The Annals of Mathematical Statistics
1970, Vol. 41, No. 1, 195–202

A SIMPLE PROOF OF A RESULT OF KESTEN AND STIGUM ON SUPERCRITICAL MULTITYPE GALTON–WATSON BRANCHING PROCESS[1]

BY KRISHNA B. ATHREYA

University of Wisconsin

0. Summary. Let $\{\mathbf{Z}_n : n \geqq 0\}$ be a supercritical p-type ($p \geqq 2$) Galton-Watson branching process with offspring probability generating functions (pgf) $h_i(\mathbf{s})$ $i = 1, 2, \cdots, p$. Assume (i) $m_{ij} \equiv \partial h_i / \partial s_j |_{s=1} < \infty$ for all i and j where $\mathbf{s} = (s_1, \cdots, s_p)$ and $\mathbf{1} = (1, 1, \cdots, 1)$, (ii) $\exists n_0 > 0 \ni$ if $M \equiv ((m_{ij}))$ then $M^{n_0} \gg 0$ (i.e. each element of M^{n_0} is > 0) and (iii) the largest real eigenvalue ρ of M is > 1. Let $\mathbf{u} \gg 0$ and $\mathbf{v} \gg 0$ be column vectors such that $M\mathbf{v} = \rho\mathbf{v}$, $\mathbf{u}'M = \rho\mathbf{u}'$, $\mathbf{u} \cdot \mathbf{1} = 1$, $\mathbf{u} \cdot \mathbf{v} = 1$ where $\mathbf{u}'$ denotes transpose of $\mathbf{u}$ and $\cdot$ refers to inner product. Kesten and Stigum [6] showed (i) there always exists a nonnegative random variable W such that $\mathbf{Z}_n \rho^{-n}$ converges almost surely (a.s.) to $\mathbf{u}W$ and (ii) $P(W = 0) < 1$ if and only if $E(Z_1{}^j \log Z_1{}^j \mid \mathbf{Z}_0 = e_i) < \infty$ for all i and j where $e_i = (\delta_{i1}, \delta_{i2}, \cdots, \delta_{ip})$, $\delta_{ij} = 1$ if $i = j$ and 0 if $i \neq j$, $Z_1{}^j$ is the jth coordinate of $\mathbf{Z}_1$.

We give here a simple proof of a modified result which is exactly the same as above except that convergence a.s. is replaced by convergence in probability. We do this by showing that without any extra assumption other than the existence of M the vector $(\mathbf{v} \cdot \mathbf{Z}_n)^{-1}\mathbf{Z}_n$ converges in probability to $\mathbf{u}$ on the set of non-extinction.

1. Introduction. Let $(\Omega, \mathscr{F}, P)$ be a probability space on which a Markov chain $\{\mathbf{Z}_n : n = 0, 1, 2, \cdots\}$ with stationary transition probabilities and the nonnegative integer lattice S in p dimension as its state space is defined. For $\mathbf{i} = (i_1, i_2, \cdots, i_p)$ and $\mathbf{j} = (j_1, j_2, \cdots, j_p)$ where i_r and j_r are nonnegative integers let the one-step transition probabilities $P(\mathbf{i}, \mathbf{j})$ satisfy

$$\sum_{j \in S} P(\mathbf{i}, \mathbf{j})\mathbf{s}^j = \mathbf{h}^i(\mathbf{s}) \tag{1}$$

where $\mathbf{h}(\mathbf{s}) = (h_1(\mathbf{s}), \cdots, h_p(\mathbf{s}))$, $\mathbf{s} = (s_1, s_2, \cdots, s_p)$, $0 \leqq s_i \leqq 1$, for $r = 1, 2, \cdots, p$, $\mathbf{s}^j = s_1{}^{j_1} \cdots s_p{}^{j_p}$ and $h_r(\mathbf{s})$ is a probability generating function (pgf) of a random variable with values in S.

The chain $\{\mathbf{Z}_n : n = 0, 1, 2, \cdots\}$ is called a p-type Galton-Watson branching process. In the picturesque language of branching processes (see [4]) $\mathbf{Z}_n$ is the vector that denotes the number of particles of various types at the nth generation starting with $\mathbf{Z}_0$ at the 0th generation. The particles are supposed to breed independently of each other and a type i particle produces particles of all types according to the pgf $h_i(\mathbf{s})$.

