The Annals of Mathematical Statistics
1968, Vol. 39, No. 2, 347–357

SOME RESULTS ON MULTITYPE CONTINUOUS TIME MARKOV BRANCHING PROCESSES[1]

By Krishna Balasundaram Athreya

Stanford University

1. Introduction. Of late there has been a lot of interest in multitype continuous time Markov branching processes(MCMBP). It was recently noted that there is a fundamental and yet simple connection between classical urn schemes like Polya's and Friedman's, etc. and multitype continuous time Markov branching processes (see [1]). Problems on urn schemes have their counterparts in branching processes and it is while attempting to solve these that the author felt the need for a systematic study of the MCMBP. The present paper is a partial answer to this need (see also [1], [2], [3]).

In this paper we develop in a systematic way some basic properties of multitype continuous time Markov branching process. Although a few of these properties are elementary and not entirely new in content we present them here for the sake of completeness. But there are two very important results which we believe are new. We prove them under minimal assumptions. (See Section 2 for notations and preliminaries.) Let $\{X(t); t \geqq 0\}$ be a MCMBP and let A be the infinitesimal generator of the mean matrix semigroup $\{M(t); t \geqq 0\}$ where $M(t) = ((m_{ij}(t)))$ and $m_{ij}(t) = ((E(X_j(t) \mid X_r(0) = \delta_{ri}, r = 1, 2, \cdots, k)))$ and δ_{ij} are Kronecker deltas. Assuming positive regularity and nonsingularity of the process we establish the following:

Theorem 1. *Needing nothing more than the existence of the first moments, we have*

$$\lim_{t\to\infty} X(t, \omega)e^{-\lambda_1 t} = W(\omega)u \qquad \textit{exists} \qquad \text{wp } 1$$

where $W(\omega)$ is a nonnegative numerical valued random variable, λ_1 is the maximal real eigenvalue of A and u is an appropriately normalized vector satisfying $u^*A = \lambda_1 u^*$ where u^* is the transpose of u.

Theorem 4. *Let $\lambda_1 > 0$ so that $P\{X(t) = 0 \text{ for some } t\} < 1$ for any nontrivial makeup and assume second moments exist. Then*

$$\lim_{t\to\infty} P\{0 < x_1 \leqq W \leqq x_2 < \infty, (v\cdot X(t) - e^{\lambda_1 t}W)(v\cdot X(t))^{-\frac{1}{2}} \leqq y\} = P\{0 < x_1 \leqq W \leqq x_2 < \infty\}\Phi(y/\sigma)$$

where v is an appropriately normalized vector satisfying $Av = \lambda_1 v$, σ^2 an appropriate constant and $\Phi(x)$ is the normal distribution function.

Here is an outline of the rest of the paper. In Section 2 we describe our set up and construct some martingales. Section 3 establishes Theorem 1. Assuming the

Received 18 October 1967.

[1] Research supported in part under contracts N0014-67-A-0112-0015 and NIH USPHS 10452 at Stanford University.

existence of second moments we study in Section 4 the growth behavior of $E|\xi \cdot X(t)|^2$ where ξ is an eigenvector of the matrix A. The last section develops a representation for $X(t) - e^{\lambda_1 t} Wu$ and uses it to prove Theorem 4.

This paper forms a minor part of the author's doctoral thesis at the Department of Mathematics, Stanford. The author is much indebted to his adviser Professor S. Karlin for help and encouragement.

2. The set up, mean matrix and martingales.

2.1. *The set up.* We start with a strong Markov, continuous time k dimensional $(2 \leqq k < \infty)$ branching process $\{X(t); t \geqq 0\}$ defined on a probability space $(\Omega, \mathfrak{F}, P)$. That is, $\{X(t, \omega); t \geqq 0\}$ is a stochastic process on $(\Omega, \mathfrak{F}, P)$ such that

(i) The state space is the nonnegative integer lattice in k dimensions.

