LIMITING DISTRIBUTIONS FOR BRANCHING PROCESSES

JOHN LAMPERTI
DARTMOUTH COLLEGE

1. Introduction

Let Z_n, $n = 0, 1, 2, \cdots$, be the number of individuals in the n-th generation of a Galton-Watson branching process with basic distribution $\{p_i, i = 0, 1, \cdots\}$. That is, $\{Z_n\}$ are the random variables of a Markov chain whose states are the nonnegative integers and whose transition probability matrix is defined by

$$(1.1) \qquad p_{ij} = \text{coeff. of } x^j \text{ in } p(x)^i, \qquad \text{where} \quad p(x) = \sum_{\ell=0}^{\infty} p_\ell x^\ell.$$

For background on these processes we refer to the monograph of T. E. Harris [3], and all statements to the effect that some property of branching processes is "well known" are hereby defined to mean that the property in question is discussed there.

The purpose of this paper is to undertake a systematic study of limit distributions for Z_n, where the initial state Z_0 is allowed to tend to infinity with n. Thus it may happen, for certain sequences of numbers a_n, $b_n > 0$, c_n = positive integer ($c_n \to \infty$), that

$$(1.2) \qquad \lim_{n\to\infty} P\left\{\frac{Z_n - a_n}{b_n} \le x \,\middle|\, Z_0 = c_n\right\} = G(x)$$

exists in the usual sense of weak convergence of distribution functions. The basic problem, which is far from being completely solved, is to determine the class of distributions G which can arise in this manner, and the conditions on $\{p_i\}$ and $\{c_n\}$ under which a particular G will appear. If the distribution $\{p_i\}$ has finite variance, these questions can be fully answered with little difficulty; this is done in section 2 below. The case of infinite variance seems much more difficult, however, and only fragmentary results have been achieved so far (section 3).

A closely related problem occurs when (1.2) is strengthened by replacing Z_n by $Z_{[nt]}$ and requiring convergence to a limit (depending on t) for each $t \ge 0$. This is essentially equivalent to asserting the existence of a limiting process, and can occur only in the "critical" case $\mu = 1$. If the convergence occurs without translating the process—that is, if $a_n \equiv 0$—it has been possible to find all of the limiting processes which can arise. This theorem is the subject of section 4, and states that the possible limits form a one-parameter family (apart

Prepared with partial support from the National Science Foundation.

from scaling) which is closely related to a class of conditional limiting distributions discovered by V. M. Zolotarev [5]. In the alternative case $a_n/b_n \to \infty$, however, matters are much less satisfactory, for only partial results similar to those for the first problem have been obtained so far. These are given in section 5, and we conclude with a discussion of some unsolved problems.

2. Finite variance

Let $\mu = \sum_{i=0}^{\infty} ip_i$ be the mean number of offspring per individual, always assumed finite, and let $p(x)$ be the generating function of $\{p_i\}$. We shall take up in turn the cases $\mu > 1$, $\mu < 1$, $\mu = 1$. The basis of the discussion is the following.

LEMMA 2.1. *Let $F_n(x)$ be a sequence of distribution functions which converge weakly to a nondegenerate limit $F(x)$, and suppose too that the second moment of F_n converges to that of F, assumed finite. Let c_n be any sequence of positive integers tending to ∞. Then there exist sequences a_n, $b_n > 0$ such that*

$$\lim_{n\to\infty} F_n^{(c_n)}(b_n x + a_n c_n) = \Phi(x) \tag{2.1}$$

for all x, where Φ is the standard normal distribution.

(The notation $G^{(m)}(x)$, where G is a distribution function, will always mean the m-fold convolution of G with itself.)

PROOF. Choose a_n to be the mean of the distribution F_n and let $b_n^2 = c_n \operatorname{var}(F_n)$; then $F_n^{(c_n)}(b_n x + a_n c_n)$ is a distribution with mean zero and unit variance. From the form of central limit theorem for double sequences given in ([2], p. 103) we find that

$$\lim_{n\to\infty} c_n \int_{|x|>\epsilon} x^2\, dF_n(b_n x + a_n) = 0 \tag{2.2}$$

is a necessary and sufficient condition for (2.1). But

$$c_n \int_{|x|>\epsilon} x^2\, dF_n(b_n x + a_n) = \frac{c_n}{b_n^2} \int_{|y|>b_n\epsilon} y^2\, dF_n(y + a_n) \tag{2.3}$$

and $c_n/b_n^2 = 1/\operatorname{var}(F_n)$. Thus it is enough to see that $\operatorname{var}(F_n)$ is bounded away from zero and that

$$\int_{|y|>b_n\epsilon} y^2\, dF_n(y + a_n) \to 0 \qquad \text{for each} \quad \epsilon > 0. \tag{2.4}$$

The first of these conditions is immediate, since $\operatorname{var}(F_n) \to 0$ would imply that the limit F was degenerate. Because of the convergence of the second moments we have $a_n \to a$ (the mean of F), and also $b_n \to \infty$ since $c_n \to \infty$ and $\operatorname{var}(F_n)$ is bounded. Hence, for large n,

$$\begin{aligned} \int_{|y|>b_n\epsilon} y^2\, dF_n(y + a_n) &\leq \int_{-\infty}^{\infty} y^2\, dF_n(y + a_n) - \int_{-2M}^{2M} y^2\, dF_n(y + a_n) \\ &\leq \int_{-\infty}^{\infty} (z - a_n)^2\, dF_n(z) - \int_{-M}^{M} (z - a_n)^2\, dF_n(z), \end{aligned} \tag{2.5}$$

where M is any constant larger than a. If $\pm M$ are in addition continuity points of $F(z)$, we can pass to the limit $n \to \infty$ and use the hypotheses to obtain

$$\limsup_{n\to\infty} \int_{|y|>b_n\epsilon} y^2\, dF_n(y + a_n) \le \operatorname{var}(F) - \int_{-M}^{M} (z - a)^2\, dF(z), \tag{2.6}$$

which is arbitrarily small if M is chosen large; this proves the lemma.

THEOREM 2.1. *Assume $\mu > 1$ and* var $\{p_i\} = \sigma^2 < \infty$. *Let c_n be any sequence of positive integers tending to ∞. Then*

$$\lim_{n\to\infty} P\left\{\frac{Z_n - \mu^n c_n}{b_n} \le x \,\middle|\, Z_0 = c_n\right\} = \Phi(x), \tag{2.7}$$

provided $b_n \sim \mu^n\sqrt{c_n}\sigma(\mu^2 - \mu)^{-1/2}$.

