EXTENSIONS OF A LIMIT THEOREM OF EVERETT, ULAM AND HARRIS ON MULTITYPE BRANCHING PROCESSES TO A BRANCHING PROCESS WITH COUNTABLY MANY TYPES[1]

By Shu-Teh C. Moy

Syracuse University

1. Introduction and summary. Let I be a countable set with elements $i, j, k, \cdots$. Elements of I are called "types." Let T be the collection of all sequences $z = \{z(i): i \varepsilon I\}$ of non negative integers of which all, but at most finitely many, are 0. We shall identify elements of T as real valued functions defined on I. The sum of two elements of T is again an element of T. The 0 function shall be denoted by $\mathbf{0}$. The function which takes the value 1 at i and 0 elsewhere shall be designated by e_i. Our branching process is a Markov chain $Z_0, Z_1, Z_2, \cdots$ with state space T and stationary transition probabilities described as follows:

The conditional distribution of Z_{n+1} given $Z_n = e_i$ is p_i where p_i is a probability distribution on T.

Let $z \varepsilon T$. $z(i_1) = n_1, \cdots, z(i_k) = n_k$; $z(i) = 0$ if i is not one of $i_1, \cdots, i_k$. Then the conditional distribution of Z_{n+1} given $Z_n = z$ is the distribution of a sum of $n_1 + \cdots + n_k$ independent random variables, taking values in T, of which n_1 have distribution $p_{i_1}, \cdots$, and n_k have distribution p_{i_k}.

Finally, Pr $\{Z_{n+1} = \mathbf{0} \mid Z_n = \mathbf{0}\} = 1$ completes the description of the transition probabilities of the process.

Let $Z_n(i)$ be the ith component of Z_n, i.e., $Z_n(i) = z(i)$ if $Z_n = z$. $Z_n(i)$ represents the size of the population of type i in the nth generation, and $\sum_i Z_n(i)$ represents the size of the total population of the nth generation. Let (m_{ij}) be the expectation matrix of the process, i.e., $m_{ij} = E\{Z_{n+1}(j) \mid Z_n = e_i\}$. Let $m_{ij}^{(n)}$ be defined inductively by

$$m_{ij}^{(1)} = m_{ij}, \qquad m_{ij}^{(n+1)} = \textstyle\sum_k m_{ik}^{(n)} m_{kj}.$$

We shall assume that all $m_{ij}^{(n)}$ *are finite and the matrix* (m_{ij}) *is irreducible*. For an extended real valued function f defined on I, we define functions Mf, fM by

$$Mf(i) = \textstyle\sum_j m_{ij} f(j); \qquad fM(j) = \sum_i f(i) m_{ij},$$

whenever the right sides of the equalities are well defined. Note that Mf, fM are always well defined if f is non negative since m_{ij} are non negative. (We shall always adopt the following conventions concerning addition and multiplication involving ∞:

$$\begin{aligned} a + \infty &= \infty + a = \infty \quad \text{if} \quad a > -\infty, \\ a \cdot \infty &= \infty \cdot a = \infty \quad \text{if} \quad a > 0, \\ &= 0 \quad \text{if} \quad a = 0. \end{aligned}$$

Received 23 September 1966.

[1] This work was supported in part by N.S.F. grant GP 5270.

Therefore, we have

$$E\{Z_{k+n}(j) \mid Z_k = e_i\} = m_{ij}^{(n)} \tag{1}$$

and

$$E\{Z_{k+n}(j) \mid Z_k = z\} = zM^n(j). \tag{2}$$

If f is a non negative function, then

$$E\{\textstyle\sum_j f(j)Z_{k+n}(j) \mid Z_k = z\} = \textstyle\sum_j zM^n(j)f(j) = \textstyle\sum_i z(i)M^n f(i). \tag{3}$$

In [6] it is shown that there is a unique non negative number r which is the common radius of convergence of the power series $\sum_{n=1}^{\infty} m_{ij}^{(n)} s^n$. *We shall always assume* $r > 0$.