Received June 30, 1969.

[1] Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No. DA-31-124-ARO-D-462.

We now assume the following:

(2)
(i) $M_{ij} \equiv \partial h_i / \partial s_j \big|_{s=1} < \infty$ for all i and j where $\mathbf{1} = (1, 1, \cdots, 1)$.
(ii) $\exists n_0 > 0 \ni$ if $M = ((m_{ij}))$ then $M^{n_0} \gg 0$. That is all entries of M^{n_0} are strictly positive.
(iii) Not all $h_r(\mathbf{s})$ are linear in $\mathbf{s}$.

The property (ii) is called positive regularity while (iii) is referred to as non-singularity.

In view of Perron-Frobenius theory of positive matrices (see [5]) there exists an eigenvalue ρ of M, vectors $\mathbf{u}$ and $\mathbf{v}$ such that

(a) $|\lambda| < \rho$ for any eigenvalue λ of M.
(b) $\mathbf{u} \gg 0$, $\mathbf{v} \gg 0$ (i.e. all entries are strictly positive).
(c) $M\mathbf{v} = \rho\mathbf{v}$, $\mathbf{u}'M = \rho\mathbf{u}'$, $\mathbf{u} \cdot \mathbf{1} = 1$, $\mathbf{u} \cdot \mathbf{v} = 1$.

It can be shown (see [4]) that if $\rho \leqq 1$ the process gets extinct a.s. i.e. $P(\mathbf{Z}_n = 0$ for some $n) = 1$ for any initial distribution of $\mathbf{Z}_0$.

We now assume that the process is *supercritical*, i.e. $\rho > 1$ and this ensures (see [4]) that $P(\mathbf{Z}_n = 0$ for some $n) < 1$ for any initial distribution of $\mathbf{Z}_0$ except $P(\mathbf{Z}_0 = \mathbf{0}) = 1$. Since all states of S other than 0 are transient (see [4]) it follows that if $\rho > 1$, $P(\mathbf{Z}_n \to \infty$ as $n \to \infty) = 1 - P(\mathbf{Z}_n = 0$ for some $n) > 0$.

This leads one to the problem of determining the growth rate of $\mathbf{Z}_n$ on the set of non-extinction. H. Kesten and B. Stigum [6] gave the following answer.

THEOREM 1. *Let the probability measure P be nontrivial i.e. let* $P\{\mathbf{Z}_0 = 0\} = 0$. *Let* $(*)$ *stand for*

$$(*) \qquad E(Z_1^{\,j} \log Z_1^{\,j} \mid \mathbf{Z}_0 = e_i) < \infty \quad \text{for all} \quad 1 \leqq i, j \leqq p$$

where $e_i = (\delta_{i1}, \delta_{i2}, \cdots, \delta_{ip})$, $\delta_{ij} = 1$ *if* $i = j$ *and* 0 *if* $i \neq j$.

Then the following holds

$$(3) \qquad (*) \ \textit{false} \Rightarrow \mathbf{Z}_n \rho^{-n} \to 0 \quad \text{a.s.} \qquad \textit{while}$$

$$(4) \qquad (*) \ \textit{true} \Rightarrow \mathbf{Z}_n \rho^{-n} \to \mathbf{u}W \quad \text{a.s.}$$

where W is a nonnegative numerical random variable and satisfies a.s. $E(W \mid \mathbf{Z}_0 = e_i) = v_i$, $P(W = 0 \mid \mathbf{Z}_0 = e_i) = P\{\mathbf{Z}_n = 0$ *for some* $n \mid \mathbf{Z}_0 = e_i\}$.

We shall give a proof of this important result (with a slight weakening namely, in (4) the convergence is in probability instead of a.s.). Our approach is completely different from that of Kesten and Stigum. They construct an auxiliary process $\mathbf{Y}_n$ which is close to $\mathbf{Z}_n$ for which (3) and (4) hold and from this deduce the result for $\mathbf{Z}_n$. The proof gets quite complicated this way.

