useful in various areas, including finance, actuarial science, hydro-climatic studies, and others (see, e.g., Arendarczyk et al. 2018; Barreto-Souza 2012; Barreto-Souza and Silva 2019; Biondi et al. 2005, 2008; Kozubowski and Panorska 2005, 2010; Kozubowski et al. 2008, 2009 and reference therein). In modeling, the vector (X, N) in (1) is often called an "episode" with magnitude X and duration N. For example, in finance we are interested in growth/decline episodes of a certain investment, which happen when the consecutive log-returns $\{X_i\}$ are positive/negative. The sequence of N consecutive positive log-returns is called a "growth/decline episode of duration N and magnitude $X = \sum_{i=1}^{N} X_i$".

The majority of work connected with bivariate distributions given by (1) concerned cases where the distributions of X and N were both *light-tailed*, with the exception of Arendarczyk et al. (2018), where N was (light tailed) geometric and X was heavy tailed with *Lomax* distribution (also known as Pareto Type II), given by the survival function (SF)

$$S(x) = P(X > x) = \left(\frac{1}{1 + x/\sigma} \right)^{1/\delta}, \; x \in \mathbb{R}_+, \; \sigma > 0, \; \delta \geq 0. \tag{5}$$

In contrast, in our generalization studied in this paper, both X and N in (1) are heavy tailed. We achieve this by keeping the $\{X_i\}$ in (1) gamma distributed with PDF (4) as in Barreto-Souza (2012), while changing N to be a *discrete version* of the continuous Pareto distribution (5) with the same survival function (see, e.g., Buddana and Kozubowski 2014; Krishna and Singh Pundir 2009). The PMF of this *discrete Pareto* (DP) distribution can be written as

$$P(N = n) = \left(\frac{1}{1 - \delta(n-1)\log(1-p)} \right)^{\frac{1}{\delta}} - \left(\frac{1}{1 - \delta n \log(1-p)} \right)^{\frac{1}{\delta}}, \; n \in \mathbb{N}, \tag{6}$$

where $p \in (0, 1)$ is a "size" parameter and $\delta \in [0, \infty)$ is a tail parameter. We note that as $\delta \to 0^+$, then the DP distribution (6) converges to a geometric distribution (2). Thus, while our model (1) with DP distributed N offers more flexibility compared with the BEG model of Kozubowski and Panorska (2005) or the BGG model of Barreto-Souza (2012), it turns into the latter two in the special case with $\delta = 0$ and exponential or gamma distributed $\{X_i\}$ in (1).

In addition to presenting basic properties of the new model described above, we also summarize fundamental properties of the entire class of such models with gamma distributed $\{X_i\}$ and arbitrary distribution of the random variable N, which drives the dependence structure of the (X, N) vector in (1). These properties include infinite divisibility, the tail behavior of X, the conditional distributions of N given $X = x$ as well as N given $X > x$, and estimation. In particular, we establish the existence and uniqueness of maximum likelihood estimators of α and β within this class of models. Our general results concerning the conditional distributions have interesting connections with *hidden truncation* (see, e.g., Arnold) and play important role in goodness-of-fit analysis for these models.

Our article is organized as follows. In Section 2 we present fundamental properties of general class of bivariate distributions (1), driven by an IID gamma sequence $\{X_i\}$. In Section 3, we summarize basic facts about the special case with discrete Pareto distributed N, including parameter estimation. An application from finance is presented in Section 4,

which is followed by concluding remarks in Section 5. The last section contains selected proofs and technical lemmas.

2 Bivariate episodes driven by an IID gamma sequence

In this section we summarize the properties which are common to all bivariate distributions describing the episodes (1) driven by a sequence $\{X_i\}$ of IID gamma variables with the PDF (4), denoted by $\mathcal{GAM}(\alpha, \beta)$. We assume that the variable N, which describes the *duration* of an episode (X, N), is independent of the $\{X_i\}$ and supported on the set of positive integers $\mathbb{N} = \{1, 2, \ldots\}$. Clearly, given $N = n \in \mathbb{N}$, the variable X is the sum of n IID gamma variables $\mathcal{GAM}(\alpha, \beta)$, so it is also gamma distributed, with parameters αn and β. Thus, the joint PDF of (X, N) in (1) is of the form

$$f_{X,N}(x, n) = \frac{\beta^{\alpha n} x^{\alpha n - 1} e^{-\beta x}}{\Gamma(\alpha n)} f_N(n), \quad x \in \mathbb{R}_+, n \in \mathbb{N}, \tag{7}$$

where $f_N(\cdot)$ is the PMF of N. This leads to the following result, whose proof is elementary.

Proposition 1 *The cumulative distribution function (CDF) and the survival functions (SF) of (X, N) in (1) are given by*

$$F_{X,N}(x, n) = \mathbb{P}(X \leq x, N \leq n) = \sum_{k=1}^{n} \frac{\gamma(\alpha k, \beta x)}{\Gamma(\alpha k)} f_N(k), \quad x \in \mathbb{R}_+, n \in \mathbb{N},$$

and

$$S_{X,N}(x, n) = \mathbb{P}(X > x, N > n) = \sum_{k=n+1}^{\infty} \left(1 - \frac{\gamma(\alpha k, \beta x)}{\Gamma(\alpha k)}\right) f_N(k), \quad x \in \mathbb{R}_+,$$
$$n \in \mathbb{N}_0 = \{0, 1, \ldots\},$$

where $f_N(\cdot)$ is the PMF of N and

$$\gamma(\alpha, x) = \int_0^x t^{\alpha - 1} e^{-t} dt, \quad \alpha, x > 0,$$

is the (lower) incomplete gamma function.

In view of the above, we conclude that the marginal distribution of X is an infinite *mixture of gamma distributions*, as the PDF of X takes on the form

$$f_X(x) = \sum_{n=1}^{\infty} \frac{\beta^{\alpha n} x^{\alpha n - 1} e^{-\beta x}}{\Gamma(\alpha n)} f_N(n), \quad x \in \mathbb{R}_+. \tag{8}$$

Remark 1 *In the special case where N is geometric (2) we recover the **B**ivariate **G**amma-**G**eometric (BGG) distribution of Barreto-Souza (2012). In this case, the PDF of X can be written in terms of 2-parameter Mittag-Leffler (ML) special function*

$$E_{\alpha,\beta}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + \beta)}, \quad z \in C, Re(\alpha) > 0, Re(\beta) > 0, \tag{9}$$

which is an entire function on the complex plain C (see, e.g., Haubold et al. 2011). Indeed, simple algebra shows that the PDF (8) takes on the form

$$f_X(x) = \beta p (\beta x)^{\alpha - 1} e^{-\beta x} E_{\alpha,\alpha}\left((1 - p)(\beta x)^{\alpha}\right), \quad x \in \mathbb{R}_+. \tag{10}$$

*Further, in the special case $\alpha = 1$, when the $\{X_i\}$ in (1) are exponentially distributed, we obtain the BEG model of Kozubowski and Panorska (2005), where the term "BEG" stands for **B**ivariate distribution with **E**xponential and **G**eometric margins. Here, the ML special function becomes the exponential function, $E_{1,1}(z) = e^z$, and the marginal density of X in (10) reduces to the exponential PDF (3) with β replaced with $p\beta$.*

Let us also account for the joint Laplace transform $\psi(t,s)$ connected with the distribution of (X, N) in (1), where

$$\psi(t,s) = \mathbb{E}\left(e^{-tX-sN}\right), \quad t, s \in \mathbb{R}_+. \tag{11}$$

Standard conditioning argument shows that this function can be conveniently expressed in terms of the probability generating function (PGF) of the random variable N, $G_N(s) = \mathbb{E}(s^N)$, $s \in [0,1]$.

Proposition 2 *The LT of (X, N) in (1) driven by an IID gamma $\mathcal{GAM}(\alpha, \beta)$ sequence $\{X_i\}$ is given by*

$$\psi(t,s) = G_N\left\{e^{-s}\left(\frac{\beta}{\beta+t}\right)^\alpha\right\}, \quad t, s \in \mathbb{R}_+, \tag{12}$$

where $G_N(\cdot)$ is the PGF of N.

Remark 2 *When the variable N is geometric (2) where we have*

$$G_N(s) = \frac{ps}{1-(1-p)s}, \quad s \in [0,1],$$

then (12) turns into the LT of the BGG model of Barreto-Souza (2012),

$$\psi(t,s) = \frac{p\beta^\alpha e^{-s}}{(\beta+t)^\alpha - (1-p)\beta^\alpha e^{-s}}, \quad t, s \in \mathbb{R}_+.$$

Remark 3 *As can be seen from the proof of Proposition 2, if the IID variables $\{X_i\}$ are non-negative but not necessarily gamma distributed, then the LT in (12) is given by $\psi(t,s) = G_N[e^{-s}\psi_{X_1}(t)]$, where $\psi_{X_1}(\cdot)$ is the LT of the $\{X_i\}$.*

2.1 Moments and tail behavior

Here we consider moments and the tail behavior of the variables X and N. Clearly, if the counting variable N is light tailed, and all of its moments $\mathbb{E}[N^\eta]$ exist for $\eta > 0$, then the mixed moments $\mathbb{E}[X^r N^\eta]$ exist as well for any $r, \eta > 0$. However, this is not so if the distribution of N is heavy tailed, and its survival function behaves like a power law,

$$S_N(x) = \mathbb{P}(N > x) \sim cx^{-\nu} \quad (x \to \infty), \tag{13}$$

where c is a positive constant and $f(x) \sim g(x)$ $(x \to \infty)$ means that the ratio $f(x)/g(x)$ converges to 1 as $x \to \infty$. Indeed, if the tail probability satisfies (13) for some $\nu > 0$, then for positive $\eta > 0$ the expectation $\mathbb{E}[N^\eta]$ exists only for $\eta < \nu$. Moreover, in this case of heavy tailed N, the variable X in (1) is heavy tailed as well, with essentially the same tail behavior as N.

Proposition 3 *If the variable N in (1) satisfies (13) with some $c, \nu > 0$ and the $\{X_i\}$ in (1) are IID gamma distributed $\mathcal{GAM}(\alpha, \beta)$, then we have*

$$S_X(x) = \mathbb{P}(X > x) \sim c(\alpha/\beta)^\nu x^{-\nu} \quad (x \to \infty).$$

Under the conditions of Proposition 3, the moments $\mathbb{E}[X^r]$ with $r > 0$ are finite only for $r < \nu$. More generally, standard calculations involving either the joint PDF (7) or conditioning arguments viz. the representation (1) show that the joint moments $\mathbb{E}[X^r N^\eta]$ with positive $r, \eta > 0$ exist only for $r + \eta < \nu$, and their values are provided in the following result.

Proposition 4 *Under the conditions of Proposition 3, the moments $\mathbb{E}[X^r N^\eta]$ with $r, \eta > 0$ are finite only for $r + \eta < \nu$, in which case we have*

$$\mathbb{E}[X^r N^\eta] = \frac{1}{\beta^r} \sum_{n=1}^{\infty} n^\eta \frac{\Gamma(r + \alpha n)}{\Gamma(\alpha n)} f_N(n), \tag{14}$$

where $f_N(\cdot)$ is the PMF of N.

If they exist, then the mean vector and the covariance matrix of (X, N) in (1) can be deduced from Proposition 4. The result below provides relevant information concerning this.

Proposition 5 *If the variable N in (1) has a finite mean $\mathbb{E}[N]$, then so does X, and we have $\mathbb{E}[X] = (\alpha/\beta)\mathbb{E}[N]$. Moreover, if the variance $\mathbb{V}ar[N]$ is finite, then so is the covariance matrix $\mathbf{\Sigma}$ of (X, N), and we have*

$$\mathbf{\Sigma} = \mathbb{E}[N] \begin{bmatrix} \alpha/\beta^2 & 0 \\ 0 & 0 \end{bmatrix} + \mathbb{V}ar[N] \begin{bmatrix} \alpha^2/\beta^2 & \alpha/\beta \\ \alpha/\beta & 1 \end{bmatrix}.$$

Remark 4 *We note that if N is geometric (2), so that $\mathbb{E}[N] = 1/p$ and $\mathbb{V}ar[N] = (1 - p)/p^2$, we recover the mean vector and the covariance matrix of the BGG model of Barreto-Souza (2012).*

Remark 5 *If the variance of N is finite, then the correlation of X and N is a simple function of the parameter α and the index of dispersion of N,*

$$\rho = \frac{1}{\sqrt{1 + \frac{1}{\alpha} \frac{\mathbb{E}[N]}{\mathbb{V}ar[N]}}},$$

which can fall anywhere in the interval $(0, \infty)$. In the BGG case, this reduces to $\rho = \sqrt{(1 - p)/(1 - p + p/\alpha)}$, which can be either greater than or less than $\sqrt{1 - p}$, as noted by Barreto-Souza (2012). The latter is the correlation in the BEG model, arising in the special case $\alpha = 1$ of the BGG.

