*The Annals of Applied Probability*
1994, Vol. 4, No. 4, 1206–1222

# GENERALISATIONS OF THE BIENAYMÉ–GALTON–WATSON BRANCHING PROCESS VIA ITS REPRESENTATION AS AN EMBEDDED RANDOM WALK

BY M. P. QUINE AND W. SZCZOTKA

*University of Sydney*

We define a stochastic process $\mathscr{X} = \{X_n,\ n = 0, 1, 2, \dots\}$ in terms of cumulative sums of the sequence $K_1, K_2, \dots$ of integer-valued random variables in such a way that if the $K_i$ are independent, identically distributed and nonnegative, then $\mathscr{X}$ is a Bienaymé–Galton–Watson branching process. By exploiting the fact that $\mathscr{X}$ is in a sense embedded in a random walk, we show that some standard branching process results hold in more general settings. We also prove a new type of limit result.

**1. Introduction.** Let $K_1, K_2, \dots$ be integer-valued random variables (rv's) defined on the same probability space $(\Omega, \mathscr{F}, P)$ and define sequences $\mathscr{X} = \{X_n,\ n = 0, 1, 2, \dots\}$ and $\mathscr{T} = \{T_n,\ n = 0, 1, 2, \dots\}$ by $X_0 = 1$, $X_1 = K_1$, $T_n = \sum_{j=0}^{n} X_j$, $n = 0, 1, 2, \dots$ and

$$X_{n+1} = \left( \sum_{j=T_{n-1}+1}^{T_{n-1}+X_n} K_j \right) I(X_n \geq 1), \qquad n = 1, 2, \dots, \tag{1.1}$$

where $I(A)$ denotes the indicator of $A$.

If $K_1, K_2, \dots$ are independent and identically distributed (iid) and nonnegative, then $\mathscr{X}$ is a Bienaymé–Galton–Watson (BGW) branching process [see, e.g., Heyde and Seneta (1977)] and $\mathscr{T}$ is the corresponding total progeny process (see Proposition 1). So the representation (1.1) allows us to extend the definition of the BGW process to the case where the offspring distribution is concentrated on all integers. In that case, that is, when $K_1, K_2, \dots$ are iid with $\mathbf{P}(K_1 < 0) > 0$, then $\mathscr{X}$ can be given a branching process interpretation in terms of a two type process with $X_n$ denoting the excess of type I over type II particles at the $n$th generation. In the case where $K_1, K_2, \dots$ are independent and nonnegative but not iid, then $\mathscr{X}$ can be regarded as a branching process with varying offspring distributions.

In this paper we investigate to what extent familiar BGW branching process results remain true when the conditions on the distribution of $\{K_1, K_2, \dots\}$ are relaxed. We focus here on three subjects, namely:

1. The probability of extinction (Theorems 2 and 3).

Received June 1993; revised March 1994.
*AMS* 1991 *subject classifications.* 60J80, 60J15, 60J10, 60J99.
*Key words and phrases.* Bienaymé–Galton–Watson branching process, embedded random walk, martingale.

2. The almost sure asymptotic behaviour of $\{X_n\}$ when the probability of extinction is less than 1 (Theorems 4–6).
3. The rate of asymptotic convergence (Theorem 7).

For BGW processes it is well known that $\mathbf{P}(X_n \to 0 \text{ or } X_n \to \infty) = 1$, $\mathbf{P}(X_n \to 0) = 1$ if $a =_{\mathrm{df}} \mathbf{E}K_1 < 1$ or if $a = 1$, $\mathbf{P}(K_1 = 1) < 1$ and $\mathbf{P}(X_n \to \infty) > 0$ if $a \in (1, \infty]$. Furthermore $W_n =_{\mathrm{df}} X_n/a^n$ converges almost surely (a.s.) to some random variable $W$, say, and if the variance $\sigma^2 =_{\mathrm{df}} \mathrm{var}(K_1) < \infty$, then $\mathbf{E}(W_n - W)^2 \to 0$ [see, e.g., Harris (1963)]. For our process $\mathscr{X}$ we show that similar results hold when the $K_i$ assume any integer values. In particular, assuming here for clarity that $K_1, K_2, \dots$ are iid, we have the same results as for the BGW process with these differences: (1) In the critical case ($a = 1$), $\mathbf{P}(X_n \to 0) = 1$ is implied by the finiteness of $\mathbf{E}(e^{-tK_1})$ for all $t > 0$, which trivially holds if $\mathbf{P}(K_1 \geq 0) = 1$ (Theorem 3); (2) when $a > 1$, for the a.s. convergence $W_n \to W$ when $\mathbf{P}(K_1 < 0) > 0$ we need the finiteness of $\sigma^2$, which seems to be a technical assumption [see Theorem 4; as a matter of fact, to prove Theorem 4 we need the boundedness of $\mathbf{E}(-\sum_{j=1}^n K_j)^+$ with respect to $n$, which is a consequence of the boundedness of $\{\sum_{j=1}^n \mathbf{E}K_j^-\}$, which trivially holds if $\mathbf{P}(K_1 \geq 0) = 1$]. Furthermore, we have obtained (Section 6, Theorem 7) some asymptotic results that seem to go beyond existing results for the BGW case, namely,

$$\frac{T_{n+1} - aT_n}{\sigma\sqrt{a^n}} \to_{\mathscr{D}} \mathscr{N}\sqrt{U'}, \qquad \frac{X_{n+1} - (a-1)T_n}{\sigma\sqrt{a^n}} \to_{\mathscr{D}} \mathscr{N}\sqrt{U'}$$

and

$$\left(\frac{T_{n+1} - aT_n}{\sigma\sqrt{T_n}} | U > 0\right) \to_{\mathscr{D}} \mathscr{N}, \qquad \left(\frac{X_{n+1} - (a-1)T_n}{\sigma\sqrt{T_n}} | U > 0\right) \to_{\mathscr{D}} \mathscr{N}$$

as $n \to \infty$, where $U = aW/(a-1)$, $U' =_{\mathscr{D}} U$, $\mathscr{N}$ has the standard normal distribution and $U'$ and $\mathscr{N}$ are mutually independent. These last results seem to be interesting from a practical point of view because they give predictions for $T_{n+1}$ or $X_{n+1}$ if we know $T_n$. For instance,

$$P\left(aT_n - z_{\alpha/2}\sigma\sqrt{T_n} \leq T_{n+1} \leq aT_n + z_{\alpha/2}\sigma\sqrt{T_n} | U > 0\right) \approx 1 - \alpha,$$

where $P(\mathscr{N} > z_{\alpha/2}) = \alpha/2$. All the above results are formulated in Theorems 2–7 in greater generality, dropping the assumption that the $K_i$'s be iid or even independent. Their formulations cover the BGW case with its standard assumptions. Of course dropping these assumptions means that the traditional methods for BGW processes based on the existence of a functional equation for probability generating functions are not easily applied (see Section 3.2). Our methods of getting the main results are based on our new representations (1.1), (2.3) and (2.4), which allow us to treat the process $\mathscr{X}$ on the event $\{T_n \to \infty\}$ as an embedded random walk and under some further conditions, treat the process $\{\max(W_n, 0), n = 0, 1, 2, \dots\}$ as a (sub)martingale.

