PRUNED DISCRETE RANDOM SAMPLES

RUDOLF GRÜBEL,* *Leibniz Universität Hannover*

PAWEŁ HITCZENKO,** *Drexel University*

Abstract

Let X_i, $i \in \mathbb{N}$, be independent and identically distributed random variables with values in $\mathbb{N}_0$. We transform ('prune') the sequence $\{X_1, \ldots, X_n\}$, $n \in \mathbb{N}$, of discrete random samples into a sequence $\{0, 1, 2, \ldots, Y_n\}$, $n \in \mathbb{N}$, of contiguous random sets by replacing X_{n+1} with $Y_n + 1$ if $X_{n+1} > Y_n$. We consider the asymptotic behaviour of Y_n as $n \to \infty$. Applications include path growth in digital search trees and the number of tables in Pitman's Chinese restaurant process if the latter is conditioned on its limit value.

Keywords: Chinese restaurant process; digital search trees; geometric distribution; maxima; tail behaviour

2010 Mathematics Subject Classification: Primary 60C05
Secondary 60F99; 60J10

1. Introduction

Let X_i, $i \in \mathbb{N}$, be independent and identically distributed random variables with values in the set $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ of nonnegative integers. We assume that

$$p_k := \mathbb{P}(X_i = k) > 0 \quad \text{for all } k \in \mathbb{N}_0. \tag{1}$$

In [7] we investigated the gaps in the set $\{X_1, \ldots, X_n\} \subset \mathbb{N}_0$ and found 'phase transitions'—in the sense of a qualitative change of behaviour—at two standard discrete distributions. Roughly, if the tails $(q_k)_{k \in \mathbb{N}_0}$,

$$q_k := \mathbb{P}(X_i \geq k) = \sum_{j=k}^{\infty} p_j, \tag{2}$$

of the underlying distribution are thinner than that of a Poisson distribution then, with probability 1, the gaps will eventually vanish as $n \to \infty$. In the range between Poissonian and geometric tail behaviour we still see that the probability of a gap tends to 0. Finally, if the tails are heavier than that of a geometric distribution then the length of the longest gap in the sample tends to ∞ in probability.

In the present paper we consider the 'pruned' samples, by which we mean the sequence of random sets $\{0, \ldots, Y_n\}$ defined inductively by $Y_1 := 0$ and

$$Y_{n+1} := \begin{cases} Y_n + 1 & \text{if } X_{n+1} > Y_n, \\ Y_n & \text{otherwise,} \end{cases} \quad \text{for all } n \in \mathbb{N}.$$

Hence, values beyond the current range are reduced so that the range remains contiguous. Clearly, this leads to a Markov chain $(Y_n)_{n \in \mathbb{N}}$ with state space $\mathbb{N}_0$, starting at 0. This chain is of pure birth type and thus fully described by the transition probabilities $p_{i,i+1}$ of moving from i to $i + 1$. Indeed, here we have $p_{i,i+1} = q_{i+1}$, and it is easy to see that, whenever $i \to p_{i,i+1}$ decreases and has limit 0, then the birth chain can be represented by a pruned discrete random sample. Our results can therefore be used to obtain the asymptotic behaviour of a wide class of pure birth processes.

For later use, and also in view of another interpretation of the Y-sequence, we need both the 'number of failures' and the 'waiting for the first success' versions of the geometric distribution, writing $Z \sim \text{Geo}_0(p)$ if $\mathbb{P}(Z = k) = q^k p$ for all $k \in \mathbb{N}_0$ and $Z \sim \text{Geo}_1(p)$ if $\mathbb{P}(Z = k) = q^{k-1} p$ for all $k \in \mathbb{N}$, $q = 1 - p$. We may then regard $(Y_n)_{n \in \mathbb{N}}$ as the (nonhomogeneous) renewal process associated with independent lifetimes G_k, $G_k \sim \text{Geo}_1(q_k)$ for all $k \in \mathbb{N}$.

As in [7], our motivation for the study of these processes comes from the analysis of algorithms and random discrete structures.

Example 1. In the digital search tree (DST) algorithm infinite zero–one sequences $x_n = (x_{n,i})_{i \in \mathbb{N}} \in \{0, 1\}^{\mathbb{N}}$, $n \in \mathbb{N}$, are stored sequentially into the nodes $u \in \bigcup_{k=0}^{\infty} \{0, 1\}^k$ of a binary tree in a 'trickle-down' manner: starting at the root we use the digits (entries) $x_{n,i}$ of x_n to choose either the left (if $x_{n,i} = 0$) or right (if $x_{n,i} = 1$) successor of an occupied node until an empty node is found. This algorithm and its properties are discussed in [4], [9], and [10].

Suppose now that the x_ns are independent and that they all have the same distribution μ. For any node $u = (u_1, \ldots, u_k)$, let μ_u be the probability that x_n has prefix u, i.e. that $x_{n,i} = u_i$ for $i = 1, \ldots, k$. Along a specified path $v = (v_i)_{i \in \mathbb{N}} \in \{0, 1\}^{\infty}$ the sequence of random trees then grows like a pruned discrete random sample, with q_k equal to μ_u, $u := (v_1, \ldots, v_k)$. Important special cases are the symmetric binomial model, with $\mu_u = 2^{-|u|}$, $|u|$ being the depth of the node u, and the general binomial model with parameter $\theta \in (0, 1)$, where $\mu_u = \theta^l (1 - \theta)^m$, l and m being the number of entries of u that are equal to 0 and 1, respectively. More general μs appear in connection with the binary search tree (BST) algorithm if we condition on the value of the limit tree; see [5].

Example 2. The following construction of a sequence $(X_n)_{n \in \mathbb{N}}$ of random partitions

$$X_n = \{B_{n,1}, \ldots, B_{n,k_n}\}$$

of the sets $\{1, \ldots, n\}$, $n \in \mathbb{N}$, is known as Pitman's Chinese restaurant process; see, e.g. [11]. Starting with $X_1 = \{\{1\}\}$ we obtain X_{n+1} from X_n by choosing the next partition, now of $\{1, \ldots, n, n + 1\}$, to be

$$X_{n+1} = \{B_{n,1}, \ldots, B_{n,k_n}, \{n + 1\}\}$$

with probability $1/(n + 1)$ and

$$X_{n+1} = \{B_{n,1}, \ldots, B_{n,j-1}, B_{n,j} \cup \{n + 1\}, \ldots, B_{n,k_n}\}$$

with probability $\#B_{n,j}/(n + 1)$, $j = 1, \ldots, k_n$. Here $\#A$ denotes the number of elements of the set A. Clearly, $(X_n)_{n \in \mathbb{N}}$ is a transient Markov chain.

