SOME ASYMPTOTIC RESULTS FOR THE BRANCHING PROCESS WITH IMMIGRATION

C.Z. WEI
Department of Mathematics, University of Maryland College Park, MD 20742, USA

J. WINNICKI
Department of Statistics, Columbia University, New York, NY 10027, USA

Received 24 November 1987
Revised 10 October 1988

Motivated by the statistical applications, the asymptotic behavior of certain functionals of a branching process with immigration, \( \{X_n\} \), is studied. The main results concern the critical case. A functional limit theorem for \( \frac{X_n}{n} \) is established. The rate of growth for \( T_n = \sum_{i=0}^{n-1} (X_i + 1)^{-1} \) is examined and found to depend on whether the process is transient or recurrent. Some convergence theorems for the supercritical case are also included.

AMS 1980 Subject Classification: Primary 60J80.

branching process with immigration * asymptotic distribution * weak convergence * martingales * diffusion approximation * Mittag-Leffler distribution

1. Introduction

Let \( \{X_n\} \) be a branching process with immigration defined by

\[
X_n = \sum_{j=1}^{X_{n-1}} Y_{n,j} + I_n, \quad n = 1, 2, \ldots,
\]

where \( \{Y_{n,j}\} \) and \( \{I_n\} \) are independent sequences of i.i.d. nonnegative, integer valued random variables and \( X_0 \) is a nonnegative, integer valued random variable which is independent of \( \{Y_{n,j}\} \) and \( \{I_n\} \). When \( m = E(Y_{1,1}) < \infty \), the process \( \{X_n\} \) is referred to as subcritical if \( m < 1 \), critical if \( m = 1 \) and supercritical if \( m > 1 \).

The study of \( \{X_n\} \) dates back to Smoluchowski (1916), and there is a substantial literature on asymptotic behavior of functionals of \( \{X_n\} \). However, in an attempt (Wei and Winnicki, 1987) to solve a long standing estimation problem raised by Heyde and Seneta (1974) we encounter several new functionals which were not investigated before. These functionals are related to the ordinary and weighted conditional least squares estimators of the means \( m \) and \( \lambda = E(I_1) \).

* Research supported by the National Science Foundation under grant NSF-DMS-8603084.
Assume that \( \lambda < \infty \) and let \( \mathcal{F}_n \) be the \( \sigma \)-field generated by \( \{X_0, Y_{i,j}, I_i; 1 \leq i \leq n, 1 \leq j < \infty \} \). Then by (1.1),

\[
X_n = E(X_n|\mathcal{F}_{n-1}) + \varepsilon_n = mX_{n-1} + \lambda + \varepsilon_n, \quad n = 1, 2, \ldots,
\]

where

\[
\varepsilon_n = \sum_{j=1}^{n} (Y_{n,j} - m) + I_n - \lambda.
\]

The least squares estimators for \( m \) and \( \lambda \) based on the regression equation (1.2) are

\[
\hat{m}_n = \left[ \sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_{i-1} - n \sum_{i=1}^{n} X_i^2 \right] \left[ \sum_{i=1}^{n} X_i^2 - n \sum_{i=1}^{n} X_i^2 \right]^{-1} \tag{1.4}
\]

and

\[
\hat{\lambda}_n = \left[ \sum_{i=1}^{n} X_{i-1}X_i \sum_{i=1}^{n} X_{i-1} - n \sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} X_i \right] \left[ \sum_{i=1}^{n} X_i^2 - n \sum_{i=1}^{n} X_i^2 \right]^{-1}. \tag{1.5}
\]

These and closely related estimators were proposed and studied for the subcritical case (Heyde and Seneta, 1972, 1974; Quine, 1976; Klimko and Nelson, 1978; and Venkataraman, 1982). The problem raised by Heyde and Seneta is to provide a unified estimation theory which would allow inference without knowing the range of \( m \). In the subcritical case the process \( \{X_n\} \) is essentially stationary and ergodic (cf. Remark 2.9) and the estimation theory in this case is completely parallel to the analogous results for linear time series. However, in general, even consistency of \( \hat{m}_n \) and \( \hat{\lambda}_n \) is in question.

Furthermore, if \( 0 < \sigma^2 = E(Y_{1,1} - m)^2 < \infty \) and \( 0 < b^2 = E(I_1 - \lambda)^2 < \infty \), then

\[
E(\varepsilon_n^2|\mathcal{F}_{n-1}) = \sigma^2 X_{n-1} + b^2. \tag{1.6}
\]

In view of the Gauss-Markov theorem this heterogeneity of the conditional variances of the "error" process \( \{\varepsilon_n\} \) suggests that \( \hat{m}_n \) and \( \hat{\lambda}_n \) may not be efficient. Therefore, we are led to consider stabilizing the conditional variances by assigning weights to (1.2). If we multiply both sides of (1.2) by the weights \( (1 + X_{n-1})^{-1/2} \), then the resulting weighted least squares estimators are

\[
\hat{m}_n = \left[ \sum_{i=1}^{n} X_i \sum_{i=1}^{n} (1 + X_{i-1})^{-1} - n \sum_{i=1}^{n} X_i (1 + X_{i-1})^{-1} \right] \left[ \sum_{i=1}^{n} (1 + X_{i-1})^{-1} - n \right]^{-1} \tag{1.7}
\]

and

\[
\hat{\lambda}_n = \left[ \sum_{i=1}^{n} X_{i-1} \sum_{i=1}^{n} X_i (1 + X_{i-1})^{-1} - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_{i-1} (1 + X_{i-1})^{-1} \right] \left[ \sum_{i=1}^{n} (1 + X_{i-1})^{-1} - n \right]^{-1}. \tag{1.8}
\]
Hence, in order to analyze $\hat{m}_n$ and $\hat{\lambda}_n$ we have to study not only the functional $\sum_{i=1}^n X_i$, which appears in (1.4)-(1.5), but also a new functional $T_n = \sum_{i=1}^n (1 + X_{i-1})^{-1}$ in (1.7) and (1.8).

Section 2 is devoted to a study of the critical case. We first prove a weak convergence result for the random element $X_{[m]}/n$ in $D^+[0, \infty)$ (Theorem 2.1). The limit is characterized as a diffusion process. Our results extend those of Kawazu and Watanabe (1971) and Mellein (1983b) on convergence of finite dimensional distributions and Lindvall (1972), who proved functional convergence for the nonimmigration process. The continuous mapping theorem is then used to derive as simple corollaries limiting results for $n^{-k-1}\sum_{i=1}^n X_i^k$ (Corollary 2.2). These functional convergence results are used in a companion paper (Wei and Winnicki 1987) to obtain asymptotic properties of the weighted least squares estimators of $m$ and $\lambda$.

The study of the functional $T_n = \sum_{i=1}^n (1 + X_{i-1})^{-1}$ is tied to another aspect of the process: in the critical case, $\{X_n\}$ is transient or null recurrent according as $\tau = 2\lambda/\sigma^2 > 1$ or $\tau \leq 1$. The normalizing factor $1/n$ in the weak convergence result mentioned above does not reflect this dichotomy. (However, the boundary behavior of the limiting diffusion does depend on $\tau$). The rate of growth of the functional $T_n$ turns out to depend on $\tau$. More precisely, in the transient case ($\tau > 1$), with the help of martingale theory, $T_n/\log n$ is shown to converge in probability to $(\lambda - \sigma^2/2)^{-1}$ (Theorem 2.12). In the case $\tau < 1$, using the recurrent Markov chain techniques (viz. Athreya and Ney 1978), $n^{\tau - 1}T_n$ is proved to converge in distribution to a multiple of a Mittag-Leffler distribution (Theorem 2.18). The case $\tau = 1$ seems to be more difficult; only some upper and lower bounds are given (Theorem 2.22).