In this paper we shall extend the far reaching theorem of Everett, Ulam and Harris for a branching process with finitely many types to a branching process with countably many types. Let us first describe the above mentioned theorem. For a branching process with k types, the expectation matrix (m_{ij}) is a $k \times k$ matrix. The matrix is assumed to be positive regular. It has a largest positive eigenvalue ρ with corresponding right and left eigenvectors u, v. If we normalize Z_n by dividing by ρ^n, the sequence $\{Z_n\rho^{-n}\}$ of vector valued random variables converges to a vector valued random variable W with probability 1, if $\rho > 1$. Furthermore, the direction of W, if $W \neq 0$, coincides with that of v. The theorem has been proven under some conditions on the existence of moments of Z_1 in [1], [2], [3]. And recently Keston and Stigum have proven the theorem under less stringent conditions [4]. In fact, the theorem has attained its best possible form in their work. All these works are based on the theory of ergodic behavior of $(m_{ij}^{(n)})$ which is described completely in the classical Perron Frobenius theory of positive matrices.

For an infinite, irreducible, non negative (m_{ij}) the number r^{-1} shall take the place of ρ of the finite case, in view of the fact that the radii of convergence of series $\sum_{n=1}^{\infty} m_{ij}^{(n)} s^n$ are all equal to ρ^{-1} for the finite case. The ergodic behavior of $(m_{ij}^{(n)})$ for an infinite matrix is far more complicated than its counterpart for a finite matrix. However, it has been worked out in great detail in [5]. There are two different but exhaustive cases:

CASE I, $\sum_{n=1}^{\infty} m_{ij}^{(n)} r^n < \infty$ for all i, j;
CASE II. $\sum_{n=1}^{\infty} m^{(n)} r^n = \infty$ for all i, j.

In Case II there are two strictly positive functions u and v such that

$$rvM(i) = v(i) \quad \text{for all } i; \tag{4}$$

$$rMu(i) = u(i) \quad \text{for all } i. \tag{5}$$

They are unique (up to constant multiples), non negative functions satisfying (4) and (5) respectively. The sum $\sum_i u(i)v(i)$ may be finite or infinite. In either Case I or Case II with $\sum_i u(i)v(i) = \infty$ we have $\lim_{n\to\infty} m_{ij}^{(n)} r^n = 0$. This being the case, if the process is initiated by finitely many particles

($\Pr\{Z_0 = z\} = 1$ for some $z \varepsilon T$),

$$E\{r^n Z_n(j)\} = r^n z M^n(j) = r^n \sum_i z(i) m_{ij}^{(n)} \to 0$$

as $n \to \infty$. Hence all sequences $\{r^n Z_n(j)\}$ converge in mean to 0. One might say that the conclusion of the theorem of E. U. H. remains valid with $W = 0$ and mean convergence instead of convergence with probability 1. The most interesting case we shall analyze in detail is Case II with $\sum_i u(i)v(i) < \infty$. In this case the infinite matrix behaves strikingly like a finite matrix. We are able to use an approach similar to that of Harris, which may be described as the mean square approach. A condition is imposed on distributions p_i so that all $Z_n(i)$ have finite second moments. We remark that there are other varieties of conditions which may serve the same purpose, but the one we have chosen has the appeal of being simple in appearance. The main result is Theorem 1 in which we prove the mean square convergence of $\{r^n Z_n(j)\}$.

2. Theorem and proof. Throughout this section, r is the common radius of convergence of power series $\sum_n m_{ij}^{(n)} s^n$. We *assume* $r < 1$ and the matrix (m_{ij}) to be *aperiodic*. We also *assume that* $\sum_n m_{ij}^{(n)} r^n = \infty$ and the functions u, v of (4) and (5) *satisfy* $\sum_i u(i)v(i) = 1$. Let us designate by $L_\infty(u)$ the collections of all real valued functions f defined on I such that $f(i)/u(i)$ are bounded, and by $L_1(u)$ the collection of functions f such that $\sum_i u(i) |f(i)| < \infty$. Similarly we define collections $L_\infty(v)$ and $L_1(v)$. We have $L_\infty(u) \subset L_1(v)$ and $L_\infty(v) \subset L_1(u)$. We now list two properties of $M^n f$ and fM^n which will be used in the sequel ([5] Theorem 5 and Theorem 6).

I. If $f \varepsilon L_1(v)$, then $Mf \varepsilon L_1(v)$ and $\lim_{n\to\infty} r^n M^n f(i) = \gamma u(i)$ for every $i \epsilon I$ where $\gamma = \sum_i f(i)v(i)$.

