PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 29, Fasc. 1 (2009), pp. 43–71

DISTRIBUTIONAL PROPERTIES OF THE NEGATIVE BINOMIAL LÉVY PROCESS

BY

TOMASZ J. KOZUBOWSKI (RENO) AND KRZYSZTOF PODGÓRSKI (LUND)

Abstract. The geometric distribution leads to a Lévy process parameterized by the probability of success. The resulting negative binomial process (NBP) is a purely jump and non-decreasing process with general negative binomial marginal distributions. We review various stochastic mechanisms leading to this process, and study its distributional structure. These results enable us to establish strong convergence of the NBP in the supremum norm to the gamma process, and lead to a straightforward algorithm for simulating sample paths. We also include a brief discussion of estimation of the NPB parameters, and present an example from hydrology illustrating possible applications of this model.

2000 AMS Mathematics Subject Classification: Primary: 60G51; Secondary: 60G50, 60E07.

Key words and phrases: Borehole data; cluster Poisson process; compound Poisson process; count data: Cox process; discrete Lévy process; doubly stochastic Poisson process; fractures; gamma-Poisson process; gamma process: geometric distribution; immigration birth process; infinite divisibility; logarithmic distribution: over-dispersion; Pascal distribution; point process; random time transformation; subordination; simulation.

1. PRELIMINARIES

The Poisson distribution and the corresponding Lévy process is the most basic and widely used stochastic model for count data. However, empirical count data often exhibit *overdispersion* – the term that is used when the sample variance is larger than the sample mean. In such a case the standard Poisson model is inappropriate (see, e.g., [38], [62], [93], [102], [103]) and a common solution to this problem, which goes back to [54], involves randomization of the Poissonian mean leading to continuous mixtures of Poisson processes. A frequent and convenient choice of the mixing measure is the gamma distribution. This leads to an explicit expression for the resulting probability distribution parameterized by $p \in [0, 1]$ and $t > 0$ with the characteristic function (ChF)

$$\phi_{NB}(u;p,t) = \left(\frac{p}{1-(1-p)e^{iu}}\right)^t, \quad u \in \mathbb{R}, \tag{1.1}$$

and the probability mass function (PMF)

$$\mathbb{P}_{NB}(k;p,t) = \binom{t+k-1}{k} p^t (1-p)^k, \quad k \in 0,1,2,\ldots$$

The resulting generalized negative binomial (NB) distribution is a popular stochastic model in many areas of research [65]. Applications of NB models include theory of accidents ([3], [46], [54], [63], [76]), population growth processes and epidemiology ([49], [70], [79], [80], [84], [89], [108]), particle physics ([4], [26], [29], [34], [49], [90]), geosciences ([67], [71]), cosmology [25], psychology [98], economics [101], library science ([17]–[21], [51]), marketing ([27], [28], [39], [52]), ecology ([40], [41], [77], [78]), entomology ([85], [105]), human geography [31], environmental science [64], software reliability ([94], [95]), and biology ([9], [30], [48], [82], [102]). For more extensive reviews of the NB distribution with many references see [7] and [65].

There are numerous stochastic mechanisms leading to an NB distribution – over ten of them can be found in [11]. The four most common ones are the following (see [2]):

• *Inverse binomial sampling* ([58], [107]). The waiting time till the t^{th} success (measured as the number of failures) in an infinite sequence of Bernoulli trials with success probability p has the NB distribution (1.1).

• *Heterogeneous Poisson sampling* [54]. This is the randomization of the Poissonian mean discussed above. If the mean of a Poisson distribution has a gamma distribution with shape parameter $t > 0$ and scale parameter $(1-p)/p$, then the resulting mixed Poisson distribution is the NB distribution (1.1).

• *Randomly distributed colonies* ([76], [89]). This is a compound Poisson representation of the NB distribution. If groups of individuals are distributed randomly in space (or time), the number of colonies has a Poisson distribution with mean $-t\ln p$, and if the numbers of individuals in the colonies are distributed independently with a logarithmic distribution given by the PMF

$$\mathbb{P}(k;p) = -\frac{(1-p)^k}{k\ln p}, \quad k \in \mathbb{N}, \tag{1.2}$$

the total number of the individuals in all colonies has the NB distribution (1.1).

• *Stationary distribution arising in Markov population processes* ([70], [79]). The equilibrium distribution of a stationary Markov population process with constant rates of birth (λ), death ($\mu > \lambda$), and immigration (ν) is NB with $p = \lambda/\mu$ and $t = \nu/\mu$.

The last three of the above examples lead to stochastic processes with the NB marginal distributions (and some generalizations). Discrete-time processes with marginal NB distributions were discussed in [74], [81], [99]. These models are based on the fact that a Poisson process stopped at a random gamma distributed time has an NB distribution. Various continuous-time NB processes follow this idea, including those considered in [6]. *Pascal processes* in [15] offer tractable models for the number of opportunities in certain investment problems (see also [14]). Continuous-time population growth models with marginal NB distributions go back to [79] and [80], followed by [42], [49], [70] and [108] (see also historical comments in [43] and more recent [32]). Although these models were originally motivated by applications in biology and spread of epidemics/contagious diseases, they have been applied in many areas of science, including particle physics, where they justify the NB model for the multiplicity distributions of high-energy particle interactions (see, e.g., [4], [29], [49], [90] in this connection). As noted by several authors (see, e.g., [23], [24], [46]), this "contagious" interpretation of the NB distribution leads to the same (non-ergodic) stochastic process as the one obtained by heterogeneous Poisson sampling. For example, the NB Pólya process (a pure birth process with intensity of birth $\lambda_n(t) = (k + n)/(1 + t)$, depending on both t and n, which is the population size at time t) and mixed Poisson process whose intensity has a standard gamma distribution with the shape parameter equal to k have the same marginal NB distributions. This is referred to as "contagion-stratification duality" in [46], where different interpretations of the NB distributions and processes are discussed. A mixed Poisson process with gamma distributed intensity is a well-known construction (see, e.g., [55], [100]), similar in spirit to random hazard rate models for heterogeneous populations in survival analysis (see, e.g., [1], [22], [59], [60], [68], [91]). For obvious reason, this model is known as *gamma-Poisson* process (see also [12], [13], [20], and [83], where more general shot-noise Cox processes are considered). However, not many authors have noticed that sample realizations of such processes look Poisson ones – the variation is not *within*, but *between* processes. Consequently, these models are not appropriate to describe spatial data with empirical distribution of counts in disjoint sets resembling the NB distribution. In [62], one of the exceptions, there was offered an alternative: use a Poisson process with a *random time scale*, that is, subordinate it to another (independent) non-decreasing process.

The process we discuss in this paper can be defined equivalently through three different stochastic mechanisms. First, it is clear from (1.1) that the NB distributions are infinitely divisible and lead to a continuous time process with independent and homogeneous increments whose one-dimensional distributions are NB (see, e.g., [44], pp. 179–182). We refer to it as the Negative Binomial (Lévy) Process (NBP) with parameter p, denoted by $NB(t)$ (or by $NB_p(t)$ to emphasize the dependence on p). The NBP is integer-valued, non-decreasing, and consequently a pure jump process, whose mean and variance are linear in t: $\mathbb{E}NB(t) = t \cdot q/p$ and $\mathbb{V}arNB(t) = t \cdot q/p^2$. Thus, in contrast to the Poisson process, here the variance exceeds the mean, which is known as overdispersion. This process appeared in [16], where its equivalent compound Poisson representation has been established as well as the limiting Poisson distribution (in a triangular scheme). This representation of the NBP as a standard Poisson process compounded with integer-valued logarithmically distributed clusters has been also discussed in [62], [75], [100], and as such is known as the *Compound Software Reliability Model* in software reliability community (see [94], [95]). However, the process hardly ever appears in the literature in an explicit form, and when it does, its equivalent representations are rarely noticed or discussed (see, e.g., [37], [100]). The third equivalent way of obtaining the NBP is through subordination of Poisson process with intensity $\lambda = 1/p - 1$ to a standard gamma process. This construction is known in the literature as the *gamma-Poisson process* (see [12], [13], [83]). Equivalently, it is a *doubly stochastic* Poisson process (Cox process; see, e.g., [53], [97]) whose intensity is an independent gamma process.

All above constructions are standard in defining Lévy processes and these have been studied extensively in recent years, in connection with financial applications, as seen in recent monographs [33] and [96]. In fact, this process has appeared as a subordinator in a compound Cox process in [69] in connection with option pricing, although it was neither defined nor studied there. Let us also note the class of Poisson processes subordinated to the *Hougaard family* studied in [75] of which the process studied in this paper is an important (limiting) special case (see [5], [60], [61], [66] for more details on the Hougaard family).

The equivalent representations of the NBP discussed above are scattered in the literature (see, e.g., [43], pp. 155–157, 271, [44], pp. 348–349, and [37], [106]). For the sake of future reference, we summarize them below.

PROPOSITION 1.1. *The following three stochastic processes are equivalent in distribution*:

(i) *Lévy process corresponding to the semigroup of the NB ChF* (1.1);

(ii) *subordinated Poisson process with intensity* $\lambda = (1-p)/p$ *with standard gamma subordinator*;

(iii) *compound Poisson process* $\sum_{n=1}^{N(t)} X_i$, *where* $N(t)$ *is a Poisson process with intensity* $\lambda = -\ln p$ *and the* $\{X_i\}$ *are IID logarithmic random variables with PMF* (1.2).

The equivalence follows easily by comparing the relevant ChFs for $t = 1$, which is enough as all three processes are Lévy ones. Let us note that the ChF of the NBP admits the representation

$$\phi_{NB(t)}(u) = \exp\left(t \int (e^{ixu} - 1) d\Lambda(x)\right),$$

where the (discrete) Lévy measure Λ is given by

$$\Lambda = \sum_{k=1}^{\infty} \frac{q^k}{k} \delta_{\{k\}}$$

and $\delta_{\{k\}}$ denotes a point mass at k.

Our main results are the distributional properties of the NBP. They lead to interesting representations of this process as well as to convenient simulation algorithms of its sample paths. In Section 2 we analyze distributional structure of the NBP. Simulation algorithms are presented in Section 3, where we also include new results on approximating the gamma process by an NBP. Brief remarks on further properties and estimation are discussed in Section 4. Finally, in Section 5 we present an example from hydrology illustrating the modeling potential of the NBP.

2. DISTRIBUTIONAL STRUCTURE OF THE NBP

The NBP is a pure jump process that has positive integer jump sizes. The following series representation of the NBP, which applies to a general Lévy process (see, e.g., [45]), is a simple consequence of its compound Poisson representation. We have

$$NB(t) = \sum_{i=1}^{\infty} J_i \mathbf{1}_{[\Gamma_i,\infty)}(t),$$

where $\Gamma_i = E_1 + \ldots + E_i$, the $\{E_j\}$ are IID exponential RVs with parameter $\lambda = \ln(1/p)$, while the $\{J_i\}$ are IID discrete logarithmic random variables, independent of the $\{E_j\}$. This implies that the jumps of the NBP occur at the same instants, Γ_i, as the jumps of the Poisson process defined through the interarrival times E_j. However, the sizes of the jumps are random, and distributed according to the logarithmic distribution.

Our next two results provide more insight into the distributional structure of the NBP.