2.2 Conditional distributions

The derivation of basic conditional distributions in this bivariate model is straightforward. Clearly, the conditional distribution of X given $N = n$ is $\mathcal{GAM}(\alpha n, \beta)$, with the PDF

$$f_{X|N=n}(x) = \frac{\beta^{\alpha n}}{\Gamma(\alpha n)} x^{\alpha n - 1} e^{-\beta x}, \quad x \in \mathbb{R}_+. \tag{15}$$

Similarly, the PMF of the conditional distribution of N given $X = x$ can be obtained by dividing the joint PDF of (X, N) in (7) by the marginal PDF of X, given by (8). Interestingly, this conditional PMF is proportional to the product of two PMFs as follows:

$$f_{N|X=x}(n) = (1/c) f_N(n) f_{\alpha, \beta x}(n - 1), \quad n \in \mathbb{N}, \tag{16}$$

where $c > 0$ is a normalizing constant that does not involve n, $f_N(\cdot)$ is the PMF of N, and

$$f_{\alpha,\lambda}(n) = \frac{\lambda^{\alpha(n+1)}}{\Gamma(\alpha(n+1))} \frac{\lambda^{-\alpha}}{E_{\alpha,\alpha}(\lambda^\alpha)}, \quad n \in \mathbb{N}_0. \tag{17}$$

The function $E_{\alpha,\alpha}(\cdot)$ in (17) is the Mittag-Leffler special function defined in (9) with $\alpha = \beta > 0$. The function in (17) is a genuine probability distribution on the set of non-negative integers $\mathbb{N}_0$ for any choice of $\alpha, \lambda > 0$, and it reduces to the classical Poisson distribution with mean λ when $\alpha = 1$. We shall denote the distribution with the PMF (17) by $\mathcal{EP}(\alpha, \lambda)$, which stands for **E**xtended **P**oisson.

Remark 6 *This distribution is an example of a "power-series" type distribution (see, e.g., Noack 1950), where we used the power series of the Mittag-Leffler special function (9) with $\beta = \alpha$ and $z = \lambda^\alpha$. A similar generalization of Poisson distribution, referred to as "hyper-Poisson" by the authors, was considered in Bardwell and Crow (1964), where the parameter α in (9) was taken to be $\alpha = 1$, leading to the PMF*

$$g_{\beta,\lambda}(n) = \frac{\lambda^n}{\Gamma(n+\beta)} \frac{1}{E_{1,\beta}(\lambda)}, \quad n \in \mathbb{N}_0.$$

This reduces to the Poisson PMF with parameter λ when $\beta = 1$.

Let us note an important stochastic interpretation of an integer-valued random variable Y with PMF proportional to the product of two PMFs.

Lemma 1 *If we have that*

$$f_Y(n) = (1/c)f_N(n)f_Q(n), \quad n \in \mathbb{N}_0, \tag{18}$$

where $f_Y(\cdot)$, $f_N(\cdot)$, and $f_Q(\cdot)$ are the PMFs of Y, N, and Q, respectively, then we have the representation

$$Y \stackrel{d}{=} N|N = Q, \tag{19}$$

where the variables N and Q are mutually independent.

In view of the above lemma, we get the following result regarding the conditional distribution of N given $X = x$.

Proposition 6 *Let (X, N) have a joint distribution with the PDF (7). Then the conditional distribution of N given $X = x$ is the same as that of the random variable Y in (19), where N and Q are independent and $Q - 1$ has extended Poisson distribution $\mathcal{EP}(\alpha, \lambda)$ with $\lambda = \beta x$, given by the PMF (17).*

Remark 7 *We note that in the special case of the BGG model, where N is geometric (2), the PDF of N given $X = x$ in (16) is proportional to $[(1-p)^{1/\alpha}\beta x]^{\alpha n} / \Gamma(\alpha n)$. We conclude that this conditional distribution is the same as that of $1 + R$, where $R \sim \mathcal{EP}(\alpha, \lambda)$ with $\lambda = (1-p)^{1/\alpha}\beta x$. In the special case of the BEG model, where we have $\alpha = 1$, the variable R is Poisson with mean $\lambda = (1-p)\beta x$, and $N|X = x$ is R shifted up by 1.*

The information about the conditional distribution of X given $N = n$ provided above is useful in goodness-of-fit analysis for this bivariate model. Indeed, given data (X_i, N_i) from the model, one can isolate the pairs with a particular value of $N_i = n$ and check the fit

of the theoretical conditional distribution of $X|N = n$ against the empirical distribution of the $\{X_i|N_i = n\}$. Unfortunately, one cannot proceed in a similar fashion using the conditional distribution of N given $X = x$, since the distribution of X is continuous. With this limitation in mind, we now derive the conditional distribution of N given $X > x$, which will be more useful in goodness-of-fit analysis.

Observe that the PMF of this conditional distribution can be written as

$$\mathbb{P}(N = n|X > x) = \frac{\mathbb{P}(N = n, X > x)}{\mathbb{P}(X > x)} = \frac{H(n)f_N(n)}{\mathbb{P}(X > x)}, \ \ n \in \mathbb{N}, \tag{20}$$

where

$$H(n) = \int_x^\infty \frac{\beta^{\alpha n}}{\Gamma(\alpha n)} t^{\alpha n-1} e^{-\beta t} dt = \mathbb{P}(X_1 + \cdots + X_n > x) \tag{21}$$

and the $\{X_i\}$ are IID gamma $\mathcal{GAM}(\alpha, \beta)$ distributed random variables. This is because the integrand above is the PDF of the sum $X_1 + \cdots + X_n$. Consider a renewal process $\{N(x), x \in R_+\}$ defined through the sequence $\{X_i\}$, so that

$$N(x) \geq n \ \text{ if and only if } \ X_1 + \cdots + X_n \leq x, \ \ n \in \mathbb{N}, x \in \mathbb{R}_+.$$

Then, the probability on the right-hand-side of (21) coincides with the probability $\mathbb{P}(N(x) + 1 \leq n)$, which in turn is the same as the $H(n)$ in (20), with the latter taking the form

$$H(n) = \mathbb{P}(N(x) + 1 \leq n) = \frac{\Gamma(\alpha n, \beta x)}{\Gamma(\alpha n)}, \ \ n \in \mathbb{N}. \tag{22}$$

The function in the numerator on the right-hand-side in (22) is the (upper) incomplete gamma function,

$$\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1} e^{-t} dt, \ \ \alpha, x > 0.$$

Thus, the function $H(n)$ above is a genuine CDF of a discrete random variable $N(x) + 1$. In turn, in view of (20), the conditional PMF of N given $X > x$ is proportional to the product of the CDF $H(n)$ and the PMF $f_N(n)$. Distributions with such a structure have an important stochastic representation related to the *hidden truncation* (see, e.g., Arnold, 2009), given below.

Lemma 2 *If we have that*

$$f_Y(n) = (1/c)f_N(n)F_Q(n), \ \ n \in \mathbb{N}_0,$$

where $f_Y(\cdot)$ and $f_N(\cdot)$ are the PMFs of Y and N, respectively, while $F_Q(\cdot)$ is the CDF of a random variable Q, then we have the representation

$$Y \overset{d}{=} N|N \geq Q, \tag{23}$$

where the variables N and Q are mutually independent.

In view of the above lemma, we have the following result concerning the conditional distribution of N given $X > x$ in our model.

Proposition 7 *Let (X, N) have a joint distribution with the PDF (7). Then the conditional distribution of N given $X > x$ is the same as that of the random variable Y in (23)*

with N and Q independent and $Q \overset{d}{=} N(x) + 1$, where $\{N(x), x \in R_+\}$ is a renewal process defined through an IID gamma $\mathcal{GAM}(\alpha, \beta)$ sequence $\{X_i\}$.

We note that in the special case where $\alpha = 1$, so that the variables $\{X_i\}$ in (1) are exponential with parameter β, the renewal process $\{N(x), x \in R_+\}$ is actually a Poisson process, with rate β. If in addition the variable N is geometrically distributed with the PMF (2), the distribution of (X, N) in (1) turns into the BEG model studied by Kozubowski and Panorska (2005). We have the following new result concerning the conditional distribution of N given $X > x$ in this special case.

Proposition 8 *Let (X, N) have a joint distribution with the PDF (7) where $\alpha = 1$ and N is geometric, so that $f_N(n) = p(1-p)^{n-1}$, $n \in \mathbb{N}$. Then the conditional distribution of N given $X > x$ is the same as that of the random variable $Y = N + T$, where T is Poisson with parameter $(1-p)\beta x$, independent of N.*

The above results are useful in goodness-of-fit analysis when our bivariate distribution is fitted to data. Indeed, given data (X_i, N_i) from the model, one can isolate the pairs where $X_i > x$ for a particular $x > 0$, and check the fit of the theoretical conditional distribution of $N|X > x$ against the empirical distribution of the $\{N_i\}$ that are part of these pairs. One simple strategy is to compare certain characteristics, such as the mean, the mode, or the variance, of the theoretical and the empirical distributions across different values of x. For example, in the BEG model, by Proposition 8, the conditional expectation

$$\mathbb{E}(N|X > x) = \mathbb{E}N + \mathbb{E}T = 1/p + (1-p)\beta x$$

is a simple linear function of x, which can be plotted against x and compared with the corresponding plot of the sample mean of $N|X > x$ for the same x. Then the goodness-of-fit of the BEG model may be judged by comparing the two plots.

We finally note that the conditional distribution of N given $X > x$ studied above is one of the two marginal distributions in the bivariate model $(X, N)|X > x$. The full joint distribution of this model can be described by the joint PDF, which can be shown to be of the form

$$g(t, n) = g(t|n)f_Y(n), \quad n \in \mathbb{N}, t \in \mathbb{R}_+, \tag{24}$$

where $f_Y(\cdot)$ is the PMF (20) of the distribution of $N|X > x$ and

$$g(t|n) = \frac{\beta^{\alpha n}}{\Gamma(\alpha n)} t^{\alpha n - 1} e^{-\beta t} I_{(x,\infty)}(t)/H(n), \quad t \in \mathbb{R}_+, \tag{25}$$

with $H(n)$ as in (21), is the PDF of gamma distribution $\mathcal{GAM}(\alpha n, \beta)$ truncated below at x. The joint CDF of this distribution can be written in terms of (upper) incomplete gamma function as follows:

$$\begin{aligned} G(y, n) &= \mathbb{P}(X \le y, N \le n|X > x) \\ &= \frac{\sum_{j=1}^{n} \left\{ [\Gamma(\alpha j, \beta x) - \Gamma(\alpha j, \beta y)] / \Gamma(\alpha j) \right\} f_N(j)}{\sum_{j=1}^{\infty} \left\{ \Gamma(\alpha j, \beta x) / \Gamma(\alpha j) \right\} f_N(j)}, \quad x \le y, n \in \mathbb{N}. \end{aligned}$$

Observe that, in view of (24), while the PDF of this distribution admits a hierarchical structure in the spirit of (7), this conditional distribution generally does not have a stochastic representation analogous to (1). On the other hand, the bivariate distribution

describing $(X, N)|N > u$, where $u \in \mathbb{N}_0$, can actually be related to (1). Indeed, it is easy to see that the joint PDF of this distribution is of the form

$$f_{X,\tilde{N}}(x, n) = \frac{\beta^{\alpha n}}{\Gamma(\alpha n)} x^{\alpha n - 1} e^{-\beta x} f_{\tilde{N}}(n), \quad x \in \mathbb{R}_+, n \in \mathbb{N}, \tag{26}$$

where

$$f_{\tilde{N}}(n) = \begin{cases} f_N(n)/\mathbb{P}(N > u) & \text{for } n = u + 1, u + 2, \dots \\ 0 & \text{for } n \leq u \end{cases} \tag{27}$$

is the PMF of $\tilde{N} \stackrel{d}{=} N|N > u$. Thus, this conditional distribution admits the representation (1) with N replaced by $\tilde{N}$. For any $x \in \mathbb{R}_+$ and $n \geq u$, the joint survival function of this distribution takes on the form

$$S_{X,\tilde{N}}(x, n) = \mathbb{P}(X > x, N > n|N > u) = \sum_{k=n+1}^{\infty} \left(1 - \frac{\gamma(\alpha k, \beta x)}{\Gamma(\alpha k)}\right) f_{\tilde{N}}(k) \tag{28}$$

with $f_{\tilde{N}}(\cdot)$ provided in (27). We note the following stochastic representation of this distribution involving the *excess* random variable $N_u \stackrel{d}{=} N - u|N > u$, defined for any $u \in \mathbb{N}_0$, which, like N, is also supported on the set of positive integers $\mathbb{N}$.

Proposition 9 *Let (X, N) have a joint distribution with the PDF (7) and let $N_u \stackrel{d}{=} N - u|N > u$ where $u \in \mathbb{N}_0$. Then we have the following stochastic representation:*

$$(X, N)|N > u \stackrel{d}{=} \left(\sum_{j=1}^{u} X_{1j}, u\right) + \left(\sum_{j=1}^{N_u} X_{2j}, N_u\right),$$

where the $\{X_{ij}\}$ are IID gamma $\mathcal{GAM}(\alpha, \beta)$ distributed random variables, independent of N_u.

Thus, if the support of N is the set of positive integers, then the conditional distribution of $(X, N)|N > u$ can be expressed as a convolution of two distributions with similar structure, one of which is degenerate and the other is also connected with a discrete distribution supported on $\mathbb{N}$. In particular, if N is geometrically distributed with the PMF (2), then due to its memoryless property, we have that $N_u \stackrel{d}{=} N$ for each $u \in \mathbb{N}_0$, so that

$$(X, N)|N > u \stackrel{d}{=} (G_u, u) + (X, N), \tag{29}$$

where G_u is a gamma $\mathcal{GAM}(\alpha u, \beta)$ distributed random variable, independent of (X, N).

We finish this section with a brief account for two other related conditional distributions, one of $(X, N)|X \leq x$, and another of $(X, N)|N \leq u$. The following result, which is straightforward to derive, can be found in Amponsah (2017) (see equations (2.36) and (2.38) therein) in its special case of discrete Pareto random variable N.