In Section 2, we discuss the ramifications of the representation of $\mathscr{X}$ on $\{T_n \to \infty\}$ as an embedded random walk and indicate the restrictions on the $K_i$'s that make $\mathscr{X}$ a BGW process. Furthermore, we present there our main results. Sections 3–6 contain proofs and further discussion.

**2. Main results.** Here, in Theorems 2–7 we present the main results of the paper dealing with the process $\mathscr{X}$. Before that it would be helpful to have some preliminary facts. One of them is the following proposition, which is a consequence of the representation (1.1) and indicates the successive degrees of specialisation of our process leading to the BGW process. We discuss the proof of this proposition briefly in Section 3. We denote by $\mathscr{F}_n$ the $\sigma$-field generated by $\{X_0, \dots, X_n\}$, write $y^+ = \max(y, 0)$, $y^- = \max(-y, 0)$ and write $=_{\text{a.s.}}$ simply as $=$. We define the conditional variance $\operatorname{var}(X|\mathscr{F})$ as $\mathbf{E}(X^2|\mathscr{F}) - (\mathbf{E}(X|\mathscr{F}))^2$.

PROPOSITION 1. (*a*) *If the rv's* $K_1, K_2, \dots$ *are such that*

$$\mathbf{E}(K_{n+1}|K_1, K_2, \dots, K_n) = a, \qquad n = 1, 2, \dots,$$

*then*

$$\mathbf{E}(X_{n+1}|\mathscr{F}_n) = aX_n^+, \qquad n = 0, 1, \dots. \tag{2.1}$$

*If in addition*

$$\operatorname{var}(K_{n+1}|K_1, K_2, \dots, K_n) = \sigma^2, \qquad n = 1, 2, \dots,$$

*then*

$$\operatorname{var}(X_{n+1}|\mathscr{F}_n) = \sigma^2 X_n^+, \qquad n = 0, 1, \dots. \tag{2.2}$$

(*b*) *If the rv's* $K_1, K_2, \dots$ *are independent, then* $\{(X_n, T_{n-1}),\ n \geq 1\}$ *is a Markov chain.*

(*c*) *If* $K_1, K_2, \dots$ *are iid, then* $\mathscr{X}$ *is a homogeneous Markov chain.*

(*d*) *If* $K_1, K_2, \dots$ *are iid and nonnegative, then* $\mathscr{X}$ *is a BGW process.*

Other valuable observations are the following representations of $\mathscr{X}$ and $\mathscr{T}$. For the process $\mathscr{T}$, on the event $\{X_m > 0\}$ the indicator function in (1.1) equals unity for $n = 1, 2, \dots, m$, so

$$T_{n+1} = 1 + \sum_{j=1}^{T_n} K_j, \qquad 1 \leq n \leq m, \tag{2.3}$$

and

$$X_{n+1} = T_{n+1} - T_n = 1 + \sum_{j=1}^{T_n} \tilde{K}_j, \qquad 1 \leq n \leq m, \tag{2.4}$$

where $\tilde{K}_j = K_j - 1$ [note, however, that (2.3) fails if $X_n < 0$ for some $n$]. It follows that on the event $\{T_n \to \infty\} = \bigcap_{m \geq 1} \{X_m > 0\}$, $\mathscr{X}$ is embedded in the random walk $\tilde{\mathscr{S}} = \{\tilde{S}_n,\ n = 0, 1, 2, \dots\}$, where $\tilde{S}_0 = 0$ and $\tilde{S}_n = \sum_{j=1}^n \tilde{K}_j$,

$n = 1, 2, \dots,$ in the sense that

$$X_{n+1} = 1 + \tilde{S}_{T_n}, \qquad n \geq 1.$$

A similar fact was observed by Harris (1952) and used by Lindvall (1976) to obtain an estimate of the right tail of the distribution of the maximum generation size in a critical BGW process.

The representations (2.3) and (2.4) of $\mathscr{T}$ and $\mathscr{X}$ are crucial in our investigations and allow us to use methods of random walk in our treatment of these processes. Suppose for instance $(1/n)\sum_{i=1}^n K_i \to a > 1$ a.s. On the event $\{T_n \to \infty\}$, from (2.3) we have $T_{n+1}/T_n \to a$ and from (2.4) we have $X_{n+1}/T_n \to a - 1$. Furthermore, using the fact that $\log T_{n+1} = \sum_{i=1}^n \log(T_{i+1}/T_i) + \log T_1$, it follows that on $\{T_n \to \infty\}$,

$$\frac{1}{n} \log T_{n+1} \to \log a. \tag{2.5}$$

This leads to the convergence $\sqrt[n]{T_n} \to a$ on $\{T_n \to \infty\}$.

Representations (2.3) and (2.4) also give the relation

$$\frac{X_{n+1} - 1}{c_{n+1}} = \frac{(1/T_n)\tilde{S}_{T_n}}{(c_n/T_n)c_{n+1}/c_n} \quad \text{on } \{T_n \to \infty\}, \tag{2.6}$$

which leads to the following result: If $c_n^{-1}c_{n+1} \to c > 1$, then $X_n/c_n \to W$ and $T_n/c_n \to cW/(a-1)$ are equivalent on $\{T_n \to \infty\} \cap \{(1/n)\sum_{i=1}^n K_i \to a > 1\}$. This equivalence is also trivially true on $\{T_n \nrightarrow \infty\} \cap \{(1/n)\sum_{i=1}^n K_i \to a > 1\}$, so we have

$$\left\{\frac{X_n}{c_n} \to W\right\} \equiv \left\{\frac{T_n}{c_n} \to \frac{cW}{a-1}\right\} \quad \text{on } \left\{\frac{1}{n}\sum_{i=1}^n K_i \to a > 1\right\}. \tag{2.7}$$

One half of (2.7) (that convergence of $X_n/c_n$ implies that of $T_n/c_n$) has been proved and used in the BGW process context by Heyde [(1970), Theorem 3]. We use this equivalence in Corollary 5.

These results are trivial consequences of the representations (2.3) and (2.4) and are based on the sole assumption that the average of the $K_i$'s converges almost surely. By assuming more and using more sophisticated methods, we are able to prove stronger results. For instance, Corollary 5 gives a rate of convergence corresponding to (2.5) in the sense that it implies that $\log T_n - n \log a$ converges almost surely to a finite rv. This means the a.s. convergence of $T_n/a^n$ to a finite rv and of course the a.s. convergence to 0 of the sizes of the increments of that process. Theorem 7 gives the speed of this last convergence in the sense of weak convergence.