(ii) It is a Markov chain with stationary transition probabilities and strong Markov with respect to the family $\mathfrak{F}_t$ of σ-algebras where $\mathfrak{F}_t = \sigma\{X(u, w); u \leqq t\}$ and $\sigma\{D\}$ stands for the sub σ-algebra of $\mathfrak{F}$ generated by the family D of real random variables on $(\Omega, \mathfrak{F})$

(iii) The transition probabilities $P_{\mathrm{i},\mathrm{j}}(t)$ satisfy

$$\left(\sum_{\mathrm{j}} P_{\mathrm{i},\mathrm{j}}(t) s^{\mathrm{j}}\right) = \prod_{r=1}^{k} \left(\sum_{\mathrm{j}} P_{e_i,\mathrm{j}}(t) s^{\mathrm{j}}\right)^{i_r}$$

where $\quad \mathrm{i} = (i_1, i_2, \cdots, i_k), \qquad e_i = (\delta_{i1}, \delta_{i2}, \cdots, \delta_{ik})$

$$\delta_{ij} = 1 \quad \text{if} \quad i = j$$
$$= 0 \quad \text{if} \quad i \neq j.$$

Let the associated infinitesimal generating functions be

$$(1) \qquad u_i(s) = a_i[h_i(s) - s_i] \qquad \text{for} \quad i = 1, 2, \cdots, k,$$

where $0 < a_i < \infty$, $s = (s_1, s_2, \cdots, s_k)$, $0 \leqq s_i \leqq 1$ and $h_i(s)$ is a probability generating function. We make the following basic assumption.

ASSUMPTION 1.

$$(2) \qquad \partial h_i(s)/\partial s_j|_{s=(1,\cdots,1)} < \infty \quad \textit{for all} \quad i \quad \textit{and} \quad j.$$

Under (2) it can be proved using either the theory of branching Markov processes developed by Ikeda, Nagasawa and Watanabe [9] or the theory of minimal processes as in Chung [5] that a branching process $\{X(t); t \geqq 0\}$ of the above specification exists. Also from the same theory one could take the sample paths to be right continuous in t wp 1.

It is very suggestive to think of $X(t) = (X_1(t), \cdots, X_k(t))$ as the vector denoting the sizes of the population at time t in a system with k types of particles where

(i) a type i particle lives an exponential length of time with mean a_i^{-1} and on death creates particles of all types according to a distribution whose generating function is given by $h_i(s)$,

(ii) all particles engender independent lines of descent. This property is the basic feature of a branching process.

A moment's reflection is enough to justify the following representation of $X(t+u, \omega)$ as

(3a) $$X(t+u, \omega) = \sum_{i=1}^{k} \sum_{j=1}^{X_i(t,\omega)} X^{ij}(u, \omega)$$

where $\{X^{ij}(u, \omega), u \geqq 0\}$ is the vector denoting the line of descent from the jth particle of type i living at time t. Without loss of generality (see Chapter 6 in [8]) one can assume Ω and $\mathfrak{F}$ to be "big" enough to make $X^{ij}(u, \omega)$ a $\mathfrak{F}$-measurable function and conditionally (conditioned on $X(t)$) mutually independent.

In terms of generating functions (3) becomes

(3b) $$f(s, t+u) = f(f(s, u), t)$$

where

(3c) $$f(s, t) = (f_1(s, t), f_2(s, t), \cdots, f_k(s, t))$$

and

$$f_i(s, t) = E(s_1^{X_1(t)} s_2^{X_2(t)} \cdots s_k^{X_k(t)} \mid X(0) = e_i).$$

One relates the $u_i(s)$ to $f(s, t)$ as follows: The Kolmogorov forward and backward differential equations for $P_{i,j}(t)$ lead to the following in terms of $f(s, t)$.

(3d) (forward) $$\partial f_i(s, t)/\partial t = \sum_{j=1}^{k} (\partial f_i(s, t)/\partial s_j) u_j(s), \quad i = 1, 2, \cdots, k,$$

(3e) (backward) $$\partial f_i(s, t)/\partial t = u_i(f(s, t)), \quad i = 1, 2, \cdots, k,$$

the initial conditions for both the systems being $f_i(s, 0) = s_i$, $i = 1, 2, \cdots, k$.

2.2. *Mean matrix.* From (2) it follows (see Chapter 5 in [8]) that

(4) $$m_{ij}(t) = E(X_j(t) \mid X(0) = e_i)$$

is finite for all $0 \leqq t < \infty$, i and j.

If we set

(5) $$M(t) = ((m_{ij}(t)))_{k \times k}, \quad \text{for} \quad t \geqq 0,$$

we get, using (3b), the semigroup property

(6) $$M(t+u) = M(t)M(u) \quad \text{for all} \quad t, u \geqq 0.$$

We also have using (3d) or (3e) the continuity condition

(7) $$\lim_{t \to 0} M(t) = I.$$

Let

(8) $$b_{ij} = \partial h_i(s)/ds_j|_{s=(1,1,\cdots,1)} - \delta_{ij},$$

$$a_{ij} = a_i b_{ij}, \qquad A = ((a_{ij}))_{k \times k} \quad \text{for} \quad 1 \leqq i, j \leqq k.$$