LEMMA 2.1. *For each* $t, s \geqslant 0$, *define the increment process by* $\Delta_t(s) = NB(t+s) - NB(t)$. *For* $u > 0$, *the conditional distribution of* $\Delta_t(s)$, $s \in [0,u]$, *given* $\Delta_t(u)$, *is free of the parameter* p. *Further, for each* $r \in \mathbb{N}$, $0 = s_0 \leqslant s_1 \leqslant \ldots \leqslant s_r \leqslant s_{r+1} = u$, *and* $0 = n_0 \leqslant n_1 \leqslant \ldots \leqslant n_r \leqslant n_{r+1} = n$, *we have*

$$\mathbb{P}\big(\Delta_t(s_1) = n_1, \ldots, \Delta_t(s_r) = n_r | \Delta_t(u) = n\big) = \prod_{i=1}^{r+1} \binom{d_i + k_i - 1}{k_i} \Big/ \binom{u+n-1}{n}, \tag{2.1}$$

where $d_i = s_i - s_{i-1}$, $k_i = n_i - n_{i-1}$, $i = 1, \ldots, r+1$.

P r o o f. By the homogeneity and independence of the increments, the conditional probability in (2.1) is given by

$$\frac{\mathbb{P}\big(\Delta_{s_0}(d_1) = k_1, \ldots, \Delta_{s_r}(d_{r+1}) = k_{r+1}\big)}{\mathbb{P}\big(\Delta_{s_0}(s_{r+1}) = n_{r+1}\big)} = \prod_{i=1}^{r+1} \frac{\mathbb{P}\big(\Delta_{s_{i-1}}(d_i) = k_i\big)}{\mathbb{P}\big(\Delta_{s_0}(s_{r+1}) = n_{r+1}\big)}$$

$$= \frac{\prod_{i=1}^{r+1} \binom{d_i + k_i - 1}{k_i} p^{d_i} q^{k_i}}{\binom{s_{r+1} + n_{r+1} - 1}{n_{r+1}} p^{s_{r+1}} q^{n_{r+1}}} = \prod_{i=1}^{r+1} \binom{d_i + k_i - 1}{k_i} \Big/ \binom{u+n-1}{n}. \quad \blacksquare$$

For each $u > 0$ and $n = 0, 1, 2, \dots$, equation (2.1) defines a non-decreasing, integer-valued stochastic process $X_n(s)$, $s \in [0, u]$, such that $X_n(0) = 0$ and $X_n(u) = n$. From now on we assume that X_n denotes such a process for $u = 1$. The following result is a consequence of the above lemma.

COROLLARY 2.1. *For any sequence $\mathbf{k} = (k_1, \dots, k_r)$ of non-negative integers that add up to n and a set $\mathbf{I} = \{I_1, \dots, I_r\}$ of disjoint intervals in $[0, 1]$ of the respective lengths $|I_1|, \dots, |I_r|$, let $X_n(\cdot) \in A(\mathbf{k}, \mathbf{I})$ denote the event that the process $X_n(\cdot)$ jumps by exactly k_i over the intervals I_i, $i = 1, \dots, r$. Then*

$$\mathbb{P}\big(X_n(\cdot) \in A(\mathbf{k}, \mathbf{I})\big) = \prod_{i=1}^{r} \binom{|I_i| + k_i - 1}{k_i}. \tag{2.2}$$

Proof. Note that since $k_1 + \ldots + k_r = n$, there are no jumps inside the set $[0, 1] \setminus \bigcup_{i=1}^{r} I_i$ and

$$\mathbb{P}\big(X_n(\cdot) \in A(\mathbf{k}, \mathbf{I})\big) = \mathbb{P}\big(NB(s_1) = n_1, \dots, NB(s_l) = n_l | NB(1) = n\big),$$

where $\{s_k, k = 0, \dots, l + 1\}$ is the set of ordered endpoints of the intervals I_i that also includes $s_0 = 0$ and $s_{l+1} = 1$. The factors on the right-hand side of (2.1) for which $k_i = 0$ are equal to one, and thus can be dropped from the formula. The remaining factors produce the right-hand side of (2.2). ∎

Independent copies of the processes X_n describe the behavior of the NBP on intervals between integer values as shown in the following representation.

THEOREM 2.1. *Let G_k be a sequence of IID geometric random variables and let $\{Z_k(s), s \in [0, 1]\}$ be a sequence of processes defined by*

$$Z_k(s) = X^{(k)}_{G_{k+1}}(s),$$

where, conditionally on $G_{k+1} = n_k$, the processes $X^{(k)}_{n_k}(s)$ are mutually independent versions of the processes $X_n(s)$ as defined by (2.1), *and independent of all the variables G_i with $i \neq k + 1$. Then the NBP can be written as*

$$Y(t) = \sum_{i=1}^{[t]} G_i + Z_{[t]}(t - [t]). \tag{2.3}$$

Proof. First, note that $Y(k) = \sum_{i=1}^{k} G_i$ coincides with the NBP for each $k \in \mathbb{N}$. Further, it follows from the definition and Lemma 2.1 that the conditional distribution of the increments $Y(k + s) - Y(k)$, $s \in [0, 1]$, given $Y(k + 1) = n$, is the same as that of the NBP. Finally, by independence of the terms of the sequence $X^{(k)}_{n_k}$, $k \in \mathbb{N}$, the distribution of arbitrary increments of $Y(t)$ is the same as that of the NBP. ∎

The above representation of the NBP consists of two different components: a parametric one, defined at integer time points (this is, the sum of geometric variables) and a non-parametric portion, represented by the processes $Z_k(\cdot)$. The latter describes the behavior of the NBP between integers. Our next result provides a more explicit description of the stochastic process $X_n(\cdot)$. Here, we use the following notation and terminology. A random sequence $\boldsymbol{\delta}_k = (\delta_1^k, \ldots, \delta_k^k)$ of 0-1 vectors such that exactly one coordinate is equal to one is called a *uniform selector* if $\mathbb{P}(\delta_i^k = 1) = 1/k$, $i = 1, 2, \ldots, k$, $k = 1, 2, \ldots$ The inner product $\boldsymbol{\delta}_k \cdot \mathbf{x} = \delta_1^k x_1 + \ldots + \delta_k^k x_k$ can be thought of as a uniformly random selection of a single coordinate of the vector $\mathbf{x}$.

THEOREM 2.2. *Let $(U_k)_{k\in\mathbb{N}}$ be a sequence of IID standard uniform random variables. Then the process $X_n(s)$, $s \in [0,1]$, has the same distribution as*

$$Y_n(s) = \sum_{k=1}^{n} \mathbf{1}_{[V_k,1]}(s), \tag{2.4}$$

where the sequence of jump positions, $(V_k)_{k\in\mathbb{N}}$, is defined recursively as follows:

$$V_k = \boldsymbol{\delta}_k \cdot (V_1, \ldots, V_{k-1}, U_k), \qquad k = 1, 2, \ldots, n,$$

and $(\boldsymbol{\delta}_k)$ is a sequence of independent uniform selectors, independent of the sequence $(U_k)_{k\in\mathbb{N}}$.

Proof. Since the processes X_n and Y_n are non-decreasing and integer valued on $[0,1]$, their distributions are uniquely defined by the probabilities of the events $A(\mathbf{k},\mathbf{I})$. We show by induction that these probabilities coincide.

For $n = 1$, it is enough to take $A(k_1, I_1)$ with $k_1 = 1$. We have

$$\mathbb{P}\big(X_n(\cdot) \in A(1, I_1)\big) = |I_1| = \mathbb{P}\big(Y_n(\cdot) \in A(1, I_1)\big).$$

Assume now that $X_n(\cdot)$ has the same distribution as $Y_n(\cdot)$ for each $n \leqslant l$. Let $n = l+1$. Since the processes X_n and Y_n have exactly n jumps, it is enough to consider $A(\mathbf{k},\mathbf{I})$ with coordinates of $\mathbf{k}$ non-zero. Then for $n = l+1$ we have

$$\begin{aligned}
\mathbb{P}\big(Y_n(\cdot) \in A(\mathbf{k},\mathbf{I})\big) &= \mathbb{P}\big(Y_{n-1}(\cdot) + \mathbf{1}_{[V_n,1]}(\cdot) \in A(\mathbf{k},\mathbf{I})\big) \\
&= \sum_{j=1}^{r} \mathbb{P}\big(Y_{n-1}(\cdot) + \mathbf{1}_{[V_n,1]}(\cdot) \in A(\mathbf{k},\mathbf{I}), V_n \in I_j\big) \\
&= \sum_{j=1}^{r} \mathbb{P}\Big(Y_{n-1}(\cdot) \in A\big((k_1, \ldots, k_j - 1, \ldots, k_r), \mathbf{I}\big), V_n \in I_j\Big) \\
&= \sum_{j=1}^{r} \mathbb{P}\big(Y_n(\cdot) \in A(\mathbf{k},\mathbf{I}), U_n \in I_j, V_n = U_n\big) \\
&\quad + \sum_{j=1}^{r} \mathbb{P}\big(Y_n(\cdot) \in A(\mathbf{k},\mathbf{I}), V_n \in I_j, V_n = V_k, \exists k < n\big).
\end{aligned}$$

Observe that, by definition, V_n is selected as U_n independently of everything else with probability $1/n$, and the probability that U_n is in I_j is equal to $|I_j|$. On the other hand, if V_n is one of the $\{V_k\}$, $k = 1, \ldots, n-1$, then it is in I_j with probability $(k_j - 1)/n$ (as there is $k_j - 1$ jumps of Y_{n-1} in I_j and the selector $\boldsymbol{\delta}_n$ selects any one of them with probability $1/n$). Thus, using (2.2), we infer that the above probability is equal to

$$\sum_{j=1}^{r} \mathbb{P}\Big(Y_{n-1}(\cdot) \in A\big((k_1, \ldots, k_j - 1, \ldots, k_r), \mathbf{I}\big)\Big) \frac{|I_j| + k_j - 1}{n}$$
$$= \sum_{j=1}^{r} \frac{|I_j| + k_j - 1}{n} \binom{|I_1| + k_1 - 1}{k_1} \cdots \binom{|I_j| + k_j - 2}{k_j - 1} \cdots \binom{|I_r| + k_r - 1}{k_r}$$
$$= \sum_{j=1}^{r} \frac{k_j}{n} \binom{|I_1| + k_1 - 1}{k_1} \cdots \binom{|I_j| + k_j - 1}{k_j} \cdots \binom{|I_r| + k_r - 1}{k_r}$$
$$= \Big(\sum_{j=1}^{r} \frac{k_j}{n}\Big) \cdot \binom{|I_1| + k_1 - 1}{k_1} \cdots \binom{|I_r| + k_r - 1}{k_r}$$
$$= \binom{|I_1| + k_1 - 1}{k_1} \cdots \binom{|I_r| + k_r - 1}{k_r},$$

which proves the induction step. This concludes the proof. ■

2.1. An immigration and birth process. Although $X_n(\cdot)$ is a conditional process arising from the NBP, it has its own merit. It can be viewed as the following immigration and birth process. Let the interval $[0, 1]$ represent a habitat for a population of individuals that can either immigrate from the outside and then locate randomly in $[0, 1]$ according to the uniform distribution, or can give birth to a child that stays with the parent to build a cluster (family) at some point in $[0, 1]$. Assuming that the chances of an individual to give birth are the same as that for immigration of a newcomer, the process $X_n(\cdot)$ represents the spatial distribution of families at the moment when the total population is n. Similar models related to the negative binomial distribution and their applications to the theory of avalanches were discussed in [8]. Notice that the number of clusters, K_n, their sizes, $X_1^{(n)}, \ldots, X_{K_n}^{(n)}$, and their locations, $W_1^{(n)} < \ldots < W_{K_n}^{(n)}$, are random, and

$$X_n(s) = \sum_{i=1}^{K_n} X_i^{(n)} \mathbf{1}_{[W_i^{(n)}, 1]}(s).$$

In the next result, we use the compound Poisson representation of the NBP to obtain the joint distribution of $K_n, X_1^{(n)}, \ldots, X_{K_n}^{(n)}$ and $W_1^{(n)} < \ldots < W_{K_n}^{(n)}$.