II. Of $|f(i)| \leqq \alpha u(i)$ for all $i \varepsilon I$, then $r\,|Mf(i)| \leqq \alpha u(i)$ for all $i \varepsilon I$. Similarly, if $|f(i)| \leqq \beta v(i)$ for all $i \varepsilon I$, then $r\,|fM(i)| \leqq \beta v(i)$ for all $i \varepsilon I$. It follows that $f \varepsilon L_\infty(u)$ implies $Mf \varepsilon L_\infty(u)$, and $f \varepsilon L_\infty(v)$ implies $fM \varepsilon L_\infty(v)$.

Let z be a non negative function defined on I and f be an extended real valued function defined on I. We shall use the following notation, provided the right side of (6) is meaningful:

$$z[f] = \sum_i z(i) f(i). \tag{6}$$

In particular, $z[f]$ is always well defined for a non negative f. In using this notation, $zM[f]$ is equal to $z[Mf]$ and is also equal to the sum $\sum_{i,j} z(i) m_{ij} f(j)$ if f is non negative. The usual product of two functions z, f is written as zf. Thus, if Z is a random variable which takes values in T, then $Z[f]$ is a real valued random variable which takes the value $\sum_i z(i)f(i)$ when Z takes the value z. In using (6), (3) becomes

$$E\{Z_{k+n}[f] \mid Z_k = z\} = z[M^n f]. \tag{7}$$

For two non negative functions f, g defined on I, we let

$$C_{f,g}(i) = E\{Z_1[f] Z_1[g] \mid Z_0 = e_i\}.$$

$C_{f,g}$ is a non negative function (possibly taking on the value ∞) defined on I.

LEMMA 1. *Let f, g be two non negative functions defined on I and $z \varepsilon T$. Then*

(8) $E\{Z_{n+1}[f]Z_{n+1}[g] \mid Z_n = z\} + z[(Mf)(Mg)] = z[Mf]z[Mg] + z[C_{f,g}]$.

PROOF. Let $z(i_1) = n_1, \cdots, z(i_k) = n_k$; $z(i) = 0$ if i is not one of $i_1, \cdots, i_k$. Let random variables $X_1^{(1)}, \cdots, X_{n_1}^{(1)}, \cdots, X_1^{(k)}, \cdots, X_{n_k}^{(k)}$ be independent and have values in T of which the first n_1 have distribution $p_{i_1}, \cdots$, the last n_k have distribution p_{i_k}. Then

$$E\{Z_{n+1}[f]Z_{n+1}[g] \mid Z_0 = z\} = E\{(\sum_{q=1}^{k}\sum_{l=1}^{n_q} X_l^{(q)}[f])(\sum_{q=1}^{k}\sum_{l=1}^{n_q} X_l^{(q)}[g])\} = \sum_{q,p=1}^{k} \pi_{q,p}$$

where $\pi_{q,p} = E\{(\sum_{l=1}^{n_q} X_l^{(q)}[f])(\sum_{l=1}^{n_p} X_l^{(p)}[g])\}$.

If $q \neq p$,

$$\pi_{q,p} = n_q n_p Mf(i_q)Mg(i_p),$$

$$\text{while} \quad \pi_{q,q} = (n_q^2 - n_q)Mf(i_q)Mg(i_q) + n_q C_{f,g}(i_q).$$

Hence

$$E\{Z_{n+1}[f]Z_{n+1}[g] \mid Z_0 = z\} + \sum_{q=1}^{k} n_q Mf(i_q)Mg(i_q) = \sum_{q,p=1}^{k} n_q n_p Mf(i_q)Mg(i_p) + \sum_{q=1}^{k} n_q C_{f,g}(i_q)$$

and (8) follows immediately.

By repeatedly applying (8) we obtain the following:

LEMMA 2. *Let f, g be two non negative functions defined on I and $z \varepsilon T$. Then*

(9) $$E\{Z_n[f]Z_n[g] \mid Z_0 = z\} + \sum_{k=0}^{n-1} zM^k[(M^{n-k}f)(M^{n-k}g)] = z[M^n f]z[M^n g] + \sum_{k=0}^{n-1} zM^k[C_{M^{n-k-1}f, M^{n-k-1}g}].$$

The following lemma follows immediately from Lemma 2.

