BRANCHING PROCESSES AND THE THEORY OF EPIDEMICS

ROBERT BARTOSZYŃSKI
MATHEMATICAL INSTITUTE, POLISH ACADEMY OF SCIENCES

1. Introduction

In the present paper we shall discuss the extinction problem for certain branching processes. Our main purpose is to study those branching processes which can serve as models of epidemics; that is, spreads of an infectious disease. The phenomenon of epidemics is fairly complex, and all models necessarily have to be based on a certain compromise. This compromise consists of taking into account some of the (presumably important) factors governing the spread of the disease at the cost of neglecting others. Our main idealization will consist of assuming such mechanisms of infection which yield a branching process. Using informal language, it means that all infectives present in the population at a given time infect the susceptibles independently of each other. More precisely, if there are k infectives in the nth generation of the epidemics, then the distribution of the next $(n + 1)$st generation can be represented as the distribution of a sum of k independent, identically distributed random variables, with a specified distribution. These random variables represent the "progeny" of k infectives of nth generation.

It is debatable whether the above assumption is justified in the sense that models based on it provide an adequate description of reality. The main objection is that when the population gets "saturated" with infectives, the assumption of independence becomes violated. We shall not attempt to defend our models; we only want to show that it is possible to fit within the framework of branching processes some of the factors governing the spread of the disease; such as, effects of vaccination and other preventive methods, variable lengths of the period of incubation and infectiousness, effects due to random movements of infectives within the habitat, and also, to some extent, the effects due to the exhaustion of susceptibles. We do obtain meaningful results within our models; whether or not these results bear any relation to reality lies beyond our present interest.

In section 2 we consider a simple case of a Galton-Watson process, and we show that even in this relatively simple case one can take into consideration the variable length of the period of incubation and infectiousness. In section 3 we generalize the model of section 2 to include the random movements of individuals.

Finally, in section 4 we deal with generalized branching processes, in an

attempt to approximate the epidemics on finite populations. The definition of a branching process as presented there differs considerably from the standard one: we include the possibility that the distribution of the number of descendants may vary in a random manner.

2. A model of epidemics based on a Galton-Watson process

Since the present section is mathematically trivial, we may dispose for a while with formalities and use the intuitive terminology. We make the following assumptions about the mechanisms of infection and the duration of the disease.

ASSUMPTION 1. *Every individual who becomes infected passes first through an incubation period of the length X followed by the period of infectiousness of the length Y. During the period of incubation he is harmless to others, whereas during the period of infectiousness he may infect those with whom he comes in contact. We shall measure time in appropriate discrete units (calling them days, for simplicity), and we assume that X and Y are random variables with the joint probability distribution $p_{m,n} = \Pr\{X = m, Y = n\}$.*

Let $F(s, t) = \sum_{m,n} p_{m,n} s^m t^n$ be the probability generating function of the pair (X, Y).

ASSUMPTION 2. *At each day during the period of incubation every individual has a probability $1 - \alpha$, $(0 < \alpha \leqq 1)$ of being discovered and isolated. During the period of infectiousness the probability of being discovered and isolated equals $1 - \beta$, $(0 < \beta \leqq 1)$ at each day. These probabilities do not depend on the number of days during which the individual in question remains undiscovered.*

ASSUMPTION 3. *Every infected individual who has not been previously discovered makes a certain number of contacts with noninfected members of the population during each day of his period of infectiousness. The number of contacts made on different days are assumed to be independent and identically distributed with the distribution $\{r_k\}$, $k = 0, 1, 2, \cdots$. Let $R(t) = \sum_k r_k t^k$ be the probability generating function of the number of daily contacts.*

ASSUMPTION 4. *We assume that each contact with an infectious individual yields an infection with probability γ, $(0 < \gamma \leqq 1)$ independently of the results of other contacts.*

ASSUMPTION 5. *The events occurring to an individual on a given day are independent of the events that occurred to him or other members of the population on previous days, and they are independent of the events that occur to other members of the population on the same day.*

ASSUMPTION 6. *The expected number of daily contacts is finite.*

Comment. Obviously, the assumptions 1–4 are partly superfluous: one could incorporate 2 into 1 by considering at once the number of days of "effective infectiousness." One could also incorporate 4 into 3 by considering only "successful" contacts. We hope, however, that by separating the assumptions we may be able to trace the influences of various factors which could, perhaps, have biological significance. Thus, for instance, γ may represent the degree of "con-

tagiousness" of the disease, or the resistance of the population due to vaccination, climate, and so on. Similarly, α may represent the efficiency of health service in its attempts to trace the individuals who may have had contact with the disease in order to isolate them, while β may measure the efficiency of periodic checkups and health control.