THEOREM 2.3. *Let C_k^n be the set of* 0-1 *sequences $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ with exactly k ones and $n - k$ zeros* (*combinations of k out of n*). *The distribution of* *the number of clusters, K_n, is given by*

$$\mathbb{P}(K_n = k) = \frac{1}{n!} \sum_{\boldsymbol{\sigma} \in \mathcal{C}_k^n} \prod_{i=1}^{n} (i-1)^{1-\sigma_i}, \qquad k = 1, 2, \ldots, n. \tag{2.5}$$

The vectors $\mathbf{X}^{(n)} = (X_1^{(n)}, \ldots, X_k^{(n)})$ and $\mathbf{W}^{(n)} = (W_1^{(n)}, \ldots, W_k^{(n)})$, conditionally on $K_n = k$, are independent. Further, given $K_n = k$, $\mathbf{W}^{(n)}$ has the same distribution as that of the order statistics of k IID standard uniform random variables, while the distribution of $\mathbf{X}^{(n)}$ is given by

$$\mathbb{P}(X_1^{(n)} = i_1, \ldots, X_k^{(n)} = i_k | K_n = k) = C \prod_{j=1}^{k} \frac{1}{i_j},$$

where $i_1 + \ldots + i_k = n$ and $C = \sum_{r_1+\ldots+r_k=n} \prod_{j=1}^{k} 1/r_j$.

P r o o f. The number of clusters K_n is equal to the number of times the independent selectors $\boldsymbol{\delta}_k$, $k = 1, \ldots, n$, have the value one at the last coordinate. The probability of such an event is given by

$$\sum_{\boldsymbol{\sigma} \in \mathcal{C}_k^n} \prod_{i=1}^{n} \frac{1}{i^{\sigma_i}} \left(1 - \frac{1}{i}\right)^{1-\sigma_i},$$

which is equivalent to (2.5). Next, using the representation (iii) of Proposition 1.1, we have

$$\begin{aligned}
\mathbb{P}\big(X_n(\cdot) \in A | K_n = k\big) &= \mathbb{P}\big(NB(\cdot) \in A | NB(1) = n, N(1) = k\big) \\
&= \mathbb{P}\big(\sum_{i=1}^{k} X_i \mathbf{1}_{[E_i,1]}(\cdot) \in A | X_1 + \ldots + X_k = n, N(1) = k\big) \\
&= \mathbb{P}\big(\sum_{i=1}^{k} X_i \mathbf{1}_{[U_i,1]}(\cdot) \in A | X_1 + \ldots + X_k = n\big),
\end{aligned}$$

where the $\{X_i\}$ are IID logarithmic random variables and the $\{U_i\}$ are IID standard uniform random variables, independent of everything else. This shows that $\mathbf{W}^{(n)}$ has the same distribution as the order statistics connected with k IID standard uniform variables. The above implies that, conditionally on $K_n = k$, the vectors $\mathbf{X}^{(n)}$ and $\mathbf{W}^{(n)}$ are independent, and the distribution of $\mathbf{X}^{(n)}$ is equivalent to the distribution of $(X_1, \ldots, X_k)$ given $X_1 + \ldots + X_k = n$. Let $i_1 + \ldots + i_k = n$. Then

$$\begin{aligned}
&\mathbb{P}(X_1 = i_1, \ldots, X_k = i_k | X_1 + \ldots + X_n = n) \\
&\qquad = \frac{\prod_{j=1}^{k} \ln^{-1}(1/p) q^{i_j} / i_j}{\sum_n \prod_{j=1}^{k} \ln^{-1}(1/p) q^{r_j} / r_j} = \frac{\prod_{j=1}^{k} 1/i_j}{\sum_n \prod_{j=1}^{k} 1/r_j},
\end{aligned}$$

where the summation runs over $n = r_1 + \ldots + r_k$. This concludes the proof. ■

3. APPROXIMATIONS OF GAMMA PROCESS AND SIMULATIONS

3.1. Convergence to the gamma process. The probability distribution of a negative binomial random variable (1.1) multiplied by p converges weakly to a standard gamma distribution with shape parameter t. Similarly, the finite-dimensional distributions of $pNB_p(\cdot)$, where $NB_p(\cdot)$ is the NBP, converge to those of the standard gamma process (see [73]). Using general convergence results for Lévy processes (see, for example, Theorem V.19 and Example VI.18 in [86]) one can establish weak convergence of $pNB_p(\cdot)$ to $\Gamma(\cdot)$ in the Skorokhod J_1 metric. However, stronger convergence results can be obtained by using the distributional representation of the NBP discussed above. Namely, a proper version of the NBP is convergent *in the supremum norm* with *probability one* over a compact set to the so-called shot noise representation of the gamma process (see [10]). We also provide the rate of convergence and discuss the upper bound for the norm.

We start with the description of a version of the NBP for which the almost sure convergence holds. Recall that a logarithmic random variable X can be represented in the form

$$X \stackrel{d}{=} [1 - W/\ln(1 - p^{1-V})],$$

where $[x]$ is the integer part of x, and W and V are mutually independent variables with standard exponential and uniform distributions, respectively (see [36]). Thus, the compound Poisson interpretation of the NBP leads to the following representation of the NBP on the interval $[0, 1]$:

$$NB(t;p) = \sum_{k=1}^{N_p} \left[1 - \frac{W_k}{\ln(1 - p^{1-\Gamma_k/\Gamma_{N_p+1}})} \right] \mathbf{1}_{[U_k,1)}(t). \tag{3.1}$$

Here, $N_p = N(\ln 1/p)$, where $N(t)$ is a standard Poisson process, the $\{W_k\}$ are IID standard exponential variables, the $\{\Gamma_k\}$ are the arrivals of another standard Poisson process, and the $\{U_k\}$ are IID standard uniform variables, all mutually independent. We used the fact that the order statistics of uniform variables V_k, $k = 1, \ldots, n$, have the same distribution as that of Γ_k/Γ_{n+1}, $k = 1, \ldots, n$. Since the number of jumps in this representation is equal to n_p, we describe this representation conditional on the number of jumps, as opposed to the one based on the positions of the jumps described in Theorems 2.2 and 2.3.

Using the same notation, we can write the so-called shot noise representation of a standard gamma process (see [92]):

$$G(t) = \sum_{k=1}^{\infty} e^{-\Gamma_k} W_k \mathbf{1}_{[U_k,1)}(t), \qquad t \in [0, 1]. \tag{3.2}$$

THEOREM 3.1. *Let $NB(t;p)$ and $G(t)$ be defined as in* (3.1) *and* (3.2), *respectively. Then for $p \to 0$ with probability one we have*

$$\sup_{t\in[0,1]} |G(t) - pNB(t;p)| = O\biggl(\sqrt{\frac{\ln\ln\ln 1/p}{\ln 1/p}}\biggr).$$

P r o o f. Consider the process

$$N(t;p) = p\sum_{k=1}^{N_p}\biggl(1 - \frac{W_k}{\ln(1-p^{1-\Gamma_k/\Gamma_{N_p+1}})}\biggr)\mathbf{1}_{[U_k,1)}(t).$$

We obviously have

$$|G(t) - pNB(t;p)| \leqslant |G(t) - N(t;p)| + p.$$

We examine the asymptotics of the following 4 components of the difference

$$\begin{aligned}
&G(t) - N(t;p)\\
&= \sum_{k=N_p+1}^{\infty}\frac{W_k}{e^{\Gamma_k}}\mathbf{1}_{[U_k,1)}(t) + \sum_{k=1}^{N_p}\biggl[\frac{W_k}{e^{\Gamma_k}} - p + \frac{pW_k}{\ln(1-p^{1-\Gamma_k/\Gamma_{N_p+1}})}\biggr]\mathbf{1}_{[U_k,1)}(t)\\
&= S_1(t;p) + p\sum_{k=1}^{N_p}(W_k - 1)\mathbf{1}_{[U_k,1)}(t) + \sum_{k=1}^{N_p}\frac{W_k}{e^{\Gamma_k}}\biggl[1 - \frac{\lambda(pe^{\Gamma_k(1+o_p)})}{e^{\Gamma_k o_p}}\biggr]\mathbf{1}_{[U_k,1)}(t)\\
&= S_1(t;p) + S_2(t;p) + \sum_{k=1}^{N_p}W_k\biggl[\frac{1-e^{-\Gamma_k o_p}}{e^{\Gamma_k}} - \frac{\lambda(pe^{\Gamma_k(1+o_p)}) - 1}{e^{\Gamma_k(1+o_p)}}\biggr]\mathbf{1}_{[U_k,1)}(t)\\
&= S_1(t;p) + S_2(t;p) + S_3(t;p) - p\sum_{k=1}^{N_p}W_k\frac{\lambda(pe^{\Gamma_k(1+o_p)}) - \lambda(0)}{pe^{\Gamma_k(1+o_p)}}\mathbf{1}_{[U_k,1)}(t)\\
&= S_1(t;p) + S_2(t;p) + S_3(t;p) - S_4(t;p),
\end{aligned}$$

where

$$o_p = \frac{\ln 1/p}{\Gamma_{N_p+1}} - 1, \qquad \lambda(x) = x\frac{\ln(1-x) - 1}{\ln(1-x)},$$

$$S_1(t;p) = \sum_{k=N_p+1}^{\infty} e^{-\Gamma_k}W_k\mathbf{1}_{[U_k,1)}(t),$$

$$S_2(t;p) = p\sum_{k=1}^{N_p}(W_k - 1)\mathbf{1}_{[U_k,1)}(t),$$

$$S_3(t;p) = \sum_{k=1}^{N_p}\frac{W_k}{e^{\Gamma_k}}(1 - e^{-\Gamma_k o_p})\mathbf{1}_{[U_k,1)}(t),$$

$$S_4(t;p) = p\sum_{k=1}^{N_p}W_k\frac{\lambda(pe^{\Gamma_k(1+o_p)}) - \lambda(0)}{pe^{\Gamma_k(1+o_p)}}\mathbf{1}_{[U_k,1)}(t).$$

First, we show that, for each $\epsilon > 0$, with probability one we have

$$S_1(p) \stackrel{\text{def}}{=} \sup_{t\in[0,1]} S_1(t;p) = o\left(p \cdot \exp\left(\ln^{1/2+\epsilon}\frac{1}{p}\right)\right) \quad \text{as } p \to 0. \tag{3.3}$$

By Lemma 6.4 in the Appendix, there is a set of probability one on which for sufficiently large t we have

$$t - t^{1/2+\epsilon/2} < N(t) < t + t^{1/2+\epsilon/2}.$$

By Lemma 6.3 in the Appendix with $a_n = n^{\epsilon/2}$, on this set and for sufficiently small p we have

$$\begin{aligned}
&\frac{S_1(p)}{p\cdot\exp(\ln^{1/2+\epsilon}1/p)}\\
&\quad\leqslant \frac{\exp\left(-N(\ln 1/p) + N^{1/2+\epsilon/2}(\ln 1/p)\right)}{p\cdot\exp(\ln^{1/2+\epsilon}1/p)}\\
&\quad\leqslant \frac{\exp\left(-\ln 1/p + \ln^{1/2+\epsilon/2}1/p + (\ln 1/p + \ln^{1/2+\epsilon/2}1/p)^{1/2+\epsilon/2}\right)}{p\cdot\exp(\ln^{1/2+\epsilon}1/p)}\\
&\quad= \exp\left(-\ln^{1/2+\epsilon}\frac{1}{p} + \ln^{1/2+\epsilon/2}\frac{1}{p} + \left(\ln\frac{1}{p} + \ln^{1/2+\epsilon/2}\frac{1}{p}\right)^{1/2+\epsilon/2}\right)\\
&\quad\leqslant \exp\left(-\ln^{1/2+\epsilon}\frac{1}{p} + 3\ln^{1/2+\epsilon/2}\frac{1}{p}\right),
\end{aligned}$$

where in the last inequality we assumed additionally that $\epsilon \leqslant 1$. The convergence to zero of the last term is obvious.

