A LIMIT THEOREM FOR MULTIDIMENSIONAL GALTON-WATSON PROCESSES[1]

By H. Kesten and B. P. Stigum

Cornell University

1. Introduction. In this paper we consider a positively regular, nonsingular, vector-valued Galton-Watson process. Specifically we consider a temporally homogeneous, k-vector-valued Markov chain, $\{Z_n ; n = 0, 1, \cdots\}$, with among others the following properties:

1. Z_0 is taken to be one of the vectors,

$$e_i = (\delta_{i,1}, \cdots, \delta_{i,k}), \qquad 1 \leqq i \leqq k;$$

2. if P denotes the probability measure of the process, if $Z_n = (Z_n^1, \cdots, Z_n^k)$, $n = 0, 1, \cdots$, and if for each n, $F_{i,j}(x) = P\{Z_{n+1}^j \leqq x \mid Z_n = e_i\}$, $1 \leqq i, j \leqq k$; $x \geqq 0$, then Z_n^j, $1 \leqq j \leqq k$, $0 \leqq n < \infty$, takes on only non-negative integer values and

$$P\{Z_{n+1}^j \leqq x \mid Z_0, \cdots, Z_n\} = F_{1,j}^{Z_n^1} * F_{2,j}^{Z_n^2} * \cdots * F_{k,j}^{Z_n^k}(x),$$

where the right hand side is the convolution of Z_n^i times $F_{i,j}$ for $i = 1, \cdots, k$;

3. if E denotes the expectation functional, if $m_{i,j} = E\{Z_1^j \mid Z_0 = e_i\}$, $1 \leqq i, j \leqq k$, and if M denotes the matrix $(m_{i,j})$, then

$$m_{i,j} = \int_0^\infty x \, dF_{i,j}(x) < \infty, \qquad 1 \leqq i, j \leqq k; \tag{1.1}$$

and there exists a finite positive integer t such that

$$(M^t)_{i,j} > 0, \qquad 1 \leqq i, j \leqq k; \tag{1.2}$$

4. if ρ denotes the largest positive characteristic root associated with M, then

$$\rho > 1. \tag{1.3}$$

We will prove a limit theorem for these processes that we state succinctly below. In the statement of this theorem u and v will be positive right and left eigenvectors of M corresponding to ρ, normalized such that their inner product is 1. (For the existence and properties of ρ, u, and v see our comments below and for a more detailed description of Galton-Watson processes see Chapter II of [3]).

Theorem. *There exists a random vector W and a onedimensional random variable w such that*

$$\lim_{n\to\infty} (Z_n/\rho^n) = W \quad \textit{with probability } 1, \tag{1.4}$$

Received 3 February 1966.

[1] Financial support for this research from the National Science Foundation is gratefully acknowledged.

and

$$W = w \cdot v \quad \textit{with probability } 1. \tag{1.5}$$

Also one has either

$$E\{w \mid Z_0 = e_i\} = u_i, \qquad 1 \leqq i \leqq k, \tag{1.6}$$

or

$$w = 0 \quad \textit{with probability } 1. \tag{1.7}$$

Moreover (1.6) *holds if and only if*

$$E\{Z_1^j \log Z_1^j \mid Z_0 = e_i\} < \infty \quad \textit{for all} \quad 1 \leqq i, j \leqq k. \tag{1.8}$$

Finally if $Z_0 = e_i$, $1 \leqq i \leqq k$, *if* (1.8) *holds, and if there is at least one* j_0, $1 \leqq j_0 \leqq k$, *such that*

(1.9) $\sum_{i=1}^{k} Z_1^i u_i$ *can take at least two values with positive probability, given* $Z_0 = e_{j_0}$,

then the distribution of w *has a jump of magnitude* q_i *at the origin and a continuous density function on the set of positive real numbers.* (*The constants* q_i *will be defined later.*) *If* (1.9) *fails to hold for all* $1 \leqq j_0 \leqq k$, *then the distribution of* w *is concentrated on one point.*[2]

To give the reader a better perspective we point out that (1.5) and (1.6) were proved from (1.8) in the onedimensional case by Levinson [6] and in the multidimensional case from stronger conditions than (1.8) by Harris [2]. The fact that w either satisfies (1.6) or (1.7) and that (1.8) is necessary and sufficient for (1.6) is new even in the onedimensional case. The fact that we have here a necessary and sufficient condition seems to be the main novelty of our result, even though it could already have been obtained by sharpening Levinson's argument. Also the absolute continuity of the distribution of w in the onedimensional case was proved in [6] (see also [7]). Finally we point out that while our proof in the case of multidimensional processes is somewhat involved, the proof of the theorem for onedimensional processes is quite simple.

In the proof of our theorem we will make heavy use of certain important properties of positive matrices which are subsumed under the Perron-Frobenius theorem, [5], Section 2 of Appendix. Specifically, the matrix M has a positive eigenvalue, denoted here by ρ, that is simple and exceeds all other eigenvalues in absolute value. Moreover there exist row vectors u and v with positive components which are eigenvectors of M corresponding to ρ; i.e.

$$vM = \rho v, \qquad v_i > 0,\ 1 \leqq i \leqq k, \tag{1.10a}$$

[2] Our argument below shows indirectly that if $q_i > 0$ for some i, then (1.9) must hold for some j_0.

and

$$Mu' = \rho u', \qquad u_i > 0,\ 1 \leqq i \leqq k, \tag{1.10b}$$

where u' denotes the transpose of u. We assume throughout that u and v are normalized so that $vu' = 1$. Finally

$$((M^n)_{i,j}/\rho^n) - u_i v_j = O(|\rho_1/\rho|^n) \quad \text{as} \quad n \to \infty, \tag{1.11}$$

for some $|\rho_1| < \rho$.

The proof of our theorem will be obtained in several distinct steps. We begin by constructing an auxiliary process, $\{Y_n\,; n = 0, 1, \cdots\}$, that is not appreciably different from the Z-process. Thereafter we use this process to show firstly that if (1.8) does not hold, then (1.4), (1.5), and (1.7) must hold, and secondly that if (1.8) is satisfied, then (1.4), (1.5), and (1.6) must hold. Lastly, we establish several important properties of the characteristic function of W that enable us to give a simple proof that the distribution of w has the required properties.