Proposition 10 *Let $u, n \in \mathbb{N}$ and $x > 0$, with $n \leq u$. Then the conditional CDF of (X, N) given $N \leq u$ is given by*

$$\mathbb{P}(X \leq x, N \leq n|N \leq u) = \frac{\sum_{j=1}^{n} \frac{\gamma(\alpha j, \beta x)}{\Gamma(\alpha j)} f_N(j)}{\mathbb{P}(N \leq u)}. \tag{30}$$

Similarly, for $0 < y \leq x$ *and* $n \in \mathbb{N}$, *the conditional CDF of* (X, N) *given* $X \leq x$ *is given by*

$$\mathbb{P}(X \leq y, N \leq n | X \leq x) = \frac{\sum\limits_{j=1}^{n} \frac{\gamma(\alpha j, \beta y)}{\Gamma(\alpha j)} f_N(j)}{\sum\limits_{j=1}^{\infty} \frac{\gamma(\alpha j, \beta x)}{\Gamma(\alpha j)} f_N(j)}. \tag{31}$$

2.3 Divisibility properties

It was shown in Barreto-Souza (2012) and Kozubowski and Panorska (2005) that the BGG and BEG models are *infinitely divisible* (ID). Below, we show that this is a general property of this model, as long as the ID property holds for N as well as the underlying IID sequence $\{X_i\}$.

Recall that a random vector $\mathbf{X}$ (and its probability distribution) is said to be ID if for each $n \in \mathbb{N}$ the vector $\mathbf{X}$ can be "decomposed" into the sum

$$\mathbf{X} \overset{d}{=} \mathbf{X}_1^{(n)} + \cdots + \mathbf{X}_n^{(n)},$$

where the random vectors $\{\mathbf{X}_i^{(n)}\}$ are IID. We assume that the non-negative variables $\{X_i\}$ in (1) are ID, so that $X_1 \overset{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}$ for some IID sequence $\{Y_i^{(n)}\}$ for each $n \in \mathbb{N}$. In this situation, the variables $Y_i^{(n)}$ are also non-negative and their common LT is given by $\psi_n(t) = [\psi_{X_1}(t)]^{1/n}$, where $\psi_{X_1}(\cdot)$ is the LT of the $\{X_i\}$. We also assume that the variable $N - 1$ is *discrete* ID, so that the relation $N - 1 \overset{d}{=} N_1^{(n)} + \cdots + N_n^{(n)}$ holds for each $n \in \mathbb{N}$ with IID variables $\{N_i^{(n)}\}$ whose distribution is supported on the set $\mathbb{N}_0$ of non-negative integers. This implies that N itself is ID as well, and can be decomposed into the sum of n IID terms of the form $1/n + N_i^{(n)}$, $i = 1, \ldots n$. Moreover, the common PGF of the $\{N_i^{(n)}\}$ is of the form $G_n(s) = [G_N(s)/s]^{1/n}$, where $G_N(\cdot)$ is the PGF of N. Under this notation, we have the following result.

Proposition 11 *Let N be an integer-valued random variable supported on $\mathbb{N}$, given by the PGF $G_N(\cdot)$. Further, suppose that (X, N) is defined viz. (1), where the $\{X_i\}$ are nonnegative, IID, and ID random variables, given by LT $\psi_{X_1}(\cdot)$, independent of N. Suppose also that the distribution of $N - 1$ is discrete ID. Then, the distribution of (X, N) is ID and for each $n \in \mathbb{N}$ we have the stochastic representation*

$$(X, N) \overset{d}{=} \sum_{j=1}^{n} (R_j^{(n)}, V_j^{(n)}), \tag{32}$$

where the $\{(R_j^{(n)}, V_j^{(n)})\}$ are IID. Moreover, we have $V_j^{(n)} \overset{d}{=} 1/n + T_j^{(n)}$, where the $\{T_j^{(n)}\}$ are IID, integer-valued random variables with the PGF $G_n(s) = [G_N(s)/s]^{1/n}$, and

$$R_j^{(n)} = \sum_{j=1}^{T_j^{(n)}} X_{ij} + G_j^{(n)}, \tag{33}$$

where all the variables on the right-hand-side of (33) are mutually independent, the $\{G_j^{(n)}\}$ are IID non-negative variables given by the LT $\psi_n(t) = [\psi_{X_1}(t)]^{1/n}$, and the $\{X_{ij}\}$ are IID with the LT $\psi_{X_1}(\cdot)$.

When we restrict the above result to the case where the variables $\{X_i\}$ are gamma $\mathcal{GAM}(\alpha, \beta)$ distributed (and thus ID), we obtain the following generalizations of the results in Barreto-Souza (2012).

Corollary 1 *Let N be an integer-valued random variable supported on $\mathbb{N}$, given by the PGF $G_N(\cdot)$. Further, suppose that (X, N) is defined viz. (1), where the $\{X_i\}$ are IID gamma $\mathcal{GAM}(\alpha, \beta)$ distributed, and independent of N. Suppose also that the distribution of $N - 1$ is discrete ID. Then, the distribution of (X, N) is ID and for each $n \in \mathbb{N}$ we have the stochastic representation (32) where the $\{(R_j^{(n)}, V_j^{(n)})\}$ are IID. Moreover, we have $V_j^{(n)} \stackrel{d}{=} 1/n + T_j^{(n)}$, where the $\{T_j^{(n)}\}$ are IID, integer-valued random variables with the PGF $G_n(s) = [G_N(s)/s]^{1/n}$, and the $\{R_j^{(n)}\}$ are given by (33), where all the variables on the right-hand-side of (33) are mutually independent, the $\{G_j^{(n)}\}$ are IID gamma $\mathcal{GAM}(\alpha/n, \beta)$ distributed, and the $\{X_{ij}\}$ are IID gamma $\mathcal{GAM}(\alpha, \beta)$ distributed.*

Remark 8 *In Kozubowski and Panorska (2005) and Barreto-Souza (2012), the variable N was geometric (2) while the $\{X_i\}$ were exponentially or gamma distributed, respectively. In this case, the variables $T_j^{(n)}$ in (33) have negative binomial $\mathcal{NB}(r, p)$ distribution with $r = 1/n$, whose PDF and PGF are given by*

$$f(n) = \binom{r + n - 1}{n} p^r (1 - p)^n, \ \ n \in \mathbb{N}_0,$$

and

$$G(s) = \left(\frac{p}{1 - (1 - p)s} \right)^r, \ \ s \in [0, 1],$$

respectively. We note that in the gamma case the LT of the $\{(R_j^{(n)}, V_j^{(n)})\}$ in (32) is of the form $[\psi(t, s)]^{1/n}$, where $\psi(t, s)$ is the LT of (X, N), given by (12). More generally, the ID property of the distribution of (X, N) established above implies that $[\psi(t, s)]^r$ is a genuine LT for any $r \in \mathbb{R}_+$, and describes the marginal distribution of a bivariate Lévy motion $\{(R(r), V(r)), r \in \mathbb{R}_+\}$, which is a stochastic process with independent and stationary increments started at the origin, built upon the distribution of (X, N). Such processes connected with the BEG and BGG models were studied in Kozubowski et al. (2008) and Barreto-Souza (2012), respectively. Their analogs in the more general case described above can be studied along the same lines.

2.4 Parameter estimation

In this section we consider maximum likelihood estimation of the parameters of the model (1) where the $\{X_i\}$ are IID with gamma $\mathcal{GAM}(\alpha, \beta)$ distribution. We shall assume that the distribution of N depends on k unknown parameters $\omega_1, \ldots, \omega_k$, where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_k) \in \Omega \subset \mathbb{R}^k$. Under this notation, our model has $k + 2$ unknown parameters $\theta_i, i = 1, \ldots, k + 2$, where $\theta_1 = \alpha$, $\theta_2 = \beta$, and $\theta_i = \omega_{i-2}$ for $i = 3, \ldots, k + 2$, with $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{k+2}) \in \Theta = \mathbb{R}_+^2 \times \Omega \subset \mathbb{R}^{k+2}$. Under this notation, the PMF of the random variable N will be denoted by $f_N(\cdot; \boldsymbol{\omega})$.

Let us start with the Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ corresponding to this bivariate model. Note that the logarithm of the joint PDF in (7),

$$\log f_{X,N}(x, n) = \alpha n \log \beta + (\alpha n - 1) \log x - \beta x - \log \Gamma(\alpha n) + \log f_N(n; \boldsymbol{\omega}),$$

is the sum of two terms,

$$\nu_1(x, n; \alpha, \beta) = \alpha n \log \beta + (\alpha n - 1) \log x - \beta x - \log \Gamma(\alpha n) \quad \text{and} \quad \nu_2(n; \boldsymbol{\omega}) = \log f_N(n; \boldsymbol{\omega}),$$

with each of them depending on different parameters. Consequently, if standard regularity conditions for the family $\{f_N(\cdot; \boldsymbol{\omega}), \boldsymbol{\omega} \in \Omega\}$ are satisfied and appropriate expectations exist, then the Fisher information $\mathbf{I}(\boldsymbol{\theta})$ will have a block-diagonal structure,

$$I(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}(\boldsymbol{\omega}) \end{bmatrix}, \tag{34}$$

where $\mathbf{W} = [w_{ij}]$ is a 2 by 2 matrix where

$$w_{ij} = -\mathbb{E}\left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \nu_1(X, N; \theta_1, \theta_2)\right], \quad i, j = 1, 2, \tag{35}$$

and $\mathbf{J}(\boldsymbol{\omega})$ is the Fisher information matrix connected with the family $\{f_N(\cdot; \boldsymbol{\omega}), \boldsymbol{\omega} \in \Omega\}$. Straightforward algebra shows that the entries of $\mathbf{W}$ are as follows:

$$w_{11} = \mathbb{E}[N^2 \psi'(\alpha N)], \quad w_{12} = w_{21} = -\mathbb{E}[N]/\beta, \quad w_{22} = \alpha \mathbb{E}[N]/\beta^2, \tag{36}$$

where $\psi(x) = d/dx \log \Gamma(x)$ is the digamma function.

Remark 9 *In the special case of the BGG model studied in Barreto-Souza (2012) the variable N is geometric (2) so that $k = 1$ and $\omega_1 = p \in \Omega = (0, 1)$. Here, $\mathbf{J}(\boldsymbol{\omega})$ in (34) is a 1 by 1 matrix $[1/(p^2(1-p))]$ and $\mathbb{E}[N] = 1/p$, so that by (36) we have*

$$I(\boldsymbol{\theta}) = \frac{1}{p} \begin{bmatrix} w(\alpha, p) & \frac{-1}{\beta} & 0 \\ \frac{-1}{\beta} & \frac{\alpha}{\beta^2} & 0 \\ 0 & 0 & \frac{1}{p}\frac{1}{1-p} \end{bmatrix},$$

where $w(\alpha, p) = p^2 \sum_{n=1}^{\infty} n^2 (1-p)^{n-1} \psi'(\alpha n)$ (cf., Barreto-Souza 2012, eq. (15), p. 135).

Next, we let $(X_1, N_1), \ldots, (X_n, N_n)$ be a random sample from the bivariate model with the PDF (7). By straightforward calculations, we find the log-likelihood function to be of the form

$$l(\boldsymbol{\theta}) = R(\alpha, \beta) + Q(\boldsymbol{\omega}) - \sum_{i=1}^{n} \log(X_i), \tag{37}$$

where

$$R(\alpha, \beta) = n\alpha \bar{N}_n \log(\beta) - n\beta \bar{X}_n + \alpha \sum_{i=1}^{n} N_i \log(X_i) - \frac{1}{n} \sum_{i=1}^{n} \log[\Gamma(\alpha N_i)], \tag{38}$$

$$Q(\boldsymbol{\omega}) = \sum_{i=1}^{n} \log[f_N(N_i; \boldsymbol{\omega})], \tag{39}$$

and $\bar{N}_n$, $\bar{X}_n$ are the sample means of the $\{N_i\}$ and $\{X_i\}$, respectively. In order to find the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$, we need to maximize the log-likelihood function (37) with respect to $\boldsymbol{\theta} \in \Theta = \mathbb{R}_+^2 \times \Omega \subset \mathbb{R}^{k+2}$. Since $\boldsymbol{\theta} = (\alpha, \beta, \boldsymbol{\omega})$, by the special structure of the log-likelihood function (37), we can proceed by maximizing the functions $R(\alpha, \beta)$ and $Q(\boldsymbol{\omega})$ with respect to their arguments independently of each other. Consequently, the MLE of $\boldsymbol{\omega}$ is only dependent on the $\{N_i\}$, and is connected with the family $\{f_N(\cdot; \boldsymbol{\omega}), \boldsymbol{\omega} \in \Omega\}$. Additionally, the MLEs of α and β will be of the same form regardless of the particular distribution of N, although their asymptotic properties may depend on the latter.

