THE DISTRIBUTION OF GENERATIONS AND OTHER ASPECTS OF THE FAMILY STRUCTURE OF BRANCHING PROCESSES

WOLFGANG J. BÜHLER
UNIVERSITY OF HEIDELBERG

1. Introduction

Throughout this paper branching processes will be viewed as models for the development of biological populations. Unless stated otherwise, it will be assumed that the population starts at time $t = 0$ with one individual in generation 0 and of age 0. Each member of the population will live for a random life time. Then he will be replaced by a random number of new individuals, his sons. These will be in generation $k + 1$ if their father was a member of the kth generation. Also we shall allow for an individual to "survive," that is, he may have himself as one offspring in generation k and start a new life. In the branching process model, it is further assumed that the lifetimes of all individuals have a common probability distribution with distribution function G, that the probability β for survival is the same for all individuals, that all individuals have the same distribution of the number of offspring given by a probability generating function h, and further, that all the random variables introduced so far are independent.

We shall follow the notations of Harris [10] who has studied many aspects of this model. Until recently the emphasis has been on studying the total population size; the possibility of an individual giving birth to offspring more than once has not usually been considered. The only exception seems to be the papers by Crump and Mode [5], [6], who consider a case somewhat more general than ours.

The questions with which this paper is concerned are about the distribution of generations in a population at a given time, the time pattern according to which generations appear and disappear, the degree of relationship between different individuals, the number of relatives of a certain degree, and so forth.

The first mention of distribution of generations with the present meaning in the literature was by Harris [10] who used the number $Z^{(k)}(t)$ of individuals in generations $0, 1, \cdots, k - 1$ alive at time t as an approximation to the total

population size $Z(t)$. Thus, the probability generating function $F(\cdot, t)$ of $Z(t)$ could be approximated by that of $Z^{(k)}(t)$. Conditioning on the first event in the process, that is, on life length and number of offspring of the original ancestor, led to the recursion formula

$$(1.1) \quad F^{(k+1)}(s, t)$$

$$= s[1 - G(t)] + \int_0^t h[F^{(k)}(s, t - u)] [\beta F^{(k)}(s, t - u) + 1 - \beta] \, dG(u)$$

with $F^{(0)}(s, t) = 1$. Formula (1.1) was given in [10] for the case $\beta = 0$. Applying the same kind of argument to the sequence $Z_k^*(t)$, defined as $Z^{(k)}(t)$ if $Z^{(k)}(t) = Z(t)$ and as ∞ otherwise, yields a sequence of probability generating functions F_k^* which also converges to F and satisfies a relation similar to (1.1), however with $F_0^*(s, t) = 0$. This sequence of probability generating functions (p.g.f.) was first introduced without a probabilistic interpretation by Levinson [16]. The asymptotic behavior of the distribution of generations after a long time and the development in time of high generations in supercritical processes (that is, in processes with $h'(1) + \beta > 1$) was studied by Martin-Löf [18] and then, independently and using different methods, by Kharlamov [13], [14], [15], Samuels [22], [23], and Bühler [4]. While Martin-Löf and Samuels considered general age dependent branching processes, Kharlamov and Bühler restricted themselves to the case of Markov branching processes (with exponential distribution of the life length) and Galton-Watson processes (with constant life length). Kharlamov allows for individuals of more than one type (admitting in [15] even an uncountable set of types). Methods similar to those of Kharlamov were used by Greig [8] in the study of chains of bacteria multiplying according to a birth and death process, where a chain is broken into two when a bacterium forming an "interior link" of the chain dies.

In all papers except [4], the survival probability β is zero. These asymptotic results will be described in Section 2. Section 3 considers the same questions in the subcritical case. While these questions themselves are not definitely answered, Theorems 3.1 and 3.2, dealing with the behavior of subcritical processes "long before extinction" may be of independent interest.

We shall then return to the supercritical case and study the family and relationship structure of the population, first extending the results given in [4] about the distribution of distant relatives, then considering the finer structure as, for example, the sizes of sibships, cousinships, and so forth. Finally, we shall reconsider and generalize a result first given by Stratton and Tucker [25] about the emergence of the first generation if the population is started by a large number N of individuals in generation zero.

2. The distribution of generations

It will be assumed throughout that the lifetime of an individual is positive with probability one, that is, that $G(0+) = 0$, and that its distribution is not con-

centrated on a lattice, except when we consider Galton-Watson processes. If $1 < m = h'(1) < \infty$, then there exists a unique positive α satisfying,

$$(2.1) \qquad m \int_0^\infty e^{-\alpha t} \, d\big(G(t) = 1.$$

The subcritical case with $m < 1$ will be considered only when a (necessarily negative) solution α of (2.1) exists and when the mean $\bar{\mu}$ and variance $\bar{\sigma}^2$ of the distribution

$$(2.2) \qquad \bar{F}(t) = m \int_0^t e^{-\alpha x} \, dG(x)$$

are finite. In the critical case $m = 1$, we have $\alpha = 0$ and $\bar{F} = G$. Denoting the ratio $Z_k(t)/Z(t)$ by $R_k(t)$, Samuels [23] obtained the following result.

THEOREM 2.1. *If G is not a lattice distribution, if $G(0+) = 0, \bar{\sigma}^2 < \infty$, $h(0) = 0, 1 < m < \infty, h''(1) < \infty$, and $\beta = 0$, then*

$$(2.3) \qquad \lim_{t \to \infty} \sum_{k=0}^{K_t} R_k(t) = \Phi(x)$$

in probability, whenever K_t is such that

$$(2.4) \qquad \left(K_t - \frac{t}{\bar{\mu}} \right) \left(\frac{t\bar{\sigma}^2}{\bar{\mu}^3} \right)^{-1/2} \to x.$$

Here, as usual, $\Phi(t)$ denotes the distribution function of the standard normal distribution.

Kharlamov [14] proved the same result and its extension to the branching process with individuals of different types [15] for the Markov case with lifetime distribution $G(t) = 1 - \exp\{-\lambda t\}$. He did not assume $h(0) = 0$, and thus he obtained (2.3) in conditional probability given $W > 0$, where $W = \lim_{t \to \infty} Z(t) \exp\{-at\}$ with $a = \lambda(m - 1)$.

Both Kharlamov's and Samuels' proofs proceed by first showing

$$(2.3') \qquad \lim_{t \to \infty} \sum_{k=0}^{K_t} \rho_k(t) = \Phi(x)$$

with $\rho_k(t) = EZ_k(t)/EZ(t)$, where the only condition on the offspring distribution needed to establish (2.3') is $m < \infty$. Showing then that $\sum_{k=0}^{K_t} Z_k(t)/E \sum_{k=0}^{K_t} Z_k(t)$ is close to $Z(t)/EZ(t)$ for large t and hence close to W, they obtain (2.3) by taking the ratio of these two expressions, cancelling W which is positive, and using (2.3').

Using essentially the same method, Samuels establishes a local version of the above result.