The asymptotics of $S_2(t;p)$ can be obtained directly from the law of large numbers. Namely, we show that with probability one

$$S_2(p) \stackrel{\text{def}}{=} \sup_{t\in[0,1]} S_2(t;p) = O\left(p\ln 1/p\right). \tag{3.4}$$

Indeed,

$$\frac{|S_2(p)|}{p\ \ln 1/p} \leqslant \frac{N(\ln 1/p)}{\ln 1/p}\frac{\sum_{k=1}^{N_p}|W_k-1|}{N_p},$$

with the right-hand side converging with probability one to $\mathbb{E}|W_1 - 1|$.

Next we turn to the asymptotics of $S_3(t;p)$. We will show that

$$S_3(p) \stackrel{\text{def}}{=} \sup_{t\in[0,1]} S_3(t;p) = o\left(\sqrt{\frac{\ln\ln\ln 1/p}{\ln 1/p}}\right). \tag{3.5}$$

From Taylor's first order approximation, there is a random variable $\rho_{p,k} \in [0,1]$ such that

$$|S_3(p)| \leqslant |o_p| \cdot \sum_{k=1}^{N_p} \frac{W_k}{e^{\Gamma_k}} \Gamma_k e^{\rho_{p,k}\Gamma_k |o_p|} \leqslant |o_p| \cdot \sum_{k=1}^{N_p} \frac{W_k}{e^{\Gamma_k}} \Gamma_k e^{\Gamma_k |o_p|}.$$

Note that by Lemma 6.5 in the Appendix, on a set of probability one we have

$$\limsup_{p\to\infty} \sqrt{\frac{o_p^2 \ln 1/p}{\ln\ln\ln 1/p}} \leqslant 2^{3/2}.$$

On this set, for each $\delta > 0$ and for sufficiently small p we have

$$\sqrt{\frac{\ln 1/p}{\ln\ln\ln 1/p}} \cdot |S_3(p)| \leqslant \sum_{k=1}^{\infty} \frac{W_k \Gamma_k}{e^{(1-\delta)\Gamma_k}} \Big(2^{2/3} + \delta\Big), \tag{3.6}$$

which concludes the proof of (3.5).

Finally, we turn to the asymptotics of $S_4(t;p)$. We shall show that for each $\delta > 0$

$$S_4(p) \overset{\text{def}}{=} \sup_{t\in[0,1]} S_4(t;p) = o(p \ln^{1+\delta} 1/p). \tag{3.7}$$

To see this, first note that, by the first order Taylor expansion applied to function λ there exists a random variable $\rho_{p,k} \in [0,1]$ such that

$$\begin{aligned}\frac{|S_4(p)|}{p\ln^{1+\delta} 1/p} &\leqslant \frac{1}{\ln^{1+\delta} 1/p} \sum_{k=1}^{N_p} W_k \frac{\Big|\lambda\Big(p\exp\big(\Gamma_k(1+o_p)\big)\Big) - \lambda(0)\Big|}{p\exp\big(\Gamma_k(1+o_p)\big)} \\ &= \frac{1}{\ln^{1+\delta} 1/p} \sum_{k=1}^{N_p} W_k \Big|\lambda'\Big(\rho_{p,k}\, p\, \exp\big(\Gamma_k(1+o_p)\big)\Big)\Big|.\end{aligned}$$

Further, the properties of λ listed in Lemma 6.1 in the Appendix produce

(3.8)

$$\begin{aligned}\frac{|S_4(p)|}{p\ln^{1+\delta} 1/p} &\leqslant \frac{\sum_{k=1}^{N_p} W_k \Big(\frac{1}{2} + \Big|\lambda'\Big(p\, \exp\big(\Gamma_{N_p}(1+o_p)\big)\Big)\Big|\Big)}{\ln^{1+\delta} 1/p} \\ &= \frac{U_p}{\ln^{\delta} 1/p} + \frac{\Big|\lambda'\Big(\exp\big(-[(\ln 1/p)/\Gamma_{N_p+1}]T_{N_p+1}\big)\Big)\Big|}{\ln^{\delta} 1/p} \, \frac{\sum_{k=1}^{N_p} W_k}{\ln 1/p},\end{aligned}$$

where $\Gamma_n = T_1 + \ldots + T_n$, the $\{T_i\}$ are independent standard exponential random variables and

$$U_p = \frac{1}{2} \frac{\sum_{k=1}^{N_p} W_k}{N_p} \frac{N_p}{\ln 1/p}.$$

Clearly, U_p converges with probability one to $1/2$. Moreover, by the zero-one law for each a_n diverging to infinity and $\rho > 0$, we have with probability one

$$\lim_{n\to\infty} \frac{\lambda'(e^{-\rho T_n})}{a_n} = 0.$$

Consequently, for a fixed $\epsilon \in (0,1)$, there is a set of probability one on which both

$$\lim_{p\to 0} \frac{\lambda'\big(\exp\big(-(1-\epsilon)T_{N_p+1}\big)\big)}{a_{N_p+1}} = \lim_{p\to 0} \frac{\lambda'\big(\exp\big(-(1+\epsilon)T_{N_p+1}\big)\big)}{a_{N_p+1}} = 0$$

and

$$\lim_{p\to 0} \frac{\ln 1/p}{\Gamma_{N_p+1}} = 1.$$

On this set, for sufficiently small p we have

$$1+\epsilon \geqslant \frac{\ln 1/p}{\Gamma_{N_p+1}} \geqslant 1-\epsilon,$$

and by the properties of λ we get

$$\begin{aligned}&\lambda'\big(\exp\big(-(1+\epsilon)T_{N_p+1}\big)\big)\\ &\qquad\leqslant \lambda'\left(\exp\left(-\frac{\ln 1/p}{\Gamma_{N_p+1}}T_{N_p+1}\right)\right) \leqslant \lambda'\big(\exp\big(-(1-\epsilon)T_{N_p+1}\big)\big).\end{aligned}$$

Using the above (take $a_n = n^{\delta}$) and applying (3.8) we obtain (3.7).

Comparing the different asymptotics that are obtained in (3.3), (3.4), (3.5), and (3.7), we conclude that the dominating one is (3.5). This completes the proof. ∎

REMARK 3.1. In the representation (3.1), p enters stochastic components only through $\ln 1/p$. For example, the number of jumps in this representation is equal to $N(\ln 1/p)$, which asymptotically behaves as $\ln 1/p$. Thus, from the point of view of computational intensity, the above rate of convergence of $NB(t;p)$ to $G(t)$ should be viewed as the rate with respect to $\ln(1/p)$ rather than p.

REMARK 3.2. From the proof we can also obtain an asymptotic upper bound for the convergence. Namely, it follows from (3.6) that for each $\delta > 0$ we have

$$\limsup_{p\to\infty} \sqrt{\frac{\ln 1/p}{\ln\ln\ln 1/p}} \sup_{t\in[0,1]} |G(t) - pNB(t;p)| \leqslant 2^{3/2} \sum_{k=1}^{\infty} \frac{W_k\Gamma_k}{e^{(1-\delta)\Gamma_k}}.$$

3.2. Simulation. The presented results lead to two general methods of simulating sample paths of an NBP: one based on the representation of the NBP as a compound Poisson process, and another one based on the representation (2.3) and Theorem 2.2.

Suppose we are interested in the process on the interval $[0, N]$. The first method requires a sample path of a Poisson process with intensity $\lambda = -\ln p$ over the interval $[0, N]$. This is equivalent to generating a random sample, $E_1, \ldots, E_n, E_{n+1}$, from the standard exponential distribution, such that

$$E_1 + \ldots + E_n < -N/\ln p \leqslant E_1 + \ldots + E_{n+1},$$

followed by another random sample, $X_1, \ldots, X_n$, from the logarithmic distribution, which will provide jump sizes at the points $E_1 + \ldots + E_k$, $k = 1, \ldots, n$. An algorithm to generate a logarithmic random variate is available in [36]. It is based on the representation of a logarithmic random variable X as

$$X \stackrel{d}{=} [1 + \ln V/\ln(1 - p^U)],$$

where U and V are IID standard uniform variables and $[x]$ denotes the integer part of x.

The second method, based on the representation (2.3), requires first to generate a random sample from a geometric distribution with parameter p, $G_1, \ldots, G_N$. This can be achieved using the probability integral transformation (see, e.g., [36]), $G_i \stackrel{d}{=} [\ln(U_i)/\ln(1-p)] + 1$, where the $\{U_i\}$ are IID standard uniform variables. Given the variables G_i, the values of the process at *positive integers* are given by $NB(k) = G_1 + \ldots + G_k$, $\;k = 1, 2, \ldots, N$. Next, the values between integers are obtained by generating uniform random variables and selectors as discussed in Theorem 2.2. Since given the process values at the integers, the process between integers is parameter free, this method is dependent on p only through the generation of the geometric variables. This makes it attractive when sample paths are required for various values of p, and is well suited for parallel computing algorithms. Selected sample paths of the NBP are presented in Figure 1.

4. FURTHER PROPERTIES

Here we provide a brief account of further properties and generalizations of the NBP, including parameter estimation connected with this model.