In Section 3, devoted to the supercritical case, a weak convergence result in $\mathbb{R}^\infty$ with metric $d(x, y) = \sum_{i=1}^\infty 2^{-i}||x_i - y_i|/(1 + |x_i - y_i|)$ is given (Theorem 3.1). Our approach differs from the one used by Heyde and Brown (1971) in the study of the ordinary, nonimmigration branching process. In order to apply the continuous mapping theorem to functionals such as $h(x) = \sum_{i=1}^\infty m^{-i}x_i$, they introduce a new metric. As a consequence, they need a tightness argument. In our approach, this is replaced by a simple probabilistic lemma. In particular, we are able to rederive an analogue of the Central Limit Theorem (Heyde and Seneta 1971 and Corollary 3.4 below) as well as some weak convergence results of statistical relevance.

### 2. Critical case

Let $D^+[0, \infty)$ be the space of nonnegative functions on $[0, \infty)$ which are right continuous and have left limits. We equip $D^+[0, \infty)$ with the Skorohod topology. For the definition of this topology and the associated theory of weak convergence, the reader is referred to Ethier and Kurtz (1986). Now let $Y_n(t) = X_{[m]}/n$. It is clear that $Y_n$ is a sequence of random elements that take values in $D^+[0, \infty)$. We also denote by $Z$ a nonnegative diffusion process with generator

$$Af(x) = \frac{1}{2}\sigma^2 x^2f''(x) + \lambda f'(x), \quad f \in C_c^\infty[0, \infty), \quad (2.1)$$
where \( C^\infty_c(0, \infty) \) is the space of infinitely differentiable functions on \([0, \infty)\) which have compact supports (cf. Ikeda and Watanabe (1981)).

**Theorem 2.1.** Assume that \( m = 1, \sigma^2 < \infty \) and \( b^2 < \infty \). Then \( Y_n \to Y \) (weakly in \( D^+[0, \infty) \)), where \( Y \) is a diffusion with generator (2.1) and \( Y(0) = 0 \).

**Proof.** The theorem can be shown by an argument similar to that in Ethier and Kurtz (1986, Chapter 9, Theorem 1.3). We only indicate below the necessary modifications. Observe that \( \{X_j/n, j \geq 0\} \) is a Markov chain with values in \( E_n = \{l/n: l = 0, 1, \ldots\} \). For each \( f \in C^\infty_c([0, \infty)), \) define

\[
A_n f(x) = \frac{E f \left( n^{-1} \left( \sum_{j=1}^{n} Y_{1,j} + I_j \right) \right)}{x \in E_n.} \tag{2.2}
\]

Since \( Y_n(0) = X_0/n \to 0 \) a.s., it is sufficient to show that

\[
\lim_{n \to \infty} \sup_{x \in E_n} |\varepsilon_n(x)| = 0, \tag{2.3}
\]

where

\[
\varepsilon_n(x) = n(A_n f(x) - f(x)) - \frac{1}{2} \sigma^2 x f''(x) - \lambda f'(x).
\]

If we define \( S_n = n^{-1/2} \left( \sum_{j=1}^{n} (Y_{1,j} - 1) + I_j \right) \), then

\[
\varepsilon_n(x) = E \left[ \int_0^1 S_{nx}^2 x (1 - v) \{ f''(x + v \sqrt{x/n} S_{nx}) - f''(x) \} \, dv \right] \tag{2.4}
\]

and

\[
x ES_{nx}^2 = \sigma^2 x + (b^2 + \lambda^2)/n. \tag{2.5}
\]

Using (2.4), (2.5) and an argument similar to that on pages 388–389 of Ethier and Kurtz (1986), (2.3) follows.

**Remark 2.2.** Convergence of finite dimensional distributions of the sequence of processes \( Y_n \) was considered by Kawazu and Watanabe (1971) and Mellein (1982a, 1983b). Functional convergence for the nonimmigration branching process conditioned on nonextinction was first proved by Lindvall (1972). Our result can be viewed as an extension of these earlier works. It is well known (cf. Ikeda and Watanabe (1981, p. 222)) that the diffusion process \( Z \) determined by (2.1) satisfies

\[
E(e^{-s Z(t)}|Z(0) = x) = (1 + \frac{1}{2} \sigma^2 + s)^{-2\lambda/\sigma^2} e^{-sx/(1+(1/2)\sigma^2 s)}, \quad s \geq 0.
\]

We note that this is a special case of the Laplace transform of the transition function of the limiting process proposed by Kawazu and Watanabe (1971, Theorem 2.3). Their limiting process reduces to ours when \( \sigma^2 < \infty \) and \( \lambda < \infty \).
Furthermore, applying the Markov property and an induction argument, it is not difficult to see that the multivariate Laplace transform \( \phi_m(s_1, \ldots, s_m; t_1, \ldots, t_m) \) of the joint distribution of \((Y(t_1), \ldots, Y(t_m))\), \(0 \leq t_1 < \cdots < t_m\), is given by the recursive relation

\[
\phi_m(s_1, \ldots, s_m; t_1, \ldots, t_m) = \phi_{m-1}(s_1, \ldots, s_{m-1} + s_m / [1 + \frac{1}{2}\sigma^2(t_m - t_{m-1})s_m]; t_1, \ldots, t_{m-1}) \cdot [1 + \frac{1}{2}\sigma^2(t_m - t_{m-1})s_m]^{-2m/s^2},
\]

where \( s_1 \geq 0, \ldots, s_m \geq 0 \) and \( m > 1 \). Mellein (1983b, Lemma 5) characterized the limiting diffusion process by giving its joint finite-dimensional densities and it can be checked that their corresponding Laplace transforms are given by (2.6).

**Corollary 2.3.** Assume that \( m = 1, \sigma^2 < \infty \) and \( b^2 < \infty \). Then for \( k = 1, 2, \ldots, \)

\[
n^{-k-1} \sum_{i=0}^{n} X_i^k \xrightarrow{d} \int_0^1 Y^k(t) \, dt.
\]

**Proof.** This follows immediately by the continuous mapping theorem.

**Remark 2.4.** The invariance principle in \( \mathbb{R}^{k+1} \) can be used to prove joint convergence

\[
\left(\frac{X_n}{n}, n^{-2} \sum_{i=0}^{n} X_i, \ldots, n^{-k-1} \sum_{i=0}^{n} X_i^k\right) \xrightarrow{d} \left(Y(1), \int_0^1 Y(t) \, dt, \ldots, \int_0^1 Y^k(t) \, dt\right).
\]

This form of Corollary 2.3 is most useful for statistical applications.

**Remark 2.5.** Pakes (1972) proved that \( n^{-2} \sum_{i=1}^{n} X_i \) converges in distribution and gave the Laplace transform of the limiting distribution. Mellein (1983b) showed that this limiting Laplace transform coincides with the Laplace transform of \( \int_0^1 Y(t) \, dt \). Corollary 2.3 for \( k = 1 \) explains the connection between these results.

**Theorem 2.1** provides a functional limit result that enables one to obtain limiting distributions for certain functionals of the process \( \{X_n\} \). However, this theorem is not applicable to the functional \( T_n = \sum_{k=0}^{n-1} 1/(1 + X_k) \). In order to study the asymptotic behavior of \( T_n \), we will need a few preparatory lemmas. The first lemma characterizes \( \{X_n\} \) as a Markov chain. We introduce some notation. Let \( \kappa = \inf\{j \geq 0: P[I_j = j] > 0\} \), let \( p_n(i, j) = P(X_n = j|X_0 = i) \) be the \( n \)-step transition probabilities, let \( C_i = \{j: p_n(i, j) > 0 \text{ for some } n > 0\} \) and denote by \( G_{i, j} = \sum_{n=0}^{\infty} p_n(i, j) \) the Green function of \( \{X_n\} \).