Our approach, on the other hand, is simpler and more natural for the following reasons. It is essentially the same as in the one-type case namely the use of Lemma 1 of Section 2 below. Further, although we establish only convergence in probability in (4) we prove in Theorem 3 the convergence in probability of $(\mathbf{v} \cdot \mathbf{Z}_n)^{-1}\mathbf{Z}_n$ to $\mathbf{u}$ on the set of non-extinction without any extra assumptions. We then use this and the martingale convergence of $(\mathbf{v} \cdot \mathbf{Z}_n)\rho^{-n}$ to prove Theorem 1. In earlier works (see

[2], [4]) one first proves the convergence of $\mathbf{Z}_n \rho^{-n}$ under assumptions strong enough to ensure that the limit random variable W was not zero and then deduce the convergence of $(\mathbf{v} \cdot \mathbf{Z}_n)^{-1}\mathbf{Z}_n$. It is clear from the above that this is not natural in the sense that convergence of $(\mathbf{v} \cdot \mathbf{Z}_n)^{-1}\mathbf{Z}_n$ depends solely on the fact that on the set of non-extinction $\mathbf{Z}_n \to \infty$ and has nothing to do with the validity of $*$ or any other stronger assumptions.

We now describe in some detail the main steps of our approach. The proofs are given in the subsequent sections. We start with an easily verified result (see [4]).

THEOREM 2. *Let for* $n \geqq 0$, $W_n = \mathbf{v} \cdot \mathbf{Z}_n \rho^{-n}$, $\mathscr{F}_n \equiv$ *sub* σ*-algebra of* $\mathscr{F}$ *generated by* $\mathbf{Z}_0, \mathbf{Z}_1, \cdots, \mathbf{Z}_n$.
Then $\{W_n, \mathscr{F}_n, n \geqq 0\}$ *is a nonnegative martingale and hence*

$$\lim_{n \to \infty} W_n = W \text{ exists a.s.} \tag{5}$$

The next result describes the nature of ratios of Z_n^j to $\mathbf{v} \cdot \mathbf{Z}_n$ on the set of non-extinction.

THEOREM 3. *Let* $A = \{\omega : \mathbf{Z}_n(\omega) \to \infty \quad as \quad n \to \infty\}$ *and*

$$\mathbf{x}_n = (\mathbf{v} \cdot \mathbf{Z}_n)^{-1}\mathbf{Z}_n. \tag{6}$$

Then just assuming (2) *holds we have for any* $\varepsilon > 0$

$$\lim_{n \to \infty} P(\omega : \omega \in A, \|\mathbf{x}_n - \mathbf{u}\| > \varepsilon) = 0 \tag{7}$$

where $\|\cdot\|$ *refers to the usual Euclidean distance.*

We give a proof of Theorem 3 in Section 3.

Next, let for $x \geqq 0$, $\varphi_i(x) = E(e^{-xW} \mid \mathbf{Z}_0 = e_i)$ where e_i is as in $(*)$ and W is defined in (5).

By virtue of (1) (see [4]) we note that $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \cdots, \varphi_p)$ is a solution of the functional equation in $\boldsymbol{\psi}$

$$\boldsymbol{\psi}(x) = \mathbf{h}(\boldsymbol{\psi}(x/\rho)). \tag{8}$$

Let C stand for the set $C = \{\boldsymbol{\psi}(x) = (\psi_1(x), \psi_2(x), \cdots, \psi_p(x))$ where for any i, ψ_i maps $[0, \infty)$ into $(0, 1]$ and $\lim_{x \downarrow 0} x^{-1}(1 - \psi_i(x)) > 0\}$.

We then have the following result which is proved in Section 2.

THEOREM 4. $\varphi \in C$ *if and only if* $(*)$ *holds.*

Noting that $\rho^{-n}\mathbf{Z}_n = W_n \mathbf{x}_n$ we now see that Theorem 1 with convergence a.s. replaced by convergence in probability in (4) is immediate.