LEMMA 3. *Let $W_n = r^n Z_n$ and $f_n = r^n M^n f$, $g_n = r^n M^n g$ where f, g are two non negative functions defined on I. Then*

(10) $$E\{W_n[f]W_n[g] \mid Z_0 = z\} + \sum_{k=0}^{n-1} r^{2k} zM^k[f_{n-k}g_{n-k}] = z[f_n]z[g_n] + r^2\sum_{k=0}^{n-1} r^{2k} zM^k[C_{f_{n-k-1}, g_{n-k-1}}].$$

Let

(11) $$V(i) = C_{u,u}(i) - r^{-2}u^2(i).$$

The first term in the right side of (11) is $E\{(Z_1[u])^2 \mid Z_0 = e_i\}$ and the second term (without the minus sign) is equal to the square of $E\{Z_1[u] \mid Z_0 = e_i\}$. Hence $V(i)$ is the *conditional variance* of $Z_1[u]$ given that $Z_0 = e_i$, and is therefore non negative. It follows that $C_{u,u}(i) \geqq r^{-2}u^2(i)$ for every i so that $C_{u,u} \varepsilon L_1(v)$ implies that u^2 and V also belong to $L_1(v)$.

Now, let

(12) $$\theta_i(j) = \sum_{n=0}^{\infty} r^{2n} e_i M^n(j).$$

By II, $\theta_i(j)$ is finite and $\theta_i \varepsilon L_\infty(v)$ for every $i \varepsilon I$. Furthermore, for any $z \varepsilon T$, $\sum_{n=0}^{\infty} r^{2n} zM^r$ is an element of $L_\infty(v)$ and

$$\sum_{n=0}^{\infty} r^{2n} zM^n(j) = \sum_i z(i)\theta_i(j).$$

LEMMA 4. *Let* $\Pr\{Z_0 = z\} = 1$ *where* $z \varepsilon T$. *If* f, g *are two functions belonging to* $L_\infty(u)$ *and if* $C_{u,u} \varepsilon L_1(v)$, *then*

$$E\{(W_n[f])^2\} < \infty, \qquad E\{(W_n[g])^2\} < \infty$$

and

$$\lim_{m,n\to\infty} E\{W_n[f]W_m[g]\} = \gamma\eta\{(z[u])^2 + r^2 \sum_i z(i)\theta_i[V]\} \tag{13}$$

where θ_i *are given by* (12), V, *by* (11) *and* $\gamma = \sum_i v(i)f(i)$, $\eta = \sum_i v(i)g(i)$.

PROOF. Since $\Pr\{Z_0 = z\} = 1$, various conditional expectations given $Z_0 = z$ become ordinary expectations. Let $z_k = r^{2k}zM^k$, $k = 0, 1, 2, \cdots$. Then, in letting $f = u$, $g = u$ in (10), we obtain

$$E\{(W_n[u])^2\} + \sum_{k=0}^{n-1} z_k[u^2] = (z[u])^2 + r^2 \sum_{k=0}^{n-1} z_k[C_{u,u}]. \tag{14}$$

Since $z_k \varepsilon L_\infty(v)$, $C_{u,u} \varepsilon L_1(v)$, the second term in the right hand side of (14) is finite. It follows that the second term in the left hand side of (14) is also finite since $r^{-2}u^2 \leqq C_{u,u}$. The first term in the right hand side of (14) is obviously finite. Hence $E\{(W_n[u])^2\}$ is finite. Let $|f(i)| \leqq \alpha u(i)$, $|g(i)| \leqq \beta u(i)$ for all $i \varepsilon I$. Then $E\{(W_n[f])^2\} \leqq \alpha^2 E\{(W_n[u])^2\}$, $E\{(W_n[g])^2\} \leqq \beta^2 E\{(W_n[u])^2\}$ so that both $E\{(W_n[f])^2\}$ and $E\{(W_n[g])^2\}$ are finite.

To prove (13), we shall prove for non negative f, g with $\gamma = \eta = 1$. The extension to general f, g is immediate. Let us assume $m = n + p$ where $p \geqq 0$. Then, applying (7), we have

$$E\{W_n[f]W_m[g]\} = E\{E\{W_n[f]W_m[g] \mid Z_n\}\}$$
$$= E\{W_n[f]E\{W_m[g] \mid Z_n\}\} = E\{W_n[f]W_n[g_p]\} \quad \text{where } g_p = r^pM^pg.$$

Hence, by (10),

$$E\{W_n[f]W_m[g]\} + \sum_{k=0}^{n-1} z_k[f_{n-k}g_{m-k}] = z[f_n]z[g_m] + r^2 \sum_{k=0}^{n-1} z_k[C_{f_{n-k-1},g_{m-k-1}}].$$