**Lemma 2.6.** Assume that

\[
0 < P(Y_{1,1} = 0) < 1 \quad \text{and} \quad P(I_1 = 0) < 1.
\]
Then $\kappa \in C_i$ and $C_\kappa = C_i$ for all $i \geq 0$. Consequently, $C_\kappa$ is an irreducible, aperiodic set and for $n \geq 1$, $X_n$ has its support on $C_\kappa$.

Remark 2.7. Parts of this result were stated without proof by Seneta (1969), Pakes (1975b) and Quine (1976). The lemma implies that in the non-supercritical case ($m \leq 1$) a nondegenerate branching process with immigration is an irreducible aperiodic Markov chain on $C_\kappa$. It is not difficult to see that Lemma 2.6 can be obtained by an argument similar to that in Lemma 2 of Kesten, Ney and Spitzer (1966). We also omit the proof.

We next formulate a coupling result for the critical branching process with immigration.

Lemma 2.8. Assume that $m = 1$ and $P(Y_{1,1} = 1) < 1$. Then for any probability distribution $\{q_i, i = 0, 1, \ldots\}$ concentrated on nonnegative integers, there exists a branching process with immigration $\{Z_n\}$ such that $P(Z_0 = i) = q_i$, $i = 0, 1, \ldots$, and $P(X_n = Z_n$ eventually) = 1.

Proof. Let $Z_0$ be an independent of $\{Y_{n,i}\}$ and $\{I_n\}$ random variable with distribution $\{q_i\}$ and define the process $\{Z_n\}$ by

$$Z_{n+1} = \sum_{i=1}^{Z_n} Y_{n+1,i} + I_{n+1} \quad \text{for } n = 0, 1, \ldots$$

Consider first the case $Z_0 = 0$. In this case, $X_n - Z_n$ is an ordinary Galton-Watson process with offspring numbers $Y_{n,i}$ and the initial population size $X_0$. Under our assumptions, it is well known that the extinction probability of $\{X_n - Z_n\}$ is 1. Hence

$$P(X_n - Z_n = 0 \text{ eventually}) = P(X_n = Z_n \text{ eventually}) = 1.$$ 

The case of general $Z_0$ now follows since both $X_n$ and $Z_n$ eventually coincide with the process started at zero.

Remark 2.9. Lemma 2.8 remains valid when $m < 1$. It is well known that in this case (under the assumption that $E(\log^+(I_1)) < \infty$) the process has a unique limiting stationary distribution. Lemma 2.8 is then a consequence of the more general coupling property true of any aperiodic positive recurrent Markov chain (cf. Griffeath (1975), Nummelin (1984)). However, this simple fact has important applications in the study of asymptotic properties of the statistics related to $\{X_n\}$. In particular, it shows that no loss of generality is entailed by assuming that $X_0$ has the limiting, stationary distribution of the process. This results in a considerable simplification and clarification of the proofs of the Central Limit Theorem and the Law of the Iterated Logarithm for the estimators of $m$ and $\lambda$ in the subcritical case. Earlier authors (Heyde and Seneta (1972, 1974), Quine (1976)) had to use special arguments to get around the problem that given an arbitrary initial distribution, the process $\{X_n\}$ is, in general, not stationary. The above observation removes this
obstacle and ensures that the usual limit theorems for the sum of stationary martingale increments are applicable. For another application see Wei and Winnicki (1987). In the context of the present paper (\(m = 1\)), Lemma 2.8 will allow us to assume, without loss of generality, that \(X_0 = \text{const.} = \kappa = \inf\{j \geq 0: P[I_1 = j] > 0\}\).

Recall the notation \(\tau = 2\lambda / \sigma^2\).

**Lemma 2.10.** Assume that \(m = 1, 0 < \sigma^2 < \infty, b^2 < \infty, \) and \(E( Y_{1,1}^2 \log^+ Y_{1,1}) < \infty. \) Then

\[
p_n(i, j) \sim \mu_j n^{-\tau} \quad \text{as } n \to \infty,
\]

where \(\{\mu_j\}\) is an invariant measure for \(\{X_n\}\) with \(\mu_\kappa > 0\) and

\[
\mu_j \sim ((\sigma^2/2)\Gamma^2(\tau))^{-1} n^{\tau - 1} \quad \text{as } j \to \infty.
\]

If \(\tau > 1,\)

\[
G_{i,j} \to (\lambda - \sigma^2/2)^{-1} \quad \text{as } j \to \infty.
\]

**Proof.** When \(\kappa = 0,\) (2.10) is due to Pakes (1972) while (2.11) and (2.12) are due to Mellein (1982b, 1983a). For the general case, observe that \(X_n \geq \kappa \) a.s. for \(n \geq 1\) and

\[
X_{n+1} - \kappa =\sum_{j=1}^{X_{n+1} - \kappa} Y_{n+1, j} + \left( I_{n+1} - \kappa + \sum_{j=1}^{n} Y_{n+1, j} \right).
\]

Hence \(\{X_n - \kappa, n \geq 1\}\) is a branching process with immigration. The immigration process \(I_n = I_n - \kappa + \sum_{j=1}^{\kappa} Y_{n, j}\) has mean \(\lambda - \kappa + m\kappa = \lambda.\) Thus the value of \(\tau\) is not affected under the transformation (2.13). Since \(P(I_n = 0) \geq P(I_n = \kappa) \prod_{i=1}^{n} P(Y_{n, j} = 0) > 0,\) (2.10)-(2.12) are valid for the process \(\{X_n - \kappa, n \geq 1\}\) by Pakes' and Mellein's results, and consequently also for the process \(\{X_n\}\).

**Corollary 2.11.** Assume that

\[
m = 1, \quad 0 < \sigma^2 < \infty, \quad 0 < b^2 < \infty \quad \text{and} \quad E( Y_{1,1}^2 \log^+ Y_{1,1}) < \infty.
\]

Then \(C_\kappa\) is a null recurrent class when \(\tau \leq 1\) and transient otherwise.

**Proof.** The assumptions \(m = 1, \sigma^2 > 0\) and \(b^2 > 0\) imply (2.9). Hence Lemma 2.6 is applicable and \(C_\kappa\) is an irreducible and aperiodic class. Now by (2.10), \(\sum_{n=1}^{\infty} p_n(\kappa, \kappa) = \infty\) or \(< \infty\) according as \(\tau \leq 1\) or \(\tau > 1.\) This completes our proof.

By Lemma 2.8 (cf. Remark 2.9) we can (and shall) assume for the rest of this section that \(X_0 = \kappa.\) We also recall that \(\mathcal{F}_n\) is the \(\sigma\)-field generated by \(\{X_0, Y_{i, j}, I_i, 1 \leq i \leq n, 1 \leq j\}.\) We will now formulate a fundamental but far from trivial property of \(T_n = \sum_{i=0}^{n-1} 1/(X_i + 1).\)
**Theorem 2.12.** Suppose that (2.14) holds. Then $T_n \to \infty$ a.s.

**Proof.** By Corollary 2.11, \( \{X_n\} \) is either null recurrent or transient. In the recurrent case, it is obvious that $T_n \to \infty$ a.s. In the transient case, we have that

\[ X_n \to \infty \quad \text{a.s.} \quad (2.15) \]

Define

\[ M_n = \frac{(X_n + 1)^{n-1}}{\prod_{i=0}^{n-1} \left[ 1 + \lambda/(X_i + 1) \right]} . \]

It is not difficult to check that \( \{M_n, \mathcal{F}_n\} \) is a positive martingale. By the martingale convergence theorem (Hall and Heyde, 1980, p. 58), \( M_n \) converges a.s. Using (2.15) this in turn implies that

\[ \prod_{i=0}^{n} \left[ 1 + \lambda/(X_i + 1) \right] \to \infty \quad \text{a.s.} \]

From the theory of infinite product, we have that

\[ \sum_{i=0}^{n} \lambda/(X_i + 1) \to \infty \quad \text{a.s.} \]

This completes our proof.