2.4.1   The MLEs of α and β

Our focus are the MLEs of α and β. By the above discussion, to find them one needs to maximize the function $R(\alpha, \beta)$ in (38) with respect to $(\alpha, \beta) \in \mathbb{R}_+^2$. To do this, we first maximize the function $h(\beta) = R(\alpha, \beta)$ with respect to $\beta \in \mathbb{R}_+$ when $\alpha > 0$ is held fixed. Since $h(\beta)$ is a continuous, differentiable function of β with the derivative of the form

$$h'(\beta) = \frac{n\alpha\bar{N}_n}{\beta} - n\bar{X}_n, \ \ \beta \in \mathbb{R}_+,$$

it is easy to see that $h'(\beta) > 0$ iff $\beta < \beta(\alpha)$ and $h'(\beta) < 0$ iff $\beta > \beta(\alpha)$, where

$$\beta(\alpha) = \frac{\alpha\bar{N}_n}{\bar{X}_n}. \tag{40}$$

Thus, the function $h(\beta)$ has a unique maximum value on $(0, \infty)$, which is attained at $\beta = \beta(\alpha)$ given by (40). Next, we optimize the function $R(\alpha, \beta(\alpha))$ with respect to $\alpha \in \mathbb{R}_+$. Straightforward algebra shows that $R(\alpha, \beta(\alpha)) = n\bar{N}_n g(\alpha)$, where for $\alpha \in \mathbb{R}_+$ we have

$$g(\alpha) = \alpha \log(\alpha) + \alpha \log\left(\frac{\bar{N}_n}{\bar{X}_n}\right) - \alpha + \frac{\alpha}{n\bar{N}_n} \sum_{i=1}^{n} N_i \log(X_i) - \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} \log\left[\Gamma(\alpha N_i)\right].$$

This is a continuous, differentiable function on $(0, \infty)$, with the derivative of the form

$$g'(\alpha) = \log(\alpha) - \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} N_i \psi\left[\alpha N_i\right] + \log\left(\frac{\bar{N}_n}{\bar{X}_n}\right) + \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} N_i \log(X_i), \tag{41}$$

where $\psi(x) = d/dx \log \Gamma(x)$ is the digamma function. The following result provides key properties of the above function

Proposition 12 *Let* $(X_1, N_1), \ldots, (X_n, N_n)$ *be IID copies of* (X, N) *in (1). Then, the function (41) has the following properties:*

(i) $g'(\alpha)$ *is continuous and monotonically decreasing in* $\alpha \in (0, \infty)$;
(ii) $\lim_{\alpha \to 0^+} g'(\alpha) = \infty$;
(iii) $\lim_{\alpha \to \infty} g'(\alpha) = \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} N_i \left[\log\left(\frac{X_i}{N_i}\right)\right] - \log\left(\frac{\bar{X}_n}{\bar{N}_n}\right) \leq 0$;
(iv) *For* $n = 1$, $\lim_{\alpha \to \infty} g'(\alpha) = 0$;
(v) *For* $n \geq 2$, $\lim_{\alpha \to \infty} g'(\alpha) < 0$ *almost surely.*

In view of the above result, we conclude that whenever the limit in Part (iii) is negative, which occurs with probability 1 for any finite sample of size $n \geq 2$, the function $g(\alpha)$ in monotonically increasing on the interval $(0, \hat{\alpha}_n)$ and monotonically decreasing on the interval $(\hat{\alpha}_n, \infty)$, where $\hat{\alpha}_n$ is a unique solution for α of the equation $g'(\alpha) = 0$. This establishes the existence and uniqueness of the MLEs of α and β for this model.

Corollary 2 *Let* $(X_1, N_1), \ldots, (X_n, N_n)$ *be IID copies of* (X, N) *in (1), where* $n \geq 2$. *Then, the MLEs of* α *and* β *exist and are unique, where the MLE of* α, $\hat{\alpha}_n$, *solves the equation*

$$\log(\hat{\alpha}_n) - \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} N_i \psi\left[\hat{\alpha}_n N_i\right] + \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} N_i \log(X_i) + \log\left(\frac{\bar{N}_n}{\bar{X}_n}\right) = 0 \tag{42}$$

while the MLE of β *is given by* $\hat{\beta}_n = \hat{\alpha}_n \bar{N}_n / \bar{X}_n$.

Remark 10 *When the parameter α is known, the MLE of β always exists, is unique, and its value is provided by (40). In particular this recovers the results of the BEG model of Kozubowski and Panorska (2005), where we have $\alpha = 1$.*

Remark 11 *Since the function in (42) is continuous and monotonic, the solution of this equation is straightforward to compute using standard numerical methods.*

Remark 12 *In the special case when the sample size is $n = 1$, the equation (42) does not admit a finite solution. However, the equation is satisfied in the limiting sense, with $\hat{\alpha}_n = \infty$. What the above analysis shows is that the function $R(\alpha, \beta)$ in (38) is maximized "at infinity", when $\alpha \to \infty$ and $\beta = \beta(\alpha)$ is given by (40). A practical interpretation of this is that the gamma variables that drive the model of (X, N) in (1) reduce to a point mass at c, where $c = X_1/N_1$, so that the distribution of (X, N) is a degenerate one, with $(X, N) = (cN, N)$. Similar interpretation applies when $n \geq 2$ and the limit in Part (iii) in Proposition 12 is zero, in which case the solution of (42) is also $\hat{\alpha}_n = \infty$. As can be seen from the proof of this result, this can only happen if all the ratios X_i/N_i in the sample are equal to some $c > 0$. However, this degenerate case is not of a concern in practice, as this event occurs with probability zero.*

Remark 13 *Compared with the results of Barreto-Souza (2012), who considered a special case in (1) with geometrically distributed N (see eq. (14) therein), our results are stronger and more general, as they ensure the existence and uniqueness of the MLEs of α and β in any model of the form (1), going beyond geometrically distributed N.*

Remark 14 *If the family of distributions $\{f_N(\cdot; \boldsymbol{\omega}), \boldsymbol{\omega} \in \Omega\}$ is such that the function $\boldsymbol{\Psi} = (\psi_1, \ldots, \psi_{k+2})$ with $\psi_i = -\partial/\partial\theta_i \log f_{X,N}(x, n)$, $i = 1, \ldots, k+2$, satisfies standard regularity conditions of large-sample theory (see, e.g., conditions A0-A6 in Bickel and Doksum (2015), pp. 384–385) then the MLEs of α and β are consistent, asymptotically normal, and efficient (see, e.g., Theorem 6.2.2 in Bickel and Doksum (2015), p. 386). In particular, due to the special structure of the Fisher information matrix (34), we will have the following convergence in distribution*

$$\sqrt{n}[\,(\hat{\alpha}_n, \hat{\beta}_n) - (\alpha, \beta)\,] \xrightarrow{d} \boldsymbol{Z} \sim N_2(\boldsymbol{0}, \boldsymbol{W}^{-1}) \ \ as \ n \to \infty, \tag{43}$$

where $N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes d-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ and $\boldsymbol{W}$ is the 2 by 2 matrix with entries given by (36). This asymptotic normal distribution can be used to derive (approximate) confidence intervals for the parameters.

3   Bivariate episodes with discrete Pareto duration

Here, we illustrate the general results above with a special case studied in Amponsah (2017), where the random variable N in (1) has discrete Pareto (DP) distribution with parameters $\delta \geq 0$ and $p \in (0, 1)$, given by the PMF (6) and denoted by $\mathcal{DP}(\delta, p)$. Virtually all the results discussed in Section 2 carry-over to this special case, where we use the PMF (6) in place of a generic PMF $f_N(\cdot)$. First, we provide a formal definition of this particular stochastic model.

Definition 1 *A random vector (X, N) with the stochastic representation (1), where the $\{X_i\}$ are IID gamma variables with the PDF (4) and N is DP with the PMF (6), independent of the $\{X_i\}$, is said to have a GMDP distribution with parameters $\alpha > 0$, $\beta > 0$, $\delta \geq 0$, and $p \in (0, 1)$, denoted by $\mathcal{GMDP}(\alpha, \beta, \delta, p)$. The term "GMDP" stands for distribution with* **G**amma-Mixture and **D**iscrete **P**areto marginals.

Remark 15 *In the special case with $\delta = 0$, the DP distribution turns into a geometric distribution (2), and the GMDP distribution reduces to the BGG model proposed by Barreto-Souza (2012). If in addition to $\delta = 0$ we also have $\alpha = 1$, so that the $\{X_i\}$ in (1) are IID exponential with parameter $\beta > 0$, the GMDP model turns into the BEG distribution studied in Kozubowski and Panorska (2005). The GMDP model offers more flexibility, as it has an additional parameter, and can be useful for modeling heavy tailed data, as its marginal distributions of X and N are both heavy tailed for any $\delta > 0$, as will be seen below.*

The joint PDF of $(X, N) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p)$ is given explicitly by (7), with the DP PMF (6) in place of $f_N(\cdot)$. In turn, the CDF and the SF are given by the expressions in Proposition 1. Further, by (8), we obtain the PDF of X for $x \in \mathbb{R}_+$,

$$f_X(x) = \sum_{n=1}^{\infty} \frac{\beta^{n\alpha} x^{n\alpha-1}}{e^{\beta x} \Gamma(n\alpha)} \left[\left(\frac{1}{1 - \delta(n-1)\log(1-p)} \right)^{\frac{1}{\delta}} - \left(\frac{1}{1 - \delta n \log(1-p)} \right)^{\frac{1}{\delta}} \right], \tag{44}$$

which takes a particularly simple form in case of the BEG model ($\delta = 0$, $\alpha = 1$), where X is exponential with parameter βp. We also note that (for $x \in \mathbb{R}_+$) the CDF of X is of the form

$$F_X(x) = \mathbb{P}(X \leq x)$$
$$= \sum_{n=1}^{\infty} \frac{\gamma(\alpha n, \beta x)}{\Gamma(\alpha n)} \left[\left(\frac{1}{1 - \delta(n-1)\log(1-p)} \right)^{\frac{1}{\delta}} - \left(\frac{1}{1 - \delta n \log(1-p)} \right)^{\frac{1}{\delta}} \right].$$

3.1 Laplace transform

The LT of the GMDP model follows from general theory, discussed in Proposition 2 and the remark following it. When we take into account the particular form of the PGF connected with DP distribution (see Proposition 2.6 in Buddana and Kozubowski 2014), we obtain the following result.

Proposition 13 *The LT (11) of $(X, N) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p)$ is of the form*

$$\psi(t, s) = \left(e^{-s} \psi_{X_1}(t) \right) - \left[1 - e^{-s} \psi_{X_1}(t) \right] \sum_{n=1}^{\infty} \frac{\left[e^{-s} \psi_{X_1}(t) \right]^n}{\left[1 - \delta n \log(1-p) \right]^{1/\delta}}, \quad t, s \in \mathbb{R}_+,$$

where $\psi_{X_1}(t) = [\beta/(\beta + t)]^{\alpha}$, $t \in \mathbb{R}_+$, is the LT of gamma $\mathcal{GAM}(\alpha, \beta)$ distribution.

3.2   Moments and tail behavior

Recall that the SF of $N \sim \mathcal{DP}(\delta, p)$ is of the form (see Buddana and Kozubowski 2014, Proposition 2.4)

$$S(x) = \mathbb{P}(N > x) = \left(\frac{1}{1 - \delta \log(1 - p) \lfloor x \rfloor} \right)^{\frac{1}{\delta}}, \quad x \geq 1,$$

where $\lfloor x \rfloor$ is the floor function (the largest integer that is less than or equal to x). This shows that DP distribution is heavy tailed with tail index $\nu = 1/\delta$, that is the SF satisfies (13) with $\nu = 1/\delta$. Further the constant c in (13) is equal to $c = [-\delta \log(1 - p)]^{-1/\delta}$. Thus, by Proposition 3, we have the following result.

Proposition 14 *If* $(X, N) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p)$ *with* $\delta > 0$ *then* X *is heavy tailed with index* $1/\delta$, *and we have*

$$S_X(x) = \mathbb{P}(X > x) \sim [-\delta \log(1 - p) \beta / \alpha]^{-1/\delta} x^{-1/\delta} \quad (x \to \infty).$$

In turn, by Proposition 4, the joint moments $\mathbb{E}[X^r N^\eta]$ with positive $r, \eta > 0$ exist only for $r + \eta < 1/\delta$, and their values are provided in (14) with $f_N(\cdot)$ replaced by the PMF (6) of DP distribution. Further, we also have an analog of Proposition 5, concerning the mean vector and the covariance matrix of $(X, N) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p)$. To formulate the result, we need the mean and the variance of $N \sim \mathcal{DP}(\delta, p)$, which can be deduced from Proposition 2.7 in Buddana and Kozubowski (2014). Namely, the mean of N exists whenever $\delta < 1$, in which case we have

$$\mathbb{E}[N] = \sum_{n=0}^{\infty} \left(\frac{1}{1 - \delta \log(1 - p)n} \right)^{\frac{1}{\delta}} = C\zeta \left(\frac{1}{\delta}, -\frac{1}{-\delta \log(1 - p)} \right), \tag{45}$$

where $C = [-\delta \log(1 - p)]^{-1/\delta}$ and

$$\zeta(s, q) = \sum_{n=0}^{\infty} \left(\frac{1}{n + q} \right)^s$$

is the Hurwitz-zeta special function. Similarly, the variance of N exists whenever $\delta < 1/2$, in which case we have

$$\mathbb{V}ar[N] = \sum_{n=1}^{\infty} \left\{ (2n + 1) \left[\frac{1}{1 - \delta n \log(1 - p)} \right]^{\frac{1}{\delta}} \right\} - \left[\sum_{n=0}^{\infty} \left(\frac{1}{1 - \delta n \log(1 - p)} \right)^{\frac{1}{\delta}} \right]^2. \tag{46}$$

Proposition 15 *Let* $(X, N) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p)$. *The mean vector* $\boldsymbol{\mu}$ *of* $(X, N) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p)$ *exists only if* $\delta < 1$, *in which case we have* $\boldsymbol{\mu} = (\alpha/\beta, 1)\mathbb{E}[N]$ *with* $\mathbb{E}[N]$ *given by (45). Moreover, the covariance matrix* $\boldsymbol{\Sigma}$ *of* (X, N) *exists only if* $\delta < 1/2$, *in which case we have*

$$\boldsymbol{\Sigma} = \mathbb{E}[N] \begin{bmatrix} \alpha/\beta^2 & 0 \\ 0 & 0 \end{bmatrix} + \mathbb{V}ar[N] \begin{bmatrix} \alpha^2/\beta^2 & \alpha/\beta \\ \alpha/\beta & 1 \end{bmatrix},$$

where $\mathbb{E}[N]$ *and* $\mathbb{V}an[N]$ *are given by (45) and (46), respectively.*

3.3   Conditional distributions

The conditional distributions of $(X, N) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p)$ follow general theory laid out in Section 2. First, we note that the conditional distribution of X given $N = n$ is

gamma $\mathcal{GAM}(\alpha n, \beta)$, with the PDF in (15). This conditional distribution is the same in any model of the form (1) driven by an IID gamma sequence $\{X_i\}$, regardless of a particular distribution of N. Second, the conditional distribution of N given $X = x$ has the PMF given by (16) with $f_N(\cdot)$ replaced by the DP PMF (6), the function $f_{\alpha,\lambda}(\cdot)$ given by (17), and the normalizing constant c is given by

$$c = \frac{(\beta x)^{-\alpha}}{E_{\alpha,\alpha}[(\beta x)^{\alpha}]} \sum_{n=1}^{\infty} \frac{(\beta x)^{\alpha n}}{\Gamma(\alpha n)} \left[\left(\frac{1}{1 - \delta(n-1)\log(1-p)} \right)^{\frac{1}{\delta}} - \left(\frac{1}{1 - \delta n \log(1-p)} \right)^{\frac{1}{\delta}} \right].$$

Moreover, according to Proposition 6, this conditional distribution is the same as that of the random variable $N|N = Q$, where N and Q are independent and $Q-1$ has extended Poisson distribution $\mathcal{EP}(\alpha, \lambda)$ with $\lambda = \beta x$, given by the PMF (17). Further, the conditional distribution of $N|X > x$ has the PMF of the form (20), and, in view of Proposition 7, is the same as the distribution of $N|N \geq Q$ with independent N and Q, where this time $Q \overset{d}{=} N(x) + 1$ and $\{N(x), x \in R_+\}$ is a renewal process defined through an IID gamma $\mathcal{GAM}(\alpha, \beta)$ sequence $\{X_i\}$. The latter distribution is one of the two marginal distributions in the bivariate model $(X, N)|X > x$. The full joint distribution of this conditional model has the PDF of the form (24) where $f_Y(\cdot)$ is the PMF (20) of the distribution of $N|X > x$ and $g(\cdot|n)$ is the PDF of gamma distribution $\mathcal{GAM}(\alpha n, \beta)$ truncated below at x, with the PDF given by (25). As discussed in Section 2, the two conditional distributions, one of $X|N = n$ and one for $N|X > x$, provide useful aids in goodness-of-fit analysis when the GMDP model is fitted to data.