THEOREM 2.2. *If the conditions of Theorem 2.1 are satisfied and if $\int x^3 \, dG(x) < \infty$, then as $n \to \infty$,*

$$(2.5) \qquad \bar{\sigma}\sqrt{n} R_n(t_n) \to \bar{\mu}\varphi(x)$$

in probability provided the sequence t_n satisfies $(t_n - n\bar{\mu})/\bar{\sigma}\sqrt{n} \to x$. Again the corresponding relation

$$(2.5') \qquad\qquad \bar{\sigma}\sqrt{n}\rho_n(t_n) \to \bar{\mu}\varphi(x)$$

requires no condition on the offspring distribution but $m < \infty$.

These theorems, loosely speaking, state that at time t the generations with generation numbers around $t/\bar{\mu}$ are the most frequent in the population or that generation n will never make up a higher proportion of the total population than around time $n\bar{\mu}$. If we denote the mean and variance of G by μ and σ^2, Samuels [23] has pointed out that for the supercritical process, while the above is true, the absolute number $Z_n(t)$ of living individuals of the nth generation achieves its maximum at a much later time, namely, around $t = n\mu$, and in fact at time $n\bar{\mu}$ is still relatively small. More precisely, she obtained the following two theorems, in which $U_n(t)$ denotes the number of generation n objects born by time t and $V_n = U_n(\infty)$ is the total size of the nth generation.

THEOREM 2.3.   *If $G(0+) = 0$ and $\sigma^2 < \infty$, and if $h(0) = 0, 1 < m < \infty$, $\beta = 0$, and $h''(1) < \infty$, then as $n \to \infty$ and $(t_n - n\mu)/\sigma\sqrt{n} \to x$,*

$$(2.6) \qquad\qquad \frac{U_n(t_n)}{V_n} \to \Phi(x)$$

in probability.

THEOREM 2.4.   *If in addition to the assumptions of Theorem 2.3 we have $\int x^3 \, dG(x) < \infty$ and if G is not a lattice distribution, then*

$$(2.7) \qquad\qquad \frac{\sigma\sqrt{n}Z_n(t_n)}{\mu V_n} \to \varphi(x)$$

in probability.

These theorems again are established by first proving their expectation counterparts without the restrictions on h (except for $m < \infty$) and then showing that a certain sequence of differences of random variables converges to zero in quadratic mean. If we substitute $m^n W$ for V_n in the denominator of the left side of (2.6), we obtain the quantity which was shown by Martin-Löf [18] to converge to $\Phi(x)$, not only in probability but also in square mean.

The contrast between Theorems 2.1 and 2.2 on one hand and Theorems 2.3 and 2.4 on the other may be illustrated by the following corollaries, the first being a consequence of Theorem 2.3, the second following from Theorem 2.1.

COROLLARY 2.1.   *Under the conditions of Theorem 2.3, the time T_n of the birth of a randomly chosen individual of the nth generations has, as $n \to \infty$, asymptotically a normal distribution with mean $n\mu$ and variance $n\sigma^2$.*

COROLLARY 2.2.   *Under the conditions of Theorem 2.1, as $t \to \infty$, the generation G_t of a random individual alive at time t has asymptotically a normal distribution with mean $t/\bar{\mu}$ and variance $t\bar{\sigma}^2/\bar{\mu}^3$.*

Corollary 2.2 for the Markov case and its counterpart for the Galton-Watson process were also obtained by Bühler [4]. They will be stated separately here and proof will be indicated since it gives additional insight into the processes and the idea underlying it will again be used in later sections.

THEOREM 2.5. (i) *If in a Galton-Watson sequence with* $\beta > 0$ *and* $1 < \beta + m < \infty$, *and if* G_n *denotes the generation of an individual chosen randomly at time* n, *then*

$$(2.8) \qquad \lim_{n \to \infty} P\left\{\frac{G_n - nm/(m + \beta)}{(nm\beta/(m + \beta)^2)^{1/2}} \leq x \,\middle|\, W > 0\right\} = \Phi(x),$$

where $W = \lim_{n \to \infty} Z(n)(m + \beta)^{-n}$ *almost surely.*

(ii) *In a Markov branching process with* $G(t) = 1 - e^{-t}$, *if* $0 \leq \beta \leq 1$, $1 < \beta + m < \infty$ *and* $h''(1) < \infty$, *denoting by* G_t *the generation of a random individual alive at time* t, *we have*

$$(2.9) \qquad \lim_{t \to \infty} P\left\{\frac{G_t - mt}{(mt)^{1/2}} \leq x \,\middle|\, W > 0\right\} = \Phi(x),$$

where $W = \lim_{t \to \infty} Z(t) \exp\{-(\beta + m - 1)t\}$ *almost surely.*

To prove the continuous time version, one conditions first on the number $N(t)$ of "splits" that have occurred up to time t. Then given $N(t) = n$, say, we further condition on the sizes of the families produced at each split, that is, on the variables $U_i, V_i, i = 1, 2, \cdots, n$, representing the numbers of survivors (necessarily equal to zero or one) and the number of individuals newly born at the ith split. Then after choosing one of the n individuals at random following a suggestion of P. Clifford, we trace his line of ancestry back to the original ancestor present at time zero defining random variables $L_i, i = 1, 2, \cdots, n$, as follows. If the ith split produces (as a newborn individual) an ancestor of the individual considered, we let $L_i = 1$, otherwise L_i will be equal to zero. The generation G_t of our individual thus is just the sum of all the L_i. Under our conditioning the variables L_i, representing the "loss" in generation number as we go back beyond the ith split, will be independent. Let S_k denote the sum $u_1 + u_2 + \cdots + u_k + v_1 + v_2 + \cdots + v_k - k$. Then, given $(U_i, V_i) = (u_i, v_i)$, $i = 1, 2, \cdots, n$, the conditional probability of the event $L_k = 1$ is v_k/s_k and since the L_i are bounded we can apply the central limit theorem to their sum whenever the variance

$$(2.10) \qquad V(n) = \sum_{k=1}^{n} \frac{v_k}{s_k}\left(1 - \frac{v_k}{s_k}\right)$$

converges to infinity. This, however, will be the case since for almost all realizations of the process V_k/S_k is of the order $1/k$ as $k \to \infty$. More precisely, invoking Toeplitz' lemma, we can replace both the conditional expectation $E(n) = \sum_{k=1}^{n} v_k/s_k$ and $V(n)$ by $[m/(\beta + m - 1)] \log n$. This at the same time removes the conditioning on the family sizes (U_i, V_i). The final step is unconditioning

with respect to $N(t)$ using the fact, established by Athreya and Karlin [2], that $N(t)(\beta + m - 1) \exp\{-(\beta + m - 1)t\}$ converges to W almost surely.