The next two lemmas, needed to obtain the rate of growth of $T_n$, are of independent interest.

**Lemma 2.13.** Assume that (2.14) holds. If $\tau > 1$, then for any nonnegative and nonincreasing function \( f \) defined on \( \{0, 1, \ldots\} \), such that $\sum_{n=0}^{\infty} f(n) < \infty$, we have

\[ \sum_{n=0}^{\infty} Ef(X_n) < \infty \quad (2.16) \]

and

\[ \sum_{n=0}^{\infty} f(X_n) < \infty \quad \text{a.s.} \quad (2.17) \]

**Proof.** It is clear that (2.17) follows from (2.16). Now,

\[ \sum_{n=0}^{\infty} Ef(X_n) = \sum_{j=0}^{\infty} f(j) \sum_{n=0}^{\infty} P(X_n = j) = \sum_{j=0}^{\infty} f(j) G_{\kappa, j} < \infty \]

using (2.12) of Lemma 2.10.

**Lemma 2.14.** Assume that (2.14) holds. If $\tau > 1$ and for some $\delta > 0$, $E(Y_{1,1}^{\kappa, \delta}) < \infty$, then

\[ X_n/X_{n-1} \to 1 \quad \text{a.s.} \quad (2.18) \]
Proof. Since $X_n \to \infty$ a.s., it is sufficient to show that

$$\left(\frac{X_n - X_{n-1}}{X_{n-1} + 1}\right) = \sum_{i=1}^{X_{n-1}} \frac{Y_{n,i}}{X_{n-1} + 1} + I_n/(X_{n-1} + 1) \to 0 \quad \text{a.s.}$$

Note that for any $\epsilon > 0$, by the conditional Markov's inequality and Lemma 2.13,

$$\sum_{n=1}^{\infty} P\left(\frac{I_n}{X_{n-1} + 1} > \epsilon \frac{\mathcal{F}_{n-1}}{X_{n-1} + 1}\right) \leq \epsilon^{-2} \sum_{n=1}^{\infty} E\left(\frac{I_n^2}{(1 + X_{n-1})^2}\right) < \infty \quad \text{a.s.}$$

By the conditional Borel–Cantelli lemma (Hall and Heyde, 1980, p. 32),

$$\frac{I_n}{X_{n-1} + 1} \to 0 \quad \text{a.s.}$$

Now, using the fact that $E\left[\sum_{i=1}^{I_n} (Y_{n,i} - 1)^2\right] = O(I_n^{1+\delta}/2)$ (Chow and Teicher, 1978, p. 357), there is a constant $K > 0$ such that

$$\sum_{n=1}^{\infty} \frac{1}{X_{n-1} + 1} \leq K \sum_{n=1}^{\infty} \frac{X_{n-1}^{1+\delta/2}}{(1 + X_{n-1})^{2+\delta}} < \infty \quad \text{a.s.}$$

by Lemma 2.13.

Applying the conditional Borel–Cantelli lemma again,

$$\sum_{i=1}^{X_{n-1}} \frac{Y_{n,i} - 1}{X_{n-1} + 1} \to 0 \quad \text{a.s.}$$

This completes our proof.

The following theorem, not to be used in the sequel, is a refinement of Lemma 2.14 under a higher moment assumption. It is an analogue of the law of the Iterated Logarithm studied by Heyde and Seneta (1971) for the supercritical case.

Theorem 2.15. Assume that (2.14) holds. If $\tau > 1$ and $E(Y_{1,1}^{4+\delta}) < \infty$ for some $0 < \delta < 1$, then

$$\limsup_{n \to \infty} \frac{|X_n - X_{n-1}|}{(2\sigma^2 X_{n-1} \log X_{n-1})^{1/2}} = 1 \quad \text{a.s.}$$

(To avoid ambiguity when $X_n = 1$, we interpret $X_n$ as $X_n \vee 2$. This applies also to the proof below.)

Proof. By an argument similar to that in the proof of Lemma 2.14, it is enough to show that

$$\limsup_{n \to \infty} \frac{\sum_{j=1}^{X_{n-1}} (Y_{n,j} - 1)}{(2\sigma^2 X_{n-1} \log X_{n-1})^{1/2}} = 1 \quad \text{a.s.} \quad (2.19)$$
Using a result on moderate deviations due to Rubin and Sethuraman (1965, theorem 4), we obtain that for \(-\delta/8 < \varepsilon < \delta/8,\)
\[
P\left( \left| \sum_{j=1}^{n-1} \left( Y_{n,j} - 1 \right) \right| / (2\sigma^2 X_{n-1} \log X_{n-1})^{1/2} > 1 + \varepsilon \right| \mathcal{F}_{n-1} \right) 
\sim \{X_{n-1}^{(1+\varepsilon)^2}(1+\varepsilon)(\pi \log X_{n-1})^{1/2}\}^{-1} \text{ a.s. as } n \to \infty.
\]

By Theorem 2.12 and Lemma 2.13,
\[
\sum_{n=1}^{\infty} \{X_{n-1}^{(1+\varepsilon)}(1+\varepsilon)(\pi \log X_{n-1})^{1/2}\}^{-1}
\]
converges or diverges a.s. according to whether \(\varepsilon > 0\) or \(-1 < \varepsilon < 0\). Thus, (2.19) follows from the conditional Borel-Cantelli lemma.

We will now resume our study of the functional \(T_n = \sum_{i=0}^{n-1} (1 + X_i)^{-1}\). We first study the rate of growth of \(T_n\) in the transient case.

**Theorem 2.16.** Assume that (2.14) holds. If \(\tau > 1\) and \(E(Y_{1,1}^{2+\delta}) < \infty\) for some \(0 \leq \delta < 1\), then
\[
T_n / \log n \overset{p}{\to} \left( \lambda - \frac{\sigma^2}{2} \right)^{-1}.
\]

Before we prove this theorem, let us recall a local martingale convergence theorem due to Chow (1965). The proof of this result can be found in Hall and Heyde (1980, Theorem 2.17).

**Lemma 2.17.** Let \(1 \leq p \leq 2\), let \(\{\varepsilon_n\}\) be a sequence of martingale differences, with respect to an increasing sequence of \(\sigma\)-fields \(\{\mathcal{G}_n\}\) and \(S_n = \sum_{i=1}^{n} \varepsilon_i\). Then
\[
S_n \text{ converges a.s. on the set } \left\{ \sum_{i=1}^{\infty} E(|\varepsilon_i|^p | \mathcal{G}_{i-1}) < \infty \right\}
\]
and
\[
S_n = o\left( \sum_{i=1}^{n} E(\varepsilon_i^2 | \mathcal{G}_{i-1}) \right) \text{ a.s. on the set } \left\{ \sum_{i=1}^{\infty} E(\varepsilon_i^2 | \mathcal{G}_{i-1}) = \infty \right\}.
\]

**Proof of Theorem 2.16.** By Theorem 2.1, \(X_n / n \overset{d}{\to} Y(1)\). Since \(Y(1)\) has a gamma distribution by (2.6), \(Y(1) > 0\) a.s. and consequently
\[
\log(X_{n+1})/\log n \overset{p}{\to} 1.
\]

Observe that
\[
\sum_{i=1}^{n} \log[(X_i + 1)/(X_{i-1} + 1)] = \log(X_n + 1) - \log(X_0 + 1).
\]
In view of this identity, (2.23) and Theorem 2.12, in order to prove (2.20), it is sufficient to show that
\[
\sum_{i=1}^{n} \log[(X_i + 1)/(X_{i-1} + 1)]/T_n \rightarrow \lambda - \sigma^2/2 \quad \text{a.s.}
\] (2.24)

