A generalized Sibuya distribution

Kozubowski, Tomasz J; Podgórski, Krzysztof

2016

Working Papers in Statistics
No 2016:8

Department of Statistics
School of Economics and Management
Lund University

A generalized Sibuya distribution

TOMASZ J. KOZUBOWSKI, UNIVERSITY OF NEVADA

KRZYSZTOF PODGÓRSKI, LUND UNIVERSITY

A generalized Sibuya distribution

Tomasz J. Kozubowski

Department of Mathematics & Statistics, University of Nevada, Reno NV 89557, USA

Krzysztof Podgórski

Department of Statistics, Lund University, Sweden

Abstract The Sibuya distribution arises as the distribution of the waiting time for the first success in Bernoulli trials, where the probabilities of success are inversely proportional to the number of a trial. We study a generalization that is obtained as the distribution of the excess random variable $N - k$ given $N > k$, where N has the Sibuya distribution. We summarize basic facts regarding this distribution and provide several new results and characterizations, shedding more light on its origin and possible applications. In particular, we emphasize the role Sibuya distribution plays in the extreme value theory and point out its invariance property with respect to random thinning operation.

Keywords Discrete Pareto distribution, Distribution theory, Extreme value theory, Infinite divisibility, Mixed Poisson Process, Power law, Pure death process, Records, Yule distribution, Zipf's law

1. Introduction

Let X_i, $i = 1, 2, \ldots$, be a sequence of independent and identically distributed (IID) continuous random variables. The first X_i that exceeds all previous values is called the first *record* value. Let I_j, $j = 1, 2, \ldots$, be the associated sequence of Bernoulli random variables, indicating wether or not a particular X_j is a record. It follows from the random records theory [see, e.g., Rényi (1962, 1976)] that the variables I_j are mutually independent and

$$\mathbb{P}(I_j = 1) = \frac{1}{1+j}, \quad j \in \mathbb{N} = \{1, 2, \ldots\}. \tag{1.1}$$

Accordingly, if N denotes the waiting time for the first record to occur, then

$$\mathbb{P}(N = n) = \frac{1}{n(n+1)} = \frac{1}{1+(n-1)} - \frac{1}{1+n}, \quad n \in \mathbb{N}. \tag{1.2}$$

The probability distribution given by (1.2) is a special case with $\alpha = 1$ and $\sigma = 1$ of *discrete Pareto distribution*, which in general has the probability mass function (PMF)

$$\mathbb{P}(N = n) = \left(\frac{1}{1+\frac{n-1}{\sigma}}\right)^{\alpha} - \left(\frac{1}{1+\frac{n}{\sigma}}\right)^{\alpha}, \quad n \in \mathbb{N}, \tag{1.3}$$

and arises by discretization of continuous Pareto Type II (Lomax) distribution with tail parameter $\alpha > 0$ and scale parameter $\sigma > 0$ [see Krishna and Singh Pundir (2009), Buddana and Kozubowski (2014)]. The distribution given by (1.2) is also a special case $\alpha = 1$ of *Yule distribution* [see Yule (1925)], which in general case $\alpha > 0$ is given by the PMF

$$\mathbb{P}(N = n) = \frac{\alpha\Gamma(\alpha+1)\Gamma(n)}{\Gamma(\alpha+n+1)}, \quad n \in \mathbb{N}. \tag{1.4}$$

Both Yule and discrete Pareto distributions are *heavy tailed*, with *power law* behavior of their PMFs (and tails),

$$\mathbb{P}(N = n) = O\left(\frac{1}{n^{\alpha+1}}\right) \quad \text{as } n \to \infty. \tag{1.5}$$

Along with the Zipf's law, whose PMF has the same asymptotics [see, e.g., Zipf (1949) or Johnson et al. (1993)], these distributions provide important modeling tools whenever empirical distributions display power-law tails. Such scaling behavior has been observed across many fields, including biology, chemistry, computer science, economics, finance, geo-sciences, and social science [see, e.g., Aban et al. (2006), Clauset and Newman (2009), Gabaix (2009), Newman (2005), Sornette (2006), Stumpf and Porter (2012)].

In this paper we study another generalization of (1.2), which is directly related to its interpretation through the record process described above. Namely, we define a discrete variable N to be the waiting time for the first success in a sequence of independent Bernoulli trials $\{I_j, j \in \mathbb{N}\}$, where the probabilities of success are given by

$$\mathbb{P}(I_j = 1) = \frac{\alpha}{k+j}, \quad k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\} \quad \text{and} \quad 0 < \alpha < k+1. \tag{1.6}$$

We observe that the record times correspond to $\alpha = k = 1$. If N is the *number of trials* until the first success, then

$$\mathbb{P}(N = n) = \left(1 - \frac{\alpha}{k+1}\right) \cdots \left(1 - \frac{\alpha}{k+n-1}\right) \frac{\alpha}{k+n}, \quad n \in \mathbb{N}. \tag{1.7}$$

It can be shown that, asymptotically, the probabilities (1.7) are also power laws of the form (1.5). Moreover, in the special case $k = 0$ and $\alpha \in (0,1)$, we obtain the *Sibuya* distribution with the PMF

$$\mathbb{P}(N = n) = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}(-1)^{n+1} = \binom{\alpha}{n}(-1)^{n+1}, \quad n \in \mathbb{N}, \tag{1.8}$$

which first appeared in Sibuya (1979) and was later studied in connection with discrete stable, Linnik, and Mittag-Leffler distributions [see, e.g., Christoph and Schreiber (1998, 2000), Devroye (1993), Pakes (1995), Pillai and Jayakumar (1995), Satheesh and Nair (2002)]. Due to this connection, we name the distribution with the PMF (1.7) *generalized Sibuya*.

The main goal of the paper is to account for basic properties of the generalized Sibuya distribution (1.7), and to show how it interrelates with its special case of Sibuya distribution (1.8). Additionally, we wish to emphasize the importance of the Sibuya distribution in the distribution theory, and provide its new characterization which goes beyond the class of generalized Sibuya variables. To this end, let us first comment on the importance of the Sibuya model in the extreme value theory. It is well known that, for any $n \in \mathbb{N}$, the quantity $[F(x)]^n$, where F is a cumulative distribution function, is also a distribution function, corresponding to the random variable

$$X = \max\{X_1, \ldots, X_n\} = \bigvee_{j=1}^{n} X_j, \tag{1.9}$$

where the $\{X_j\}$ are IID with the CDF F. For a non-integer exponent $\alpha > 0$, the quantity $[F(x)]^\alpha$ is a genuine CDF as well, although in this case we no longer have the interpretation (1.9) through the maximum. Similarly, the quantity $[S(x)]^n$, where $S(x) = 1 - F(x)$ is the survival function (SF) of the $\{X_j\}$, is the survival function corresponding to

$$Y = \min\{X_1, \ldots, X_n\} = \bigwedge_{j=1}^{n} X_j, \tag{1.10}$$

although $[S(x)]^\alpha$, still being a genuine survival function, lacks such an interpretation for fractional $\alpha > 0$. It turns out that the Sibuya distribution (1.8) provides a missing link, allowing an interpretation through *stochastic* maxima and minima as presented in the following result.

Proposition 1.1. *Let F be a distribution function on $\mathbb{R}$ and S be the corresponding survival function, $S(x) = 1 - F(x)$. Further, let X and Y have SF and CDF given by $[S(x)]^\alpha$ and $[F(x)]^\alpha$, respectively, where $\alpha \in (0,1]$. Then X and Y admit the stochastic representations*

$$X \stackrel{d}{=} \bigvee_{j=1}^{N_\alpha} X_j \ \ \text{and} \ \ Y \stackrel{d}{=} \bigwedge_{j=1}^{N_\alpha} X_j, \tag{1.11}$$

where N_α has the Sibuya distribution (1.8) and is independent of the IID $\{X_j\}$ with the CDF F.

We refer to Kozubowski and Podgórski (2016) for the proof, further results on random maxima and minima with Sibuya number of terms, and generalizations to random processes. As we shall see in Section 5, one can define a pure jump random process with Sibuya marginal distributions. The laws of the jumps are related to the generalized Sibuya distribution. In particular, the size of the first jump of this process has the generalized Sibuya distribution (1.7) with $k = 1$. Such relations between the Sibuya and generalized Sibuya distributions, along with the importance of the former, provide additional motivation for studying the latter.

