# ASYMPTOTIC NORMALITY FOR THE NUMBER OF RECORDS FROM GENERAL DISTRIBUTIONS

RAUL GOUET,* *Universidad de Chile*

F. JAVIER LÓPEZ ** *** AND

GERARDO SANZ,** **** *Universidad de Zaragoza*

## Abstract

We provide necessary and sufficient conditions for the asymptotic normality of $N_n$, the number of records among the first $n$ observations from a sequence of independent and identically distributed random variables, with general distribution $F$. In the case of normality we identify the centering and scaling sequences. Also, we characterize distributions for which the limit is not normal in terms of their discrete and continuous components.

*Keywords:* Extremes; records; central limit theorem

2010 Mathematics Subject Classification: Primary 60G70
Secondary 60F05

## 1. Introduction

Consider a sequence $\{X_n\}$ of independent and identically distributed (i.i.d.) random variables, whose common distribution function $F$ is general, in the sense that it has continuous and discrete components, restricted to have only finitely many atoms (discontinuities) on bounded intervals. A record is defined as an observation which is strictly greater than the current maximum. More precisely, for $n \geq 2$, $X_n$ is a record if $X_n > M_{n-1}$, where $M_n = \max\{X_1, \ldots, X_n\}$, and $X_1$ is always a record by convention.

Records are intrinsically interesting objects. Their mathematical theory has reached maturity and has not only found application in natural fields such as sports or meteorology but has also shown usefulness, for instance, in the mathematical analysis of data structures. See [1], [3], and [8] for general theory and applications.

There are three main random sequences related to records: the record values themselves, the indices of record observations, called record times, and the counting process of records. In this paper we focus our attention on the latter, giving necessary and sufficient conditions for its asymptotic normality. See [4] and [9] for applications related to record counts in computer data structures analysis.

Let $N_n$ be the number of records among $X_1, \ldots, X_n$ for $n \geq 1$; that is, $N_n = \sum_{k=1}^{n} I_k$, where $I_k = \mathbf{1}_{\{X_k > M_{k-1}\}}$ is the indicator of $X_k$ being a record. The asymptotic behavior of $N_n$ is well understood for distributions $F$ which are either continuous or lattice, but, to the authors'

knowledge, no result has been reported for general distributions, with infinitely many atoms. Of course, when the number of atoms is finite, the situation is either trivial or it reduces to the problem with continuous $F$, which is very well known. Indeed, Rényi [10] discovered that, in the latter case, the indicators $I_n$ are independent, with $\mathrm{E}[I_n] = 1/n$, $n \geq 1$. This structural result yields asymptotic normality for $N_n$, as well as other types of convergence, rather easily. The case of discrete distributions was not addressed until decades later, probably because of the lack of a result similar to Rényi's. Vervaat [12] started exploring records for integer-valued random variables. Later, Bai *et al.* [2] used a generating function approach to asymptotic normality, for observations from the geometric distribution. More recently, a martingale method was used in [6] to obtain a central limit theorem for $N_n$, applicable to a large range of discrete distributions. In the present paper we provide a complete answer to the question of the asymptotic normality of record counts from i.i.d. observations. As a consequence, for lattice distributions, we improve the results of [6] and solve a conjecture posed in [2].

A useful way to look at records is to consider $\xi(x)$, equal to the number of records in the whole sequence $\{X_n\}$ with values less than or equal to $x$. Shorrock [11] discovered that $\xi = \{\xi(x); x \geq 0\}$ is a point process on the line that has independent increments and decomposes as the sum of two independent processes $\xi_c$ and $\xi_d$. The continuous component $\xi_c$ is a nonhomogeneous Poisson process and the discrete component $\xi_d$ is a singular, independent increments process, concentrated on the set of atoms of $F$. This result applies to general distributions only restricted to have a finite number of atoms on compact sets. We will refer to $\xi$ as the Shorrock process.

Observe that, in terms of the asymptotic behavior of $N_n$, there is no loss of generality in restricting attention to positive random variables, because we deal with upper extremes. On the other hand, for discontinuous $F$, it is relevant to consider only the case of infinitely many atoms. Hence, in the sequel we work with distribution functions $F$ such that $F(x) = 0$ for all $x \leq 0$, whose set of atoms is $A = \{a_n; n \geq 1\}$, with $a_n \uparrow \infty$. Note that $a_n \uparrow \infty$ implies that $1 - F(x) > 0$ for all $x \geq 0$, which in turn implies the almost-sure divergence of $N_n$ as $n \to \infty$. For convenience, we define $a_0 = 0$, although 0 is not an atom.

We use the following notation throughout the paper. Sequences of real numbers $x_1, x_2, \dots$ are indexed by the positive integers, unless otherwise stated, and are denoted by $\{x_n, n \geq 1\}$ or $\{x_n\}$. Braces are sometimes omitted for simplicity. Convergence (divergence) of $x_n$ to a finite (infinite) limit $x$, as $n \to \infty$, is denoted by $\lim_n x_n = x$, or $x_n \to x$, or $x_n \uparrow x$ when $x_n$ is increasing. We write $x_n \sim y_n$ if either both sequences diverge to $\infty$ or converge to 0, with $\lim_n x_n/y_n = 1$, or both converge to nonzero, possibly different, finite limits. The $O(\cdot)$ and $o(\cdot)$ symbols have their usual meanings. Analogously, convergence (divergence) of a real function $g$, defined on $[0, \infty)$, to a finite (infinite) limit $y$, as $x \to \infty$, is denoted by $\lim_{x \to \infty} g(x) = y$, $g(x) \to y$, or $g(x) \uparrow y$ if $g$ is increasing.

Random variables are defined on a common probability space $(\Omega, \mathcal{F}, \mathrm{P})$; probability, expectation, and variance are respectively denoted by $\mathrm{P}[\cdot]$, $\mathrm{E}[\cdot]$, and $\mathrm{var}[\cdot]$. Convergence in distribution and in probability are respectively denoted by '$\xrightarrow{\mathrm{D}}$' and '$\xrightarrow{\mathrm{P}}$'. The Bernoulli and Poisson distributions, with parameters $p \in [0, 1]$ or $\lambda \geq 0$, are denoted by $\mathrm{Ber}(p)$ and $\mathrm{Po}(\lambda)$, respectively. The notation $X \hookrightarrow \mathrm{Ber}(p)$ and $X \hookrightarrow \mathrm{Po}(\lambda)$ mean that $X$ respectively has a $\mathrm{Ber}(p)$ or $\mathrm{Po}(\lambda)$ distribution. The normal distribution, with mean 0 and variance $\sigma^2$, is denoted by $N(0, \sigma^2)$. For a generic random variable $X$ with distribution $F$, let $F^-(x) = \mathrm{P}[X < x]$ for $x \in (0, \infty)$.

Also, let $m(t) = \sup\{x \geq 0: \mathrm{P}[X \geq x] \geq 1/t\}$ for $t \geq 1$, and observe that, for $\varepsilon > 0$,

$$1 - F^-(m(t)) \geq \frac{1}{t} \quad \text{and} \quad 1 - F^-(m(t) + \varepsilon) < \frac{1}{t}. \tag{1}$$

For $k \geq 1$, let $p_k = \mathrm{P}[X = a_k] = F(a_k) - F^-(a_k)$, $\mu_k = \mathrm{P}[X \in (a_{k-1}, a_k)] = F^-(a_k) - F(a_{k-1})$, $r_k = \mathrm{P}[X = a_k \mid X \geq a_k]$ (the hazard rate), and $l_k = \log((1 - F(a_{k-1}))/(1 - F^-(a_k)))$. Note that $r_k = \mathrm{P}[X = a_k]/\mathrm{P}[X \geq a_k] = p_k/(\sum_{i \geq k} p_i + \sum_{i > k} \mu_i)$. Finally, let

$$\theta(x) = \sum_{k=1}^{n} (r_k + l_k) + \log \frac{1 - F(a_n)}{1 - F(x)} \quad \text{and} \quad v(x) = \theta(x) - \sum_{k=1}^{n} r_k^2 \tag{2}$$

for $x \in [a_n, a_{n+1})$ and $n \geq 0$.