We shall study conditions under which a single individual who becomes infected has a positive probability of originating epidemics. The word "epidemics" will be somewhat arbitrarily understood to mean that all generations of a Galton-Watson process originated by this individual are nonempty, and consequently, their sizes tend to infinity with probability one (see, for instance, Harris [1]).

It is well known that all information about the behavior of a Galton-Watson branching process is contained in the probability generating function of the number of "direct descendants" of a single element. To compute this generating function in our case, let us suppose that we have a single individual who became infected at a certain moment, and let w_n be the probability that he will be "effectively infectious" (infectious and not isolated) during exactly n days ($n \geqq 1$). He must, therefore, remain undiscovered during all his period of incubation and during the first n days of his infectiousness period. If $X = m$, $Y = n$, this probability is $\alpha^m\beta^n$; in case $X = m$, $Y = k$ ($k > n$), this probability equals $\alpha^m\beta^n(1 - \beta)$. Denoting for simplicity

$$q_{m,n} = p_{m,n+1} + p_{m,n+2} + \cdots = \Pr\{X = m,\ Y > n\}, \tag{1}$$

we get

$$w_n = \beta^n \sum_m p_{m,n}\alpha^m + (1 - \beta)\beta^n \sum_m q_{m,n}\alpha^m. \tag{2}$$

The last formula holds for all $n \geqq 1$; for $n = 0$ we must add a term which accounts for the possibility of being discovered during the period of incubation. Adding all w_n for $n \geqq 1$ one could easily obtain $\sum_n w_n = F(\alpha, 1)$.

Now let v_k be the probability of exactly k contacts leading to the disease during one day of the period of infectiousness. We find easily

$$v_k = \sum_{j \geqq k} \binom{j}{k} \gamma^k(1 - \gamma)^{j-k} r_j. \tag{3}$$

For the generating function of the sequence $\{v_k\}$, we obtain

$$\begin{aligned} \sum_k v_k t^k &= \sum_k \sum_{j \geqq k} \binom{j}{k} \gamma^k(1 - \gamma)^{j-k} r_j t^k \\ &= \sum_k \frac{(\gamma t)^k}{k!} \sum_{j \geqq k} j(j - 1) \cdots (j - k + 1)(1 - \gamma)^{j-k} r_j \\ &= \sum_k \frac{(\gamma t)^k}{k!} R^{(k)}(1 - \gamma) = R(\gamma t + 1 - \gamma); \end{aligned} \tag{4}$$

the inner series in the antepenultimate expression is the kth derivative of the function R at the point $1 - \gamma$, and the penultimate expression is the expansion

of $R(\gamma t + 1 - \gamma)$ into the Taylor series in the neighborhood of the point $1 - \gamma$.

By assumption of independence, the generating function of the joint number of infected during n days of effective infectiousness is $R(\gamma t + 1 - \gamma)^n$. Thus, for the generating function of the number of direct descendants of an infected individual we get

$$
\begin{aligned}
(5) \qquad G(t) &= 1 - F(\alpha, 1) + \sum_n w_n[R(\gamma t + 1 - \gamma)]^n \\
&= 1 - F(\alpha, 1) + \sum_{n,m} \beta^n \alpha^m p_{m,n}[R(\gamma t + 1 - \gamma)]^n \\
&\qquad + (1 - \beta) \sum_{n,m} \beta^n \alpha^m q_{m,n}[R(\gamma t + 1 - \gamma)]^n \\
&= 1 - F(\alpha, 1) + F(\alpha, \beta R(\gamma t + 1 - \gamma)) \\
&\qquad + (1 - \beta) Q(\alpha, \beta R(\gamma t + 1 - \gamma)),
\end{aligned}
$$

where $Q(s, t) = \sum_{m,n} q_{m,n} s^m t^n$.