Then it is well known that $M(t)$ has the representation

(9) $$M(t) = e^{At} \quad \text{for} \quad t \geqq 0.$$

We next make the basic assumption of *positive regularity,* viz., that there exists $t_0 > 0$ and finite such that

$$(10) \qquad m_{ij}(t_0) > 0 \qquad \text{for all } \; i, j.$$

By Frobenius-Perron theory of positive matrices (see appendix in [10]), there exists a strictly positive eigenvalue $\rho_1(t_0)$ of $M(t_0)$ whose algebraic and geometric multiplicities are both one and any other eigenvalue $\rho(t_0)$ of $M(t_0)$ satisfies $|\rho(t_0)| < \rho_1(t_0)$. This implies that the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_k$ of A can be taken to be such that λ_1 is real, $\rho_1(t_0) = e^{\lambda_1 t_0}$ and

$$(11) \qquad \lambda_1 > \operatorname{Re} \lambda_i \qquad \text{for } \; i \neq 1.$$

Further if u and v satisfy

$$(12) \qquad u^* A = \lambda_1 u^*, \qquad Av = \lambda_1 v,$$

then they also satisfy

$$(13) \qquad u^* M(t_0) = \rho_1(t_0) u^*, \qquad M(t_0) v = \rho_1(t_0) v.$$

By the Frobenius-Perron theory of positive matrices u and v can be taken to satisfy

$$(14) \qquad u_i > 0, v_j > 0 \quad \text{for all } \; i, j, \qquad \textstyle\sum_{i=1}^k u_i v_i = 1.$$

2.3. *Martingales.* For any collection D of random variables on $(\Omega, \mathfrak{F}, P)$ we denote by $\sigma(D)$ the smallest σ-algebra contained in $\mathfrak{F}$, with respect to which all members of D are measurable.

Set

$$(15) \quad \mathfrak{F}_t = \sigma\{X(u, \omega); u \leqq t\}, \quad Y(t) = X(t) e^{-At}, \quad Y(t) = (Y_1(t), \cdots, Y_k(t)).$$

Then we have the following generalization of the well known martingale result of the simple branching process.

PROPOSITION 1. *For every* $i = 1, 2, \cdots, k$, *the family* $\{Y_i(t); \mathfrak{F}_t; t \geqq 0\}$ *is a martingale.*

PROOF. Trivial using Markov property, (3) and (6). q.e.d.

The same argument also yields

PROPOSITION 2. *Let* ξ *by any right eigenvector of* A *with eigenvalue* λ. *Then the family* $\{\xi \cdot X(t) e^{-\lambda t}; \mathfrak{F}_t; t \geqq 0\}$ *is a martingale (possibly complex valued).*

We have the following important

COROLLARY 1. *Let* v *be as in* (12). *Then*

$$(16) \qquad \lim_{t \to \infty} v \cdot X(t) e^{-\lambda_1 t} = W \qquad \textit{exists} \text{ wp } 1$$

and

$$E(W) \leqq v_i \qquad \textit{if} \;\; X(0) = e_i.$$

PROOF. Immediate since the family $\{v \cdot X(t) e^{-\lambda_1 t}; \mathfrak{F}_t; t \geqq 0\}$ is a nonnegative martingale and Fatou's lemma applies. q.e.d.

3. Almost sure convergence of $X(t)e^{-\lambda_1 t}$. To avoid trivial degeneracies we henceforth exclude the singular case

$$h_i(s) = \sum_{j=1}^{k} p_{ij}s_j,$$

where $p_{ij} \geqq 0$, $\sum_{j=1}^{k} p_{ij} = 1$. It is easy to see that in the singular case $\{X(t); t \geqq 0\}$ with $\sum_{i=0}^{k} X_i(0) = 1$ is essentially a Markov chain with state space $\{1, 2, \cdots, k\}$, transition probabilities p_{ij} and the sojourn time in i exponential with parameter a_i.

Using a recent result of Kesten and Stigum [9] on discrete time Galton-Watson process we establish the following:

THEOREM 1. *Under positive regularity and nonsingularity*

$$\lim_{t\to\infty} X(t, \omega)e^{-\lambda_1 t} = W(\omega)u \qquad \textit{exists} \text{ wp } 1 \tag{17}$$

where $W(\omega)$ is a nonnegative numerical valued random variable and u is defined by (12).