Our main result is Theorem 1, where we give necessary and sufficient conditions for the asymptotic normality of $N_n$ and, in that case, calculate explicit expressions of the centering and scaling sequences. The proof is based on a central limit theorem for the Shorrock process $\xi$, coupled with a random change of time scale.

At first, it was not obvious to conjecture the asymptotic behavior of $N_n$ for a general distribution $F$. One may think that perhaps the form of the continuous component of $F$ is irrelevant, since record counts from continuous distributions have the classical $\log n$ speed regardless of the distribution. Our results revealed a more complex situation of interaction between the tails of both components of $F$.

The paper is organized as follows. The main result, a corollary, and examples are presented in Section 2, followed by intermediate results and proofs in Section 3. In Section 4, distributions for which $N_n$ is not asymptotically normal are characterized in terms of their discrete and continuous components.

## 2. Main results and examples

**Theorem 1.** (a) *If* $\lim_{x \to \infty} v(x) = \infty$ *then*

$$\frac{N_n - \theta(m(n))}{\sqrt{v(m(n))}} \xrightarrow{\mathrm{D}} N(0, 1).$$

(b) *If* $\lim_{x \to \infty} v(x) < \infty$ *then* $N_n - \theta(m(n))$ *is tight and there exist no (deterministic) sequences* $b_n, c_n > 0$ *such that* $(N_n - b_n)/c_n \xrightarrow{\mathrm{D}} N(0, 1)$.

**Corollary 1.** *Let* $F$ *be concentrated on the positive integers. There exist sequences* $b_n, c_n > 0$ *such that* $(N_n - b_n)/c_n \xrightarrow{\mathrm{D}} N(0, 1)$ *if and only if* $\sum_{k=1}^{\infty}(1 - r_k) = \infty$, *in which case they can be taken as* $b_n = \sum_{k=1}^{m(n)} r_k$ *and* $c_n = \sqrt{\sum_{k=1}^{m(n)} r_k(1 - r_k)}$.

Observe that in Corollary 1 no conditions are imposed on the hazard rates $r_n$ so this result extends Theorem 1 of [6]. Also, Corollary 1 gives a positive answer to a question raised in [2, Section 1, p. 323].

In the first example below we consider a discrete distribution with hazard rates, having both 0 and 1 as limiting points, which can now be analyzed using Corollary 1. In the second example we study a mixture of the exponential and geometric distributions.

**Example 1.** Consider a discrete distribution on the positive integers, with hazard rates given by $r_{2k-1} = 1 - 1/(2k+1)$ and $r_{2k} = 1/2k$ for $k \geq 1$. The following approximations are easily obtained: $b_n = \sum_{k=1}^{m(n)} r_k = m(n)/2 + O(1)$ and $c_n^2 = \sum_{k=1}^{m(n)} r_k(1 - r_k) = \log m(n) + O(1)$. Then, from Corollary 1,

$$\frac{N_n - m(n)/2}{\sqrt{\log m(n)}} \xrightarrow{\mathrm{D}} N(0, 1).$$

A slightly improved presentation of the result above follows from the inequalities in (1). Note that $A_{m(n)} := -\log(1 - F^-(m(n))) = -\sum_{k=1}^{m(n)-1} \log(1 - r_k) \sim \frac{1}{2} m(n) \log m(n)$ as $n \to \infty$. Then, taking the logarithm in (1), we have $A_{m(n)} \leq \log n < A_{m(n)+1}$ and, also, $\log A_{m(n)} \leq \log \log n < \log A_{m(n)+1}$. From both inequalities we obtain (for large enough $n$)

$$\frac{1}{m(n)} \frac{A_{m(n)}}{\log A_{m(n)+1}} < \frac{1}{m(n)} \frac{\log n}{\log \log n} < \frac{1}{m(n)} \frac{A_{m(n)+1}}{\log A_{m(n)}}.$$

Finally, it is easy to see that both extremes above converge to $\frac{1}{2}$. Thus, $m(n) \sim 2 \log n / \log \log n$ as $n \to \infty$, and so

$$\frac{N_n - m(n)/2}{\sqrt{\log \log n}} \xrightarrow{\text{D}} N(0, 1).$$

**Example 2.** Consider the mixture $F = \alpha F_e + (1 - \alpha) F_g$, where $F_e$ and $F_g$ stand for the exponential and geometric distribution functions, with parameters $\lambda$ and $p$, respectively. That is, $F_e(x) = 1 - e^{-\lambda x}$ and $F_g(x) = 1 - q^{\lfloor x \rfloor}$, $x \geq 0$, with $\lambda > 0$ and $q = 1 - p \in (0, 1)$. Then, for $k \geq 1$, $p_k = (1 - \alpha) q^{k-1} p$ and $\sum_{i \geq k} p_i = (1 - \alpha) q^{k-1}$. Also, $\mu_k = \alpha \int_{k-1}^{k} \lambda e^{-\lambda y} \, dy$ and $\sum_{i \geq k} \mu_i = \alpha e^{-\lambda(k-1)}$. So, letting $\rho = e^{-\lambda}/q$,

$$r_k = \frac{(1-\alpha)q^{k-1}p}{(1-\alpha)q^{k-1} + \alpha e^{-\lambda k}} = \frac{p}{1 + \alpha q \rho^k/(1-\alpha)}. \tag{3}$$

On the other hand,

$$l_k = \log \frac{1 - F(a_{k-1})}{1 - F^-(a_k)} = \log \frac{(1-\alpha)q^{k-1} + \alpha e^{-\lambda(k-1)}}{(1-\alpha)q^{k-1} + \alpha e^{-\lambda k}} = \log \frac{1 + \alpha \rho^{k-1}/(1-\alpha)}{1 + \alpha q \rho^k/(1-\alpha)}. \tag{4}$$

For the asymptotic behavior, we consider the following possibilities of relative tail behavior of $F_e$ and $F_g$. First, if $\rho > 1$ then $r_k \sim p(1 - \alpha)\rho^{-k}/(\alpha q)$ and so, $\sum_{k=1}^{\infty} r_k < \infty$. Also, $l_k = \lambda + O(\rho^{-k})$. Thus, $\sum_{k=1}^{n}(r_k + l_k) = \lambda n + O(1)$ and, hence, $v(x) = \theta(x) + O(1) = \lambda x + O(1)$ for $x \to \infty$. From Theorem 1 we conclude that

$$\frac{N_n - \lambda m(n)}{\sqrt{\lambda m(n)}} \xrightarrow{\text{D}} N(0, 1). \tag{5}$$

We now estimate $m(n) = \sup\{x \geq 0 : 1 - F^-(x) \geq 1/n\}$, observing that $1 - F^- = \alpha(1 - F_e^-) + (1 - \alpha)(1 - F_g^-)$, with $1 - F_e^-(x) = 1 - F_e(x) = e^{-\lambda x}$ and $1 - F_g^-(x) = q^{\lceil x \rceil - 1}$. From (1) we have, for all $\varepsilon > 0$,

$$-\log(\alpha e^{-\lambda m(n)} + (1-\alpha)q^{\lceil m(n) \rceil - 1}) \leq \log n \leq -\log(\alpha e^{-\lambda(m(n)+\varepsilon)} + (1-\alpha)q^{\lceil m(n)+\varepsilon \rceil - 1}).$$