Now we state the main results dealing with the subjects 1–3 described in Section 1. Concerning subject 1, extinction means the event $\{X_n \to 0\} = \bigcup_{n \geq 1} \{X_n \leq 0\}$. The next two theorems, which contain the BGW criticality theorem, concern conditions under which the following results hold:

$$\mathbf{P}(X_n \to \infty \text{ or } X_n \to 0 \text{ as } n \to \infty) = 1, \tag{2.8}$$

$$\mathbf{P}(X_n \to 0) < 1 \tag{2.9}$$

and

$$\mathbf{P}(X_n \to 0) = 1. \tag{2.10}$$

THEOREM 2. *Suppose* $K_1, K_2, \ldots$ *are such that, as* $n \to \infty$,

$$\frac{\tilde{S}_n}{n} \to \lambda \quad a.s. \tag{2.11}$$

*for some* $\lambda \in [-\infty, \infty]$. *Then*:

(a) *If* $\lambda \neq 0$, *then* (2.8) *holds.*
(b) *If* $\lambda > 0$ *and* $\mathbf{P}(\tilde{S}_n \geq 0,\ n = 1, 2, \ldots) > 0$, *then* (2.9) *holds.*
(c) *If* $\lambda < 0$, *then* (2.10) *holds.*

THEOREM 3. *If* $K_1, K_2, \ldots$ *are iid with* $\mathbf{E}(K_1) = 1$ [*so that* (2.11) *holds with* $\lambda = 0$] *and* $\mathbf{P}(K_1 = 1) < 1$, *then a sufficient condition for* (2.10) *to hold is that there exists a number* $C < \infty$ *such that*

$$-\mathbf{E}\big(K_1 + \cdots + K_j \,|\, K_1 + \cdots + K_j \leq 0\big) \leq C \quad \text{for all } j. \tag{2.12}$$

*Furthermore*, (2.12) *is implied by*

$$\phi(t) = \mathbf{E}e^{-K_1 t} < \infty \quad \text{for all } t > 0. \tag{2.13}$$

It is worth adding that without the assumption of Theorem 3 that the $K_i$'s are iid, it can happen that (2.11) holds with $\lambda = 0$, yet $X_n \to X$ a.s. for a proper rv $X$ or that $X_n \to \infty$ a.s. (see Examples 1 and 2, Section 3.2). What happens when (2.12) fails is an open question.

Concerning subject 2, we need to make the basic assumption that for some $a > 0$ the $K_i$'s are such that (2.1) holds. A sufficient condition for this is given in Proposition 1(a). Condition (2.1) allows us to exploit the (sub)martingale structure of the process $\{(W_n^+, \mathscr{F}_n)\}$ and prove (when $a > 1$) its almost sure convergence (Theorem 4) as well as its convergence in mean square (Theorem 6). We have the following results concerning subject 2.

THEOREM 4. *Assume* (2.1) *holds with* $a > 1$. *If* $K_1, K_2, \ldots$ *are such that*

$$\mathbf{E}\left(\left(-\sum_{j=r+1}^{r+s} K_j\right)^+ \,\middle|\, K_1, K_2, \ldots, K_r\right) < D < \infty \tag{2.14}$$

*for all* $r, s \geq 1$, *then*

$$W_n = \frac{X_n}{a^n} \to W \quad a.s.\ as\ n \to \infty, \tag{2.15}$$

*where* $W \geq 0$ *is an a.s. finite rv. Furthermore* (2.15) *is equivalent to* $W_n^+ = X_n^+/a^n \to W$ *a.s.*

From Theorem 4 and (2.7) we get the following corollary.

COROLLARY 5. *Under the conditions of Theorem* 4,

$$\frac{T_n}{a^n} \to \frac{aW}{a-1} \quad a.s. \tag{2.16}$$

*whenever* $(1/n)\sum_{i=1}^{n} K_i \to a$ *a.s.*

THEOREM 6. *Let the conditions of Theorem* 4 *hold and assume that*

$$\operatorname{var}(X_{n+1}|\mathscr{F}_n) = \sigma^2 X_n^+,$$

*where* $\sigma^2 \in (0,\infty)$. *Then*

$$\mathbf{E}(W_n^+ - W)^2 \to 0 \quad as\ n \to \infty. \tag{2.17}$$

*Furthermore*, $\mathbf{E}W \geq 1$ *and* $\operatorname{var}(W) > 0$.

In the last part of the paper, dealing with subject 3, we give the speed of the convergence $T_{n+1}/a^{n+1} - T_n/a^n \to 0$ a.s. Roughly speaking, we show that $a^{(n/2)+1}(T_{n+1}/a^{n+1} - T_n/a^n)$ converges in distribution to a mixture of the standard normal distribution and the limiting distribution of $\{T_n/a^n\}$. The conditions and formulation of this result involve the random functions

$$Z_n(t) = \frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{[nt]}(K_j - a), \qquad 0 \leq t < \infty,\ n = 1,2,\dots, \tag{2.18}$$

where $\sigma > 0$ and $a > 1$ are some constants and $\{\mathscr{W}(t),\ t \geq 0\}$ is a standard Wiener process.

THEOREM 7. *Assume that the following conditions hold*:

$$\begin{aligned}\mathbf{P}(Z_n(t_1) \in A_1,\dots,Z_n(t_k) \in A_k; B)&\\ \to \mathbf{P}(\mathscr{W}(t_1) \in A_1,\dots,\mathscr{W}(t_k) \in A_k)\mathbf{P}(B),&\end{aligned} \tag{2.19}$$

*for any* $B \in \mathscr{F}$ *and any sets* $A_i \in \mathscr{B}(R)$ *which are continuity sets of* $\mathscr{W}(t)$;

$$\{Z_n\}\ \text{is tight}; \tag{2.20}$$

$$U_n = \frac{T_n}{a^n} \to U \quad a.s. \tag{2.21}$$

*Then*

$$\frac{T_{n+1} - aT_n}{\sigma\sqrt{a^n}} \to_{\mathscr{D}} \mathscr{N}\sqrt{U'} \quad as\ n \to \infty, \tag{2.22}$$

*where* $\mathscr{N}$ *is a standard normal rv*, $U' =_{\mathscr{D}} U$ *and* $U'$ *and* $\mathscr{N}$ *are mutually independent.*

Hence as an immediate consequence of $X_{n+1} = T_{n+1} - T_n$ we get the following corollary.

COROLLARY 8. *Under the same conditions,*

$$\frac{X_{n+1} - (a-1)T_n}{\sigma\sqrt{a^n}} \to_{\mathscr{D}} \mathscr{N}\sqrt{U'} \quad \text{as } n \to \infty. \tag{2.23}$$

Roughly speaking, assumption (2.19) says that $Z_n$ and $U$ are asymptotically independent, which holds, for example, when the $K_i$'s are iid with mean $a$ and variance $\sigma^2 > 0$. An extension of Theorem 7 is given at the end of the paper where we drop (2.19).