2. Proof of the Theorem. To construct the Y-process mentioned above we proceed as follows: We observe first that it is well known (see [3], p. 49) that $T_n = (1/\rho^n)\sum_{j=1}^k Z_n^j u_j$, $n = 0, 1, \cdots$, is a non-negative martingale and that

$$E\{(1/\rho^n)Z_n^j \mid Z_0 = e_i\} = ((M^n)_{i,j}/\rho^n) \to u_i v_j \quad \text{as} \quad n \to \infty. \tag{2.1}$$

Hence T_n converges with probability one and since each $u_j > 0$ we can for each $\delta > 0$ find an $A = A(\delta)$ such that

$$P\{|Z_n| \geqq A\rho^n \text{ for any } n\} \leqq (\delta/2), \tag{2.2}$$

where as usual $|Z_n|^2 = \{\sum_{j=1}^k (Z_n^j)^2\}$. For an arbitrarily chosen $\delta > 0$ and an A satisfying (2.2) we shall now define an auxiliary Markov chain Y_n, $n = 0, 1, \cdots$, by truncating Z_n. Of course $Y_r = (Y_r^1, \cdots, Y_r^k)$ is also a k-vector and Y_r^j is the number of particles of type j in the rth generation of the Y-process. The definition of the Y_n's is obtained by induction. Let B be a large positive number to be chosen later, let $Y_0 = Z_0$, and assume that Y_r has already been defined for a given $r \geqq 0$. We then define Y_{r+1} as follows. Consider any particle in the rth generation of the Y-process. If it has less than $B\rho^r$ descendants of type j in the $(r+1)$st generation, these all survive. If, however, the number of descendants of type j is $B\rho^r$ or more, these descendants are all killed off. This process is carried out for each type of descendants of each particle in the rth generation separately. We now let $Y_{r+1}^{i,j}$ denote the total number of particles of type j surviving in the $(r+1)$st generation which descend from a particle of type i in the rth generation of the Y-process and define Y_{r+1}^j to be the total number of particles of type j surviving in the $(r+1)$st generation. Thus $Y_{r+1}^j = \sum_{i=1}^k Y_{r+1}^{i,j}$, $j = 1, \cdots, k$.

From the preceding description of the Y-process we can deduce immediately the following facts:

$$Y_r^j \leqq Z_r^j, \qquad 1 \leqq j \leqq k, \qquad \text{with probability one;} \tag{2.3}$$

$$P\{\text{any particle is killed off in the } (r+1)\text{st generation} \mid Y_r\} \leqq |Z_r| \sum_{1 \leqq i,j \leqq k} \int_{B\rho^r}^{\infty} dF_{i,j}(x); \quad \text{and} \tag{2.4}$$

$$E\{Y_{r+1}^{i,j} \mid Y_r\} = Y_r^{\,i} \int_0^{B\rho^r} x\, dF_{i,j}(x), \qquad 1 \leqq i, j \leqq k. \tag{2.5}$$

Thus if we introduce the matrix

$$M(r) = (m_{i,j}(r)) = \left(\int_0^{B\rho^r} x\, dF_{i,j}(x)\right), \tag{2.6}$$

then it follows from the fact that $Y_0, Y_1, \cdots,$ is a Markov chain that

$$E\{Y_{r+1} - Y_r M(r) \mid Y_0, \cdots, Y_r\} = 0, \qquad r \geqq 0. \tag{2.7}$$

In particular, for each fixed r_0, the sequence of random variables

$$\sum_{r=r_0}^{n} (Y_{r+1} - Y_r M(r)), \qquad n = r_0, r_0 + 1, \cdots,$$

is a martingale.

Next we will show that we can choose B so large that the Y-process is not appreciably different from the Z-process.

Lemma 1. *For each $\delta > 0$, B can be chosen so large that*

$$P\{Y_n \neq Z_n \text{ for any } n \geqq 0\} \leqq \delta. \tag{2.8}$$

Proof. Let $\delta > 0$ be fixed and let $A = A(\delta)$ be as defined above. Then $P\{Y_n \neq Z_n \text{ for any } n\} \leqq P\{|Z_n| \geqq A\rho^n \text{ for any } n\} + \sum_{n=0}^{\infty} P\{Y_r = Z_r \text{ for } 0 \leqq r \leqq n, |Z_n| < A\rho^n, Y_{n+1} \neq Z_{n+1}\} \leqq (\delta/2) + \sum_{n=0}^{\infty} P\{|Z_n| < A\rho^n$ and some particle in the nth generation has at least $B\rho^n$ descendants of one type in the $(n + 1)$st generation$\}$

$$\leqq (\delta/2) + \sum_{n=0}^{\infty} A\rho^n \sum_{1 \leqq i,j \leqq k} \int_{B\rho^n}^{\infty} dF_{i,j}(x)$$
$$\leqq (\delta/2) + A \sum_{1 \leqq i,j \leqq k} \int_B^{\infty} dF_{i,j}(x) \left(\sum_{B\rho^n \leqq x} \rho^n\right)$$
$$\leqq (\delta/2) + [A\rho/B(\rho - 1)] \sum_{1 \leqq i,j \leqq k} \int_B^{\infty} x\, dF_{i,j}(x).$$

It follows from this observation and from (1.1) that we can find a B so large that the inequality, (2.8), is satisfied. Q.E.D.

Having shown that the Y-process represents a good approximation of the Z-process, we will proceed to give a sufficient condition that $w = 0$ with probability one.

Lemma 2. *If for some pair, (i_0, j_0), $1 \leqq i_0, j_0 \leqq k$,*

$$\int_1^{\infty} x \log x\, dF_{i_0,j_0}(x) = \infty, \tag{2.9}$$

then $\lim_{n\to\infty} (Z_n/\rho^n) = 0$ *with probability one.*

Proof. It suffices to show that (2.9) implies that

$$\lim_{r\to\infty} E\{(Y_r/\rho^r) \mid Z_0 = e_i\} = 0. \tag{2.10}$$

For then (Y_r/ρ^r) converges to 0 in probability and because of (2.8) (Z_r/ρ^r) converges to 0 in probability as well. Since we already know that $(1/\rho^r) \sum_{i=1}^{k} Z_r^{\,i} u_i$ converges with probability one and that $u_i > 0$, $Z_r^{\,i} \geqq 0$, $1 \leqq i \leqq k$, $r \geqq 0$, it will then follow that $\lim_{r\to\infty} (Z_r^{\,j}/\rho^r) = 0$ with probability one, $1 \leqq j \leqq k$.