2. Existence of a solution to (8) **in** C. We need the following lemmas.

LEMMA 1. *Let* $f(s)$ *be a* pgf *of a nonnegative integer valued random variable* X *with* $0 < m = Ex < \infty$. *Let*

$$A(s) = m - [1 - f(1-s)]s^{-1} \quad \text{for} \quad 0 < s \leqq 1, \tag{9}$$
$$= 0 \qquad\qquad s = 0.$$

Then (i) $A(s)$ *is nonnegative, increasing and continuous in* [0, 1], *and* (ii) *for any* $0 < c, r < 1$,

$$\sum_{n=0}^{\infty} A(cr^n) < \infty \quad \textit{and} \quad \lim_{c \downarrow 0} \sum_{n=0}^{\infty} A(cr^n) = 0 \tag{10}$$

if and only if

$$EX\,|\log X| < \infty. \tag{11}$$

PROOF. See [1].

LEMMA 2. *For each* i, *and* $\mathbf{s} = (s_1, s_2, \cdots, s_p)$, *with* $0 \leqq s_j \leqq 1$ *for all* j *let*

$$h_i(1-s_1, 1-s_2, \cdots, 1-s_p) \equiv 1 - \sum_{j=1}^{p} s_j m_{ij} + A_i(s_1, s_2, \cdots, s_p). \tag{12}$$

If $\mathbf{s} = (s_1, s_2, \cdots, s_p)$ *and* $\mathbf{s}^* = (s_1^*, s_2^*, \cdots, s_p^*)$ *are such that* $s_j \geqq s_j^*$ *for all* j *then* $A_i(\mathbf{s}) \geqq A_i(\mathbf{s}^*)$.

Also the function

$$\bar{A}_i(s) \equiv s^{-1} A_i(s, s, \cdots, s), \qquad 0 < s \leqq 1, \tag{13}$$

is nonnegative, increasing and continuous in (0, 1] *and* $\lim_{s \downarrow 0} \bar{A}_i(s) = 0$.

PROOF. The first part is obvious since $\partial A_i / \partial s_j \geqq 0$ for $\mathbf{s} = (s_1, s_2, \cdots, s_p)$ such that $0 \leqq s_i \leqq 1$. The second part follows easily from Lemma 1 by noting that for $0 < s \leqq 1$

$$\bar{A}_1(s) = [m_i - s^{-1}\{1 - \bar{h}_i(1-s)\}] \tag{14}$$

where

$$m_i = \sum_{j=1}^{p} m_{ij} \qquad \bar{h}_i(s) = E[s^{Z_1{}^1 + Z_1{}^2 + \cdots + Z_1{}^p} \,|\, \mathbf{Z}_0 = \mathbf{e}_i]. \qquad \square$$

PROOF OF THE "ONLY IF" PART OF THEOREM 4. Let us assume $\varphi \in C$ and (*) is false. We shall reach a contradiction. Define $\varphi^*(x)$ for $x \geqq 0$ by

$$\varphi^*(x) = \sum_{i=1}^{p} u_i \varphi_i(x) \tag{15}$$

and $g(x) = x^{-1}(1 - \varphi^*(x))$ for $x > 0$.

Using (8) and (12) we get $1 - \varphi^*(x) = \rho[1 - \varphi^*(x/\rho)] - \sum_{i=1}^{p} u_i A_i(\mathbf{1} - \varphi(x/\rho))$ which yields

$$g(x) = g(x/\rho)\{1 - \rho^{-1}[1 - \varphi^*(x/\rho)]^{-1}[\sum_{i=1}^{p} u_i A_i(\mathbf{1} - \varphi(x/\rho))]\}. \tag{16}$$

Since $\varphi \in C$ there exists constants $c > 0$, $\delta > 0$ such that

$$x \leqq \delta \Rightarrow 1 - \varphi_j(x) \geqq cx \qquad \text{for all } j. \tag{17}$$

By Lemma 2 we get $A_i(\mathbf{1} - \varphi(x/\rho)) \geqq A_i(cx\rho^{-1}\mathbf{1})$. Also for $x > 0$

$$\begin{aligned} 1 - \varphi^*(x/\rho) &\leqq (x/\rho) \sum_{i=1}^{p} u_i E(W \,|\, \mathbf{Z}_0 = e_i) \\ &\leqq (x/\rho) \sum_{i=1}^{p} u_i v_i \qquad \text{(Fatou's lemma)} \\ &= (x/\rho). \end{aligned}$$