4.1. Stochastic self-similarity. The NBP plays an important role in connection with the property of *stochastic self-similarity* introduced in [72]. Let $\mathcal{T} = \{T_c(t), t \geqslant 0\}$, $c \geqslant 1$, be a family of random time changes with $\mathbb{E}T_c(t) = ct$. Then a process $X(t)$, $t \geqslant 0$, is said to be *stochastically self-similar* (SSS) with index H with respect to $\mathcal{T}$ if $X\big(T_c(\cdot)\big) \stackrel{d}{=} c^H X(\cdot)$ for each $c \geqslant 1$. Since it involves stochastic renormalization in *time*, this notion of stochastic self-similarity is different than

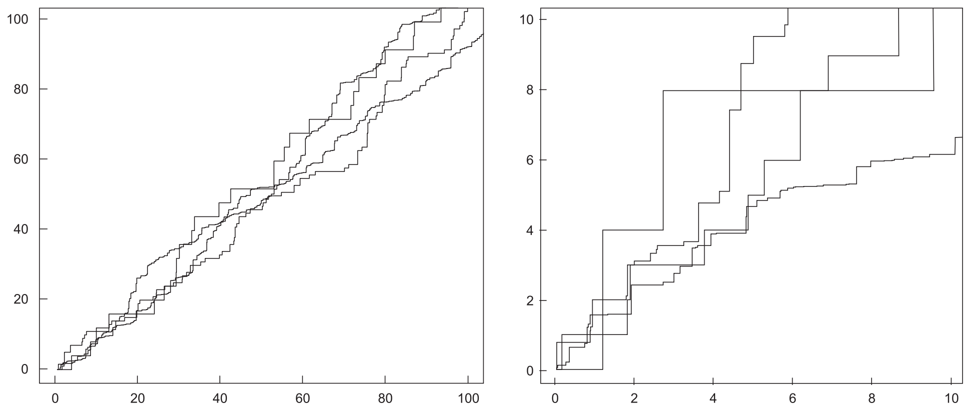

FIGURE 1. Sample paths of the NBP for various values of p ($p = 0.8, 0.5, 0.1, 0.01$). The processes have been normalized, so their mean values are the same. The larger values of p correspond to larger and scarcer jumps in the trajectories. The sample paths are over the interval $[0, 100]$ *(left)* and the interval $[0, 10]$ *(right)*

that considered in [56] and [104], which is based on stochastic renormalization in *space*. The family of negative binomial processes with drift,

$$\mathcal{T} = \{T_c(t) = t + NB_p(t),\, t \geqslant 0\}, \qquad c \geqslant 1, \tag{4.1}$$

where $NB_p(t)$ is an NBP with parameter $p = 1/c$, is an example of stochastic times changes with respect to which large classes of stochastic processes are SSS. As shown in [72], the standard gamma process is SSS with respect to (4.1), and so is any self-similar process in the classical sense subordinated to an independent gamma process. For example, a process with correlated increments, termed a *fractional Laplace motion* in [72], obtained by subordinating a fractional Brownian motion to the gamma process, is an SSS process with respect to (4.1). Further examples and more information on stochastic self-similarity can be found in [73].

4.2. Inverse process. Let $\{V_p(x),\, x \geqslant 0\}$ be the inverse of an NBP, defined as

$$V_p(x) = \inf\{t : NB_p(t) \geqslant x\}, \qquad x \geqslant 0.$$

Similarly, let $\{W_p(x),\, x \geqslant 0\}$ be the inverse of $\{NB_p(t) + t,\, t \geqslant 0\}$ (the NBP with drift). The one-dimensional distributions of these two inverse processes are quite different. Indeed, the CDF of $V_p(x)$ is of the form:

$$F_{V_p(x)}(t) = \mathbb{P}\big(V_p(x) \leqslant t\big) = P\big(NB_p(t) \geqslant x\big) = P\big(NB_p(t) \geqslant \lceil x \rceil\big), \qquad x \geqslant 0,$$

where $\lceil x \rceil$ denotes the smallest integer that is greater than or equal to x. We see that this process starts at zero with probability one, and for $x > 0$ we have

$$F_{V_p(x)}(t) = 1 - \sum_{k=0}^{\lceil x \rceil - 1} P\big(NB_p(t) = k\big) = 1 - p^t \sum_{k=0}^{\lceil x \rceil - 1} \binom{t}{k} (1-p)^k.$$

In particular, for all $x \in (n, n+1]$ $(n = 0, 1, 2, \ldots)$ we have the same distribution with mean

$$\mathbb{E}V_p(x) = \int_0^\infty p^t \sum_{k=0}^{n} \binom{t}{k} (1-p)^k dt.$$

In contrast with $V_p(x)$, the distribution of $W_p(x)$ is concentrated on the interval $[0, x]$ with the CDF

$$F_{W_p(x)}(t) = 1 - p^t \sum_{k=0}^{\lceil x-t \rceil - 1} \binom{t+k-1}{k} (1-p)^k, \quad t \leqslant x.$$

Here, the distribution at each x is different, and no longer continuous. For $x \in (n-1, n]$ $(n = 1, 2, \ldots)$, the CDF of $W_p(x)$ has n discontinuities occurring at the points $t_j = x - (n-1) + j$, $j = 0, 1, 2, \ldots, n-1$, with respective jump sizes

$$p_j = p^{x-n+1} \binom{x-1}{x-n+j} p^j (1-p)^{n-1-j}, \quad j = 0, 1, 2, \ldots, n-1.$$

Note that the mean of each inverse process is not linear in x, so that $V_p(\cdot)$ and $W(\cdot)$ *are not* Lévy processes.

4.3. Some generalizations. Our results show that on the unit interval the NBP has the same distribution as

$$X(t) = \sum_{j=1}^{G} I_{[V_j,1]}(t), \quad t \in [0, 1], \tag{4.2}$$

where G is a geometric variable given by the PMF (1.1) with $t = 1$ and the $\{V_j\}$ are identically distributed but *dependent* standard uniform variables, defined in Theorem 2.2. Various generalizations can be obtained by changing the dependence structure (or distribution) of the $\{V_j\}$, or by changing the distribution of G, or even replacing it by an increasing stochastic process $\{G(t),\, t \in [0, 1]\}$. The resulting processes will go beyond the negative binomial and may allow for dependent increments. One simple example is the process of the form (4.2) with geometric G and IID standard uniform variables V_j. This process has geometric marginal distributions and dependent, stationary increments. Let us note that by reversing the time via

$$\tilde{X}(s) = X(e^{-s}) = \sum_{j=1}^{G} I_{[V_j,1]}(e^{-s}) = \sum_{j=1}^{G} I_{[0,-\ln V_j]}(s) = \sum_{j=1}^{G} I_{[0,E_j]}(s), \tag{4.3}$$

where $E_j = -\ln V_j$ are IID standard exponential variables, we obtain a pure-death process. Here, $\tilde{X}(s)$ represents a number of individuals still alive at time $s > 0$, assuming that at time $s = 0$ there is a random number G of individuals with IID lifetimes E_j.

4.4. Maximum likelihood estimation. Suppose that $X_1, X_2, \ldots, X_n$ are the increments of an NB stochastic process taken at some lag $t > 0$. Then the $\{X_i\}$ can be viewed as a random sample from an NB distribution with parameters $p \in (0,1)$ and $t > 0$, given by (1.1). Thus, the log-likelihood function is

$$L(t,p) = n\left(\frac{1}{n}\sum_{j=1}^{n} \ln \binom{t+X_j-1}{X_j} + t\ln p + \bar{X}\ln(1-p)\right), \tag{4.4}$$

where $\bar{X}$ is the sample mean. Fixing $t > 0$ and maximizing the function L with respect to p leads to

$$p = p(t) = \frac{t}{t+\bar{X}}. \tag{4.5}$$

Incidentally, the same expression for p follows from the moment equation $\mathbb{E}\big(NB(t)\big) = t(1-p)/p$, when the sample mean is used in place of the expectation. To find the maximum likelihood estimator (MLE) of t, substitute (4.5) into (4.4) and maximize the resulting expression $L\big(t, p(t)\big)$, with respect to $t > 0$. This leads to the MLE $\hat{t}$ as the value that maximizes the function

$$g(t) = \frac{1}{n}\sum_{j=1}^{n} \ln \binom{t+X_j-1}{X_j} + t\ln t - (t+\bar{X})\ln(t+\bar{X}), \tag{4.6}$$

which has to be done numerically. In turn, the MLE of p is obtained from (4.5), $\hat{p} = p(\hat{t})$.

5. AN ILLUSTRATION

To illustrate the modeling potential of the NBP, we present an example from hydrology, taken from [71], where this model is successfully applied to borehole data from fractured granite at the Aspo Hard Rock laboratory in Sweden. Modeling of groundwater or solute transport in fractured rock [35] requires information on fractures and their transmissivities (see [57]). Fractures in rock often appear in clusters (see [50]) with spacings between the clusters following the exponential law (see [87], [88]). Since these are precisely the features of the NBP, this model appears to be well suited for such applications.

As in [71], we consider the cored borehole KLX 01 data, taken between 106 and 691 m depth, discussed in [57]. There are two data sets, consisting of 3 m and 30 m interval measurements, with sample sizes of 195 and 30, respectively. The variable of interest is discrete and measures the number of fractures in successive 3 m (and 30 m) intervals along the borehole. The summary statistics of the 3 m data yield the sample mean and variance of 8.5 and 36.3, respectively. The corresponding figures of the 30 m data are 84.2 and 666.9. This over-dispersion suggests the data might follow the negative binomial distribution.

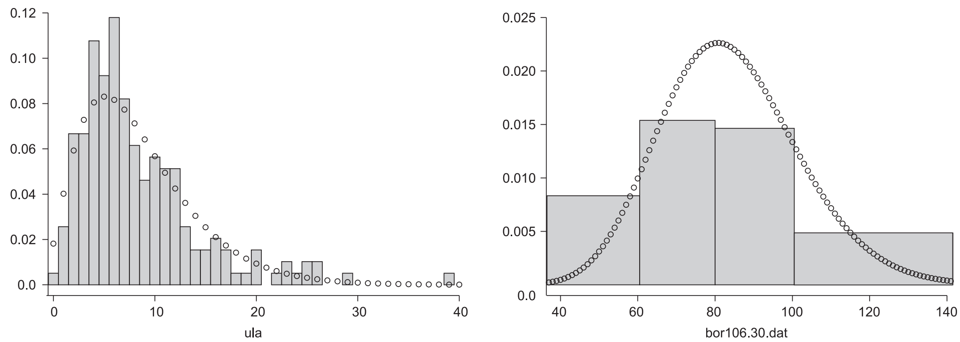

FIGURE 2. Frequencies of 3 m (the *left* panel) and 30 m (the *right* panel) intervals with different numbers of fractures (represented by bars) in KLX 01 106–691 borehole, along with estimated NB probabilities (represented by dots)

When fitting the NB distribution to the fracture count data, $X_1, X_2, \ldots, X_n$, along the borehole KLX 01 (3 m intervals with $n = 195$ data points), one can think of the data as the increments of an NB stochastic process, where $t = 0$ corresponds to the initial depth of 106 m. The method of maximum likelihood discussed in Section 4 produces $\hat{t} = 3.1$, which rounded to 3 is in almost perfect agreement with the interval size of 3 m. This value of t coupled with (4.5) leads to $\hat{p} = 0.263$.

When fitting the NB distribution to the 30 m intervals along the same borehole, one can now assume that $p = 0.263$ (since under this model all increments must have the same value of p), which in view of (4.5) leads to the MLE t equal to $(84.150)(0.263)/(1 - 0.263) = 30.0291$. Rounded to 30, this value is precisely 10 times the value of t corresponding to the 3 m intervals. Figure 2 shows the empirical distributions of the 3 m and 30 m data along with estimated binomial probabilities. This model, which postulates that the number of fractures in the interval $(0, t)$ (where t is the depth and zero corresponds to the initial level of 106 m) has an NB distribution with parameters t and $p = 0.263$, is in a very good agreement with the data.