The left-hand side above can be written as

$$\lambda m(n) - \log(\alpha + (1-\alpha)\rho^{-m(n)}q^{\lceil m(n) \rceil - m(n) - 1}) = \lambda m(n) - \log \alpha + O(\rho^{-m(n)}).$$

As the same analysis can be carried out on the right-hand side, for all $\varepsilon > 0$, we finally obtain $m(n) = (1/\lambda) \log n + O(1)$. We can now replace the estimate of $m(n)$ in (5) to obtain, when $\rho > 1$,

$$\frac{N_n - \log n}{\sqrt{\log n}} \xrightarrow{\text{D}} N(0, 1).$$

Suppose now that $\rho < 1$, that is, the exponential tail is lighter than the geometric tail. From (3) and (4), we have $r_k = p + O(\rho^k)$ and $l_k = O(\rho^k)$. Thus, $\sum_{k=1}^{n} r_k = np + O(1)$ and $\sum_{k=1}^{n} l_k = O(1)$, and, hence, $\theta(x) = px + O(1)$ and $v(x) = pqx + O(1)$ as $x \to \infty$. It follows from Theorem 1 that

$$\frac{N_n - pm(n)}{\sqrt{pqm(n)}} \xrightarrow{\text{D}} N(0, 1). \tag{6}$$

The estimation of $m(n)$ can be calculated as above. We have $m(n) = -\log n / \log q + O(1)$, which can be substituted into (6) to obtain, when $\rho < 1$,

$$\frac{N_n + p \log n / \log q}{\sqrt{\log n}} \xrightarrow{\text{D}} N\left(0, -\frac{pq}{\log q}\right).$$

Finally, if $\rho = 1$ then $r_k = (1 - \alpha)p / (1 - \alpha p)$ and $l_k = -\log(1 - \alpha p)$, $k \geq 1$. Hence, $\theta(x) = C(\alpha, p)x + O(1)$ and $v(x) = D(\alpha, p)x + O(1)$, where $C(\alpha, p) = (1 - \alpha)p / (1 - \alpha p) - \log(1 - \alpha p)$ and $D(\alpha, p) = q / (1 - \alpha p) - \log(1 - \alpha p)$. Theorem 1 yields

$$\frac{N_n - C(\alpha, p)m(n)}{\sqrt{D(\alpha, p)m(n)}} \xrightarrow{\text{D}} N(0, 1).$$

For $m(n)$, we have, as before, $m(n) = -\log n / \log q + O(1)$, thus getting

$$\frac{N_n + C(\alpha, p) \log n / \log q}{\sqrt{\log n}} \xrightarrow{\text{D}} N\left(0, -\frac{D(\alpha, p)}{\log q}\right).$$

## 3. Intermediate results and proofs

The Shorrock process $\xi$ is an independent increment point process on $\mathbb{R}_+ = [0, \infty)$, which admits the representation $\xi = \xi_c + \xi_d$, where $\xi_c$ and $\xi_d$ are the continuous and discrete (independent) components.

The intensity measure of $\xi$ is given by the so-called hazard measure $H(dx) = F(dx)/(1 - F^-(x))$, which we consider in detail in order to develop the central limit theorem for $N_n$. Note that $\mathrm{E}[\xi(x)] = \int_{[0,x]} H(dy)$ for all $x \geq 0$. On the other hand, the continuous component $\xi_c$ has intensity measure $\Lambda_c$ given by

$$\Lambda_c(-\infty, x] = -\log(1 - F(x)) + \sum_{j=1}^{k} \log(1 - r_j), \qquad x \in [a_k, a_{k+1}),$$

while the discrete component is the Bernoulli process $\xi_d(x) = \sum_{j \geq 1} Z_j \mathbf{1}_{\{a_j \leq x\}}$, where the random variables $Z_j$ are independent $\mathrm{Ber}(r_j)$.

The following proposition is an easy consequence of Shorrock's results.

**Proposition 1.** *Let $\xi$ be the Shorrock process, and let $\theta(x)$ and $v(x)$ be as defined in (2). Then, for $x > 0$,*

(a) *$\xi(x) = \sum_{k=1}^{n}(B_k + P_k) + P'_x$ for $x \in [a_n, a_{n+1})$, where $B_k \hookrightarrow \mathrm{Ber}(r_k)$, $P_k \hookrightarrow \mathrm{Po}(l_k)$, $k \geq 1$, and $P'_x \hookrightarrow \mathrm{Po}(\log((1 - F(a_n))/(1 - F(x))))$ are independent,*

(b) *$\mathrm{E}[\xi(x)] = \theta(x)$,*

(c) *$\mathrm{var}[\xi(x)] = v(x)$,*

(d) $\lim_{x\to\infty} \theta(x) = \infty$,

(e) $\lim_{x\to\infty} v(x) < \infty$ *if and only if*

$$\sum_{k=1}^{\infty} r_k(1 - r_k) < \infty \quad and \quad \sum_{k=1}^{\infty} \frac{\mu_k}{1 - F^-(a_k)} < \infty.$$

*Proof.* Parts (a), (b), and (c) follow from Theorem 1 of [11]. For part (d), note that $\lim_{x\to\infty} \theta(x) < \infty$ would imply that $\sum_{k=1}^{\infty} P[B_k + P_k > 0] < \infty$ so, by the Borel–Cantelli lemma and (a), $\lim_{x\to\infty} \xi(x) < \infty$ almost surely (a.s.); that is, there would be a finite number of records in the sequence $\{X_n\}$, which is a contradiction since $F(x) < 1$ for all $x > 0$. Part (e) is immediate upon noting that $l_k = \log(1 + \mu_k/(1 - F^-(a_k)))$ and that $\sum_{k=1}^{\infty} l_k < \infty$ if and only if $\sum_{k=1}^{\infty} \mu_k/(1 - F^-(a_k)) < \infty$. $\square$

Our proof of Theorem 1 is based on a central limit theorem for $\xi(x)$. To that end, the representation of $\xi(x)$ in Proposition 1(a), as a sum of independent random variables, is very well suited. However, a drawback is the lack of control over the parameters $l_k$ which makes Lyapunov's condition arduous to check. We overcome this difficulty by enriching the sequence of atoms with extra points which allow the $l_k$s to be split into manageable pieces.

Let $x_0 = 0$, and, for $n \geq 1$, let

$$x_n = \inf\left\{x > x_{n-1} : x \in A \text{ or } \phi_{n-1}(x) > \tfrac{1}{2}\right\}, \tag{7}$$

where $\phi_n(x) = \log((1 - F(x_n))/(1 - F^-(x)))$ for $n \geq 0$. That is, the sequence $\{x_n\}$ is formed by all the atoms of $F$ together with continuity points selected in such a way that $\phi_n(x_{n+1})$ is bounded above (see Lemma 1 below). Note that $x_n$ is well defined because $a_n \uparrow \infty$ implies that the set over which the infimum is taken is nonempty. Some properties are collected in the following lemma.