PROOF. Let

$$F_n(x) = P\left\{\frac{Z_n}{\mu^n} \le x \,\middle|\, Z_0 = 1\right\}. \tag{2.8}$$

It is well known that for $Z_0 = 1$, Z_n/μ^n converges in mean square to a non-constant random variable ω, and so F_n satisfies the hypotheses of lemma 2.1. Moreover, by the definition of a branching process, letting $Z_0 = c_n$ has the effect of convolving the distribution of Z_n (given $Z_0 = 1$) with itself c_n times. Hence, (2.1) takes the form (2.7) in this case; the normalizing constant we have specified is asymptotic to the one chosen for the lemma.

REMARK. In this case there is also a limit distribution when c_n tends to a finite "limit" c, and it is just the c-th power of the law of ω. These distributions are numerous and somewhat difficult to study; we will return to the law of ω in section 3 in a different context. Of course, when $c_n \to \infty$ only infinitely divisible limiting distributions can occur.

Turning to the case $\mu < 1$, still with $\sigma^2 < \infty$, we will need to use the known facts that

$$P\{Z_n \ne 0 | Z_0 = 1\} \sim c\mu^n, \tag{2.9}$$

where c is a positive constant, and that

$$\lim_{n\to\infty} P\{Z_n = j | Z_0 = 1, Z_n \ne 0\} = f_j, \qquad \sum_{j=1}^{\infty} f_j = 1. \tag{2.10}$$

THEOREM 2.2. *If $c_n \sim d\mu^{-n}$, $d > 0$, then*

$$\lim_{n\to\infty} P\{Z_n = j | Z_0 = c_n\} = g_j, \qquad j \ge 0, \tag{2.11}$$

where $\{g_j\}$ is the distribution with generating function

$$\sum_{j=0}^{\infty} g_j x^j = e^{cd[f(x)-1]}, \qquad f(x) = \sum_{j=1}^{\infty} f_j x^j. \tag{2.12}$$

If instead $c_n\mu^n \to \infty$, (2.7) holds with $b_n \sim [c_n\mu^n\sigma^2/(\mu - \mu^2)]^{1/2}$, and a normal limiting distribution is obtained. These are the only possibilities for nondegenerate limits.

(Strictly speaking, these laws and anything of the same type are the only possibilities.)

PROOF. In case $c_n \sim d\mu^{-n}$, since the c_n original "ancestors" act independently, it is clear from (2.9) that as $n \to \infty$ the number of "lines of descent" surviving at time n tends to have a Poisson distribution with mean cd. But by (2.10) each such surviving line tends to have the distribution $\{f_j\}$. Thus the limiting distribution of Z_n should be that of the sum of a Poisson-distributed number of independent random variables with distribution $\{f_j\}$; the resulting law has generating function (2.12). This informal argument can easily be made rigorous; we omit further details.

In the case $c_n\mu^n \to \infty$ it is again possible to apply lemma 2.1. Instead of (2.8), we will define

$$F_n(x) = P\{Z_n \leq x | Z_0 = [\mu^{-n}]\}, \tag{2.13}$$

which converges to the compound Poisson law with generating function $\exp [cf(x) - c]$ as described above. The limiting first and second factorial moments of F_n are 1 and $\sigma^2/(\mu - \mu^2)$ respectively. Using the known first and second factorial moments of $\{f_j\}$, namely c^{-1} and $(\sigma^2 - \mu + \mu^2)/c(\mu - \mu^2)$, it is easy to verify that the first and second factorial moments of the limiting law are also 1 and $\sigma^2/(\mu - \mu^2)$ so that the hypothesis of the lemma is satisfied. Its conclusion (2.1) then becomes the current version of (2.7).

Finally, we must show that these are the only possibilities. But if, for any subsequence n', $c_{n'} = o(\mu^{-n'})$, it is clear by (2.9) that the probability that any of the $c_{n'}$ ancestors have descendants in the n' generation tends to zero, so the limit, if it exists, is degenerate. Similarly, any subsequence for which $c_{n'}\mu^{n'} \to \infty$ produces a normal law in the limit. In any other case, then, there is a subsequence for which $c_{n'} \sim d\mu^{-n'}$ and we have (2.11). That *any* limiting law must belong to the same type as one of these follows from ([2], theorem 1, p. 40).

The case $\mu = 1$ is very similar to that above. The discussion is based on the formulas

$$P\{Z_n \neq 0 | Z_0 = 1\} \sim \frac{2}{\sigma^2 n}, \tag{2.14}$$

$$\lim_{n\to\infty} P\left\{\frac{Z_n}{n} \leq x | Z_0 = 1, Z_n \neq 0\right\} = 1 - e^{-2x/\sigma^2}; \tag{2.15}$$

these results have long been known to hold under the assumption that $\sum i^3 p_i < \infty$, but as F. Spitzer has recently pointed out (not yet published), the existence of the second moment is enough. From (2.14) and (2.15), by an argument entirely analogous to that used when $\mu < 1$ we can derive the following.

THEOREM 2.3. *Assume* $\mu = 1$ *and* $\operatorname{var}\{p_i\} = \sigma^2 < \infty$. *If* $c_n \sim dn$, $d > 0$, *then*

$$\lim_{n\to\infty} P\left\{\frac{Z_n}{n} \leq x | Z_0 = c_n\right\} = G(x) \tag{2.16}$$

exists, where

$$\int_0^\infty e^{-\lambda x}\, dG(x) = e^{-\frac{2d}{\sigma^2}\frac{\lambda}{\lambda + 2/\sigma^2}}. \tag{2.17}$$

If instead $c_n/n \to \infty$, (2.7) *holds with* $b_n \sim \sigma\sqrt{nc_n}$; *there are no other limit distributions except the degenerate laws.*

3. Infinite variance

In this section we will consider cases where $\mu > 1$ and $\sigma^2 = \infty$. It is natural to try and use the general limit theorems for sums of independent random variables in a manner analogous to lemma 2.1, but this does not seem easy to carry out. Instead, we will discuss a method which allows some, at least, of the nonnormal limiting distributions to be identified. I do not know if they can all be obtained in this manner, even in principle.

When $\mu > 1$, even if $\sigma^2 = \infty$, $\omega = \lim_{n\to\infty} Z_n/\mu^n$ exists a.s. by the martingale convergence theorem; let $F(x) = P\{\omega \le x\}$ be its distribution function. The approach we will use is based on the following theorem.

