The Annals of Probability
1975, Vol. 3, No. 5, 790-801

THE FIRST BIRTH PROBLEM FOR AN AGE-DEPENDENT BRANCHING PROCESS

BY J. F. C. KINGMAN

University of Oxford

If B_n denotes the time of the first birth in the nth generation of an age-dependent branching process of Crump-Mode type, then under a weak condition there is a constant γ such that $B_n/n \to \gamma$ as $n \to \infty$, almost surely on the event of ultimate survival. This strengthens a result of Hammersley, who proved convergence in probability for the more special Bellman–Harris process. The proof depends on a class of martingales which arise from a 'collective marks' argument.

1. The problem. Hammersley [9] has considered the following problem for an age-dependent Bellman–Harris process (see [11] for definitions and basic theory). Let B_n be the time of the first birth in the nth generation, or equivalently of the first death in the $(n-1)$th generation. If the expected number of children born to an individual exceeds 1, then the event S of ultimate survival has positive probability, and B_n is defined on S for all $n \geqq 1$. What can be said about the asymptotic properties of B_n as $n \to \infty$? He shows that there is a constant γ, which may be calculated from the distributions defining the process, such that

$$B_n/n \to \gamma \tag{1.1}$$

in conditional probability given S.

Hammersley's techniques depend on an ingenious theory of 'superconvolutive' sequences of distributions. They can be supplemented by a result of Kesten [12], which appears to be sufficiently powerful to establish the convergence (1.1) almost everywhere on S. The purpose of this paper is to present a quite different approach to (1.1), which seems to me somewhat more transparent, and may be useful for other problems about branching processes. This approach applies without extra cost to the more general processes defined by Crump and Mode [4], and the analysis will therefore be set within this broader context.

In the Crump–Mode model, an initial ancestor is born at $t = 0$ and then produces children at random throughout his lifetime. We shall not be concerned with death, and thus it makes no difference whether his lifetime is finite or infinite, or whether (as in Doney's generalisation [6] of the Crump–Mode model) he is allowed to produce posthumous children. If $Z_1(t)$ denotes the number of children born to this ancestor before time t, then $Z_1(t)$ will be an arbitrary counting process (i.e. a positive, increasing, right continuous, integer-valued random process. Words like 'positive' and 'increasing' are to be understood in

Received December 2, 1974; revised March 12, 1975.

AMS 1970 *subject classifications*. Primary 60J80; Secondary 60F10, 60F15, 60G45.

Key words and phrases. Age-dependent branching processes, almost sure convergence, subadditive processes, martingales.

the weak sense unless qualified by 'strictly'.) In the next section a condition (2.3) will be imposed on Z_1, which will imply that $Z_1(t)$ is finite for each finite t, but the possibility is not excluded that the ancestor will live for ever and have an infinite family.

The important feature of the Crump–Mode model is that the children of the ancestor, from their several births, behave independently of one another and of their parent, producing children at random according to random processes with the same joint distributions as Z_1. Their children produce children in the same way, and so on. For a formal definition of this process, we refer to [4].

Let $Z_n(t)$ denote the number of nth generation individuals born before time t, so that Z_n is again a counting process, and write

$$\nu = \sup\{n; Z_n(t) > 0 \text{ for some } t\}, \tag{1.2}$$

so that $0 \leqq \nu \leqq \infty$ and S, the event of ultimate survival, is given by

$$S = \{\nu = \infty\}. \tag{1.3}$$

For each finite $n \leqq \nu$, denote by

$$B_{n1} \leqq B_{n2} \leqq B_{n3} \leqq \cdots \tag{1.4}$$

the instants at which the births in the nth generation occur, arranged in ascending order, so that

$$Z_n(t) = \#\{r; B_{nr} \leqq t\} \tag{1.5}$$

and

$$B_n = B_{n1}. \tag{1.6}$$

The aim of this analysis is to show that, under the weak condition (2.3), the limiting equation (1.1) holds almost everywhere on S, and to exhibit the constant γ in terms of the distributions of the process Z_1.