3. The subcritical case

In the subcritical case $\beta + m < 1$, it is known that $Z(t)$ almost surely becomes zero as t becomes large. It is also well known that at least in the Galton-Watson case with $h''(1) < \infty$ and in the continuous time Markovian case the conditional distribution of $Z(t)$ given $Z(t) > 0$ has a limit as $t \to \infty$. The probability generating function of this limit will be denoted by g. Since the total population size (even after conditioning) does not go to infinity, Theorems 2.1 through 2.4 cannot possibly be true if $\beta + m < 1$ even though the assertions about ratios of expected values made in them or in connection with their proofs remain valid. It seems, however, reasonable to expect that Theorem 2.5 and possibly Corollary 2.1 should be valid in the subcritical case. If we try to imitate the proof of Theorem 2.5 in the case $\beta + m < 1$ in order to establish relation (2.10) for the process conditioned on $Z(t) \neq 0$, we might want to show that the relative frequencies of the possible values of v_k/s_k to appear up to time t have reasonable limits. A first step in this direction is to prove that in a sense we are dealing with a stationary process. While this result is relatively obvious, it seems not to have been published before. Let us first consider the Galton-Watson case, that is, let $\mathbf{Z}_n = (Z_n^1, Z_n^2, \cdots, Z_n^k)$ be a k type Galton-Watson process with basic probability generating function $\mathbf{f} = (f^1, f^2, \cdots, f^k)$,

$$(3.1) \qquad f^j(\mathbf{s}) = f^j(s_1, s_2, \cdots, s_k) = E\{s_1^{Z_1^1} s_2^{Z_1^2}, \cdots, s_k^{Z_1^k} \,|\, \mathbf{Z}_0 = \mathbf{e}_j\}.$$

Assume that the expectations $m_{ij} = E(Z_1^j \,|\, \mathbf{Z}_0 = \mathbf{e}_i)$ are finite and positive; then the matrix $\mathbf{M} = \{m_{ij}\}_{1 \leq i, j \leq k}$ has a largest positive eigenvalue ρ.

THEOREM 3.1. *If* $\mathbf{Z}_0, \mathbf{Z}_1, \cdots$ *is a k type Galton-Watson process with positive expectation matrix* $\mathbf{M}$ *and finite second moments and if* $\rho < 1$, *then as* $n \to \infty$ *and* $N \to \infty$, *the joint conditional distributions of* $(\mathbf{Z}_{n+0}, \mathbf{Z}_{n+1}, \cdots, \mathbf{Z}_{n+r})$ *given* $\mathbf{Z}_0 = \mathbf{e}_i$ *and* $\mathbf{Z}_{N+n+r} \neq \mathbf{0}$ *converge to those of a stationary process* $\mathbf{X}_0, \mathbf{X}_1, \cdots$.

The stationarity of the limit distribution, if it exists, is immediate. To prove the convergence, we first state a lemma which follows immediately from Theorem 4.3 of Jiřina [11] and its proof.

LEMMA 3.1. *Under the assumptions of Theorem 3.1, there exist constants* C_i *such that*

$$(3.2) \qquad \lim_{n \to \infty} \frac{f_n^i(\mathbf{s}) - 1}{c_i \rho^n} = g^i(\mathbf{s}) - 1.$$

$$(3.3) \qquad \lim_{n \to \infty} \frac{\partial^{j_1 + \cdots + j_k}}{\partial s_1^{j_1} \cdots \partial s_k^{j_k}} \frac{f_n^i(\mathbf{s})}{c_i \rho^n} = \frac{\partial^{j_1 + \cdots + j_k}}{\partial s_1^{j_1} \cdots \partial s_k^{j_k}} g^i(\mathbf{s})$$

for all $(j_1, j_2, \cdots, j_k)$. *Here* $g^i(\mathbf{s})$ *is the limiting probability generating function of* $\mathbf{Z}_n$ *given* $\mathbf{Z}_0 = \mathbf{e}_i$ *and* $\mathbf{Z}_n \neq \mathbf{0}$, *and the convergence is uniform on compact subsets of the region* $S = \{\mathbf{s}: |s_i| < 1, i = 1, 2, \cdots k\}$.

To avoid unnecessarily complex notation, the proof of Theorem 3.1 will be given only for the one dimensional case, $k = 1$. Let H be the conditional joint p.g.f. of $Z_n, Z_{n+1}, \cdots, Z_{n+r}, Z_{N+n+r}$ given $Z_{N+n+r} \neq 0$, then

$$(3.4) \quad H_{n,N}(s_0, s_1, \cdots, s_r. W)$$
$$= E\big(s_0^{Z_n} s_1^{Z_{n+1}} \cdots s_r^{Z_{n+r}} W^{Z_{N+n+r}} \big| Z_0 = 1. Z_{N+n+r} > 0\big)$$
$$= \{1 - f_{N+n+r}(0)\}^{-1} \{f_n(s_0 f(s_1 f(s_2 \cdots f(s_r f_N(W)) \cdots)$$
$$- f_n(s_0 f(s_1 \cdots f(s_r f_N(0)) \cdots))\}.$$

Using $f_N(w) - f_N(0) \sim c_1 m^N g(w)$ and $1 - f_N(0) \sim c_1 m^N$ and applying Lemma 3.1, we obtain

$$(3.5) \quad \lim_{n, N \to \infty} H_{n,N}(s_0, s_1, \cdots, s_r, w)$$
$$= c_1 m^{-r} g(w) s_0 s_1 \cdots s_r g'\big(s_0 f(s_1 f(\cdots f(s_r)) \cdots)\big)$$
$$f'\big(s_1 f(s_2 \cdots f(s_r)) \cdots\big) f'\big(s_2 \cdots f(s_r)\big) \cdots f'(s_r).$$

REMARK 3.1.   It is intuitively clear and confirmed by (3.5) that as $n, N \to \infty$ the variables $(Z_n, Z_{n+1}, \cdots, Z_{n+r})$ and Z_{N+n+r} become independent (conditionally given $Z_{N+n+r} > 0$).

REMARK 3.2.   The joint probability distributions of $X_0 - 1, X_1 - 1, \cdots$ are the same as those of a branching process $Y_0, Y_1, \cdots$ with immigration, where Y_0 has the p.g.f. g', where the p.g.f. of the number of offspring of each individual is f, and where the number I_j of immigrants at time j has the probability generating function $f'(s)/m$.

Let us now look at Theorem 3.1 for the two type Galton-Watson process in which the two types represent the numbers of surviving and newborn individuals, respectively. Thus, the limiting process $\mathbf{X}_0, \mathbf{X}_1, \cdots$ can be written as (S_0, N_0), $(S_1, N_1), \cdots$, where obviously the number S_i of survivors at time i can be at most equal to the total number $X_{i-1} = S_{i-1} + N_{i-1}$ of individuals living at time $i - 1$. If for the process $\mathbf{X}_0. \mathbf{X}_1, \cdots$ we define the variables $L_0. L_1, \cdots$ as having the values 1 and 0 with probabilities (conditional on $\mathbf{X}_0, \mathbf{X}_1, \cdots$) N_i/X_i and S_i/X_i, respectively, then (see, for example, Doob [7], Ch. X) almost surely the sequences $(1/r) \Sigma_{i=1}^r N_i/X_i$ and $(1/r) \Sigma_{i=1}^r N_i S_i/X_i^2$ converge to the conditional expectations $M = E(N_0/X_0 \big| \mathscr{F})$ and $V = E(N_0 S_0/X_0^2 \big| \mathscr{F})$, where $\mathscr{F}$ is the σ-field of invariant events for the stationary sequence $\mathbf{X}_i, i = 0, 1, 2, \cdots$. Thus, we can apply the central limit theorem analogously as in the corresponding step in the proof of Theorem 2.5.