Thus for $x \leqq \delta$

$$\rho^{-1}[1-\varphi^*(x/\rho)]^{-1}[\textstyle\sum_{i=1}^p u_i A_i(\mathbf{1}-\boldsymbol{\varphi}(x/\rho))] \geqq c' \sum_{i=1}^p u_i \bar{A}_i(cx\rho^{-1}) \qquad (c' = c\rho^{-1})$$

and using the trivial inequality $1-x \leqq e^{-x}$ for $x > 0$ we obtain for $x \leqq \delta$

$$c \leqq g(x) \leqq \exp[-c' \textstyle\sum_{i=1}^p u_i \bar{A}_i(cx\rho^{-1})]\, g(x/\rho). \tag{18}$$

Iteration of (18) yields for $x \leqq \delta$

$$0 < c \leqq g(x) \leqq \exp[-\textstyle\sum_{i=1}^p u_i \sum_{n=1}^\infty \bar{A}_i(cx\rho^{-n})]. \tag{19}$$

Since (∗) is false there exists an i_0 and j_0 such that

$$E(Z_1^{j_0} \log Z_1^{j_0} \mid \mathbf{Z}_0 = e_{i_0}) = \infty$$

and $x \log x$ being increasing and nonnegative for $x \geqq 1$ this implies

$$E((Z_1^1 + \cdots + Z_1^p) \log (Z_1^1 + \cdots + Z_1^p) \mid \mathbf{Z}_0 = e_{i_0}) = \infty$$

and thus from (14) and Lemma 1 we conclude that $\sum_{n=1}^\infty \bar{A}_{i_0}(cx\rho^{-n}) = \infty$. This with (19) yields the absurd relation $0 < c \leqq g(x) \leqq 0$ and we have reached the needed contradiction.

COROLLARY 1. *If* (∗) *is false then for any* P

$$\mathbf{Z}_n \rho^{-n} \to \mathbf{0} \qquad \text{a.s.} \tag{20}$$

PROOF. Use the preceding result, Theorem 2 and the fact that $v_i > 0$ for all i. □

We now turn to the

PROOF OF THE "IF" PART OF THEOREM 4. Let for $x \geqq 0$

$$\begin{aligned} \varphi_{n,i}(x) &= E(e^{-xW_n} \mid \mathbf{Z}_0 = e_i) \\ \varphi_n^*(x) &= \textstyle\sum_{i=1}^p u_i \varphi_{n,i}(x) \\ g_n(x) &= 1 - x^{-1}[1-\varphi_n^*(x)] \end{aligned} \tag{21}$$

Then as before using (1) (see [4]) we get for $x > 0$, $\boldsymbol{\varphi}_n(x) = \mathbf{h}(\boldsymbol{\varphi}_{n-1}(x/\rho))$ which on using (12) implies

$$0 < g_n(x) = g_{n-1}(x/\rho) + \textstyle\sum_{i=1}^p u_i x^{-1} A_i(\mathbf{1} - \boldsymbol{\varphi}_{n-1}(x/\rho)) \tag{22}$$

where $\boldsymbol{\varphi}_n(x) = (\varphi_{n,1}(x), \varphi_{n,2}(x), \cdots, \varphi_{n,p}(x))$.

Observe that for any n

$$x^{-1}(1-\varphi_{n,i}(x)) \leqq E(W_n \mid Z_0 = e_i) = v_i \leqq \bar{v} \equiv \max_{1 \leqq i \leqq p} v_i.$$

Lemma 2 now yields

$$x^{-1} A_i(1 - \boldsymbol{\varphi}_n(x/\rho)) \leqq \bar{v}\rho^{-1} \bar{A}_i(\bar{v}x\rho^{-1})$$

and hence from (22) we have

(23) $$0 < g_n(x) \leqq g_1(x) + \sum_{i=1}^p u_i \bar{v} \rho^{-1} \sum_{r=1}^{n-1} \bar{A}_i(\bar{v} x \rho^{-r}).$$