THEOREM 3.1. *If* $G(x)$ *is any distribution function containing* $F(x)$ *in its domain of partial attraction, then* G *is a possible limit; that is, there exist sequences of constants for which* (1.2) *holds.*

PROOF. By the definition of partial attraction ([2], § 37) there exist sequences α_i, $\beta_i > 0$, γ_i = positive integer $\to \infty$, such that

$$\lim_{i\to\infty} P\left\{\frac{\omega_1 + \cdots + \omega_{\gamma_i} - \alpha_i}{\beta_i} \le x\right\} = G(x), \tag{3.1}$$

where ω_i are independent random variables with the same distribution as ω. For each γ_i, however, we have

$$\lim_{n\to\infty} P\left\{\frac{Z_n - \mu^n\alpha_i}{\mu^n\beta_i} \le x | Z_0 = \gamma_i\right\} = P\left\{\frac{\omega_1 + \cdots + \omega_{\gamma_i} - \alpha_i}{\beta_i} \le x\right\} \tag{3.2}$$

by the convergence theorem and the independence of different lines of descent. The idea of the proof is to let the limiting process in (3.2) be "almost complete" before switching from α_{i-1}, β_{i-1}, γ_{i-1} to α_i, β_i, γ_i.

To carry this out precisely, define N_i to be the first integer ($\ge N_{i-1}$ if $i > 1$) such that for all $n \ge N_i$, the Lévy distance between the distribution on the left side of (3.2) and its limit is at most ϵ_i, where $\{\epsilon_i\}$ is any positive sequence tending to zero. Then for $N_i \le n < N_{i+1}$ we can choose

$$c_n = \gamma_i, \qquad b_n = \beta_i\mu^n, \qquad a_n = \alpha_i\mu^n. \tag{3.3}$$

Combining the defining properties of this construction with (3.1), it is easy to see that (1.2) holds.

In order to apply this result, it is necessary to determine the laws to which $F(x)$ is partially attracted. Since F is not known very explicitly, this appears to be difficult in general. However, there is one tractable situation, and the result may have some independent interest.

THEOREM 3.2. *The distribution F is in the domain of normal attraction of a stable law with index $1 < \alpha < 2$ if and only if the distribution $\{p_i\}$ belongs to the same domain of attraction.*

REMARK. The result is probably true without the restriction to normal attraction ([2], p. 181), but its proof would be more complicated. Of course the stable law must be one which is capable of attracting positive random variables, so that $c_1 = 0$ or $\beta = -1$ in the usual representation of the characteristic function of the law.

PROOF. The requirement that $\{p_i\}$ be normally attracted to a stable law is expressed in terms of the behavior of its generating function as $x \to 1$ by the condition

$$1 - p(x) = \mu(1 - x) - c(1 - x)^\alpha + o(1 - x)^\alpha, \qquad c > 0. \tag{3.4}$$

If $\varphi(\lambda)$ is the Laplace transform of F, the condition that as $\lambda \to 0$

$$\varphi(\lambda) = 1 - \lambda + d\lambda^\alpha + o(\lambda^\alpha), \qquad d > 0, \tag{3.5}$$

has the same interpretation; thus we wish to show the equivalence of (3.4) and (3.5). The connection between p and φ is known to be

$$\varphi(\mu\lambda) = p(\varphi(\lambda)), \tag{3.6}$$

which (with the requirement that they be generating functions) determines either φ or p uniquely in terms of the other.

It is easy to see that (3.5) implies (3.4). Calling $\varphi(\lambda) = x$, we have from (3.5) and (3.6) $1 - p(x) = \mu\lambda - d\mu^\alpha\lambda^\alpha + o(\lambda^\alpha)$. But from (3.5), λ can be expressed in terms of x as $\lambda = (1 - x) + d(1 - x)^\alpha + o(1 - x)^\alpha$, and combining these two equations gives (3.4) with $c = d(\mu^\alpha - \mu)$.

The converse is a little more difficult. Let us write $\varphi(\lambda) = 1 - \lambda + a(\lambda)$, where $a(\lambda) = o(\lambda)$ since ω has mean 1. The function φ must satisfy (3.6), and using (3.4) this becomes

$$a(\mu\lambda) = \mu a(\lambda) + c\lambda^\alpha + o(\lambda). \tag{3.7}$$

It is convenient to make the substitutions $a(\lambda) = A(w)$, $\lambda = e^w$, $\mu = e^p$ which convert (3.7) into the first order difference equation

$$A(w + p) - e^p A(w) = ce^{\alpha w} + \psi(w), \tag{3.8}$$

where $\psi(w) = o(e^{\alpha w})$ as $w \to -\infty$. It should be noted that the $o(\lambda^\alpha)$ term in (3.7), and so $\psi(w)$ in (3.8), depend on the unknown $a(\lambda)$, and the growth condition on ψ is a consequence of the requirement that $a(\lambda) = o(\lambda)$. We shall now study the solutions of

$$B(w + p) - e^p B(w) = ce^{\alpha w} + \psi(w), \tag{3.9}$$

and attempt to single out the desired one.

The homogeneous part of (3.9) has the general solution $P(w)e^w$, where $P(w)$ is any function with period p. It is clear, therefore, that (3.9) has at most one solution which is $o(e^w)$ as $w \to -\infty$, for if there were two, their difference would be of the form $P(w)e^w$ which does not satisfy the condition unless $P(w) \equiv 0$. But it is not very difficult to construct such a solution $B_0(w)$, which must in fact be just $A(w)$ by the uniqueness. We can then see directly that $B_0(w)$ has the desired behavior at $-\infty$.

To carry out this plan, first make the further change of variable $B(w) = e^{w-p}C(w)$, so that (3.9) becomes

$$C(w+p) - C(w) = ce^{(\alpha-1)w} + e^{-w}\psi(w), \tag{3.10}$$

and the particular solution of interest is that one which tends to zero at $-\infty$. Any solution of (3.10) can be extended by summation from its values in, say, $[0, p)$ to the whole line:

$$C(w - np) = C(w) - ce^{-(\alpha-1)p} \sum_{i=0}^{n-1} e^{(\alpha-1)(w-ip)}(1 + o(1)), \tag{3.11}$$

where "$o(1)$" tends to 0 as $w - ip \to -\infty$. If we choose $C(w)$ in $[0, p)$ so that $C(w) \to 0$ as $w \to -\infty$, which must be possible since the infinite series converges for each fixed w ($\alpha > 1$), then the resulting particular solution can be written

$$C_0(w - np) = ce^{-(\alpha-1)p} \sum_{i=n}^{\infty} e^{(\alpha-1)(w-ip)}(1 + o(1)). \tag{3.12}$$

This is the solution we want; $e^{w-p}C_0(w) = B_0(w) = A(w)$. The dominant parts of the terms of the series are in geometric progression, so that (3.12) easily yields

$$C_0(w) = \frac{c}{e^{(\alpha-1)p} - 1} e^{(\alpha-1)w} + o(e^{(\alpha-1)w}), \qquad w \to -\infty \tag{3.13}$$

But reversing the substitutions we have made, (3.13) becomes

$$a(\lambda) = \frac{c}{\mu^\alpha - \mu} \lambda^\alpha + o(\lambda^\alpha), \qquad \lambda \to 0, \tag{3.14}$$

which is equivalent to (3.5) with $d = c/(\mu^\alpha - \mu)$; q.e.d.