Let us finally provide yet another result on the Sibuya distribution, which appears to be new. It relates to the theory of birth/death Markov processes. Consider a sequence $\{X_i\}$, $i \in \mathbb{N}$, of IID random variables having continuous distribution on $\mathbb{R}_+ = (0, \infty)$. Suppose that at time $t = 0$, a population consists of a random number $N \in \mathbb{N}$ of individuals, whose future lifetimes are given by X_i, $i = 1, \ldots, N$. Then

$$N(t) = \sum_{i=1}^{N} I_{(t,\infty)}(X_i), \ \ t \geq 0, \tag{1.12}$$

is a pure death process, describing the number of individuals alive at time t (the quantity I_A is an indicator of the set A). It turns out that if N has the Sibuya distribution (1.8) with some $\alpha \in (0, 1)$, then, regardless of a choice for the distribution of X_i's, the conditional distribution of $N(t)|N(t) > 0$ is the same as that of N. In other words, the Sibuya distribution provides a *stationary* conditional distribution of $N(t)$ for each $t \in [0, \infty)$: if it is known that the population is still alive at time $t > 0$, its size is described by *the same*

Sibuya distribution. This in fact is a characterization of the Sibuya distribution, as stated in the following result, which is proven in the Appendix.

Proposition 1.2. *Let $\{X_i\}$, $i \in \mathbb{N}$, be a sequence of IID random variables having continuous distribution on $\mathbb{R}_+ = (0, \infty)$, and let N be a random variable on $\mathbb{N}$, independent of the X_i's. Then N has a Sibuya distribution (1.8) with some $\alpha \in (0,1)$ if and only if for each $t \in [0, \infty)$ we have the equality in distribution*

$$N \stackrel{d}{=} N(t) | N(t) > 0, \tag{1.13}$$

where $N(t)$ is a pure death process defined by (1.12).

The rest of the paper is a careful account of the properties of the generalized Sibuya model. We start with Section 2, where we introduce the model and derive its basic characteristics. Various stochastic representations of the model appear in Section 3. They are followed by account of divisibility properties in Section 4. In Section 5 we define a Sibuya random process on $[0,1]$ and study the structure of its sample paths. We conclude with the Appendix, containing (selected) proofs and auxiliary results.

2. Definition and basic properties

We begin with a definition of the generalized Sibuya stochastic model.

Definition 2.1. A random variable N with the PMF (1.7) is said to have a generalized Sibuya distribution with parameters $\alpha \in \mathbb{R}_+$ and $k \in \mathbb{N}_0$, denoted by $GS_1(\alpha, k)$. The two parameters are restricted by the relation $0 < \alpha < k + 1$.

The subscript in the notation indicates that the distribution is supported on the set $\mathbb{N}$ of positive integers. Another version of this distribution, which is defined as the *number of failures* before the first success, shall be denoted by $GS_0(\alpha, k)$, i.e.

$$N \sim GS_1(\alpha, k) \quad \text{if and only if} \quad N - 1 \sim GS_0(\alpha, k). \tag{2.1}$$

The properties provided in the sequel shall be stated in terms of either one of the two distributions, and can be easily re-formulated in terms of the other if needed.

2.1. **Special cases.** Note that at the boundary of the parameter space, where $0 < \alpha = k+1$, the distribution collapses to a point mass at 1. This exceptional case shall be omitted from most considerations. The Sibuya distribution (1.8) arises as a special case of $GS_1(\alpha, k)$ with $\alpha \in (0,1)$ and $k = 0$. This distribution is often described through its probability generating function (PGF), which, compared with the general case discussed in the sequel, is of a particularly simple form:

$$G_N(s) = \sum_{n=1}^{\infty} \binom{\alpha}{n} (-1)^{n+1} s^n = 1 - (1-s)^{\alpha}, \;\; 0 < s < 1. \tag{2.2}$$

In the further special case $\alpha = 1/2$, we have that $GS_0(1/2, 0)$ is a *discrete Mittag-Leffler* distribution with the PGF $G(s) = [1 + (1-s)^{\alpha}]^{-1}$ [see, e.g., Pillai and Jayakumar (1995)].

We have already noted the special case $\alpha = k = 1$ of the $GS_1(\alpha, k)$ distribution, where the PMF simplifies to (1.2) and we obtain a particular case of the discrete Pareto and the Yule distributions. For $\alpha = 1$ and general $k \in \mathbb{N}$, the GS PMF (1.7) becomes

$$\mathbb{P}(N = n) = \left(1 - \frac{1}{k+1}\right) \cdots \left(1 - \frac{1}{k+n-1}\right) \frac{1}{k+n} = \frac{k}{k+n-1} - \frac{k}{k+n}, \tag{2.3}$$

and we also obtain a case of discrete Pareto distribution (1.3) with $\alpha = 1$ and $\sigma = k$.

2.2. **Distribution and survival functions.** The CDF and the SF of a GS random variable $N \sim GS_1(\alpha, k)$ are straightforward to derive. Indeed, for any $n \in \mathbb{N}$:

$$\mathbb{P}(N > n) = \mathbb{P}(I_j = 0, j = 1, \ldots, n) = \left(1 - \frac{\alpha}{k+1}\right) \cdots \left(1 - \frac{\alpha}{k+n}\right). \tag{2.4}$$

It follows that the SF and the PMF of $N \sim GS_1(\alpha, k)$ are linked as follows:

$$\mathbb{P}(N = n) = \frac{\alpha}{n+k-\alpha} \mathbb{P}(N > n), \;\; n \in \mathbb{N}. \tag{2.5}$$

We now consider the conditional distribution of $N - m$ given $N > m$. Straightforward algebra incorporating the above results shows that

$$\mathbb{P}(N - m = n | N > m) = \left(1 - \frac{\alpha}{k+m+1}\right) \cdots \left(1 - \frac{\alpha}{k+m+n-1}\right) \frac{\alpha}{k+m+n}, \;\; n \in \mathbb{N}. \tag{2.6}$$

The above is recognized as a GS probability as well, with parameters α and $k + m$. In particular, if N has Sibuya distribution (1.8), then the corresponding excess $N - m$ conditionally on $N > m$ is generalized Sibuya $GS_1(\alpha, m)$. Thus the class of generalized Sibuya distributions is closed with respect to the operation of taking the excess, as summarized in the result below.

Proposition 2.2. *If $m \in \mathbb{N}_0$ and $N \sim GS_1(\alpha, k)$ then $N - m | N > m \sim GS_1(\alpha, k+m)$.*

2.3. **Moments and tail behavior.** As shown in Christoph and Schreiber (2000), Sibuya probabilities (1.8) admit the asymptotic representation

$$\mathbb{P}(N = n) \sim \frac{1}{\pi} \sin(\alpha\pi)\Gamma(1+\alpha)\frac{1}{n^{\alpha+1}} \text{ as } n \to \infty, \tag{2.7}$$

where $f(x) \sim g(x)$ means that $f(x)/g(x) \to 1$ as $x \to \infty$. Thus, if $N \sim G_1(\alpha, 0)$ (ordinary Sibuya), where necessarily $\alpha \in (0, 1)$, then we have (1.5). As shown below, the latter asymptotic relation holds for generalized Sibuya distribution as well.

Proposition 2.3. *If $N \sim GS_1(\alpha, k)$ then*

$$\mathbb{P}(N = n) \sim \frac{\Gamma(k+1)}{\Gamma(k+1-\alpha)} \frac{\alpha}{n^{\alpha+1}} \text{ as } n \to \infty. \tag{2.8}$$

Remark 2.4. Note that if we set $k = 0$ in (2.8) and use two well-known properties of the gamma function,

$$\Gamma(\alpha)\Gamma(1-\alpha) = \frac{\pi}{\sin(\alpha\pi)}, \quad \Gamma(1+\alpha) = \alpha\Gamma(\alpha),$$

then we recover (2.7).

In view of the link (2.5) between generalized Sibuya survival function and its probabilities, the above result immediately provides the asymptotics of the tail, stated below.

Corollary 2.5. *If $N \sim GS_1(\alpha, k)$ then*

$$\mathbb{P}(N > n) \sim \frac{\Gamma(k+1)}{\Gamma(k+1-\alpha)} \frac{1}{n^{\alpha}} \text{ as } n \to \infty. \tag{2.9}$$

Because of the power-law asymptotics (2.9) of its tail, the moments of order α and above of generalized Sibuya distribution do not exist.