Pitman also introduced parametrized extensions of this process. Of these, the processes with an infinite base set are h-transforms of each other, and the Doob–Martin compactification can be given explicitly [5]. The elements of the boundary may be represented by probability distributions $\rho = (\rho_k)_{k \in \mathbb{N}}$ on $\mathbb{N}$, and, conditionally on the limit being ρ, the random sequence $(K_n)_{n \in \mathbb{N}}$ of partition sizes (number of occupied tables in the restaurant model) is identical in

distribution to $(Y_n + 1)_{n\in\mathbb{N}}$, where Y_n is the rightmost value in the pruned sample sequence associated with the distribution $\tilde{\rho} = (\tilde{\rho}_k)_{k\in\mathbb{N}_0}$ on $\mathbb{N}_0$ given by $\tilde{\rho}_k := \rho_{k+1}$, $k \in \mathbb{N}_0$. This is also related to Kingman's paint box construction; again, we refer the reader to [11] for details and further references.

We are interested in the asymptotic behaviour of Y_n, especially in the way the asymptotics depend on the behaviour of the tail probabilities defined in (2). We will see that, again, there are phase transitions when the tail behaviour moves from the thin to the heavy end of the spectrum.

In the next section we list and discuss our results; proofs are deferred to Section 3.

2. Results

We begin at the thin end, comparing the Y_ns to the maxima

$$M_n := \max\{X_1, \ldots, X_n\}$$

of the samples. We have $\mathbb{P}(M_n < k) = (1 - q_k)^n$ for all $k, n \in \mathbb{N}$, i.e. the distribution of the M-variables depends on the distribution of the X-variables in a simple and explicit way.

Clearly, $Y_n \le M_n$ for all $n \in \mathbb{N}$. Also, a jump in the M-sequence entails a jump of size 1 in the Y-sequence, but note that a jump in the M-sequence may be larger than 1, and that the Y-sequence may jump where the M-sequence does not. The following result shows that, for rapidly decreasing tails, we have, with probability 1, $Y_n = M_n$ eventually.

Theorem 1. *The sequence $(M_n - Y_n)_{n\in\mathbb{N}}$ converges to 0 with probability 1 as $n \to \infty$ if and only if*

$$\sum_{k=0}^{\infty} \frac{q_{k+1}}{q_k} < \infty. \tag{3}$$

We also have a sufficient condition for convergence in probability in terms of $(q_k)_{k\in\mathbb{N}}$. For a partial converse, i.e. in order to deduce this condition from the in-probability behaviour of the differences, we use a specific random subsequence of the full sequence.

Theorem 2. *If*

$$\lim_{k\to\infty} \frac{q_{k+1}}{q_k} = 0 \tag{4}$$

then the sequence $(M_n - Y_n)_{n\in\mathbb{N}}$ converges to 0 in probability as $n \to \infty$. Furthermore, with

$$\tau_m := \inf\{n \in \mathbb{N}: Y_n = m\}, \qquad m \in \mathbb{N},$$

convergence of the sequence $(M_{\tau_m} - Y_{\tau_m})_{m\in\mathbb{N}}$ to 0 in probability as $m \to \infty$ implies (4).

We note in passing that Theorems 1 and 2 together imply that, for samples from a Poisson distribution, the differences $M_n - Y_n$ converge to 0 in probability but not almost surely.

As in [7], the geometric case turns out to be of special interest, in view of its appearance in connection with the DST algorithm, and also in view of its role as a borderline case. Because of our general assumption (1), we use the 'number of failures' version $\mathrm{Geo}_0(p)$; see also Remark 1 of [7]. We also put $r := 1/q$. It is well known that, asymptotically, the sequence of maxima for samples from this distribution, after an appropriate shift, exhibits periodic distributional fluctuations. Indeed, with a suitable background construction, see, e.g. [6, p. 454], we see that $M_{n_m} - \lfloor \log_r m_n \rfloor$ converges in distribution along subsequences $(n_m)_{m\in\mathbb{N}}$ that are such that

$$\lim_{m\to\infty} \{\log_r n_m\} = \eta \quad \text{for some } \eta \in [0, 1). \tag{5}$$

Here we have written $\lfloor x \rfloor := \max\{k \in \mathbb{Z}: k \leq x\}$ for the 'floor' of x and $\{x\} := x - \lfloor x \rfloor$ for its fractional part. The limit distribution is the same as the distribution of $\lfloor (\log r)Z_1 + \eta \rfloor$, with $\mathbb{P}(Z_1 \leq z) = e^{-e^{-z}}$, the Gumbel distribution. A similar result holds for $(Y_n)_{n \in \mathbb{N}}$; see [3]. Along subsequences $(n_m)_{m \in \mathbb{N}}$ that satisfy (5), the sequence $(Y_{n_m} - \lfloor \log_r m_n \rfloor)_{m \in \mathbb{N}}$ converges in distribution, where the distribution of the limit is the same as the distribution of $\lfloor (-\log r)Z_2 + \eta \rfloor$, with $Z_2 := \sum_{j=1}^{\infty} q^j V_j$ and $(V_j)_{j \in \mathbb{N}}$ an independent and identically distributed sequence of standard exponentials (Dennert and Grübel [3] only treated the case $p = \frac{1}{2}$, but the extension to general p is straightforward). Remarkably, for the differences, the periodicities cancel.

Theorem 3. *If the random variables X_i are independent and $X_i \sim \text{Geo}_0(p)$ for all $i \in \mathbb{N}$, then*

$$\lim_{n \to \infty} \mathbb{P}(M_n - Y_n = k) = \pi_k := \frac{a_k}{\sum_{l=0}^{\infty} a_l} \quad \text{for all } k \in \mathbb{N}_0,$$

where the sequence $(a_k)_{k \in \mathbb{N}_0}$ is given recursively by $a_0 := 1$ and

$$a_{k+1} := \frac{1}{1 - q^{k+1}}\left(q a_k - p\sum_{j=0}^{k-1} q^{k-j} a_j\right) \quad \text{for all } k \in \mathbb{N}_0.$$

In the proof we will see that $a_k > 0$ for all $k \in \mathbb{N}$, and that indeed $\sum_{l=0}^{\infty} a_l < \infty$.