To show that (2.10) holds we observe that $M(s) = M - \epsilon(s)$, where $\epsilon(s)$ is the matrix with entries

$$\epsilon(s)_{i,j} = \int_{B\rho^s}^{\infty} x\, dF_{i,j}(x), \qquad 1 \leqq i, j \leqq k.$$

Clearly, $0 \leqq \epsilon(s)_{i,j} \leqq m_{i,j}$, $\epsilon(s)_{i,j} \to 0$ as $s \to \infty$ for all $1 \leqq i, j \leqq k$, and thus by (1.2) we can find a positive constant C such that

$$\begin{aligned}((M - \epsilon(r))(M - \epsilon(r+1)) \cdots (M - \epsilon(r+2t)))_{i,j} \\ \leqq m_{i,j}^{2t+1} - m_{i,i_0}^{t}\epsilon_{i_0,j_0}(r+t)m_{j_0,j}^{t} \\ \leqq m_{i,j}^{2t+1}(1 - C\epsilon_{i_0,j_0}(r+t)) \leqq m_{i,j}^{2t+1} \exp\{-C\epsilon_{i_0 j_0}(r+t)\}, \quad 1 \leqq i, j \leqq k.\end{aligned}$$

Hence for all N, $1 \leqq N \leqq r$,

$$\begin{aligned}(1/\rho^r)E\{Y_r^j \mid Z_0 = e_i\} &= (1/\rho^r)[M(0)M(1) \cdots M(r-1)]_{i,j} \\ &\leqq (1/\rho^r)[M^N M(N)M(N+1) \cdots M(r-1)]_{i,j} \\ &\leqq (1/\rho^r)(M^r)_{i,j} \\ &\quad \cdot \exp\{-C\textstyle\sum_{s=0}^{[(r-N-2t-1)/2t+1]} \epsilon_{i_0,j_0}(N + t + s(2t+1))\}.\end{aligned} \tag{2.11}$$

Since

$$\begin{aligned}\textstyle\sum_{r=0}^{\infty} \epsilon_{i_0,j_0}(r) &= \textstyle\sum_{r=0}^{\infty} \int_{B\rho^r}^{\infty} x\, dF_{i_0,j_0}(x) \\ &= \textstyle\int_B^{\infty} x\, dF_{i_0,j_0}(x)(\sum_{B\rho^r \leqq x} 1) \geqq \lambda \int_B^{\infty} x \log x\, dF_{i_0,j_0}(x)\end{aligned}$$

for suitable $\lambda = \lambda(B) > 0$, (2.9) implies that there exists an integer a such that $\sum_{s=0}^{\infty} \epsilon_{i_0,j_0}(N + t + s(2t+1)) = \infty$ whenever $N \equiv a \pmod{2t+1}$. This together with (1.11) and (2.11) in turn implies (2.10) and the proof is completed. Q.E.D.

Next we will use the Y-process to show that if (1.8) is satisfied, then (1.4) and (1.5) must hold as well.

LEMMA 3. (1.8) *implies* (1.4) *and* (1.5).

PROOF. To prove this lemma we begin by observing that the sequence of random variables, $|\sum_{r=r_0}^{n}((Y_{r+1} - Y_r M(r))/\rho^{r+1})|^2$, $n = r_0, r_0 + 1, \cdots$, form a sub-martingale (or semi-martingale). This is an immediate consequence of (2.7). From this observation it follows (see [1], Th. VII, 3.2) that

$$\begin{aligned}P\{\max_{r_0 \leqq n \leqq r_1} |\textstyle\sum_{r=r_0}^{n}(Y_{r+1} - Y_r M(r))/\rho^{r+1}| > \epsilon\} \\ \leqq (1/\epsilon^2)E\{|\textstyle\sum_{r=r_0}^{r_1}(Y_{r+1} - Y_r M(r))/\rho^{r+1}|^2\} \\ = (1/\epsilon^2)\textstyle\sum_{r=r_0}^{r_1} E\{|(Y_{r+1} - Y_r M(r))/\rho^{r+1}|^2\}.\end{aligned} \tag{2.12}$$

Moreover since

$$\begin{aligned}E\{|Y_{r+1} - Y_r M(r)|^2 \mid Y_r\} &= \textstyle\sum_{j=1}^{k} E\{(Y_{r+1}^j - \sum_{i=1}^{k} Y_r^i m_{i,j}(r))^2 \mid Y_r\} \\ &= \textstyle\sum_{j=1}^{k} E\{(\sum_{i=1}^{k}(Y_{r+1}^{i,j} - Y_r^i m_{i,j}(r)))^2 \mid Y_r\} \\ &= \textstyle\sum_{i,j=1}^{k} E\{(Y_{r+1}^{i,j} - Y_r^i m_{i,j}(r))^2 \mid Y_r^i\} \\ &\leqq \textstyle\sum_{i,j=1}^{k} Y_r^i \int_0^{B\rho^r} x^2\, dF_{i,j}(x),\end{aligned} \tag{2.13}$$

we deduce from (2.3), (2.1), and (1.1) that

$$\begin{aligned}(2.14)\qquad &\textstyle\sum_{r=0}^{\infty} E\{|(Y_{r+1} - Y_r M(r))/\rho^{r+1}|^2 \mid Z_0\} \\ &\leqq \textstyle\sum_{i,j=1}^{k} \sum_{r=0}^{\infty} E\{Y_r^{\,i} \mid Z_0\}(1/\rho^{2r+2}) \int_0^{B\rho^r} x^2\, dF_{i,j}(x) \\ &= O(\textstyle\sum_{i,j=1}^{k} |Z_0| \int_B^{\infty} x^2\, dF_{i,j}(x) \sum_{B\rho^r \geqq x} (1/\rho^{r+2})) \\ &= O(|Z_0| \textstyle\sum_{i,j=1}^{k} \int_B^{\infty} x^2\, dF_{i,j}(x)(\min(B/x, 1))) < \infty .\end{aligned}$$

Thus, as $r_0 \to \infty$, the right and hence left hand side of (2.12) tends to zero.