PROOF. Step 1. For $\forall\delta > 0$ the discrete skeleton $\{X(n\delta); n = 0, 1, 2, \cdots\}$ is a discrete time Galton-Watson process with mean matrix $M(\delta)$. By (16) and (10), if $\delta n_0(\delta) > t_0$, then $M(n_0(\delta)\delta) = [M(\delta)]^{n_0(\delta)} \gg 0$. Further the process is nonsingular. Now appealing to Kesten and Stigum's result [9] we conclude that there exists $A_\delta \,\varepsilon\, \mathfrak{F}$ such that $P(A_\delta) = 1$ and $\omega \,\varepsilon\, A_\delta \Rightarrow X(n\delta, \omega)e^{-\lambda_1 n\delta} \to W(\omega, \delta)u$. It follows that $\lim_{n\to\infty} v \cdot X(n\delta, \omega)e^{-\lambda_1 n\delta} = W(\omega, \delta)$ since $v \cdot u = 1$.

But by Corollary 1 we know $\lim_{t\to\infty} v \cdot x(t, \omega)e^{-\lambda_1 t} = W(\omega)$ exists wp 1. Thus we have $W(\omega, \delta) = W(\omega)$, wp 1. Hence given any sequence of δ_i we conclude that

$$\exists A \,\varepsilon\, \mathfrak{F} \ni P(A) = 1$$

and

$$\omega \,\varepsilon\, A \Rightarrow \quad \text{for} \quad \forall_i X(n\delta_i, \omega)e^{-\lambda_1 n\delta_i} \to W(\omega)u \quad \text{as} \quad n \to \infty. \tag{18}$$

We write λ for λ_1 in the remainder of this proof.

STEP 2. Let $D = \{\omega: W(\omega) > 0\}$. Then on $D^c \lim v \cdot X(t)e^{-\lambda t} = 0$ wp 1. But $v_i > 0$ for all i. This implies on D^c,

$$X_i(t)e^{-\lambda t} \to 0 \qquad \text{wp } 1. \tag{19}$$

It is known that $P\{D\} = 0$ if $\lambda_1 \leqq 0$. So we shall consider only the case $\lambda_1 > 0$. To establish (17) in this case we have to examine now only D. We shall use (18) to show that on D wp 1,

$$\liminf X_j(t, \omega)e^{-\lambda t} \geqq u_j W(\omega) \qquad \text{for all } j. \tag{20}$$

However,

$$\limsup \sum_{i=1}^{k} v_i X_i(t)e^{-\lambda t} \geqq \Big(\sum_{i\neq j} \liminf v_i X_i(t)e^{-\lambda t}\Big) + \limsup v_j X_j(t)e^{-\lambda t}. \tag{21}$$

But by (16)

$$\lim \sum_{i=1}^{k} v_i X_i(t)e^{-\lambda t} = W(\omega) \qquad \text{wp } 1.$$

Thus (14), (20) and (21) $\Rightarrow$ wp 1 on D

$$\limsup X_j(t)e^{-\lambda t} \leqq u_j W(\omega) \qquad \text{for all } j. \tag{22}$$

Putting (19), (20) and (22) together we get (17).

It remains only to establish (20) and this is our

STEP 3. By (16) $P\{W < \infty\} = 1$. Let $D_r = \{\omega : 1/r < W(\omega) < r\}$. Then $D_r \uparrow D$. To establish (20) it suffices to show $\forall r$ wp 1. on D_r

$$\liminf X_j(t)e^{-\lambda t} \geqq u_j W(\omega). \tag{23}$$

Fix $\delta > 0$ and $\forall t > 0$ define $n(t) = n(t, \delta)$ by

$$n(t)\delta \leqq t < (n(t) + 1)\delta.$$

Fix j and let $N(t) \equiv N(t, \omega)$ be defined as

$$\begin{aligned} N(t, \omega) &= 0 \quad \text{if} \quad X_j(n(t)\delta) = 0 \\ &= \text{number among the} \quad X_j(n(t)\delta) \quad \text{that split during} \\ &\quad\ [n(t)\delta, (n(t) + 1)\delta]. \end{aligned}$$

Clearly

$$X_j(t) \geqq X_j(n(t)\delta) - N(t). \tag{24}$$

We need only to show

$$P\{\omega : \omega \varepsilon D_r ; e^{-\lambda n\delta} N(n\delta, \omega) \geqq p(\delta) \quad \text{for infinitely many} \quad n\} = 0 \tag{25}$$

for some $p(\delta) \to 0$ as $\delta \downarrow 0$, because we then could use (18) to get (23) from (24) and(25).