2. The martingales. The basic tool of the present analysis is the function

$$\phi(\theta) = \mathbf{E}\{\textstyle\sum_r \exp(-\theta B_{1r})\} \qquad \theta \geqq 0, \tag{2.1}$$

where the summation runs over all the children of the initial ancestor. This sum may have a finite or infinite number of terms, and is to be taken as 0 if it is empty. By Fubini's theorem, when $\theta > 0$,

$$\phi(\theta) = \int_0^\infty e^{-\theta t}\, d\{\mathbf{E}Z_1(t)\} = \theta \int_0^\infty e^{-\theta t}\mathbf{E}\{Z_1(t)\}\, dt. \tag{2.2}$$

Note that $\phi(0)$ is the expected number of children of any individual, and in order that S should have positive probability it is sufficient (and almost necessary [11]) that $\phi(0) > 1$. We do not exclude the possibility that $\phi(0) = \infty$, but it is then necessary to suppose that $\phi(\theta)$ is finite for some (and then for all larger) θ. These two conditions are neatly combined by requiring that

$$1 < \phi(\theta_0) < \infty \tag{2.3}$$

for some $\theta_0 > 0$, and this will be assumed in all that follows.

THEOREM 1. *If $\mathscr{F}_n$ is the σ-field generated by the births in the first n generations, then for all $\theta \geqq \theta_0$,*

(2.4) $$\mathrm{E}\{\sum_r \exp(-\theta B_{n+1,r}) \mid \mathscr{F}_n\} = \phi(\theta) \sum_r \exp(-\theta B_{nr}) .$$

Hence

(2.5) $$\mathrm{E}\{\sum_r \exp(-\theta B_{nr})\} = \phi(\theta)^n .$$

and

(2.6) $$W_n(\theta) = \phi(\theta)^{-n} \sum_r \exp(-\theta B_{nr}) = \phi(\theta)^{-n} \int_0^\infty e^{-\theta t}\, dZ_n(t)$$

defines a martingale with respect to the σ-fields $\mathscr{F}_n$.

PROOF. Each birth in the sum

$$\sum_r \exp(-\theta B_{n+1,r})$$

comes from one of the individuals in the nth generation, so that the sum may be written as

$$\sum_r \exp(-\theta B_{nr})\{\sum_s \exp[-\theta(B_{n+1,r,s} - B_{nr})]\} ,$$

where $B_{n+1,r,s}$ is the birth time of the sth child of the rth individual in the nth generation. The inner sum is of the form

$$\sum_s \exp(-B'_{1s}) ,$$

where the sequence (B'_{1s}) is independent of $\mathscr{F}_n$ and has the same joint distributions as (B_{1s}). Hence, since B_{nr} is $\mathscr{F}_n$-measurable,

$$\begin{aligned}\mathrm{E}\{\sum_r \exp(-\theta B_{n+1,r}) \mid \mathscr{F}_n\} &= \sum_r \exp(-\theta B_{nr})\mathrm{E}\{\sum_s \exp(-\theta B_{1s})\} \\ &= \sum_r \exp(-\theta B_{nr})\phi(\theta) .\end{aligned}$$

Thus (2.4) is proved, and it follows that $W_n(\theta)$ is a martingale. Since $\mathrm{E}W_1(\theta) = 1$, $\mathrm{E}W_n(\theta) = 1$, which establishes (2.5) and completes the proof.

When $\theta = 0$, and if $\phi(0) < \infty$, the martingale W_n is an important tool in the theory of branching processes [11], [1], and it is possible that it may have other uses for general values of θ. We shall need to use the fact that, because $W_n(\theta) \geqq 0$, the martingale convergence theorem shows that the limit

(2.7) $$W(\theta) = \lim_{n\to\infty} W_n(\theta)$$

exists, and satisfies

(2.8) $$\mathrm{E}\{W(\theta)\} \leqq 1 \qquad (\theta \geqq \theta_0) .$$

The identity (2.5) is in the spirit of van Dantzig's 'method of collective marks' [5]. Suppose that the individuals are subject to a disease, which kills any individual in a time interval $(t, t + h)$ with probability $\theta h + o(h)$, and acts independently on different individuals and in disjoint intervals. The result is a new Crump–Mode process, in which the mean family size is reduced from $\phi(0)$ to $\phi(\theta)$. Conditional on the original process, the probability that the nth generation individual due to be born at B_{nr} is not preempted by the death of an ancestor is $\exp(-\theta B_{nr})$. Thus (2.5) expresses the obvious fact that the expected number in the nth generation of the modified process is $\phi(\theta)^n$.