Now let \( u_i = (X_i - X_{i-1})/(X_{i-1} + 1) \). Note that by Lemma 2.14, \( u_i \rightarrow 0 \) a.s. The Taylor expansion gives
\[
\log[(X_i + 1)/(X_{i-1} + 1)] = \log[1 + u_i] = u_i - \frac{1}{2}u_i^2 + v_i,
\] (2.25)
where \( v_i = \frac{1}{2}u_i^3(1 + \theta_i)^3 \) and \(|\theta_i| \leq |u_i|\). We are going to prove (2.24) by showing that
\[
\sum_{i=1}^{n} u_i/T_n \rightarrow \lambda \quad \text{a.s.,}
\] (2.26)
\[
\sum_{i=1}^{n} u_i^2/T_n \rightarrow \sigma^2 \quad \text{a.s.,}
\] (2.27)
and
\[
\sum_{i=1}^{n} v_i/T_n \rightarrow 0 \quad \text{a.s.}
\] (2.28)

To show (2.26) is equivalent to proving that
\[
\sum_{i=1}^{n} [(X_i - X_{i-1} - \lambda)/(1 + X_{i-1})]/T_n \rightarrow 0 \quad \text{a.s.}
\] (2.29)

Note that \( \{\sum_{i=1}^{n} (X_i - X_{i-1} - \lambda)/(1 + X_{i-1}), \mathcal{F}_n\} \) is a martingale. Hence, (2.29) follows from (2.22) of Lemma 2.17, theorem 2.12 and the fact that
\[
\sum_{i=1}^{n} E((X_i - X_{i-1} - \lambda)^2/(1 + X_{i-1})^2)
\]
\[
= \sum_{i=1}^{n} (X_{i-1}\sigma^2 + b^2)/(1 + X_{i-1})^2 = O(T_n) \quad \text{a.s.}
\]

For (2.27), we first note that
\[
(X_i - X_{i-1})^2 = \sum_{j=1}^{X_{i-1}} (Y_{i,j} - 1)^2 + 2 \sum_{l=2}^{X_{i-1}} \sum_{j=1}^{l-1} (Y_{i,j} - 1)(Y_{i,l} - 1)
\]
\[
+ 2I_i \sum_{j=1}^{X_{i-1}} (Y_{i,j} - 1) + I_i^2
\]
\[
= Q_{i1} + 2Q_{i2} + 2Q_{i3} + I_i^2, \quad \text{say.}
\] (2.30)

Since by Lemma 2.13,
\[
E \sum_{i=1}^{\infty} I_i^2/(1 + X_{i-1})^2 = b^2 \sum_{i=1}^{\infty} E(1 + X_{i-1})^{-2} < \infty,
\]
we have that \( \sum_{i=1}^{\infty} I_i^2/(1 + X_{i-1})^2 < \infty \) a.s. and
\[
\sum_{i=1}^{n} I_i^2/(1 + X_{i-1})^2 = o(T_n) \quad \text{a.s.} \tag{2.31}
\]

It is not difficult to check that \( \{ \sum_{i=1}^{\infty} Q_{i3}/(1 + X_{i-1})^2, \mathcal{F}_n \} \) is a martingale with
\[
\sum_{i=2}^{\infty} E(Q_{i3}^2 | \mathcal{F}_{i-1})/(1 + X_{i-1})^4 = (\lambda^2 + b^2) \sigma^2 \sum_{i=2}^{\infty} X_{i-1}^4/(1 + X_{i-1})^4 < \infty \quad \text{a.s.,}
\]
by Lemma 2.13. Consequently,
\[
\sum_{i=1}^{n} Q_{i3}/(1 + X_{i-1})^2 = o(T_n) \quad \text{a.s.} \tag{2.32}
\]

Now, by the independence of \( Y_{i,j} \) and \( Y_{i,i} \) for \( j \neq i \), \( \{ \sum_{i=1}^{n} Q_{i2}^2/(1 + X_{i-1})^2, \mathcal{F}_n \} \) is an \( L^2 \)-martingale such that
\[
\sum_{i=2}^{\infty} E(Q_{i2}^2 | \mathcal{F}_{i-1})/(1 + X_{i-1})^4 = \frac{\sigma^4}{2} \sum_{i=2}^{\infty} X_{i-1}(X_{i-1} - 1)/(X_{i-1} + 1)^4 < \infty \quad \text{a.s.,}
\]
by Lemma 2.13. Thus,
\[
\sum_{i=2}^{n} Q_{i2}^2/(1 + X_{i-1})^2 = o(T_n) \quad \text{a.s.} \tag{2.33}
\]

In view of (2.30)-(2.33), in order to obtain (2.27) it is sufficient to show that
\[
\sum_{i=2}^{n} (Q_{i1} - \sigma^2 X_{i-1})^2/(1 + X_{i-1})^2 = o(T_n) \quad \text{a.s.}
\]
or, more strongly,
\[
\sum_{i=2}^{n} (Q_{i1} - \sigma^2 X_{i-1})^2/(1 + X_{i-1})^2 \quad \text{converges a.s.} \tag{2.34}
\]

Clearly, \( \{ \sum_{i=2}^{n} (Q_{i1} - \sigma^2 X_{i-1})^2/(1 + X_{i-1})^2, \mathcal{F}_n \} \) is a martingale. By the fact that
\[
E \left| \sum_{i=1}^{n} [(Y_{1,i} - 1)^2 - \sigma^2] \right|^{1+\delta/2} = o(n)
\]
(Chow and Teicher, 1978, p. 361),
\[
\sum_{i=2}^{\infty} E[|Q_{i1} - \sigma^2 X_{i-1}|^{1+\delta/2} | \mathcal{F}_{i-1}]/(1 + X_{i-1})^{2+\delta} \leq K \sum_{i=2}^{\infty} X_{i-1}/(1 + X_{i-1})^{2+\delta} < \infty \quad \text{a.s.}
\]
by Lemma 2.13 and \( K \) denoting a constant. By Lemma 2.17 with \( p = 1 + \delta/2 \), (2.34) follows.
Finally, let us prove (2.28). As observed earlier, $u_n \to 0$ a.s. Consequently, $\theta_n \to 0$ a.s. and by Toeplitz' lemma and (2.27),

$$\sum_{i=1}^{n} v_i = \frac{1}{3} \sum_{i=1}^{n} u_i^2 [u_i(1 + \theta_i)^3] = o\left(\sum_{i=1}^{n} u_i^2\right) = o(T_n) \quad \text{a.s.}$$

This completes our proof.

We now turn to the recurrent case.

**Theorem 2.18.** Assume that (2.14) holds. If $\tau < 1$, then for some constant $c > 0$,

$$n^{\tau-1} T_n \xrightarrow{d} cW_{1-\tau}, \quad (2.35)$$

where $W_q$ ($0 \leq q < 1$) is a random variable having the Mittag–Leffler distribution with parameter $q$, i.e.

$$P(W_q < x) = \frac{1}{\Gamma(q)} \sum_{j=0}^{\infty} (-1)^{j-1} \sin[q\pi(qj + 1)]y^{-j}/j! \, dy.$$  

Before proving this theorem, we will formulate a lemma from Pakes (1971). Note that although Pakes deals only with the case $\kappa = 0$, the same proof works for the general case using Lemma 2.10.