To prove (25) notice

$$\lim_n X_j(n\delta)_e^{-\lambda n\delta} = u_j W(\omega) \qquad \text{wp 1.}$$

Therefore, by Egoroff's theorem [7] $\forall \eta_1 > 0$, $\eta_2 > 0, \exists\, A \varepsilon \mathfrak{F}$ such that $P(A^c) < \eta_1$ and

$$\omega \varepsilon A, n > N \Rightarrow |X_j(n\delta)e^{-\lambda n\delta} - u_j W(\omega)| < \eta_2 .$$

Let $E_n = \{\omega : \omega \varepsilon D_r : e^{-\lambda n\delta} N(n\delta, \omega) \geqq p(\delta)\}$ where $p(\delta)$ will be specified later and $E = \{\omega : \omega \varepsilon E_n$ for infinitely many $n\}$. Now $P(E) = P(EA^c) + P(EA)$. For $n > N$, $P(E_n A) \leqq P\{E_n A_n\}$ where

$$A_n = \{\omega : \omega \varepsilon D_r , |X_j(n\delta)e^{-\lambda n\delta} - u_j W(\omega)| < \eta_2\}. \tag{26}$$

But

$$\begin{aligned} P\{E_n A_n\} \leqq P\{\omega : \omega \varepsilon A_n , N(n\delta)/X_j(n\delta) - g(\delta) \\ \geqq (p(\delta) - g(\delta)X_j(n\delta)e^{-\lambda n\delta})/(X_j(n\delta)e^{-\lambda n\delta})\} \end{aligned} \tag{27}$$

where

$$\begin{aligned} g(\delta) &= 1 - e^{-a_j\delta} \\ &= P\{\text{a type } j \text{ particle splits in } [n\delta, (n+1)\delta) \mid \text{it is alive at } n\delta\}. \end{aligned}$$

Now on A_n

$$(28)\quad (p(\delta) - g(\delta)X_j(n\delta)e^{-\lambda n\delta})/X_j(n\delta)e^{-\lambda n\delta} \geqq (p(\delta) - g(\delta)(u_jr + \eta_2))/(u_jr + \eta_2) = g(\delta)$$

if we choose $p(\delta) = 2g(\delta)(u_jr + \eta_2)$ where η_2 will be specified in (31).

From (27) and (28)

$$(29)\quad P(E_nA_n) \leqq P\{\omega;\ \omega \varepsilon A_n,\ |(X_j(n\delta))^{-1} \textstyle\sum_{r=1}^{X_j(n\delta)} (\xi_r - g(\delta))| \geqq g(\delta)\}$$

where

$\xi_n = 1$ if the rth particle among the $X_j(n\delta)$ splits during $[n\delta, (n+1)\delta)$

0 otherwise.

But on A_n

$$(30)\quad X_j(n\delta)e^{-\lambda n\delta} > (u_jr^{-1} - \eta_2).$$

Now choose $\eta_2 > 0$ such that

$$(31)\quad u_jr^{-1} - \eta_2 = c > 0.$$

Then by Chebychev's inequality (29), (30) and (31) yield

$$(32)\quad P(E_nA_n) \leqq (\tfrac{1}{4}cg^2(\delta))^{-1}e^{-\lambda n\delta}.$$

Since (32) holds for $n > N$ we get $\sum_1^\infty P(E_nA_n) < \infty$ and hence by Borel-Cantelli we have $P(EA) = 0$. But $P(EA^c) < \eta_1$ and η_1 is arbitrary. This establishes (25). Notice $p(\delta) \to 0$ as $\delta \downarrow 0$ since η_2 does not depend on δ, r is fixed and $g(\delta) \to 0$ as $\delta \downarrow 0$. q.e.d.

A REMARK. Let $\xi_i = (\xi_{i1}, \xi_{i2}, \cdots, \xi_{ik})$ for $i = 1, 2, \cdots, k$ be random variables taking values on the k-dimensional nonnegative integer lattice with generating functions respectively $h_i(s)$. Then it can be shown that

$$(33)\quad E(\xi_{ij} \log \xi_{ij}) < \infty \qquad \text{for all } i \text{ and } j$$

if and only if

$$(34)\quad E(X_j(t) \log X_j(t) | X(0) = e_i) < \infty \qquad \text{for all } i, j \text{ and } t.$$

Now again appealing to Kesten and Stigum's result [12] yields the following:

THEOREM 2. *Either*

$$(35a)\quad W(\omega) \equiv 0$$

or

$$(35b)\quad P\{W(\omega) = 0 | X(0) = e_i\} = q_i, \qquad E(W(\omega)| X(0) = e_i) = v_i$$

where $q_i = P\{X(t) = 0$ *for some* $t \mid X(0) = e_i\}$ *for* $i = 1, 2, \cdots, k$ *are the extinction probabilities and the unique root in the unit hypercube* $\{x = (x_1, \cdots, x_k);\ 0 \leqq x_i \leqq 1\}$ *satisfying* $u_i(x) = 0$ *for* $i = 1, 2, \cdots, k$ *and*

$x_i \neq 1$ *for at least one* i. *We have* (35b) *if and only if* (33) *holds. Further if* (33) *holds* $W(\omega)$ *has an absolutely continuous distribution on* $(0, \infty)$.