At the heavy end of the spectrum of tail behaviour, M_n and Y_n drift apart as $n \to \infty$; hence, we deal with the Y_ns directly. The following result on sums of geometric random variables is our main tool for this range of tail behaviour; we believe that it is of interest in its own right.

Theorem 4. *Assume that the random variables G_k, $k \in \mathbb{N}$, are independent and that $G_k \sim \text{Geo}_1(q_k)$, with $q_k = \mathbb{P}(X_1 \geq k)$ as in (2). Set $S_n = \sum_{k=1}^{n} G_k$. Assume further that*

$$\frac{q_{k+1}}{q_k} \to 1 \quad \text{as } n \to \infty. \tag{6}$$

(a) *Then*

$$\frac{S_n}{\mathbb{E}S_n} \xrightarrow{\mathbb{P}} 1 \quad \text{as } n \to \infty.$$

(b) *For all $m \in \mathbb{N}$, let*

$$r_m := \frac{\mathbb{E}S_m}{\mathbb{E}G_m} = q_m \sum_{k=1}^{m} \frac{1}{q_k}.$$

Then

$$\frac{S_n}{\mathbb{E}S_n} \xrightarrow{\text{a.s.}} 1 \quad \text{as } n \to \infty$$

if and only if the following condition holds:

$$\sum_{m=1}^{\infty} e^{-\delta r_m} < \infty \quad \text{for all } \delta > 0. \tag{7}$$

In Section 3.5 we will show that (6) implies that $r_m \to \infty$, but that the rate may be arbitrarily slow. In particular, condition (7) is not vacuous.

In order to use this for the pruned samples, we first recall that a function $L \colon \mathbb{R}_+ \to \mathbb{R}$ is said to be slowly varying (at ∞) if

$$\lim_{x \to \infty} \frac{L(\alpha x)}{L(x)} = 1 \quad \text{for all } \alpha > 0.$$

A standard example is $L(x) = (\log x)^\beta$, $\beta \in \mathbb{R}$. Clearly, if L is slowly varying then so is $x \mapsto 1/L(x)$. Such a function L has a de Bruijn conjugate $L^\#$, unique up to asymptotic equivalence, which is again slowly varying, and satisfies

$$\lim_{x \to \infty} L(x) L^\#(x L(x)) = 1 \quad \text{and} \quad \lim_{x \to \infty} L^\#(x) L(x L^\#(x)) = 1.$$

This is Theorem 1.5.13 of [1], our basic reference for this area. For $L(x) = (\log x)^\beta$, we obtain $L^\# = 1/L$, i.e. $L^\#(x) = (\log x)^{-\beta}$.

Note that $(Y_n)_{n \in \mathbb{N}}$ is the renewal process associated with the lifetimes $(G_k)_{k \in \mathbb{N}}$ from Theorem 4. We now obtain our main result for the heavy-tailed case by a process known as renewal inversion. We write $a_n \sim b_n$ for two real sequences $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ if $\lim_{n \to \infty}(a_n/b_n) = 1$.

Theorem 5. *Suppose that, for some $\alpha > 0$,*

$$q_k \sim k^{-\alpha} L(k) \quad \text{as } k \to \infty, \tag{8}$$

with L slowly varying. Then, with probability 1,

$$Y_n \sim n^{1/(1+\alpha)} L_0(n) \quad \text{as } n \to \infty,$$

where L_0 is the de Bruijn conjugate of $1/L$.

3. Proofs

For very thin tails, we use a pathwise argument that builds on the corresponding result in [7]. For moderately thin tails, we use suitable bounds for the probability in question, together with direct arguments. The geometric case is approached via a suitable bivariate Markov chain. For heavy tails, we first prove the result on sums of geometrically distributed random variables and then obtain the result on the asymptotics of pruned samples.

In addition to the processes $X = (X_n)_{n \in \mathbb{N}}$, $Y = (Y_n)_{n \in \mathbb{N}}$, and $M = (M_n)_{n \in \mathbb{N}}$, we define another process $\Delta = (\Delta_n)_{n \in \mathbb{N}}$ by $\Delta_n := M_n - Y_n$, $n \in \mathbb{N}$.

3.1. Proof of Theorem 1

Suppose first that (3) is satisfied. Let $\nu_1 := 1$ and, for $k > 1$,

$$\nu_k := \inf\{n > \nu_{k-1} \colon X_n > X_{\nu_{k-1}}\}$$

be the jump times of the process M. With probability 1, these are all finite. Let

$$A_k := \{\omega \in \Omega \colon X_n(\omega) = M_{\nu_k}(\omega) \text{ for some } n \in \{\nu_k(\omega) + 1, \ldots, \nu_{k+1}(\omega) - 1\}\}$$

be the event that the X-process visits the old maximum before it achieves the next new maximum. Let $A_\infty := \limsup_{k \to \infty} A_k$ be the event that infinitely many of the A_ks, $k \in \mathbb{N}$, happen. Finally, let

$$B_\infty := \{\omega \in \Omega \colon \text{there exists a } k_0 \text{ such that, for all } k \geq k_0,\ M_{\nu_{k+1}}(\omega) = M_{\nu_k}(\omega) + 1\} \tag{9}$$

be the event that, eventually, the jumps of M all have size 1.

Let $\omega \in A_\infty \cap B_\infty$, with $k_0 = k_0(\omega)$ as in (9). Suppose that $\Delta_{\nu_{k_0}}(\omega) = j$. Consider the evolution of the associated M- and Y-paths from $n = \nu_{k_0}(\omega)$ onwards: M jumps by at most 1, and then Y also jumps by 1; furthermore, if Y visits the old maximum between two jump times, then Δ decreases by 1 if it is greater than 0. By the definition of A_∞, there will be infinitely many of these visits; hence, $\Delta_n(\omega) = 0$ from some $n_0 = n_0(\omega)$ onwards.