We now turn to the conditional distribution of $(X, N)|N > u$, where $u \in \mathbb{N}_0$. As in the general case discussed in Section 2, this conditional distribution admits representation (1) with N replaced by $\tilde{N}$, where the latter has a *truncated* DP distribution, with the PMF given by (27). The PDF and the SF of this distribution are given by (26) and (28), respectively. This case is very special, since the corresponding *excess* random variable, $N_u \overset{d}{=} N-u|N > u$ is also DP distributed (see Proposition 3.3 in Buddana and Kozubowski 2014). In view of this result, we have the following analog of Proposition 9.

Proposition 16 *If $(X, N) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p)$ then we have*

$$(X, N)|N > u \overset{d}{=} (G_u, u) + (X_u, N_u), \tag{47}$$

where G_u and (X_u, N_u) are independent, $G_u \sim \mathcal{GAM}(\alpha u, \beta)$, and $(X_u, N_u) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p_u)$ with

$$p_u = 1 - e^{\frac{\log(1-p)}{1-u\delta\log(1-p)}}.$$

Remark 16 *By the above result, if (X, N) has GMDP distribution then the conditional distribution of $(X, N)|N > u$ is a convolution of two distributions, of which one is degenerate and the other is also a GMDP distribution, with different "size" parameter p. In the special case when $\delta = 0$ (so that N is geometrically distributed) the variable (X_u, N_u) in (47) has actually the same distribution as (X, N) itself, due to the memoryless property of the geometric distribution.*

Additional two conditional distributions, $(X, N)|X \leq x$ and $(X, N)|N \leq u$, are described in the next result which follows from Proposition 10.

Proposition 17 *Suppose that* $(X, N) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p)$.

(i) Let $u, n \in \mathbb{N}$ *and* $x > 0$, *with* $n \leq u$. *Then the conditional CDF of* $(X, N) \sim \mathcal{GMDP}(\alpha, \beta, \delta, p)$ *given* $N \leq u$ *is given by (30), where* $f_N(\cdot)$ *is the DP PMF (6) and*

$$\mathbb{P}(N \leq u) = 1 - \left(\frac{1}{1 - \delta n \log(1 - p)} \right)^{\frac{1}{\delta}}.$$

(ii) Similarly, for $0 < y \leq x$ *and* $n \in \mathbb{N}$, *the conditional CDF of* (X, N) *given* $X \leq x$ *is given by (31), where again* $f_N(\cdot)$ *is given by (6).*

3.4 Infinite divisibility and stochastic representations

In this section we discuss an important stochastic representation of the GMDP distribution as well as its infinite divisibility property. Both of these are related to the fact that the DP distribution is a mixture of geometric distributions. Indeed (see, e.g., shown in Buddana and Kozubowski 2014, Proposition 3.4), the DP distribution of $N \sim \mathcal{DP}(\delta, p)$ admits the following hierarchical structure:

- $Z \sim \mathcal{GAM}(1/\delta, \gamma)$, where $\gamma = [-\delta \log(1 - p)]^{-1}$,
- $N|Z = z$ has a geometric distribution (2) with parameter $p = 1 - e^{-z}$.

Since a DP distributed random variable N is an important component of the GMDP distribution, a similar hierarchical structure carries over to the latter, as shown in the result below.

Proposition 18 *The distribution of* $(X, N) \sim GMDP(\alpha, \beta, \delta, p)$ *admits the following hierarchical structure:*

- $Z \sim \mathcal{GAM}(1/\delta, \gamma)$, where $\gamma = [-\delta \log(1 - p)]^{-1}$,
- $(X, N)|Z = z \sim GMDP(\alpha, \beta, 0, q)$ with $q = 1 - e^{-z}$.

Remark 17 *According to the above result, a GMDP distribution can be thought of as being conditionally a BGG distribution, which is a GMBD distribution with* $\delta = 0$, *so that the variable N in the stochastic representation is a geometric one. In other words, this N is conditionally geometric, so that its unconditional distribution is DP, as seen above, which is precisely why we recover the GMDP model. Another way to describe this is through the density functions. Namely, given a gamma* $\mathcal{GAM}(1/\delta, \gamma)$ *distributed Z in Proposition 18, with density given by*

$$f_Z(z) = \frac{[-\delta \log(1 - p)]^{-1/\delta}}{\Gamma(1/\delta)} z^{1/\delta - 1} e^{z/[\delta \log(1 - p)]}, \quad z \in \mathbb{R}_+,$$

the PMF of $(X, N)|Z = z$ *in Proposition 18 will be*

$$f_{X,N|Z}(x, n) = \frac{\beta^{\alpha n}}{\Gamma(\alpha n)} x^{\alpha n - 1} e^{-\beta x} (1 - e^{-z}) e^{-zn}, \quad x \in \mathbb{R}_+, n \in \mathbb{N}.$$

Then, the distribution of $(X, N) \sim GMDP(\alpha, \beta, \delta, p)$ *can be seen as the marginal distribution of* (X, N) *in a trivariate model* (X, N, Z) *given by the PDF*

$$f_{X,N,Z}(x, n, z) = f_{X,N|Z}(x, n) f_Z(z), \quad x, z \in \mathbb{R}_+, n \in \mathbb{N},$$

so that

$$f_{X,N}(x, n) = \int_0^\infty \frac{\beta^{\alpha n}}{\Gamma(\alpha n)} x^{\alpha n - 1} e^{-\beta x} (1 - e^{-z}) e^{-zn} f_Z(z) dz, \quad x \in \mathbb{R}_+, n \in \mathbb{N}.$$

Further properties of the GMDP model can be developed by means of this hierarchical structure coupled with standard conditioning arguments.

Another consequence of the hierarchical structure of the DP distribution is its infinite divisibility (cf., Corollary 3.5 in Buddana and Kozubowski 2014), which, in view of Corollary 1, immediately implies the same property of the GMDP model. In the following result we need the PGF of $N \sim \mathcal{DP}(\delta, p)$, which is known to be of the form (see Buddana and Kozubowski 2014, Proposition 2.6):

$$G_N(s) = s - (1-s) \sum_{n=1}^{\infty} \left(\frac{1}{1 - \delta \log(1-p)n} \right)^{\frac{1}{\delta}} s^n, \quad s \in (0,1). \tag{48}$$

Corollary 3 *The distribution of $(X, N) \sim GMDP(\alpha, \beta, \delta, p)$ is ID and for each $n \in \mathbb{N}$ we have the stochastic representation (32), where the $\{(R_j^{(n)}, V_j^{(n)})\}$ are IID. Moreover, we have $V_j^{(n)} \stackrel{d}{=} 1/n + T_j^{(n)}$, where the $\{T_j^{(n)}\}$ are IID, integer-valued random variables with the PGF $G_n(s) = [G_N(s)/s]^{1/n}$ and $G_N(\cdot)$ given by (48). Further, the $\{R_j^{(n)}\}$ are given by (33), where all the variables on the right-hand-side of (33) are mutually independent, the $\{G_j^{(n)}\}$ are IID gamma $\mathcal{GAM}(\alpha/n, \beta)$ distributed, and the $\{X_{ij}\}$ are IID gamma $\mathcal{GAM}(\alpha, \beta)$ distributed.*

3.5 Parameter estimation

The discussion of estimation for the general case in Section 2 generally carries over to the special case of the GMDP distribution. Here, the variable $N \sim \mathcal{DP}(\delta, p)$ depends on a 2-dimensional vector-parameter $\boldsymbol{\omega} = (\delta, p) \in \Omega$, where we take the parameter space Ω to be an *open set* $\Omega = (0, \infty) \times (0, 1)$. The PDF of N, which is given by (6), will be denoted by $f_N(\cdot; \boldsymbol{\omega})$. Clearly, our model has four unknown parameters, which can be conveniently described as a vector-parameter $\boldsymbol{\theta}$, where $\theta_1 = \alpha$, $\theta_2 = \beta$, $\theta_3 = \delta$, and $\theta_4 = p$. With this notation, the parameter space is an open set $\Theta = (0, \infty)^3 \times (0, 1)$.

3.5.1 Fisher information

According to general results in Section 2, if the appropriate expectations exist, the Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ connected with the $\mathcal{GMDP}(\alpha, \beta, \delta, p)$ distribution has a block-diagonal structure,

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} w_{11} & w_{12} & 0 & 0 \\ w_{21} & w_{22} & 0 & 0 \\ 0 & 0 & j_{11} & j_{12} \\ 0 & 0 & j_{21} & j_{22} \end{bmatrix}, \tag{49}$$

where the entries $\{w_{ij}\}$, generally given by (35), depend only on the parameters α and β, and the 2×2 matrix

$$\mathbf{J}(\boldsymbol{\omega}) = \begin{bmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{bmatrix} \tag{50}$$

is the Fisher information matrix connected with the $\mathcal{DP}(\delta, p)$ model. Since in our situation the expectation of N is given by (45), according to (36), we have

$$
w_{11} = \sum_{n=1}^{\infty} n^2 \psi'(\alpha n) \left[\left(\frac{1}{1 - \delta(n-1)\log(1-p)} \right)^{\frac{1}{\delta}} - \left(\frac{1}{1 - \delta n \log(1-p)} \right)^{\frac{1}{\delta}} \right],
$$

$$
w_{12} = w_{21} = -\frac{1}{\beta} \sum_{n=0}^{\infty} \left(\frac{1}{1 - \delta \log(1-p)n} \right)^{\frac{1}{\delta}},
$$

$$
w_{22} = \frac{\alpha}{\beta^2} \sum_{n=0}^{\infty} \left(\frac{1}{1 - \delta \log(1-p)n} \right)^{\frac{1}{\delta}},
$$

where $\psi(\cdot)$ is the digamma function. We note that the above quantities are well defined only when $\delta < 1$, due to the heavy tail of the DP distribution. Indeed, clearly the two series describing $w_{12} = w_{21}$ and w_{22} converge only if $1/\delta > 1$, so that we must have $\delta < 1$. In addition, we have the asymptotic relation $\psi'(z) \sim 1/z$ as $z \to \infty$ (see, e.g., formula 6.4.12 in Abramowitz and Stegun 1972, p. 260) and the PMF of the DP distribution given by the expression in the square bracket in the series describing w_{11} behaves asymptotically as the power $n^{-(1+1/\delta)}$ at infinity. Thus, that series also converges only if $\delta < 1$. Further, standard calculations produce the information matrix (50) of the DP model, which is given in the result below.