4. Second moments. In the rest of this paper unless stated explicitly to the contrary (or is clear from the context) we make the following additional assumptions.

ASSUMPTION 3.

$$\lambda_1 > 0. \tag{36a}$$

ASSUMPTION 4.

$$\partial^2 h_i(s)/\partial s_j \partial s_k \,|_{s=(1,1,\cdots,1)} < \infty \quad \textit{for all} \quad i, j \quad \textit{and} \quad k. \tag{36b}$$

It is known that $\lambda_1 > 0$ means $P\{X(t) = 0 \text{ for some } t\} < 1$. Harris [8] has shown (36b) implies $E(X_i(t)X_j(t) | X(0) = e_r) < \infty$ for all i, j, r and $t \geqq 0$. Let

$$C_r^{ij}(t) = \operatorname{cov}(X_i(t), X_j(t) | X(0) = e_r). \tag{37}$$

Using (3) we can assert

$$C_r(t + s) = (M(s))^* C_r(t) M(s) + \textstyle\sum_{i=1}^k m_{ri}(t) C_i(s) \tag{38}$$

where $C_r(t) = ((C_r^{ij}(t)))$ and $(M(s))^*$ is the transpose of $M(s)$. Let

$$D_r^{ij}(t) = E(X_i(t)X_j(t) | X(0) = e_r). \tag{39}$$

From (38) it will follow that

$$D_r(t + s) = (M(s))^* D_r(t) M(s) + \textstyle\sum_{i=1}^k m_{ri}(t) C_i(s). \tag{40}$$

This yields the differential equations

$$D_r'(t) = A^* D_r(t) + D_r(t) A + \textstyle\sum_{i=1}^k m_{ri}(t) C_i'(0) \tag{41}$$

where $'$ denotes differentiation.

It can be verified that the unique solution of (41) is

$$\begin{aligned} D_r(t) = {} & (M(t))^* D_r(0)(M(t)) \\ & + \textstyle\int_0^t (M(t - \tau))^* (\sum_{i=1}^k m_{ri}(\tau) C_i'(0))(M(t - \tau))\, d\tau. \end{aligned} \tag{42}$$

Let ξ be an eigenvector of A with eigenvalue λ, i.e., $A\xi = \lambda\xi$. Let

$$V_r(t) = E(|\xi \cdot X(t)|^2 | X(0) = e_r) = \bar{\xi}^* D_r(t) \xi. \tag{43}$$

Using (42) we conclude that

$$V_r(t) = e^{2at} \xi_r^2 + \textstyle\sum_{i=1}^k c_i (\int_0^t e^{2a(t-\tau)} m_{ri}(\tau)\, d\tau), \tag{44}$$

where $c_i = \bar{\xi}^* C_i'(0) \xi$, $a = \operatorname{Re} \lambda$.

Now appealing to Frobenius-Perron theory [10] we conclude

$$\lim_{t \to \infty} M(t) e^{-\lambda_1 t} = P = vu^*. \tag{45}$$

Hence we get the following:

PROPOSITION 3. *With the above notations*

(46a) $$\lim_t V_r(t)e^{-2at} = \xi_r^2 + \sum_{i=1}^k c_i \int_0^\infty e^{-(2a-\lambda_1)\tau} m_{ri}(\tau)e^{-\lambda_1\tau}\, d\tau \quad \textit{if} \quad 2a > \lambda_1,$$

(46b) $$\lim_t V_r(t)e^{-\lambda_1 t}t^{-1} = v_r(\textstyle\sum_{i=1}^k c_i u_i) \quad \textit{if} \quad 2a = \lambda_1,$$

(46c) $$\lim_t V_r(t)e^{-\lambda_1 t} = v_r(\textstyle\sum_{i=1}^k u_i c_i)(\lambda_1 - 2a)^{-1} \quad \textit{if} \quad 2a < \lambda_1.$$

This leads us to the following:

THEOREM 3. *Let ξ be any vector such that $A\xi = \lambda\xi$ and let $a = \mathrm{Re}\,\lambda$ be $>\lambda_1/2$. Then the martingale $\{Y(t) = \xi\cdot X(t)e^{-\lambda t}; \mathfrak{F}_t ; t \geqq 0\}$ satisfies (for any initial set up),*

$$\sup_t E\,|Y(t)|^2 < \infty$$

and hence there exists a random variable Y such that

(47) $$Y(t) \quad \textit{tends to} \quad Y \quad \text{wp } 1 \textit{ and in means square as} \quad t \to \infty.$$

PROOF. Immediate from (43), (46a) and Doob's convergence theorem [1]. We make use of (42) in [2], [3].