For the proof of sufficiency of (3), it is therefore enough to show that $\mathbb{P}(B_\infty) = 1$ and $\mathbb{P}(A_\infty) = 1$. The first of these is an immediate consequence of Theorem 1 of [7]. For the second, we consider A_k: on $M_{\nu_k} = j$, this is the event that the post-ν_k process $(X_{\nu_k+n})_{n\in\mathbb{N}}$ hits j when it first leaves the set $\{0, \ldots, j-1\}$. A straightforward calculation leads to

$$\mathbb{P}(A_k \mid M_{\nu_k} = j) = \frac{p_j}{q_j}. \tag{10}$$

From (1) and (3), it follows that

$$\alpha := \inf_{j\in\mathbb{N}_0} \frac{p_j}{q_j} > 0. \tag{11}$$

The post-ν_k process $(X_{\nu_k+n})_{n\in\mathbb{N}}$ and the pre-ν_k process $(X_{\nu_k\wedge n})_{n\in\mathbb{N}}$ are independent; hence, (10) and (11) together lead to

$$\mathbb{P}(A_k \mid \mathcal{F}_{\nu_k}) \geq \alpha \quad \text{for all } k \in \mathbb{N},$$

with $(\mathcal{F}_n)_{n\in\mathbb{N}}$ the natural filtration associated with the process X. Clearly, $A_k \in \mathcal{F}_{\nu_{k+1}}$, and $\mathbb{P}(A_\infty) = 1$ now follows with Lévy's generalization of the Borel–Cantelli lemma; see, e.g. [8, Corollary 6.20].

For the proof of the necessity of (3), we note that, as Δ_n has nonnegative integer values, if $\Delta_n \to 0$ almost surely then there is a set A with $\mathbb{P}(A) = 1$ such that,

for all $\omega \in A$, there exists an $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$, $\Delta_n(\omega) = 0$.

On this set, the gaps will eventually vanish as the Y-process can only increase by steps of size 1; hence, we can apply the other direction of Theorem 1 of [7].

3.2. Proof of Theorem 2

We first show that condition (4) is sufficient. For this, we need the following auxiliary statement.

Lemma 1. *If $q_{k+1}/q_k \to 0$ as $k \to \infty$ then*

$$q_m \sum_{j=1}^{m-1} \frac{1}{q_j} \to 0 \quad \text{as } m \to \infty.$$

Proof. Let $\varepsilon > 0$ be given. For $\eta > 0$, pick $k_0 \in \mathbb{N}$ such that $q_{k+1}/q_k < \eta$ for $k \geq k_0$. Then, for $m > k_0$, we have

$$q_m \sum_{j=1}^{m-1} \frac{1}{q_j} = q_m \sum_{j=1}^{k_0-1} \frac{1}{q_j} + \sum_{j=k_0}^{m-1} \frac{q_m}{q_j} < q_m \sum_{j=1}^{k_0-1} \frac{1}{q_j} + \sum_{j=k_0}^{m-1} \eta^{m-j} < q_m \sum_{j=1}^{k_0-1} \frac{1}{q_j} + \frac{\eta}{1-\eta}.$$

If η is such that $\eta < \min\{\varepsilon/3, \frac{1}{2}\}$ then $\eta/(1-\eta) < 2\eta < 2\varepsilon/3$. Since $q_m \to 0$, we can find an m_0 such that, for $m \geq m_0$, $q_m \sum_{j=1}^{k_0-1} 1/q_j < \varepsilon/3$. This proves the lemma. $\square$

We next show that (4) implies that $M_n - Y_n \to 0$ in probability, which is the first part of Theorem 2.

Take $\varepsilon > 0$, and consider $\mathbb{P}(M_n - Y_n > \varepsilon)$. Since both random variables are integer valued we have, for all $m_0 \in \mathbb{N}$,

$$
\begin{aligned}
\mathbb{P}(M_n - Y_n > \varepsilon) &\leq \mathbb{P}(M_n \geq Y_n + 1) \\
&= \sum_{m=1}^{\infty} \mathbb{P}(M_n \geq m, \ Y_n = m - 1) \\
&\leq \mathbb{P}(Y_n \leq m_0) + \sum_{m=m_0}^{\infty} \mathbb{P}(M_n \geq m, \ Y_n = m - 1) \\
&= \mathbb{P}(Y_n \leq m_0) + \sum_{m=m_0}^{\infty} \mathbb{P}(M_n \geq m \mid Y_n = m - 1)\mathbb{P}(Y_n = m - 1).
\end{aligned}
$$

The key step is the following bound on the conditional probability.

Lemma 2. *For any $m \geq 1$, we have*

$$
\mathbb{P}(M_n \geq m \mid Y_n = m - 1) = 1 - \prod_{j=1}^{m-1}\left(1 - \frac{q_m}{q_j}\right) \leq q_m \sum_{j=1}^{m-1}\frac{1}{q_j}. \tag{12}
$$

Once we prove this lemma the proof of the theorem may be completed as follows. For $\eta > 0$, first pick $m_0 \in \mathbb{N}$ such that, for $m \geq m_0$, $q_m \sum_{j=1}^{m-1} 1/q_j < \eta/2$. This is possible by Lemma 1. Then, since $Y_n \to \infty$ in probability, there exists $n_0 \in \mathbb{N}$ such that, for $n \geq n_0$, we have $\mathbb{P}(Y_n \leq m_0) < \eta/2$. It now follows that, for $n \geq n_0$, we have

$$
\mathbb{P}(M_n - Y_n > \varepsilon) \leq \mathbb{P}(Y_n \leq m_0) + \frac{\eta}{2}\sum_{m=m_0}^{\infty}\mathbb{P}(Y_n = m - 1) \leq \eta.
$$

Proof of Lemma 2. Describe the event $\{Y_n = m - 1\}$ in terms of the sequence (X_k). First, with $\tau_0 := 0$ and τ_m ($m \in \mathbb{N}$) as in the theorem,

$$
\{Y_n = m - 1\} = \{\tau_{m-1} \leq n < \tau_m\} = \bigcup_{1 \leq k_1 < \cdots < k_{m-1} \leq n} \bigcap_{j=1}^{m-1}\{\tau_j = k_j\} \cap \{\tau_m > n\}.
$$

Note that the union is over pairwise disjoint events. Therefore, it is enough to show that, for all $1 \leq k_1 < \cdots < k_{m-1} \leq n$ and $m \geq 1$,

$$
\mathbb{P}\left(M_n \geq m \ \Bigg| \ \bigcap_{j=1}^{m-1}\{\tau_j = k_j\} \cap \{\tau_m > n\}\right) = 1 - \prod_{j=1}^{m-1}\left(1 - \frac{q_m}{q_j}\right). \tag{13}
$$