Noticing that $x \log x$ is convex for $x \geqq 1$ and that if $x_i \geqq 1$ for all i

$$(x_1 + \cdots + x_p) \log (x_1 + \cdots + x_p) \leqq [\sum_1^p x_i \log x_i + (\log p) \sum_{i=1}^p x_i]$$

we see that $(*)$ implies

$$E((Z_1{}^1 + \cdots + Z_1{}^p) \log (Z_1{}^1 + \cdots + Z_1{}^p) \mid \mathbf{Z}_0 = e_i) < \infty \qquad \text{for all } i.$$

Now appealing to Lemma 1 and letting $n \to \infty$ first and then $x \downarrow 0$ we get from (23) that $\lim_{x \downarrow 0} \{1 - x^{-1}[1 - \varphi^*(x)]\} = 0$.

But since $1 - x^{-1}[1 - \varphi^*(x)] = \sum_{i=1}^p u_i \{v_i - x^{-1}[1 - \varphi_i(x)]\}$ it follows that

(24) $$\lim_{x \downarrow 0} \{v_i - x^{-1}[1 - \varphi_i(x)]\} = 0$$

thus proving that $\varphi \in C$. □

COROLLARY 2. *If* $(*)$ *holds then for every* i

$$E(W \mid \mathbf{Z}_0 = e_i) = v_i \qquad \textit{and}$$

$$P(W = 0 \mid \mathbf{Z}_0 = e_i) = P(\mathbf{Z}_n = 0 \textit{ for some } n \mid \mathbf{Z}_0 = e_i).$$

PROOF. The first part is precisely (24) while the second part follows by letting $x \to \infty$ in (8) and noting that $P(W = 0 \mid \mathbf{Z}_0 = e_i) < 1$ for all i. □

3. Convergence of $x_n \equiv \mathbf{Z}_n(\mathbf{v} \cdot \mathbf{Z}_n)^{-1}$ on the set of non-extinction. We now prove Theorem 3. We shall use the following result on positive matrices. (For a proof see [5].)

LEMMA 3. *If* $K \equiv \{\mathbf{x} = (x_1, x_2, \cdots, x_p), x_i > 0, \sum_{i=1}^p x_i v_i = 1\}$ *then*

$$\lim_{m \to \infty} \sup_{\mathbf{x} \in K} \|\mathbf{x} M^m \rho^{-m} - \mathbf{u}\| = 0$$

where M, ρ, $\mathbf{v}$, $\mathbf{u}$ *and* $\|\cdot\|$ *are as defined in Section* 1.

PROOF OF THEOREM 3.[2] Exploiting the basic feature of Galton-Watson branching processes namely that lines of descent of different particles are stochastically independent we can write

(25) $$\mathbf{Z}_{n+m} = \sum_{i=1}^p \sum_{l=1}^{Z_n{}^i} \mathbf{Z}_m^{(il)}$$

where for each $i \mathbf{Z}_m{}^{(il)}$ for $l = 1, 2, \cdots, Z_n{}^i$ are independent copies of $\mathbf{Z}_m$ when $\mathbf{Z}_0 = e_i$ and the sets of random variables $\{\mathbf{Z}_m{}^{(il)} : l = 1, 2, \cdots, Z_n{}^i\}$ for $i = 1, 2,$

[2] This proof is a joint work of Dr. T. G. Kurtz and the author.

$\cdots, p$, are conditionally (given $\mathbf{Z}_n$) are independent. Dividing both sides of (25) on the set $A = \{\omega : \mathbf{Z}_n(\omega) \to \infty\}$ by $\mathbf{v} \cdot \mathbf{Z}_{n+m}$ we get

$$\mathbf{x}_{n+m} = [\mathbf{x}_n M^m + \textstyle\sum_{i=1}^p x_n{}^i (Z_n{}^i)^{-1} \sum_{l=1}^{Z_n{}^i} (\mathbf{Z}_m^{(il)} - e_i{}' M^m)] \times [\rho^m + \textstyle\sum_{i=1}^p x_n{}^i (Z_n{}^i)^{-1} \sum_{l=1}^{Z_n{}^i} (\mathbf{v} \cdot \mathbf{Z}_m{}^{(il)} - \rho^m v_i)]^{-1}$$

where $\mathbf{x}_n = (\mathbf{v} \cdot \mathbf{Z}_n)^{-1} \mathbf{Z}_n$ and $x_n{}^i$ is the ith coordinate of $\mathbf{x}_n$. Thus