All these theorems are proved and discussed in the following sections: Section 3 deals with Proposition 1 and Theorems 2 and 3; Section 4 derives moment results based on assumption (2.1); Section 5 contains proofs of Theorems 4 and 6 together with Corollary 5; Section 6 contains a proof of Theorem 7.

## 3. Probability of extinction.

3.1. *Proofs.* First we consider briefly Proposition 1. We use the notation $\mathbf{E}(X; B) = \mathbf{E}XI(B)$ (expectation on the event $B$). Obviously

$$\mathbf{E}(X_{n+1} | X_1 = r_1, X_2 = r_2, \dots, X_n = r_n) = a r_n^+ \tag{3.1}$$

for $r_n \le 0$, so we prove it for $r_n > 0$. Let $i_j = 1 + r_1 + \cdots + r_j$, $j = 1, \dots, n$, so that

$$\begin{aligned} B_n &= B_n(i_1, i_2, \dots, i_{n-1}, r_n) =_{\text{df}} \{T_1 = i_1, T_2 = i_2, \dots, T_{n-1} = i_{n-1}, X_n = r_n\} \\ &= \{X_1 = r_1, X_2 = r_2, \dots, X_{n-1} = r_{n-1}, X_n = r_n\}. \end{aligned}$$

Then

$$\mathbf{E}(X_{n+1}; X_1 = r_1, X_2 = r_2, \dots, X_n = r_n) = \mathbf{E}\left( \sum_{s=i_{n-1}+1}^{i_{n-1}+r_n} K_s; B_n \right). \tag{3.2}$$

Using (1.1) and (2.3) we get

$$\begin{aligned} B_n = \Bigg\{ & K_1 = i_1 - 1, \sum_{j=1}^{i_1} K_j = i_2 - 1, \\ & \sum_{j=1}^{i_2} K_j = i_3 - 1, \dots, \sum_{j=1}^{i_{n-2}} K_j = i_{n-1} - 1, \sum_{j=i_{n-2}+1}^{i_{n-1}} K_j = r_n \Bigg\}. \end{aligned}$$

It therefore follows from the assumption that for $i_{n-1} < s \le i_{n-1} + r_n$, $\mathbf{E}(K_s; B_n) = a \times P(B_n)$ and (3.1) now follows from (3.2). The second part of Proposition 1(a) can be proved similarly using conditional expectation properties such as, for $i_{n-1} < s_1 < s_2 \le i_{n-1} + r_n$,

$$\mathbf{E}\big(K_{s_1} K_{s_2} | B_n\big) = \mathbf{E}\big(K_{s_1} \mathbf{E}\big(K_{s_2} | B_n, K_{s_1}\big)\big) = \mathbf{E}\big(a K_{s_1} | B_n\big) = a^2.$$

Proofs of the other parts of the proposition are straightforward.

PROOF OF THEOREM 2. (a) It is clear that the only three possibilities are

$$\{X_n \to \infty, T_n \to \infty\}, \qquad \{X_n \nrightarrow \infty, T_n \to \infty\} \quad \text{or} \quad \{X_n \to 0, T_n \to T < \infty\}.$$

Taking any positive integers $b < c$, we have

$$\begin{aligned}\{T_i \to \infty; X_n \in [b, c] \text{ i.o.}\} &= \left\{T_i \to \infty; b - 1 \le \tilde{S}_{T_n} \le c - 1 \text{ i.o.}\right\} \\ &= \left\{T_i \to \infty; \frac{b-1}{T_n} \le \frac{\tilde{S}_{T_n}}{T_n} \le \frac{c-1}{T_n} \text{ i.o.}\right\}.\end{aligned}$$

Using (2.11), it follows that if $\lambda \neq 0$, then $\mathbf{P}(T_i \to \infty; X_n \in [b, c] \text{ i.o.}) = 0$. So (2.8) holds in this case.

(b) We have

$$\begin{aligned}\mathbf{P}(T_n \to \infty) &= \mathbf{P}\left(T_n \to \infty; \tilde{K}_1 \ge 0, \tilde{S}_{T_1} \ge 0, \tilde{S}_{T_2} \ge 0, \dots\right) \\ &\ge \mathbf{P}\left(\tilde{S}_n \ge 0, n = 1, 2, \dots\right) \\ &> 0\end{aligned}$$

by assumption.

(c) We have

$$\begin{aligned}\mathbf{P}(T_i \to \infty) &= \mathbf{P}(T_i \to \infty, X_n > 0 \text{ i.o.}) \\ &\le \mathbf{P}\left(T_i \to \infty, \frac{1 + \tilde{S}_{T_n}}{T_n} > 0 \text{ i.o.}\right) \\ &= 0.\end{aligned}$$

Thus (2.10) is true, which completes the proof of Theorem 2. □

Because (2.12) is trivially satisfied with $C = 0$ when $\mathbf{P}(K_1 \ge 0) = 1$ and $\mathbf{P}(K_1 = 1) < 1$, our main concern in proving Theorem 3 is with the case $\mathbf{P}(K_1 < 0) > 0$. We need two preliminary lemmas.

LEMMA 3.1. *Assume $K_1, K_2, \dots$ are iid and let*

$$p_i = \mathbf{P}\left(\sum_{j=1}^{i} K_j > 0\right), \qquad i = 1, 2, \dots.$$

*Then*

$$p_i \le 1 - (1 - p_1)^i, \qquad i = 1, 2, \dots. \tag{3.3}$$

PROOF. We have

$$\mathbf{P}\left(\sum_{j=1}^{i} K_j \le 0\right) \ge \mathbf{P}\left(\bigcap_{j=1}^{i} \{K_j \le 0\}\right) = (1 - p_1)^i.$$

□

LEMMA 3.2. *Under the conditions of Theorem* 3, $\mathscr{X}$ *is tight; that is, any increasing sequence* $\{n'\}$ *contains an increasing subsequence* $\{n''\}$ *such that*

$$\mathbf{P}(X_{n''} = i) \to \pi_i \quad \text{for } i = 0,1,2,\dots \tag{3.4}$$

*and* $\sum_{i=0}^{\infty} \pi_i = 1$.