Since $f_n(i) \leqq \alpha u(i)$, $g_n(i) \leqq \beta u(i)$ by II, and since $f_n(i) \to u(i)$, $g_n(i) \to u(i)$ by I, and since $C_{u,u}(i)$ is finite, we conclude

$$\lim_{m,n\to\infty} C_{f_n,g_m}(i) = C_{u,u}(i) \tag{15}$$

and $C_{f_n,g_m}(i) \leqq \alpha\beta C_{u,u}(i)$ for all $i \varepsilon I$. Since $r < 1$, $\sum_{k=0}^{\infty} z_k \varepsilon L_\infty(v)$, hence $\sum_{k=0}^{\infty} z_k[C_{u,u}] = \sum_{k=0}^{\infty} z_k(i)C_{u,u}(i) < \infty$. For an arbitrarily given $\epsilon > 0$, there is a finite subset F of I and a positive integer K such that

$$\sum_{k>K, i\notin F} z_k(i)C_{u,u}(i) < \epsilon.$$

Therefore, for $m, n > K$

$$\begin{aligned} &\left|\sum_{k=0}^{n-1} z_k[C_{f_{n-k-1},g_{m-k-1}}] - \sum_{k=0}^{\infty} z_k[C_{u,u}]\right| \\ &\leqq \left|\sum_{k\leqq K, i\varepsilon F} z_k(i)C_{f_{n-k-1},g_{m-k-1}}(i) - \sum_{k\leqq K, i\varepsilon F} z_k(i)C_{u,u}(i)\right| \\ &\qquad + (\alpha\beta + 1)\epsilon. \end{aligned} \tag{16}$$

Applying (15), we conclude that the second term in the right hand side of (16) has a limit of 0 as $m, n \to \infty$. Since ϵ is arbitrary, we obtain

$$\lim_{m,n\to\infty} \sum_{k=0}^{n-1} z_k[C_{f_{n-k-1}, g_{m-k-1}}] = \sum_{k=0}^{\infty} z_k[C_{u,u}].$$

Similarly, we prove

$$\lim_{m,n\to\infty} \sum_{k=0}^{n-1} z_k[f_{n-k} g_{m-k-1}] = \sum_{k=0}^{\infty} z_k[u^2].$$

Hence

$$\begin{aligned}\lim_{m,n\to\infty} E\{W_n[f]W_m[g]\} &= (z[u])^2 + r^2 \sum_{k=0}^{\infty} z_k[C_{u,u}] - \sum_{k=0}^{\infty} z_k[u^2] \\ &= (z[u])^2 + r^2 \sum_{k=0}^{\infty} z_k[V] \\ &= (z[u])^2 + r^2 \sum_i z(i)\theta_i[V].\end{aligned}$$

In the above proof, it seems that $m \geqq n$ is assumed; but the situation may be easily remedied by replacing the upper limit $n - 1$ of many summations which appear in the argument by $\min(m, n) - 1$.

THEOREM 1. *If $C_{u,u} \varepsilon L_1(v)$ and if $\Pr\{Z_0 = z\} = 1$ where $z \varepsilon T$, then there is a non negative, real random variable Y with $E\{Y^2\} < \infty$ such that for every $f \varepsilon L_\infty(u)$ the sequence $\{W_n[f]\}$ converges in mean square to $\{\sum_i v(i)f(i)\}Y$; in particular, the sequence $\{W_n(i)\}$ converges in mean square to $v(i)Y$ for every $i \epsilon I$. Furthermore,*

$$E\{Y^2\} = (z[u])^2 + r^2 \sum_i z(i)\theta_i[V] \quad \textit{and} \quad E\{Y\} = z[u].$$

PROOF. Let $C = (z[u])^2 + r^2 \sum_i z(i)\theta_i[V]$ and $D = \{\sum_i v(i)f(i)\}^2 C$. We have

$$\begin{aligned}E\{W_n[f] - W_m[f]\}^2 &= E\{W_n[f]\}^2 - 2E\{W_n[f]W_m[f]\} + E\{W_m[f]\}^2 \\ &= (E\{W_n[f]\}^2 - D) - 2(E\{W_n[f]W_m[f]\} - D) + (E\{W_m[f]\}^2 - D).\end{aligned}$$