Corollary 2.6. *Let $\gamma \in \mathbb{R}_+$. If $N \sim GS_1(\alpha, k)$ then $\mathbb{E}N^\gamma < \infty$ if and only if $\gamma \in (0, \alpha)$.*

In particular, the expectation of $N \sim GS_1(\alpha, k)$ exists whenever $\alpha > 1$ (so that necessarily $k \geq 1$), while the variance exists if and only if $\alpha > 2$ (so that $k \geq 2$). Perhaps the most convenient way to obtain these, along with other moments, is through the mixture representation of generalized Sibuya distribution provided by Proposition 3.5. Using tower property for conditional expectations, that result allows us to write

$$\mathbb{E}N^j = \mathbb{E}f_j(Y), \quad j \in \mathbb{N}_0, \tag{2.10}$$

where $N \sim GS_0(\alpha, k)$, Y has a beta distribution with parameters α and $\beta = 1 - \alpha + k$, and $f_j(p)$ denotes the jth (raw) moment of geometric distribution with parameter p and PMF given by (3.10). In order to compute the expectation and the second moment of N, one would substitute

$$f_1(p) = \frac{1-p}{p}, \quad f_2(p) = \frac{1-p}{p^2} + \frac{(1-p)^2}{p^2},$$

respectively, into (2.10). Routine integration of the resulting expressions, details of which shall be omitted, leads to the following result.

Proposition 2.7. *Let $N \sim GS_0(\alpha, k)$. If $\alpha > 1$ and $k \in \mathbb{N}$, then the mean of N is finite and is given by*

$$\mathbb{E}N = \frac{k}{\alpha - 1} - 1. \tag{2.11}$$

Further, the variance of N exists only if $\alpha > 2$, in which case we have

$$\mathbb{V}ar N = \frac{\alpha k(1 - \alpha + k)}{(\alpha - 1)^2(\alpha - 2)}. \tag{2.12}$$

Remark 2.8. We note that the expectation of a generalized Sibuya distribution is straightforward, as it does not involve any special functions or infinite series, which is not the case with discrete Pareto distribution (1.3), which has the same asymptotics of the tail. For example, the expectation of $N \sim GS_0(\alpha, k)$ with $k = 1$ and $\alpha = 1 + p$, where $0 < p < 1$, is equal to $\mathbb{E}N = (1 - p)/p$, and coincides with that of geometric variable with parameter p.

Remark 2.9. The mean of $N \sim GS_1(\alpha, k)$ also exists whenever $\alpha > 1$, in which case we have $\mathbb{E}N = k/(\alpha - 1)$. When $\alpha > 2$, its variance exists as well, and coincides with (2.12).

2.4. **The probability generating function.** The probability generating function of generalized Sibuya distribution can be obtained via the mixed Poisson representation (3.3), coupled with the relation (3.6). The following result provides relevant details.

Proposition 2.10. *If $N \sim GS_0(\alpha, k)$ then the PGF of N is*

$$G_N(t) = \frac{k!}{\Gamma(\alpha)\Gamma(1-\alpha+k)} \frac{\pi}{\sin(\pi\alpha)} \frac{1}{s^{1+k}} \left\{ \sum_{j=0}^{k} \frac{s^j (-\alpha)_j}{j!} - (1-s)^\alpha \right\}, \quad 0 < s < 1, \tag{2.13}$$

where $(a)_j$ denotes the Pochhammer's symbol (3.8).

Remark 2.11. If $N \sim GS_1(\alpha, k)$, then, due to the relation (2.1), its PGF is given by (2.13) with s^{1+k} replaced by s^k. In particular, when $k = 0$, we obtain the PGF (2.2) of Sibuya distribution with the PMF (1.8).

Remark 2.12. Note that for integer values of α, the quantity $\sin(\pi\alpha)$ in the denominator of the right-hand-side of (2.13) becomes zero. In this case the expression for the PGF is understood in the limiting sense. For example, by taking the limit as $\alpha \to 1$ of the right-hand-side of (2.13) with $k = 1$, we find that the PGF of $GS_0(1, 1)$ distribution is given by

$$G_N(t) = \frac{s + (1-s)\log(1-s)}{s}, \quad 0 < s < 1. \tag{2.14}$$

Remark 2.13. When $\alpha > 1$, the expectation of generalized Sibuya distribution can also be computed via the relation

$$\mathbb{E}N = \frac{d}{ds} G_N(s) \bigg|_{s=1}, \tag{2.15}$$

where G_N is the PGF of N. However, this is not a convenient way of getting the mean. For example, for non integer values of $\alpha > 1$, this leads the expression

$$\mathbb{E}N = \frac{k!}{\Gamma(\alpha)\Gamma(1-\alpha+k)} \frac{\pi}{\sin(\pi\alpha)} \left\{ \sum_{j=0}^{k-1} \frac{(-\alpha)_{j+1}}{j!} - (k+1) \sum_{j=0}^{k} \frac{(-\alpha)_j}{j!} \right\}, \tag{2.16}$$

which is not immediately seen to coincide with (2.11).

3. Stochastic representations

Below we provide an account of several stochastic representations of generalized Sibuya random variables, involving a randomly stopped Poisson process, mixtures of geometric distributions, and a discretization scheme.

3.1. **Randomly stopped Poisson process.** Consider a random variable

$$X \stackrel{d}{=} \frac{E}{T_{\alpha,k}} \tag{3.1}$$

where E and $T_{\alpha,k}$ are independent, E is standard exponential, and $T_{\alpha,k}$ has a beta distribution of the second kind, given by the PDF

$$f(x) = \frac{k!}{\Gamma(\alpha)\Gamma(1-\alpha+k)} x^{\alpha-1}\left(\frac{1}{1+x}\right)^{k+1}, \quad x \in \mathbb{R}_+ \;\; (k \in \mathbb{N}_0, \;\; 0 < \alpha < k+1). \tag{3.2}$$

Proposition 3.1. *If $N \sim GS_0(\alpha, k)$, then*

$$N \stackrel{d}{=} N(X), \tag{3.3}$$

where X is given by (3.1) and is independent of a standard Poisson process $\{N(t), t > 0\}$.

Remark 3.2. Note that $T_{\alpha,k}$ can be obtained as

$$T_{\alpha,k} \stackrel{d}{=} \frac{X_\alpha}{X_{\alpha-1+k}}, \tag{3.4}$$

where the variables on the right-hand-side of (3.4) are independent and X_β denotes standard gamma variable with shape parameter β (and unit scale). This shows that generalized Sibuya distribution is a special case $a = 1, b = 1-\alpha+k, c = \alpha$ of the *generalized hyperbolic distribution* of Type B3, defined via the stochastic representation (3.3) with

$$X \stackrel{d}{=} \frac{X_a X_b}{X_c}, \tag{3.5}$$

where the three variables on the right-hand-side of (3.5) are independent and have standard gamma distributions [see, e.g., Sibuya (1979), Sibuya and Shimizu (1981), or Devroye (1993)].

Remark 3.3. The result of Proposition 3.1 in case of (shifted) Sibuya distribution $GS_0(\alpha, 0)$ was noted in Devroye (1993) in connection with the problem of random variate generation from this distribution. Note that if $N \sim GS_1(\alpha, 0)$, then N does not admit the representation (3.3) with any X; instead, in this case we have $N \stackrel{d}{=} 1 + N(X)$.

Remark 3.4. It can be easily seen that if N admits the stochastic representation (3.3), then the PGF of N must be of the form

$$G_N(s) = \mathbb{E}s^N = \phi_X(1-s), \;\; s \in (0,1), \tag{3.6}$$

where $\phi_X(\cdot)$ is the Laplace transform of X [see, e.g., Steutel and van Harn (2004)]. This allows for a derivation of one of the functions, $G_N(\cdot)$ or $\phi_X(\cdot)$, from the other one. It can be shown [see the proof of Proposition 2.10 in the Appendix] that the Laplace transform of X defined by (3.1) is of the form

$$\phi_X(t) = \frac{k!}{\Gamma(\alpha)\Gamma(1-\alpha+k)} \frac{\pi}{\sin(\pi\alpha)} \frac{1}{(1-t)^{1+k}} \left\{ \sum_{j=0}^{k} \frac{(-\alpha)_j (1-t)^j}{j!} - t^\alpha \right\}, \;\; t \in \mathbb{R}_+, \tag{3.7}$$

where $(a)_n$ denotes Pochhammer's symbol defined as

$$(a)_n = \begin{cases} a(a+1)\cdots(a+n-1) & \text{for } n \geq 1 \\ 1 & \text{for } n = 0. \end{cases} \tag{3.8}$$

This leads to the PGF of generalized Sibuya distribution $GS_0(\alpha, k)$, given in Proposition 2.10. In case of (shifted) Sibuya distribution $GS_0(\alpha, 0)$, the function (3.7) reduces to

$$\phi_X(t) = \frac{1-t^\alpha}{1-t}, \;\; t \in \mathbb{R}_+, \tag{3.9}$$

which can also be recovered from the PGF of $N \sim GS_0(\alpha, 0)$ via $\phi_X(t) = G_N(1-t)$. However, if $N \sim GS_1(\alpha, 0)$ with the PGF (2.2), then $G_N(1-t)$ does not lead to a valid Laplace transform, as noted by Satheesh and Nair (2002).