3. The lower bound. Equation (2.5) shows that, for $\theta \geqq \theta_0$ and any constant a,

$$\phi(\theta)^n \geqq \mathbf{E}\{\exp(-\theta B_n); n \leqq \nu\} \geqq \mathbf{E}\{\exp(-\theta B_n); B_n \leqq na, S\}$$
$$\geqq e^{-\theta na}\mathbf{P}\{B_n \leqq na, S\},$$

so that

$$\mathbf{P}\{B_n \leqq na, S\} \leqq [\phi(\theta)e^{\theta a}]^n .$$

The Borel–Cantelli lemma therefore shows that, as $n \to \infty$,

$$\liminf B_n/n \geqq a \tag{3.1}$$

almost everywhere on S, so long as a satisfies

$$\phi(\theta)e^{\theta a} < 1$$

for some $\theta \geqq \theta_0$. Writing

$$\mu(a) = \inf\{\phi(\theta)e^{\theta a}; \theta \geqq \theta_0\},\tag{3.2}$$

this shows that (3.1) is true if $\mu(a) < 1$. It is clear that $\mu(a)$ is an increasing function of $a \geqq 0$, and since $\phi(\theta) \to 0$ as $\theta \to \infty$ (by monotone convergence),

$$\mu(a) \to 0 \quad (a \to 0).$$

Hence $\mu(a) < 1$ for small values of $a > 0$. If we define

$$\gamma = \sup\{a; \mu(a) < 1\},\tag{3.3}$$

then (3.1) yields the following theorem.

THEOREM 2. *If* (2.3) *holds, then as* $n \to \infty$,

$$\liminf B_n/n \geqq \gamma \tag{3.4}$$

almost everywhere on S, *where* γ *is defined by* (3.3) *and* (3.2).

Suppose now that γ is finite. The increasing function

$$\log \mu(a) = \inf\{\log \phi(\theta) + \theta a; \theta \geqq \theta_0\}\tag{3.5}$$

is clearly concave, and therefore continuous, on $\{a; \mu(a) > 0\}$. Moreover, $\mu(a) = 0$ only if $\mathbf{P}\{Z_1(a) = 0\} = 1$. Hence, if we exclude the case

$$\mathbf{P}\{Z_1(a) = 0, \text{ all } a < \gamma\} = 1,\tag{3.6}$$

we have $\mu(\gamma) = 1$. Moreover, if (3.6) is excluded, then

$$\phi(\theta_0)e^{\theta_0\gamma} > 1, \qquad \lim_{\theta\to\infty} \phi(\theta)e^{\theta\gamma} = \infty .$$

Hence $\mu(\gamma)$ must be attained at an interior point ϑ of (θ_0, ∞), and we have the equations

$$\phi(\vartheta) = e^{-\vartheta\gamma}, \qquad \phi'(\vartheta) = -\gamma\phi(\vartheta)\tag{3.7}$$

satisfied by ϑ and γ. (In particular cases, (3.7) would usually be the starting-point for the evaluation of γ.) Note also that since $\phi(\theta) \geqq e^{-\theta\gamma}$ for all $\theta \geqq \theta_0$, it follows that $\mu(a) \geqq \inf e^{\theta(a-\gamma)} \geqq e^{\theta_0(a-\gamma)} > 1$ when $a > \gamma$;

$$\mu(a) > 1 \quad (a > \gamma) . \tag{3.8}$$

In the special case of the Bellman–Harris process (when all the birthtimes B_{1r} are equal) it is trivial to check that the definition of γ is equivalent to that of Hammersley ([9], equation (3.97)). Hence this argument gives a simple proof of one side of Hammersley's result.

Equation (2.5) shows that the increasing function

$$\mathbf{E}\{Z_n(t)\} \tag{3.9}$$

is the n-fold Stieltjes convolution of

$$\mathbf{E}\{Z_1(t)\} \tag{3.10}$$

with itself, and this suggests the estimation of (3.9) by standard techniques for sums of independent random variables. For example, although Chernoff's theorem [3] is not directly applicable because (3.10) is not a distribution function, the simple proof given by Bahadur and Ranga Rao [2] applies without change, to show that

$$\mathbf{E}\{Z_n(na)\} \geqq m^n \tag{3.11}$$

for large n, whenever $a > \gamma$ and $1 < m < \mu(a)$. If we could deduce from (3.11) that

$$\mathbf{P}\{Z_n(na) > 0\} \not\to 0$$

as $n \to \infty$, this would suffice to prove (1.1), but such a deduction would not be valid without further argument. However, (3.11) is of some interest in its own right, and may be made more precise by using the methods of [2] in a rather obvious way to show that, under slight further conditions,

$$\mathbf{E}\{Z_n(na)\} \sim c(a)n^{-\frac{1}{2}}\mu(a)^n \tag{3.12}$$

as $n \to \infty$, for a suitable constant $c(a)$.