Now applying Lemma 4, we obtain $\lim_{m,n\to\infty} E\{W_n[f] - W_m[f]\}^2 = 0$. Hence the sequence $\{W_n[f]\}$ converges in mean square. Let us designate the limit by $W[f]$. (We shall see later on that the notation is well justified.) Then by Lemma 4

$$E\{W[f]W[g]\} = C\{\sum_i v(i)f(i)\}\{\sum_i v(i)g(i)\} \tag{17}$$

if g is another function belonging to $L_\infty(u)$. It follows, as a special case, that the sequence $\{W_n(i)\}$ converges in mean square to a limit which we shall designate by $W(i)$ and

$$E\{W(i)W(j)\} = Cv(i)v(j). \tag{18}$$

For a non negative function h, integrating term by term and applying (18), we obtain

$$E\{\sum_i h(i)W(i)\}^2 = C\{\sum_i h(i)V(i)\}^2.$$

Since our f satisfies $\sum_i |f(i)|\, v(i) < \infty$, we have $E\{\sum_i |f(i)|\, W(i)\}^2 < \infty$; hence the series $\sum_i f(i)W(i)$ converges absolutely with probability 1 and also in mean square.

Now we shall prove $W[f] = \sum_i f(i)W(i)$. We shall first prove for a non negative f. Using the diagonal procedure, we obtain an increasing sequence $\{n_k\}$ of positive integers for which $\{W_{n_k}(i)\}$ converges with probability 1 to $W(i)$ for every $i \varepsilon I$ and the sequence $\{W_{n_k}[f]\}$ also converges with probability 1 to $W[f]$. Since $\sum_i f(i)W_{n_k}(i) = W_{n_k}[f]$ with probability 1, in letting $k \to \infty$, we conclude $\sum_i f(i)W(i) \leqq W[f]$ with probability 1. However, integrating term by term and applying (18), we have

$$E\{\sum_i f(i)W(i)\}^2 = C\{\sum_i f(i)v(i)\}^2 = E\{W[f]\}^2.$$

Hence $\sum_i f(i)W(i) = W[f]$ with probability 1. For an f which takes on negative values, we set $f = f^+ - f^-$. Then

$$W[f] = W[f^+] - W[f^-] = \sum_i f^+(i)W(i) - \sum_i f^-(i)W(i) = \sum_i f(i)W(i).$$

From (17) we conclude that $E\{\sum_i f(i)W(i)\}^2 = 0$ whenever $\sum_i f(i)v(i) = 0$. It follows that $\sum_i f(i)W(i) = 0$ with probability 1 if f satisfies the same condition. Now for every ordered pair (j, k) of elements of I we define a function f_{jk} as follows:

$$f_{jk}(j) = v(j)^{-1},$$
$$f_{jk}(k) = -v(k)^{-1},$$
$$f_{jk}(i) = 0 \quad \text{if} \quad i \neq j, \qquad i \neq k.$$

Then $\sum_i f_{jk}(i)v(i) = 0$. Since the totality of all such functions is a countable collection $\mathcal{E}$, there is a set Ω' of probability 1 such that $\sum_i f(i)W(i) = 0$ on Ω' for all $f \varepsilon \mathcal{E}$. Hence $W(j)/v(j) - W(k)/v(k) = 0$ for all j, k on Ω'. Let j_0 be a fixed element of I and let $Y = W(j_0)/v(j_0)$. Then $W(k) = v(k)Y$ with probability 1 for all $k \varepsilon I$. It is clear that $E\{Y^2\}$ is finite and equal to C by (18) and $W[f] = \sum_i f(i)W(i) = \{\sum_i f(i)v(i)\}Y$. Since mean square convergence implies mean convergence, we have, for every $f \varepsilon L_\infty(u)$, $\{W_n[f]\}$ converging in mean to $\{\sum_i f(i)v(i)\}Y$. It follows that $E\{W_n[u]\} \to E[Y]$. Since $E\{W_n[u]\} = z[u]$ for $n = 1, 2, \cdots$, we conclude that $E\{Y\} = z[u]$.