**Lemma 2.19.** Assume that (2.14) holds and $\tau \leq 1$. Let $a_n = n^{\tau-1}$ if $\tau < 1$ and $\log n$ otherwise. Then for some $c > 0$,

$$a_n^{-1} \sum_{i=0}^{n} I[X_i = \kappa] \xrightarrow{d} cW_{1-\tau}. \quad (2.36)$$

**Proof of Theorem 2.18.** By Corollary 2.11, $\tau < 1$ implies that $\{X_n\}$ is null recurrent. Let $S_0 = 0$ and $S_i = \inf\{n: n > S_{i-1}, X_n = \kappa\}$. Since we assumed that $X_0 = \kappa$ a.s., by Theorem 6.1 of Athreya and Ney (1978) the sequence $\{\phi_i\}$ defined by

$$\phi_i = E\left\{\frac{S_{j-1}}{\sum_{j=0}^{S_{j-1}} I[X_j = i]}\right\}$$

is an invariant measure for $\{X_n\}$. This invariant measure is unique up to a multiplicative constant. By (2.11) of Lemma 2.10, for some positive constant $K$

$$\phi_i \sim Ki^{\tau-1} \quad \text{as } i \to \infty. \quad (2.37)$$

For $n = 1, 2, \ldots$, let $\xi_n = \sum_{i=S_{n-1}}^{S_n} (X_i + 1)^{-1}$. It is known (Chung, 1967, p. 84) that $\{\xi_n\}$ are i.i.d. random variables. By (2.37),

$$E(\xi_1) = \sum_{i=0}^{\infty} (i + 1)^{-1} \phi_i < \infty.$$
Now define

\[ N_n = \sup\{i: S_i \leq n\} = \sum_{k=1}^{n} I[X_k = \kappa]. \]

Then by the Strong Law of Large Numbers and the recurrence of \( \{X_n\} \), we have that \( N_n \to \infty \) a.s. and

\[ \sum_{i=1}^{N_n} \xi_i / N_n \to E(\xi_1) \quad \text{a.s.} \quad (2.38) \]

Observe that

\[ N_n \xi_i \leq \sum_{i=0}^{n} (X_i + 1)^{-1} \leq \sum_{i=1}^{N_n+1} \xi_i. \]

This and (2.38) imply that

\[ \sum_{i=0}^{n} (X_i + 1)^{-1} / N_n \to E(\xi_1) \quad \text{a.s.} \quad (2.39) \]

The conclusion (2.35) therefore follows from Lemma 2.19 and the fact that

\[ n^{r-1} T_n = (n^{r-1} N_n) \left( \frac{\sum_{i=1}^{n} (X_i + 1)^{-1}}{N_n} \right). \]

**Corollary 2.20.** Assume that (2.14) holds and \( \tau < 1 \). Let \( \{\phi_i\} \) be an invariant measure of \( \{X_n\} \). Suppose that \( f \) and \( g \) are two functions on \( \{0, 1, \ldots\} \) such that

\[ \sum_{i=0}^{\infty} |f(i)| \phi_i < \infty, \sum_{i=0}^{\infty} |g(i)| \phi_i < \infty \quad \text{and} \quad \sum_{i=0}^{\infty} g(i) \phi_i \neq 0. \]

Then

\[ \sum_{i=0}^{n} f(X_i) / \sum_{i=0}^{n} g(X_i) \to \sum_{i=0}^{\infty} f(i) \phi_i / \sum_{i=0}^{\infty} g(i) \phi_i \quad \text{a.s.} \]

**Proof.** The proof used to show (2.39) yields

\[ \sum_{i=0}^{n} f(X_i) / \sum_{i=0}^{n} g(X_i) = \left( \sum_{i=0}^{n} f(X_i) / N_n \right) \left( \sum_{i=0}^{n} g(X_i) / N_n \right)^{-1} \]

\[ \to \sum_{i=0}^{\infty} f(i) \phi_i / \sum_{i=0}^{\infty} g(i) \phi_i \quad \text{a.s.} \]

**Corollary 2.21.** Assume that (2.14) holds and \( \tau < 1 \). Let \( \{\phi_i\} \) be an invariant measure of \( \{X_n\} \) and let \( f \) be a function on \( \{0, 1, \ldots\} \) such that \( \sum_{i=0}^{\infty} |f(i)| \phi_i < \infty \). Then for any \( \alpha > 1 - \tau \) we have

\[ n^{-\alpha} \sum_{i=0}^{n} f(X_i) \to 0 \quad \text{a.s.} \quad (2.40) \]

In particular, \( n^{-\alpha} T_n \to 0 \) a.s.
Proof. Using Corollary 2.20 with \( g(i) = \delta_i \) and (2.11),
\[
\sum_{i=0}^{n} f(X_i) / \sum_{i=0}^{n} I[X_i = \kappa] \to \sum_{i=0}^{\infty} f(i) \phi_i / \phi_\kappa \quad \text{a.s.}
\]
Hence it is sufficient to show that
\[
\sum_{i=0}^{n} I[X_i = \kappa] = o(n^\alpha) \quad \text{a.s.}
\]
By Kronecker's Lemma, this would follow if we can prove that
\[
\sum_{i=0}^{\infty} i^{-\alpha} I[X_i = \kappa] < \infty \quad \text{a.s.}
\]
But
\[
\sum_{i=0}^{\infty} i^{-\alpha} E(I[X_i = \kappa]) < \infty,
\]
by (2.10).

While Theorems 2.16 and 2.18 provide exact growth rates for \( T_n \) in the cases \( \tau > 1 \) and \( \tau < 1 \), the case \( \tau = 1 \) appears to be more difficult. The following theorem gives only a partial result.

**Theorem 2.22.** Assume that (2.14) holds and \( \tau = 1 \). Then
\[
T_n / \log n \overset{p}{\to} \infty \quad (2.41)
\]
and for any \( \alpha > 0 \),
\[
n^{-\alpha} T_n \to 0 \quad \text{a.s.} \quad (2.42)
\]
Proof. We will show (2.42) first. Let \( U_n \) be a sequence of i.i.d. random variables which are uniformly distributed over \([0, 1]\) and are independent of \( \{ Y_{n,i}, I_n, n \geq 1, i \geq 1 \} \). (If necessary, we can enlarge the probability space so that \( \{U_n\} \) can be defined.) Given \( \alpha > 0 \), choose \( 0 < p < 1 \) such that \( p + \alpha > 1 \). Define
\[
I'_n = I_n I[U_n \leq p]
\]
and
\[
X'_n = \sum_{i=1}^{X'_n} Y_{n,i} + I'_n, \quad X'_0 = 0.
\]
Clearly, \( X_n \geq X'_n \) a.s. and \( \tau' = [2E(I'_n)] / \var(Y_{n,i}) = 2p\lambda / \sigma^2 = p \). Thus, by Corollary 2.20
\[
T_n \leq \sum_{i=0}^{n} (1 + X'_i)^{-1} = o(n^\alpha) \quad \text{a.s.}
\]
Now let us show (2.41). Given $M > 0$, we can find $0 < q < 1$ such that $q^{-1} > M$. Define a sequence of i.i.d. random variables $J_n$ such that $\{J_n\}$ is independent of $\{Y_{n,i}, I_n, n \geq 1, i \geq 1\}$ and $P\{J_n = 1\} = 1 - P\{J_n = 0\} = q$. Also define

\[ Z_n = \sum_{i=1}^{Z_{n-1}} Y_{n,i} + I_n + J_n, \quad Z_0 = X_0. \]

Clearly, $Z_n \geq X_n$ a.s. and $2E(I_n + J_n)/\text{var}(Y_{n,i}) = 2(\lambda + q)/\sigma^2 > 1$. By Theorem 2.15,

\[ \frac{\sum_{i=0}^{n} (1+Z_i)^{-1}/\log n}{\log n} \rightarrow (\lambda + q - \sigma^2/2)^{-1} = q^{-1} > M. \]

Hence

\[ P\{T_n/\log n > M\} \geq P\left(\frac{\sum_{i=0}^{n} (1+Z_i)^{-1}}{\log n} > M\right) \rightarrow 1. \]

This completes our proof.