PROOF. Condition (2.1) is satisfied with $a = 1$, so

$$\begin{aligned} \mathbf{E}X_n^+ &= \mathbf{E}X_n + \mathbf{E}X_n^- \\ &= \mathbf{E}X_{n-1}^+ + \mathbf{E}X_n^- \\ &= 1 + \mathbf{E}\sum_{i=1}^{n} X_i^- \end{aligned}$$

on iterating. Because $\sum_{i=1}^{n} X_i^-$ is nondecreasing in $n$ and nonnegative, we get

$$\mathbf{E}X_n^+ \le \lim_{m\to\infty} \mathbf{E}X_m^+ = 1 + \mathbf{E}\sum_{i=1}^{\infty} X_i^- . \tag{3.5}$$

However, from the definition of $\mathscr{X}$, this sum is just the absolute value of $X_n$ when it first visits $(-\infty, 0]$, if it ever does, and is 0 otherwise. So if we define the Markov time $N$ by

$$\{N = n\} = \{X_i > 0, 1 \le i \le n-1;\ X_n \le 0\}$$

for $n < \infty$ and $N = \infty$ on $\{X_i > 0,\ i = 1,2,\dots\}$, then writing $X_\infty = 0$,

$$\begin{aligned} \mathbf{E}\sum_{i=1}^{\infty} X_i^- &= \mathbf{E}(-X_N) \\ &= \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} \frac{\mathbf{E}\left(\sum_{j=1}^{i} K_j\right)^-}{\mathbf{P}\left(\sum_{j=1}^{i} K_j \le 0\right)} \mathbf{P}(X_{n-1} = i, N = n) \\ &\le \sup_{i\ge 1} \frac{\mathbf{E}\left(\sum_{j=1}^{i} K_j\right)^-}{\mathbf{P}\left(\sum_{j=1}^{i} K_j \le 0\right)} \\ &= \sup_{j\ge 1} \mathbf{E}\left(-(K_1 + \cdots + K_j) | K_1 + \cdots + K_j \le 0\right) \\ &\le C \end{aligned} \tag{3.6}$$

by assumption. Thus for any $A < \infty$, for all sufficiently large $n$ we have, using (3.5) and (3.6),

$$\begin{aligned} A\mathbf{P}(X_i \to \infty) &\le \mathbf{E}X_n I(X_i \to \infty) \\ &\le \mathbf{E}X_n^+ \\ &\le 1 + C. \end{aligned}$$

Because $A$ can be taken arbitrarily large, it follows that $\mathbf{P}(X_i \to \infty) = 0$, which concludes the proof of Lemma 3.2. □

PROOF OF THEOREM 3. Because $\mathbf{E}K_1 = 1$ and $\mathbf{P}(K_1 = 1) < 1$, we can use Lemma 3.1 to get

$$\begin{aligned} \mathbf{P}(X_{n+1} > 0) &= \sum_{i=1}^{\infty} p_i \mathbf{P}(X_n = i) \\ &\le \sum_{i=1}^{\infty} \big(1 - (1-p_1)^i\big) P(X_n = i) \\ &= \mathbf{P}(X_n > 0) - \sum_{i=1}^{\infty} (1-p_1)^i P(X_n = i). \end{aligned} \tag{3.7}$$

Since $\mathbf{P}(X_n > 0)$ is nonincreasing in $n$, by letting $n \to \infty$ through $n''$ in (3.4) we get, using dominated convergence (noting that our assumptions imply $0 < 1 -$
$p_1 < 1$),

$$1 - \pi_0 \le 1 - \pi_0 - \sum_{i=1}^{\infty} (1-p_1)^i \pi_i.$$

It follows that $\pi_i = 0$ for all $i \ge 1$, which together with Lemma 3.2 implies that the only possible limit distribution of $\mathscr{X}$ has mass 1 on zero, that is, (2.10) is true.

Finally we prove that (2.12) is implied by (2.13). Our assumptions guarantee the existence of a positive number $\tau$ at which $\phi$ attains a minimum value $\rho \in (0,1)$; see Bahadur [(1971), Lemma 2.3]. Let $F$ be the distribution function of $-K_1$ and put $dG(z) = \rho^{-1} e^{\tau z}\, dF(z)$. Then $G$ is a distribution function with zero mean and finite variance, which we denote $\sigma_G^2$. Let $\Gamma_n$ denote the $n$-fold convolution of $G$ with itself. Writing $S_n = -(K_1 + \cdots + K_n)$,

$$\begin{aligned} \mathbf{E}(S_n I(S_n \ge 0)) &= \int_{x_1 + \cdots + x_n \ge 0} (x_1 + \cdots + x_n)\, dF(x_1) \cdots dF(x_n) \\ &= \int_{x_1 + \cdots + x_n \ge 0} \rho^n (x_1 + \cdots + x_n) \\ &\qquad \times \exp(-\tau(x_1 + \cdots + x_n))\, dG(x_1) \cdots dG(x_n) \\ &= \int_0^{\infty} \rho^n t e^{-\tau t}\, d\Gamma_n(t) \\ &\le \frac{2}{\tau} \rho^n \int_0^{\infty} e^{-\tau t/2}\, d\Gamma_n(t) \\ &= \rho^n \int_0^{\infty} e^{-\tau t/2} (\Gamma_n(t) - \Gamma_n(0))\, dt \end{aligned} \tag{3.8}$$

on integration by parts. Now, using the Berry–Esseen theorem,

$$
\begin{aligned}
|\Gamma_n(t) - \Gamma_n(0)| &\le \left|\Gamma_n(t) - \Phi\left(\frac{t}{\sigma_G\sqrt{n}}\right)\right| \\
&\quad + |\Gamma_n(0) - \Phi(0)| + \left|\Phi\left(\frac{t}{\sigma_G\sqrt{n}}\right) - \Phi(0)\right| \qquad (3.9)\\
&\le \frac{2\gamma}{\sigma_G^3\sqrt{n}} + \frac{t}{\sigma_G\sqrt{2\pi n}},
\end{aligned}
$$

where $\gamma$ is the (finite) third absolute moment of $G$. It follows from (3.8) and (3.9) that

$$
\mathbf{E}(S_n I(S_n \ge 0)) = O\left(\frac{\rho^n}{\sqrt{n}}\right). \qquad (3.10)
$$

Theorem 1 of Bahadur and Ranga Rao [(1960); see also their (46)] implies that

$$
\mathbf{P}(S_n \ge 0) \sim \frac{\rho^n b_n}{\sqrt{2\pi n}},
$$

where $\liminf_{n\to\infty} b_n = \delta > 0$, which together with (3.10) gives $\mathbf{E}(S_n|S_n \ge 0) = O(1)$, which is equivalent to (2.12). □

3.2. *Discussion of results.* The iid case of Theorem 2 includes results (2.8)–(2.10) for BGW processes with $\mathbf{E}K_1 \ne 1$. In particular, Theorem 2(b) relates to any supercritical $(1 < \mathbf{E}K_1 \le \infty)$ BGW process; it is a standard random walk result that $\mathbf{P}(\tilde{S}_n \ge 0,\ n = 1,2,\dots) = \mathbf{P}(V = \infty) > 0$, where $V = \inf\{n\colon \tilde{S}_n < 0\}$, whenever $\mathbf{E}\tilde{K}_1^- < \mathbf{E}\tilde{K}_1^+ \le \infty$ [see Feller (1971), pages 396–397]. Jagers [(1992), Theorem 2] considered a process $\{X_n,\ n \ge 1\}$ that has the properties that $X_n \ge 0$ and that $X_n = 0$ implies $X_{n+1} = 0$. He gives a sufficient condition for (2.8) in terms of the "generation sizes" $X_n$; however, it is not clear how to compare this to our condition (a) in Theorem 2, which is given in terms of the "litter sizes" $K_j$.