5. A Representation of $X(t) - e^{\lambda_1 t}Wu$. In Section 3 we saw

$$\lim_t X(t)\, e^{-\lambda_1 t} = Wu \text{ wp } 1.$$

The question arises as to what can be said about the order of magnitude $Y(t) = (X(t)e^{-\lambda_1 t} - Wu)$. The following lemma helps in answering this.

LEMMA 1. *There exists a set $A \in \mathfrak{F}$ such that*

(1) $P(A) = 1$.

(2) $\omega \in A \Rightarrow$ *for every t, there exists random variables $W_t^{ij}(\omega)$, $j = 1, 2, \cdots,$ $X_i(t, \omega)$, $i = 1, 2, \cdots, k$, such that*

(48) $$X(t, \omega) - e^{\lambda_1 t}W(\omega)u = \sum_{i=1}^k \sum_{j=1}^{X_i(t,\omega)} (e_i - W_t^{ij}(\omega)u).$$

PROOF. Fix a t. Then for any $l > 0$ recall the representation (3)

$$X(t + l, \omega) = \sum_{i=1}^k \sum_{j=1}^{X_i(t)} X^{ij}(l, \omega).$$

Now

$$uW(\omega) = \lim_{s\to\infty} X(s)e^{-\lambda_1 s} = \lim_{l\to\infty} X(t + l, \omega)e^{-\lambda_1(t+l)}.$$

Thus

$$e^{\lambda_1 t}uW(\omega) = \lim_{l\to\infty} \sum_{i=1}^k \sum_{j=1}^{X_i(t)} X^{ij}(l, \omega)e^{-\lambda_1 l}.$$

For a fixed t, $\lim_{l\to\infty} X^{ij}(l, \omega)e^{-\lambda_1 l} = uW_t^{ij}(\omega)$ exists wp 1 for $j = 1, 2, \cdots, X_i(t)$, $i = 1, 2, \cdots, k$. Hence for each fixed t, there exists a set $A_t \varepsilon \mathfrak{F}$ such that $P(A_t) = 1$, for $\omega \varepsilon A_t$, $W_t^{ij}(\omega)$ exist and

$$e^{\lambda_1 t}W(\omega) = \sum_{i=1}^{k} \sum_{j=1}^{X_i(t)} W_t^{ij}(\omega)$$

and

$$X(t) - e^{\lambda_1 t}uW(\omega) = \sum_{i=1}^{k} \sum_{j=1}^{X_i(t)} (e_i - uW^{ij}(\omega)).$$

Set $A = \bigcap_{t \text{ rational}} A_t$. Then using the right continuity of the sample paths we have for $\omega \varepsilon A$, for every t

$$X(t, \omega) - e^{\lambda_1 t}W(\omega) = \sum_{i=1}^{k} \sum_{j=1}^{X_i(t,\omega)} (e_i - uW^{ij}(\omega)). \qquad \text{q.e.d.}$$

We now get the following theorem which says something about the order of magnitude of $Y(t) = (X(t)e^{-\lambda_1 t} - uW)$.

THEOREM 4. *Let λ_1 be >0 so that $P\{X(t) = 0$ for some $t\}$ is less than one for any nontrivial initial make up. Assume second moments exist. Then*

$$(49) \quad \lim_{t\to\infty} P\{0 < x_1 \leqq W \leqq x_2 < \infty, v\cdot Y(t)e^{\lambda_1 t}(v\cdot X(t))^{-\frac{1}{2}} \leqq y\} = P\{0 < x_1 \leqq W \leqq x_2 < \infty\}\Phi(y/\sigma)$$

where v is as in (12),

$$(49a) \quad \sigma^2 = \sum_{i=1}^{k} u_i\sigma_i^2, \qquad \sigma_i^2 = \text{Var}\,(W \mid X(0) = e_i), \quad \textit{for} \quad i = 1, 2, \cdots, k.$$