**Lemma 1.** *Let $\{x_n\}$ be as defined in (7). Then the following statements hold.*

(a) $x_n > x_{n-1}$ *for all $n \geq 1$.*

(b) $\phi_{n-1}(x_n) \leq \tfrac{1}{2}$ *for all $n \geq 1$. Also, if $x_n \notin A$ then $\phi_{n-1}(x_n) = \tfrac{1}{2}$.*

(c) *There are a finite number of $x_n$ values in every bounded interval $(a, b)$.*

(d) $x_n \uparrow \infty$.

*Proof.* (a) By definition, $x_n \geq x_{n-1}$, so we prove that the strict inequality holds. Let $t_n = \min\{a \in A : a > x_{n-1}\}$ be the smallest atom greater than $x_{n-1}$.

As $\lim_{x\to x_{n-1}^+} \phi_{n-1}(x) = 0$ and $\phi_{n-1}(x)$ is continuous and nondecreasing in $(x_{n-1}, t_n)$, then $(x_{n-1}, t_n) \cap \{x : \phi_{n-1}(x) > \tfrac{1}{2}\}$ is either empty or an interval $(b, t_n)$, with $b > x_{n-1}$. In the former case $x_n = t_n > x_{n-1}$ and in the latter case, $x_n = b > x_{n-1}$.

(b) We consider two cases depending on whether $x_n \in A$ or not. If $x_n \notin A$, as in part (a), $\phi_{n-1}(x)$ is continuous in $(x_{n-1}, t_n)$ and $x_n = b \in (x_{n-1}, t_n)$, so $\phi_{n-1}(x_n) = \tfrac{1}{2}$. If $x_n \in A$ then $(x_{n-1}, t_n) \cap \{x : \phi_{n-1}(x) > \tfrac{1}{2}\} = \varnothing$ so $\phi_{n-1}(x) \leq \tfrac{1}{2}$ in $(x_{n-1}, t_n)$ and, since $\phi_{n-1}$ is left continuous, we have $\phi_{n-1}(x_n) \leq \tfrac{1}{2}$.

(c) It suffices to see that there are finitely many values $x_n$ in the interval $(a_k, a_{k+1})$ for all $k \geq 0$. Suppose on the contrary that, for some $n \geq 1$, $x_{n-1} = a_k$ and $x_{n+m} \in (a_k, a_{k+1})$ for

every $m \geq 0$. Then, by part (b),

$$
\begin{aligned}
\log \frac{1 - F(a_k)}{1 - F^-(a_{k+1})} &= \log \frac{1 - F(a_k)}{1 - F^-(x_n)} + \log \frac{1 - F^-(x_n)}{1 - F^-(a_{k+1})} \\
&= \frac{1}{2} + \log \frac{1 - F^-(x_n)}{1 - F^-(a_{k+1})} \\
&= \frac{m}{2} + \log \frac{1 - F^-(x_{n+m-1})}{1 - F^-(a_{k+1})} \\
&> \frac{m}{2}
\end{aligned}
$$

for every $m \geq 0$, so $\log((1 - F(a_k))/(1 - F^-(a_{k+1}))) = \infty$, which is a contradiction. Part (d) follows immediately from (a) and (c).

We introduce additional notation. For $k \geq 1$, let $s_k = \mathrm{P}[X = x_k]/\mathrm{P}[X \geq x_k]$ (that is, $s_k = r_n$ if $x_k = a_n \in A$ and $s_k = 0$ otherwise) and $\lambda_k = \phi_{k-1}(x_k) = \log((1 - F(x_{k-1}))/(1 - F^-(x_k)))$. For $x \in (0, \infty)$, let

$$
\pi(x) = \max\{k \colon x_k \leq x\} \quad \text{and} \quad \tau(x) = \min\{k \colon x_k \geq x\}, \tag{8}
$$

which are well defined by Lemma 1. Also, for $n \geq 0$, let

$$
\hat{\theta}(n) = \sum_{k=1}^{n}(s_k + \lambda_k) \quad \text{and} \quad \hat{v}(n) = \sum_{k=1}^{n}(s_k(1 - s_k) + \lambda_k). \tag{9}
$$

The following is a rewriting of Proposition 1 in terms of $x_k$, $s_k$, and $\lambda_k$.

**Corollary 2.** *Let $\xi$ be the Shorrock process, and let $\pi$, $\tau$, $\hat{\theta}$, and $\hat{v}$ be as defined in (8) and (9). Then, for $x > 0$,*

  (a) $\xi(x) = \sum_{k=1}^{\pi(x)}(C_k + Q_k) + Q'_x$, *where $C_k \hookrightarrow \mathrm{Ber}(s_k)$, $Q_k \hookrightarrow \mathrm{Po}(\lambda_k)$, $k \geq 1$, and $Q'_x \hookrightarrow \mathrm{Po}(\log((1 - F(x_{\pi(x)}))/(1 - F(x))))$ are independent,*

  (b) $\mathrm{E}[\xi(x)] = \theta(x) = \hat{\theta}(\pi(x)) + \log((1 - F(x_{\pi(x)}))/(1 - F(x)))$,

  (c) $\mathrm{var}[\xi(x)] = v(x) = \hat{v}(\pi(x)) + \log((1 - F(x_{\pi(x)}))/(1 - F(x)))$.

Here begins a series of technical results leading to the proof of Theorem 1. The idea is to establish a central limit theorem for $\hat{\xi}(n)$ (Lemma 2) and apply Kubacki and Szynal's result (Lemma 4) in order to replace $n$ by the random variable $M_n$. Then the random centering sequences are replaced by deterministic sequences and convergence for $\xi(M_n)$ is obtained thanks to the asymptotic equivalence of $\xi$ and $\hat{\xi}$. The final conclusion follows from the simple identity which connects the Shorrock process with $N_n$, namely $N_n = \xi(M_n)$.

**Lemma 2.** *Let $\hat{\xi}(n) = \sum_{k=1}^{n}(C_k + Q_k)$. If $\sum_{n=1}^{\infty}(s_k(1 - s_k) + \lambda_k) = \infty$ then*

$$
\frac{\hat{\xi}(n) - \hat{\theta}(n)}{\sqrt{\hat{v}(n)}} \xrightarrow{\mathrm{D}} N(0, 1).
$$

*Proof.* Since the random variables $Z_k = C_k + Q_k - (s_k + \lambda_k)$ are independent, with $E[Z_k] = 0$ and $\text{var}[\sum_{k=1}^n Z_k] = \hat{v}(n)$, it suffices to check Lyapunov's condition

$$\sum_{k=1}^n \frac{E[|Z_k|^3]}{(\hat{v}(n))^{3/2}} \to 0.$$

Since $\lambda_k \leq \frac{1}{2}$, simple calculations yield $E[|Z_k|^3] \leq 2(s_k(1-s_k) + \lambda_k)$ and the result follows.