Fix $1 \leq k_1 < \cdots < k_{m-1} \leq n$. The event $\bigcap_{j=1}^{m-1}\{\tau_j = k_j\} \cap \{\tau_m > n\}$ means that

$$
\begin{aligned}
&X^*_{[1:k_1-1]} = 0, \ X_{k_1} \geq 1, \ X^*_{[k_1+1,k_2-1]} \leq 1, \ X_{k_2} \geq 2, \dots, \\
&X_{k_{m-1}} \geq m - 1, \ X^*_{[k_{m-1}+1,n]} \leq m - 1,
\end{aligned}
$$

where, for a set $A \subset \mathbb{N}$, $X_A^* = \max_{k \in A} X_k$, and, for nonnegative integers $a \leq b$, $[a, b] = \{a, a + 1, \ldots, b\}$. Therefore, conditionally on this event, $M_n \geq m$ means that $\max_{1 \leq j \leq m-1} X_{k_j} \geq m$. Thus, the conditional probability in (13) is

$$
\mathbb{P}\left(\max_{1 \leq j < m} X_{k_j} \geq m \mid \bigcap_{j=1}^{m-1} \{\tau_j = k_j\} \cap \{\tau_m > n\} \right)
$$

$$
= 1 - \mathbb{P}\left(\bigcap_{j=1}^{m-1} \{X_{k_j} < m\} \mid \bigcap_{j=1}^{m-1} \{\tau_j = k_j\} \cap \{\tau_m > n\} \right).
$$

By the above description in terms of random variables (X_k) of the event we are conditioning on, the independence of the X_ks, and the definition of the conditional probability, we see that the last probability is

$$
\frac{\mathbb{P}(j \leq X_{k_j} < m; \ j = 1, \ldots, m-1)}{\mathbb{P}(X_{k_j} \geq j; \ j = 1, \ldots, m-1)} = \prod_{j=1}^{m-1} \frac{q_j - q_m}{q_j} = \prod_{j=1}^{m-1} \left(1 - \frac{q_m}{q_j} \right).
$$

This shows (13) and completes the proof as the inequality in (12) is clear.

For the proof of the second part of Theorem 2, we begin with

$$
\mathbb{P}(X_{\tau_m} = j) = \sum_{l=1}^{\infty} \mathbb{P}(X_l = j \mid \tau_m = l)\mathbb{P}(\tau_m = l).
$$

In view of $\{\tau_m = l\} = \{X_l \geq m\} \cap \{Y_{l-1} = m - 1\}$ and the fact that Y_{l-1} is a function of $X_1, \ldots, X_{l-1}$ we have $\mathbb{P}(X_l = j \mid \tau_m = l) = \mathbb{P}(X_1 = j \mid X_1 \geq m)$, which leads to

$$
\mathbb{P}(\Delta_{\tau_m} > 0) \geq \mathbb{P}(X_{\tau_m} \geq m + 1) = \mathbb{P}(X_1 \geq m + 1 \mid X_1 \geq m) = \frac{q_{m+1}}{q_m}.
$$

Hence, if (4) does not hold, i.e. if $\limsup_{m \to \infty}(q_{m+1}/q_m) > 0$, then $\mathbb{P}(\Delta_{\tau_m} > 0)$ cannot converge to 0.

3.3. Proof of Theorem 3

We require the following auxiliary result which roughly says that the limit distribution of a process does not change if its clock is slowed down in a specific manner.

Proposition 1. *Suppose that $Y = (Y_n)_{n \in \mathbb{N}}$ and $D = (D_n)_{n \in \mathbb{N}_0}$ are independent Markov chains, where Y is of pure birth type with birth probabilities*

$$
\mathbb{P}(Y_{n+1} = i + 1 \mid Y_n = i) = 1 - \mathbb{P}(Y_{n+1} = i \mid Y_n = i) = \mu_i > 0 \quad \text{for all } i \in \mathbb{N}_0,
$$

and start at 0, and that D is irreducible and aperiodic, with state space E, transition probabilities

$$
\mathbb{P}(D_{n+1} = j \mid D_n = i) = p_D(i, j) \quad \text{for all } i, j \in E,
$$

and stationary distribution $\pi = (\pi_i)_{i \in E}$. Let $Z = (Z_n)_{n \in \mathbb{N}}$ be defined by

$$
Z_n := (Y_n, D_{Y_n}) \quad \text{for all } n \in \mathbb{N}.
$$

Then Z is a Markov chain with state space $\mathbb{N}_0 \times E$ and transition probabilities

$$p_Z((i, j), (i, j)) = 1 - \mu_i \quad \text{for all } i \in \mathbb{N}_0, \; j \in E, \tag{14}$$

$$p_Z((i, j), (i + 1, k)) = \mu_i \, p_D(j, k) \quad \text{for all } i \in \mathbb{N}_0, \; j, k \in E. \tag{15}$$

Furthermore,

$$\lim_{n \to \infty} \mathbb{P}(D_{Y_n} = k) = \pi_k \quad \text{for all } k \in E. \tag{16}$$

Proof. We first show that

$$\mathbb{P}(Z_{n+1} = (i_{n+1}, j_{n+1}) \mid Z_0 = (i_0, j_0), \dots, Z_n = (i_n, j_n)) \tag{17}$$

is the same as

$$\mathbb{P}(Z_{n+1} = (i_{n+1}, j_{n+1}) \mid Z_n = (i_n, j_n)). \tag{18}$$

For this, it is enough to consider paths with $i_{n+1} \in \{i_n, i_n + 1\}$ and $j_{n+1} = j_n$ if $i_{n+1} = i_n$. Also, $(i_0, \dots, i_n)$ has to be of the form

$$(\underbrace{0, \dots, 0}_{\nu_0}, \underbrace{1, \dots, 1}_{\nu_1}, \dots, \underbrace{l, \dots, l}_{\nu_l})$$

with $l \in \mathbb{N}_0$ and $\nu_0 + \cdots + \nu_l = n$. Then

$$(j_0, \dots, j_n) = (\underbrace{k_0, \dots, k_0}_{\nu_0}, \underbrace{k_1, \dots, k_1}_{\nu_1}, \dots, \underbrace{k_l, \dots, k_l}_{\nu_l}).$$

In such cases (17) reduces to either

$$\mathbb{P}(Y_{n+1} = l, \; D_l = k_l \mid Y_0 = 0, \dots, Y_n = l, \; D_0 = k_0, \dots, D_l = k_l)$$

or

$$\mathbb{P}(Y_{n+1} = l + 1, \; D_{l+1} = k_{l+1} \mid Y_0 = 0, \dots, Y_n = l, \; D_0 = k_0, \dots, D_l = k_l),$$

which equals $1 - \mu_l$ or $\mu_l \, p_D(k_l . k_{l+1})$, respectively. The same values appear for (18). This shows that Z is a Markov chain and also confirms (14) and (15).