3.2. **Randomly mixed geometric variable.** Our second representation shows that a generalized Sibuya distribution can be thought of as a mixed geometric distribution. The result below, which follows from the theory of generalized hypergeometric distributions of Type B3 [see, e.g, Sibuya (1979), Sibuya and Shimizu (1981)], can be proven directly from the representation (3.3) and standard conditioning arguments.

Proposition 3.5. *Let Y have a beta distribution with parameters α and $\beta = 1 - \alpha + k$, where $k \in \mathbb{N}_0$ and $0 < \alpha < k + 1$. Further, assume that, conditionally on $Y = p$, N has a geometric distribution with parameter p, i.e.*

$$\mathbb{P}(N = n | Y = p) = p(1 - p)^n, \;\; n \in \mathbb{N}_0. \tag{3.10}$$

Then, unconditionally, $N \sim GS_0(\alpha, k)$.

Remark 3.6. The $GS_1(\alpha, k)$ version of generalized Sibuya distribution is also mixed geometric with the same stochastic probability of success, but with a shifted-by-one version of the geometric variable.

Remark 3.7. This result can aid derivation of properties of generalized Sibuya distribution, such as its moments, by "mixing" the corresponding results for geometric distribution.

3.3. **Discretization scheme.** A generalized Sibuya variable arises also by a discretization scheme of the form $N = [W]$, where a discrete counterpart of a continuously distributed W is the integer part of W. A discrete counterpart of exponential distribution in this scheme is a geometric variable, while discretization of continuous Pareto II (Lomax distribution) leads to discrete Pareto distribution [see, e.g., Buddana and Kozubowski (2014)].

Proposition 3.8. *If W is a mixed exponential variable of the form*

$$W \stackrel{d}{=} \frac{E}{V_{\alpha,k}}, \tag{3.11}$$

where E and $V_{\alpha,k}$ are independent, E is standard exponential, and $V_{\alpha,k}$ has the PDF

$$g(x) = \frac{k! e^{-kx}}{\Gamma(\alpha)\Gamma(1 - \alpha + k)} (e^x - 1)^{\alpha - 1}, \;\; x \in \mathbb{R}_+ \;\; (k \in \mathbb{N}_0, \; 0 < \alpha < k + 1), \tag{3.12}$$

then $N = [W] \sim GS_0(\alpha, k)$.

4. Divisibility properties

4.1. **Infinite divisibility.** Recall that a random variable X (and its distribution) is *infinitely divisible* (ID) if for each $n \in \mathbb{N}$ it can be decomposed into the sum

$$X \stackrel{d}{=} X_{n,1} + \cdots + X_{n,n} \tag{4.1}$$

of IID random variables $\{X_{n,j}\}$ $(1 \le j \le n)$. Further, an integer-valued random variable X supported on $\mathbb{N}_0$ is *discrete infinitely divisible* if it is ID and the variables $\{X_{n,j}\}$ in (4.1) are integer-valued and supported on $\mathbb{N}_0$ as well. It is well-known that (shifted) Sibuya distribution $GS_0(\alpha, 0)$ is discrete ID [see, e.g., Christoph and Schreiber (2000)], implying that Sibuya distribution $GS_1(\alpha, 0)$ is ID (but not discrete ID). Similar properties hold for generalized Sibuya distribution, and follow from their representations as mixtures of geometric distributions, as the latter are ID [see, e.g., Steutel and van Harn (2004), Theorem 7.8, p. 381]. The following result summarizes these facts.

Proposition 4.1. *If $N \sim GS_0(\alpha, k)$, then the distribution of N is discrete ID (and thus ID). Further, the distribution of $N + 1 \sim GS_1(\alpha, k)$ is ID (but not discrete ID).*

This property allows us to build a continuous-time discrete-value stochastic processes based on the generalized Sibuya distribution. In particular, we can define a *Lévy motion* $\{N(t),\ \ t > 0\}$, a process with stationary, independent increments, where $N(1)$ is $GS_0(\alpha, k)$ with PGF G given by (2.13), and, for each $t > 0$, the PGF of $N(t)$ is G^t.

4.2. **Self-decomposability.** A discrete-valued random variable N supported on $\mathbb{N}_0$ is discrete self-decomposable (DSD) if for each $c \in (0, 1)$ it can be decomposed as

$$N \stackrel{d}{=} c \odot N + N_c, \tag{4.2}$$

where the variable N_c is also discrete-valued and supported on $\mathbb{N}_0$, and is independent of $c \odot N$ [see, e.g., Steutel and van Harn (1979)]. The dot product $c \odot N$ is the *discrete multiplication* (also known as *thinning*), defined as

$$c \odot N \stackrel{d}{=} \sum_{j=1}^{N} I_j, \ \ c \in (0, 1), \tag{4.3}$$

where the $\{I_j\}$ are IID Bernoulli variables with parameter c, independent of N. In terms of the PGFs, the condition (4.2) can be stated as

$$G_N(s) = G_N(1 - c + cs)G_c(s), \ \ s \in (0, 1), \tag{4.4}$$

where G_N is the PGF of N, $G_N(1 - c + cs)$ is the PGF of the dot product (4.3), and G_c is the PGF of N_c. It was shown by Christoph and Schreiber (2000) that the (shifted)

Sibuya distribution $GS_0(\alpha, 0)$ is DSD for each $\alpha \in (0,1)$. The following result provides an extension to the generalized Sibuya case.

Proposition 4.2. *If $N \sim GS_0(\alpha, k)$, then the distribution of N is discrete self decomposable.*

Remark 4.3. Let us note that if $N \sim GS_1(\alpha, k)$ then N is not DSD, since $\mathbb{P}(N = 0) = 0$. In particular, Sibuya distribution (1.8) is not DSD. However, for $c \in (0,1)$, the *scaled* Sibuya variable

$$N^{(c)} \stackrel{d}{=} c \odot N, \tag{4.5}$$

where $N \sim GS_1(\alpha, 0)$, may be DSD, depending on the value of c. Indeed, as shown in Christoph and Schreiber (2000), the variable (4.5) is DSD if and only if

$$0 < c \leq \left(\frac{1-\alpha}{1+\alpha}\right)^{1/\alpha}.$$

Moreover, it is also shown in Christoph and Schreiber (2000), that $N^{(c)}$ is (discrete) infinitely divisible if and only if $0 < c \leq (1-\alpha)^{1/\alpha}$.

4.3. **Invariance properties.** In this section we present an important new characterization of the Sibuya distribution, which is connected with the thinning operation (4.3) and (partially) explains the characterization of this distribution stated in Proposition 1.2. Let N have Sibuya distribution $GS_1(\alpha, 0)$, given by the PMF (1.8). As been observed by several authors [see, e.g., Christoph and Schreiber (2000)], the probability distribution corresponding to the scaled Sibuya variable $N^{(c)}$, defined by (4.5), is a mixture of a point mass at zero (with probability $1 - c^\alpha$) and the original distribution of N (with probability c^α). In other words, we can write

$$c \odot N \stackrel{d}{=} I^{(c)} \cdot N, \;\; c \in (0,1), \tag{4.6}$$

where $I^{(c)}$ is a Bernoulli random variable with parameter $p_c = c^\alpha$, independent of N. A natural question is whether the property (4.6) is unique to Sibuya distribution, that is whether there is any other variable N supported on $\mathbb{N}$ for which we have (4.6) with *some* $p_c \in (0,1)$. As shown below, there is no such distribution other than the Sibuya distribution.