COROLLARY. *If a branching process has $\mu > 1$ and if $\{p_i\}$ is normally attracted to a particular stable law, then there exist constants a_n, b_n, c_n for which* (1.2) *holds with that stable law in the role of G.*

This is an immediate consequence of theorems 3.1 and 3.2. It is natural to conjecture that *only* the stable law which attracts $\{p_i\}$ can appear in (1.2) if $c_n \to \infty$, but I have not proved this; perhaps a careful approach via functional iteration would succeed. The possibilities when $\{p_i\}$ is not attracted to any stable distribution remain completely unexplored.

4. Limit processes

In this section we will assume that there are sequences $b_n > 0$, $c_n =$ positive integer, such that for each $t \geq 0$

$$\lim_{n\to\infty} P\left\{\frac{Z_{[nt]}}{b_n} \le x | Z_0 = c_n\right\} = G_t(x) \tag{4.1}$$

exists (in the usual sense of weak convergence), where $G_t(x)$ is always a distribution function and is nondegenerate for $t > 0$. This hypothesis is, in appearance at least, somewhat more general than asserting that the finite-dimensional distributions of $(Z_{[nt]}/b_n)$, given $Z_0 = c_n$, converge to those of a "reasonable" limiting Markov process. Our goal will be to find all such processes which are possible, and it turns out that existence of the one-dimensional limits $G_t(x)$ already provides adequate information.

Let us consider examples. If $\mu = 1$ and the second moment of $\{p_i\}$ is finite, then as we have already noted in connection with theorem 2.3,

$$P\{Z_n \ne 0 | Z_0 = 1\} = \frac{c}{n} + o\left(\frac{1}{n}\right), \tag{4.2}$$

and

$$\lim_{n\to\infty} P\left\{\frac{Z_n}{n} \le x | Z_0 = 1, Z_n \ne 0\right\} = 1 - e^{-cx}, \tag{4.3}$$

where c is a positive constant. From these facts it follows as before that if we choose $b_n = n$, $c_n = dn$, then (4.1) holds and $G_t(x)$ becomes the law of the sum of a Poisson-distributed number of independent exponentially-distributed random variables. In this case, therefore, we have

$$\int_0^\infty e^{-\lambda x}\, dG_t(x) = \exp\left[\frac{-d\lambda}{1 + \frac{t}{c}\lambda}\right]. \tag{4.4}$$

It is not difficult to identify the limiting process with the diffusion whose backward equation is

$$\frac{\partial u}{\partial t} = \frac{x}{c}\frac{\partial^2 u}{\partial x^2}, \tag{4.5}$$

and whose initial state is $x = d$ (see [1]).

A larger class of examples results from work of V. M. Zolotarev, who showed in [5] that in certain cases (with $\mu = 1$ but with infinite variance) (4.2) and (4.3) can be replaced by

$$P\{Z_n \ne 0 | Z_0 = 1\} = \frac{c}{n^\alpha} + o\left(\frac{1}{n^\alpha}\right) \tag{4.6}$$

and

$$\lim_{n\to\infty} E\{e^{-\lambda Z_n/n^\alpha} | Z_0 = 1, Z_n \ne 0\} = 1 - \frac{\lambda}{c}\left\{1 + \left(\frac{\lambda}{c}\right)^{1/\alpha}\right\}^{-\alpha}, \tag{4.7}$$

where $c > 0$ and $\alpha \ge 1$ are constants. (The law whose transform is given in (4.7) has a density which can be expressed in terms of a stable law of index α^{-1} when $\alpha > 1$.) When these formulas are valid, (4.1) again holds if we choose $b_n = n^\alpha$, $c_n = dn^\alpha$, and this time we obtain the compound Poisson law determined by

$$\int_0^\infty e^{-\lambda x}\, dG_t(x) = \exp\left[-d\lambda\left\{1 + t\left(\frac{\lambda}{c}\right)^{1/\alpha}\right\}^{-\alpha}\right]. \tag{4.8}$$

The example above which arises in the case of finite variance is obtained in the special case $\alpha = 1$. However, the nature of the limiting processes is quite different when $\alpha > 1$ than when $\alpha = 1$; in particular, although they are still martingales, they do not have continuous path functions.

The main result of this section is that these examples are the only ones possible.

THEOREM 4.1. *If* (4.1) *holds as described in the first paragraph of this section, then* $G_t(x)$ *must satisfy* (4.8) *for some* $\alpha \geq 1$; $d, c > 0$. *In addition, it must be the case that* c_n/b_n *tends to a positive, finite limit, and that*

$$c_n = n^\alpha L(n), \tag{4.9}$$

where L is slowly varying in the sense of Karamata.

(That is, $L(cx)/L(x) \to 1$ as $x \to \infty$ for every fixed $c > 0$.)

The proof is quite lengthy, so a brief outline may be worthwhile. We will use the Laplace transform of G_t in the form

$$\int_0^\infty e^{-\lambda x}\, dG_t(x) = e^{-\psi_t(\lambda)}, \tag{4.10}$$

and the function $\psi_t(\lambda)$ will be shown to satisfy two functional equations. The first of these merely expresses the branching property, and is easily obtained (lemma 4.3). The second equation is a consequence of (4.9) (lemma 4.7), and it is the derivation of (4.9) which is the hardest part of the proof. This will be done in lemmas 4.4 to 4.6. Finally, the functional equations can be solved, yielding (4.8).

LEMMA 4.1. *If* (4.1) *holds, the branching process satisfies* $\mu \leq 1$.