LEMMA 3.2.   Let $\mathbf{X}_0, \mathbf{X}_1, \cdots$ and $L_0. L_1, \cdots$ be defined as above and let $G_r = L_0 + L_1 + \cdots + L_{r-1}$, then almost surely

$$(3.6) \quad \lim_{r \to \infty} P\left\{\frac{G_r - rM}{(rV)^{1/2}} \leq x \,\Big|\, \mathscr{F}\right\} = \Phi(x).$$

The asymptotic independence of remote variables (as expressed, for example, in Remark 3.1) should mean that actually, as $\mathbf{X}_0$ will be independent of $\mathscr{F}. M$

and V are nonrandom and the conditioning on $\mathscr{F}$ can be removed. Furthermore, going back to the original process $\mathbf{Z}_k = (U_k, V_k)$, $k = 0, 1, 2, \cdots$, we could let N, n, and r all be large in such a way that the first n and the last N variables do not contribute much to G_{N+n+r} and that the stretch of r variables in between would essentially behave like the stationary process $\mathbf{X}_0, \mathbf{X}_1, \cdots$.

CONJECTURE 3.1.   *If $Z_0, \mathbf{Z}_1, \mathbf{Z}_2, \cdots$ is a Galton-Watson process with $\beta + m < 1$ and finite second moments, then as $I \to \infty$, the generation G_I of a random individual living at time I has asymptotically a normal distribution with expectation IM and variance IV.*

REMARK 3.3.   Whereas in the supercritical case the proportion of newly born individuals almost surely converges to $m/(m + \beta)$ leading to the expressions for M and V in (2.7), this is not the case in subcritical processes where, indeed, M and V will depend on β and on the offspring distribution in a more complex way.

In the case of continuous time, we can proceed similarly. Let $\{Y_t, t \geqq 0\}$ be a subcritical Markov branching process with lifetime distribution function $1 - e^{-t}$, with h as the p.g.f. of the offspring distribution and with immigration. Assume the immigration to be compound Poisson; more precisely, let times between immigrations be independent exponential variables (having the same distribution as the lifetimes) and let the numbers of immigrants at each immigration be distributed according to $h'(s)/m$ independently of each other and of the other variables defining the process. Let finally Y_0 be distributed according to the p.g.f. $c_1 g'$, where

$$(3.7) \qquad g(s) = 1 - \exp\left\{-(1 - m) \int_0^s \frac{du}{h(u) - u}\right\},$$

and $c_1 = 1/g'(1)$.

THEOREM 3.2.   *Let $\{Z(t), t \geqq 0\}$ be a Markov branching process with $G(t) = 1 - e^{-t}$ whose offspring distribution is given by $h(s)$ with $h'(1) = m < 1$ and with $\beta = 0$. Then, as t and T tend to infinity, the finite dimensional distributions of the process $\{Z(t + \tau), \tau \geqq 0\}$ conditioned on $Z(t + T) > 0$ converge to those of the process $\{Y_\tau + 1, \tau \geqq 0\}$. The joint p.g.f. of $Y_0, Y_{\tau_1}, Y_{\tau_1 + \tau_2}, \cdots, Y_{\tau_1 + \cdots + \tau_r}$ is given by*

$$(3.8) \qquad c_1 \exp\left\{-m(\tau_1 + \tau_2 + \cdots + \tau_r)\right\}$$
$$\cdot g'\big(s_0 F\big(s_1 F\big(s_2 F \cdots s_{r-1}\big(F(s_r, \tau_r), \tau_{r-1}\big), \cdots\big), \tau_1\big)\big)$$
$$\cdot F_t(s_r, \tau_r) F_t\big(s_{r-1} F(S_r, \tau_r), \tau_{r-1}\big)$$
$$\cdots F_t\big(s_1 F\big(s_2 F\big(\cdots s_{r-1} F(s_r, \tau_r), \tau_{r-1}\big), \cdots\big), \tau_1\big),$$

where $F(s, t)$ denotes the p.g.f. of $Z(t)$ and F_t indicates the partial derivative of F with respect to its second (time) argument.

The proof will be omitted as it basically follows that of Theorem 3.1. It can also be extended to the multiple type case.

The arguments which led to Lemma 3.2 and from there to Conjecture 3.1 have to be appropriately modified. We now define a process $\{L_\tau, \tau \geqq 0\}$ as follows. Let L_τ be constant except for those points in time at which Y_τ increases; if at $\tau = t$ the Y process jumps by the positive amount V_t, then, with probability $V_t/Y_{t+} + 1$, the process L_τ will jump by unity, otherwise it will remain constant. Thus, the increments of the process $\{L_\tau, \tau \geqq 0\}$ will be stationary. Also, given the process $\{Y_\tau, \tau \geqq 0\}$, they will be conditionally independent. An analogue of Lemma 3.1 will hold. This will allow us to argue in the same way as in the discrete time case to obtain an analogue to Conjecture 3.1.

CONJECTURE 3.2. *If $Z(t)$ is a Markov branching process with $\beta + m < 1$, then there exist constants M and V depending on the offspring distribution such that the generation of a random individual alive at time t has asymptotically, as $t \to \infty$, a normal distribution with mean Mt and variance Vt.*

4. The relationship structure, distant relatives

In this section, we shall again study supercritical branching processes in discrete or continuous time. For these we shall investigate how closely related individuals taken randomly from the population will be and how many relatives of a given large degree a random individual will have. If we take two individuals at time t, these will be in generations G_t^1 and G_t^2 and have a last common ancestor (at some time prior to t) whose generation will be denoted G_t^{12}. The number $R_t^{12} = (G_t^1 - G_t^{12}) + (G_t^2 - G_t^{12})$ will then be called the degree of relationship.

THEOREM 4.1. *Let R_n^{ij}, $1 \leqq i < j \leqq k$, be the degrees of relationship of k random individuals at time n in a Galton-Watson process with $\beta > 0$ and $1 < \beta + m < \infty$. Then asymptotically, as $n \to \infty$, given the population does not become extinct, the random variables*

$$(4.1) \qquad V_n^{ij} = \left\{ R_n^{ij} - \frac{2nm}{m + \beta} \right\} \left(\frac{2nm\beta}{(m + \beta)^2} \right)^{-1/2}$$

have a joint normal distribution. Asymptotically, expectations are zero, variances are unity, and the covariance of V_n^{ij} and $V_n^{\ell q}$ is $\frac{1}{2}$ if $\{i, j\}$ and $\{\ell, q\}$ have one index in common and zero otherwise.