For the proof of (16), let $k \in E$ and $\varepsilon > 0$. A decomposition with respect to the value of the clock together with the independence of Y and D results in

$$\mathbb{P}(D_{Y_n} = k) = \sum_{m=0}^{\infty} \mathbb{P}(D_m = k)\mathbb{P}(Y_n = m).$$

The ergodic theorem for Markov chains gives $\lim_{m \to \infty} \mathbb{P}(D_m = k) = \pi_k$, so, for a suitable m_0,

$$|\mathbb{P}(D_m = k) - \pi_k| \leq \tfrac{1}{2}\varepsilon \quad \text{for all } m \geq m_0.$$

As the birth probabilities are all strictly greater than 0, it follows further that Y_n converges to ∞ with probability 1, and, hence, in probability. This implies that, for some n_0,

$$\mathbb{P}(Y_n < m_0) \leq \tfrac{1}{2}\varepsilon \quad \text{for all } n \geq n_0.$$

Combining these we obtain, for all $n \geq n_0$,

$$
\begin{aligned}
|\mathbb{P}(D_{Y_n} = k) - \pi_k| &\leq \left(\sum_{m=0}^{m_0-1} + \sum_{m=m_0}^{\infty} \right) |\mathbb{P}(D_m = k) - \pi_k| \mathbb{P}(Y_n = m) \\
&\leq \mathbb{P}(Y_n < m_0) + \sup_{m \geq m_0} |\mathbb{P}(D_m = k) - \pi_k| \\
&\leq \varepsilon.
\end{aligned}
$$

This completes the proof of the proposition.

We remark in passing that, in contrast to Y and D, the process $(D_{Y_n})_{n \in \mathbb{N}}$ is not a time-homogeneous Markov chain; indeed, if $\lim_{i \to \infty} \mu_i = 0$ then the 'stay where you are' probabilities $\mathbb{P}(D_{Y_{n+1}} = k \mid D_{Y_n} = k)$ tend to 1 as $n \to \infty$ for each $k \in E$.

For the proof of Theorem 3, we now define the bivariate process $V = (V_n)_{n \in \mathbb{N}}$ by $V_n := (Y_n, \Delta_n)$ for all $n \in \mathbb{N}$. The dynamics of V are given by $V_{n+1} = \Phi(V_n, X_{n+1})$ with

$$
\Phi((i, j), k) = \begin{cases} (i, j) & \text{if } k \leq i, \\ (i + 1, j - 1) & \text{if } i < k \leq i + j, \\ (i + 1, j + l - 1) & \text{if } k = i + j + l \text{ for some } l \in \mathbb{N}. \end{cases}
$$

This implies that V is a Markov chain with state space $\mathbb{N}_0 \times \mathbb{N}_0$ and transition probabilities

$$
\begin{aligned}
p_V((i, j), (i, j)) &= 1 - q_{i+1}, \\
p_V((i, j), (i + 1, j - 1)) &= q_{i+1} - q_{i+j+1}, \\
p_V((i, j), (i + 1, j + k)) &= p_{i+j+k+1}, \quad k \in \mathbb{N}_0,
\end{aligned}
$$

in the general case. For $\mathrm{Geo}_0(p)$, as in the theorem, we obtain

$$
\begin{aligned}
p_V((i, j), (i, j)) &= 1 - q^{i+1}, \\
p_V((i, j), (i + 1, j - 1)) &= q^{i+1}(1 - q^j), \\
p_V((i, j), (i + 1, j + k)) &= q^{i+1} p q^{j+k}, \quad k \in \mathbb{N}_0.
\end{aligned}
$$

Now let $D = (D_n)_{n \in \mathbb{N}_0}$ be a Markov chain, independent of Y, with state space $E = \mathbb{N}_0$, starting at 0, and transitions

$$
p_D(j, j - 1) = 1 - q^j, \qquad p_D(j, j + k) = p q^{j+k} \quad \text{for all } k \in \mathbb{N}_0. \tag{19}
$$

Clearly, D is irreducible and aperiodic, and positive recurrence follows with Foster's theorem [2, p. 167].

Let $Z = (Z_n)_{n \in \mathbb{N}}$ with $Z_n := (Y_n, D_{Y_n})$ for all $n \in \mathbb{N}$. By Proposition 1, Z is a Markov chain and, on the basis of the above computations, it is easily checked that $p_Z = p_V$. Markov chains with the same transition probabilities and the same initial value are equal in distribution; in particular, $(\Delta_n)_{n \in \mathbb{N}}$ has the same distribution as $(D_{Y_n})_{n \in \mathbb{N}}$. The final statement in Proposition 1 now implies that Δ_n converges in distribution to the stationary distribution π of D. Invariance of π under the transition mechanism (19) means that π solves

$$
\pi_k = (1 - q^{k+1}) \pi_{k+1} + \sum_{j=0}^{k} p q^j \pi_{k-j} \quad \text{for all } k \in \mathbb{N}_0,
$$

which implies that it is of the form given in the theorem.

3.4. Proof of Theorem 4

We begin with a lemma.

Lemma 3. *Assume that $q_{k+1}/q_k \to 1$ as $k \to \infty$. Then, as $m \to \infty$,*

$$q_m \sum_{k=1}^{m} \frac{1}{q_k} \to \infty, \tag{20}$$

but it may go to ∞ at an arbitrarily slow rate.

Proof. For the first assertion, take an arbitrary $M > 0$. Since $q_{k+1}/q_k \to 1$, there exists an m_0 such that, for $m \geq m_0$, we have $q_{m+1}/q_m > 1 - 1/M$. Then, for $m > m_0$,

$$
\begin{aligned}
q_m \sum_{k=1}^{m} \frac{1}{q_k} &\geq \sum_{k=m_0+1}^{m} \frac{q_m}{q_k} \\
&\geq \sum_{k=m_0+1}^{m} \left(1 - \frac{1}{M}\right)^{m-k} \\
&= \sum_{j=0}^{m-m_0-1} \left(1 - \frac{1}{M}\right)^{j} \\
&= M\left(1 - \left(1 - \frac{1}{M}\right)^{m-m_0}\right) \\
&\geq M(1 - e^{-1}),
\end{aligned}
$$

provided $m - m_0 \geq M$. This shows (20).