4. The limit $W(\theta)$. It is a commonplace of martingale theory that, if the limit of a positive martingale has the same expectation as the martingale, then it may be used to close it [7]. In particular, if (2.8) can be replaced by the equality

$$\mathbf{E}\{W(\theta)\} = 1 , \tag{4.1}$$

then

$$W_n(\theta) = \mathbf{E}\{W(\theta) \mid \mathscr{F}_n\} . \tag{4.2}$$

In case $\theta = 0$ (and $\phi(0) < \infty$) this question has been the subject of much research (see [11] and [1]), culminating in the Kesten–Stigum theorem that (4.1) holds if and only if $\mathbf{E}\{W_1(0)\,|\log W_1(0)|\} < \infty$. Fortunately, we shall not have to generalise this deep result, and it will suffice to use a simpler (and earlier) technique used by Harris to deal with the problem when $\mathbf{E}\{W_1(0)^2\} < \infty$. Harris estimated $\mathbf{E}\{W_n(0)^2\}$, and this works with $W_n(\theta)$ provided that θ is sufficiently small. This is not enough, and further values of θ can be dealt with by estimating $\mathbf{E}\{W_n(\theta)^\alpha\}$ for values of α slightly greater than 1. By suitable choice of α, this proves (4.1) whenever $\theta < \vartheta$, and this result is best possible since (4.1) is false for $\theta > \vartheta$ (as will become clear in the proof of Theorem 5). I do not know what happens in the critical case $\theta = \vartheta$.

THEOREM 3. *Suppose that* (2.3) *holds, and that* α *and* θ *satisfy* $\theta \geqq \theta_0$, $1 < \alpha < 2$,

$$\mathbf{E}\{W_1(\theta)^2\} < \infty \tag{4.3}$$

and

$$\phi(\alpha\theta) < \phi(\theta)^\alpha . \tag{4.4}$$

Then

$$\sup \mathbf{E}\{W_n(\theta)^\alpha\} < \infty , \tag{4.5}$$

and therefore (4.1) *and* (4.2) *hold.*

LEMMA. *Let* λ_r $(r \geqq 1)$ *be positive constants, with finite sum* Λ. *Let* Y_r $(r \geqq 1)$ *be uncorrelated positive random variables, each with nonzero mean* μ *and finite variance* σ^2. *Then, for* $1 < \alpha < 2$,

$$\mathbf{E}\{(\textstyle\sum_r \lambda_r Y_r)^\alpha\} \leqq \Lambda^\alpha\mu^\alpha + (\alpha - 1) \sum_r \lambda_r{}^\alpha \mu^{\alpha-2}\sigma^2 . \tag{4.6}$$

PROOF OF LEMMA. For any positive random variable X, Hölder's inequality implies that

$$\mathbf{E}(X^\alpha) \leqq [\mathbf{E}(X)]^{2-\alpha}[\mathbf{E}(X^2)]^{\alpha-1} .$$

If $X = \sum \lambda_r Y_r$, then $\mathbf{E}(X) = \Lambda\mu$ and

$$\begin{aligned}\mathbf{E}(X^2) &= \Lambda^2\mu^2 + \textstyle\sum_r \lambda_r{}^2\sigma^2 \leqq \Lambda^2\mu^2 + \sum_r \lambda_r{}^\alpha\Lambda^{2-\alpha}\sigma^2 \\ &= \Lambda^2\mu^2(1 + \Lambda^{-\alpha}\mu^{-2} \textstyle\sum_r \lambda_r{}^\alpha\sigma^2) .\end{aligned}$$

Hence

$$\begin{aligned}\mathbf{E}(X^\alpha) &\leqq (\Lambda\mu)^\alpha(1 + \Lambda^{-\alpha}\mu^{-2} \textstyle\sum_r \lambda_r{}^\alpha\sigma^2)^{\alpha-1} \\ &\leqq (\Lambda\mu)^\alpha(1 + (\alpha - 1)\Lambda^{-\alpha}\mu^{-2} \textstyle\sum_r \lambda_r{}^\alpha\sigma^2) \\ &= \Lambda^\alpha\mu^\alpha + (\alpha - 1) \textstyle\sum_r \lambda_r{}^\alpha\mu^{\alpha-2}\sigma^2 .\end{aligned}$$

PROOF OF THEOREM 3. In the notation of the proof of Theorem 1,

$$\mathbf{E}\{[\textstyle\sum_r \exp(-\theta B_{n+1,r})]^\alpha \,|\, \mathscr{F}_n\} = \mathbf{E}\{(\sum_r \lambda_r Y_r)^\alpha \,|\, \mathscr{F}_n\} ,$$

where $\lambda_r = \exp(-\theta B_{nr})$ is constant conditional on $\mathscr{F}_n$, and

$$Y_r = \textstyle\sum_s \exp[-\theta(B_{n+1,r,s} - B_{nr})]$$

satisfies the conditions of the lemma with

$$\mu = \phi(\theta) , \qquad \sigma^2 = \phi(\theta)^2 \operatorname{Var}[W_1(\theta)] .$$