6. APPENDIX

We collect here some basic results that were used in the proof of Theorem 3.1. Most of these are rather standard facts, and they are included here only for completeness of the presentation.

LEMMA 6.1. *The function*

$$\lambda(x) = x\frac{\ln(1-x) - 1}{\ln(1-x)}$$

is well defined for $x < 1$ *with* $\lambda(0) = 1$, *and its first two derivatives are continuous* *and given by*

$$\lambda'(x) = 1 + \frac{-x-(1-x)\ln(1-x)}{(1-x)\ln^2(1-x)}, \quad \lambda'(0) = \frac{1}{2},$$
$$\lambda''(x) = \frac{(x-2)\ln(1-x)-2x}{(1-x)^2\ln^3(1-x)}, \quad \lambda''(0) = -\frac{1}{6}.$$

Moreover, *the first derivative of* λ *is a decreasing function in* $x \in (0,1)$.

Proof. The result is a consequence of standard evaluations of derivatives and proper limits. ∎

In the course of establishing convergence results it is convenient to use some properties of the stationary time series Y_n that is defined by

$$Y_n = \sum_{k=-\infty}^{n} V_k U_k \dots U_n, \quad n \in \mathbb{Z}, \tag{6.1}$$

where $\mathbf{Z}_k = (V_k, U_k)$ are independent, identically distributed bivariate random variables such that $\mathbb{E}V_k = v$, $U_k \in (0,1)$, and $\mathbb{E}U_k = u$. We further assume that the first and the second moments of $W_k = -\ln U_k$ exist, and let

$$\mathbb{C}ov(\mathbf{Z}_k) = \begin{bmatrix} \sigma_V^2 & \sigma_V\sigma_U\rho \\ \sigma_V\sigma_U\rho & \sigma_U^2 \end{bmatrix}$$

denote the covariance matrix of $\mathbf{Z}_k$. We also assume the existence of the two covariances $r_1 = \mathbb{C}ov(U_n^2, V_n)$ and $r_2 = \mathbb{C}ov(U_n^2, V_n^2)$.

In the following result we list most important properties of this series and then we specify (V_k, U_k) that are considered in this work.

PROPOSITION 6.1. *The discrete time series* (Y_n) *defined by* (6.1) *is strictly stationary and satisfies the following relations*:

(i) (Y_n) *is first order autoregressive process with random coefficients*, *where*

$$Y_{n+1} = \rho_n Y_n + \epsilon_n,$$
$$Y_{n+1} = U_1 \dots U_{n+1}\left(Y_0 + \sum_{k=1}^{n+1} \frac{V_k}{U_1 \dots U_{k-1}}\right).$$

Here, $\rho_n = U_{n+1}$ *and* $\epsilon_n = \rho_n V_{n+1}$ *are independent of* Y_n. *The mean and the variance of* ϵ_n, *as well as the covariance of* ρ_n *and* ϵ_n, *are*

$$\mathbb{E}\,\epsilon_n = \sigma_V\sigma_U\rho + u\cdot v,$$
$$\mathbb{V}ar\,\epsilon_n = r_2 + (1-\rho)\big(\sigma_U^2\sigma_V^2(1+\rho) + \sigma_U^2 v^2 + \sigma_V^2 u^2\big) + \rho(\sigma_V u + \sigma_U v)^2,$$
$$\mathbb{C}ov(\rho_n, \epsilon_n) = r_1 + \sigma_U(\sigma_U v - \rho\sigma_V u).$$

(ii) *The first moment and the covariance structure of* (Y_n) *are as follows*:

$$\mathbb{E}Y_n = \frac{\sigma_U \sigma_V \rho + uv}{1-u},$$
$$\mathbb{V}ar\ Y_n = \frac{\sigma_U^2 \mathbb{E}^2 \epsilon_n + 2\mathbb{E}\epsilon_n \cdot \mathbb{C}ov(\rho_n, \epsilon_n) + \mathbb{V}ar\ \epsilon_n}{1-u^2-\sigma_U^2},$$
$$\mathbb{C}orr(Y_n, Y_0) = u^n.$$

(iii) *For each sequence of measurable functions* f_n, *with probability one we have*

$$\limsup_{n\to\infty} f_n(Y_{-n}) = \text{const.}$$

In particular, *if* $\lim_{n\to\infty} f_n(y) = 0$, *then the constant is equal to zero.*

(iv) *For each monotone*, *positive sequence* a_n *converging to infinity*, *we have with probability one*

$$\limsup_{n\to\infty} \frac{Y_n}{a_n} = 0.$$

P r o o f. Let us note the obvious equality in distribution

$$(\mathbf{Z}_{k+n})_{k\in\mathbb{Z}} \stackrel{d}{=} (\mathbf{Z}_k)_{k\in\mathbb{Z}}.$$

For a sequence $\mathbb{W} = (Y_k, X_k)_{k\in\mathbb{Z}}$ define

$$g(\mathbb{W}) = \sum_{k=-\infty}^{0} Y_k\, X_k \ldots X_0$$

under the assumption that the series is convergent.

First, let us note that $g\big((\mathbf{Z}_k)_{k\in\mathbb{Z}}\big)$ is well defined with probability one. Indeed, the series of independent random variables $\sum_{k=-\infty}^{0} V_k e^{-k/2}$ is almost surely convergent. Moreover, for ω in the set Ω_0 on which both this convergence holds along with the strong law of large numbers for $W_k = -\ln U_k$, for sufficiently large n we have

$$\sum_{k=n}^{\infty} |V_k(\omega)| \exp\big(-W_k(\omega) - \ldots - W_0(\omega)\big) \leqslant \sum_{k=n}^{\infty} |V_k(\omega)| e^{-k/2}.$$

Since the right-hand side is finite, we have the convergence. The strict stationarity of the series Y_n follows easily when we note that $Y_n = g\big((\mathbf{Z}_{k+n})_{k\in\mathbb{Z}}\big)$.

Part (i) is obtained after standard calculations based on the following relations:

$$Y_{n+1} = U_1 \ldots U_{n+1} \sum_{k=-\infty}^{0} V_k\, U_k \ldots U_0 + \sum_{k=1}^{n+1} V_k\, U_k \ldots U_{n+1}$$
$$= U_1 \ldots U_{n+1}\, Y_0 + U_1 \ldots U_{n+1} \sum_{k=1}^{n+1} \frac{V_k}{U_1 \ldots U_{k-1}}.$$

The computation of the moments and covariances is obvious, and thus is omitted.

Part (ii) follows easily from (i). In particular, for the covariance structure we have

$$\begin{aligned}\mathbb{C}ov(Y_{k+1}, Y_0) &= \mathbb{C}ov(\rho_k Y_k + \epsilon_k, Y_0) = \mathbb{C}ov(\rho_k Y_k, Y_0)\\ &= \mathbb{E}\big(\rho_k \mathbb{C}ov(Y_k, Y_0)\big) = u \cdot \mathbb{C}ov(Y_k, Y_0).\end{aligned}$$

Parts (iii) and (iv) can be obtained from Kolmogorov's zero-one law as follows. First, Y_n and thus $f_n(Y_n)$ are measurable with respect to $\mathcal{F}_n$, where $\mathcal{F}_n$ is the natural filtration σ-field of measurable sets generated by V_k, W_k, $k \leqslant n$. Thus, $X_{-n} = \sup_{k \leqslant n} f_k(Y_k)$ is measurable with respect to $\mathcal{F}_n$ as well. We have

$$\limsup_{n\to-\infty} f_n(Y_n) = \lim_{n\to\infty} X_n$$

and the limit is measurable with respect to $\mathcal{F}_{-\infty} = \bigcap_{n\in -\mathbb{N}} \mathcal{F}_n$. The tail σ-field $\mathcal{F}_{-\infty}$ is made of the zero-one sets, which concludes the first part of (iii). The second part follows from stationarity as we have

$$\mathbb{P}\big(f_n(Y_n) > \epsilon\big) = \mathbb{P}\big(f_n(Y_1) > \epsilon\big)$$

and $f_n(Y_1)$ converges to zero with probability one (and thus in probability), so $f_n(Y_n)$ converges in probability to zero.

Notice that for the σ-field $\mathcal{G}_n$ of measurable sets generated by W_k, $k \geqslant n$, the limit

$$\begin{aligned}\limsup_{n\to\infty} Y_n/a_n &= \lim_{n\to\infty} \exp\big(-(W_1 + \ldots + W_n)\big) Y_0/a_n\\ &\quad + \limsup_{n\to\infty} \sum_{k=1}^{n} V_k \exp\big(-(W_k + \ldots + W_n)\big)/a_n\\ &= \limsup_{n\to\infty} \sum_{k=1}^{n} V_k \exp\big(-(W_k + \ldots + W_n)\big)/a_n\end{aligned}$$

is $\mathcal{G}_0$-measurable. Since

$$\limsup_{n\to\infty} Y_n/a_n = \limsup_{n\to\infty} Y_{n+k}/a_{n+k}$$

is also $\mathcal{G}_k$-measurable for each $k \in \mathbb{N}$, the limit belongs to the tail σ-field $\mathcal{G}_\infty = \bigcap_{n\to\infty} \mathcal{G}_n$. The rest of the proof is the same as for (iii). ∎

LEMMA 6.2. *Let* Γ_k *be the arrivals of a standard Poisson process and* $a_n > 0$ *be such that for some* $\delta > 0$ (*and thus for all* $\delta' \in (0, \delta)$) *we have*

$$\limsup_{n\to\infty} \sqrt{n}\big((1+\delta)\sqrt{2\ln\ln n} - a_n\big) < \infty.$$

Then, *with probability one*

$$\exp(-\Gamma_n) = o\big(\exp(-n + \sqrt{n}a_n)\big).$$

P r o o f. Let Ω_0 be a set of probability one on which the law of iterated logarithm for the $\{\Gamma_n\}$ holds. For $\omega \in \Omega_0$ and $\delta' < \delta$ there exists an n_0 such that for $n > n_0$ we have

$$\Gamma_n(\omega) > n - (1+\delta')\sqrt{2n \ln\ln n}.$$

By assumptions, n_0 could be chosen so that

$$\sqrt{n}\big((1+\delta)\sqrt{2\ln\ln n} - a_n\big) < M.$$

Consequently, for $n > n_0$

$$\frac{\exp\big(-\Gamma_n(\omega)\big)}{\exp(-n+\sqrt{n}a_n)} \leqslant \exp\Big(\sqrt{n}\big((1+\delta')\sqrt{2\ln\ln n} - a_n\big)\Big) \leqslant e^M,$$

which proves that $\exp(-\Gamma_n) = O\big(\exp(-n+\sqrt{n}a_n)\big)$. The assumption of the lemma holds also with a_n replaced by $\tilde{a}_n = (1-\delta_0)a_n$ (for example take $\delta_0 = \delta/\big(2(1+\delta)\big)$ and replace δ by $\delta/2$). Thus we also obtain

$$\exp(-\Gamma_n) = O\Big(\exp\big(-n+\sqrt{n}(1-\delta_0)a_n\big)\Big),$$

and since $\sqrt{n}a_n \to \infty$, we eventually have the assertion. ∎

LEMMA 6.3. *With the notation and assumptions of Lemma* 6.2, *we have with probability one*

$$\sum_{k=n+1}^{\infty} W_k \exp(-\Gamma_k) = o\big(\exp(-n+\sqrt{n}a_n)\big).$$

P r o o f. We obviously have

$$\sum_{k=n+1}^{\infty} \exp(-\Gamma_{n+k})W_{n+k} = \exp(-\Gamma_n)\sum_{k=1}^{\infty}\exp\big(-(e_{n+1}+\ldots+e_{n+k})\big)W_{n+k}.$$