This implies that $\sum_{r=0}^{\infty}((Y_{r+1} - Y_r M(r))/\rho^{r+1})$, converges with probability one. (This convergence can also be derived from the fact that each component of $\sum_{r=r_0}^{n}(Y_{r+1} - Y_r M(r))/\rho^{r+1}$, $n = r_0, r_0 + 1, \cdots$, is a martingale.)

Next we will show that the series

$$(2.15)\qquad \textstyle\sum_{r=0}^{\infty}((Y_{r+1} - Y_r M)/\rho^{r+1}),$$

also converges with probability one. This assertion is an immediate consequence of (1.8) and the following inequality which holds (see (2.2) and (2.3)) with probability one:

$$\begin{aligned}(2.16)\qquad &\textstyle\sum_{r=0}^{\infty} |(Y_r M(r) - Y_r M)/\rho^{r+1}| \\ &\leqq \sup_{r\geqq 0} |(Y_r/\rho^{r+1})| \textstyle\sum_{r=0}^{\infty} \sum_{i,j=1}^{k} |m_{i,j}(r) - m_{i,j}| \\ &\leqq \sup_{r\geqq 0} |(Y_r/\rho^{r+1})| \textstyle\sum_{i,j=1}^{k} \sum_{r=0}^{\infty} \int_{B\rho^r}^{\infty} x\, dF_{i,j}(x) \\ &\leqq \sup_{r\geqq 0} |(Y_r/\rho^{r+1})| \textstyle\sum_{i,j=1}^{k} \int_B^{\infty} x\, dF_{i,j}(x)(\sum_{B\rho^r \leqq x} 1) \\ &= O(\sup_{r\geqq 0} |(Y_r/\rho^{r+1})| \textstyle\sum_{i,j=1}^{k} \int_B^{\infty} x \log x\, dF_{i,j}(x)) < \infty .\end{aligned}$$

The preceding result can be used to show that there exists a random vector, $\tilde{W}$, and a random variable $\tilde{w}$, such that

$$(2.17)\qquad \lim_{r\to\infty}(Y_r/\rho^r) = \tilde{W} \quad \text{with probability one,}$$

and such that

$$(2.18)\qquad \tilde{W} = \tilde{w}\cdot v \quad \text{with probability one.}$$

In view of (2.8), (2.17) and (2.18) will then imply (1.4) and (1.5).

To establish the a.e. convergence of the (Y_r/ρ^r)'s we observe that (1.11) implies that for any given $r_0 < r_1$,

$$\begin{aligned}(2.19)\qquad &\textstyle\sum_{r=r_0}^{r_1}((Y_{r+1}M^{r_1-r} - Y_r M^{r_1-r+1})/\rho^{r_1+1}) \\ &= \textstyle\sum_{r=r_0}^{r_1}((Y_{r+1} - Y_r M)/\rho^{r+1})(M^{r_1-r}/\rho^{r_1-r}) \\ &= \textstyle\sum_{r=r_0}^{r_1}((Y_{r+1} - Y_r M)/\rho^{r+1})\cdot u'v \\ &\quad + O(\sup_{r\geqq r_0} |(Y_{r+1} - Y_r M)/\rho^{r+1}|\, |\rho_1/\rho|^{r_0}).\end{aligned}$$

The convergence of the series (2.15), and $|\rho_1| < \rho$ imply that the right hand side of (2.19) tends to zero as r_0 tends to ∞, $r_1 \geqq r_0$. But the left hand side of (2.19) telescopes and we obtain

$$\lim_{r_0\to\infty, r_1 \geqq r_0} (Y_{r_1+1}/\rho^{r_1+1}) - (Y_{r_0}/\rho^{r_0})(M^{r_1-r_0+1}/\rho^{r_1-r_0+1}) = 0$$

with probability one. Finally when invoking (1.11) once more we find that

$$\text{(2.20)} \qquad \lim_{r_0\to\infty, r_1-r_0\to\infty} (Y_{r_1+1}/\rho^{r_1+1}) - (Y_{r_0}/\rho^{r_0})\cdot u'v = 0$$

with probability one. Clearly (2.17) and (2.18) are immediate consequences of (2.20). Q.E.D.

We will next conclude the proof of the first half of our theorem by showing that (1.8) implies (1.6).

LEMMA 4. (1.8) *implies* (1.6).

PROOF. To establish the proof of this lemma it will suffice to show that for each fixed $\delta > 0$ we can find a B so large that

$$\text{(2.21)} \quad E\{\lim_{r\to\infty} (Y_r^j/\rho^r) \mid Z_0 = e_i\} \geqq (1-\delta)u_i v_j, \quad \text{for all} \quad 1 \leqq i, j \leqq k,$$

since then by (2.1), Lemma 3, Fatou's lemma, and (2.3),

$$\lim_{r\to\infty} E\{Z_r^j/\rho^r \mid Z_0 = e_i\} = u_i v_j \geqq E\{\lim_{r\to\infty} (Z_r^j/\rho^r) \mid Z_0 = e_i\}$$
$$\geqq E\{\lim_{r\to\infty} (Y_r^j/\rho^r) \mid Z_0 = e_i\} \geqq (1-\delta)u_i v_j .$$

To show that (1.8) implies (2.21) we will first show that the (Y_r/ρ^r)'s are uniformly integrable. This fact is an immediate consequence of the following inequality which in turn can be deduced from (2.7), (1.11) and (2.14).

$$\begin{aligned}
&E\{|(Y_{r_1+1}/\rho^{r_1+1}) - (Y_{r_0}/\rho^{r_0})M(r_0)\cdots M(r_1)/\rho^{r_1-r_0+1}|^2 \mid Z_0\} \\
\text{(2.22)} \quad &= E\{|\textstyle\sum_{r=r_0}^{r_1} [(Y_{r+1} - Y_r M(r))/\rho^{r+1}]M(r+1)\cdots M(r_1)/\rho^{r_1-r}|^2 \mid Z_0\} \\
&= \textstyle\sum_{r=r_0}^{r_1} E\{|[(Y_{r+1} - Y_r M(r))/\rho^{r+1}]M(r+1)\cdots M(r_1)/\rho^{r_1-r}|^2 \mid Z_0\} \\
&\leqq C_1 \textstyle\sum_{r=r_0}^{r_1} E\{|(Y_{r+1} - Y_r M(r))/\rho^{r+1}|^2 \mid Z_0\} \leqq C_2 < \infty
\end{aligned}$$

for certain constants C_1 and C_2 that are independent of r_0 and r_1. In view of (1.11) and $0 \leqq m_{i,j}(r) \leqq m_{i,j}$ for all r and all $1 \leqq i, j \leqq k$ this implies that $E\{|(Y_r/\rho^r|^2 \mid Z_0\}$ is bounded, which is even stronger than uniform integrability.