Hence

$$\mathbf{E}\{[\sum_r \exp(-\theta B_{n+1,r})]^\alpha \mid \mathscr{F}_n\}$$
$$\leqq [\sum_r \exp(-\theta B_{nr})]^\alpha \phi(\theta)^\alpha + (\alpha - 1) \sum_r [\exp(-\theta B_{nr})]^\alpha \phi(\theta)^\alpha \operatorname{Var}[W_1(\theta)],$$

so that

$$\mathbf{E}\{W_{n+1}(\theta)^\alpha \mid \mathscr{F}_n\} \leqq W_n(\theta)^\alpha + (\alpha - 1)\phi(\theta)^{-n\alpha}\phi(\alpha\theta)^n W_n(\alpha\theta) \operatorname{Var}[W_1(\theta)].$$

Taking expectations and summing over n, we have

$$\mathbf{E}\{W_n(\theta)^\alpha\} \leqq 1 + (\alpha - 1) \operatorname{Var}[W_1(\theta)] \sum_{m=0}^{n-1} \{\phi(\alpha\theta)/\phi(\theta)^\alpha\}^m,$$

so that (4.5) follows from (4.4). Moreover, (4.5) implies that the martingale $W_n(\theta)$ is uniformly integrable, and (4.1) and (4.2) are consequences of standard martingale theory ([7], Theorem 4.1).

Corollary 1. *Excluding* (3.6), (4.1) *and* (4.2) *hold for all values of* θ *in* $\theta_0 \leqq \theta < \vartheta$ *for which* (4.3) *is satisfied.*

Proof. It is necessary only to show that α exists in $1 < \alpha < 2$ such that (4.4) is satisfied. If $\phi(\theta) \geqq 1$ this is trivially so for any such α, since ϕ is strictly decreasing. Hence we may suppose that $\phi(\theta) < 1$, and this means that θ lies in the interval (θ_0, ∞) in which ϕ has derivatives of all orders. Then (4.4) will hold with α greater than, but sufficiently near, 1 if

$$\frac{\partial}{\partial\alpha}[\phi(\alpha\theta) - \phi(\theta)^\alpha] < 0 \quad at \quad \alpha = 1,$$

that is, if

$$\theta\phi'(\theta) < \phi(\theta)\log\phi(\theta),$$

or equivalently, if

$$\psi(\theta) = \frac{\theta\phi'(\theta)}{\phi(\theta)} - \log\phi(\theta) < 0.$$

From (3.7), $\psi(\vartheta) = 0$, and for $\theta > \theta_0$,

$$\psi'(\theta) = \theta\left\{\frac{\phi''(\theta)}{\phi(\theta)} - \left(\frac{\phi'(\theta)}{\phi(\theta)}\right)^2\right\} > 0$$

as a consequence of (2.2) and Schwarz's inequality (the case when ψ' is identically zero being ruled out as a sub-case of (3.6)). Hence $\psi(\theta) < 0$ for all $\theta_0 \leqq \theta < \vartheta$, as required.

Corollary 2. *If* (3.6) *is true, then* (4.1) *and* (4.2) *hold for all values of* $\theta \geqq \theta_0$ *for which* (4.3) *is satisfied.*

Proof. The function

(4.7) $$\phi(\theta)e^{\theta\gamma} = \int_\gamma^\infty e^{-\theta(t-\gamma)}\, d\{\mathbf{E}Z_1(t)\}$$

is either constant (in which case (2.3) shows that (4.4) is trivially satisfied), or is strictly decreasing and (by (3.3)) $\geqq 1$, so that

$$\phi(\alpha\theta)e^{\alpha\theta\gamma} < \phi(\theta)e^{\theta\gamma} \leqq (\phi(\theta)e^{\theta\gamma})^{\alpha},$$

so that (4.4) holds for any $\alpha > 1$.

5. Subadditivity. In [8] Hammersley adduced the first birth problem as evidence that a strengthening which I had adopted in [13] and [14] of the original Hammersley–Welsh postulates [10] for subadditive processes was such as to rule out interesting applications. Later ([9], Note 9) he withdrew this objection in the light of remarks of Joshi, which made it clear that the problem could not be fitted directly into subadditive ergodic theory. Nevertheless, the methods of [10] and [13] do play a part in the proof of the main result.