$$\mathbf{x}_{n+m} - \mathbf{u} = [(\mathbf{x}_n M^m \rho^{-m} - \mathbf{u}) + A_{n,m}] \times [1 + R_{n,m}]^{-1}$$

where

(26)
$$A_{n,m} = \textstyle\sum_{i=1}^p x_n{}^i (Z_n{}^i)^{-1} \sum_{l=1}^{Z_n{}^i} \rho^{-m} (\mathbf{Z}_m{}^{(il)} - e_i{}' M^m) - u R_{n,m},$$
$$R_{n,m} = \textstyle\sum_{i=1}^p x_n{}^i (Z_n{}^i)^{-1} \sum_{l=1}^{Z_n{}^i} \rho^{-m} (\mathbf{v} \cdot \mathbf{Z}_m{}^{(il)} - v_i \rho^m)$$

Let $\varepsilon > 0$ be arbitrary. By Lemma 3 there exists an m such that $\sup_{\mathbf{x} \in K} \|\mathbf{x} M^m \rho^{-m} - \mathbf{u}\| < \varepsilon$. Fix this m. Then by weak law of large numbers both the following hold. For any $\eta > 0$

(27)
$$\lim_{n \to \infty} P\{\omega : \omega \in A, |R_{n,m}| > \eta\} = 0$$
$$\lim_{n \to \infty} P\{\omega : \omega \in A, \|A_{n,m}\| > \eta\} = 0.$$

From (26) we have

(28)
$$\|\mathbf{x}_{n+m} - \mathbf{u}\| \leqq [\varepsilon + \|A_{n,m}\|] \times [1 + R_{n,m}]^{-1}$$

and so

$$P\{\omega : \omega \in A, \|\mathbf{x}_{n+m} - \mathbf{u}\| \leqq (\varepsilon + \eta)(1 - \eta)^{-1}\}$$
$$\geqq 1 - P\{\omega : \omega \in A, |R_{n,m}| > \eta\} - P\{\omega : \omega \in A, \|A_{n,m}\| > \eta\}.$$

This with (27) implies for any $\varepsilon > 0, \eta > 0$,

$$\limsup_{N \to \infty} P\{\omega : \omega \in A, \|\mathbf{x}_N - \mathbf{u}\| > (\varepsilon + \eta)(1 - \eta)^{-1}\} = 0,$$

which proves the theorem. □

4. Some remarks.

(a) Although it has not been done here we suspect one can strengthen (27) to assert that for each fixed m almost surely on A the following hold

$$\lim_{n \to \infty} |R_{n,m}| = 0, \qquad \lim_{n \to \infty} \|A_{n,m}\| = 0.$$

This with (28) will imply the almost sure convergence of $\mathbf{x}_n$ to $\mathbf{u}$.

(b) Kesten and Stigum [6] had also shown that when (*) holds there exists for any nontrivial P (provided not all $h_i(\mathbf{s})$ are degenerate), a continuous function $w(x)$ for $x > 0$ such that

$$P(x_1 \leqq W \leqq x_2) = \textstyle\int_{x_1}^{x_2} w(x)\, dx, \qquad 0 < x_1 < x_2 < \infty.$$

We have a different and we believe simpler proof of this fact too. For this and a complete discussion of related aspects of supercritical multitype branching processes we refer the reader to a forthcoming book by Peter Ney and the author [3].

REFERENCES

[1] Athreya, K. B. (1969). On the supercritical one dimensional age dependent branching processes. *Ann. Math. Statist.* **39** 743–763.

[2] Athreya, K. B. and Karlin, S. (1967). Limit theorems for the split times of branching processes. *J. Math. Mech.* **17** 257–278.

[3] Athreya, K. B. and Ney, P. (1970). *Branching Processes*, to appear.

[4] Harris, T. E. (1963). *The Theory of Branching Processes*, Springer-Verlag, Berlin.

[5] Karlin, S. (1966). *A First Course in Stochastic Processes.* Academic Press, New York.

[6] Kesten, H. and Stigum, B. (1966). A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.* **37** 1211–1223.