Proposition 19 *The Fisher information matrix (50) of the $\mathcal{DP}(\delta, p)$ model is well defined for any $\delta > 0$ and $p \in (0, 1)$, and is given by*

$$
\mathbf{J}(\boldsymbol{\omega}) = \boldsymbol{D}^\top \boldsymbol{A} \boldsymbol{D} \tag{51}
$$

where

$$
\boldsymbol{D} = \begin{bmatrix} -\frac{1}{\delta^2} & 0 \\ -\log(1-p) & \frac{\delta}{1-p} \end{bmatrix}, \tag{52}
$$

$\boldsymbol{A} = [a_{ij}]$ is a symmetric matrix with entries of the form

$$
a_{11} = \mathbb{E}\left\{ \frac{[g(\delta,p,N) - g(\delta,p,N-1)]^2}{[f_N(N;\delta,p)]^2} \right\},
$$

$$
a_{12} = a_{21} = \frac{1}{\delta} \mathbb{E}\left\{ \frac{[g(\delta,p,N) - g(\delta,p,N-1)][v(\delta,p,N) - v(\delta,p,N-1)]}{[f_N(N;\delta,p)]^2} \right\},
$$

$$
a_{22} = \frac{1}{\delta^2} \mathbb{E}\left\{ \frac{[v(\delta,p,N) - v(\delta,p,N-1)]^2}{[f_N(N;\delta,p)]^2} \right\},
$$

the function $f_N(\cdot; \delta, p)$ is the PDF (6) of the $\mathcal{DP}(\delta, p)$ distribution, and

$$
g(\delta,p,n) = \frac{\log[\,1 - \delta n \log(1-p)]}{[\,1 - \delta n \log(1-p)]^{1/\delta}}, \quad v(\delta,p,n) = \frac{n}{[\,1 - \delta n \log(1-p)]^{1+1/\delta}},
$$

$$
\delta \in (0, \infty), p \in (0, 1), n \in \mathbb{N}_0.
$$

3.5.2 Maximum likelihood estimation

Let $(X_1, N_1), \ldots, (X_n, N_n)$ be a random sample from a GMDP distribution. According to general results on estimation discussed in Section 2, maximum likelihood estimation can be carried out separately for the two pairs, α, β, and δ, p. Regarding the former, the MLEs of α and β always exist and are unique, and their values are provided in Corollary 2. Moreover, while the calculation of the MLEs involves solving a one-dimensional non-linear

equation, stated in (42), the relevant function is rather well behaved and finding the solution by standard numerical methods is straightforward. Further, since the family of DP distributions $\{f_N(\cdot; \boldsymbol{\omega}), \boldsymbol{\omega} \in (0, \infty) \times (0, 1)\}$ is such that the function $\boldsymbol{\Psi} = (\psi_1, \ldots, \psi_4)$ with $\psi_i = -\partial/\partial\theta_i \log f_{X,N}(x, n)$, $i = 1, \ldots, 4$, satisfies standard regularity conditions of large-sample theory, the MLEs of α and β are consistent, asymptotically normal, and efficient. Due to the special structure of the Fisher information matrix (49), we will have the convergence in distribution (43) where $\mathbf{W}$ is the 2×2 matrix with entries $\{w_{ij}\}$ given above. This asymptotic normal distribution can be used to derive (approximate) confidence intervals for the two parameters.

Let us now move to estimation of parameters δ and p. According to general results on estimation discussed in Section 2, this requires optimization of the function $Q(\boldsymbol{\omega})$ in (39) with respect to $\boldsymbol{\omega} = (\delta, p) \in (0, \infty) \times (0, 1)$. Since the function

$$Q(\delta, p) = \sum_{i=1}^{n} \log \left[\left(\frac{1}{1 - \delta(N_i - 1)\log(1 - p)} \right)^{\frac{1}{\delta}} - \left(\frac{1}{1 - \delta(N_i)\log(1 - p)} \right)^{\frac{1}{\delta}} \right] \quad (53)$$

is the log-likelihood (LL) function of an univariate random sample $\{N_i\}$ from a DP distribution, one can use methods available for this distribution in order to get the MLEs. In particular, we can proceed by solving the likelihood equations, obtained by setting to zero the two partial derivatives of the LL function in (53). The calculation of the derivatives is straightforward; after some algebra, we obtain the following likelihood equations:

$$\sum_{i=1}^{n} \frac{N_i}{h(\delta, p, N_i)} \left\{ \left[\frac{h(\delta, p, N_i)}{h(\delta, p, N_i - 1)} \right]^{\frac{1}{\delta}} - 1 \right\}^{-1} = \sum_{i=1}^{n} \frac{N_i - 1}{h(\delta, p, N_i - 1)} \left\{ 1 - \left[\frac{h(\delta, p, N_i - 1)}{h(\delta, p, N_i)} \right]^{\frac{1}{\delta}} \right\}^{-1},$$

$$\sum_{i=1}^{n} \log h(\delta, p, N_i) \left\{ \left[\frac{h(\delta, p, N_i)}{h(\delta, p, N_i - 1)} \right]^{\frac{1}{\delta}} - 1 \right\}^{-1} = \sum_{i=1}^{n} \log h(\delta, p, N_i - 1) \left\{ 1 - \left[\frac{h(\delta, p, N_i - 1)}{h(\delta, p, N_i)} \right]^{\frac{1}{\delta}} \right\}^{-1},$$

where

$$h(\delta, p, n) = 1 - \delta n \log(1 - p), \quad \delta \in (0, \infty), p \in (0, 1), n \in \mathbb{N}_0.$$

While the above equations can be solved for δ and p by standard numerical methods, the actual calculations may be computationally intensive. An alternative approach for finding the MLEs in this case is the standard EM algorithm, as the DP distribution is conveniently represented as a mixture of geometric distributions. Such an algorithm has been recently developed and tested in Amponsah and Kozubowski (2021), and was found to be working quite well in this setting. We recommend that it be used for this estimation problem.

Remark 18 *A careful analysis of the case with $n = 1$ shows that the likelihood equations above do not admit any solution within the parameter space. However, the function $Q(\delta, p)$ is maximized by $\hat{\delta}_n = 0$ and $\hat{p}_n = 1/\bar{N}$. These values point towards the limiting geometric distribution of N, which arises when $\delta = 0$ in the DP model.*

Table 1 Frequencies, relative frequencies, and discrete Pareto model probabilities of all observed growth periods' durations

Growth period N in days	1	2	3	4	5	6	7 +
Frequency	272	106	61	31	15	12	11
Relative frequency	0.5354	0.2087	0.1201	0.0610	0.0295	0.02362	0.0217
Model probability	0.5205	0.2359	0.1136	0.0576	0.0306	0.0169	0.0250

4   An application from finance

In this section, we show an application of the GMDP model to a financial data set. We chose the daily closing price of Bahrain all share Index, quoted in Bahrain dinar. The data set was obtained from Investing.com, and consisted of the daily closing prices, from May 24, 2010, to February 20, 2019. We converted our daily data to ($n = 2171$) daily log-returns, which are the logarithms of the ratios of the closing prices for every two consecutive days. Next, we computed the growth episodes and obtained ($n = 508$) observations (X_i, N_i), $i = 1, 2, \ldots, 508$, where the N_i is the duration and X_i is the magnitude (total log-return over the period of the ith episode) of the ith growth episode. Duration data from this data set was previously fitted with the DP model in Amponsah and Kozubowski (2021). In this work we fit the bivariate GMDP model to the duration and magnitude of the growth episodes.
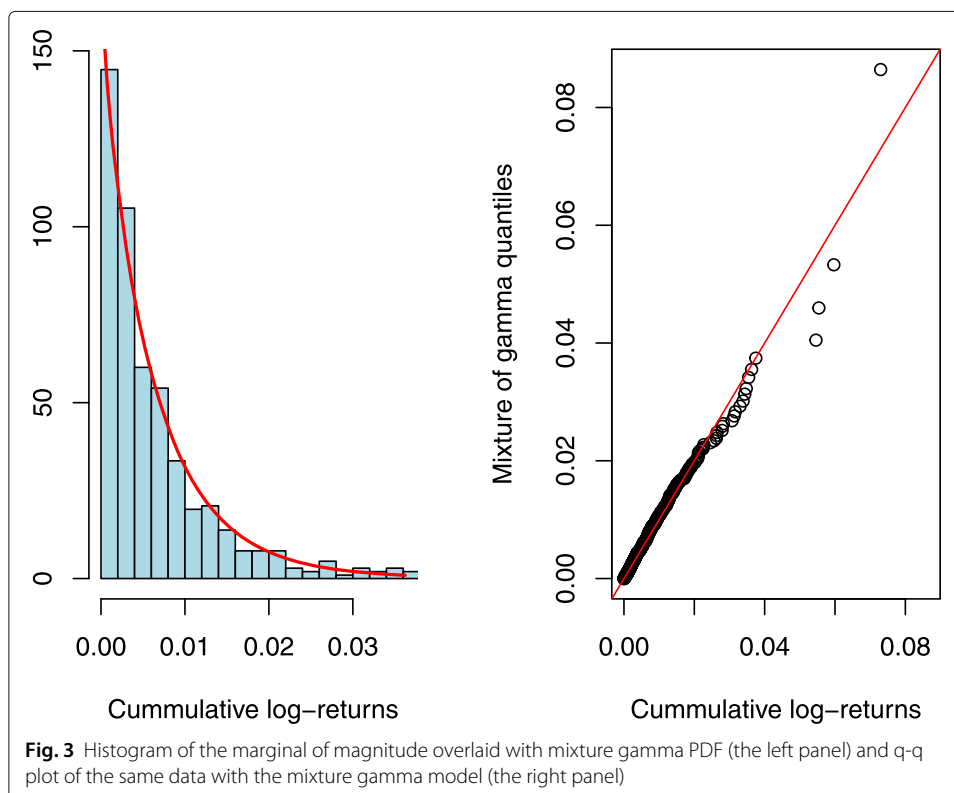
Assuming the data are IID $\mathcal{GMDP}(\alpha, \beta, \delta, p)$, the MLEs computed by solving likelihood equations (see Section 2.4) are $\hat{\alpha} = 0.9420$, $\hat{\beta} = 281.4498$, $\hat{\delta} = 0.1153$, and $\hat{p} = 0.5355$. The EM algorithm of Amponsah and Kozubowski (2021) was used to compute $\hat{\delta}$ and $\hat{p}$. To assess the goodness-of-fit of our model to the growth episodes of the Bahrain all share Index, we used the MLEs described above to compute the parameters of all the marginal and conditional distributions we fit below. Considering this fact, the generally
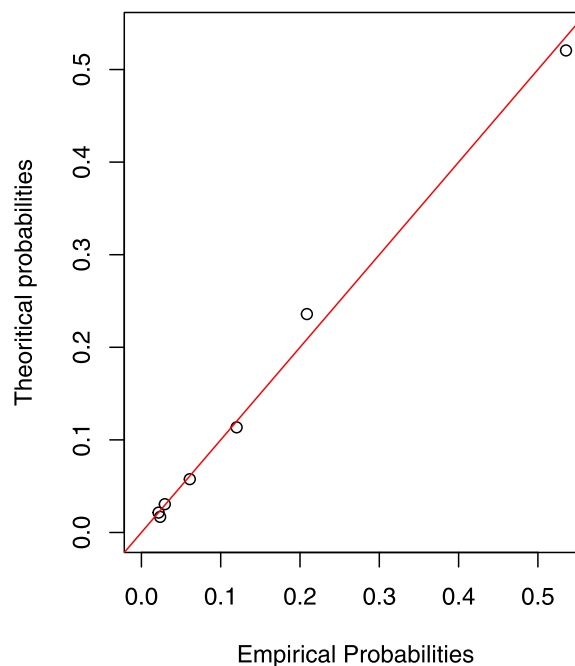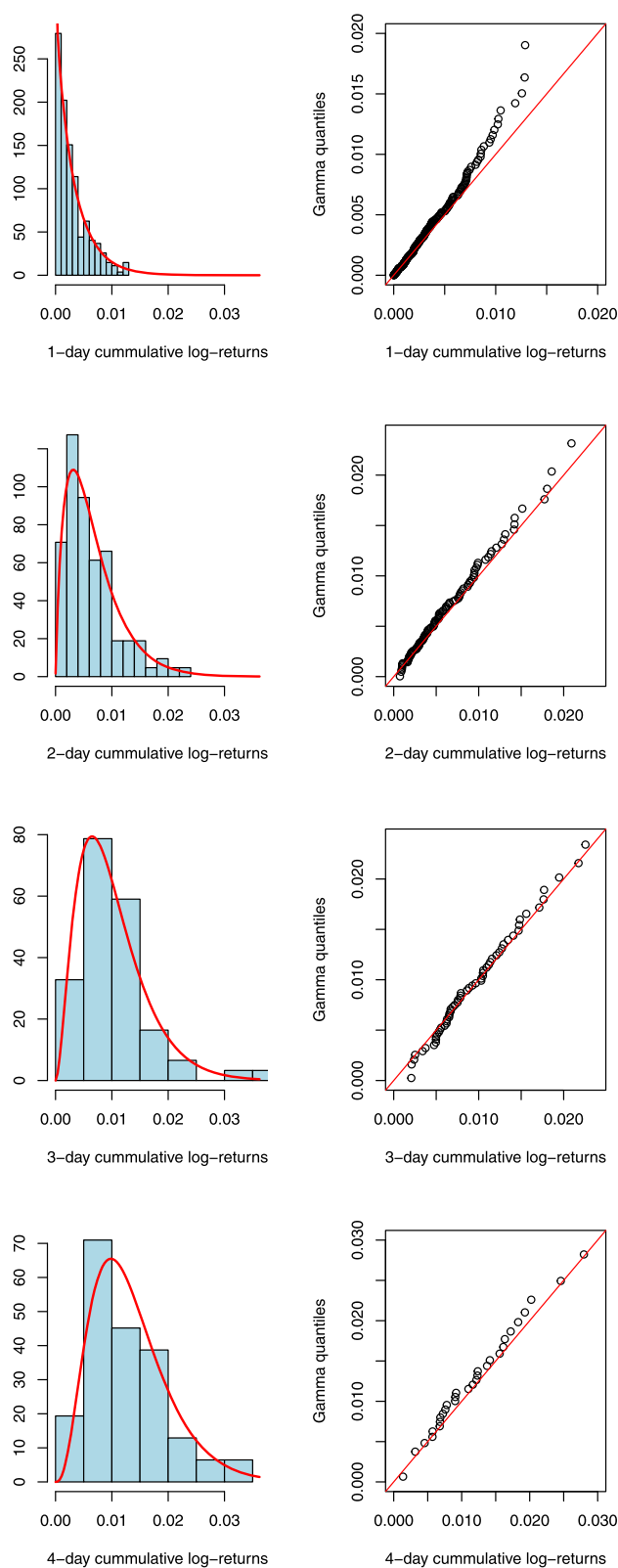

Fig. 1 The plot of relative frequency (x-axis) versus model probability (y-axis) of durations $N = 1, 2, 3, 4, 5, 6$ and $N \geq 7$