Proposition 4.4. *If a random variable N supported on $\mathbb{N}$ satisfies the relation (4.6), where $I^{(c)}$ is a Bernoulli random variable with some parameter $p_c \in (0,1)$, independent of N, then N must have Sibuya distribution $GS_1(\alpha, 0)$ and $p_c = c^\alpha$.*

Observe that whenever we have (4.6), then for $n \in \mathbb{N}$

$$\mathbb{P}(c \odot N = n | c \odot N > 0) = \frac{\mathbb{P}(I^{(c)} \cdot N = n)}{\mathbb{P}(I^{(c)} \cdot N > 0)} = \frac{\mathbb{P}(I^{(c)} = 1)\mathbb{P}(N = n)}{1 - \mathbb{P}(I^{(c)} = 0)} = \mathbb{P}(N = n), \tag{4.7}$$

so that

$$c \odot N | c \odot N > 0 \stackrel{d}{=} N, \quad c \in (0,1). \tag{4.8}$$

In other words, the distribution of the thinned random variable $c \odot N$, conditioned on being positive, is the same as that of N, regardless of the thinning parameter $c \in (0,1)$. Note that, for $c \in (0,1)$ and any integer-valued variable N supported on $\mathbb{N}_0$, we have $\mathbb{P}(c \odot N = 0) = G_N(1-c) = 1 - \mathbb{P}(c \odot N > 0)$. Thus, if an integer-valued variable N supported on $\mathbb{N}$ satisfies (4.8), then it also satisfies (4.6) with

$$p_c = \mathbb{P}(I^{(c)} = 1) = 1 - G_N(1-c), \quad c \in (0,1). \tag{4.9}$$

Thus, in view of Proposition 4.4, the only distributions that are *stable with respect to the operation of thinning*, in the sense of (4.8), are Sibuya distributions.

Corollary 4.5. *Within the class of all probability distributions supported on $\mathbb{N}$, the stability property (4.8) is unique to Sibuya distributions $GS_1(\alpha, 0)$, defined by the PMF (1.8).*

Let us relate these properties to the characterization of the Sibuya distribution given in Proposition 1.2. Consider again the pure death process (1.12), connected with the population of N individuals, whose lifetimes $\{X_j\}$ are IID with the common CDF F. In terms of the operation of thinning, we have

$$N(t) \stackrel{d}{=} c(t) \odot N, \tag{4.10}$$

where $c(t) = \mathbb{P}(X_j > t) = 1 - F(t)$ is a function on $\mathbb{R}_+$ with the range coinciding with the interval $(0,1)$. In view of this, the condition (1.13) is essentially a restatement of (4.8), which, according to Corollary 4.5, is known to characterize the Sibuya distribution.

5. A Sibuya random process on $[0, 1]$

The Sibuya distribution with parameter α less than one arises as the marginal distribution of the *Sibuya random process* that we define as follows. Consider a sequence of IID uniform random variables U_n, $n \in \mathbb{N}$, and set

$$N(t) = \min\{n \in \mathbb{N}_0 : nU_n \leq t\}, \ t \in [0, 1], \tag{5.1}$$

with the convention that the minimum over an empty set is infinity, so that $N(0) = \infty$. Since for each $n \in \mathbb{N}$ we have $\mathbb{P}(nU_n \leq t) = t/n$, the variable $N(t)$ has the Sibuya distribution $GS_1(\alpha, 0)$ given by the PMF (1.8) with $\alpha = t$.

This process can be conveniently described throughout the classical concept of records. Let $\mathbf{x} = \{x_n\}$, $n \in \mathbb{N}$, be a sequence of positive numbers, and consider the pairs

$$(k_i, r_i) = (k_i(\mathbf{x}), r_i(\mathbf{x})), \ \ i \in \mathbb{N},$$

where the k_i is the time (index) at which the ith record occurs among the $\{x_i\}$, while $r_i = x_{k_i}$ is the size of that record. Here, a value that is *smaller* than all the previous values sets a new record, and x_1 is also considered to be a record, so that $k_1 = 1$ and $r_1 = x_1$. Further, assume that $x_1 \leq 1$, so that all the $\{r_i\}$ are smaller than one (while the $\{x_i\}$ are not required to be such). Moreover, let $\delta_i = r_{i-1} - r_i$ (with $r_0 = 1$) represent the differences between successive record values and let $\tau_i = k_i - k_{i-1}$ (with $k_0 = 0$) be the inter-arrival times between successive records. Under this notation, define

$$N_{\mathbf{x}}(t) = 1 + \sum_{i=1}^{\infty} \tau_{i+1} I_{(t,1]}(r_i), \ \ t \in [0, 1], \tag{5.2}$$

where, as before, I_A is an indicator function of the set A.

Clearly, the $N(t)$ defined by (5.1) is the same as the $N_{\mathbf{x}}(t)$ above if we take $\mathbf{x} = \{nU_n\}$. We see that, looking from right to left, the random process $N(t)$ initially "starts" with the value of one at $t = 1$ and then jumps up at every record value r_i, with the size of the jump being τ_{i+1}. Further, by the definition of the process, the values of $N(t)$ are constant on the intervals $[r_n, r_{n-1})$, and $N(r_n) = k_n$. The following result provides basic properties of the Sibuya random process $\{N(t), \ t \in [0, 1]\}$ discussed above.

Proposition 5.1. *For each $t \in [0,1]$, the marginal distribution $N(t)$ is Sibuya given by (1.8) with $\alpha = t$. Further, $N(t)$ is a right-continuous, pure jump, and non-increasing random process. Moreover, for any $\delta \in (0,1)$, the number of jumps is finite on the interval $[\delta, 1]$ and infinite on the interval $[0, \delta]$.*

Remark 5.2. One application of the Sibuya process is a construction of an extremal process on $[0,1]$ (and beyond) via Proposition 1.1, as discussed in Kozubowski and Podgórski (2016). For example, if $\{X_n\}$ is a sequence of IID random variables with the common CDF F and we let

$$Y(t) = \bigwedge_{n=1}^{N(t)} X_n, \quad t \in [0,1], \tag{5.3}$$

where $N(t)$ is the Sibuya process defined above, independent of the $\{X_n\}$, then the CDF of $Y(t)$ is given by F^t for each $t \in [0,1]$. This extends the notion of an extreme value of n IID random variables to fractional values of n.

We now look at the sample path structure of the Sibuya process. For convenience, we will look at a time-reversed process $S(t) = N(1-t)$, as it is more natural to follow the evolution of the sample paths from left to right. In Figure 1, we schematically present a part of a sample path of $S(t)$.

By Proposition 5.1, $S(t)$ is a pure-jump process whose sample paths (which start at $S(0) = 1$ almost surely) are continuous from the left and non-decreasing. We already know from the above construction, that the jumps of this process and the waiting times between them, are closely related to record inter-arrival times and record sizes connected with a random sequence $\{nU_n\}$. Here, the locations of the jumps occur at an increasing sequence $\{T_i\}$, where $T_i = 1 - R_i$ and the R_i is the location of the ith jump of the process $N(t)$ (counted from right to left). Our first result describes the joint distribution of the locations of the jumps.

Proposition 5.3. *Let $S(t) = N(1-t)$, where $\{N(t)\ t \in [0,1]\}$ is a Sibuya process defined by (5.1), and let $\{\Gamma_n\}$ be the successive arrival times of a standard Poisson process. Then for each $n \in \mathbb{N}$ we have*

$$(T_1, \ldots, T_n) \stackrel{d}{=} (H(\Gamma_1), \ldots, H(\Gamma_n)), \tag{5.4}$$

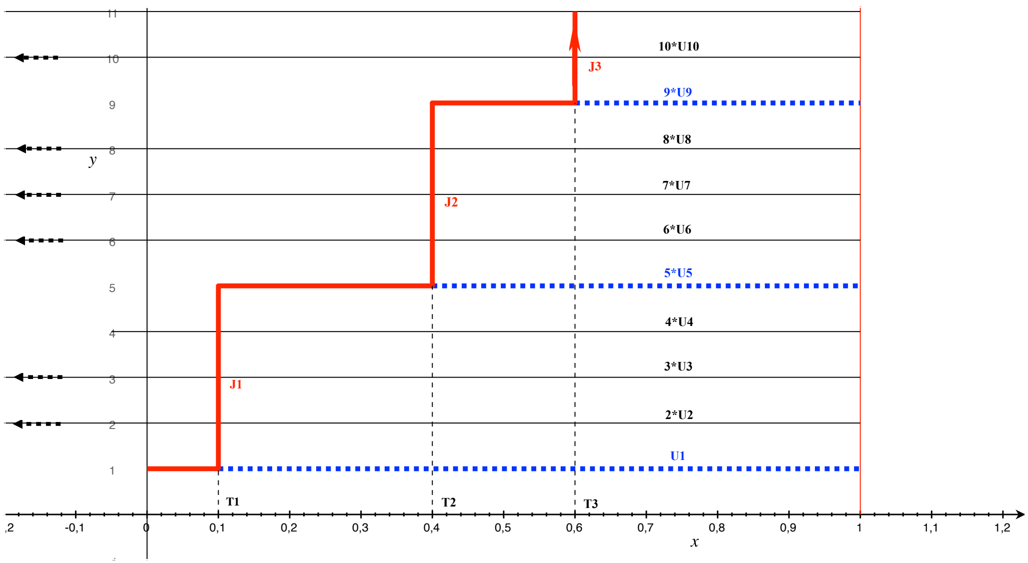

FIGURE 1. A schematic sketch of an individual sample of the Sibuya process $S(t)$. The horizontal lines represent the values of nU_n, with the doted lines marking records at which the process jumps. Three locations (T_1, T_2, T_3) and corresponding jumps (J_1, J_2, J_3) are shown and the trajectory is marked by a thick line. The arrows indicate that the lengths of lines extend beyond the figure due to large values of nU_n and the jumps size J_3.