THEOREM 4.2. *Let R_t^{ij}, $1 \leqq i < j \leqq k$, be the degrees of relationship of k individuals chosen randomly at time t in a Markov branching process satisfying the conditions of Theorem 2.5 (ii). Then, given $W > 0$, asymptotically as $t \to \infty$, the random variables*

$$(4.2) \qquad V_t^{ij} = (R_t^{ij} - 2mt)(2mt)^{-1/2} \qquad\qquad 1 \leqq i < j \leqq k,$$

have the joint normal distribution with expectations zero and covariance matrix as in Theorem 4.1.

For the case $k = 2$, Theorems 4.1 and 4.2 are stated and proved in Bühler [4]. The essential step in the proof is to show that the generation G_n^A or G_t^A of

the last individual who is a common ancestor of at least two of the k individuals chosen has itself a limiting distribution (the tail of which is studied in [4] for $k = 2$). The variables $G_t^1 - G_t^A$, $G_t^2 - G_t^A$, $\cdots$, $G_t^k - G_t^A$ are then conditionally independent given G_t^A and the R_t^{ij} are sums of two such variables plus something that can be neglected as $t \to \infty$.

We now want to choose a random individual from the population. Then, if we follow his line of ancestry back by a given number n of generations or by a given time t, how many descendents will the ancestor have and how closely will they be related to the individual chosen? To attack this question for large n or t, respectively, we shall assume that we have a supercritical process in which the age distribution is the stationary one. We shall then make use of the following lemma about renewal processes, of which part (i) is a version of the central limit theorem and part (ii) is due to Takács [26].

LEMMA 4.1. *Let $X_1, X_2, X_3, \cdots$ be nonnegative independent random variables with a common distribution function G. Let*

$$(4.3) \qquad S_n = X_1 + X_2 + \cdots + X_n, \qquad N(t) = \max \{K; S_K \leqq t\}.$$

(i) *If $EX_1 = m$ and $0 < \operatorname{Var} X_1 = \sigma^1 < \infty$, then $(S_n - nm)(n\sigma^2)^{-1/2}$ has asymptotically, as $n \to \infty$, a standard normal distribution.*

(ii) *Under the same conditions $(N(t) - t/m)(t\sigma^2/m^3)^{-1/2}$ asymptotically as $t \to \infty$, has a standard normal distribution.*

Now, as we select an individual at random, his age (under suitable conditions on G) will be a random variable X_0 with distribution function

$$(4.4) \qquad A(x) = \frac{\int_0^x e^{-\alpha t}[1 - G(t)]\, dt}{\int_0^\infty e^{-\alpha t}[1 - G(t)]\, dt}.$$

At the time the individual was born he "selected his father" among the individuals present, whose ages were distributed according to A, with the risk of selecting an individual of age x proportional to the "failure rate" $g(x)/(1 - G(x))$. Thus, we are led to the following lemma.

LEMMA 4.2. *If in a supercritical branching process, $h''(1) < \infty$, and if G has a density g with $\int_0^\infty [g(t)]^p\, dt < \infty$ for some $p > 1$, then tracing the line of ancestry of a random individual, the life lengths $X_1, X_2, \cdots$ of his ancestors are independent and have the common probability distribution function*

$$(4.5) \qquad B(x) = m \int_0^x e^{-\alpha t} g(t)\, dt.$$

If we choose an individual at random after a long time τ, denote his ancestor living at time $\tau - t$ his t ancestor, and let L_t be the number of generations that we "lose" when we go back to him. Two individuals alive at time τ will be called t relatives if their last common ancestor was present at time $\tau - t$. Corres-

pondingly, we shall define the (n) ancestor of an individual by going back n generations in his line of ancestry. The (n) relatives will be those descendents of the (n) ancestor who are not also descendents of the $(n - 1)$ ancestor.

THEOREM 4.3. *If $h(0) = 0$, $1 < m < \infty$, $h''(1) < \infty$, $\beta = 0$, and G has a density g with $\int_0^\infty [g(t)]^p \, dt < \infty$ for some $p > 1$ such that the distribution B of (4.5) has finite expectation $\bar{\mu}$ and variance $\bar{\sigma}^2$, and if at time zero the population consists of a (fixed or random) number of individuals whose ages are independently distributed according to stationary age distribution of the process, then asymptotically as $t \to \infty$,*

(i) *L_t has a normal distribution with mean $t/\bar{\mu}$ and variance $\sigma^2 t/\bar{\mu}^3$;*

(ii) *if D_t denotes the number of t relatives of the individual chosen, then*

$$(4.6) \qquad \lim_{t \to \infty} P(D_t e^{-(m-1)t} \leqq x) = \sum_{k=1}^\infty \frac{k p_k}{m} F^{(k-1)}(x),$$

where F is the distribution function of $W = \lim_{t \to \infty} Z(t) e^{-(m-1)t}$ and $F^{(k)}$ is the kth convolution of F with itself, $F^{(0)}(x) = 1_{[0,\infty)}(x)$;

(iii) *the degree of relationship R_t of a random individual with a random t relative is asymptotically normally distributed with expectation $2t/\bar{\mu}$ and variance $2\bar{\sigma}^2 t/\bar{\mu}^3$.*

PROOF. Part (i) follows from Lemma 4.1 (ii) using (4.5). Part (ii) is established by conditioning on the number (k) of children of the t ancestor, where the factor $k p_k/m$ in (4.6) will be justified in Section 5. Finally, part (iii) follows from (i) and Corollary 2.2.

Note that part (iii) is closely related to Theorem 4.2. Since the ages of individuals in the Markovian situation are immaterial, for $k = 2$, Theorem 4.2 becomes a special case of Theorem 4.3 (iii).

The study of (n) relatives is not quite as simple as that of t relatives. We shall denote by T_n the time of birth of the (n) ancestor and by R_n the degree of relationship between a random individual and a random (n) relative.

THEOREM 4.4. *Under the conditions of Theorem 4.3, we have*

(i) *asymptotically, as $n \to \infty$, $(T_n - n\bar{\mu})/\bar{\sigma}\sqrt{n}$ has a standard normal distribution;*

(ii) *if D_n is the number of (n) relatives of a random individual, then for all $x > 0$*

$$(4.7) \qquad \lim_{n \to \infty} P(e^{-bn} D_n \leqq x) = \begin{cases} 0 & \text{if } b < \bar{\mu}\alpha, \\ \frac{1}{2} & \text{if } b = \bar{\mu}\alpha, \\ 1 & \text{if } b > \bar{\mu}\alpha; \end{cases}$$

(iii) *the degree of relationship R_n between a random individual and a random (n) relative is asymptotically, as $n \to \infty$, normally distributed with expectation $2n$ and variance $2n\bar{\sigma}^2$.*