Next we observe that (1.2) and (1.11) together with the facts $\epsilon(s)_{i,j} \geqq 0$ and $\epsilon(s)_{i,j} \to 0$ as $B \to \infty$ uniformly in s, imply that we can find a positive constant C_3 such that with $C_4 = \max_{i,j} (C_3/(M^t)_{i,j})$ and B sufficiently large,

$$\begin{aligned}
&[(M-\epsilon(r))(M-\epsilon(r+1))\cdots(M-\epsilon(r+t-1))]_{i_0,j_0} \\
&\qquad \geqq (M^t)_{i_0,j_0} - C_3 \textstyle\sum_{1\leqq i,j\leqq k}\sum_{s=r}^{r+t-1} \epsilon_{i,j}(s) \\
&\qquad \geqq (M^t)_{i_0,j_0}[1 - C_4 \textstyle\sum_{1\leqq i,j\leqq k}\sum_{s=r}^{r+t-1} \epsilon_{i,j}(s)] \\
&\qquad \geqq \exp\{-2C_4 \textstyle\sum_{1\leqq i,j\leqq k}\sum_{s=r}^{r+t-1} \epsilon_{i,j}(s)\}(M^t)_{i_0,j_0}
\end{aligned}$$

and hence

$$\begin{aligned}
&[M(0)\cdots M(mt-1)]_{i_0,j_0} \\
\text{(2.23)} \quad &\geqq \exp\{-2C_4 \textstyle\sum_{1\leqq i,j\leqq k}\sum_{s=0}^{t-1} \epsilon_{i,j}(s)\}[M^t M(t)\cdots M(mt-1)]_{i_0,j_0} \\
&\geqq \cdots \geqq \exp\{-2C_4 \textstyle\sum_{1\leqq i,j\leqq k}\sum_{s=0}^{mt-1} \epsilon_{i,j}(s)\}(M^{mt})_{i_0,j_0} .
\end{aligned}$$

Finally it follows from (2.7), (2.23), and (1.11) that

$$\begin{aligned} E\{Y^{j_0}_{mt}/\rho^{mt} \mid Z_0 = e_{i_0}\} &= [M(0)M(1) \cdots M(mt-1)]_{i_0,j_0} \\ &\geqq \exp\{-2C_4 \textstyle\sum_{1\leqq i,j\leqq k}\sum_{s=0}^{\infty} \epsilon_{i,j}(s)\}(M^{mt})_{i_0,j_0}/\rho^{mt} \\ &\geqq \exp\{-2C_4 \textstyle\sum_{1\leqq i,j\leqq k}\sum_{s=0}^{\infty} \epsilon_{i,j}(s)\}u_{i_0}v_{j_0}(1 + O(|\rho_1/\rho|^{mt})). \end{aligned} \tag{2.24}$$

Since under (1.8)

$$\begin{aligned} \textstyle\sum_{s=0}^{\infty} \epsilon_{i,j}(s) &= \textstyle\sum_{s=0}^{\infty} \int_{B\rho^s}^{\infty} x\, dF_{i,j}(x) = \int_B^{\infty} x\, dF_{i,j}(x)(\sum_{B\rho^s \leqq x} 1) \\ &= O(\textstyle\int_B^{\infty} x \log x\, dF_{i,j}(x)) = o(1) \quad \text{as} \quad B \to \infty, \end{aligned}$$

we conclude that for sufficiently large B

$$\liminf_{m\to\infty} E\{Y^{j_0}_{mt}/\rho^{mt} \mid Z_0 = e_{i_0}\} \geqq (1-\delta)u_{i_0}v_{j_0}.$$

This result together with the uniform integrability of the (Y_r/ρ^r)'s implies (2.21). Q.E.D.

The results obtained in Lemmas 1 through 4 can be conveniently summarized as follows:

If conditions, (1.1)–(1.3), are satisfied, then (1.4) and (1.5) must hold. Moreover, the limiting variable w satisfies either (1.6) or (1.7). Finally w satisfies (1.6) if and only if (1.8) holds.

Having established the a.e. convergence of the (Z_r/ρ^r)'s and the properties of the mean value of the limiting variable w, we shall now proceed to show that if (1.8) holds and if there is at least one j_0, $1 \leqq j_0 \leqq k$, such that (1.9) holds, then for $Z_0 = e_i$, $1 \leqq i \leqq k$, the distribution of w has a jump of magnitude q_i at the origin and a continuous density function on the set of positive real numbers. The constant q_i is the ith component of the vector q, defined in (2.29) below.

To prove this part of our theorem we introduce some more notation. Let $s = (s_1, \cdots, s_k)$ be a complex k-vector with $|s_i| \leqq 1$, $1 \leqq i \leqq k$, and let $r = (r_1, \cdots, r_k)$ run through k-vectors with $r_i \geqq 0$, r_i integer, $1 \leqq i \leqq k$. Then we define the moment generating function, $f_n(s) = (f_n^1(s), \cdots, f_n^k(s))$ by

$$f_n^i(s) = \textstyle\sum_r P\{Z_n = r \mid Z_0 = e_i\}s^r, \qquad 1 \leqq i \leqq k, n \geqq 1,$$

where s^r stands for $s_1^{r_1}s_2^{r_2} \cdots s_k^{r_k}$. For real k-vectors, $t = (t_1, \cdots, t_k)$, $\varphi(t) = (\varphi^1(t), \cdots, \varphi^k(t))$ will denote the characteristic function of W; i.e.

$$\varphi^j(t) = E\{\exp i \textstyle\sum_{m=1}^k t_m W^m \mid Z_0 = e_j\}, \qquad 1 \leqq j \leqq k.$$