THEOREM 4. *Suppose that every individual has at least one child, and that*

$$\mathbf{E}(B_1) < \infty . \tag{5.1}$$

Then there is a constant Γ *such that*

$$E\left|\frac{B_n}{n} - \Gamma\right| \to 0 . \tag{5.2}$$

as $n \to \infty$, *and if* (2.3) *holds, then*

$$\Gamma \geqq \gamma . \tag{5.3}$$

PROOF. For any strictly positive integer k, we define by induction on m a sequence of individuals ${}_kA_m$ ($m \geqq 0$), such that ${}_kA_0$ is the initial ancestor and ${}_kA_m$ is in the mkth generation. If ${}_kA_m$ has been defined, then by hypothesis he has at least one descendant in the $(m+1)k$th generation; let ${}_kA_{m+1}$ be the first-born of these. Define ${}_k\beta_m$ as the time between the births of ${}_kA_{m-1}$ and ${}_kA_m$. Then, for fixed k, the random variables ${}_k\beta_m$ are independent, with the same distribution as

$${}_k\beta_1 = B_k . \tag{5.4}$$

Moreover, since ${}_kA_m$ cannot be born before $B_{(mk)}$, we have

$$B_{(mk)} \leqq \textstyle\sum_{r=1}^{m} {}_k\beta_r . \tag{5.5}$$

From (5.1), (5.4) and (5.5), $\mathbf{E}({}_k\beta_m) < \infty$ for all k and m, and we can apply the strong law of large numbers to (5.5) to conclude that, as $m \to \infty$,

$$\limsup B_{(mk)}/m \leqq \mathbf{E}({}_m\beta_1) = \mathbf{E}(B_k)$$

with probability one. Since B_n increases with n, it follows that, as $n \to \infty$,

$$\limsup B_n/n \leqq \mathbf{E}(B_k/k) , \tag{5.6}$$

with probability one.

Consider the positive random variables

$$X_n = n^{-1}(\textstyle\sum_{r=1}^{n} {}_1\beta_r - B_n) ,$$

which have finite expectations

$$x_n = \mathbf{E}(X_n) = \mathbf{E}(B_1 - B_n/n)\ .$$

Then (5.6) implies that

$$X_* = \liminf X_n \geqq x_k$$

with probability one. By Fatou's lemma,

$$\liminf x_n \geqq \mathbf{E}(X_*) \geqq x_k$$

for all k, which shows that

$$x = \lim x_n$$

exists. Moreover, since

$$P(X_* \geqq x) = 1\ , \qquad \mathbf{E}(X_*) \leqq x\ ,$$

we must have

$$P(X_* = x) = 1\ .$$

The random variables

$$Y_n = \inf\{X_m;\ m \geqq n\}$$

increase to $X_* = x$ as $n \to \infty$, so that

$$\mathbf{E}|Y_n - x| \to 0$$

by monotone convergence. Since $Y_n \leqq X_n$,

$$\mathbf{E}|X_n - Y_n| = \mathbf{E}(X_n) - \mathbf{E}(Y_n) \to x - x = 0\ ,$$

so that

$$\mathbf{E}|X_n - x| \to 0\ .$$

Since

$$\mathbf{E}|n^{-1} \textstyle\sum_{r=1}^{n} {}_1\beta_r - \mathbf{E}(B_1)| \to 0$$

by the strong law, this establishes (5.2) with $\Gamma = \mathbf{E}(B_1) - x$. Letting $k \to \infty$ in (5.6), we have

(5.7) $$\limsup B_n/n \leqq \Gamma$$

with probability one. If (2.3) holds, then comparison of (5.7) and (3.4) establishes (5.3), and the proof is complete.

6. The main theorem.

THEOREM 5. *Under condition* (2.3),

(6.1) $$\lim_{n\to\infty} B_n/n = \gamma$$

holds almost everywhere on S.

PROOF. If $\gamma = \infty$, the result is a consequence of Theorem 2, and we therefore suppose γ finite. We prove the theorem first under three unecessary restrictions, which will then be successively removed. Thus we suppose that each individual has at least one child,

(6.2) $$\mathbf{P}\{Z_1(\infty) \geqq 1\} = 1\ ,$$

that

$$\mathrm{E}(B_1) < \infty \,, \tag{6.3}$$

and that, for some θ_1 in $\theta_0 < \theta_1 < \vartheta$,

$$\mathrm{E}\{W_1(\theta_1)^2\} < \infty \,. \tag{6.4}$$

Note that, since $\phi(\theta)W_1(\theta)$ decreases with θ, (6.4) implies that (4.3) holds for all $\theta \geqq \theta_1$. Hence the corollaries to Theorem 3 show that (4.1) and (4.2) are true whenever $\theta_1 \leqq \theta < \vartheta$, or in case (3.6) whenever $\theta \geqq \theta_1$.