PROOF. Part (i) follows from Lemma 4.2 together with Lemma 4.1 (i). To prove (4.7), we condition on T_n and then use (4.6) and part (i). As T_n becomes large

$$(4.8) \quad P\{e^{-bn}D_n \leq x \,|\, T_n\} = P\{D_n \exp\{-\alpha T_n\} \leq x \exp\{bn - \alpha T_n\} \,|\, T_n\}$$

is approximately equal to

$$(4.9) \qquad \sum_{k=1}^{\infty} \frac{kp_k}{m} F^{(k-1)}[x \exp\{bn - \alpha T_n\}].$$

Thus,

$$(4.10) \qquad P(e^{-bn}D_n \leq x) = EP\{e^{-bn}D_n \leq x \,|\, T_n\}$$

can be approximated by

$$(4.11) \qquad \frac{1}{\bar{\sigma}\sqrt{n}} \int_{-\infty}^{+\infty} \sum_{k=1}^{\infty} \frac{kp_k}{m} F^{(k-1)}[x \cdot \exp\{bn - \alpha t\}] \varphi\left(\frac{t - n\bar{\mu}}{\bar{\sigma}\sqrt{n}}\right) dt$$

which equals

$$(4.12) \qquad \int_{-\infty}^{+\infty} \sum_{k=1}^{\infty} \frac{kp_k}{m} F^{(k-1)}[x \cdot \exp\{n(b - \alpha\bar{\mu}) + \alpha\bar{\sigma}\tau\sqrt{n}\}] \varphi(\tau) \, d\tau.$$

If $b - (m-1)\bar{\mu} > 0$, the argument of $F^{(k-1)}$ will tend to $+\infty$ and therefore $F^{(k-1)}$ will tend to 1 uniformly on every finite τ interval for $k \leq K$ for any K. Similarly, $b - (m-1)\bar{\mu} < 0$ implies the convergence of the $F^{(k-1)}$ to 0 uniformly for $k \leq K$ on every finite τ interval. This proves the first and last statements of (4.7). If $b = (m-1)\bar{\mu}$, then the argument of $F^{(k-1)}$ will go to $+\infty$ or $-\infty$ according to whether $\tau > 0$ or $\tau < 0$, thus making $F^{(k-1)}$ converge to 0 on the negative halfline and to 1 on the positive halfline. With the corresponding uniformity, the whole expression converges to $\frac{1}{2}$.

Part (iii) is also proved by conditioning on T_n. We approximate for large T_n the probability

$$(4.13) \qquad P\left\{\left(R_n - n - \frac{T_n}{\bar{\mu}}\right)\left(\frac{\bar{\sigma}^2 T_n}{\bar{\mu}^3}\right)^{-1/2} \leq y \,\Big|\, T_n\right\}$$

by $\Phi(y)$ using Corollary 2.2. Therefore, we can approximate

$$(4.14) \quad P\left\{(R_n - 2n)\left(\frac{\bar{\sigma}^2 n}{\bar{\mu}^2}\right)^{-1/2} \leq x \,\Big|\, T_n\right\}$$

$$= P\left\{\left(R_n - n - \frac{T_n}{\bar{\mu}}\right)\left(\frac{\bar{\sigma}^2 T_n}{\bar{\mu}^3}\right)^{-1/2}\right.$$

$$\left. \leq x\left(\frac{n\bar{\mu}}{T_n}\right)^{1/2} - \left(\frac{T_n}{\bar{\mu}} - n\right)\left(\frac{\bar{\sigma}^2 T_n}{\bar{\mu}^3}\right)^{-1/2} \,\Big|\, T_n\right\}$$

by $\Phi\big(x(n\bar{\mu}/T_n)^{1/2} - (T_n/\bar{\mu} - n)(\bar{\sigma}^2 T_n/\bar{\mu}^3)^{-1/2}\big)$. Unconditioning shows that $P\{(R_n - 2n)(\bar{\sigma}^2 n/\bar{\mu}^2)^{-1/2} \leq x\}$ is close to

$$(4.15) \qquad \int_{-\infty}^{\infty} \Phi\left(x\left(\frac{n\bar{\mu}}{t}\right)^{1/2} - (t - n\bar{\mu})\left(\frac{\bar{\sigma}^2 t}{\bar{\mu}}\right)^{-1/2} \right.$$
$$\left. \cdot \varphi\big((t - n\bar{\mu})(n\bar{\sigma}^2)^{-1/2}\big) (n\bar{\sigma}^2)^{-1/2} \, dt \right.$$

which can be rewritten as

$$(4.16) \qquad \int_{-\infty}^{+\infty} \Phi\left\{ \left(\frac{n\bar{\mu}}{n\bar{\mu} + \tau(n\bar{\sigma}^2)^{1/2}}\right)^{1/2} (x - \tau) \right\} \varphi(\tau) \, d\tau.$$

As $n \to \infty$, (4.16) converges to $\int_{-\infty}^{\infty} \Phi(x - \tau)\varphi(\tau) \, d\tau$ which is the normal distribution function with variance 2.

5. The relationship structure, close relatives

In this section, we shall study the sizes of sibships, the numbers of cousins of an individual chosen at random, and related questions. There are two methods of approaching these questions, the first makes use of the theory of multitype branching processes and yields the distribution of the number of brothers, cousins, and so forth, that a random individual will ever have; the second, which gives a more complicated method of going back to the corresponding (n) ancestor and viewing his progeny, will also enable us to determine how many close relatives are alive at the given point in time.

We shall illustrate the latter method with one example only. Suppose we want to find the joint probability distribution of the numbers S of sibs and U of uncles of a random individual. We assume that we are dealing with a supercritical process which has been developing for a long time, so that the population size is big and the distribution of ages is the stationary age distribution with distribution function

$$(5.1) \qquad A(x) = \frac{\displaystyle\int_0^x e^{-\alpha t}[1 - G(t)] \, dt}{\displaystyle\int_0^\infty e^{-\alpha t}[1 - G(t)] \, dt}.$$

First, condition on the age A_0 of the individual and on the life length B_0 of his father, which is distributed according to the distribution function B defined in (4.5). Given $A_0 = a$ and $B_0 = b$ our individual has s live sibs and u live uncles if his grandfather has $k \geq u$ children out of which u survive to the age of $a + b$ and his father has $j > s$ children and s of the $j - 1$ brothers survive to the age of a. To carry our argument through, we need the conditional probability $P_k(x)$ that an individual has $k - 1$ brothers given his age is x. Thus, we select an individual at random among those of age x (or, since we are considering the

continuous time case, with ages in a small interval around x). Hence, $P_k(x)$ will be proportional to the number of k sibship individuals of age x present. This is expected to be proportional to

$$(5.2) \qquad p_k E\{\text{number of survivors to age } x \text{ out of a } k \text{ sibship}\} = p_k \cdot k[1 - G(x)].$$

Therefore, $P_k(x) = kp_k/m$ for all $x \geq 0$. Now we can simply write down the joint distribution of the numbers of uncles and sibs.