To prove the second assertion of the lemma, let $k_1, k_2, \ldots \in \mathbb{N}$ and $\varepsilon_1, \varepsilon_2, \ldots$ decreasing to 0 be given. Let the first k_1 q_ms be defined by

$$(1 - \varepsilon_1), (1 - \varepsilon_1)^2, \ldots, (1 - \varepsilon_1)^{k_1},$$

the next k_2 by

$$(1 - \varepsilon_1)^{k_1}(1 - \varepsilon_2), (1 - \varepsilon_1)^{k_1}(1 - \varepsilon_2)^2, \ldots, (1 - \varepsilon_1)^{k_1}(1 - \varepsilon_2)^{k_2},$$

and so on, with the last group of k_m q_ks given by

$$\prod_{j=1}^{m-1}(1 - \varepsilon_j)^{k_j}(1 - \varepsilon_m), \prod_{j=1}^{m-1}(1 - \varepsilon_j)^{k_j}(1 - \varepsilon_m)^2, \ldots, \prod_{j=1}^{m-1}(1 - \varepsilon_j)^{k_j}(1 - \varepsilon_m)^{k_m}.$$

Set $K_r = \sum_{i=1}^{r} k_i$. Then, for $K_{r-1} \leq k \leq K_r$, we have

$$\prod_{j=1}^{r}(1 - \varepsilon_j)^{k_j} \leq q_k \leq \prod_{j=1}^{r-1}(1 - \varepsilon_j)^{k_j}.$$

Hence,

$$\sum_{k=K_{r-1}+1}^{K_r} \frac{1}{q_k} \leq k_r \prod_{j=1}^{r}(1 - \varepsilon_j)^{-k_j},$$

and, therefore, provided that the k_js are increasing,

$$\sum_{k=1}^{K_m} \frac{1}{q_k} \le k_m \sum_{k=1}^{m} \prod_{j=1}^{k} (1 - \varepsilon_j)^{-k_j}.$$

It follows that, for $K_{m-1} \le n \le K_m$,

$$q_n \sum_{k=1}^{n} \frac{1}{q_k} \le k_m \prod_{j=1}^{m-1} (1 - \varepsilon_j)^{k_j} \sum_{k=1}^{m} \prod_{j=1}^{k} (1 - \varepsilon_j)^{-k_j}$$

$$= k_m \sum_{k=1}^{m} \prod_{j=1}^{k} (1 - \varepsilon_j)^{-k_j} \prod_{j=1}^{m-1} (1 - \varepsilon_j)^{k_j}$$

$$= k_m \left\{ \sum_{k=1}^{m-1} \prod_{j=k+1}^{m-1} (1 - \varepsilon_j)^{k_j} + (1 - \varepsilon_m)^{-k_m} \right\},$$

where, as usual, we adopt the convention that a product over an empty range is 1. Now, upon setting $\varepsilon_i = 1/k_i$ we see that the term in the curly braces is bounded by an absolute constant $C \le e/(e-1) + 4$ (this may be seen by choosing $k_1 \ge 2$ and using the elementary inequality $1 - x \le e^{-x}$ to bound the first sum). Thus,

$$q_n \sum_{k=1}^{n} \frac{1}{q_k} \le C k_m,$$

which can go to ∞ arbitrarily slowly. This proves the lemma.

We now turn to the proof of Theorem 4.

Recall that if $G \sim \text{Geo}_1(p)$ then its moment generating function is

$$\mathbb{E} e^{\lambda G} = \frac{p e^\lambda}{1 - (1-p) e^\lambda} = \left(1 - \frac{1 - e^{-\lambda}}{p} \right)^{-1},$$

whenever $e^\lambda (1-p) < 1$, i.e. $\lambda < -\log(1-p)$. Therefore, for any $t > 0$, we have

$$\mathbb{P}(S_n > t) = \mathbb{P}(\exp(\lambda S_n) > e^{\lambda t}) \le e^{-\lambda t} \mathbb{E} e^{\lambda S_n} = e^{-\lambda t} \prod_{k=1}^{n} \left(1 - \frac{1 - e^{-\lambda}}{q_k} \right)^{-1}.$$

Note that, for any $\beta > 1$, if $0 \le x \le (\beta - 1)/\beta$ then $1/(1-x) \le e^{\beta x}$. We will choose a suitable β in a moment. For now, note that if $\lambda > 0$ is such that $(1 - e^{-\lambda})/q_k < (\beta - 1)/\beta$ for each $k = 1, \dots, n$, we obtain

$$\mathbb{P}(S_n > t) \le \exp\left\{ -\lambda t + \beta(1 - e^{-\lambda}) \sum_{k=1}^{n} \frac{1}{q_k} \right\} \le \exp\left\{ -\lambda \left(t - \beta \sum_{k=1}^{n} \frac{1}{q_k} \right) \right\},$$

where in the last step we used the inequality $1 - e^{-\lambda} \le \lambda$ valid for any λ. Now, fix arbitrary $\varepsilon > 0$ and set $t = (1 + \varepsilon) \sum_{k=1}^{n} 1/q_k$. Then, choose $\beta > 1$ so that $\delta := 1 + \varepsilon - \beta > 0$.

Finally, set $\lambda = q_n(\beta - 1)/\beta$. Then all the earlier requirements are satisfied (note that since (q_k) is decreasing, we do have $1 - e^{-\lambda} \le q_k(\beta - 1)/\beta$ for every $k \le n$) and we obtain

$$\mathbb{P}(S_n > (1 + \varepsilon)\mathbb{E}S_n) \le \exp\left\{-\delta q_n \sum_{k=1}^{n} \frac{1}{q_k}\right\}.$$

But this means that

$$\mathbb{P}\left(\frac{S_n}{\mathbb{E}S_n} - 1 > \varepsilon\right) \le \exp\left\{-\delta q_n \sum_{k=1}^{n} \frac{1}{q_k}\right\}. \tag{21}$$

To bound the lower tail

$$\mathbb{P}\left(\frac{S_n}{\mathbb{E}S_n} - 1 < -\varepsilon\right),$$

we proceed in a similar fashion: namely, for any $\lambda > 0$, this last probability is equal to

$$\mathbb{P}(e^{-\lambda S_n} > e^{-\lambda(1-\varepsilon)ES_n}) \le e^{\lambda(1-\varepsilon)ES_n} \prod_{k=1}^{n} \mathbb{E}e^{-\lambda G_k}$$

$$= e^{\lambda(1-\varepsilon)ES_n} \prod_{k=1}^{n}\left(1 + \frac{e^{\lambda} - 1}{q_k}\right)^{-1}.$$