By simple generalization of the well-known formula that gives the relation between the generating functions of a sequence and the sequence of its "tails," we get

$$
(6) \qquad Q(s, t) = \frac{F(s, 1) - F(s, t)}{1 - t}.
$$

In fact, let $p_{m\cdot} = \sum_n p_{m,n}$. Then $\sum_m p_{m\cdot} s^m = F(s, 1)$ and

$$
\begin{aligned}
(7) \qquad Q(s, t) &= \sum_m \sum_n q_{m,n} s^m t^n = \sum_m \sum_n \sum_{j>n} p_{m,j} s^m t^n \\
&= \sum_m s^m \sum_n t^n \left(p_{m\cdot} - \sum_{j=0}^{n} p_{m,j} \right) \\
&= \sum_n t^n \sum_m p_{m\cdot} s^m - \sum_m \sum_n \sum_{j=0}^{n} p_{m,j} s^m t^n \\
&= (1 - t)^{-1} F(s, 1) - \sum_m s^m \sum_j p_{m,j} \sum_{n \geq j} t^n \\
&= (1 - t)^{-1} F(s, 1) - (1 - t)^{-1} \sum_m \sum_j p_{m,j} s^m t^j \\
&= (1 - t)^{-1} [F(s, 1) - F(s, t)].
\end{aligned}
$$

Combining (6) with (5), we finally obtain

$$
\begin{aligned}
(8) \qquad G(t) &= 1 - F(\alpha, 1) \\
&\quad + \frac{(1 - \beta) F(\alpha, 1) + \beta(1 - R(\gamma t + 1 - \gamma)) F(\alpha, \beta R(\gamma t + 1 - \gamma))}{1 - \beta R(\gamma t + 1 - \gamma)}.
\end{aligned}
$$

Differentiating, we obtain

$$
(9) \qquad G'(1) = \frac{\beta}{1 - \beta} \gamma R'(1) [F(\alpha, 1) - F(\alpha, \beta)],
$$

provided that $\beta \neq 1$. For $\beta \to 1$ we find

$$
(10) \qquad G'(1) = \gamma R'(1) \frac{\partial}{\partial t} F(s, t)|_{s=\alpha, t=1}.
$$

The last value may be infinite.

We shall now discuss the condition under which $G'(1) > 1$. We shall use the theorem which asserts that for a Galton-Watson process with the generating function $f(t)$, the probability of extinction is smaller than one if and only if $f'(1) > 1$ and equals to the smallest positive root of the equation $x = f(x)$. Denoting for simplicity $R'(1) = r$ (r being the average number of contacts per day), we see easily that $G'(1) > 1$ if and only if

$$F(\alpha, \beta) < F(\alpha, 1) - \frac{1}{\gamma r} \cdot \frac{1 - \beta}{\beta}. \tag{11}$$

Consider now, for a fixed α, two functions $h_1(\beta) = F(\alpha, \beta)$ and $h_2(\beta) = F(\alpha, 1) - (1 - \beta)\gamma r\beta$. We see that $h_1(1) = h_2(1)$; both h_1' and h_2' are positive, while h_1'' and h_2'' are of different signs. Thus h_1 and h_2 are increasing; one of them is convex and the other concave. Two possible modes of behavior of h_1 and h_2 are presented in figures 1 and 2.