**Lemma 3.** *Let $0 \leq x \leq y$. Then*

$$\frac{1 - F^-(x)}{1 - F^-(y)} \geq \exp\{\hat{\theta}(\pi(y) - 1) - \hat{\theta}(\tau(x))\}. \tag{10}$$

*Proof.* Inequality (10) trivially holds if $\tau(x) \geq \pi(y) - 1$, so assume that $\tau(x) < \pi(y) - 1$. Write, for $k \geq 1$,

$$1 - F^-(x_k) = \frac{1 - F^-(x_k)}{1 - F(x_{k-1})} \prod_{i=1}^{k-1} \frac{1 - F(x_i)}{1 - F^-(x_i)} \frac{1 - F^-(x_i)}{1 - F(x_{i-1})} = e^{-\lambda_k} \prod_{i=1}^{k-1} (1 - s_i) e^{-\lambda_i}.$$

Then, for $x \in (0, \infty)$,

$$\prod_{i=1}^{\tau(x)} (1 - s_i) e^{-\lambda_i} \leq 1 - F^-(x_{\tau(x)}) \leq 1 - F^-(x) \leq 1 - F^-(x_{\pi(x)}) \leq \prod_{i=1}^{\pi(x)-1} (1 - s_i) e^{-\lambda_i}.$$

Therefore, recalling that $1 - t \leq e^{-t}$ for any $t \geq 0$, we have

$$\frac{1 - F^-(x)}{1 - F^-(y)} \geq \prod_{i=\tau(x)+1}^{\pi(y)-1} \frac{1}{1 - s_i} e^{\lambda_i} \geq \prod_{i=\tau(x)+1}^{\pi(y)-1} e^{s_i + \lambda_i} = e^{\hat{\theta}(\pi(y)-1) - \hat{\theta}(\tau(x))}.$$

Note that, for $\theta$ given in (2), its generalized inverse $\theta^{\leftarrow}(t) = \inf\{x : \theta(x) \geq t\}$ for $t \in (0, \infty)$ is well defined since $\lim_{x \to \infty} \theta(x) = \infty$ by Proposition 1(d). The function $\theta^{\leftarrow}$ has the following properties, whose proofs are immediate since $\theta$ is nondecreasing, right continuous, and $\theta(x) - \theta^-(x) < 1$ for all $x > 0$: (i) $\theta(a) \geq b$ if and only if $a \geq \theta^{\leftarrow}(b)$, (ii) $\theta(\theta^{\leftarrow}(t)) \in [t, t+1)$.

**Proposition 2.** *The sequence $\theta(M_n) - \theta(m(n))$ is tight.*

*Proof.* We must prove that $P[|\theta(M_n) - \theta(m(n))| \geq \varepsilon c_n] \to 0$ for any deterministic sequence $c_n \uparrow \infty$ and every $\varepsilon > 0$, which is equivalent to $P[M_n \geq \theta^{\leftarrow}(\theta(m(n)) + \varepsilon c_n)] \to 0$ and $P[M_n \geq \theta^{\leftarrow}(\theta(m(n)) - \varepsilon c_n)] \to 1$, also respectively equivalent (see [5, Proposition 3.1.1]) to

$$n\,P[X \geq \theta^{\leftarrow}(\theta(m(n)) + \varepsilon c_n)] \to 0 \tag{11}$$

and

$$n\,P[X \geq \theta^{\leftarrow}(\theta(m(n)) - \varepsilon c_n)] \to \infty. \tag{12}$$

For (11), let $\delta_n = (x_{\tau(m(n))+1} - m(n))/2$ for $n \geq 1$. Note that $\delta_n > 0$. By the properties of $m(n)$ displayed in (1) and Lemma 3 we obtain, for large enough $n$,

$$n\,P[X \geq \theta^{\leftarrow}(\theta(m(n)) + \varepsilon c_n)]$$
$$< \frac{P[X \geq \theta^{\leftarrow}(\theta(m(n)) + \varepsilon c_n)]}{P[X \geq m(n) + \delta_n]}$$
$$\leq \exp\{-(\hat{\theta}(\pi(\theta^{\leftarrow}(\theta(m(n)) + \varepsilon c_n)) - 1) - \hat{\theta}(\tau(m(n) + \delta_n)))\}.$$

Thus, we must show that $\hat{\theta}(\pi(\theta^{\leftarrow}(\theta(m(n)) + \varepsilon c_n)) - 1) - \hat{\theta}(\tau(m(n) + \delta_n)) \to \infty$, which follows by noting that $0 < \hat{\theta}(n) - \hat{\theta}(n-1) < 2$, $0 \leq \tau(m(n) + \delta_n) - \tau(m(n)) \leq 1$ for $n \geq 1$, $0 \leq \theta(x) - \hat{\theta}(\pi(x)) < 1$, $0 \leq \hat{\theta}(\tau(x)) - \theta(x) < 1$, and $\theta(\theta^{\leftarrow}(x)) \in [x, x+1)$ for $x > 0$.

The proof of (12) is similar, noting that

$$
n\,\mathrm{P}[X \geq \theta^{\leftarrow}(\theta(m(n)) - \varepsilon c_n)] \geq \frac{\mathrm{P}[X \geq \theta^{\leftarrow}(\theta(m(n)) - \varepsilon c_n)]}{\mathrm{P}[X \geq m(n)]}
$$
$$
\geq \exp\{\hat{\theta}(\pi(m(n)) - 1) - \hat{\theta}(\tau(\theta^{\leftarrow}(\theta(m(n)) - \varepsilon c_n)))\}.
$$

**Proposition 3.** *If* $\lim_{x \to \infty} v(x) = \infty$ *then*

$$
\frac{\hat{v}(\pi(M_n))}{\hat{v}(\pi(m(n)))} \xrightarrow{\mathrm{P}} 1, \qquad \frac{\hat{v}(\tau(M_n))}{\hat{v}(\tau(m(n)))} \xrightarrow{\mathrm{P}} 1. \tag{13}
$$

*Proof.* By (9),

$$
|\hat{v}(\pi(M_n)) - \hat{v}(\pi(m(n)))| = \sum_{k=\pi(M_n \wedge m(n))+1}^{\pi(M_n \vee m(n))} (s_k(1 - s_k) + \lambda_k)
$$
$$
\leq \sum_{k=\pi(M_n \wedge m(n))+1}^{\pi(M_n \vee m(n))} (s_k + \lambda_k)
$$
$$
= |\hat{\theta}(\pi(M_n)) - \hat{\theta}(\pi(m(n)))|.
$$

Then, observing that $0 \leq \theta(x) - \hat{\theta}(\pi(x)) < 1$ and $0 \leq v(x) - \hat{v}(\pi(x)) < 1$ for $x > 0$, and using Proposition 2, we have $|\hat{\theta}(\pi(M_n)) - \hat{\theta}(\pi(m(n)))|/\hat{v}(\pi(m(n))) \xrightarrow{\mathrm{P}} 0$. Also, since $0 \leq \hat{\theta}(\tau(x)) - \theta(x) < 1$ and $0 \leq \hat{v}(\tau(x)) - v(x) < 1$ for $x > 0$, the second limit in (13) is obtained analogously.

The result below, which follows from Theorem 1 of [7], is a useful generalization of Rényi's well-known central limit theorem for the sum of a random number of independent random variables.

**Lemma 4.** *Let* $\{X_n\}$ *be a sequence of centered and independent random variables, with* $\mathrm{var}[X_n] = \sigma_n^2 < \infty$ *for all* $n \geq 1$. *Let* $s_n^2 = \sum_{k=1}^{n} \sigma_k^2 \uparrow \infty$, *and suppose that* $\sum_{k=1}^{n} X_k/s_n \xrightarrow{\mathrm{D}} Y$ *for some random variable* $Y$. *Let* $\{T_n\}$ *and* $\{v_n\}$ *be two sequences of positive integer-valued random variables, with* $\{v_n\}$ *independent of* $\{X_n\}$, *such that* $v_n \xrightarrow{\mathrm{P}} \infty$ *and* $s_{T_n}/s_{v_n} \xrightarrow{\mathrm{P}} 1$. *Then* $\sum_{k=1}^{T_n} X_k/s_{v_n} \xrightarrow{\mathrm{D}} Y$.