Fig. 2 Histogram of positive daily log-returns overlaid with gamma PDF (the left panel) and a q-q plot of the same data with gamma model (the right panel)
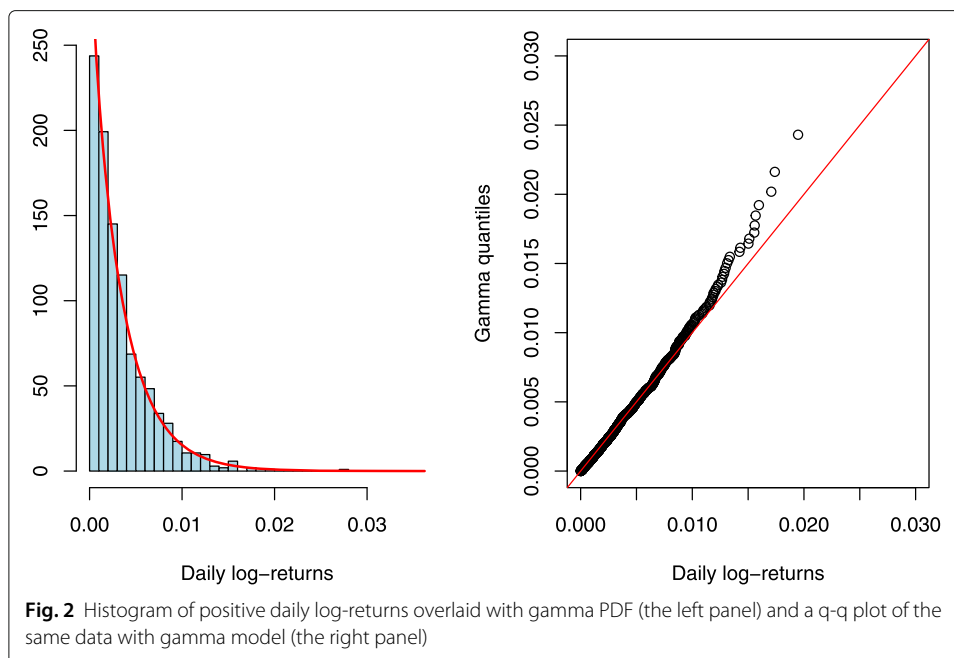
reasonable fit of the marginals and analyzed conditional models seems quite remarkable. We analyzed the fit of the following marginal and conditional distributions: (1) marginal of duration, (2) marginal of magnitude, (3) conditional distribution of $X|N = n$ for $n = 1, 2, 3, 4$, and (4) conditional distribution of $N|X > x$ for thresholds $x = 0.0001, 0.0015$, and 0.0025.

Fig. 3 Histogram of the marginal of magnitude overlaid with mixture gamma PDF (the left panel) and q-q plot of the same data with the mixture gamma model (the right panel)

Fig. 4 Histogram of positive cumulative log-returns overlaid with gamma PDFs and q-q plots of the same data with gamma models for $N = 1$ (top row), 2 (second row) 3 (third row), and 4 (bottom row)

Table 2 Results of the KS test of goodness-of-fit of the conditional distributions of $X|N = k$

Growth period *N* in days	1	2	3	4
KS statistic	0.081	0.132	0.180	0.226
p-value	0.336	0.314	0.276	0.414

4.1 Goodness-of-fit analysis

We started with the fit of DP model to the marginal of duration. Table 1 contains frequency, relative frequency, and model probabilities for the values of N observed in the data, namely $N = 1, 2, ..., 6$, and $N \geq 7$.

Figure 1 shows a plot of the DP model probabilities (the last row of Table 1) versus the observed relative frequencies (the third row of Table 1), with a diagonal line $y = x$ added for visualization purposes. The DP probabilities are very close to the relative frequencies, and the mean (absolute) percentage error is about 10.598%. A formal chi-square goodness-of-fit test based on the information provided in Table 1 was performed as well, resulting in the *p*-value of 0.27. Thus, the null hypothesis of DP Pareto distribution of N is not rejected at 5% significance level. Overall, this analysis shows a reasonable fit of the DP model to the marginal of duration.

Next, we checked if the positive daily log-returns are well fit by the gamma distribution with parameter α and β. We used graphical tools (histogram, q-q plots), and formal Kolmogorov-Smirnov (KS) test to confirm the goodness-of-fit. Figure 2 shows a histogram of positive daily log-returns overlayed with fitted gamma PDFs and a q-q plot of the model and empirical quantiles. Both plots show a good fit. We also tested the fit of the model to the positive daily log-returns. The KS statistic and its *p*-value are 0.0397 and 0.3905, respectively. Therefore, the null hypothesis that the positive daily log-returns follow gamma distributions with parameters α and β, is not rejected at significance level of 5%. Overall, we believe that the positive log-returns may be assumed to have come from gamma distribution with parameters α and β.

Further, we investigated the fit of the model marginal distribution of magnitude to the data. The model marginal distribution of X is a mixture of gamma distributions with scale parameter β and shape parameters αn, $n \in \mathbb{N}$, with the mixing weights given by $P(N = n)$, where $N \sim DP(\delta, p)$. The plot in Fig. 3 shows a histogram of the marginal of magnitude overlaid with the gamma mixture model and the corresponding q-q plot. The KS test statistic and *p*-value are 0.0453 and 0.6751, respectively. Thus, the null hypothesis that the model marginal follows a mixture of gamma distributions with PDF (44) is not rejected at significance level of 5%. We believe the results show a reasonable fit of the marginal distribution.

Next, we turned to goodness-of-fit analysis of the GMDP model's conditional distributions of X given $N = k$, which are gamma with parameters $k\alpha$ and β, to the corresponding data. Figure 4 shows graphical illustration of fit via histograms overlaid with fitted gamma

Table 3 Frequencies, relative frequencies, and estimated model probabilities connected with the conditional distribution of *N* given that $X > 0.0001$

Growth period *N*|*X* > 0.0001 in days	1	2	3	4	5	6	7 +
Frequency	262	106	61	31	15	12	11
Relative frequencies	0.5261	0.2129	0.1225	0.0622	0.0301	0.0241	0.0221
Model probabilities	0.5118	0.2401	0.1157	0.0587	0.0311	0.0172	0.0255

Table 4 Frequencies, relative frequencies, and estimated model probabilities connected with the conditional distribution of N given that $X > 0.0015$

Growth period $N\|X > 0.0015$ in days	1	2	3	4	5	6	7 +
Frequency	168	98	61	30	15	12	11
Relative frequencies	0.4253	0.2481	0.1544	0.0759	0.0380	0.0304	0.0278
Model probabilities	0.4161	0.2756	0.1428	0.0733	0.0389	0.0215	0.0318

PDFs and q-q plots for durations $N = 1, 2, 3$, and 4. The plots show a reasonable fit of the conditional distributions of the GMDP model to the corresponding data.

In addition, we performed the Kolmogorov-Smirnov goodness-of-fit test of these conditional models to the data. The results of the KS tests of the null hypothesis that the conditional distributions of X given $N = k$ are gamma with parameters $k\alpha$ and β, are given in Table 2. None of the tests rejected the null hypothesis on the significance level of 5%. We believe the results confirm a reasonable fit of the conditional distributions.

Similarly, we performed a goodness-of-fit analysis of the GMDP model's conditional distributions of $N|X > x$ for thresholds $x = 0.0001, 0.0015$ and 0.0025, which should have a hidden truncated discrete Pareto distributions as described in Proposition 6.

Tables 3, 4 and 5 show the frequencies, relative frequencies and estimated model probabilities of all the observed growth periods' durations given that $X > x$, for the thresholds of $x = 0.0001, 0.0015$, and 0.0025, respectively. Figure 5 shows plots of the model probabilities versus the corresponding relative frequencies reported in these tables. The estimated conditional model probabilities are very close to the relative frequencies, and the mean of the (absolute) percentage errors for the thresholds $x = 0.0001, 0.0015$, and 0.0025, are 10.5728%, 10.0313% and 10.0817%, respectively. The chi-square goodness-of-fit tests based on the information in Tables 3 - 5 resulted in the p-values of 0.45, 0.52, 0.49, respectively. We conclude that these conditional theoretical models show reasonable fit to the data.
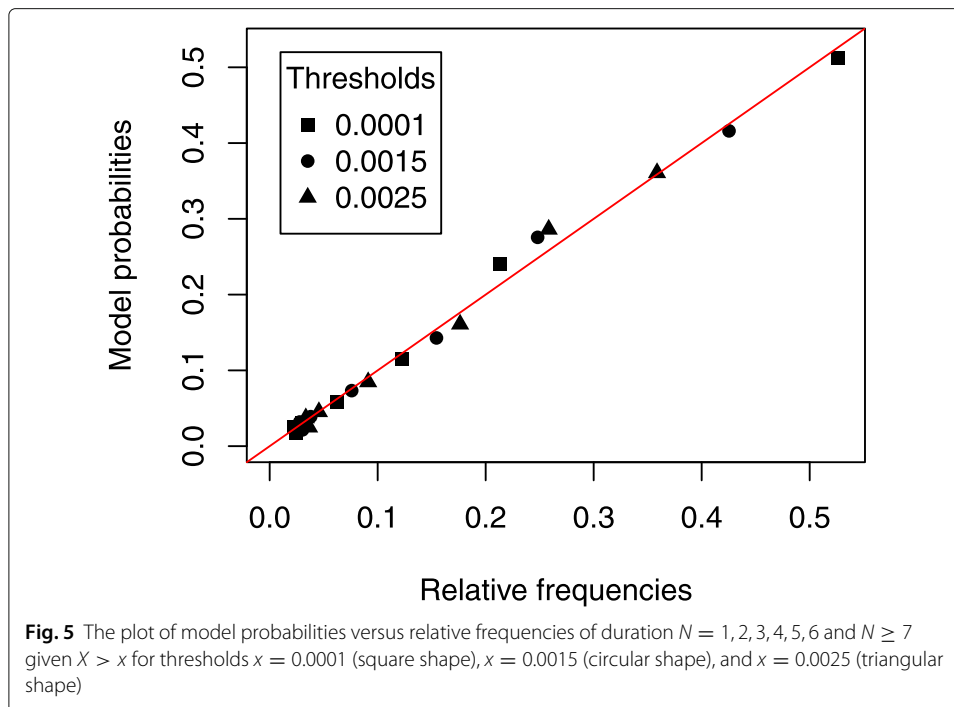
Overall, we believe the GMDP model fits the growth periods of the Bahrain index rather well. This example provides evidence for the considerable modeling potential of the GMDP model.

5 Concluding remarks

This work provides fundamental properties of a general stochastic model describing the joint distribution of (X, N), where N is a counting variable while X is the sum of N independent gamma random variables. These properties include marginal and conditional distributions, Laplace transform, moments and tail behavior, and divisibility properties. The existence and uniqueness of maximum likelihood estimators of the relevant parameters are established as well, along with their asymptotic properties. A particular example with a heavy tailed discrete Pareto distributed N illustrates general properties of this

Table 5 Frequencies, relative frequencies, and estimated model probabilities connected with the conditional distribution of N given that $X > 0.0025$

Growth period $N\|X > 0.0025$ in days	1	2	3	4	5	6	7 +
Frequency	118	85	58	30	15	12	11
Relative frequencies	0.3587	0.2584	0.1763	0.0912	0.0456	0.0365	0.0334
Model probabilities	0.3607	0.2861	0.1609	0.0848	0.0453	0.0250	0.0372

Fig. 5 The plot of model probabilities versus relative frequencies of duration $N = 1, 2, 3, 4, 5, 6$ and $N \geq 7$ given $X > x$ for thresholds $x = 0.0001$ (square shape), $x = 0.0015$ (circular shape), and $x = 0.0025$ (triangular shape)

stochastic model, while a data example from finance shows the modeling potential of this new mixed bivariate distribution.

Appendix

This section contains (selected) proofs and auxiliary results.

Lemma 3 *For $x > 0$, let $h(x) = \log(x) - \psi(x)$ where $\psi(x)$ is the digamma function, which is the derivative of $\log \Gamma(x)$. Then $h(x)$ is continuous and strictly decreasing on the interval $(0, \infty)$, with*

$$\lim_{x \to 0^+} h(x) = \infty \text{ and } \lim_{x \to \infty} h(x) = 0.$$

Proof of Lemma 3. Since the trigamma function $\psi'(x)$ satisfies the inequality (see, e.g., Corollary 5.5 in Laforgia and Natalini 2013, p. 501)

$$\psi'(x) > \frac{1}{x} \text{ for } x > 0,$$

we conclude that

$$h'(x) = \frac{1}{x} - \psi'(x) < 0 \text{ for } x > 0,$$

which proves the monotonicity of $h(x)$ on $(0, \infty)$. The limits follow easily from the inequality (see, e.g., Remark 5.3 in Laforgia and Natalini 2013, p. 500)

$$\frac{1}{2x} < \log(x) - \psi(x) < \frac{1}{x}, \ x > 0.$$

This concludes the proof.

Proof of Proposition 2. By standard conditioning argument, we have

$$\psi(t, s) = \mathbb{E}\left(e^{-tX - sN}\right) = \mathbb{E}\left\{\mathbb{E}\left(e^{-sN}e^{-tX}|N\right)\right\} = \mathbb{E}\left\{e^{-sN}H(N)\right\},$$

where

$$H(n) = \mathbb{E}\left(e^{-t\sum_{i=1}^{N} X_i} \,|\, N = n\right) = \mathbb{E}\left(e^{-t\sum_{i=1}^{n} X_i}\right) = \prod_{i=1}^{n} \mathbb{E}(e^{-tX_i}) = [\psi_{X_1}(t)]^n$$

and $\psi_{X_1}(\cdot)$ is the LT of the $\{X_i\}$. This leads to

$$\psi(t,s) = \mathbb{E}\left\{\left[e^{-s}\psi_{X_1}(t)\right]^N\right\} = G_N\left(e^{-s}\psi_{X_1}(t)\right).$$

Since for gamma distributed $\{X_i\}$ we have $\psi_{X_1}(t) = 1/(1 + t/\beta)^\alpha$, the result follows.