Similarly we let

$$\varphi_n^j(t) = E\{\exp i \textstyle\sum_{m=1}^k t_m(Z_n^m/\rho^n) \mid Z_0 = e_j\}, \qquad 1 \leqq j \leqq k, n \geqq 0,$$

and let $\varphi_n(t) = (\varphi_n^1(t), \cdots, \varphi_n^k(t))$. It is easily shown that

$$\varphi_n(t) = f(\varphi_{n-1}(t/\rho)). \tag{2.25}$$

(Compare [3], Equation I.8.6.) Since (Z_n/ρ^n) converges to W, it follows that $\varphi_n(t) \to \varphi(t)$ as $n \to \infty$. Hence

$$\varphi(t) = f(\varphi(t/\rho)). \tag{2.26}$$

Moreover since $W = w \cdot v$ with probability one, it is also true that

$$\varphi^j(t) = g^j(\textstyle\sum_{i=1}^k t_i v_i), \qquad 1 \leqq j \leqq k, \tag{2.27}$$

where $g^j(t) = E\{\exp itw \mid Z_0 = e_j\}$.

To establish the fact that the distribution of w has a jump at the origin and that its magnitude is exactly q_i if $Z_0 = e_i$, $1 \leqq i \leqq k$, we begin by recording the following properties of the $f_n(\cdot)$'s which can be found in [3], Chapter II:

$$f_{n+k}(s) = f_n(f_k(s)), \qquad n, k \geqq 1; \tag{2.28}$$

there exists a unique vector, $q = (q_1, \cdots, q_k)$, with $0 \leqq q_i < 1$, $1 \leqq i \leqq k$, for which

$$f(q) = q. \tag{2.99}$$

As in [3], p. 14, Remark 1, this implies that

$$P\{W = 0 \mid Z_0 = e_i\} = q_i ; \qquad 1 \leqq i \leqq k; \tag{2.30}$$

for if q_i^* denotes the left hand side of (2.30), one must have $q^* = f(q^*)$ and since $E\{w \mid Z_0 = e_i\} = u_i > 0$, $1 \leqq i \leqq k$, by (1.6), $q_i^* < 1$ for each i. Thus q^* equals the unique solution q of (2.29).

Next we will show that for some $\delta > 0$,

$$|\varphi^i(t)| < 1, \qquad 1 \leqq i \leqq k, \tag{2.31}$$

whenever

$$|t| = (\textstyle\sum_{i=1}^k |t_i|^2)^{\frac{1}{2}} \leqq \delta, \qquad \sum_{i=1}^k t_i v_i \neq 0. \tag{2.32}$$

To establish this fact we will need to use assumption (1.9). Clearly (1.9) implies the existence of two vectors, r^1 and r^2 such that

$$\textstyle\sum_{i=1}^k r_i^1 u_i \neq \sum_{i=1}^k r_i^2 u_i \tag{2.33}$$

and such that

$$P\{Z_1 = r^1 \mid Z_0 = e_{j_0}\} > 0 \tag{2.34a}$$

and

$$P\{Z_1 = r^2 \mid Z_0 = e_{j_0}\} > 0. \tag{2.34b}$$

On the other hand, since as is easily shown from (1.6),

$$E\{w \mid Z_0 = e_{j_0}, Z_1 = r\} = (1/\rho)E\{w \mid Z_0 = r\} = (1/\rho) \textstyle\sum_{i=1}^k r_i u_i,$$

(2.33), and (2.34a and b) imply that the conditional distribution of w given $Z_0 = e_{j_0}$ cannot be concentrated on one point. As is well known ([4], Section 14, Theorem 2) this in turn implies that for sufficiently small $t \neq 0$, $|g^{j_0}(t)| < 1$ and hence that (2.31) holds for $i = j_0$ and sufficiently small t with $\sum_{j=1}^k t_i v_i \neq 0$. To prove (2.31) for all i we observe that

$$\varphi^i(t) = f_n^{\,i}(\varphi(t/\rho^n)), \qquad 1 \leqq i \leqq k, \tag{2.35}$$

because of (2.26) and (2.28). Also by (1.2) there must exist an n with

$$P\{Z_n^{\,j_0} > 0 \mid Z_0 = e_i\} > 0 \quad \text{for all} \quad i.$$

Since

$$\begin{aligned} |\varphi^i(t)| &= |\textstyle\sum_r P\{Z_n = r \mid Z_0 = e_i\}\varphi(t/\rho^n)^r| \\ &\leqq \textstyle\sum_{r_1 \geqq 0, \cdots, r_k \geqq 0} P\{Z_n = r \mid Z_0 = e_i\}|\varphi^1(t/\rho^n)|^{r_1} \cdots |\varphi^k(t/\rho^n)|^{r_k} \\ &< 1 \end{aligned} \tag{2.36}$$

if there is a vector r with positive j_0th component, such that $P\{Z_n = r \mid Z_0 = e_i\} > 0$ and if t is sufficiently small so that $|\varphi^{j_0}(t)| < 1$, we can conclude from the preceding remarks that there exists a $\delta > 0$ such that (2.32) implies (2.31).

Our assumption (1.9) played an essential role in establishing the existence of a neighborhood around the origin in which the absolute values of the characteristic functions of w are different from one everywhere except at the origin itself. Before continuing our proof of the absolute continuity of the probability distribution of w on the set of positive real numbers we want to point out here that if (1.9) fails to hold, then the distribution of w is concentrated at one single point. To establish this fact we assume that there exist constants b_i, $1 \leqq i \leqq k$, such that if $Z_0 = e_i$, $\sum_{j=1}^k Z_1^{\,j}u_j = b_i$ with probability one. Then $b_i = \rho u_i$ and for all $n > 1$, $\sum_{j=1}^k Z_n^{\,j}u_j = \rho^n u_i$ with probability one if $Z_0 = e_i$, $1 \leqq i \leqq k$. Indeed, since $E\{\sum_{j=1}^k Z_1^{\,j}u_j \mid Z_0 = e_i\} = \rho u_i$, $1 \leqq i \leqq k$, the first part of the assertion is obvious. As for the second part for each pair, $1 \leqq i, j \leqq k$, let $Z_n^{\,i,j}$ denote the total number of particles of type j descending from a particle of type i in the $(n-1)$st generation. Then, given Z_{n-1},

$$\textstyle\sum_{j=1}^k Z_n^{\,j}u_j = \sum_{i=1}^k \sum_{j=1}^k Z_n^{\,i,j}u_j = \sum_{i=1}^k Z_{n-1}^i b_i = \rho \sum_{i=1}^k Z_{n-1}^i u_i$$

with probability one.