When (2.9) holds for a BGW process, it is also known that $q = \mathbf{P}(X_n \to 0)$ is the unique root in $[0,1)$ of the equation $\mathbf{E}(q^{K_1}) = q$. It does not seem possible to give a precise result like this even in the general iid case, part of the reason being that the result $\mathbf{P}(\text{extinction}|X_1 = j) = (\mathbf{P}(\text{extinction}|X_1 = 1))^j$ is not generally true outside the BGW process context.

Theorem 3 involves much stronger assumptions than Theorem 2. We give two examples to show that (2.11) with $\lambda = 0$ does not imply (2.10) or even (2.8). That is, the "critical" $(\mathbf{E}K_1 = 1)$ BGW process results cannot be extended so widely.

EXAMPLE 1. Suppose $K_1, K_2, \dots$ are independent with $\mathbf{P}(K_j = 1) = 1 - 1/j^2$ and $\mathbf{P}(K_j = j) = 1/j^2$. Then the Borel–Cantelli lemma implies $\mathbf{P}(K_j = j$ i.o.$) = 0$. Thus $\tilde{S}_n/n \to 0$ a.s. and $X_n \to X$ a.s., where $\mathbf{P}(1 \le X < \infty) = 1$, that is, there exists a proper limiting stationary distribution.

EXAMPLE 2. Suppose $K_1, K_2, \dots$ are independent with $\mathbf{P}(K_j = 1) = 1 - 1/j$ and $\mathbf{P}(K_j = [\sqrt{j}]) = 1/j$. Then $\operatorname{var}(K_j) \sim 1$, $\mathbf{E}\tilde{S}_n \sim 2\sqrt{n}$, $\tilde{S}_n/n \to 0$ a.s. and

$$\sum_j \mathbf{P}\big(K_j = [\sqrt{j}]\big) = \sum_j 1/j = \infty.$$

Thus $\mathbf{P}(K_j = [\sqrt{j}]\text{ i.o.}) = 1$ and hence $\mathbf{P}(X_n \to \infty) = 1$.

**4. Moments.** In this section, we give some results about $W_n^+ = a^{-n}X_n^+$ under condition (2.1).

LEMMA 4.1. *Assume that condition* (2.1) *holds. Then* $\{(W_n^+, \mathscr{F}_n),\ n = 1, 2, \dots\}$ *is a submartingale. Furthermore, if* $K_j \geq 0$ *for all* $j \geq 1$, *then* $\{(W_n^+, \mathscr{F}_n),\ n = 1, 2, \dots\}$ *is a martingale and* $\mathbf{E}W_n^+ = 1$ *for each* $n \geq 1$.

PROOF. Using Jensen's inequality for conditional expectations and condition (2.1) we get

$$\mathbf{E}\big(X_{n+1}^+ | \mathscr{F}_n\big) \geq \big(\mathbf{E}\big(X_{n+1} | \mathscr{F}_n\big)\big)^+ = aX_n^+,$$

which gives the first assertion of Lemma 4.1. The second part follows immediately from assumption (2.1), the fact that $X_n = X_n^+$ in that case and the definition of $W_n^+$. □

LEMMA 4.2. *Assume that condition* (2.1) *holds with* $a > 1$ *and that* (2.14) *holds. Then* $\mathbf{E}W_n^+$ *is bounded above uniformly in* $n$.

PROOF. From assumption (2.1) it follows by iterating as in the proof of Lemma 3.2 that

$$\mathbf{E}X_{n+1}^+ = a^{n+1} + \sum_{i=1}^{n+1} a^{n+1-i} M_i, \tag{4.1}$$

where $M_i = \mathbf{E}X_i^-$. If (2.14) holds, then

$$\begin{aligned} M_{n+1} &= \sum_{i_1} \sum_{i_2} \cdots \sum_{i_{n-1}} \sum_{r_n} \mathbf{E}\left(\left(- \sum_{j=i_{n-1}+1}^{i_{n-1}+r_n} K_j\right)^+ ; B_n(i_1, i_2, \dots, i_{n-1}, r_n)\right) \\ &\leq \sum_{i_1} \sum_{i_2} \cdots \sum_{i_{n-1}} \sum_{r_n} DP(B_n(i_1, i_2, \dots, i_{n-1}, r_n)) \leq D, \end{aligned}$$

where $B_n(i_1, i_2, \dots, i_{n-1}, r_n)$ was defined in the proof of Proposition 1(a). Hence

$$\begin{aligned} \mathbf{E}W_{n+1}^+ &= \mathbf{E}\frac{X_{n+1}^+}{a^{n+1}} = 1 + \sum_{i=1}^{n+1} a^{-i} M_i \\ &\leq 1 + \frac{D}{a-1}. \end{aligned} \tag{4.2}$$

□

For the next lemma we use the decomposition

$$\operatorname{var}(X) = \mathbf{E}(\operatorname{var}(X|\mathscr{F})) + \operatorname{var}(\mathbf{E}(X|\mathscr{F})),$$

which is easily verified.

LEMMA 4.3. *If the conditions of Lemma* 4.2 *hold and in addition* (2.2) *holds, then* $\operatorname{var}(W_n^+)$ *is uniformly bounded in n. Furthermore, we have*

$$\operatorname{var}(W_n^+) \to \sigma^2 A_1 - 2A_2 - A_3 + \frac{\operatorname{var}(K_1^+)}{a^2} \quad \textit{as } n \to \infty, \tag{4.3}$$

*where*

$$A_1 = \sum_{i=1}^{\infty} \frac{1}{a^{2(1+i)}} \mathbf{E}X_i^+, \qquad A_2 = \sum_{i=1}^{\infty} \frac{1}{a^{2i}} \mathbf{E}X_{i+1}^+ \mathbf{E}X_{i+1}^-$$

*and*

$$A_3 = \sum_{i=1}^{\infty} \frac{1}{a^{2i}} \operatorname{var}(X_{i+1}^-).$$

PROOF. We have

$$\begin{aligned} \operatorname{var}(X_{n+1}) &= \mathbf{E}(\operatorname{var}(X_{n+1}|\mathscr{F}_n)) + \operatorname{var}(\mathbf{E}(X_{n+1}|\mathscr{F}_n)) \\ &= \mathbf{E}(X_n^+ \sigma^2) + \operatorname{var}(X_n^+ a) \\ &= \sigma^2 \mathbf{E}(X_n^+) + a^2 \operatorname{var}(X_n^+). \end{aligned} \tag{4.4}$$