PROOF. If $\mu > 1$, it is well known that (for $Z_0 = 1$) $Z_n \to +\infty$ with probability $q > 0$. For each k there is, therefore, an $N(k)$ such that

$$P(Z_{n+N(k)} \geq k \text{ for all } n \geq 0 \mid Z_0 = 1) \geq q/2, \tag{4.11}$$

and by the law of large numbers we have

$$\lim_{m\to\infty} P\left(Z_{n+N(k)} \geq \frac{kq}{3}\, m \text{ for all } n \geq 0 \mid Z_0 = m\right) = 1 \text{ uniformly in } k. \tag{4.12}$$

Thus the probability on the left has a positive lower bound δ as k and m vary independently. But if (4.1) holds, so that Z_n/b_n has limit law G_1, for each k we will eventually have $n > N(k)$, and so for large n

$$P\left\{\frac{Z_{2n}}{b_n} \geq \frac{kq}{3}\frac{Z_n}{b_n} \,\middle|\, \frac{Z_n}{b_n}\right\} \geq \delta. \tag{4.13}$$

This situation (with δ independent of k) makes it impossible for Z_{2n}/b_n to have a limiting distribution at all unless G_1 is degenerate at 0.

LEMMA 4.2. *If* (4.1) *holds, then* $\lim_{n\to\infty} c_n/b_n = d$ *exists,* $0 < d < \infty$, *and* $\mu = 1$.

PROOF. The existence of the limit follows at once from (4.1) upon setting $t = 0$; this also yields $d \geq 0$. But since $\mu \leq 1$ by lemma 4.1,

$$\lim_{n\to\infty} E\left\{\frac{Z_{nt}}{b_n}\,\middle|\, Z_0 = c_n\right\} = \lim_{n\to\infty} \frac{c_n}{b_n}\mu^{[nt]} \leq d \lim \mu^{nt}, \tag{4.14}$$

and either $d = 0$ or $\mu < 1$ would imply that the limiting mean is 0 and so make the distribution of Z_{nt}/b_n tend to become concentrated at 0, contrary to the hypothesis of nondegeneracy. (From now on we will write Z_{nt} in place of $Z_{[nt]}$.)

COROLLARY. *Without loss of generality, we can assume* $c_n = b_n$.

PROOF. If (4.1) holds, b_n can be replaced by any sequence asymptotic to it without changing the limit; in particular, the sequence c_n/d can be used. But then replacing x by x/d yields an equation like (4.1) in which $b_n = c_n$.

We now derive the first of the functional equations which will determine $\psi_t(\lambda)$. It will be assumed hereafter that $b_n = c_n$ unless something else is explicitly asserted.

LEMMA 4.3. *If* (4.1) *holds (and* $b_n = c_n$*), then for all* $\lambda \geq 0$

$$\psi_{t+s}(\lambda) = \psi_t\{\psi_s(\lambda)\}. \tag{4.15}$$

PROOF. If $p(x) = \sum p_i x^i$ is the generating function of the basic distribution of descendants, then it is well known that

$$E(x^{Z_n}|Z_0 = k) = p_n(x)^k, \tag{4.16}$$

where $p_n(x)$ is the n-th functional iterate of p. It follows that

$$\exp\{-\psi_t(\lambda)\} = \lim_{n\to\infty} E\{\exp[-\lambda(Z_{nt}/c_n)]|Z_0 = c_n\} = \lim_{n\to\infty} p_{[nt]}(e^{-\lambda/c_n})^{c_n}, \tag{4.17}$$

which is equivalent to

$$\psi_t(\lambda) = -\lim_{n\to\infty} c_n \log p_{[nt]}\,(e^{-\lambda/c_n}) \tag{4.18}$$

Writing (4.18) with $t + s$ in place of t, and doing the functional iteration in two stages, we have

$$\psi_{t+s}(\lambda) = -\lim_{n\to\infty} c_n \log p_{[nt]}\{p_{[ns]}(e^{-\lambda/c_n})\}, \tag{4.19}$$

or, which is the same thing,

$$\psi_{t+s}(\lambda) = -\lim_{n\to\infty} c_n \log p_{[nt]}\left\{\exp\frac{-1}{c_n}[-c_n \log p_{[ns]}(e^{-\lambda/c_n})]\right\}. \tag{4.20}$$

But the quantity in square brackets tends to $\psi_s(\lambda)$ for each $\lambda \geq 0$. Combining this with (4.18), where the convergence is uniform, shows that the limit in (4.20) must be $\psi_t\{\psi_s(\lambda)\}$ and proves the lemma.

REMARK. Relation (4.15) is related to the assertion that there is a limiting transition probability function such that $p_t(1, [0, x]) = G_t(x)$ which satisfies the Chapman-Kolmogorov equation.

COROLLARY. *If* (4.1) *holds, then there exists* $a \geq 0$ *such that*

$$\int_0^\infty x\, dG_t(x) = e^{-at} \qquad \text{for all} \quad t \geq 0. \tag{4.21}$$

PROOF. Because of lemma 4.2, G_t is the limit of distributions with mean 1 (we are again assuming $b_n = c_n$). It follows that the left side of (4.21) exists and is ≤ 1 for each t; call it $m(t)$. But from (4.10)

$$\int_0^\infty x\, dG_t(x) = \left.\frac{\partial\psi_t(\lambda)}{\partial\lambda}\right|_{\lambda=0}, \tag{4.22}$$

and so (4.15) yields

$$m(t+s) = m(t)m(s). \tag{4.23}$$

Since m is not identically zero (G_t is not degenerate), this equation implies that $m(t) = e^{-at}$ for some a, and $a \geq 0$ is necessary since $m(t) \leq 1$.

The next three lemmas will prove (4.9). The first is a kind of substitute for the weak law of large numbers. If at this point we knew that $a = 0$ in (4.21), as will eventually appear, then a form of the w.l.l.n. itself could be used instead.

LEMMA 4.4. *For each n, let X_i^n, $i = 1, \cdots, m(n)$, be nonnegative, independent random variables with distribution F_n. Assume that $F_n \Rightarrow F$ (weak convergence), where F is nondegenerate, and that $E(X_i^n) \leq 1$ for each n. Let $\nu = \int x\, dF$; necessarily $\nu \leq 1$. Assume also that $m(n) \to \infty$ and that*

$$\lim_{n\to\infty} P\left\{\frac{X_1^n + \cdots + X_{m(n)}^n}{m(n)} \leq x\right\} = H(x) \tag{4.24}$$

exists. Then $H(\nu-) = 0$.