Let us define the stationary time series Y_n by

$$Y_{-(n+1)} = \sum_{k=1}^{\infty}\exp\big(-(e_{n+1}+\ldots+e_{n+k})\big)W_{n+k}.$$

This corresponds to the definition (6.1) with $U_k = -\ln e_{-k}$ and $V_k = W_{-k}$. Using the same argument as in the last part of the proof of Lemma 6.2, we obtain

$$\frac{\sum_{k=n+1}^{\infty}\exp(-\Gamma_n)}{\exp(-n+\sqrt{n}a_n)} \leqslant \frac{\exp(-\Gamma_n)}{\exp\big(-n+(1-\delta_0)\sqrt{n}a_n\big)}\left(\frac{1+Y_{-n-1}}{\exp(\delta_0\sqrt{n}a_n)}\right).$$

The result now follows from Lemma 6.2 and Proposition 6.1 (iii). ∎

LEMMA 6.4. *Let $N(t)$ be a Poisson process. Then, with probability one we have*

$$\limsup_{t\to\infty} \frac{N(t)-t}{\sqrt{2t\ln\ln t}} = 1.$$

P r o o f. Let $\{\Gamma_k\}$ be the arrival times corresponding to $N(t)$. Since

$$\frac{\Gamma_{N(t)}}{N(t)} - 1 \leqslant \frac{t}{N(t)} - 1 \leqslant \frac{\Gamma_{N(t)+1}}{N(t)} - 1,$$

by the law of iterated logarithm we have

$$\limsup_{t\to\infty} \left| \frac{t-N(t)}{\sqrt{2N(t)\ln\ln N(t)}} \right| = 1$$

with probability one. We also note that for a certain random variable $\lambda_t \in [0,1]$ we have

$$\begin{aligned}\left|\frac{\ln\ln N(t)}{\ln\ln t} - 1\right| &= \left|\frac{\ln\big(\ln t + \ln\big(N(t)/t\big)\big) - \ln\ln t}{\ln\ln t}\right| \\ &= \frac{1}{\ln\ln t}\left|\frac{\ln\big(N(t)/t\big)}{\ln t + \lambda_t \ln\big(N(t)/t\big)}\right|,\end{aligned}$$

which demonstrates that $\sqrt{\ln\ln N(t)/\ln\ln t}$ converges to one. This completes the proof, since

$$\left|\frac{N(t)-t}{\sqrt{2t\ln\ln t}}\right| = \left|\frac{t-N(t)}{\sqrt{2N(t)\ln\ln N(t)}}\,\frac{\sqrt{2N(t)\ln\ln N(t)}}{\sqrt{2t\ln\ln t}}\right|. \quad \blacksquare$$

LEMMA 6.5. *Let $\{\Gamma_k\}$ be the arrival times of a Poisson process independent of $N(t)$. Then, with probability one we have*

$$\limsup_{t\to\infty} \frac{|\Gamma_{N(t)} - t|}{\sqrt{2t\ln\ln t}} \leqslant 2.$$

P r o o f. We have

$$\frac{|\Gamma_{N(t)} - t|}{\sqrt{2t\ln\ln t}} \leqslant \frac{|\Gamma_{N(t)} - N(t)|}{\sqrt{2N(t)\ln\ln N(t)}} \cdot \frac{\sqrt{2N(t)\ln\ln N(t)}}{\sqrt{2t\ln\ln t}} + \frac{|N(t)-t|}{\sqrt{2t\ln\ln t}}.$$

By the same argument as that used in the proof of Lemma 6.4, we have with probability one

$$\lim_{t\to\infty} \frac{\sqrt{2N(t)\ln\ln N(t)}}{\sqrt{2t\ln\ln t}} = 1.$$

Thus, the result follows from Lemma 6.4 and the law of the iterated logarithm applied to Γ_n. $\blacksquare$

REFERENCES

[1] O. O. Aalen, *Modelling heterogeneity in survival analysis by the compound Poisson distribution*, Ann. Appl. Probab. 2 (1992), pp. 951–972.

[2] F. J. Anscombe, *Sampling theory of the negative binomial and logarithmic series distributions*, Biometrika 37 (3/4) (1950), pp. 358–382.

[3] A. G. Arbous and J. E. Kerrich, *Accident statistics and the concept of accident proneness*, Biometrics 7 (1951), pp. 340–432.

[4] N. Arley, *On the Theory of Stochastic Processes and Their Application to the Theory of Cosmic Radiation*, Copenhagen 1943.

[5] S. K. Bar-Lev and P. Enis, *Reproducibility and natural exponential families with power variance functions*, Ann. Statist. 14 (1986), pp. 1507–1522.

[6] O. Barndorff-Nielsen and G. F. Yeo, *Negative binomial process*, J. Appl. Probab. 6 (1969), pp. 633–647.

[7] J. J. Bartko, *The negative binomial distribution*: *A review of properties and applications*, The Virginia Journal of Science 12 (1961), pp. 18–37.

[8] J. Betancort-Rijo, *Generalized negative binomial distributions*, J. Statist. Phys. 98 (2000), pp. 917–933.

[9] C. I. Bliss and R. A. Fisher, *Fitting the negative binomial distribution to biological data and note on the efficient fitting of the negative binomial*, Biometrics 9 (1953), pp. 176–200.

[10] L. Bondesson, *On simulation from infinitely divisible distributions*, Adv. in Appl. Probab. 14 (1982), pp. 855–869.

[11] M. T. Boswell and G. P. Patil, *Chance mechanisms generating the negative binomial distributions*, in: *Random Counts in Scientific Work*, 1: *Random Counts in Models and Structures*, G. P. Patil (Ed.), Pennsylvania State University Press, University Park, 1970, pp. 3–22.

[12] A. Brix, *Generalized gamma measures and shot-noise Cox processes*, Adv. in Appl. Probab. 31 (4) (1999), pp. 929–953.

[13] A. Brix and W. S. Kendall, *Simulation of cluster point processes without edge effects*, Adv. in Appl. Probab. 34 (2002), pp. 267–280.

[14] F. T. Bruss and T. S. Ferguson, *High-risk and competitive investment models*, Ann. Appl. Probab. 12 (4) (2002), pp. 1202–1226.

[15] F. T. Bruss and L. C. G. Rogers, *Pascal processes and their characterization*, Stochastic Process. Appl. 37 (2) (1991), pp. 331–338.

[16] R. T. Burnett and M. T. Wasan, *The negative binomial point process and its inference*, in: *Multivariate Statistical Analysis*, R. P. Gupta (Ed.), North-Holland Publishing Co., Amsterdam–New York 1980, pp. 31–45.

[17] Q. L. Burrell, *A note on ageing in a library circulation model*, Journal of Documentation 41 (1985), pp. 100–115.

[18] Q. L. Burrell, *Stochastic modelling of the first-citation distribution*, Scientometrics 52 (2001), pp. 3–12.

[19] Q. L. Burrell, *On the nth-citation distribution and obsolescence*, Scientometrics 53 (2002), pp. 309–323.

[20] Q. L. Burrell, *Predicting future citation behavior*, Journal of the American Society for Information Science and Technology 54 (5) (2003), pp. 372–378.

[21] Q. L. Burrell and V. R. Cane, *The analysis of library data*, J. Roy. Statist. Soc. Ser. A 145 (1982), pp. 439–471.

[22] J. Burridge, *Empirical Bayes analysis of survival time data*, J. Roy. Statist. Soc. Ser. B 43 (1) (1981), pp. 65–75.

[23] V. R. Cane, *The concept of accident proneness*, Bull. Inst. Math. Bulgaria 15 (1972), pp. 183–189.

[24] V. R. Cane, *A class of non-identifiable stochastic models*, J. Appl. Probab. 14 (1976), pp. 475–482.

[25] P. Carruthers and D.-V. Minh, *A connection between galaxy probabilities in Zwicky clusters counting distributions in particle physics and quantum optics*, Phys. Lett. 131B (1, 2, 3) (1983), pp. 116–120.
[26] P. Carruthers and C. C. Shih, *Correlations and fluctuations in hadronic multiplicity distributions: the meaning of KNO scaling*, Phys. Lett. 127B (3–4) (1983), pp. 242–250.
[27] C. Chatfield, *A marketing application of a characterization theorem*, in: *Statistical Distributions in Scientific Work,* 2: *Model Building and Model Selection*, G. P. Patil, S. Kotz and J. K. Ord (Eds.), Reidel, Dordrecht 1975, pp. 175–185.
[28] C. Chatfield, A. S. C. Ehrenberg and G. J. Goodhardt, *Progress on a simplified model of stationary purchasing behaviour (with discussions)*, J. Roy. Statist. Soc. Ser. A 129 (1966), pp. 317–367.
[29] S. Chaturvedi, V. Gupta and S. K. Soni, *A unified view of multiplicity distributions based on a pure birth process*, Modern Phys. Lett. A 9 (36) (1994), pp. 3359–3366.
[30] S. J. Clark and J. N. Perry, *Estimation of the negative binomial parameter k by maximum quasi-likelihood*, Biometrics 45 (1989), pp. 309–316.
[31] A. D. Cliff and J. K. Ord, *Spatial Processes: Models and Applications*, Pion Ltd., London 1981.
[32] P. Clifford and G. Wei, *The equivalence of the Cox process with squared radial Ornstein–Uhlenbeck intensity and the death process in a sample population model*, Ann. Appl. Probab. 3 (3) (1993), pp. 863–873.
[33] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman and Hall / CRC, Boca Raton 2004.
[34] S. Dahiya and M. Kaur, *Multifractal multiplicity distribution in e^+e^- and $\bar{p}p$ interactions from nonlinear Markov process*, J. Phys. G: Nucl. Part. Phys. 24 (1998), pp. 2027–2035.
[35] W. Dershowitz, G. Lee, J. Geier, T. Foxford, P. La Pointe and A. Thomas, *User documentation, FracMan, Interactive discrete data analysis, geometric modeling, and exploration simulation*, Version 2.6 Golder Associates Inc., Seattle 1998.
[36] L. Devroye, *Non-Uniform Random Variate Generation*, Springer, New York 1986.
[37] P. J. Diggle and R. K. Milne, *Negative binomial counts and point processes*, Scand. J. Statist. 10 (1983), pp. 257–267.
[38] J. B. Douglas, *Analysis with Standard Contagious Distributions*, Internat. Coop. Publishing House, Burtonsville, MD, 1980.
[39] A. S. C. Ehrenberg, *The pattern of consumer purchases*, Applied Statistics 8 (1) (1959), pp. 26–41.
[40] O. Eneroth, *On the quantity of seed in patch sowing, and the connection between the number of seedlings and the stand density in natural regeneration* (in Swedish), Norrlands Skogsvardsförbunds Tidskrift 1945, pp. 161–222.
[41] D. A. Evans, *Experimental evidence concerning contagious distributions in ecology*, Biometrika 40 (1953), pp. 186–211.
[42] W. Feller, *Die Grundlagen der Volterraschen Theorie*, Acta Biotheoretica 5 (1939), pp. 11–40.
[43] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, Wiley, New York 1960.
[44] W. Feller, *Introduction to the Theory of Probability and Its Applications*, Vol. 2, 2nd edition, Wiley, New York 1971.
[45] T. S. Ferguson and M. J. Klass, *A representation of independent increment processes without Gaussian components*, Ann. Math. Statist. 43 (1972), pp. 1634–1643.
[46] C. Ferreri, *On the extended Pólya process and some its interpretations*, Metron 41 (1–2) (1983), pp. 11–27.
[47] C. Ferreri, *On a hyperbinomial process*, Comm. Statist. Theory Methods 25 (1996), pp. 83–103.
[48] R. A. Fisher, *The negative binomial distribution*, Ann. Eugen 11 (1941–2), pp. 182–187.