THEOREM 5.1. *In a supercritical branching process with lifetime distribution function G, with stationary age distribution according to (5.1) and distribution of life lengths of ancestors given by (4.5), the joint probability distribution of the number U of uncles and the number S of sibs of a random individual is given by*

$$(5.3) \qquad P(U = u, S = s)$$

$$= \int_0^\infty dA(a) \int_0^\infty dB(b) \sum_{k=s+1}^\infty \sum_{j=a+1}^\infty \frac{kp_k jp_j}{m^2 \binom{k-1}{s}\binom{j-1}{u}}$$

$$\cdot [1 - G(a)]^s [G(a)]^{k-s-1} [1 - G(a+b)]^u [G(a+b)]^{j-u-1}.$$

Using the same method, one can find the joint distribution of the ages of the individuals under consideration. As the expressions are rather complicated only a simple example will be given. If we consider the Markovian binary split process (where each individual after an exponential lifetime is replaced by two new individuals), picking a random individual given that he has a cousin, his own age A, the lifetime L of his father, and the age $\tilde{A}$ of a random live cousin have the joint probability density

$$(5.4) \qquad p(a, \ell, \tilde{a}) = \frac{432}{43} \exp\{-3(a + \ell)\}(2 - \exp\{-\tilde{a}\}).$$

This density is not symmetric in a and $\tilde{a}$, since our random individual is more likely to have a brother and one cousin than no brother and two cousins.

We now turn to the consideration of multitype processes. First, we shall identify the type of an individual with the size of the sibship to which he belongs. Thus, any individual, no matter what his own type, will produce k offspring individuals all of type k, with probability p_k. The expectation matrix M then has $\mathbf{p} = (p_1, 2p_2, 3p_3, \cdots)$ in each of its rows. Therefore, $\mathbf{p}$ is also its left eigenvector with corresponding eigenvalue m. Thus, appealing to a result of Moy [20], we can find the limit distribution of sibship sizes.

LEMMA 5.1. *In a supercritical Galton-Watson process with $h''(1) < \infty$, the relative frequencies f_k of the individuals whose sibships are of size k converge in square mean to kp_k/m and the frequency of sibships of size k converges to $p_k/(1 - p_0)$.*

Since all types of individuals have the same offspring distribution, we can combine all sibship sizes higher than $K - 1$, say, into one type. The expectation matrix of the offspring then has rows

$$(5.5) \qquad \left(p_1, 2p_2, \cdots, (K-1)p_{K-1}, m - \sum_{k=1}^{K-1} kp_k \right) = \mathbf{p}(K).$$

Again, $\mathbf{p}(K)$ is the left eigenvector corresponding to the eigenvalue m. For those processes with finitely many types, the convergence in square mean can be replaced by convergence almost surely (Harris [10], Theorem II, 9.2); also the continuous time Markovian case has been studied (Athreya [1]). As K is arbitrary, we have almost sure convergence for all $k \geq 1$.

THEOREM 5.2. (i) *In a Galton-Watson process with expectation $m > 1$ and with $h''(1) < \infty$, the relative frequencies of the individuals belonging to sibships of size k converge almost surely to kp_k/m and the relative frequencies of sibships of size k converge to $p_k/(1 - p_0)$.*

(ii) *In a Markov branching process with expectation $m > 1$ and with $h''(1) < \infty$, the relative frequencies of individuals belonging to sibships whose size at birth was k converge almost surely to kp_k/m and the frequencies of the sibships of size k represented in the population by at least one live member converge almost surely to $p_k/(1 - p_0)$.*

The relationship between the frequency of sibships of size k and the frequency of individuals from such sibships is the one usually encountered when sampling individuals or sampling families in a population. That it holds in the continuous time case, where at the time of sampling families need not be complete, is essentially due to the independence of $P_k(x)$ of x.

If we want to study different types of relatives at the same time, the corresponding expectation matrix and their eigenvectors will not be quite as easy to find. However, one can manage most cases of interest. Theorem 5.3 is an example of such a case. As we shall see in its proof, it is now not possible to pool several types of individuals; therefore, we impose an additional condition probably not needed for the conclusion to hold.

THEOREM 5.3. *If in a supercritical Markov branching process, the number of offspring of an individual is less than or equal to K almost surely, then the proportion $f(n_0, n_1, n_2, \cdots, n_r)$ of individuals in the population with $n_0 - 1$ brothers, n_1 cousins, n_2 second cousins, $\cdots$, n_r rth cousins converges almost surely to*

$$(5.6) \qquad p(n_0, n_1, \cdots, n_r) = \frac{n_0 p_{n_0}}{m} \sum_{i=1}^{K} \frac{ip_i}{m} P_{n_1}^{(i-1)}(1) \cdot \sum_{j=1}^{K} \frac{jp_j}{m} P_{n_2}^{(j-1)}(2)$$

$$\cdots \sum_{k=1}^{K} \frac{kp_k}{m} P_{n_r}^{(k-1)}(r),$$

where $P_i^{(j)}(k)$ is the probability that j individuals of generation zero will ever have i descendants of generation k, that is, $\Sigma_{i=0}^{\infty} P_i^{(j)}(k)s^i = [h_k(s)]^j$.

PROOF.   The proof shall be given only for the case $r = 1$ to simplify notation. We shall then classify an individual as being of type (i, j) when he has $i - 1$ brothers and j cousins, $i = 1, \cdots, K - 1$ and $j = 0, 1, \cdots, K(K - 1)$. Now not all rows of the expectation matrix M can be expected to be equal any more, as obviously the number of my sons' cousins will not be independent of the number of my own sibs. However, the type of my descendents will be independent of the number of my cousins and the number of sibs of my descendents will not depend on my type at all. Thus, M has a relatively simple structure and a left eigenvector can be easily found. In fact,

$$(5.7) \qquad\qquad m_{(i,j)(k,\ell)} = k p_k P_\ell^{(i-1)}(1)$$

and the left eigenvector corresponding to the eigenvalue m is given by (5.6). Of course, $p(n_0, n_1, \cdots, n_r)$ is the probability for an individual to have a father with n_0 children, a grandfather with (i children, $i - 1$ of whom have a total of n_1 children), $n_0 + n_1$ grandchildren, $\cdots$, and an r ancestor with $n_0 + n_1 + \cdots + n_r$ r descendents.

6. Processes with limited rebranching

This section is concerned with a property of continuous time branching processes first discovered by Stratton and Tucker [25], subsequently put in a context which considers generations by Bühler [2], and discussed in a way closest to the present treatment by Savage and Shimi [24]. Here we consider a whole sequence of branching processes $Z_N(t)$, $t \geqq 0$, $N = 1, 2, \cdots$, with $Z_N^{(k)}(t)$ individuals in generation k, $k = 0, 1, 2, \cdots$. Assuming $Z_N(0) = Z_N^{(0)}(0) = N$, the result of [25], [2], and [24] can be stated essentially as follows. As $N \to \infty$, there will be no second generation individuals yet at time t/N: however, there will be a Poisson number of independent first generation families whose sizes will be distributed according to the probability generating function h underlying the process. Furthermore, in the limit the increments of $Z_N^{(1)}(t/N)$ become independent. We shall now extend this result and also consider the times of emergence of higher generations than the first.