PROOF. Let

$$(50) \qquad Z(t) = v\cdot Y(t)(v\cdot X(t))^{-\frac{1}{2}}e^{\lambda_1 t}.$$

It suffices to show

$$(51) \quad \lim_{t\to\infty} E(e^{i\theta Z(t)}; 0 < x_1 \leqq W \leqq x_2 < \infty) = e^{-(\sigma^2\theta^2)/2}P\{0 < x_1 \leqq W \leqq x_2 < \infty\}.$$

Now since $v\cdot X(t)e^{-\lambda_1 t} \to W$ a.s. it suffices to show

$$(52) \quad \lim_{t\to\infty} E(e^{i\theta Z(t)}; 0 < x_1 \leqq v\cdot X(t)e^{-\lambda_1 t} \leqq x_2 < \infty) = e^{-(\sigma^2\theta^2)/2}P\{0 < x_1 \leqq W \leqq x_2 < \infty\}.$$

But

$$(53) \quad \begin{aligned} &E(e^{i\theta Z(t)}; 0 < x_1 \leqq v\cdot x(t)e^{-\lambda_1 t} \leqq x_2 < \infty) \\ &\quad = E\{\textstyle\prod_{j=1}^{k} \Phi_j^{X_j(t)}(\theta(X_j(t))^{-\frac{1}{2}}(X_j(t)/v\cdot X(t))^{\frac{1}{2}}); \\ &\qquad 0 < x_1 \leqq v\cdot X(t)e^{-\lambda_1 t} \leqq x_2 < \infty\} \end{aligned}$$

where $\Phi_j(\theta) = E(e^{i\theta(v_j - W_j)})$ and $W_j = \lim_{t\to\infty} v\cdot X(t)e^{-\lambda t}$ when $X(0) = e_j$.

To justify (53) we observe that from (48) we can write

(54) $e^{\lambda_1 t} v \cdot Y(t)(v \cdot X(t))^{-\frac{1}{2}} = \sum_{j=1}^{k} (X_j(t)/v \cdot X(t))^{\frac{1}{2}} (X_j(t))^{-\frac{1}{2}} \sum_{r=1}^{X_j(t)} \eta_t^{jr}$

where $\eta_t^{jr} = (v_j - W_t^{jr})$, and for $r = 1, 2, \cdots, X_j(t)$ are independently and identically distributed and further independent of $X(t)$. Now $E(\eta_t^{jr}) = 0$, $E(|\eta_t^{jr}|^2) = \sigma_j^2 < \infty$. If we now appeal to the classical central limit theorem (53) yields immediately (52) since

$$(X_j(t)/v \cdot X(t))^{\frac{1}{2}} \to u_j \text{ a.s.} \qquad \text{q.e.d.}$$

REMARKS. 1. From (48) using the argument of Theorem 4 it would seem tempting to conclude convergence of $X(t) - e^{\lambda_1 t} W$ appropriately normalized, to a multivariate normal. But such a convergence does not hold since after all the limit random variable W is really one dimensional.

2. Theorem 4 says something about the behavior in law of $(X(t) - We^{\lambda_1 t})$. One would like to use (48) more strongly to assert some sample path behavior, like proving some law of the iterated logarithm etc., but this has not been done yet. (See [11].)

REFERENCES

[1] ATHREYA, KRISHNA B. (1967). Limit theorems for multitype continuous time Markov branching processes and some classical urn schemes, unpublished doctoral dissertation, Department of Mathematics, Stanford University.

[2] ATHREYA, KRISHNA B. (1967). Limit theorems for multitype continuous time Markov branching processes, I. The case of an eigenvector linear functional, (to appear).

[3] ATHREYA, KRISHNA B. (1967). Limit theorems for multitype continuous time Markov branching processes, II. The case of an arbitrary linear functional, (to appear).

[4] ATHREYA, KRISHNA B. and KARLIN, SAMUEL (1966). Limit theorems for the split times of branching processes. *J. Math. Mech.* **17** 257–278.

[5] CHUNG, K. L. (1967). *Markov Chains* (Second Edition). Springer-Verlag, Berlin.

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

[7] HALMOS, P. R. (1959). *Measure Theory.* D. Van Nostrand, Princeton.

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

[9] IKEDA, N., NAGASAWA, M. and WATANABE, S. (1966). A construction of branching Markov processes. *Proc. Japan Acad.* **42** No. 4.

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

[11] KENDALL, D. G. (1966). Branching processes. *J. London Math Soc.* 1966, 385–406.

[12] KESTEN, H. and STIGUM, B. P. (1966). A limit theorem for multidimensional Galton-Watson process *Annals of Math. Statist.* **37** 1211–1223.