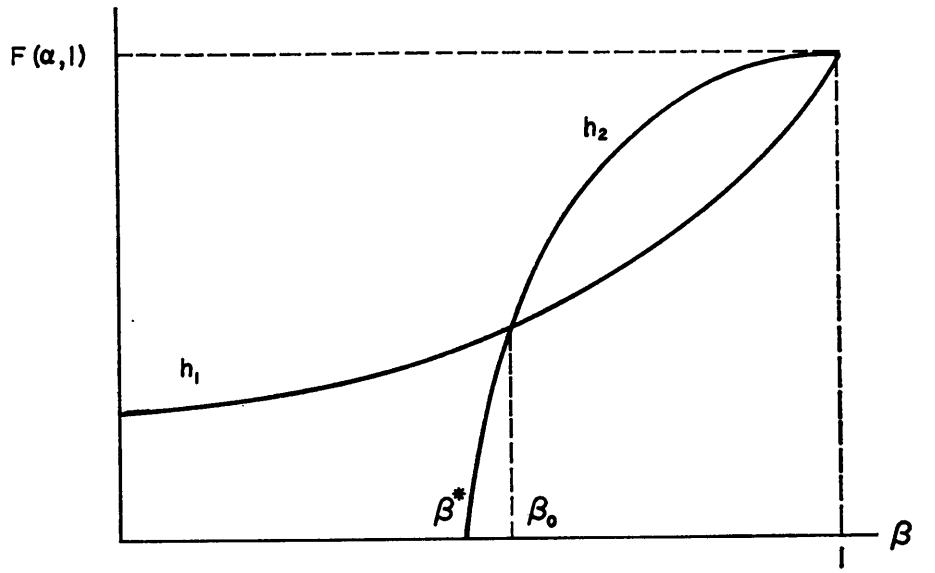

FIGURE 1

Possible behavior of h_1 and h_2.

The inequality $G'(1) > 1$ or, equivalently, inequality (11), holds if and only if the situation is such as presented in figure 1, and only for values $\beta_0 < \beta < 1$, where β_0 is the (unique) smaller-than-one root of the equation $h_1(\beta) = h_2(\beta)$. Now, in figure 1 we have $h'(1) > h_2'(1)$, whereas in figure 2 we have $h_1'(1) \leqq h_2'(1)$. Further, $h_1'(1) = \partial F(s, t)/\partial t|_{s=\alpha, t=1} = D(\alpha)$, where $D(\alpha)$ is the average length of the period of infectiousness of individuals who remained undiscovered during their incubation period, while $h_2'(1) = 1/\gamma r$. We are now in the position to state our criterion.

A single individual has a positive probability of originating epidemics if and only if $D(\alpha) > 1/\gamma r$ *and* $\beta_0 < \beta < 1$, *where* β_0 *is the smallest root of the equation*

$$\gamma r x F(\alpha, x) - \gamma r x F(\alpha, 1) + 1 - x = 0. \tag{12}$$

If the above conditions are satisfied, then the probability of epidemics equals $1 - z$, *where* z *is the smallest positive root of the equation* $x = G(x)$ *with* $G(x)$ *given by* (8).

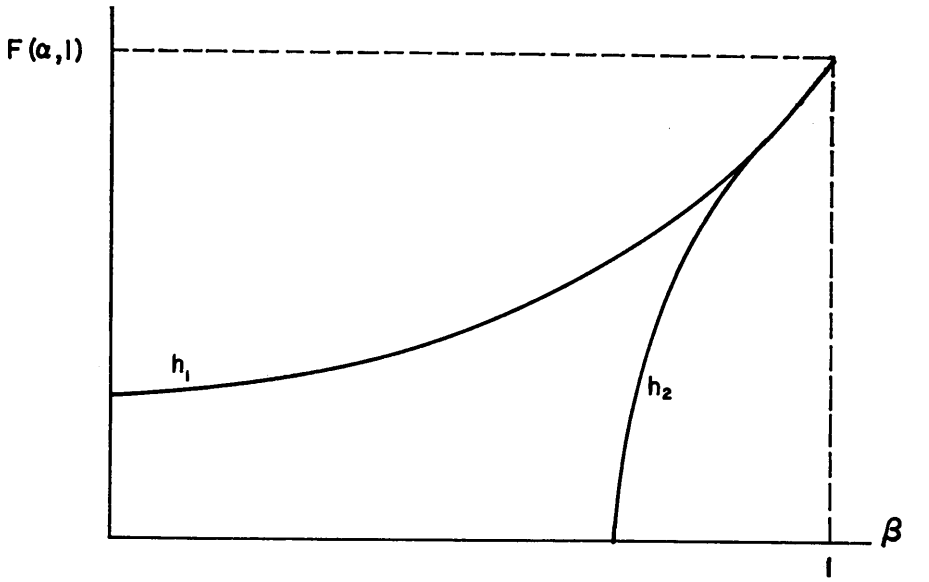

FIGURE 2

Other possible behavior of h_1 and h_2.