Proof of Proposition 3. The result follows from Theorem 3.2 in Robert and Segers (2008), as the conditions (3.10) or (3.11) of that theorem hold for our N due to the fact that N satisfies (13) while the $\{X_i\}$ are light tailed gamma distributed.

Proof of Proposition 5. The result follows from straightforward application of Proposition 4.

Proof of Lemma 1. We proceed by showing that the PMF of the random variable on the right-hand-side in (19) coincides with the expression on the right-hand-side in (18). Indeed, by independence of N and Q, for any $n \in \mathbb{N}_0$, we have

$$\mathbb{P}(N = n | N = Q) = \frac{\mathbb{P}(N = n, N = Q)}{\mathbb{P}(N = Q)} = \frac{\mathbb{P}(N = n, Q = n)}{\mathbb{P}(N = Q)} = \frac{f_N(n)f_Q(n)}{\mathbb{P}(N = Q)},$$

so that the above probability must coincide with the PMF of Y in (18) with $c = \mathbb{P}(N = Q)$.

Proof of Proposition 8. We know, by Proposition 7, that N given $X > x$ has the same distribution as N given $N \geq N(x) + 1$, where $N(x)$ is independent of N, and, in our case, Poisson with parameter $\lambda = \beta x$. The PMF of this conditional distribution is of the form

$$\begin{aligned}
h(n) = \mathbb{P}(N = n | N \geq N(x) + 1) &= \frac{\mathbb{P}(N = n, N \geq N(x) + 1)}{\mathbb{P}(N - 1 \geq N(x))} \\
&= \frac{\mathbb{P}(N = n)\mathbb{P}(N(x) \leq n - 1)}{\mathbb{P}(N - 1 \geq N(x))}.
\end{aligned} \tag{54}$$

By the independence of N and $N(x)$, the denominator on the right-hand-side in (54) is equal to

$$\begin{aligned}
\mathbb{P}(N - 1 \geq N(x)) &= \sum_{n=1}^{\infty} \mathbb{P}(N = n, N - 1 \geq N(x)) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(N = n, n - 1 \geq N(x)) = \sum_{n=0}^{\infty} p(1-p)^n \mathbb{P}(N(x) \leq n).
\end{aligned}$$

By inserting

$$\mathbb{P}(N(x) \leq n) = \sum_{j=0}^{n} \mathbb{P}(N(x) = j) = \sum_{j=0}^{n} \frac{(\beta x)^j e^{-\beta x}}{j!} \tag{55}$$

into the above expression and interchanging the order of the summation, followed by summing-up the resulting geometric series, we arrive at

$$\mathbb{P}(N - 1 \geq N(x)) = \sum_{j=0}^{\infty} \frac{(\beta x)^j e^{-\beta x}}{j!}(1-p)^j = e^{(1-p)\beta x - \beta x}. \tag{56}$$

We now calculate the PGF of $Y \stackrel{d}{=} N | N \geq N(x) + 1$, which, in view of (54) and (56), can be expressed as

$$G(s) = \mathbb{E}[s^Y] = \sum_{n=1}^{\infty} s^n h(n) = \frac{sp \sum_{n=0}^{\infty} s^n (1-p)^n \mathbb{P}(N(x) \leq n)}{e^{(1-p)\beta x - \beta x}}, \quad s \in [0,1]. \tag{57}$$

By proceeding as above, and calculating the numerator in (57) by inserting there the expression (55), followed by interchanging the order of the summation and subsequently summing-up the resulting geometric series, we find the numerator in (57) to be

$$sp \sum_{n=0}^{\infty} s^n (1-p)^n \mathbb{P}(N(x) \le n) = \frac{sp}{1-(1-p)s} e^{s(1-p)\beta x - \beta x}.$$

Thus, we obtain

$$G(s) = \frac{sp}{1-(1-p)s} e^{-(1-p)\beta x(1-s)}, \quad s \in [0,1].$$

Since this is the product of the PGF of N and the PGF of T, we obtain the result.

Proof of Proposition 9. The result follows by direct computation of the LTs on each side in (29), and subsequence verification that they coincide.

Proof of Proposition 11. We proceed by showing that the joint LT of the expression on the right-hand-side in (32) coincides with the LT of (X, N), with the latter given by $\psi(t, s) = G_N[e^{-s}\psi_{X_1}(t)]$, as stated in the second remark following Proposition 2. We start with the LT of the pair $(R_j^{(n)}, V_j^{(n)})$. By standard conditioning argument, we have

$$\mathbb{E}\left(e^{-tR_j^{(n)} - sV_j^{(n)}}\right) = \mathbb{E}\left\{\mathbb{E}\left(e^{-tR_j^{(n)} - sV_j^{(n)}} | T_j^{(n)}\right)\right\} = \mathbb{E}\left\{H(T_j^{(n)})\right\},$$

where

$$H(k) = \mathbb{E}\left(e^{-t\left(\sum_{i=1}^{k} X_{ij} + G_j^{(n)}\right)} e^{-s(k+1/n)}\right) = [\psi_{X_1}(t)]^{k+1/n} e^{-s(k+1/n)}$$

$$= [e^{-s}\psi_{X_1}(t)]^k [e^{-s}\psi_{X_1}(t)]^{1/n}.$$

Since the PGF of $\{T_j^{(n)}\}$ is equal to $G_n(s) = [G_N(s)/s]^{1/n}$, we have

$$\mathbb{E}\left\{H(T_j^{(n)})\right\} = [e^{-s}\psi_{X_1}(t)]^{1/n} \mathbb{E}\left([e^{-s}\psi_{X_1}(t)]^{T_j^{(n)}}\right)$$

$$= [e^{-s}\psi_{X_1}(t)]^{1/n} G_n[e^{-s}\psi_{X_1}(t)] = \left\{G_N[e^{-s}\psi_{X_1}(t)]\right\}^{1/n}.$$

Finally, since variables under the sum in (32) are IID, the LT of the sum becomes

$$\left[\mathbb{E}\left(e^{-tR_j^{(n)} - sV_j^{(n)}}\right)\right]^n = \left[\left\{G_N[e^{-s}\psi_{X_1}(t)]\right\}^{1/n}\right]^n = G_N[e^{-s}\psi_{X_1}(t)],$$

as desired. This completes the proof.

Proof of Theorem 12. First, we re-write the function $g'(\alpha)$ as follows:

$$g'(\alpha) = \log(\alpha) - \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} N_i \psi [\alpha N_i] + \log\left(\frac{\bar{N}_n}{\bar{X}_n}\right) + \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} N_i \log(X_i)$$

$$= \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} \left[N_i \left(\log(\alpha) - \psi [\alpha N_i] + \log\left(\frac{\bar{N}_n}{\bar{X}_n}\right) + \log(X_i)\right)\right]$$

$$= \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} \left[N_i \left(\log(\alpha N_i) - \psi [\alpha N_i] + \log\left(\frac{\bar{N}_n}{\bar{X}_n}\right) + \log\left(\frac{X_i}{N_i}\right)\right)\right]$$

$$= \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} \left[N_i \left(\log(\alpha N_i) - \psi [\alpha N_i]\right)\right] + \log\left(\frac{\bar{N}_n}{\bar{X}_n}\right) + \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} N_i \left[\log\left(\frac{X_i}{N_i}\right)\right]$$

$$= \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} \left[N_i \left(\log(\alpha N_i) - \psi [\alpha N_i]\right)\right] + \frac{1}{n\bar{N}_n} \sum_{i=1}^{n} N_i \left[\log\left(\frac{X_i}{N_i}\right)\right] - \log\left(\frac{\bar{X}_n}{\bar{N}_n}\right).$$

In view of the above and Lemma 3, we immediately obtain Parts (i) and (ii) as well as the limit in Part (iii). To see that the latter is less then or equal to zero, it is enough to use Jensen's inequality,

$$\sum_{i=1}^{n} a_i \varphi(b_i) \leq \varphi \left(\sum_{i=1}^{n} a_i b_i \right), \tag{58}$$

where $\varphi(\cdot)$ is a concave function and the positive weights $\{a_i\}$ sum up to one. Indeed, we get the result upon setting $\varphi(x) = \log x$, $a_i = N_i/(n\bar{N}_n)$, and $b_i = X_i/N_i$ in (58). Part (iv) is straightforward. Finally, Part (v) is obtained by noting that for a non-linear function $\varphi(\cdot)$, we have an equality in (58) if and only if $b_1 = \cdots = b_n$, which in our case becomes $X_i/N_i = c$, $i = 1, \ldots, n$, for some $c > 0$. However, this event has probability zero since the $\{X_i/N_i\}$ are IID and their distribution is continuous. This concludes the proof.

Proof of Proposition 19. Consider a re-parameterization of a DP model, where the new parameters are $\tau_1 = 1/\delta$ and $\tau_2 = -\delta \log(1 - p)$. Then, the Fisher information matrix of the DP model in the δ, p parameterization is given by (51), where $\mathbf{D}$ is the Jacobian matrix of the transformation $(\delta, p) \to (\tau_1, \tau_2)$, which is given by (52), and $\mathbf{A}$ is the Fisher information matrix of the DP model in the τ_1, τ_2 parameterization. Routine calculations show that the entries of the latter matrix, written in terms of δ and p, are the quantities $\{a_{ij}\}$ specified in the proposition, and all the relevant expectations are finite for any $\delta \in (0, \infty)$ and $p \in (0, 1)$.

Abbreviations

BEG: Bivariate (distribution) with exponential and geometric margins; BGG: Bivariate gamma geometric; CDF: Cumulative distribution function; DP: Discrete Pareto; $\mathcal{EP}(\alpha,\beta)$: Extended Poisson distribution with parameters α, β; $\mathcal{GAM}(\alpha,\beta)$: gamma distribution with parameters α, β; GMDP: (Distribution) with gamma-mixture and discrete Pareto margins; ID: Infinitely divisible; IID: Independent, identically distributed; LT: Laplace transform; ML: Mittag-Leffler; MLE: maximum likelihood estimator; PDF: probability density function; PGF: Probability generating function; PMF: Probability mass function; SF: Survival function

Availability of data and materials

The data set used in this study was obtained from Investing.com and is available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

References

Abramowitz, M., Stegun, I. A.: Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. Dover Publications, New York (1972)

Arnold, B. C.: Flexible univariate and multivariate models based on hidden truncation. J. Statist. Plann. Infer. **139**, 3741–3749 (2009)

Amponsah, C. K.: A Bivariate Gamma Mixture Discrete Pareto Distribution. Thesis, University of Nevada (2017)

Amponsah, C. K., Kozubowski, T. J.: A computational approach to estimation of discrete Pareto parameters, working paper (2021)

Arendarczyk, M., Kozubowski, T. J., Panorska, A. K: A bivariate distribution with Lomax and geometric margins. J. Korean Statist. Soc. **47**, 405–422 (2018)

Bardwell, G. E., Crow, E. L.: A two-parameter family of hyper-Poisson distributions. J. Amer. Statist. Assoc. **59**(305), 133–141 (1964)

Barreto-Souza, W.: Bivariate gamma-geometric law and its induced Lévy process. J. Multivar. Anal. **109**, 130–145 (2012)

Barreto-Souza, W., Silva, R. B.: A bivariate infinitely divisible law for modeling the magnitude and duration of monotone periods of log-returns. Statist. Neerlandica. **73**, 211–233 (2019)

Bickel, P. J., Doksum, K. A.: Mathematical Statistics: Basic Ideas and Selected Topics, Vol I, 2nd ed. Chapman & Hall/CRC Press, Boca Raton (2015)

Biondi, F., Kozubowski, T. J., Panorska, A. K.: A new model for quantifying climate episodes. Intern. J. Climatol. **25**, 1253–1264 (2005)

Biondi, F., Kozubowski, T. J., Panorska, A. K., Saito, L: A new stochastic model of episode peak and duration for eco-hydro-climatic applications. Ecol. Model. **211**, 383–395 (2008)

Buddana, A., Kozubowski, T. J.: Discrete Pareto distribution. Econ. Qual. Control. **29**, 143–156 (2014)

Haubold, H. J., Mathai, A. M., Saxena, R. K.: Mittag-Leffler functions and their applications. J. Appl. Math. **2011**, ID 298628 (2011)

Kozubowski, T. J., Panorska, A. K.: A Mixed bivariate distribution with exponential and geometric marginals. J. Statist. Plann. Infer. **134**, 501–520 (2005)

Kozubowski, T. J., Panorska, A. K., Biondi, F.: Mixed multivariate models for random sums and maxima. In: SenGupta, A. (ed.) Advances in Multivariate Statistical Methods, Vol 4, pp. 145–171. World Scientific, Singapore, (2009)

Kozubowski, T. J., Panorska, A. K., Podgórski, K.: A bivariate Lévy process with negative binomial and gamma marginals. J. Multivar. Anal. **99**, 1418–1437 (2008)

Kozubowski, T. J., Panorska, A. K.: A note on the joint distribution involving Poissonian sum of exponential variables. Adv. Appl. Statist. Sci. **1**(2), 157–165 (2010)

Krishna, H., Singh Pundir, P.: Discrete Burr and discrete Pareto distributions. Statist. Meth. **6**, 177–188 (2009)

Laforgia, A., Natalini, P.: Exponential, gamma and polygamma functions: Simple proofs of classical and new inequalities. J. Math. Anal. Appl. **407**, 495–500 (2013)

Noack, A.: A class of random variables with discrete distributions. Ann. Math. Statist. **21**(1), 127–132 (1950)

Robert, C. Y., Segers, J.: Tails of random sums of a heavy-tailed number of light-tailed terms. Insur. Math. Econ. **43**, 85–92 (2008)

Publisher's Note