*Proof of Theorem 1.* If $\lim_{x \to \infty} v(x) = \infty$, we conclude, from Lemma 2, Proposition 3, and Lemma 4, that

$$
\frac{\hat{\xi}(\pi(M_n)) - \hat{\theta}(\pi(M_n))}{\sqrt{\hat{v}(\pi(m(n)))}} \xrightarrow{\mathrm{D}} N(0, 1).
$$

Moreover, since $0 \leq \theta(x) - \hat{\theta}(\pi(x)) < 1$ for $x > 0$, the random centering above can be replaced by $\theta(m(n))$ by Proposition 2. Now, using $0 \leq v(x) - \hat{v}(\pi(x)) < 1$ for $x > 0$, we obtain

$$
\frac{\hat{\xi}(\pi(M_n)) - \theta(m(n))}{\sqrt{v(m(n))}} \xrightarrow{\mathrm{D}} N(0, 1). \tag{14}
$$

In a similar way it can be shown that

$$\frac{\hat{\hat{\xi}}(\tau(M_n)) - \theta(m(n))}{\sqrt{v(m(n))}} \xrightarrow{\mathrm{D}} N(0, 1). \tag{15}$$

By Corollary 2(a) we have $\hat{\hat{\xi}}(\pi(M_n)) \le \xi(M_n) \le \hat{\hat{\xi}}(\tau(M_n))$ a.s. Hence, from (14) and (15), we conclude that

$$\frac{\xi(M_n) - \theta(m(n))}{\sqrt{v(m(n))}} \xrightarrow{\mathrm{D}} N(0, 1).$$

The proof of part (a) of Theorem 1 is now complete, recalling that $N_n = \xi(M_n)$.

We now prove part (b). The random variables $Z_k = C_k + Q_k - (s_k + \lambda_k)$, $k \ge 1$, are centered and independent, with $\sum_{k=1}^{\infty} \mathrm{var}[Z_k] < \infty$. Therefore, $\sum_{k=1}^{\infty} Z_k < \infty$ a.s., which implies that $\lim_n(\hat{\hat{\xi}}(\pi(M_n)) - \hat{\theta}(\pi(M_n))) < \infty$ and $\lim_n(\hat{\hat{\xi}}(\tau(M_n)) - \hat{\theta}(\tau(M_n))) < \infty$ a.s. Consequently, for any sequence $c_n \uparrow \infty$, Proposition 2 yields

$$\frac{\hat{\hat{\xi}}(\pi(M_n)) - \theta(m(n))}{c_n} = \frac{\hat{\hat{\xi}}(\pi(M_n)) - \hat{\theta}(\pi(M_n)) + \hat{\theta}(\pi(M_n)) - \theta(m(n))}{c_n} \xrightarrow{\mathrm{P}} 0 \tag{16}$$

and

$$\frac{\hat{\hat{\xi}}(\tau(M_n)) - \theta(m(n))}{c_n} \xrightarrow{\mathrm{P}} 0. \tag{17}$$

From (16) and (17), tightness of $N_n - \theta(m(n))$ follows.

Finally, suppose that there exist sequences $b_n$, $c_n > 0$ such that $(N_n - b_n)/c_n \xrightarrow{\mathrm{D}} N(0, 1)$. Then, since the distribution of $N_n - b_n$ is lattice of span 1 for each $n \ge 1$, $(N_n - b_n)/c_n$ cannot converge to any continuous distribution unless $c_n \to \infty$. But because of the tightness of $N_n - \theta(m(n))$, $c_n \to \infty$ implies that

$$\frac{\theta(m(n)) - b_n}{c_n} = \frac{N_n - b_n}{c_n} - \frac{N_n - \theta(m(n))}{c_n} \xrightarrow{\mathrm{D}} N(0, 1),$$

which is a contradiction.

## 4. Finite variance

We give necessary and sufficient conditions for $\lim_{x \to \infty} v(x) < \infty$ in terms of the discrete and continuous components of $F$.

Let $F = F_{\mathrm{d}} + F_{\mathrm{c}}$, where $F_{\mathrm{d}}$ is purely discrete and $F_{\mathrm{c}}$ is continuous, with $F_{\mathrm{d}}(\infty) := \lim_{x \to \infty} F_{\mathrm{d}}(x) \le 1$ and $F_{\mathrm{c}}(\infty) := \lim_{x \to \infty} F_{\mathrm{c}}(x) < 1$. For $k \ge 1$, let

$$r_k^d = \frac{\mathrm{P}[X = a_k]}{F_{\mathrm{d}}(\infty) - F_{\mathrm{d}}^-(a_k)} = \frac{p_k}{\sum_{j \ge k} p_j}$$

be the hazard rate of the discrete part of $F$.

Recall that, by Proposition 1(e), $\lim_{t \to \infty} v(t) < \infty$ if and only if

$$\sum_{k=1}^{\infty} r_k(1 - r_k) < \infty, \tag{18}$$

$$\sum_{k=1}^{\infty} \frac{\mu_k}{1 - F^-(a_k)} < \infty. \tag{19}$$

Note that, since $p_k = r_k(1 - F^-(a_k))$, condition (19) can be written equivalently as

$$\sum_{k=1}^{\infty} \frac{\mu_k}{p_k} r_k < \infty. \tag{20}$$

We present next a simple but useful technical lemma, followed by a result showing that $r_k$ can be replaced by $r_k^d$ in conditions (18) and (20).

**Lemma 5.** *Suppose that* $\sum_{k=1}^{\infty} \mu_k/(1 - F^-(a_k)) < \infty$. *Then,*

$$\frac{\sum_{j>i} \mu_j}{1 - F(a_i)} \to 0, \tag{21}$$

$$\frac{1 - F^-(a_i)}{F_d(\infty) - F_d^-(a_i)} \to 1, \tag{22}$$

$$\frac{1 - F(a_i)}{F_d(\infty) - F_d(a_i)} \to 1. \tag{23}$$

*Proof.* For (21), note that

$$\frac{\sum_{j>i} \mu_j}{1 - F(a_i)} \leq \sum_{j>i} \frac{\mu_j}{1 - F^-(a_j)} \to 0.$$

Now, (22) follows from (21) and the identity

$$1 - F^-(a_i) = F_d(\infty) - F_d^-(a_i) + \sum_{j>i} \mu_j.$$

Analogously, (23) follows from (21) and

$$1 - F(a_i) = F_d(\infty) - F_d(a_i) + \sum_{j>i} \mu_j.$$

**Proposition 4.** *We have* $\lim_{x \to \infty} v(x) < \infty$ *if and only if*

$$\sum_{k=1}^{\infty} r_k^d (1 - r_k^d) < \infty, \tag{24}$$

$$\sum_{k=1}^{\infty} \frac{\mu_k}{p_k} r_k^d < \infty. \tag{25}$$

*Proof.* Suppose first that $\lim_{x \to \infty} v(x) < \infty$. Then, by (19) and Lemma 5, we have $1 - F^-(a_k) \sim F_d(\infty) - F_d^-(a_k)$ and $1 - F(a_k) \sim F_d(\infty) - F_d(a_k)$, which imply that $r_k \sim r_k^d$ and $1 - r_k \sim 1 - r_k^d$. Therefore, (18) implies (24) and (20) implies (25).

For the converse, as $r_k \leq r_k^d$, we have

$$\sum_{k=1}^{\infty} \frac{\mu_k}{1 - F^-(a_k)} = \sum_{k=1}^{\infty} \frac{\mu_k}{p_k} r_k \leq \sum_{k=1}^{\infty} \frac{\mu_k}{p_k} r_k^d < \infty,$$

so, by Lemma 5, we have $r_k \sim r_k^d$ and $1 - r_k \sim 1 - r_k^d$, and the conclusion follows.