As before, we note that if $0 < \beta < 1$ then, for $0 \le x \le (1 - \beta)/\beta$, we have $1/(1 + x) \le e^{-\beta x}$. Therefore, the right-hand side above is further bounded by

$$\exp\left\{\lambda(1-\varepsilon)\mathbb{E}S_n - \beta(e^{\lambda} - 1)\sum_{k=1}^{n}\frac{1}{q_k}\right\} \le \exp\left\{\lambda(1-\varepsilon-\beta)\sum_{k=1}^{n}\frac{1}{q_k}\right\},$$

provided of course that, for a chosen β, λ satisfies $(e^{\lambda}-1)/q_k \le (1-\beta)/\beta$ for every $1 \le k \le n$. It is clear that, for a given β, we may choose $\lambda = cq_n$ for a suitable constant c dependent only on β. (In fact, we may choose $c = \log(1 + (1 - \beta)/\beta)$.) And now, as earlier, given $\varepsilon > 0$, it is enough to choose $\beta < 1$ so that $\delta := \beta - (1 - \varepsilon) > 0$ to obtain

$$\mathbb{P}\left(\frac{S_n}{\mathbb{E}S_n} - 1 < -\varepsilon\right) \le \exp\left\{-\delta q_n \sum_{k=1}^{n} \frac{1}{q_k}\right\}. \tag{22}$$

Clearly, (21), (22), and (20) imply the first part and the sufficiency of (7) for the almost-sure convergence of $S_n/\mathbb{E}S_n$ in the second part of our theorem.

To prove the necessity of (7), suppose that $S_n/\mathbb{E}S_n$ converges almost surely. We will show that (7) holds (which in turn will imply that the limit has to be 1). Pick any $\delta > 0$. Note that

$$\frac{G_n}{\mathbb{E}S_n} = \frac{S_n}{\mathbb{E}S_n} - \frac{\mathbb{E}S_{n-1}}{\mathbb{E}S_n}\frac{S_{n-1}}{\mathbb{E}S_{n-1}},$$

and that, by (20),

$$\frac{\mathbb{E}S_{n-1}}{\mathbb{E}S_n} = 1 - \frac{\mathbb{E}G_n}{\mathbb{E}S_n} = 1 - \frac{1}{q_n\sum_{k=1}^{n}q_k^{-1}} \to 1.$$

It follows that $G_n/\mathbb{E}S_n \to 0$ almost surely. Take any δ_1 such that $0 < \delta_1 < \delta$. Since the events $\{G_m > \delta_1 \mathbb{E}S_m\}$ are independent, it follows by the Borel–Cantelli lemma that

$$\sum_{m=1}^{\infty} \mathbb{P}(G_m \geq \delta_1 \mathbb{E}S_m) < \infty.$$

But, since G_m is geometric with parameter q_m, we obtain

$$
\begin{aligned}
\mathbb{P}(G_m \geq \delta_1 \mathbb{E}S_m) &\geq \mathbb{P}(G_m \geq \lceil \delta_1 \mathbb{E}S_n \rceil) \\
&= (1 - q_m)^{\lceil \delta_1 \mathbb{E}S_m \rceil - 1} \\
&\geq (1 - q_m)^{\delta_1 \mathbb{E}S_m} \\
&= ((1 - q_m)^{q_m^{-1}\delta_1\delta^{-1}})^{\delta q_m \sum_{k=1}^{m} q_k^{-1}}.
\end{aligned}
$$

Since $\delta_1/\delta < 1$, for all m sufficiently large

$$(1 - q_m)^{q_m^{-1}\delta_1\delta^{-1}} \geq e^{-1}.$$

Thus,

$$\sum_{m=1}^{\infty} \exp\left\{ -\delta q_m \sum_{k=1}^{m} \frac{1}{q_k} \right\} < \infty,$$

which proves (7).

3.5. Proof of Theorem 5

From [1, Theorem 1.5.11] we obtain

$$\mathbb{E}S_m = \sum_{k=1}^{m} \frac{1}{q_k} \sim \frac{m^{\alpha+1}}{L(m)} \quad \text{as } m \to \infty. \tag{23}$$

Together with (8) this shows that (7) is satisfied. By Theorem 4 and (23),

$$S_m \sim \frac{m^{\alpha+1}}{L(m)} \quad \text{as } m \to \infty,$$

almost surely.

For a function $f \colon \mathbb{R}_+ \to \mathbb{R}$, we define its 'inverse' $f^{\leftarrow}$ by

$$f^{\leftarrow}(y) := \inf\{x \in \mathbb{R} \colon f(x) \geq y\}.$$

Then $n \to Y_n$ is the inverse of $m \to S_m$ in this sense, and the statement of the theorem now follows with [1, Proposition 1.5.15].

References

[1] BINGHAM, N. H., GOLDIE, C. M. AND TEUGELS, J. L. (1987). *Regular Variation* (Encyclopedia Math. Appl. **27**). Cambridge University Press.

[2] BRÉMAUD, P. (1999). *Markov Chains* (Texts Appl. Math. **31**). Springer, New York.

[3] DENNERT, F. AND GRÜBEL, R. (2007). Renewals for exponentially increasing lifetimes, with an application to digital search trees. *Ann. Appl. Prob.* **17,** 676–687.

[4] DRMOTA, M. (2009). *Random Trees.* SpringerWienNewYork, Vienna.

[5] Evans, S. N., Grübel, R. and Wakolbinger, A. (2012). Trickle-down processes and their boundaries. *Electron J. Prob.* **17,** 58pp.

[6] Grübel, R. (2007). Distributional asymptotics in the analysis of algorithms: periodicities and discretization. In *AofA '07 Conf. on Analysis of Algorithms* (Discrete Math. Theoret. Comput. Sci. Proc. **AH**), pp. 451–462.

[7] Grübel, R. and Hitczenko, P. (2009). Gaps in discrete random samples. *J. Appl. Prob.* **46,** 1038–1051.

[8] Kallenberg, O. (1997). *Foundations of Modern Probability.* Springer, New York.

[9] Knuth, D. E. (1973). *The Art of Computer Programming*, Vol. 3. Addison-Wesley, Reading, MA.

[10] Mahmoud, H. M. (1992). *Evolution of Random Search Trees.* John Wiley, New York.

[11] Pitman, J. (2006). *Combinatorial Stochastic Processes* (Lecture Notes Math. **1875**). Springer, Berlin.