PROOF. Suppose the random variables X_i^n are truncated at A:

$$X_i^n = Y_i^n + Z_i^n, \qquad Y_i^n = \begin{cases} X_i^n & \text{if } X_i^n \leq A, \\ 0 & \text{otherwise.}\end{cases} \tag{4.25}$$

Then since $Z_i^n \geq 0$, we have

$$H(x) \leq \lim_{n\to\infty} P\left\{\frac{Y_1^n + \cdots + Y_{m(n)}^n}{m(n)} \leq x\right\}. \tag{4.26}$$

But the convergence $F_n \Rightarrow F$ implies that the distribution of Y_i^n converges to F truncated at A in the same manner—at least if we take A to be a continuity point of F. Moreover, because of the uniform bounds $0 \leq Y_i^n \leq A$, it is clear that

$$\lim_{n\to\infty} E(Y_i^n) = \lim_{n\to\infty} \int_0^A x\, dF_n(x) = \int_0^A x\, dF(x). \tag{4.27}$$

It follows that the ordinary weak law of large numbers applies to $\{Y_i^n\}$, and so that the right-hand side of (4.26) is a degenerate distribution with its mass at $\int_0^A x\, dF$. Since A can be chosen to make this arbitrarily close to ν, (4.26) yields the desired conclusion.

REMARK. If any additional positive random variables, independent of X_i^n or not, are added to the sum in (4.24), and the limit $H(x)$ still exists, the conclusion $H(\nu-) = 0$ remains true a fortiori. We will use the lemma in this modified form below.

LEMMA 4.5. *If* (4.1) *holds, then $\{c(tn)/c(n)\}$ is bounded as $n \to \infty$ both from ∞ and from 0 for each $t > 0$.*

(It is convenient, here and below, to write $c(u)$ instead of $c_{[u]}$.)

PROOF. Choose any $t > 0$ for which the boundedness from above fails; then for some subsequence $k(n)$ we can write

$$c(tk(n)) = m(n)c(k(n)) + r(n), \tag{4.28}$$

where $m(n)$ is an integer, $m(n) \to \infty$, and $0 \leq r(n) < c(k(n))$. Now for each n, let $Z_0 = c(tk(n))$; we have by (4.1)

$$\lim_{n\to\infty} P\left\{\frac{Z_{tk(n)}}{c(tk(n))} \leq x|Z_0 = c(tk(n))\right\} = G_1(x). \tag{4.29}$$

But by the independence property of branching processes, the quantity $Z_{tk(n)}$ in (4.29) can be considered as the sum of $m(n) + 1$ independent random variables, which are respectively the descendants of $m(n) + 1$ disjoint sets of "original ancestors"; the first $m(n)$ sets contain $c(k(n))$ members and the last $r(n)$ by (4.28). Let us denote the first $m(n)$ of these random variables, divided by $c(k(n))$, by $X_1^n, \cdots, X_{m(n)}^n$, and the last one (also so divided) by R^n. Then clearly

$$\frac{Z_{tk(n)}}{c(tk(n))} = \frac{m(n)c(k(n))}{c(tk(n))}\left\{\frac{X_1^n + \cdots + X_{m(n)}^n}{m(n)} + \frac{R^n}{m(n)}\right\}. \tag{4.30}$$

Finally, note that by (4.1)

$$P\{X_i^n \leq x\} = P\left\{\frac{Z_{tk(n)}}{c(k(n))} \leq x|Z_0 = c(k(n))\right\} \Rightarrow G_t(x). \tag{4.31}$$

We are now ready to apply lemma 4.4. The limit F of the distributions of X_i^n is G_t from (4.31); that the means of X_i^n are ≤ 1 was shown in lemma 4.1. The mean of G_t—the ν of the lemma—is e^{-at}. From (4.29) and (4.30) we obtain (4.24), since the factor multiplying the sum in braces tends to 1. Moreover, we have $G_1(x)$ playing the role of H in (4.24). The conclusion is that

$$G_1(e^{-at}-) = 0 \tag{4.32}$$

for each $t > 0$ such that $c(tn)/c(n)$ is not bounded above. But it is clear that such unboundedness must occur for t arbitrarily close to 1 if it occurs at all, and so $G_1(e^{-a}-) = 0$. Since e^{-a} is the mean of G_1 this implies that G_1 is degenerate contrary to hypothesis; the contradiction proves that $c(tn)/c(n)$ is bounded above for each $t > 0$. The fact that no subsequence can tend to zero follows too, since replacing tn by m and n by m/t in such a subsequence leads to a contradiction with what has just been proved.

LEMMA 4.6. *If* (4.1) *holds, then*

$$c(n) = n^{\alpha}L(n), \tag{4.33}$$

where $\alpha \geq 0$ *and* L *is a slowly varying function.*

PROOF. It is enough to show that $\lim_{n\to\infty} c(tn)/c(n)$ exists for every $t > 0$. Since we have established above that this ratio is bounded from ∞ and 0, we can select a subsequence n' for which there is a finite limit $r > 0$. Now from (4.1) in its equivalent Laplace-transformed version we have

$$(4.34) \qquad \exp\{-\psi_{t^{-1}}(\lambda)\} = \lim_{n'\to\infty} E\left\{\exp\left[-\lambda \frac{Z_{n'}}{c(tn')}\right] \middle| Z_0 = c(tn')\right\}$$
$$= \lim_{n'\to\infty} E\left\{\exp\left[-\lambda \frac{Z_{n'}}{c(n')} \cdot \frac{c(n')}{c(tn')}\right] \middle| Z_0 = c(n')\right\}^{\frac{c(tn')}{c(n')}},$$

where in the last step we use the fact that the value Z_0 appears as an exponent (see (4.17)). But carrying out the last limit, we obtain

$$(4.35) \qquad \exp\{-\psi_{t^{-1}}(\lambda)\} = \lim_{n'\to\infty} E\left\{\exp\left[-\frac{\lambda}{r}\frac{Z_{n'}}{c(n')}\right] \middle| Z_0 = c(n')\right\}^{r}$$
$$= \exp\left\{-r\psi_1\left(\frac{\lambda}{r}\right)\right\}.$$

Now if convergence fails, there must be subsequences tending to distinct limits $r_1 \neq r_2$, $0 < r_1, r_2 < \infty$. In this event (4.35) holds for both r_1 and r_2, and equating the right-hand sides yields

$$(4.36) \qquad \left\{\int_0^\infty \exp\left[-\frac{\lambda}{r_1}x\right] dG_1(x)\right\}^{r_1} = \left\{\int_0^\infty \exp\left[-\frac{\lambda}{r_2}x\right] dG(x)\right\}^{r_2}.$$

Putting $\lambda = r_1$, this equation asserts that the function e^{-x} has the same L_1 and L_p $(p = r_1/r_2)$ norms with respect to the measure dG_1. It follows that e^{-x} is constant a.e. (dG_1), so that G_1 is a degenerate law. This contradiction proves the lemma.