In addition, notice that we obtain a reasonable estimate of β_0 by considering the point β^* where h_2 intersects the β-axis. We easily obtain

$$\beta_0 > \beta^* = (1 + \gamma r F(\alpha, 1))^{-1}. \tag{13}$$

We may reformulate our result, stressing its possible applications to reality as follows: there are five factors ($F(s, t)$, r, α, β, and γ) involved in our model. The first two (characteristic $F(s, t)$ of the disease and the average number of contacts per day r) are more or less beyond our control, and they could possibly be determined in practical situations. The remaining three constants measure the efficiency of preventive methods: α in discovering the infectives during their incubation period, β in discovering the infectious individuals, and γ in increasing the individual resistance to the disease. The question arises: given F and r, which values of α, β, and γ yield a "safe" domain (that is, such a domain of values, for which the epidemics do not occur with probability one)? One can see that in order to prevent epidemics one must attempt to have either $\gamma D(\alpha) \leqq 1/r$ for arbitrary β, or, if for some reason the above inequality is unattainable, one should attempt to have $\beta \leqq (1 + \gamma r F(\alpha, 1))^{-1}$ $(< \beta_0)$.

The case in which the incubation period is of the constant length M gives $D(\alpha) = D\alpha^M$, and $F(\alpha, 1) = \alpha^M$, D being the average length of the period of infectiousness. Then the "safe" domain would be at least either $\gamma\alpha^M \leqq 1/rD$, β arbitrary, or $\beta \leqq (1 + \gamma r\alpha^M)^{-1}$.

Finally, let us remark that even if the average length of the period of infectiousness D is infinite, we may still prevent the epidemics by a suitable selection of β. This problem may not only be of academic interest: in some diseases (for example, venereal) the average length may be of the order of the average length of the human life; that is, it may be considered "practically infinite" as compared with the average length of the infectiousness period. By lowering β (by increasing the efficiency of periodic checkups), we may be able to prevent the epidemics for such diseases.

3. Epidemics with random movements of individuals

The model of the preceding section can be generalized easily so as to include the possibility of random movements of individuals. Thus, we imagine that the whole habitat is partitioned into N subhabitats (which we shall call "zones"), and we assume that all individuals, independently of one another, may move from one zone to another.

We assume that the disease is characterized by the joint distribution of the vector (X, Y) as in the preceding section. The corresponding probabilities of getting discovered and isolated during the periods of infectiousness and incubation will be $1 - \alpha_i$ and $1 - \beta_i$, respectively, for individuals in the ith zone. The distribution of the number of daily contacts for the ith zone will have the generating function $R_i(t)$, and the corresponding probability of the contact being "successful" will be γ_i. We allow one transition from one zone to another per day, and we assume that all individuals travel and infect independently of each other, and independently of their past histories. The transition probability matrix will be $[q_{i,j}]$ for the transitions during the period of incubation and $[p_{i,j}]$ during the period of infectiousness.

Let $G_k(t_1, \cdots, t_N)$ be the probability generating function of the vector $(X_1^{(k)}, \cdots, X_N^{(k)})$, where $X_i^{(k)}$ is the total number of individuals infected in the ith zone by an individual who himself became infected in the kth zone. Let $G(t_1, \cdots, t_N)$ be the vector whose kth component is $G_k(t_1, \cdots, t_N)$. Continuing in the same way as in the preceding section we easily arrive at the formula

$$G(t_1, \cdots, t_N) = \Psi(F(U, V) + Q(U, V)W + C), \tag{14}$$

where

$$U = [q_{i,j}\alpha_j], \qquad W = [p_{i,j}(1 - \beta_j)], \qquad V = [p_{i,j}\beta_j R_j(\gamma_j t_j + 1 - \gamma_j)], \tag{15}$$

C is a constant matrix accounting for the possibility of leaving no descendants due to the discovery prior to the end of the incubation period (the exact form of C is of minor interest, since we deal mostly with derivatives of vector G), and, finally, Ψ is a function which assigns to a matrix the vector whose components

are equal to the sums of corresponding rows of the matrix. Since all powers of matrices U and V have entries bounded by one, no question of the existence of the generating function with matrix-valued variables arises.