In the next theorem we characterize distributions with $\lim_{x\to\infty} v(x) < \infty$. For that purpose, fix $\tilde{a} \in (0, a_1)$, let $B = \{k : r_k^d < \frac{1}{2}\}$, and define the random variable $\tilde{X}$ by

$$\tilde{X}(\omega) = \begin{cases} \tilde{a} & \text{if } X(\omega) \in \{a_k : k \in B\}, \\ X(\omega) & \text{otherwise.} \end{cases}$$

Note that $\tilde{X}$ is obtained from $X$ by removing all atoms with hazard rates $r_k^d < \frac{1}{2}$ and assigning their probabilities to $\tilde{a}$. That is, $P[\tilde{X} = a_k] = 0$ for all $k \in B$ and $P[\tilde{X} = \tilde{a}] = \sum_{k \in B} p_k$.

**Theorem 2.** *We have $\lim_{x\to\infty} v(x) < \infty$ if and only if any of the following conditions hold:*

(a) $r_k^d \to 1$,

$$\sum_{k=1}^{\infty} (1 - r_k^d) < \infty \tag{26}$$

*and*

$$\sum_{k=1}^{\infty} \frac{\mu_k}{p_k} < \infty, \tag{27}$$

(b) *the limit points of $\{r_k^d\}$ are 0 and 1,*

$$\sum_{k \in B} r_k^d < \infty, \tag{28}$$

*and $\tilde{X}$ satisfies condition (a).*

*Proof.* Suppose that $\lim_{x\to\infty} v(x) < \infty$, so that (18), (20), (24), and (25) hold, and let us prove that either (a) or (b) holds. If $r_k^d \to 1$ then (26) and (27) follow from (24) and (25).

Now suppose that $r_k^d$ does not converge to 1 (recall that $0 < r_k^d < 1$ for all $k \geq 1$). Since (24) holds, $r_k^d(1 - r_k^d) \to 0$, so the only possible subsequential limits of $r_k^d$ are 0 and 1. Now, 1 must be a subsequential limit, since otherwise, $r_k^d \to 0$ and (24) would imply that $\sum_{k=1}^{\infty} r_k^d < \infty$, which is a contradiction since every discrete random variable with infinitely many atoms has $\sum_{k=1}^{\infty} r_k^d = \infty$; 0 must also be a subsequential limit of $r_k^d$ because otherwise $r_k^d \to 1$. Moreover, the definition of $B$ and (24) imply that $\sum_{k \in B} r_k^d < 2 \sum_{k \in B} r_k^d(1 - r_k^d) < \infty$, so (28) holds.

We now prove that $\tilde{X}$ verifies condition (a).

Let $\tilde{F}$ be the distribution function of $\tilde{X}$, let $\tilde{F}_d$ be the discrete part of $\tilde{F}$, and let $\tilde{F}_d(\infty) = \lim_{x\to\infty} \tilde{F}_d(x)$. Let $\{\alpha_k\}$ be the sequence of atoms of $\tilde{X}$; that is, $\alpha_1 = \tilde{a}$ and, for $k \geq 2$, $\alpha_k = a_{n_k}$ with $n_k \in \{1, 2, \ldots\} \setminus B$, $n_k \uparrow \infty$. Observe that $F_d(\infty) = \tilde{F}_d(\infty)$ since the total mass due to atoms is the same for $X$ and $\tilde{X}$. Note also that 0 is not a limit point of $\{r_{n_k}^d\}$ because in this subsequence we are excluding atoms with discrete hazard rates less than $\frac{1}{2}$. Therefore, the fact that $\{r_k^d\}$ has 0 and 1 as unique limit points implies that $r_{n_k}^d \to 1$. Define, for $k \geq 2$,

$$\tilde{r}_k = \frac{P[X = \alpha_k]}{1 - \tilde{F}^-(\alpha_k)} = \frac{p_{n_k}}{1 - \tilde{F}^-(a_{n_k})}, \qquad \tilde{r}_k^d = \frac{P[X = \alpha_k]}{\tilde{F}_d(\infty) - \tilde{F}_d^-(\alpha_k)} = \frac{p_{n_k}}{\tilde{F}_d(\infty) - \tilde{F}_d^-(a_{n_k})}.$$

Let us first check that

$$\tilde{F}_d(\infty) - \tilde{F}_d^-(a_k) \sim F_d(\infty) - F_d^-(a_k). \tag{29}$$

To this end, consider

$$\frac{\sum_{j\geq k,\, j\in B} p_j}{F_{\mathrm{d}}(\infty) - F_{\mathrm{d}}^-(a_k)} \leq \sum_{j\geq k,\, j\in B} \frac{p_j}{F_{\mathrm{d}}(\infty) - F_{\mathrm{d}}^-(a_j)} = \sum_{j\geq k,\, j\in B} r_j^d \to 0, \tag{30}$$

where convergence to 0 above follows from $\sum_{k\in B} r_k^d < \infty$. Finally, the identity

$$F_{\mathrm{d}}(\infty) - F_{\mathrm{d}}^-(a_k) = \tilde{F}_{\mathrm{d}}(\infty) - \tilde{F}_{\mathrm{d}}^-(a_k) + \sum_{j\geq k,\, j\in B} p_j$$

and (30) imply (29).

Now, from (29), for large $k$,

$$\tilde{r}_k^d = \frac{p_{n_k}}{\tilde{F}_{\mathrm{d}}(\infty) - \tilde{F}_{\mathrm{d}}^-(a_{n_k})} \sim \frac{p_{n_k}}{F_{\mathrm{d}}(\infty) - F_{\mathrm{d}}^-(a_{n_k})} = r_{n_k}^d \to 1,$$

so $\tilde{r}_k^d \to 1$.

Let us see that $\tilde{X}$ verifies (26). By (29) we have, for large $k$,

$$1 - \tilde{r}_k^d = \frac{\tilde{F}_{\mathrm{d}}(\infty) - \tilde{F}_{\mathrm{d}}(a_{n_k})}{\tilde{F}_{\mathrm{d}}(\infty) - \tilde{F}_{\mathrm{d}}^-(a_{n_k})} \sim \frac{F_{\mathrm{d}}(\infty) - F_{\mathrm{d}}(a_{n_k})}{F_{\mathrm{d}}(\infty) - F_{\mathrm{d}}^-(a_{n_k})} = 1 - r_{n_k}^d,$$

so $\sum_{k=1}^{\infty}(1 - \tilde{r}_k^d) < \infty$ since

$$\sum_{k=2}^{\infty}(1 - r_{n_k}^d) = \sum_{k\notin B}(1 - r_k^d) \leq 2\sum_{k\notin B} r_k^d(1 - r_k^d) < \infty.$$

Let us prove (27) for $\tilde{X}$, which is clearly equivalent to

$$\sum_{k=3}^{\infty} \frac{\tilde{F}^-(\alpha_k) - \tilde{F}(\alpha_{k-1})}{\tilde{F}_{\mathrm{d}}(\infty) - \tilde{F}_{\mathrm{d}}^-(\alpha_k)} < \infty, \tag{31}$$

since $\tilde{r}_k^d \to 1$.