We can now easily obtain the second functional equation for the Laplace transform of the limiting distribution:

LEMMA 4.7. *If* (4.1) *holds, then for every* $\tau > 0$

$$(4.37) \qquad \psi_t(\lambda) = \tau^\alpha \psi_{t\tau}(\lambda\tau^{-\alpha}),$$

where α *is the constant in* (4.33).

PROOF. Consider equation (4.18) which defines $\psi_t(\lambda)$, and replace n by $m\tau$. From (4.33) we have $c(m\tau) \sim \tau^\alpha c(m)$ as $m \to \infty$ for any $\tau > 0$; using this fact (4.18) becomes

$$(4.38) \qquad \psi_t(\lambda) = -\lim_{m\to\infty} c(m\tau) \log p_{[tm\tau]}(e^{-\lambda/c(m\tau)})$$
$$= -\lim_{m\to\infty} \tau^\alpha c(m) \log p_{[t\tau m]}(e^{-\lambda/\tau^\alpha c(m)})$$
$$= \tau^\alpha \psi_{t\tau}(\lambda\tau^{-\alpha}).$$

REMARK. It is apparent from (4.37) that $\psi_t'(0)$ is a constant, and so that $a = 0$ in (4.21).

PROOF OF THEOREM 4.1. It is not difficult to solve the simultaneous equations (4.15) and (4.37). We first remark that the case $\alpha = 0$ cannot occur, for in this case (4.37) asserts that $\psi_t(\lambda)$ is independent of t. From (4.15) it then follows that $\psi(\lambda) = \psi(\psi(\lambda))$, or, since ψ is monotonic and continuous, $\psi(\lambda) \equiv \lambda$, and this is the degenerate case. Next, define

$$(4.39) \qquad h(t) = \psi_t(1), \qquad t \geq 0.$$

Putting $\tau = \lambda^{1/\alpha}$ in (4.37) then yields

$$\psi_t(\lambda) = \lambda h(t\lambda^{1/\alpha}), \tag{4.40}$$

and the branching property (4.15) becomes

$$h(t + s) = h(s)h(th(s)^{1/\alpha}). \tag{4.41}$$

Now $h(t)$ is continuously differentiable—in fact analytic—for $t > 0$, as can be seen from (4.40) and the fact that $e^{-\psi_t(\lambda)}$ is the Laplace transform of a non-negative random variable. Differentiating (4.41) with respect to t, we can let $t \to 0+$ (provided $s > 0$) to get $h'(s) = D[h(s)]^{1+(1/\alpha)}$ where $D = h'(0+)$ is a constant. This differential equation can be very easily solved:

$$h(s) = \left(\frac{-D}{\alpha} s + E\right)^{-\alpha} \qquad \text{for} \quad s > 0, \tag{4.42}$$

where E is an arbitrary constant. Substituting in (4.40) we obtain

$$\psi_t(\lambda) = \frac{\lambda}{\left[\frac{-D}{\alpha} t\lambda^{1/\alpha} + E\right]^{\alpha}}. \tag{4.43}$$

From our assumptions (including $b_n = c_n$) it follows that $\psi_t(\lambda)$ is continuous at $t = 0+$, and $\psi_0(\lambda) = \lambda$; thus $E = 1$. If we then write $c^{-(1/\alpha)}$ in place of $-D/\alpha$, (4.43) (together with (4.10)) reduces formally to (4.8). [Actually (4.43) reduces to (4.8) with $d = 1$, since we have let $c_n = b_n$. It is clear from, for instance, (4.17), that taking $Z_0 \sim dc_n$ has the effect of multiplying $\psi_t(\lambda)$ by d, which yields the general form of (4.8).] Thus it only remains to show that $-D > 0$ and $\alpha \geq 1$. Suppose first that $-D < 0$. Then there exists a $\lambda_0 > 0$ for which the denominator in (4.43) vanishes, and so $\psi_t(\lambda_0) = \infty$, which can occur only if all the mass of G_t is concentrated at $+\infty$. The case $D = 0$ is also degenerate—at 1 this time—which establishes $-D > 0$ and the legitimacy of the substitution used above. As for the possibility that $\alpha < 1$, it is easy to check by differentiating that in this case the second moment of G_t is 1. Since the mean is also 1, this again leads to a contradiction; the proof is at last complete.

5. Centered limit processes

In this section it is assumed that there are sequences a_n, $b_n > 0$, $c_n =$ positive integer, such that

$$\lim_{n\to\infty} P\left\{\frac{Z_{[nt]} - a_n}{b_n} \leq x \middle| Z_0 = c_n\right\} = H_t(x) \tag{5.1}$$

exists and is a probability distribution function for each $t \geq 0$, and is non-degenerate for $t > 0$. The difference between this situation and that of section 4 lies in the possibility of centering with the constants a_n; to insure that there is an essential difference we will suppose that $a_n/b_n \to +\infty$. As before, $\mu = 1$ is necessary for (5.1).

Putting $t = 0$ in (5.1), it becomes evident that

$$r = \lim_{n\to\infty} \frac{c_n - a_n}{b_n} \tag{5.2}$$

must exist. Since replacing a_n by $a_n + rb_n$ does not affect the existence of the limit in (5.1) but merely replaces x by $x - r$, we can and will assume that $r = 0$ so that $H_0(x)$ is degenerate at $x = 0$. It is then straightforward to establish the following theorem.

THEOREM 5.1. *If the above assumptions hold, $H_t(x)$ is the distribution at time t of a process with stationary independent increments and initial state* 0.

PROOF. By (5.2) with $r = 0$, necessarily $a_n = c_n + o(b_n)$; it is clear from (5.1) that the centering sequence $a_n = c_n$ will give the same result, and we will use it instead. In terms of characteristic functions, (5.1) becomes with these specializations

$$\lim_{n\to\infty} E\left\{\exp\left[i\lambda \frac{Z_{[nt]} - c_n}{b_n}\right]\middle| Z_0 = c_n\right\} = \int_{-\infty}^{\infty} e^{i\lambda x}\, dH_t(x) = \varphi_t(\lambda), \tag{5.3}$$

and the conclusion of the theorem is equivalent to the assertion that

$$\varphi_{t+s}(\lambda) = \varphi_t(\lambda)\varphi_s(\lambda). \tag{5.4}$$

The proof is based on the expression for $\varphi_t(\lambda)$ in terms of the basic generating function $p(x)$; in the same way as in (4.17) we have

$$\varphi_t(\lambda) = \lim_{n\to\infty} \exp\left[\frac{-i\lambda c_n}{b_n}\right] p_{[nt]}\left(\exp\left[\frac{i\lambda}{b_n}\right]\right)^{c_n}, \tag{5.5}$$

where the subscript on p refers to functional iteration. From (5.5) and basic branching property we can then write

$$\varphi_{t+s}(\lambda) = \lim_{n\to\infty} \exp\left[\frac{-i\lambda c_n}{b_n}\right] p_{[ns]}\left\{p_{[nt]}\left(\exp\left[\frac{i\lambda}{b_n}\right]\right)\right\}^{c_n}. \tag{5.6}$$