We can now use the generalization of the extinction theorem for branching processes which asserts that the branching process with generating functions $G_k(t_1, \cdots, t_N)$ $(k = 1, \cdots, N)$ is bound to expire with probability one if and only if no characteristic roots of the matrix

$$\left[\frac{\partial G_k(t_1, \cdots, t_N)}{\partial t_j}\bigg|_{t_1 = \cdots = t_N = 1}\right] \tag{16}$$

lie outside the unit circle.

4. Generalized branching processes

We shall now proceed in a more formal way and give the definition of generalized branching processes. The idea which we have in mind is to define such branching processes that can serve as approximations of epidemics on finite populations.

Let S be a finite or countable set, and suppose that to every $s \in S$ there corresponds a probability distribution π_s on the set of all nonnegative integers. Thus, every π_s is a sequence of nonnegative numbers $\{\pi(n|s)\}$, $(n = 0, 1, 2, \cdots)$ such that $\sum_n \pi(n|s) = 1$. Often, instead of speaking about the distributions π_s, we shall speak about their probability generating functions $f(x|s)$ defined as $f(x|s) = \sum_n x^n \pi(n|s)$, $|x| \leqq 1$.

Let $\mathfrak{L}$ be the class of all finite sequences $(k_0; s_1, k_1; \cdots; s_n, k_n)$, where $s_j \in S$ $(j = 1, 2, \cdots, n)$, $k_i (i = 0, 1, \cdots, n)$ are nonnegative integers, and $n = 0, 1, 2, \cdots$. Let $D(S)$ be the class of all probability distribution over the set S.

DEFINITION. *A sequence $z_0, z_1, z_2, \cdots$, of nonnegative integer-valued random variables will be called a (generalized) branching process, if there exists a function $\varphi: \mathfrak{L} \to D(S)$ such that the distribution of every finite subsequence $(z_0, z_1, \cdots, z_n)$ coincides with the marginal distribution of $(z_0^*, z_1^*, \cdots, z_n^*)$ of the process $z_0^*, \xi_1, z_1^*, \xi_2, z_2^*, \cdots$ defined as follows:*

(i) *z_0^* has the same distribution as z_0;*

(ii) *for any $n \geqq 1$ the conditional distribution of ξ_{n+1} given that $z_0^* = k_0$, $\xi_1 = s_1$, $z_1^* = k_1, \cdots, \xi_n = k_n$, $z_n^* = k_n$, is equal to $\varphi(k_0; s_1, k_1; \cdots; s_n, k_n) \in D(S)$ (thus $\xi_{n+1} \in S$, $n = 1, 2, \cdots$);*

(iii) *for any $n \geqq 1$, the conditional distribution of z_{n+1}^*, given that $z_0^* = k_0$, $\xi_1 = s_1$, $z_1^* = k_1, \cdots, z_n^* = k_n$, $\xi_{n+1} = s_{n+1}$ coincides with the distribution of the sum of k_n independent random variables, each having the distribution $\pi_{s_{n+1}}$; that is, it has the probability generating function equal to $[f(x|s_{n+1})]^{k_n}$.*

The idea of this definition is to comprise the processes, which, roughly speaking, are constructed in such a way, that given the history of the process up to the nth generation, the $(n + 1)$st generation is a sum of independently created

"offsprings" of the members of the previous generation, while the "fertility" distribution changes in a random fashion, depending on the past history of the process.

We shall identify processes $\{z_n\}$ with the "marginal" processes $\{z_n^*\}$ which generate them. If necessary, we shall refer to the auxiliary process $\{\xi_n\}$ corresponding to the process $\{z_n\}$.

Let us note that branching processes $\{z_n\}$ as defined here are, in general, non-Markovian. To find conditions under which $\{z_n\}$ is a Markov process, let us note first that one can interpret the function φ, which appears in the definition of the process as a function which maps the class of all cylinders with finite bases corresponding to the process $z_0, \xi_1, z_1, \xi_2, \cdots$ into $D(S)$. More precisely, let Ω be the class of all sequences $\omega = (k_0, s_1, k_1, s_2, \cdots)$ with $s_j \in S$ and k_i being nonnegative integers ($j = 1, 2, \cdots, i = 0, 1, 2, \cdots$). Let $\mathfrak{B}$ be the class of cylinders with finite bases in Ω, and let $\mathfrak{F}$ be the smallest σ-field containing all sets in $\mathfrak{B}$. Further, let P be the measure on $(\Omega, \mathfrak{F})$ generated by conditions (i)–(iii). The function φ maps the class $\mathfrak{B}^* \subset \mathfrak{B}$ of all cylinders with bases of the form $(k_0, s_1, \cdots, s_n, k_n)$, $n = 0, 1, 2, \cdots$ into $D(S)$.