[49] W. H. Furry, *On fluctuation phenomena in the passage of high energy electrons through lead*, Phys. Rev. 52 (1937), pp. 569–581.

[50] P. A. Gillespie, C. B. Howard, J. J. Wash and J. Watterson, *Measurement and characterization of spatial distributions of fractures*, Tectonophysics 226 (1993), pp. 113–141.

[51] W. Glänzel and A. Schubert, *Predictive aspects of a stochastic model for citation processes*, Information Processing and Management 31 (1995), pp. 69–80.

[52] G. J. Goodhardt, A. S. C. Ehrenberg and C. Chatfield, *The Dirichlet: A comprehensive model of buying behaviour (with discussions)*, J. Roy. Statist. Soc. Ser. A 147 (1984), pp. 621–655.

[53] J. Grandell, *Doubly Stochastic Poisson Processes*, Springer, New York 1976.

[54] M. Greenwood and G. U. Yule, *An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents*, J. Roy. Statist. Soc. Ser. A 83 (1920), pp. 255–279.

[55] G. Gregoire, *Negative binomial distributions for point processes*, Stochastic Process. Appl. 16 (1983), pp. 179–188.

[56] V. K. Gupta and E. C. Waymire, *Multiscaling properties of spatial rainfall in river flow distributions*, J. Geophys. Res. 95 (1990), pp. 1999–2009.

[57] G. Gustafson and A. Fransson, *The use of the Pareto distribution for fracture transmissivity assessment*, Hydrogeology J. 14 (2005), pp. 15–20.

[58] J. B. S. Haldane, *On a method of estimating frequencies*, Biometrika 33 (1945), pp. 222–225.

[59] P. Hougaard, *Life table methods for heterogeneous populations: distributions describing the heterogeneity*, Biometrika 71 (1) (1984), pp. 75–83.

[60] P. Hougaard, *Survival models for heterogeneous populations derived from stable distributions*, Biometrika 73 (1986), pp. 387–397 (correction: 75, p. 395).

[61] P. Hougaard, *Modelling multivariate survival*, Scand. J. Statist. 14 (4) (1987), pp. 291–304.

[62] P. Hougaard, M.-L. T. Lee and G. A. Whitmore, *Analysis of overdispersed count data by mixtures of Poisson variables and Poisson processes*, Biometrics 53 (1997), pp. 1225–1238.

[63] J. O. Irwin, *Discussion on Chambers and Yule's paper*, J. Roy. Statist. Soc. Ser. B 7 (1941), pp. 101–109.

[64] K. G. Janardan and D. J. Schaeffer, *Application of discrete distributions for estimating the number of organic compounds in water*, in: *Statistical Distributions in Scientific Work* 6: *Applications in Physical, Social, and Life Sciences*, C. Taillie, G. P. Patil, and B. A. Baldessari (Eds.), Reidel, Dordrecht 1981, pp. 79–94.

[65] N. L. Johnson, S. Kotz and A. W. Kemp, *Univariate Discrete Distributions*, Wiley, New York 1993.

[66] B. Jorgensen, *Exponential dispersion models*, J. Roy. Statist. Soc. Ser. B 49 (1987), pp. 127–162.

[67] Y. Y. Kagan and D. D. Jackson, *Probabilistic forecasting of earthquakes*, Geophys. J. Int. 143 (2000), pp. 438–453.

[68] J. D. Kalbfleisch, *Bayesian analysis of survival distributions*, J. Roy. Statist. Soc. Ser. B 40 (1978), pp. 214–221.

[69] D. E. Kashcheev, *Compound Cox processes and option pricing*, J. Math. Sci. 106 (2001), pp. 2682–2690.

[70] D. G. Kendall, *Stochastic processes and population growth*, J. Roy. Statist. Soc. Ser. B 11 (2) (1949), pp. 230–282.

[71] T. J. Kozubowski, M. M. Meerschaert and G. Gustafson, *A new stochastic model for fracture transmissivity assessment*, Water Resources Res. 44 (2008), W02435, doi: 10.1029/2007WR006053.

[72] T. J. Kozubowski, M. M. Meerschaert and K. Podgórski, *Fractional Laplace motion*, Adv. in Appl. Probab. 38 (2006), pp. 451–464.

[73] T. J. Kozubowski and K. Podgórski, *Invariance properties of the negative binomial Lévy process and stochastic self-similarity*, Intern. Math. Forum 2 (29–32) (2007), pp. 1457–1468.
[74] D. G. Lampard, *A stochastic process whose successive intervals between events from a first order Markov chain. I*, J. Appl. Probab. 5 (1968), pp. 648–668.
[75] M.-L. T. Lee and G. A. Whitmore, *Stochastic processes directed by randomized time*, J. Appl. Probab. 30 (1993), pp. 302–314.
[76] R. Lüders, *Die Statistik der seltenen Ereignisse*, Biometrika 26 (1934), pp. 108–128.
[77] J. S. Martin and S. K. Katti, *Fitting of some contagious distributions to some available data by the maximum likelihood method*, Biometrics 21 (1965), pp. 34–48.
[78] B. Matérn, *Doubly stochastic Poisson processes in the plane (with discussions)*, Statistical Ecology 1 (1971), pp. 195–213.
[79] A. G. McKendrick, *Studies on the theory of continuous probabilities with special reference to its bearing on natural phenomena of a progressive nature*, Proc. London Math. Soc. 2 (13) (1914), pp. 401–416.
[80] A. G. McKendrick, *Applications of mathematics to medical problems*, Proc. Edinb. Math. Soc. 44 (1925), pp. 1–34.
[81] E. McKenzie, *Autoregressive moving-average processes with negative-binomial and geometric marginal distributions*, Adv. in Appl. Probab. 18 (1986), pp. 679–705.
[82] A. Milne, *The comparison of sheep-tick populations* (*Ixodes ricinus* L.), Ann. Appl. Biol. 30 (1943), pp. 240–250.
[83] J. Moller, *Shot noise Cox processes*, Adv. in Appl. Probab. 35 (2003), pp. 614–640.
[84] J. Neyman, *Certain chance mechanisms involving discrete distributions*, in: *Classical and Contagious Discrete Distributions*, G. P. Patil (Ed.), Pergamon Press, Oxford 1965, pp. 4–14.
[85] J. N. Perry, *Negative binomial model for mosquitoes*, Biometrics 40 (1984), p. 863.
[86] D. Pollard, *Convergence of Stochastic Processes,* Springer, New York 1984.
[87] S. D. Priest and J. A. Hudson, *Discontinuity spacings in rock*, Int. J. Mech. Min. Sci. Geomech. Abstr. 13 (1976), pp. 135–148.
[88] S. D. Priest and J. A. Hudson, *Estimation of discontinuity spacing and trace length using scanline surveys*, Int. J. Mech. Min. Sci. Geomech. Abstr. 18 (1981), pp. 183–197.
[89] M. H. Quenouille, *A relation between the logarithmic, Poisson, and negative binomial series*, Biometrics 5 (1949), pp. 162–164.
[90] A. Ramakrishnan, *Some simple stochastic processes*, J. Roy. Statist. Soc. Ser. B 13 (1) (1951), pp. 131–140.
[91] D. S. Reynolds and I. R. Savage, *Random wear models in reliability theory*, Adv. in Appl. Probab. 3 (1971), pp. 229–248.
[92] J. Rosiński, *Series representations of Lévy processes from the perspective of point processes*, in: *Lévy Processes – Theory and Applications*, O. E. Barndorff-Nielsen, T. Mikosch and S. I. Resnick (Eds.), Birkhäuser, Boston 2001, pp. 401–415.
[93] K. Saha and S. Paul, *Bias-corrected maximum likelihood estimator of the negative binomial dispersion parameter*, Biometrics 61 (2005), pp. 179–185.
[94] M. Şahinoğlu, *Compound Poisson software reliability model*, IEEE Trans. Software Eng. 18 (1992), pp. 624–630.
[95] M. Şahinoğlu, *An empirical Bayesian stopping rule in testing and verification of behavioral models*, IEEE Trans. Instr. Meas. 52 (2003), pp. 1428–1443.
[96] W. Schoutens, *Lévy Processes in Finance: Pricing Financial Derivatives*, Wiley, Chichester 2003.
[97] R. F. Serfozo, *Conditional Poisson processes*, J. Appl. Probab. 9 (2) (1972), pp. 288–302.
[98] H. S. Sichel, *The estimation of the parameters of a negative binomial distribution with special reference to psychological data*, Psychometrika 16 (1951), pp. 107–127.
[99] C. H. Sim and P. A. Lee, *Simulation of negative binomial processes*, J. Stat. Comput. Simul. 34 (1989), pp. 29–42.

[100] D. L. Snyder, *Random Point Processes*, Wiley, New York 1975.
[101] R. M. Solow, *On a family of lag distributions*, Econometrica 28 (1960), pp. 392–406.
[102] L. R. Taylor, *Assessing and interpreting the spatial distributions of insect populations*, Ann. Rev. Entomol. 29 (1984), pp. 321–357.
[103] J. Vaillant, *Negative binomial distributions of individuals and spatio-temporal Cox processes*, Scand. J. Statist. 18 (1991), pp. 235–248.
[104] D. Veneziano, *Basic properties and characterization of stochastically self-similar processes in* $\mathbb{R}^d$, Fractals 7 (1) (1999), pp. 59–78.
[105] L. T. Wilson and P. M. Room, *Clumping patterns of fruit and arthropods in cotton, with implications for binomial sampling*, Environmental Entomology 12 (1983), pp. 50–54.
[106] R. L. Wolpert and K. Ickstadt, *Poisson/gamma random field models for spatial statistics*, Biometrika 85 (1998), pp. 251–267.
[107] G. U. Yule, *On the distribution of deaths with age when the causes of death act cumulatively, and similar frequency distributions*, J. Roy. Statist. Soc. 73 (1910), pp. 26–38.
[108] G. U. Yule, *A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis*, Phil. Trans. Roy. Soc., London, B 213 (1924), pp. 21–87.

University of Nevada
Department of Mathematics and Statistics
University of Nevada
Reno, NV 89557, USA
E-mail: tkozubow@unr.edu

Lund University
Centre for Mathematical Sciences
Mathematical Statistics
Lund University
Box 118, 221 00 Lund, Sweden
E-mail: krys@maths.lth.se

Received on 16.1.2008;
revised version on 5.4.2008