where $0 < T_1 < \cdots < T_n < 1$ *are the (random) locations of the first* n *jumps of* $S(t)$ *and* H *is the CDF of standard exponential distribution.*

Remark 5.4. It follows that the location of the first jump of the process $S(t)$ has a standard uniform distribution, while for $n \geq 2$ the joint distribution of the locations of its first n jumps is given by the PDF

$$g(t_1, \ldots, t_n) = \frac{1}{(1-t_1)(1-t_2)\cdots(1-t_{n-1})}, \quad 0 < t_1 < \cdots < t_n < 1. \tag{5.5}$$

An equivalent description of this is through the conditional distributions: for each $n \in \mathbb{N}$, the conditional distribution of T_n given the $n-1$ values $0 < t_1 < \cdots < t_{n-1} < 1$ of the previous jump locations has a uniform distribution on the interval $(t_{n-1}, 1)$. This distribution is known as *random division of the unit interval.*

Remark 5.5. It follows that if the time has not been reversed, the jumps of the Sibuya process (5.1), when viewed from right to left, occur at the points $\exp(-\Gamma_n)$, $n \in \mathbb{N}$. Moreover, if the time line is stretched to $(0, \infty)$ via logarithmic transformation $t \to -\log(1-t)$, the locations of the jumps of $S(t)$ will coincide with those of the standard Poisson process.

Our final result, concerning the joint conditional distribution of the jump sizes and their locations, shades light on the probabilistic structure of the time-reversed Sibuya process. Using the above notation connected with the record process, we shall look at the evolution of the sequence of random points (T_i, K_i), $i \in \mathbb{N}$, where K_i is the time of the ith record connected with the sequence $\{nU_n\}$. As illustrated in Figure 1, $S(t)$ is a pure jump process started at $S(0) = 1$, with the first jump occurring at the random location T_1. Regarding the first random point (T_1, K_1), we have $K_1 = 1$ and, by Proposition 5.3, the variable T_1 is standard uniform. We now consider the second pair (T_2, K_2), conditioned on the event $\mathcal{B}_1 = \{T_1 = t_1, K_1 = 1\}$, and consider the joint distribution of (T_2, J_2), where $J_2 = K_2 - K_1$ is the size of the jump of $S(t)$ at $t = t_1$. By the construction of the process $S(t)$, for $t_1 < t < 1$ we have

$$\mathbb{P}(T_2 > t, J_2 = n|\mathcal{B}_1) = \mathbb{P}(2U_2 > 1 - t_1, \ldots, nU_n > 1 - t_1, (n+1)U_{n+1} < 1 - t)),$$

so that

$$\mathbb{P}(T_2 > t, J_2 = n|\mathcal{B}_1) = p(r_1, 1, n)\frac{1-t}{1-t_1}, \tag{5.6}$$

where $r_1 = 1 - t_1$ and

$$p(r, k, n) = \left(1 - \frac{r}{k+1}\right)\cdots\left(1 - \frac{r}{k+n-1}\right)\frac{r}{k+n}, \quad n \in \mathbb{N}, \tag{5.7}$$

represents the probability $\mathbb{P}(S = n)$ with $S \sim GS_1(r, k)$. In view of (5.6) and the fact that the fraction on the right-hand-side in (5.6) is the probability $\mathbb{P}(T_2 > t|T_1 = t_1)$, we conclude that, conditioned on $\mathcal{B}_1$, the variables T_2 and J_2 are independent, with the latter having the generalized Sibuya distribution $GS_1(1 - t_1, 1)$. These calculations extend in a straightforward way beyond the second pair (T_2, J_2), leading to the following result.

Proposition 5.6. *In the above setting, conditioned on* $\mathcal{B}_n = \{T_1 = t_1, \ldots, T_n = t_n, K_1 = k_1, \ldots K_n = k_n\}$*, the variables* T_{n+1} *and* J_{n+1} *are independent, with* T_{n+1} *having uniform distribution on* $(t_n, 1)$ *and with* $J_{n+1} \sim GS_1(1 - t_n, k_n)$*.*

According to the above result, the conditional distributions of the jumps of the time-reversed Sibuya process $S(t)$ have generalized Sibuya distributions.

6. Appendix

Proof of Proposition 1.2. Suppose that, for some $\alpha \in (0, 1)$, N has Sibuya distribution $GS_1(\alpha, 0)$, given by the PMF (1.8). Then, for each $t \in \mathbb{R}_+$, the value of the process $N(t)$ defined by (1.12) admits the stochastic representation (4.10), where $c(t) = \mathbb{P}(X_j > t)$. Since N is Sibuya, we also have (4.6) with $c = c(t)$, which, in turn, implies (4.7). Thus $N(t)$ satisfies (1.13), as desired. Next, assume that $N(t)$ satisfies equation (1.13). Thus, for each $t \in \mathbb{R}_+$, we have

$$\mathbb{P}(X(t) = n) = \mathbb{P}(N = n)\mathbb{P}(X(t) > 0), \quad n \in \mathbb{N}. \tag{6.1}$$

Using standard conditioning argument, write

$$\mathbb{P}(X(t) = n) = \sum_{k=1}^{\infty} \mathbb{P}(X(t) = n | N = k)\mathbb{P}(N = k). \tag{6.2}$$

Noting that for $k < 0$ we have $\mathbb{P}(X(t) = n | N = k) = 0$ while for $k \geq n$ the variable $X(t) = n | N = k$ is binomial with parameters k and $p = 1 - F(t)$, where F is the CDF of the $\{X_j\}$, we conclude that

$$\mathbb{P}(X(t) = n) = \sum_{k=n}^{\infty} \binom{k}{n} [1 - F(t)]^n [F(t)]^{k-n} \mathbb{P}(N = k), \quad n \in \mathbb{N}, \quad t \in \mathbb{R}_+. \tag{6.3}$$

For $n = 0$, equation (6.2) shows that

$$\mathbb{P}(X(t) = 0) = \sum_{k=1}^{\infty} [F(t)]^k \mathbb{P}(N = k), \quad t \in \mathbb{R}_+. \tag{6.4}$$

We now write $s = F(t) \in (0, 1)$ and $p_n = \mathbb{P}(N = n)$ and substitute (6.2) and (6.4) into (6.1), which results in the following equation

$$(1 - s)^n \sum_{k=n}^{\infty} \binom{k}{n} s^{k-n} p_k = p_n \left(1 - \sum_{k=1}^{\infty} s^k p_k\right), \quad n \in \mathbb{N}, \quad s \in (0, 1). \tag{6.5}$$

Further, by expanding the term $(1-s)^n$ into a power series in s and changing the index of the summation on the left-hand-side of (6.5) to $j = k - n$, we conclude that

$$\left\{\sum_{k=0}^{n}\binom{n}{k}(-1)^k s^k\right\}\cdot\left\{\sum_{j=0}^{\infty}\binom{j+n}{n}p_{j+n}s^j\right\} = p_n - \sum_{j=1}^{\infty}p_j p_n s^j,\ \ n\in\mathbb{N},\ \ s\in(0,1). \tag{6.6}$$

Using standard result for power series, stating that the coefficients c_k of the product

$$\sum_{k=0}^{\infty}c_k s^k = \left\{\sum_{i=0}^{\infty}a_i s^i\right\}\cdot\left\{\sum_{j=0}^{\infty}b_j s^j\right\} \tag{6.7}$$

are given by

$$c_k = \sum_{i=0}^{k}a_i b_{k-i},$$

following some algebra, we conclude the left-hand-side of (6.6) is of the form (6.7) with

$$c_k = \sum_{j=0}^{k}\binom{n}{j}\binom{k-j+n}{n}p_{k-j+n}(-1)^j,\ \ 0\le k\le n,\ \ n\in\mathbb{N}. \tag{6.8}$$

Thus, in view of the above coupled with (6.6), and by the uniqueness of the power series, we conclude that

$$\sum_{j=0}^{k}\binom{n}{j}\binom{k-j+n}{n}p_{k-j+n}(-1)^j = -p_k p_n,\ \ 1\le k\le n,\ \ n\in\mathbb{N}. \tag{6.9}$$

In particular, for $k=1$, relation (6.9) reduces to

$$(n+1)p_{n+1} - np_n = -p_1 p_n,\ \ n\in\mathbb{N}, \tag{6.10}$$

leading to

$$p_{n+1} = \frac{(n-p_1)p_n}{n+1},\ \ n\in\mathbb{N}. \tag{6.11}$$

It now follows by induction that the $\{p_n\}$ coincide with Sibuya probabilities (1.8), where $\alpha = p_1 = \mathbb{P}(N=1)$. This concludes the proof.