We can now formulate the following theorem.

THEOREM 1. *The branching process $\{z_n\}$ is a Markov process if and only if the function φ has the property*

$$\varphi(k_0, s_1, k_1, \cdots, s_n, k_n) = \Phi(n, k_n) \tag{17}$$

on all cylinders in $\mathfrak{B}^*$ with positive P measure.

In fact, if (17) is not satisfied, then there exist for some n two different "histories," $H = (k_0, s_1, k_1, \cdots, k_n)$ and $H' = (k_0', s_1', k_1', \cdots, k_n')$, which may occur with positive probability and such that $k_n = k_n'$, while $\varphi(H) \neq \varphi(H')$. In this case the conditional probability distribution of z_{n+1} given H is different from the conditional distribution of z_{n+1} given H', which shows that $\{z_n\}$ is not a Markov process. Suppose now that (17) holds, and let $H = (k_0, s_1, k_1, \cdots, k_n)$ represent a history of the process up to nth generation. Given that $\xi_{n+1} = s$, the distribution of z_{n+1} has the generating function $[f(x|s)]^{k_n}$. If we denote the distribution appearing on the right-hand side of (17) by $\alpha_{n,k_n}(s)$, we can write the generating function of z_{n+1} given H as

$$\sum_{s \in S} \alpha_{n,k_n}(s)[f(x|s)]^{k_n}. \tag{18}$$

Since the last expression depends only on k_n, we conclude that $\{z_n\}$ is a Markov process.

THEOREM 2. *If the branching process $\{z_n\}$ has the property that $q = \inf_{s \in S} \pi(0|s) > 0$, then*

$$\Pr\{\lim_{n\to\infty} z_n = 0 \quad \text{or} \quad \lim_{n\to\infty} z_n = \infty\} = 1. \tag{19}$$

PROOF. Let us note that if $z_N = 0$ for some N, then $z_n = 0$ for all $n > N$. For a fixed $k > 0$ denote

$$C_N(k) = \bigcup_{n \geqq N} \{z_n = k\}, \qquad C_N^M(k) = \bigcup_{n=N}^{M} \{z_n = k\}. \tag{20}$$

To prove the assertion, it suffices to show that for every $k > 0$ we have $\Pr\{\bigcap_N C_N(k)\} = 0$. Since $C_N(k) \supset C_{N+1}(k)$, $N = 1, 2, \cdots$, it is enough to prove that $\lim_{n\to\infty} \Pr\{C_N(k)\} = 0$. We may write

$$C_N^M(k) = \{z_M = k\} \cup \bigcup_{n=N}^{M-1} \{z_n = k, z_{n+1} \neq k, \cdots, z_M \neq k\}. \tag{21}$$

The events on the right-hand side of the last formula are disjoint, hence

$$\begin{aligned} &\Pr\{C_N^M(k)\} \\ &= \Pr\{z_M = k\} + \sum_{n=N}^{M-1} \Pr\{z_n = k, z_{n+1} \neq k, \cdots, z_M \neq k\} \\ &= \Pr\{z_M = k\} + \sum_{n=N}^{M-1} \Pr\{z_M \neq k, \cdots, z_{n+1} \neq k | z_n - k\} \Pr\{z_n = k\} \\ &= \Pr\{z_M = k\} + \sum_{n=N}^{M-1} (1 - \Pr\{C_{n+1}^M(k) | z_n = k\}) \Pr\{z_n = k\}. \end{aligned} \tag{22}$$