   THEOREM 6.1.   *Let $Z_N(t)$ be a sequence of branching processes with $Z_N(0) = Z_N^{(0)}(0) = N$, with offspring probability generating function h such that $0 < m < \infty$ and with distribution of life lengths according to the distribution function G with $G(0+) = 0$ and $G'(0) = 1$. Then, as $N \to \infty$, the probability generating function $H_{k,N}(s, a_{k,N}(t))$ of the number $Z_N^{(k)}(a_{k,N}(t))$ of individuals in generation k at time $a_{k,N}(t)$ converges to*

$$(6.1) \qquad\qquad \varphi_K(s, t) = \exp\left\{ \frac{m^{(k-1)} t^k}{k!} (h(s) - 1) \right\},$$

provided $a_{k,N}(t) \cdot N^{1/k} \to t$. Furthermore, under these conditions the increments of $\{Z_N^{(k)}(a_{k,N}(t)), t \geqq 0\}$ are asymptotically independent.

PROOF. Arguing as in Section 1, the p.g.f. $F_k(s, t)$ of the number of individuals in generation k if we start with one individual in generation zero at time 0 is shown to satisfy the relation (1.1) with $\beta = 0$ and

$$(6.2) \qquad F_0(s, t) = s\big(1 - G(t)\big) + G(t).$$

From this it can be shown that

$$(6.3) \qquad F_k(s, t) = 1 - \frac{m^{(k-1)}t^k\big(1 - h(s)\big)}{k!} + O(t^k)$$

as $t \to 0$. Using the facts $H_{k,N}(s, t) = [F_k(s, t)]^N$ and $(1 - x/N)^N \to e^{-x}$, the first assertion is proved since $a_{k,N}(t)$ tends to zero at the correct rate. Repeating the same kind of argument for the joint distributions at different times, the independence statement can be established. Similarly, it can be shown, that for $i \neq j$, the variables $Z_N^{(i)}\big(a_{i,N}(t_1)\big)$ and $Z_N^{(j)}\big(a_{j,N}(t_2)\big)$ are asymptotically independent.

REMARK 6.1. Theorem 6.1 shows that, apart from a transformation of the time scale, for a large initial population, all generations emerge according to the same compound Poisson process. As for a large population, if short life lengths are possible at all, it is likely that some offspring will indeed emerge after a short time. It is not surprising that the only conditions on G are local at $t = 0$ and that we need not restrict ourselves to Markov branching processes as had been done originally.

$$\diamond \qquad \diamond \qquad \diamond \qquad \diamond \qquad \diamond$$

REFERENCES

[1] K. B. ATHREYA, "Some results on multitype continuous time Markov branching processes," *Ann. Math. Statist.*, Vol. 39 (1968), pp. 347–357.

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

[3] W. J. BÜHLER, "Slowly branching processes," *Ann. Math. Statist.*, Vol. 38 (1967), pp. 919–921.

[4] ———, "Generations and degree of relationship in supercritical Markov branching processes, *Z. Wahrscheinlichkeitstheorie und verw. Gebiete*, to appear.

[5] K. S. CRAMP and C. J. MODE, "A general age dependent branching process," *J. Math. Anal. Appl.*, Vol. 24 (1968), pp. 494–508.

[6] ———, "A general age dependent branching process II," *J. Math. Anal. Appl.*, Vol. 25 (1969), pp. 8–17.

[7] J. L. DOOB, *Stochastic Processes*, New York, Wiley, 1953.

[8] M. GREIG, "A stochastic model of the mechanism governing the distribution of lengths of chains in bacterial populations," Ph.D. thesis, Department of Statistics, University of California, Berkeley, 1972.

[9] T. E. HARRIS, "Some mathematical models for branching processes," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, Univeristy of California Press, 1951, pp. 305–328.

[10] ——, *The Theory of Branching Processes*, Berlin-Gottingen-Heidelberg, Springer-Verlag, 1963.

[11] M. JIŘINA, "The asymptotic behavior of branching stochastic processes," *Czechoslovak Math. J.*, Vol. 7 (1957), pp. 130–153. (In Russian.) (English translation has appeared in *Selected Translations in Mathematical Statistics and Probability*, Vol. 2 (1962), pp. 87–107.)

[12] D. G. KENDALL, "On super-critical branching processes with a positive chance of extinction," *Research Papers in Statistics*, Festschrift for J. Neyman (edited by F. N. David), London, Wiley, 1966, pp. 157–165.

[13] B. P. KHARLAMOV, "On properties of branching processes with an arbitrary set of particle types," *Teor. Verojatnost. i Primenen.*, Vol. 13 (1968), pp. 82–95. (In Russian.)

[14] ——, "On the generation numbers of particles in a branching process with overlapping generations," *Teor. Verojatnost. i Primenen.*, Vol. 14 (1969), pp. 44–50. (In Russian.)

[15] ——, "On the numbers of generations in a branching process with an arbitrary set of particle types," *Teor. Verojatnost. i Primenen.*, Vol. 14 (1969), pp. 452–467. (In Russian.)

[16] N. LEVINSON, "Limiting theorems for age-dependent branching processes," *Illinois J. Math.*, Vol. 4 (1960), pp. 100–118.

[17] M. LOÈVE, *Probability Theory*, Princeton, Van Nostrand, 1963 (3rd ed.).

[18] A. MARTIN-LÖF, "A limit theorem for the size of the nth generation of an age-dependent branching process," *J. Math. Anal. Appl.*, Vol. 12 (1966), pp. 273–279.

[19] SHU-TEH C. MOY, "Ergodic properties of expectation matrices of a branching process with countably many types," *J. Math. Mech.*, Vol. 16 (1967), pp. 1207–1226.

[20] ——, "Extensions of a limit theorem of Evert, Wam, and Harris on multitype branching processes with countably many types," *Ann. Math. Statist.*, Vol. 38 (1967), pp. 992–999.

[21] R. OTTER, "The multiplicative process," *Ann. Math. Statist.*, Vol. 20 (1949), pp. 206–224.

[22] M. L. SAMUELS, "Distribution of the population among generations in an age-dependent branching process," Ph.D. thesis, Department of Statistics, University of California, Berkeley, 1969.

[23] ——, "Distribution of the branching-process population among generations," *J. Appl. Probability*, Vol. 8 (1971), pp. 655–667.

[24] I. R. SAVAGE and I. N. SHIMI, "A branching process without rebranching," *Ann. Math. Statist.*, Vol. 40 (1969), pp. 1850–1851.

[25] H. H. STRATTON, JR. and H. G. TUCKER, "Limit distributions of a branching stochastic process," *Ann. Math. Statist.*, Vol. 35 (1964), pp. 557–565.

[26] L. TAKÁCS, "On a probability problem arising in the theory of counters," *Proc. Cambridge Phil. Soc.*, Vol. 52 (1956), pp. 488–498.