For any such θ, differentiate (2.5) to give

$$\begin{aligned} -n\phi(\theta)^{n-1}\phi'(\theta) &= \mathrm{E}\{\textstyle\sum_r B_{nr} \exp(-\theta B_{nr})\} \\ &\geqq \mathrm{E}\{B_n \textstyle\sum_r \exp(-\theta B_{nr})\} \,, \end{aligned}$$

so that

$$\begin{aligned} -\phi'(\theta)/\phi(\theta) &\geqq \mathrm{E}\{n^{-1}B_n W_n(\theta)\} = \mathrm{E}\{n^{-1}B_n W(\theta)\} \\ &\geqq \mathrm{E}\{n^{-1}B_n W(\theta);\ W(\theta) \leqq w\} \end{aligned}$$

for any finite w, where we have used (4.2) and the fact that B_n is $\mathscr{F}_n$-measurable. Letting $n \to \infty$ and using (5.2),

$$-\phi'(\theta)/\phi(\theta) \geqq \mathrm{E}\{\Gamma W(\theta);\ W(\theta) \leqq w\} \,.$$

Letting $w \to \infty$ and using (4.1),

$$-\phi'(\theta)/\phi(\theta) \geqq \Gamma \mathrm{E}\{W(\theta)\} = \Gamma \,,$$

and letting $\theta \to \vartheta$ (or $\theta \to \infty$ in case (3.6), remembering (4.7),)

$$\gamma \geqq \Gamma \,.$$

Hence, from (5.3), $\Gamma = \gamma$, and comparing (3.4) and (5.7), the conclusion (6.1) follows with probability one.

It remains to remove the restrictions (6.2), (6.3) and (6.4). In any Crump–Mode process, describe an individual as *fecund* if he has descendants in each succeeding generation. Then the initial ancestor is fecund if and only if S occurs, and the probability of any individual being fecund is $\mathbf{P}(S)$. An individual is fecund if and only if he has at least one fecund child. It is easy to check that, conditional on S, the fecund individuals form a new Crump–Mode process. If quantities relating to this new process are distinguished by a bar (and if $\bar{\mathbf{P}}$ and $\bar{\mathrm{E}}$ denote probability and expectation conditional on S) then although $\bar{Z}_1$ has a different structure from Z_1, it is clearly true that $\bar{\mathrm{E}}\{\bar{Z}_1(t)\} = \mathrm{E}\{Z_1(t)\}$, so that $\bar{\phi} = \phi$ and $\bar{\gamma} = \gamma$.

For the new process, (6.2) is automatically satisfied, and hence from what has already been proved we can conclude that

$$\bar{\mathbf{P}}\{\lim \bar{B}_n/n = \bar{\gamma}\} = 1$$

so long as

$$\bar{\mathrm{E}}(\bar{B}_1) < \infty \,, \qquad \bar{\mathrm{E}}\{\bar{W}_1(\theta)^2\} < \infty \,. \tag{6.5}$$

Since $B_n \leqq \bar{B}_n$, this means that

$$\limsup B_n/n \leqq \gamma$$

almost surely on S, and comparison with (3.4) then proves (6.1). Since $\bar{W}_1 \leqq W_1$, (6.5) follows from (6.4) and

$$\mathbf{E}(\bar{B}_1) < \infty . \tag{6.6}$$

Hence we have proved (6.1) under the conditions (6.4) and (6.6).

To remove (6.6), we modify the process in the manner described at the end of Section 2, but writing δ rather than θ for the parameter measuring the virulence of the disease. It is clearly possible to arrange that the modified processes for all positive values of δ are defined on the same probability space, in such a way that death for one value of δ implies death for all larger values. Then, distinguishing quantities in the modified process by a superscript δ, we see that

$$\phi^\delta(\theta) = \phi(\theta + \delta) , \tag{6.7}$$

so that (2.3) holds for all sufficiently small $\delta > 0$. The modification does not disturb (6.4), and (6.5) holds for all $\delta > 0$, because (on S^δ), $\bar{B}_1$ is less than the lifetime of the initial ancestor, which has finite expectation δ^{-1}.