Proof of Proposition 2.3. Since, in view of (2.5), the results of Proposition 2.3 and Corollary 2.5 are equivalent, it is enough to establish (2.9). First, by incorporating the well-known property of the gamma function,

$$\Gamma(\eta+k) = \Gamma(\eta)\eta(\eta+1)\cdots(\eta+k-1),\ \ \eta\in\mathbb{R}_+, k\in\mathbb{N}, \tag{6.12}$$

the generalized Sibuya SF (2.4) can be written as

$$\mathbb{P}(N > n) = \frac{1}{n^\alpha}\frac{\Gamma(k+1-\alpha+n)}{\Gamma(n)n^{k+1-\alpha}}\frac{\Gamma(n)n^{k+1}}{\Gamma(n+k+1)}\frac{\Gamma(k+1)}{\Gamma(k+1-\alpha)}. \tag{6.13}$$

Next, since for any $\gamma > 0$ we have

$$\frac{\Gamma(\gamma+n)}{\Gamma(n)n^\gamma} \to 1 \text{ as } n \to \infty,$$

the right-hand-side of (6.13) divided by the right-hand-side of (2.9) converges to 1 as $n \to \infty$, as desired.

Proof of Proposition 2.10. By Proposition 3.1, the PGF of N is given by (3.6), where $\phi_N(\cdot)$ is the LT of the variable X defined in (3.1). To prove the result, it is enough to show that the LT of X is given by (3.7). To establish the latter, we condition on $T_{\alpha,k}$ when computing the LT of X, leading to

$$\phi_X(t) = \int_0^\infty \mathbb{E}e^{-tE/x} f(x)dx,$$

where $f(x)$ is given in (3.2) and E is standard exponential with the LT

$$\mathbb{E}e^{-tE} = \frac{1}{1+t}, \ \ t \in \mathbb{R}_+.$$

Thus, after some algebra, we obtain

$$\phi_X(t) = \frac{k!}{\Gamma(\alpha)\Gamma(1-\alpha+k)}\int_0^\infty \frac{x^\alpha}{(t+x)(1+x)^{k+1}}dx.$$

The result now follows by the integration formula 6 on p. 321 of Gradshteyn and Ryzhik (2007) with $\nu = \alpha+1$, $n = 1+k$, $\gamma = 1$, and $\beta = t$.

Proof of Proposition 3.1. It is known [see, e.g., Devroye (1993)] that the generalized hypergeometric distribution of Type B3, given in (3.3) with X as in (3.5), is of the form

$$\mathbb{P}(N = n) = \frac{\Gamma(a+c)\Gamma(b+c)\Gamma(a+n)\Gamma(b+n)}{\Gamma(a)\Gamma(b)\Gamma(c)\Gamma(a+b+c+n)n!}, \ \ n \in \mathbb{N}_0. \tag{6.14}$$

Setting $a = 1$, $b = 1-\alpha+k$, and $c = \alpha$ in (6.14) produces the $GS_0(\alpha, k)$ distribution.

Proof of Proposition 3.8. We proceed by showing that the PMF of the variable $[W]$ coincides with that of the $GS_0(\alpha, k)$ distribution. First, using standard conditioning argument, write

$$\mathbb{P}([W]=n)=\int_0^\infty \mathbb{P}([E/x]=n)g(x)dx,\;\; n\in\mathbb{N}_0, \tag{6.15}$$

where E is unit exponential and g is the PDF of $V_{\alpha,k}$, given by (3.12). Since

$$\mathbb{P}([E/x]=n)=\mathbb{P}(nx\le E<(n+1)x)=e^{-nx}-e^{(n+1)x},$$

the probability (6.15) takes on the form

$$\mathbb{P}([W]=n)=\frac{k!}{\Gamma(\alpha)\Gamma(1-\alpha+k)}\left\{I_{k+n}(\alpha)-I_{k+n+1}(\alpha)\right\}, \tag{6.16}$$

where

$$I_k(\alpha)=\int_0^\infty e^{-kx}(e^x-1)^{\alpha-1}dx,\;\; k\in\mathbb{N}_0,\;\; 0<\alpha<k+1. \tag{6.17}$$

Noting that the function $g(\cdot)$ in (3.12) is a genuine PDF for each $k\in\mathbb{N}_0$ and $0<\alpha<k+1$, we conclude that

$$I_k(\alpha)=\frac{\Gamma(\alpha)\Gamma(1-\alpha+k)}{k!},\;\; k\in\mathbb{N}_0,\;\; 0<\alpha<k+1. \tag{6.18}$$

A substitution of (6.18) into (6.16), followed by some algebra, produces the $GS_0(\alpha,k)$ distribution. This concludes the proof.

Proof of Proposition 4.2. To prove the result, we shall use the following sufficient condition for this property to hold [see, Bondesson (1992), p. 28]: A strictly decreasing PMF $\{p_n\}$, $n\in\mathbb{N}_0$, is DSD if

$$\max_{0\le n\le j}\frac{p_{n+1}}{p_n}\le\frac{j+2}{j+1}\frac{p_{j+1}-p_{j+2}}{p_j-p_{j+1}},\;\; j\in\mathbb{N}_0. \tag{6.19}$$

First, we shall show that generalized Sibuya PMF is strictly decreasing in n. To see this, note that the ratio

$$\frac{p_{n+1}}{p_n}=\frac{\mathbb{P}(N=n+1)}{\mathbb{P}(N=n)}=\frac{k+n+1-\alpha}{k+n+2},\;\; n\in\mathbb{N}_0, \tag{6.20}$$

is strictly increasing in $n\in\mathbb{N}_0$. Indeed, the derivative of the function

$$g(x)=\frac{k+1-\alpha+x}{k+2+x},\;\; x\in\mathbb{R}_+, \tag{6.21}$$

is positive for all $x\in\mathbb{R}_+$, which can be checked by straightforward algebra. Since the ratio (6.20) converges to 1 as $n\to\infty$, we conclude that $p_{n+1}/p_n<1$ for all $n\in\mathbb{N}_0$, showing the monotonicity of the sequence $\{p_n\}$, $n \in \mathbb{N}_0$. This also shows that the maximum on the left-hand-side of (6.19) is attained for $n = j$, so that the condition (6.19) becomes

$$\frac{p_{j+1}}{p_j} \le \frac{j+2}{j+1}\frac{p_{j+1} - p_{j+2}}{p_j - p_{j+1}}, \quad j \in \mathbb{N}_0. \tag{6.22}$$

After some algebra, condition (6.22) can be re-stated as

$$(j+1)\left(1 - \frac{p_{j+1}}{p_j}\right) \le (j+2)\left(1 - \frac{p_{j+2}}{p_{j+1}}\right), \quad j \in \mathbb{N}_0. \tag{6.23}$$

Since

$$(j+1)\left(1 - \frac{p_{j+1}}{p_j}\right) = \frac{(j+1)(1+\alpha)}{k+1+(j+1)}, \quad j \in \mathbb{N}_0, \tag{6.24}$$

and the function

$$h(x) = \frac{x(1+\alpha)}{k+1+x} = \frac{1+\alpha}{1+\frac{k+1}{x}}$$

is non-decreasing in $x \in \mathbb{R}_+$, we obtain (6.23). This concludes the proof.

Proof of Proposition 4.4. According to the remarks following the statement of Proposition 4.4, condition (4.6) implies (4.8), which, in view of (4.10), is equivalent to (1.13). The result now follows from Proposition 1.2.

Proof of Proposition 5.3. For $n = 1$, the statement is trivial. To prove the result for general $n \in \mathbb{N}$, it is enough to show that for each $n \ge 2$, the conditional distribution of T_n given the $n-1$ values $0 < t_1 < \cdots < t_{n-1} < 1$ of the previous jump locations has a uniform distribution on the interval $(t_{n-1}, 1)$. Indeed, in this case the PDF of the joint distribution of $(T_1, \ldots, T_n)$ is easily seen to be given by (5.5). This, in turn, is the joint PDF of the random vector on the right-hand-side of (5.4), as can be verified by standard calculation.