It follows from (4.4) and the identity

$$\operatorname{var}(X_{n+1}) = \operatorname{var}(X_{n+1}^+) + \operatorname{var}(X_{n+1}^-) + 2\mathbf{E}X_{n+1}^+ \mathbf{E}X_{n+1}^-$$

that

$$\operatorname{var}(X_{n+1}^+) = \sigma^2 \mathbf{E}X_n^+ - 2\mathbf{E}X_{n+1}^+ \mathbf{E}X_{n+1}^- - \operatorname{var}(X_{n+1}^-) + a^2 \operatorname{var}(X_n^+).$$

Iterating this result gives

$$\operatorname{var}(X_{n+1}^+) = D_{n+1} + a^2 D_n + \cdots + a^{2(n-1)} D_2 + a^{2n} \operatorname{var}(K_1^+),$$

where

$$D_{n+1} = \sigma^2 \mathbf{E}X_n^+ - 2\mathbf{E}X_{n+1}^+ \mathbf{E}X_{n+1}^- - \operatorname{var}(X_{n+1}^-).$$

It therefore follows using the definition of $W_n^+$ that (4.3) holds. From Lemma 4.2, $\mathbf{E}X_i^+/a^i$ is bounded, so in (4.3), $A_1 < \infty$ and hence $\operatorname{var}(W_n^+)$ is uniformly bounded. □

REMARK. To get the boundedness of $\operatorname{var}(W_n^+)$, it is enough to assume the conditions of Lemma 4.2 and the inequality

$$\operatorname{var}(X_{n+1}|\mathscr{F}_n) \le \sigma^2 X_n^+, \qquad n = 0, 1, 2, \dots .$$

**5. Strong limit theory.** Various versions of Theorems 4 and 6 were proved for BGW processes in the 1940s; see Harris [(1963), Section 1.8.1]. A version of Corollary 5 for BGW processes appears in Heyde (1970).

PROOF OF THEOREM 4. If $K_j \geq 0$ for all $j \geq 1$, then (2.15) is a consequence of Lemma 4.1 and the fact that a nonnegative martingale with finite expectation is a.s. convergent. In other cases, we have $X_n - X_n^+ = -X_n^-$. Now $\mathbf{E}X_n^- \leq D$ so $X_n^-/a^n \to 0$ a.s. and hence the convergence (2.15) is equivalent to

$$W_n^+ = \frac{X_n^+}{a^n} \to W \quad \text{a.s.} \tag{5.1}$$

Now we have seen that $\{(W_n^+, \mathcal{F}_n)\}$ is a submartingale. According to the submartingale convergence theorem [see, e.g., Shiryayev (1984), page 476], $W_n^+$ converges a.s. if $\sup_n \mathbf{E}|W_n^+| < \infty$, which follows from (4.2). Hence (5.1) holds, which implies (2.15). □

REMARKS ABOUT THEOREM 4.

1. Condition (2.14) is trivially satisfied if $\mathbf{P}(K_i \geq 0) = 1$, $i = 1, 2, \ldots,$ and is also satisfied if

$$P\left( \sum_{j=r+1}^{r+s} (a - K_j) \geq c | K_1, K_2, \ldots, K_r \right) \leq \frac{s\sigma^2}{c^2} \tag{5.2}$$

for all $s, r, c$, because then

$$\begin{aligned} &\mathbf{E}\left( \left( -\sum_{j=r+1}^{r+s} K_j \right)^+ | K_1, \ldots, K_r \right) \\ &\quad = \sum_{i=1}^{\infty} P\left( \sum_{j=r+1}^{r+s} (a - K_j) \geq sa + i | K_1, K_2, \ldots, K_r \right) \\ &\quad \leq \sum_{i=1}^{\infty} \frac{s\sigma^2}{(sa + i)^2} \leq \frac{\sigma^2}{a}. \end{aligned}$$

2. If $\mathbf{E}(K_n | K_1, \ldots, K_{n-1}) = a$ and $\operatorname{var}(K_n | K_1, \ldots, K_{n-1}) = \sigma_n^2$, then

$$\operatorname{var}\left( \sum_{j=r+1}^{r+s} K_j | K_1, K_2, \ldots, K_{r-1} \right) = \sum_{j=r+1}^{r+s} \sigma_j^2$$

and condition (5.2) is satisfied if $\sigma_n^2 < \sigma^2$ for $n = 1, 2, \ldots$ .

PROOF OF THEOREM 6. It follows from (4.2) and (4.3) that $\{\mathbf{E}W_n^{+2}\}$ is uniformly bounded. Because in addition, $W_n^+$ is a nonnegative submartingale, (2.17) follows from, for example, Doob [(1953), Theorem 4.1s, page 325]. The fact that $\mathbf{E}W \geq 1$ follows from $\mathbf{E}W_n^+ \geq 1$ for $n = 1, 2, \ldots$ (notice that it is possible that $\mathbf{E}W > 1$). If $\mathbf{P}(X_n \to 0) > 0$, then $\mathbf{P}(W = 0) \geq \mathbf{P}(X_n \to 0) > 0$, which jointly with $\mathbf{E}W \geq 1$ implies that $\mathrm{var}(W) > 0$. If $\mathbf{P}(X_n \to 0) = 0$, then $X_n^+ = X_n$ for $n \geq 0$ and so it follows from (4.3) that $\mathrm{var}(W) = \sigma^2 A_1 + \mathrm{var}(K_1^+/a) > 0$. □

REMARKS ABOUT THEOREM 6.

1. It is not difficult to show that (2.17) of Theorem 6 remains true if $W_n^+$ is replaced by $W_n$.
2. The stated properties of the rv $W$ imply the result (2.9). That is, Theorem 6 provides conditions, possibly different from those of Theorem 2, for (2.9) to hold.
3. In the BGW case, it is well known that $\mathbf{P}(W = 0) = \mathbf{P}(X_n \nrightarrow \infty)$. We have been unable to ascertain if this result is more generally true.

## 6. Convergence in distribution.

6.1. *Discussion of Theorem* 7. Theorem 7 is only of interest when $\mathbf{P}(U > 0) > 0$. Assumption (2.19) is satisfied, for example, when the $K_i$ are iid with mean $a$ and variance $\sigma^2 > 0$ [see Rényi and Révész (1958)]. Corollary 8 provides a rate of convergence result complementing Theorem 4 and Corollary 5.