Our assertion now follows by induction.

We return to the proof of the absolute continuity of the distribution of w if (1.9) holds. Using (2.31) we shall show that the functions $g^j(\cdot)$ have absolutely integrable derivatives. For this we shall need the following lemma.

LEMMA 5. *Let*

$$m_{i,j}(s) = \partial f^i(s)/\partial s_j \quad \textit{and} \quad m_{i,j}(n; s) = \partial f_n^{\,i}(s)/\partial s_j, \qquad 1 \leqq i, j \leqq k.$$

Then for each $0 < b < 1$ *there exists a constant* C_5 *and a* λ, $0 < \lambda < 1$, *such that for all* $1 \leqq i, j \leqq k$, and all $n \geqq 1$,

$$m_{i,j}(n; s) \leqq C_5\lambda^n \quad \textit{whenever} \quad |s_1| \leqq b, \cdots, |s_k| \leqq b. \tag{2.37}$$

PROOF. Since $f_n(\cdot)$ has positive coefficients, we can only increase $m_{i,j}(n; s)$

when replacing $s_1, \cdots, s_k$ by $|s_1|, \cdots, |s_k|$. Thus we may assume $0 \leqq s_i \leqq b$, $1 \leqq i \leqq k$. It follows from the relation, $f_n(s) = f(f_{n-1}(s))$ that

$$M(n; s) = M(f_{n-1}(s))M(n-1; s), \tag{2.38}$$

where we have written $M(n; s)$ for the matrix with entries $m_{i,j}(n; s)$ and $M(\cdot)$ for the matrix with entries $m_{i,j}(\cdot)$. To study the behavior of $M(n; s)$ we begin by showing that $M(q)$ has a non-negative eigenvalue $\tilde{\rho} < 1$ which has the largest absolute value among all the eigenvalues of $M(q)$. Since $M(q)$ has non-negative entries only, the existence of a non-negative eigenvalue of largest absolute value is well known (c.f. Theorem 2.4 in Appendix of [5]). Hence we need only show that $\tilde{\rho} < 1$. In the case when $M(q) = 0$, which is not included in [5], Theorem 2.4 of Appendix, $\tilde{\rho} = 0$ and the remarks below become mostly trivial. If the entries of $M(q)$ are not all zero, we observe that (2.38) and the relation, $f_n(q) = q$, $n = 1, 2, \cdots$, imply that $M(n; q) = M(q)^n$ for all n. Hence $M(n; q)$ has the eigenvalue, $(\tilde{\rho})^n$. On the other hand the largest eigenvalue of $M(n; q)$ cannot exceed

$$k \max_{i,j} m_{i,j}(n; q) \tag{2.39}$$

(see p. 476 of [5]). However

$$m_{i,j}(n; q) = \sum_{r \neq 0} r_j P\{Z_n = r \mid Z_0 = e_i\} q_1^{r_1} \cdots q_j^{r_j - 1} \cdots q_k^{r_k},$$

and for each $r \neq 0$,

$$P\{Z_n = r \mid Z_0 = e_i\} \to 0 \quad \text{as} \quad n \to \infty, \qquad 1 \leqq i \leqq k, \tag{2.40}$$

(see [3], sect. II.6). By (2.40), $q_i < 1$, $1 \leqq i \leqq k$, and the dominated convergence theorem $\lim_{n\to\infty} m_{i,j}(n; q) = 0$, $1 \leqq i, j \leqq k$, and thus by (2.39) $(\tilde{\rho})^n \to 0$ as $n \to \infty$. This proves our statement that $\tilde{\rho} < 1$.

We return to the study of $M(n; s)$. Because eigenvalues are continuous functions of the matrix entries, there exists a $\delta > 0$ such that $0 < q_i + \delta < 1$, $1 \leqq i \leqq k$, and such that λ, the largest eigenvalue of $M(q + \delta)$ satisfies $0 < \lambda < 1$ (of course $q + \delta = (q_1 + \delta, \cdots, q_k + \delta)$). But it is easy to show (see proof of Theorem II.7.2, p. 42 of [3]) that $\lim_{n\to\infty} f_n(s) = q$ uniformly in $|s_1| \leqq b, \cdots, |s_k| \leqq b$. Consequently if we take into account that $m_{i,j}(s)$, $1 \leqq i, j \leqq k$, is non-decreasing in each argument, we can find an N such that for $n \geqq N$ and $0 \leqq s_1 \leqq b, \cdots, 0 \leqq s_k \leqq b$

$$m(f_{n-1}(s))_{i,j} \leqq m_{i,j}(q + \delta), \qquad 1 \leqq i, j \leqq k.$$

By iteration of (2.38) we finally obtain

$$M(n; s) = M(f_{n-1}(s)) \cdots M(f_{N+1}(s))M(N + 1; s) \leqq M(q + \delta)^{n-N-1} M(N + 1; s),$$

where the inequality is meant to hold for each entry of the matrices and for $0 \leqq s_i \leqq b$, $1 \leqq i \leqq k$. (2.37) now follows from $(M(q + \delta)^n)_{i,j} = O(\lambda^n)$ as

$n \to \infty$, which in turn is a consequence of the Perron-Frobenius theorem. (To see this observe that Theorem 2.3 of the Appendix of [5] is applicable because $q_i + \delta > 0$, $1 \leqq i \leqq k$, together with (1.2) shows that $(M(q + \delta)^t)_{i,j} > 0$ for all i, j.)

LEMMA 6. *If* (1.8) *and* (1.9) *hold, then* $(d/dt)g^j(t)$ *exists, is bounded and continuous and* $\int_{-\infty}^{\infty} |(d/dt)g^j(t)|\, dt < \infty$, *for all* j, $1 \leqq j \leqq k$.