If $\sum_n \Pr\{z_n = k\} < \infty$ for $k = 1, 2, 3, \cdots$, then we may pass to the limit with $M \to \infty$ obtaining

$$\lim_{M\to\infty} \Pr\{C_N^M(k)\} = \Pr\{C_N(k)\} \leqq \sum_{n=N}^{\infty} \Pr\{z_n = k\}, \tag{23}$$

which shows that $\lim_{N\to\infty} \Pr\{C_N(k)\} = 0$. Now, if we should have

$$\sum_{n=1}^{\infty} \Pr\{z_n = k\} = \infty, \tag{24}$$

then passing to the limit with $M \to \infty$ we would obtain from (22)

$$\limsup_{n\to\infty} \Pr\{C_{n+1}(k) | z_n = k\} = 1. \tag{25}$$

But $\overline{C}_{n+1}(k) \supset \{z_{n+1} = 0\}$; therefore,

$$\Pr\{\overline{C}_{n+1}(k) | z_n = k\} \geqq \Pr\{z_{n+1} = 0 | z_n = k\} \geqq q^k, \tag{26}$$

which gives the contradiction.

Using almost the same argument, one can prove a slightly more general result. For $q > 0$, let $S_q = \{s : \pi(0|s) \geqq q\}$, and let $\mathfrak{L}_n$ be the subclass of $\mathfrak{L}$ consisting of all sequences $(k_0, s_1, \cdots, k_n)$ of length $2n + 1$. Denote

$$d_{n,q} = \Pr\{s_{n+1} \in S_q | k_0, s_1, \cdots, k_n\}, \tag{27}$$

and

$$d_{n,q}^* = \inf_{\mathfrak{L}_n} d_{n,q}. \tag{28}$$

Then the following holds.

If for some $q_0 > 0$ we have

$$\limsup_{n\to\infty} d_{n,q_0}^* > 0, \tag{29}$$

then

$$\Pr\{\lim_{n\to\infty} z_n = 0 \quad \text{or} \quad \lim_{n\to\infty} z_n = \infty\} = 1. \tag{30}$$

We shall now prove a theorem which gives the sufficient conditions for the extinction.

Let $\mathcal{K}_r$ be the class of histories $(k_0, s_1, k_1, \cdots, k_n)$ of arbitrary length n such that $k_n = r$. Let $\mu_H^{(r)}$ denote the expected number of descendants of the elements of the last generation given that $H \in \mathcal{K}_r$ (thus, if $H \in \mathcal{K}_r$ is a history of the length n, then $E(z_{n+1}|H) = r\mu_H^{(r)}$). We shall prove (the idea of this proof is due to K. Urbanik) the following theorem.

THEOREM 3. *If* $\inf_{s\in S} \pi(0|s) > 0$, *and if for all sufficiently large* r *we have*

$$\sup_{H\in\mathcal{K}_r} \mu_H^{(r)} \leqq 1, \tag{31}$$

then

$$\Pr\{\lim_{n\to\infty} z_n = 0\} = 1. \tag{32}$$

PROOF. From the preceding theorem it follows that with probability one we have either $\lim_{n\to\infty} z_n = 0$ or $\lim_{n\to\infty} z_n = \infty$. To exclude the last possibility, it suffices to show that $\limsup_{n\to\infty} E(z_{n+1}) < \infty$. From the proof of the preceding theorem it follows that for every $k > 0$ we have $\sum_n \Pr\{z_n = k\} < \infty$. Now,

$$\begin{aligned} E(z_{n+1}) &= \sum_k k \Pr\{z_{n+1} = k\} \\ &= \sum_k \sum_r k \Pr\{z_{n+1} = k|z_n = r\} \Pr\{z_n = r\} \\ &= \sum_r [\sum_k k \Pr\{z_{n+1} = k|z_n = r\}] \Pr\{z_n = r\} \\ &\leqq \sum_r r(\sup_{H\in\mathcal{K}_r} \mu_H^{(r)}) \Pr\{z_n = r\} \\ &\leqq \sum_r r \Pr\{z_n = r\} + \sum_{r\in R^*} r(\sup_{H\in\mathcal{K}_r} \mu_H^{(r)} - 1) \Pr\{z_n = r\}, \end{aligned} \tag{33}$$

where R^* is the set of indices r for which $\sup_{H\in\mathcal{K}_r} \mu_H^{(r)} > 1$. Since R^* is finite, we get $E(z_{n+1}) \leqq E(z_n) + a_n$, where the series $\sum_n a_n$ converge, and consequently, $\limsup_{n\to\infty} E(z_n) < \infty$ which completes the proof.

REFERENCE

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