Hence, if (6.4) holds and $\delta > 0$, we can conclude from the fact that $B_n \leqq \bar{B}_n$ that

$$\gamma \leqq \liminf B_n/n \leqq \limsup B_n/n \leqq \gamma^\delta$$

almost everywhere on S^δ. As $\delta \to 0$, S^δ increases to an event $S^+ \subseteqq S$, and it follows easily from (6.7) that $\gamma^\delta \to \gamma$. Hence (6.1) will be proved (still assuming (6.4)) if it can be shown that $\mathbf{P}(S - S^+) = 0$. To show this, note that $s(\delta) = \mathbf{P}(S^\delta)$ is the root in (0, 1) of the equation [11]

$$f_\delta\{1 - s(\delta)\} = 1 - s(\delta) ,$$

where

$$f_\delta(x) = \textstyle\sum_{m=0}^{\infty} \mathbf{P}\{Z_1^\delta(\infty) = m\}x^m$$
$$= \mathbf{E}\{\textstyle\prod_r [1 - (1 - x)\exp(-\delta B_{1r})]\} ,$$

so long as $\phi(\delta) > 1$. From this it is immediate that

$$\mathbf{P}(S^+) = \lim_{\delta\to 0} s(\delta) = \mathbf{P}(S) ,$$

and therefore (6.1) is proved under conditions (2.3) and (6.4).

Finally, we remove (6.4) by a truncation argument. Insist that any individual is arbitrarily sterilised after producing N children, and denote this modification by a suffix N. For the modified process, (6.4) is satisfied, and B_n is increased, so that, using (3.4) again,

$$\gamma \leqq \liminf B_n/n \leqq \limsup B_n/n \leqq \gamma_N$$

almost everywhere on S_N. The survival probability $s_N = \mathbf{P}(S_N)$ satisfies

$$\textstyle\sum_{m=0}^{N} \mathbf{P}\{Z_1(\infty) = m\}(1 - s_N)^m = 1 - s_N ,$$

from which $s_N \to \mathbf{P}(S)$ as $N \to \infty$. Hence, as $N \to \infty$, S_N increases to an event S_∞ with $\mathbf{P}(S - S_\infty) = 0$. Finally, γ_N is computed from

$$\phi_N(\theta) = \mathbf{E}\{\sum_{r \leqq N} \exp(-\theta B_{1r})\},$$

from which it follows that $\gamma_N \to \gamma$. This suffices to complete the proof.

Acknowledgments. It is a pleasure to acknowledge the helpful comments of Mr. J. D. Biggins, Dr. D. J. Daley, Dr. J. M. Hammersley and a referee on the drafts of this paper. Dr. Daley points out that γ is a monotone function of ϕ in the sense that, if two processes have functions ϕ_1 and ϕ_2 with $\phi_1(\theta) \leqq \phi_2(\theta)$ for all θ, then $\gamma_1 \geqq \gamma_2$. Mr. Biggins has extended the arguments developed here to cover processes with a countable number of types, and has also attacked the corresponding 'last birth' problem.

REFERENCES

[1] ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer, Berlin.

[2] BAHADUR, R. R. and RANGA RAO, R. (1960). On deviations of the sample mean. *Ann. Math. Statist.* **31** 1015–1027.

[3] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[4] CRUMP, K. S. and MODE, C. J. (1968). A general age-dependent branching process I. *J. Math. Anal. Appl.* **24** 494–508.

[5] DANTZIG, D. van (1949). Sur la méthode des fonctions génératrices. *Colloques Internationaux du Centre National de la Récherche Scientifique* **13** 29–45.

[6] DONEY, R. A. (1974). On single and multi-type general age-dependent branching processes. *Manchester-Sheffield Reseach Report.*

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

[8] HAMMERSLEY, J. M. (1973). Contribution to the discussion of [14].

[9] HAMMERSLEY, J. M. (1974). Postulates for subadditive processes. *Ann. Probability* **2** 652–680.

[10] HAMMERSLEY, J. M. and WELSH, J. A. D. (1965). First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory, in *Bernoulli-Bayes-Laplace Anniversary Volume*. Springer, Berlin.

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

[12] KESTEN, H. (1973). Contribution to the discussion of [14].

[13] KINGMAN, J. F. C. (1968). The ergodic theory of subadditive stochastic processes. *J. Roy. Statist. Soc. Ser. B* **30** 499–510.

[14] KINGMAN, J. F. C. (1973). Subadditive ergodic theory. *Ann. Probability* **1** 883–909.

MATHEMATICAL INSTITUTE
24–29 ST. GILES
OXFORD OXI 3LB
ENGLAND