PROOF. For each j, $1 \leqq j \leqq k$, $g^j(\cdot)$ certainly has a bounded, continuous derivative since $E\{w \mid Z_0 = e_j\} < \infty$. To show that the derivative of $g^j(\cdot)$ is also absolutely integrable we proceed as follows. Let $a > 0$ be chosen so small that

$$b^i = \sup_{a \leqq t_1, \cdots, t_k \leqq \rho a} |\varphi^i(t)| < 1, \quad \text{for all} \quad i, \qquad 1 \leqq i \leqq k. \tag{2.41}$$

Such an a exists by (2.31). From (2.35) and (2.27) we deduce that

$$g^j(tv_1\rho^n) = \varphi^j(t\rho^n, 0, \cdots, 0) = f_n^{\ j}(\varphi(t, 0, \cdots, 0)),$$

which after differentiation with respect to t gives

$$v_1\rho^n g'(tv_1\rho^n) = \sum_{i=1}^k (\partial f_n^{\ j}(s)/\partial s_i)_{s=\varphi(t,0,\cdots,0)}(\partial\varphi^i(t, 0, \cdots 0)/\partial t),$$

where we have written $g'(u) = (d/du)g^j(u)$. Consequently for a suitable constant C_6,

$$\begin{aligned}\int_0^\infty |g'(u)|\, du &= \int_0^{av_1} |g'(u)|\, du + \sum_{n=0}^\infty \int_{av_1\rho^n}^{av_1\rho^{n+1}} |g'(u)|\, du \\ &= \int_0^{av_1} |g'(u)|\, du + \sum_{n=0}^\infty v_1\rho^n \int_a^{a\rho} |g'(v_1\rho^n t)|\, dt \\ &\leqq C_6\Big(1 + \int_a^{a\rho} \Big(\sum_{n=0}^\infty \sum_{i=1}^k |(\partial f_n^{\ j}(s)/\partial s_i)_{s=\varphi(t,0,\cdots,0)}|\Big)\, dt.\end{aligned} \tag{2.42}$$

On the other hand for $a \leqq t \leqq \rho a$ one has by (2.41) and Lemma 5

$$|(\partial f_n^{\ j}(s)/\partial s_i)_{s=\varphi(t,0,\cdots,0)}| \leqq C_5\lambda^n.$$

Consequently the integrand in the last member of (2.42) is bounded and thus $\int_0^\infty |g'(u)|\, du < \infty$. Since we can treat the integral over $u \leqq 0$ in a similar fashion, the proof of the lemma is complete. Q.E.D.

When using the preceding result, we can give a simple proof of the next lemma which represents the last step in the proof of our theorem.

LEMMA 7. *If* (1.8) *and* (1.9) *hold, then for* $Z_0 = e_j$ *the conditional distribution of* w *has a jump of magnitude* q_j *at the origin and has a continuous density on the set of positive real numbers.*

PROOF. The jump at the origin was already determined in (2.30). To establish the existence of a density function we define

$$g_T(x) = (1/2\pi) \int_{-T}^{+T} e^{-itx} g^j(t)\, dt.$$

Integration by parts shows that for $x \neq 0$,

$$g_T(x) = (-1/2\pi i x)\{e^{-iTx}g^j(T) - e^{iTx}g^j(-T)\} + h_T(x),$$

where $h_T(x) = (1/2\pi i x) \int_{-T}^{+T} e^{-itx} g'(t)\, dt$. In view of Lemma 6, $h_T(x)$ converges boundedly to a continuous function on $x > 0$, $h(x)$, as $T \to \infty$. Also

(2.43) $$g_+ = \lim_{T\to\infty} g^j(T) = g^j(0) + \lim_{T\to\infty} \int_0^T g'(t)\, dt$$

as well as

(2.44) $$g_- = \lim_{T\to\infty} g^j(-T)$$

exist. But then by the inversion formula, for continuity points $x_1, x_2, 0 < x_1 < x_2$, of the distribution of w,

$$\begin{aligned} P\{x_1 < w \leqq x_2 \mid Z_0 = e_j\} &= \lim_{T\to\infty} \int_{-T}^{+T} (e^{-itx_2} - e^{-itx_1}/-2\pi it) dx g^j(t)\, dt \\ &= \lim_{T\to\infty} \int_{x_1}^{x_2} g_T(x)\, dx = \int_{x_1}^{x_2} h(x) \\ &\quad + \lim_{T\to\infty} \int_{x_1}^{x_2} (-1/2\pi ix)[e^{-iTx} g^j(T) - e^{iTx} g^j(-T)]\, dx. \end{aligned}$$

But,

(2.45) $$\begin{aligned} \int_{x_1}^{x_2} (1/2\pi ix) e^{-iTx} g^j(T)\, dx &= \int_{x_1}^{x_2} (e^{-iTx}/2\pi ix) g_+\, dx \\ &\quad + \int_{x_1}^{x_2} (e^{-iTx}/2\pi ix)(g^j(T) - g_+)\, dx, \end{aligned}$$

and as $T \to \infty$ both terms on the right hand side of (2.45) tend to zero, the first term by the Riemann-Lebesgue lemma and the second term by (2.43). We conclude by applying the same argument to the integral

$$\int_{x_1}^{x_2} (e^{iTx}/2\pi ix) g^j(-T)\, dx$$

that

$$P\{x_1 < w \leqq x_2 \mid Z_0 = e_j\} = \int_{x_1}^{x_2} h(x)\, dx.$$

This completes the proof of the lemma. Q.E.D.

REFERENCES

[1] Doob, J. L. (1953). *Stochastic processes*. Wiley, New York.

[2] Harris, T. E. (1951). *Some mathematical models for branching processes. Second Berkeley Symp. Math. Statist. Prob.* 305–328.

[3] Harris, T. E. (1963). *The theory of branching processes*. Springer Verlag, Berlin.

[4] Gnedenko, B. V. and Kolmogorov, A. N. (1954). *Limit distributions for sums of independent random variables*. Addison-Wesley, Cambridge, Mass.

[5] Karlin, S. (1966). *A first course in stochastic processes*. Academic Press, New York.

[6] Levinson, N. (1959). Limiting theorems for Galton-Watson branching processes. *Illinois J. Math.* **3** 554–565.

[7] Stigum, B. P. (1966). A theorem on the Galton-Watson process. *Ann. Math. Statist.* **37** 695–698.