To establish the above we start with $n = 2$, and consider the conditional probability $\mathbb{P}(T_2 > t | T_1 = t_1)$ for $t_1 < t < 1$. Using the law of total probability we obtain

$$\mathbb{P}(T_2 > t | T_1 = t_1) = \sum_{k=2}^{\infty} \mathbb{P}(R_2 < 1 - t, K_2 = k | R_1 = r_1),$$

where (K_i, R_i) are the random pairs of record times and their sizes (with $R_i = 1 - T_i$), connected with the sequence $\{nU_n\}$ (as described in Section 5). Note that the probability under the above sum can be written in terms of the $\{U_n\}$ as

$$\mathbb{P}(R_2 < 1-t, K_2 = k | R_1 = r_1) = \mathbb{P}(2U_2 > r_1, \ldots, (k-1)U_{k-1} > r_1, kU_k < 1-t),$$

or, equivalently, as

$$\mathbb{P}(R_2 < 1-t, K_2 = k | R_1 = r_1) = p(r_1, k)\frac{1-t}{r_1},$$

where

$$p(r_1, k) = \left(1 - \frac{r_1}{2}\right) \cdots \left(1 - \frac{r_1}{k-1}\right)\frac{r_1}{k}, \;\; k \geq 2.$$

When compared with (1.7), the quantity $p(r_1, k)$ is recognized as the probability $\mathbb{P}(S = k-1)$, where $S \sim GS_1(r_1, 1)$. Consequently,

$$\mathbb{P}(T_2 > t | T_1 = t_1) = \frac{1-t}{r_1}\sum_{k=2}^{\infty} p(r_1, k) = \frac{1-t}{1-t_1},$$

since the probabilities above sum up to one. Since the quantity on the right-hand-side above is the survival function of the uniform distribution on the interval $(t_1, 1)$, the result holds for $n = 2$. The proof in the case $k > 2$ is similar. Under the same notation and using again the law of total probability, we have

$$\mathbb{P}(T_n > t | A_{n-1}) = \sum_{k=n-1}^{\infty}\sum_{m=1}^{\infty} \mathbb{P}(T_n > t, K_n = k+m | K_{n-1} = k)\mathbb{P}(K_{n-1} = k), \;\; t_{n-1} < t < 1,$$

where A_{n-1} denotes the condition $T_1 = t_1, \ldots, T_{n-1} = t_{n-1}$. Similarly as before, the conditional probabilities under the double sum above can be expressed as

$$\mathbb{P}(T_n > t, K_n = k+m | K_{n-1} = k) = p(r_{n-1}, k, m)\frac{1-t}{r_{n-1}},$$

where

$$p(r_{n-1}, k, m) = \left(1 - \frac{r_{n-1}}{k+1}\right) \cdots \left(1 - \frac{r_{n-1}}{k+m-1}\right)\frac{r_{n-1}}{k+m}, \;\; m \in \mathbb{N},$$

is recognized as the probability $\mathbb{P}(S = m)$ with $S \sim GS_1(r_1, k)$. Since these probabilities sum up to one across the values of $m \in \mathbb{N}_0$, and so do the probabilities $\mathbb{P}(K_{n-1} = k)$ across the values of $k \geq n-1$, we obtain

$$\mathbb{P}(T_n > t | A_{n-1}) = \frac{1-t}{r_{n-1}}\sum_{k=n-1}^{\infty} \mathbb{P}(K_{n-1} = k)\sum_{m=1}^{\infty} p(r_{n-1}, k, m) = \frac{1-t}{1-t_{n-1}}, \;\; t_{n-1} < t < 1.$$

Since the quantity on the right-hand-side above is the survival function of the uniform distribution on the interval $(t_{n-1}, 1)$, the result follows.

Acknowledgments

Research of the first author was partially funded by the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 318984 - RARE. The second author was partially supported by Riksbankens Jubileumsfond Grant Dnr: P13-1024.

References

[1] Aban, I.B., Meerschaert, M.M. and Panorska, A.K. (2006). Parameter estimation for the truncated Pareto distribution, *J. Amer. Statist. Assoc.* **101**, 270-277.

[2] Bondesson, L. (1992). *Generalized Gamma Convolutions and Related Classes of Distributions and Densities*, Lecture Notes in Statistics **76**, Springer, Berlin.

[3] Buddana, A. and Kozubowski, T.J. (2014). Discrete Pareto distributions, *Econ. Qual. Control* **29**(2), 143-156.

[4] Christoph, G. and Schreiber, K. (1998). Discrete stable random variables, *Statist. Probab. Lett.* **37**, 243-247.

[5] Christoph, G. and Schreiber, K. (2000). Shifted and scaled Sibuya distribution and discrete self-decomposability, *Statist. Probab. Lett.* **48**(2), 181-187.

[6] Clauset, A. and Newman, M.E.J. (2009). Power-law distributions in empirical data, *SIAM Rev.* **2**, 661-703.

[7] Devroye, L. (1993). A triptych of discrete distributions related to the stable law, *Statist. Probab. Lett.* **18**, 349-351.

[8] Gabaix, X. (2009). Power laws in economics and finance, *Annual Rev. Econ.* **1**, 255-293.

[9] Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on the new mode of determining the value of life contingencies, *Phil. Trans. Roy. Soc. London* **115**, 513-585.

[10] Gradshteyn, I.S. and Ryzhik, I.M. (2007). *Tables of Integrals, Series, and Products*, 7th Ed., Academic Press, Amsterdam.

[11] Johnson, N.L., Kotz, S. and Kemp, A.W. (1993) *Univariate Discrete Distributions*, John Wiley & Sons, New York.

[12] Kozubowski, T.J. and Podgórski, K. (2016). Certain bivariate distributions and random processes connected with maxima and minima, *Working Papers in Statistics* **2016:9**, Department of Statistics, School of Economics and Management, Lund University.

[13] Krishna, H. and Singh Pundir, P. (2009). Discrete Burr and discrete Pareto distributions, *Statist. Meth.* **6**, 177-188.

[14] Newman, M.E.J. (2005). Power laws, Pareto distributions and Zipf's law, *Contemp. Physics* **46**, 323-351.

[15] Pakes, A.G. (1995). Characterization of discrete laws via mixed sums and Markov branching processes, *Stochastic Proc. Appl.* **55**, 285-300.

[16] Pillai, R.N. and Jayakumar, K. (1995). Discrete Mittag-Leffler distributions, *Statist. Probab. Lett.* **23**, 271-274.

[17] Rényi, A. (1962), (1976). On outstanding values of a sequence of observations, In: *Selected Papers of A. Rényi*, Vol. 3, pp. 50-65, Akadémiai Kiadó, Budapest.

[18] Satheesh, S. and Nair, N.U. (2002). Some classes of distributions on the non-negative lattice, *J. Indian Statist. Assoc.* **40**(1), 41-58.

[19] Sibuya, M. (1979). Generalized hypergeometric, digamma, and trigamma ditributions, *Ann. Inst. Statist. Math.* **31**, 373-390.

[20] Sibuya, M. and Shimizu, R. (1981). The generalized hypergeometric family of distributions, *Ann. Inst. Statist. Math.* **33**, 177-190.

[21] Sornette, D. (2006). *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools*, 2nd Ed., Springer, Berlin.

[22] Steutel, F.W. and van Harn, K. (1979). Discrete analogues of self-decomposability and stability, *Ann. Probab.* **7**, 893-899.

[23] Steutel, F.W. and van Harn, K. (2004). *Infinitely Divisibility of Probability Distributions on the Real Line*, Marcel Dekker, New York.

[24] Stumpf, M.P.H. and Porter, M.A. (2012). Critical truths about power laws, *Science* **335**, 665-666.

[25] Yule, G.U. (1925). A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis, F.R.S., *Philosophical Transactions of the Royal Society of London Ser. B.* **213**, 21-87.

[26] Zipf, G.K. (1949). *Human Behavior and the Principle of Least Effort*, Addison-Wesley, Cambridge, MA.

http://journals.lub.lu.se/stat

LUND UNIVERSITY
School of Economics and Management

Working Papers in Statistics 2016
LUND UNIVERSITY
SCHOOL OF ECONOMICS AND MANAGEMENT
Department of Statistics
Box 743
220 07 Lund, Sweden