Remark 2.23. When $\tau < 1$, Theorem 2.18 and Lemma 2.19 imply that $\sum_{i=1}^{n} I[X_i = \kappa]$ and $T_n$ have the same order $n^{1-\tau}$. For the case $\tau = 1$, Lemma 2.19 implies that the order of $\sum_{i=1}^{n} I[X_i = \kappa]$ is $\log n$. However, (2.41) indicates that the order of $T_n$ should be larger. This is due to the fact that for $\tau = 1$

\[ \sum_{i=1}^{\infty} (1+i)^{-1} \mu_i = \infty, \quad (2.43) \]

where $\{\mu_i\}$ is an invariant measure for $\{X_n\}$. Note that (2.43) follows from (2.11) and (2.41) can be obtained using (2.36), (2.43) and the arguments in the proof of Theorem 2.18. Our approach has the advantage that if one is able to show that (2.20) holds a.s. then (2.41) would also hold a.s.

Remark 2.24. Pakes (1975a, p. 15) conjectures that for $\tau = 1$, there is a constant $c > 0$ such that $E(1 + X_n)^{-1} \sim c \log n/n$. If this conjecture is true then the exact order of $T_n$ would appear to be $(\log n)^2$. But the convergence type of the possible limiting result is still unknown.

3. Supercritical case

It is known (Seneta, 1970) that if

\[ m > 1, \quad 0 < \sigma^2 < \infty \quad \text{and} \quad 0 < \lambda < \infty, \quad (3.1) \]

then there is a positive random variable $V$ such that

\[ \lim_{n \to \infty} m^{-n} X_n = V \quad \text{a.s.} \quad (3.2) \]
This can be viewed as a Strong Law. In this section we will derive limiting distribution results for some functionals of \( \{X_n\} \).

Let \( \mathbb{R}^\infty \) denote the space of real sequences \( x = (x_1, x_2, \ldots) \) with metric

\[
d(x, y) = \sum_{i=1}^{\infty} 2^{-i}|x_i - y_i|/(1 + |x_i - y_i|).
\]

Let \( \mathcal{B} \) be the Borel \( \sigma \)-field generated by \( d \). For each sequence of random variables \( H_n \), consider the random elements \( \Gamma_n = \{y_n\} \) and \( \Gamma = \{y\} \) on \( (\mathbb{R}^\infty, \mathcal{B}) \) defined by

\[
y_{ni} = \begin{cases} 
(X_{n-i+1} - mX_{n-i} - H_{n-i+1})/(X_{n-i} + 1)^{1/2}, & i = 1, \ldots, n, \\
0, & \text{otherwise},
\end{cases}
\]

and

\[
y_n \text{ are i.i.d. random variables distributed as } N(0, \sigma^2).
\]

**Theorem 3.1.** Assume that (3.1) holds and \( E|H_n| = O(m^n/2) \). Then

\[
\Gamma_n \to \Gamma \text{ weakly in } \mathbb{R}^\infty.
\]

**Proof.** It is known (Billingsley, 1968, p. 19) that in \( (\mathbb{R}^\infty, \mathcal{B}) \) (3.5) is equivalent to the weak convergence of \( \{y_{ni}, 1 \leq i \leq \tau\} \) for all integers \( \tau > 0 \). Fix \( \tau > 0 \). Since \( E|I_{n-i+1}| + E|H_{n-i+1}| = O(m^n/2), \) by (3.2), \( (I_{n-i+1} - H_{n-i+1})/(X_{n-i} + 1)^{1/2} = o_p(1) \) for \( 1 \leq i \leq \tau \). Hence we only have to consider the case \( H_n = I_n \) or the case

\[
y_{ni} = I_{1 \leq i \leq n} \sum_{j=1}^{X_{n-i+1}} (Y_{n-i+1,j} - m)/(X_{n-i} + 1)^{1/2}.
\]

Now we can use the same argument as that in Theorem 1 of Heyde and Brown (1971) to complete our proof.

Note that in the proof of Theorem 1 of Heyde and Brown (1971), not only finite dimensional convergence is established but also a tightness result, which is not required in our proof of theorem 3.1. This is due to the fact that they topologize \( \mathbb{R}^\infty \) with the metric \( \rho \), where

\[
\rho(x, y) = \sup_{n \geq 1} \left\{ \left( \sum_{i=1}^{n} (x_i - y_i) \right) / \left[ 1 + \sum_{i=1}^{n} (x_i - y_i)^2 \right] \right\},
\]

and in \( (\mathbb{R}^\infty, \rho) \) finite dimensional Borel subsets do not form a convergence determining class. Their use of the metric \( \rho \) is motivated by the following application. First note that the process \( \{\tilde{X}_n\} \) they consider is a nonimmigration branching process and the random element they study is \( W_n = (W_{n1}, W_{n2}, \ldots) \) with \( W_{nj} = X_{n,j} - X_{n,j-1} \) and \( X_{n,i} = (m^{-1}\tilde{X}_{n+i} - \tilde{X}_n)/(1 + \tilde{X}_n)^{1/2} \). Letting \( W = \lim_{n \to \infty} X_n / m^n \), they use their weak convergence result to derive asymptotic normality for \( m^n(W - m^{-n}\tilde{X}_n)/(1 + \tilde{X}_n)^{1/2} \). This leads them to consider the mapping \( h(x) = \lim \sup_{n \to \infty} \sum_{j=1}^{n} x_j \). Since \( h \) is continuous under \( \rho \), the desired conclusion follows from the continuous
mapping theorem. Although $h$ is not continuous under $d$, the following simple lemma (Billingsley, 1968, p. 28) can be used to replace the tightness arguments and provide results for more general functionals (cf. Corollary 3.3).

**Lemma 3.2.** Assume that for each $n$, random variables $V_n$, $U_{1n}$, $U_{2n}$, ... are defined on the same probability space. Suppose that, for each $j$,

\[ U_{jn} \xrightarrow{d} U_j \quad \text{as } n \to \infty, \]  
\[ U_j \xrightarrow{d} U \quad \text{as } n \to \infty. \]  

and, for each $\varepsilon > 0$,

\[ \lim_{j \to \infty} \lim_{n \to \infty} \sup P(\left| U_{jn} - V_n \right| \geq \varepsilon) = 0. \]  

Then $V_n \xrightarrow{d} U$ as $n \to \infty$.

**Corollary 3.3.** Assume that (3.1) holds and $E|H_n| = o(m^{n/2})$. Let $\{c_i\}$ be a sequence of real numbers such that $\sum_{i=1}^{\infty} |c_i| < \infty$. Then as $n \to \infty$,

\[ \sum_{i=1}^{n} c_i (X_{n-i+1} - mX_{n-i} - H_{n-i+1})/(X_{n-i} + 1)^{1/2} \xrightarrow{d} N\left(0, \sigma^2 \sum_{i=1}^{\infty} c_i^2\right). \]  

**Proof.** It is not difficult to see that

\[ \sum_{i=1}^{n} |c_i| E|H_{n-i+1}|m^{(i-n)/2} = o(1). \]

Hence we only have to consider the case $H_n = I_n$. For each $j$, let

\[ U_{jn} = \sum_{i=1}^{j} c_i \gamma_{ni}, \quad U_j = \sum_{i=1}^{j} c_i \gamma_i \quad \text{and} \quad U = \sum_{i=1}^{\infty} c_i \gamma_i \]

where $\gamma_{ni}$ and $\gamma_i$ are defined by (3.6) and (3.4) respectively. By Theorem 3.1, (3.8) holds. It is obvious that (3.9) holds with $U = N(0, \sigma^2 \sum_{i=1}^{\infty} c_i^2)$. Now observe that for $j \geq n$, $U_{jn} = U_{nn}$ and for $1 \leq j < n$, by Chebyshev's inequality, we have that

\[ P(\left| U_{jn} - U_{nn} \right| > \varepsilon) \leq E(\left| U_{jn} - U_{nn} \right|^2)/\varepsilon^2 = \sigma^2 / \varepsilon^2 \sum_{i=j+1}^{n} c_i^2 E[X_{n-i}/(X_{n-i} + 1)] \]

\[ \leq \sigma^2 / \varepsilon^2 \sum_{i=j+1}^{\infty} c_i^2. \]

Hence (3.10) holds with $V_n = U_{nn}$. Applying Lemma 3.2, we obtain Corollary 3.3.