Under the conditions of Theorem 7, we have $Z_n(1) \to_{\mathscr{D}} \mathscr{N}$, so if $\{Z_n\}$ satisfies Anscombe's condition, then as a consequence of Theorem E of Csörgő and Fischler (1973), with $\nu_n = (T_n | U > 0)$ and $f(n) = a^n$, it follows that as $n \to \infty$,

$$\left( \frac{T_{n+1} - aT_n}{\sigma\sqrt{T_n}} \middle| U > 0 \right) \to_{\mathscr{D}} \mathscr{N} \tag{6.1}$$

together with a corresponding result with $X_{n+1} - (a-1)T_n$ in the numerator. Anscombe's condition is satisfied, for example, if $K_i$ are iid (or a martingale difference sequence) with mean $a$ and variance $\sigma^2$ [see also Theorem G of Csörgő and Fischler (1973)].

PROOF OF THEOREM 7. The conditions of the theorem imply the convergence results

$$(Z_n, U) \to_{\mathscr{D}} (\mathscr{W}, U'), \qquad (Z_n', U) \to_{\mathscr{D}} (\mathscr{W}, U') \tag{6.2}$$

and

$$(Z_n', U_n) \to_{\mathscr{D}} (\mathscr{W}, U'), \tag{6.3}$$

where

$$Z_n'(t) = \frac{1}{\sigma\sqrt{a^n}} \sum_{j=1}^{[ta^n]} (K_j - a), \qquad 0 \leq t < \infty,$$

whereas $U' =_{\mathscr{D}} U$ and $U'$ is independent of $\mathscr{W}$.

To see this, notice that if $B = U^{-1}B'$ for some $B' \in \mathscr{B}(R)$, then we have

$$\mathbf{P}(Z_n(t_1) \in A_1, \dots, Z_n(t_k) \in A_k; U \in B') = \mathbf{P}(Z_n(t_1) \in A_1, \dots, Z_n(t_k) \in A_k; B),$$

which jointly with (2.19) gives

$$\mathbf{P}(Z_n(t_1) \in A_1, \dots, Z_n(t_k) \in A_k; U \in B') \to \mathbf{P}(\mathscr{W}(t_1) \in A_1, \dots, \mathscr{W}(t_k) \in A_k)\mathbf{P}(U' \in B'),$$

where the Borel sets $A_i$ are continuity sets of $\mathscr{W}$. (6.2) follows using this and the tightness of $(Z_n, U)$ [which follows from (2.20)]. Now using (2.21) and the second part of (6.2) we get the convergence result

$$\sum_{i=1}^{k} b_i Z_n'(t_i) + U_n b = \sum_{i=1}^{k} b_i Z_n'(t_i) + Ub + b(U_n - U) \to_{\mathscr{D}} \sum_{i=1}^{k} b_i \mathscr{W}(t_i) + U'b,$$

for any real numbers $b, b_1, \dots, b_k$. Hence

$$(Z_n'(t_1), \dots, Z_n'(t_k), U_n) \to_{\mathscr{D}} (\mathscr{W}(t_1), \dots, \mathscr{W}(t_k), U'),$$

which jointly with the tightness of $(Z_n', U_n)$ [which follows from (2.20)] gives (6.3).

We are now in a position to prove (2.22). Notice that

$$\frac{T_{n+1} - aT_n}{\sigma\sqrt{a^n}} = \frac{1}{\sigma\sqrt{a^n}} \sum_{j=1}^{T_n} (K_j - a) = Z_n'\left(\frac{T_n}{a^n}\right).$$

Writing $\tilde{U}_n(t) = t \cdot U_n$ for $t \geq 0$ and using (6.3), we get the convergence

$$\left(Z_n', \tilde{U}_n\right) \to_{\mathscr{D}} (\mathscr{W}, \tilde{U}) \quad \text{as } n \to \infty,$$

where $\tilde{U}(t) = t \cdot U' \geq 0$, $U' =_{\mathscr{D}} U$ and $U'$ and $\mathscr{W}$ are mutually independent. Hence by the continuous mapping theorem and by random change of time [see Whitt (1980)], we get

$$\frac{1}{\sigma\sqrt{a^n}} \sum_{i=1}^{T_n} (K_j - a) \to_{\mathscr{D}} \mathscr{W}(U') \quad \text{as } n \to \infty.$$

Because $\mathscr{W}$ and $U$ are independent, it follows that $\mathscr{W}(U') =_{\mathscr{D}} \mathscr{N}\sqrt{U'}$. □

Notice that Theorem 7 can be extended to the following form: Let $Z_n(t) = (1/a_n) \sum_{j=1}^{[ta^n]} (K_j - a)$, $t \geq 0$, where $\{a_n\}$ are some constants. If $(Z_n, T_n/a^n) \to_{\mathscr{D}} (Z, V)$ as $n \to \infty$, where $Z$ has stochastically continuous paths, that is, $\mathbf{P}(Z(t) = Z(t-)) = 1$ for all $t \geq 0$, then

$$\frac{T_{n+1} - aT_n}{a_n} \to_{\mathscr{D}} Z(V) \quad \text{as } n \to \infty.$$

**Acknowledgment.** We thank two referees for their incisive comments.

## REFERENCES

BAHADUR, R. R. (1971). *Some Limit Theorems in Statistics*. SIAM, Philadelphia.

BAHADUR, R. R. and RANGA RAO, R. (1960). On deviations of the sample mean. *Ann. Math. Statist.* **31** 1015–1027.

CSÖRGŐ, M. and FISCHLER, R. (1973). Some examples and results in the theory of mixing and random-sum central limit theorems. *Period Math. Hungar.* **3** 41–57.

DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

HARRIS, T. E. (1952). First passage and recurrence distributions. *Trans. Amer. Math. Soc.* **73** 471–486.

HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin.

HEYDE, C. C. (1970). Extension of a result of Seneta for the super-critical Galton–Watson process. *Ann. Math. Statist.* **41** 739–742.

HEYDE, C. C. and SENETA, E. (1977). *I. J. Bienaymé: Statistical Theory Anticipated*. Springer, New York.

JAGERS, P. (1992). Stabilities and instabilities in population dynamics. *J. Appl. Probab.* **29** 770–780.

LINDVALL, T. (1976). On the maximum of a branching process. *Scand. J. Statist.* **3** 209–214.

RÉNYI, A. and RÉVÉSZ, P. (1958). On mixing sequences of random variables. *Acta Math. Acad. Sci. Hungar.* **9** 389–393.

SHIRYAYEV, A. N. (1984). *Probability*. Springer, New York.

WHITT, W. (1980). Some useful functions for functional limit theorems. *Math. Oper. Res.* **5** 67–85.

SCHOOL OF MATHEMATICS AND STATISTICS
UNIVERSITY OF SYDNEY
SYDNEY, NSW 2006
AUSTRALIA

MATHEMATICAL INSTITUTE
WROCŁAW UNIVERSITY
PL. GRUNWALDZKI 2/4
50-384 WROCŁAW
POLAND