But (5.5) can be rewritten in the form

$$p_{[nt]}\left(\exp\left[\frac{i\lambda}{b_n}\right]\right) = [\varphi_t(\lambda) + o(1)]^{1/c_n} \exp\left[\frac{i\lambda}{b_n}\right], \tag{5.7}$$

which we can substitute in (5.6). (Since $\varphi_t(\lambda)$ must be infinitely divisible, its roots and logarithm are unambiguously defined.) It remains to study how $p_{[ns]}$ of this quantity behaves as $n \to \infty$.

To accomplish this, we will use a slight generalization of (5.3). Let ξ_n be any sequence of numbers tending to a limit ξ; then

$$\lim_{n\to\infty} E\left\{\exp\left[i\lambda \frac{Z_{[ns]} - c_n}{b_n} + \xi_n \frac{Z_{[ns]}}{c_n}\right]\middle| Z_0 = c_n\right\} = e^{\xi}\varphi_s(\lambda). \tag{5.8}$$

The proof of (5.8) is simple: replace the second $Z_{[ns]}$ by $Z_0 + (Z_{[ns]} - Z_0)$. The Z_0 $(= c_n)$ gives rise in the limit to the factor e^{ξ}, whereas the remainder has no effect. The reason for this is that by hypothesis $(Z_{[ns]} - Z_0)/b_n$ has a limiting distribution, and $b_n/c_n \to 0$, so that $\xi_n(Z_{[ns]} - Z_0)/c_n \to 0$ in probability. In terms of the function p, (5.8) becomes

$$\lim_{n\to\infty} \exp\left[\frac{-i\lambda c_n}{b_n}\right] p_{[ns]}\left(\exp\left[\frac{i\lambda}{b_n} + \frac{\xi_n}{c_n}\right]\right)^{c_n} = e^{\xi}\varphi_s(\lambda), \tag{5.9}$$

which in turn can be written in the form

$$p_{[ns]}\left(\exp\left[\frac{i\lambda}{b_n} + \frac{\xi_n}{c_n}\right]\right) = \exp\left[\frac{i\lambda}{b_n} + \frac{\xi + \log \varphi_s(\lambda) + o(1)}{c_n}\right]. \tag{5.10}$$

It is very easy to complete the proof. To apply the function $p_{[ns]}$ to the right-hand side of (5.7), we let $\xi_n = \log[\varphi_t(\lambda) + o(1)]$ (the $o(1)$ depends on n) so that $\xi = \log \varphi_t(\lambda)$. From (5.9), then,

$$p_{[ns]}\left\{p_{[nt]}\left(\exp\left[\frac{i\lambda}{b_n}\right]\right)\right\} = \exp\left[\frac{i\lambda}{b_n} + \frac{\log \varphi_t(\lambda) + \log \varphi_s(\lambda) + o(1)}{c_n}\right], \tag{5.11}$$

and substituting this in (5.6) yields the desired conclusion (5.4).

It is natural to ask next which additive processes will in fact arise. This problem is in a state comparable to the question about limiting distributions, which was discussed in sections 2 and 3. When the variance is finite we can immediately deduce the answer from theorem 2.3. *Suppose that* $\{p_i\}$ *has mean* 1 *and variance* $\sigma^2 < \infty$, *and that* $n^{-1}c_n \to \infty$. *Then if* $a_n = c_n$ *and* $b_n = \sigma\sqrt{nc_n}$, (5.1) *and its related conditions hold, and* $H_t(x)$ *is the distribution at time* t *of a Brownian motion process.* Furthermore, it is possible to see that *stable processes of index* $1 < \alpha < 2$ *which attract nonnegative random variables can occur as limits.* The proof of this fact consists of combining the possibility of limit laws satisfying (4.8) (which are attracted to a stable law) with an argument similar to that used to prove theorem 3.1. The question of whether these are the only possibilities has not been answered.

6. Final remarks

This paper is intended as much to raise questions as to present answers. It is clear that even within the scope of the problems defined in the introduction the results obtained here are far from complete. Both sections 3 and 5 are very sketchy and end with conjectures, while the problem studied in section 3 is not attacked at all for $\mu \leq 1$. (The method used for $\mu > 1$ could be repeated for $\mu \leq 1$, but would not seem to lead very far toward a complete solution in those cases either.) Above all, except when $\sigma^2 < \infty$ nothing has been done to determine the class of $\{p_i\}$ for which a particular limit law or process actually can occur. This problem of "domains of attraction" is certainly of considerable interest.

Despite these gaps in the solution of our simple problem, we will mention two more general ones. The first is to discuss such extensions as the multidimensional case, and age-dependent processes. The second generalization applies even in the one-dimensional Markov case. In many investigations of limit properties of branching processes, it is important to make the probabilities $\{p_i\}$, as well as the initial state, functions of n; essentially new phenomena arise in this way. In [1], for instance, W. Feller finds a class of diffusions *with drift* as limiting processes, in contrast to the essentially unique diffusion (without drift) we obtained in section 4. Other results where the transition probabilities change during the passage to the limit may be found in [4]. Again, with this more general procedure, the questions of determining all possible limits and the conditions under which they occur can be raised. It appears that a complete discussion will be both interesting and difficult.

Remark added in proof. Considerable progress has now been made toward a solution of the last problem above. I hope to publish these results in the near future.

REFERENCES

[1] W. Feller, "Diffusion processes in genetics," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1951, pp. 227–246.

[2] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, Cambridge, Addison-Wesley, 1954. (Translated by K. L. Chung.)

[3] T. E. Harris, *The Theory of Branching Processes*, Berlin, Springer, 1963.

[4] B. A. Sevastyanov, "Transient phenomena for branching stochastic processes," *Theor. Probability Appl.*, Vol. 4 (1959), pp. 113–128.

[5] V. M. Zolotarev, "More exact statements of several theorems in the theory of branching processes," *Theor. Probability Appl.*, Vol. 2 (1957), pp. 245–253.