In Theorem 1, random variable Y is well defined with probability 1 while this probability measure assigns probability 1 to $\{Z_0 = z\}$. We may call this random variable Y_z. The values of Y_z on sample functions of the process which do not start from z are immaterial. Nevertheless a random variable Y which is well defined with probability 1 for all possible initial distribution of $\{Z_n\}$ may be easily obtained by defining Y piecewise in letting $Y = Y_z$ on sample functions starting from z. For this Y, using a diagonal procedure similar to that used in the proof of Theorem 1, we can extract a subsequence $\{n_k\}$ of the sequence of all positive integers such that $\{W_{n_k}(i)\}$ converges with probability 1 to $v(i)Y$ for all i and all possible initial distribution of $\{Z_n\}$. Clearly, $E\{Y \mid Z_0 = z\} = z[u]$ by Theorem 1.

A few words on the condition $C_{u,u} \varepsilon L_1(v)$. If the initial "distribution" of the process assigns "probability" $v(i)$ to $\{Z_0 = e_i\}$ for all $i \varepsilon I$, and 0 to $\{Z_0 = z\}$ for all other $z \varepsilon T$, then the "expectation" of $(Z_1[u])^2$ is $\sum_i v(i)C_{u,u}(i)$. The condition may be described thusly: the "expectation" of $(Z_1[u])^2$ is finite if the initial "distribution" is so assigned. We use quotation marks here because $\sum_i v(i)$ may not be finite; therefore, $v(i)$ may not offer a genuine probability distribution.

Finally, we shall illustrate Theorem 1 with a corollary.

Let m be a non negative function defined on I. For each $j \varepsilon I$, let $Z(j)$ be a non negative, integral valued random variable with Poisson distribution:

$$\Pr\{Z(j) = k\} = e^{-m(j)}[m(j)]^k/k!, \qquad k = 0, 1, 2 \cdots .$$

Let $Z(j), j \varepsilon I$ be all independent. It is easy to see that $\Pr\{\sum_j Z(j) = \infty\} = 1$ or $\Pr\{\sum_j Z(j) < \infty\} = 1$ according as $\sum_j m(j) = \infty$ or $\sum_j m(j) < \infty$. In the latter case, the sum $\sum_j Z(j)$ is also an integral valued random variable which is Poisson distributed with mean $\sum_j m(j)$. Thus the joint distribution of $Z(j)$, $j \varepsilon I$, defines a probability distribution on T which we shall call a *multiple Poisson distribution* with mean m.

Let us suppose that our matrix (m_{ij}) has all finite row sums. For each $i \varepsilon I$, let m_i be the function on I defined by $m_i(j) = m_{ij}$. Let p_i be the multiple Poisson distritution on T with mean m_i. The collection of distributions $\{p_i\}$ defines transition probabilities for a branching process as was described in Section 1. We shall call it a *Poisson branching process*. For this branching process, the expectation matrix is (m_{ij}), $C_{u,u}(i)$ is given by

$$C_{u,u}(i) = r^2u^2(i) + \sum_j u^2(j)m_{ij}, \tag{19}$$

and

$$\sum_i C_{u,u}(i)v(i) = (r^2 + r)\sum_i v(i)u^2(i). \tag{20}$$

Hence, in order that $C_{u,u} \varepsilon L_1(v)$, it is sufficient that the function u be bounded. Thus we have the following corollary:

COROLLARY 1. *For a Poisson branching process, if the function u is bounded and if the process is initiated by finitely many particles, then the sequences* $\{W_n(i)\}$ *converge in mean square to* $v(i)Y$ *where* Y *is a real valued random variable.*

REFERENCES

[1] EVERETT, C. J. and ULAM, S. (1948). Multiplicative systems in several variables, I, II, III. Los Alamos Scientific Laboratory LA-683, LA-690, LA-707.

[2] HARRIS, THEODORE E. (1963). *The Theory of Branching Processes.* Prentice-Hall, Englewood Cliffs.

[3] HARRIS, THEODORE E. (1951). Some mathematical models for branching processes, *Proc. Second Berkeley Symp. Math. Statist. Prob.* 305–328.

[4] KESTON, H. and STIGUM, B. P. (1966). A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.* **37** 1211–1223.

[5] MOY, SHU-TEH C. (1967). Ergodic properties of expectation matrices of a branching process with countably many types. *J. Math. and Mech.* **16** 1207–1225.

[6] VERE-JONES, D. (1962). Geometric ergodicity in denumerable Markov chains. *Quart. J. Math. Oxford Ser.* **13** 7–28.

[7] VERE-JONES, D. Ergodic properties of non-negative matrices-I. (To appear in *Pacific J. Math.*)