Observing that, for $k \geq 3$, $\tilde{F}^-(\alpha_k) - \tilde{F}(\alpha_{k-1}) = \sum_{l=n_{k-1}+1}^{n_k}(F^-(a_l) - F(a_{l-1}))$ and that

$$\tilde{F}_{\mathrm{d}}(\infty) - \tilde{F}_{\mathrm{d}}^-(\alpha_k) \sim F_{\mathrm{d}}(\infty) - F_{\mathrm{d}}^-(\alpha_k) \sim 1 - F^-(\alpha_k),$$

where the first equivalence above follows from (29) and the second from Lemma 5, (31) is equivalent to

$$\sum_{k=3}^{\infty} \sum_{l=n_{k-1}+1}^{n_k} \frac{F^-(a_l) - F(a_{l-1})}{1 - F^-(a_{n_k})} < \infty. \tag{32}$$

Consider the identity

$$1 - F^-(a_{n_{k-1}+1}) = 1 - F^-(a_{n_k}) + \sum_{l=n_{k-1}+1}^{n_k-1} p_l + \sum_{l=n_{k-1}+2}^{n_k} \mu_l$$

and

$$\frac{\sum_{l=n_{k-1}+1}^{n_k-1} p_l}{1 - F^-(a_{n_{k-1}+1})} \le \frac{\sum_{l \ge n_{k-1}+1, \, l \in B} p_l}{1 - F^-(a_{n_{k-1}+1})} \sim \frac{\sum_{l \ge n_{k-1}+1, \, l \in B} p_l}{F_{\mathrm{d}}(\infty) - F_{\mathrm{d}}^-(a_{n_{k-1}+1})} \to 0,$$

$$\frac{\sum_{l=n_{k-1}+2}^{n_k-1} \mu_l}{1 - F^-(a_{n_{k-1}+1})} \le \frac{\sum_{l=n_{k-1}+2}^{\infty} \mu_l}{1 - F^-(a_{n_{k-1}+1})} \to 0,$$

where the equivalence follows from (22) and the limits from (30) and (21), respectively. Therefore, $1 - F^-(a_{n_{k-1}+1}) \sim 1 - F^-(a_{n_k})$ and

$$\sum_{l=n_{k-1}+1}^{n_k} \frac{F^-(a_l) - F(a_{l-1})}{1 - F^-(a_{n_k})} \sim \sum_{l=n_{k-1}+1}^{n_k} \frac{F^-(a_l) - F(a_{l-1})}{1 - F^-(a_l)},$$

so (32) follows from (19). This completes the proof of the necessity of (a) or (b).

Let us now show that (a) or (b) imply that $\lim_{x \to \infty} v(x) < \infty$. Suppose first that (a) holds; then

$$\sum_{k=1}^{\infty} \frac{\mu_k}{p_k} r_k < \sum_{k=1}^{\infty} \frac{\mu_k}{p_k} < \infty,$$

by (27), so we have (20). Moreover, applying Lemma 5, we have $r_k \sim r_k^d$ and $1 - r_k \sim 1 - r_k^d$, and (18) holds.

Suppose that (b) holds. Let us prove (18) and (19). For (19), note that

$$\begin{aligned}
\sum_{k=n_2+1}^{\infty} \frac{\mu_k}{1 - F^-(a_k)} &= \sum_{k=3}^{\infty} \sum_{l=n_{k-1}+1}^{n_k} \frac{F^-(a_l) - F(a_{l-1})}{1 - F^-(a_l)} \\
&\le \sum_{k=3}^{\infty} \sum_{l=n_{k-1}+1}^{n_k} \frac{F^-(a_l) - F(a_{l-1})}{1 - F^-(a_{n_k})} \\
&= \sum_{k=3}^{\infty} \frac{\tilde{F}^-(a_{n_k}) - \tilde{F}(a_{n_{k-1}})}{1 - F^-(a_{n_k})} \\
&\le \sum_{k=3}^{\infty} \frac{\tilde{F}^-(a_{n_k}) - \tilde{F}(a_{n_{k-1}})}{1 - \tilde{F}^-(a_{n_k})} \\
&= \sum_{k=3}^{\infty} \frac{\tilde{F}^-(\alpha_k) - \tilde{F}(\alpha_{k-1})}{\mathrm{P}[X = \alpha_k]} \tilde{r}_k \\
&< \sum_{k=3}^{\infty} \frac{\tilde{F}^-(\alpha_k) - \tilde{F}(\alpha_{k-1})}{\mathrm{P}[X = \alpha_k]} \\
&< \infty
\end{aligned}$$

by (27).

For (18), observe that, by Lemma 5, the convergence of the series above implies that $\sum_{k=1}^{\infty} r_k(1 - r_k) < \infty$ if and only if $\sum_{k=1}^{\infty} r_k^d(1 - r_k^d) < \infty$. Thus, we have to prove that $\sum_{k \in B} r_k^d(1 - r_k^d) < \infty$ and $\sum_{k \notin B} r_k^d(1 - r_k^d) < \infty$. The first convergence follows from $\sum_{k \in B} r_k^d(1 - r_k^d) < \sum_{k \in B} r_k^d < \infty$ by (28). Also, (28) implies (29) so $1 - r_{n_k}^d \sim 1 - \tilde{r}_k^d$;

therefore, $\sum_{k \notin B} r_k^d (1 - r_k^d) = \sum_{k=2}^{\infty} r_{n_k}^d (1 - r_{n_k}^d) < \sum_{k=2}^{\infty} (1 - r_{n_k}^d) < \infty$, since $\sum_{k=1}^{\infty} (1 - \tilde{r}_k^d) < \infty$.

**Remark 1.** Theorem 2 shows that distributions $F$ such that $\xi(x)$ has finite asymptotic variance are very light tailed in the sense that the discrete part has light tails (the discrete hazard rates go to 1 very fast) and, moreover, the tail of the continuous part is even lighter, since it is negligible when compared with the tail of the discrete part.

## Acknowledgements

## References

[1] ARNOLD, B. C., BALAKRISHNAN, N. AND NAGARAJA, H. N. (1998). *Records*. John Wiley, New York.
[2] BAI, Z.-D., HWANG, H.-K. AND LIANG, W.-Q. (1998). Normal approximations of the number of records in geometrically distributed random variables. *Random Structures Algorithms* **13**, 319–334.
[3] BUNGE, J. AND GOLDIE, C. M. (1999). Record sequences and their applications. In *Handbook of Statistics 19*, eds D. N. Shanbhag and C. R. Rao, North-Holland, Amsterdam, pp. 277–308.
[4] DEVROYE, L. (1988). Applications of the theory of records in the study of random trees. *Acta Informatica* **26**, 123–130.
[5] EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997). *Modelling Extremal Events*. Springer, Berlin.
[6] GOUET, R., LÓPEZ, F. J. AND SANZ, G. (2005). Central limit theorems for the number of records in discrete models. *Adv. Appl. Prob.* **37**, 781–800.
[7] KUBACKI, K. S. AND SZYNAL, D. (1988). On the rate of convergence in a random central limit theorem. *Prob. Math. Statist.* **9**, 95–103.
[8] NEVZOROV, V. B. (2001). *Records. Mathematical Theory* (Trans. Math. Monogr. **194**). American Mathematical Society, Providence, RI.
[9] PRODINGER, H. (1996). Combinatorics of geometrically distributed random variables: left-to-right maxima. *Discrete Math.* **153**, 253–270.
[10] RÉNYI, A. (1962). Théorie des éléments saillants d'une suite d'observations. *Ann. Fac. Sci. Univ. Clermont-Ferrand* **8**, 7–13.
[11] SHORROCK, R. W. (1972). On record values and record times. *J. Appl. Prob.* **9**, 316–326.
[12] VERVAAT, W. (1973). Limit theorems for records from discrete distributions. *Stoch. Process. Appl.* **1**, 317–334.