The following analogue of the Central Limit Theorem was obtained using a direct method by Heyde and Seneta (1971). We derive it as a consequence of our approach.
Corollary 3.4. Assume that (3.1) holds and $b^2 < \infty$. Then as $n \to \infty$,

$$m^n(V - m^{-n}X_n)/(1 + X_{n-1})^{1/2} \overset{d}{\to} N(0, (m - 1)^{-1}\sigma^2).$$

(3.12)

Proof. Observe that

$$m^n(V - m^{-n}X_n) = \lim_{i \to \infty} m^n(m^{-(n+i)}X_{n+i} - m^{-n}X_n)$$

$$= \sum_{i=0}^{\infty} (m^{-i}X_{n+i+1} - m^{-i}X_{n+i})$$

(3.13)

$$= \sum_{i=0}^{\infty} m^{-i-1}(X_{n+i+1} - mX_{n+i} - I_{n+i+1}) + \sum_{i=0}^{\infty} m^{-i}I_{n+i+1}.$$ 

By (3.1), (3.2) and the fact that $E(I_n) = \lambda$,

$$\sum_{i=0}^{\infty} m^{-i-1}I_{n+i+1}/(1 + X_{n-1})^{1/2} = o_n(1).$$

(3.14)

Let $U = N(0, \sigma^2/(m - 1))$ and for each $n$ and $j$, define

$$U_{jn} = \sum_{i=0}^{j} m^{-i-1}(X_{n+i+1} - mX_{n+i} - I_{n+i+1})/(1 + X_{n-1})^{1/2},$$

$$U_j = N(0, \beta_j\sigma^2) \quad \text{with} \quad \beta_j = (1 - m^{-j-1})/(m - 1)$$

and

$$V_n = \sum_{i=0}^{\infty} m^{-i-1}(X_{n+i+1} - mX_{n+i} - I_{n+i+1})/(1 + X_{n-1})^{1/2}.$$ 

Since

$$U_{jn} = \sum_{i=0}^{j} m^{-i-1}y_{n+j-i}[(X_{n+i} + 1)/(X_{n-1} + 1)]^{1/2},$$

where $y_{n,i}$ is defined by (3.6), and in view of (3.2) and Theorem 3.1,

$$U_{jn} \overset{d}{\to} U_j \quad \text{as} \quad n \to \infty.$$ 

It is obvious that

$$U_j \overset{d}{\to} U \quad \text{as} \quad j \to \infty.$$ 

Now,

$$E[U_{jn} - V_n]^2 = \left(\frac{\sigma}{m}\right)^2 \sum_{i=j+1}^{\infty} m^{-2i}E[X_{n+i}/(1 + X_{n-1})]$$

and

$$E(X_{n+i}X_{n-1}) = m^{i+1}X_{n-1} + (m^{i+1} - 1)\lambda/(m - 1).$$
Thus
\[ E|U_n - V_n|^2 \leq \left( \frac{\sigma}{m} \right)^2 \left[ 1 + \lambda/(m-1) \right] \sum_{i=j+1}^{\infty} m^{-i+1} \]
and consequently (3.10) holds. By Lemma 3.2, (3.13) and (3.14), (3.12) follows.

The case \( \alpha = 1 \) of the following theorem is used by Wei and Winnicki (1987) to prove that the conditional least squares estimator of \( m \) has a normal limit law.

**Theorem 3.5.** Assume that (3.1) holds. Then for any \( \alpha > -\frac{1}{2} \), as \( n \to \infty \),
\[
\left[ \sum_{i=1}^{n} (X_{i-1} + 1) \right]^{-\alpha - 1/2} \sum_{i=1}^{n} (X_{i-1} + 1)^\alpha (X_i - mX_{i-1} - \lambda) \xrightarrow{d} N(0, \beta \sigma^2).
\]

(3.15)

where \( \beta = (m - 1)^{2\alpha + 1}/(m^{2\alpha + 1} - 1) \).

**Proof.** By (3.2) we have that
\[
\lim_{n \to \infty} m^{-n} \sum_{i=1}^{n} (X_{i-1} + 1) = \frac{V}{(m-1)} \text{ a.s.}
\]
(3.16)

Therefore, if we can show
\[
\sum_{i=1}^{n} [(X_{i-1} + 1)^{\alpha + 1/2} - (m^{-1} V)^{\alpha + 1/2}] (X_i - mX_{i-1} - \lambda)/(X_{i-1} + 1)^{1/2}
\]
\[
= o_{\mathbb{P}}(m^{n(\alpha + 1/2)}),
\]
(3.17)

then (3.15) is equivalent to
\[
\left[ \frac{m^n}{m - 1} \right]^{-\alpha - 1/2} \sum_{i=1}^{n} m^{(i-1)(\alpha + 1/2)} (X_i - mX_{i-1} - \lambda)/(X_{i-1} + 1)^{1/2} \xrightarrow{d} N(0, \beta \sigma^2),
\]
(3.18)

or
\[
(m - 1)^{\alpha + 1/2} \sum_{i=1}^{n} m^{-(n-i+1)(\alpha + 1/2)} (X_i - mX_{i-1} - \lambda)/(X_{i-1} + 1)^{1/2} \xrightarrow{d} N(0, \beta \sigma^2).
\]
(3.19)

But (3.19) is a direct consequence of Corollary 3.3 with \( H_n = \lambda \) and \( c_n = [(m - 1)m^n]^{-(\alpha + 1/2)} \).

It remains to prove (3.17). By the Cauchy–Schwarz inequality
\[
\left| \sum_{i=1}^{n} [(X_{i-1} + 1)^{\alpha + 1/2} - (m^{-1} V)^{\alpha + 1/2}] (X_i - mX_{i-1} - \lambda)/(X_{i-1} + 1)^{1/2} \right|
\]
\[
\leq A_n^{1/2} B_n^{1/2},
\]
(3.20)
where

\[ A_n = \sum_{i=1}^n m_{i-1}^{(i-1)(\alpha+1/2)} \left[ \left( \frac{X_{i-1} + 1}{m_{i-1}} \right)^{\alpha+1/2} - V^{\alpha+1/2} \right]^2 \]

and

\[ B_n = \sum_{i=1}^n m_{i-1}^{(i-1)(\alpha+1/2)} (X_{i-1} - mX_{i-1} - \lambda)^2/(X_{i-1} + 1). \]

By (3.2)

\[ A_n = o\left( \sum_{i=1}^n m_{i-1}^{(i-1)(\alpha+1/2)} \right) = o(m^{n(\alpha+1/2)}) \quad \text{a.s.} \quad (3.21) \]

Also, \( E(B_n) = O(m^{n(\alpha+1/2)}) \). Hence \( B_n = O_p(m^{n(\alpha+1/2)}) \). From this, (3.21) and (3.20), (3.17) follows.

This completes our proof.

Acknowledgement

The authors would like to thank the referee for suggesting several references and improvements to an earlier version of this paper.

References


B. Mellein, Local limit theorems for the critical Galton-Watson process with immigration, Rev. Colomb. Mat. 16 (1982b) 31-56.


