J. Appl. Prob. **20**, 11–18 (1983)
Printed in Israel

A NOTE ON MODELS USING THE BRANCHING PROCESS WITH IMMIGRATION STOPPED AT ZERO

E. SENETA,* *University of Sydney*
S. TAVARÉ,** *Colorado State University*

Abstract

The Galton–Watson process with immigration which is time-homogeneous but not permitted when the process is in state 0 (so that this state is absorbing) is briefly studied in the subcritical and supercritical cases. Results analogous to those for the ordinary Galton–Watson process are found to hold. Partly-new techniques are required, although known end-results on the standard process with and without immigration are used also. In the subcritical case a new parameter is found to be relevant, replacing to some extent the criticality parameter.

GALTON–WATSON PROCESS; IMMIGRATION; SUBCRITICAL; SUPERCRITICAL; RENEWAL EQUATION; EXTINCTION-TIME DISTRIBUTION; PLASMIDS

1. Introduction

We treat briefly the simple branching (Galton–Watson) process with time-homogeneous immigration (BPI), with the difference that no immigration is permitted when the process is in state 0, so that this state is absorbing. This variant of the usual branching process arises as a biological model; and our sketch of the theory is motivated by the need to consider the relevant mathematical structure separately from the biological problems. We do not aim, at least in this note, at a more or less complete presentation of the theory, nor at the levels of utmost mathematical generality to which such simple processes have been studied in recent years. It will be seen that while use is made of results for the ordinary Galton–Watson process (without and with immigration), the techniques of analysis are not entirely standard. The results obtained are, however, analogous to those for the ordinary absorbing process, and the

Received 11 March 1982; revision received 21 May 1982.
* Postal address: Department of Mathematical Statistics, University of Sydney, NSW 2006, Australia.
Research carried out while this author was visiting Colorado State University.
** Postal address: Department of Statistics, Colorado State University, Fort Collins CO 80523, U.S.A.

modified process provides information about the first-passage time to 0 in the ordinary Galton–Watson process with immigration.

Pakes [6] has treated the situation dual to ours, where immigration is permitted *only* when the process is in state 0, and extensions of that situation have been made; these are different in essence.

We denote our process by $\{Y_t\}$, $t \geqq 0$, where Y_0 is taken to have some initial distribution $\{\pi_j\}$, $j \geqq 0$, with p.g.f. $G_0(s) = \sum_{j=0}^{\infty} \pi_j s^j$, $s \in [0,1]$; and write the offspring and immigration p.g.f.'s as $F(s) = \sum_{j=0}^{\infty} f_j s^j$, $B(s) = \sum_{j=0}^{\infty} b_j s^j$ respectively. Our basic assumptions, which we assume to obtain throughout, are $0 < B(0) < 1$, $0 < F(0) < 1$. Then for $s \in [0,1]$

$$G_{t+1}(s) = E(s^{Y_{t+1}}) = E(E(s^{Y_{t+1}} \mid Y_t))$$

i.e.

$$G_{t+1}(s) = B(s)G_t(F(s)) + G_t(0)(1 - B(s)), \qquad t \geqq 0 \tag{1}$$

which is our fundamental equation.

One model [3], [5] for the number of plasmids in a randomly chosen daughter cell (out of two daughter cells) supposes that if there were $i \geqq 1$ plasmids in the parent cell, then (a fixed number) n of replication events take place, on each of which there is, independently, a chance p of a new plasmid being formed. If at the end of this process there are k ($\geqq i$) plasmids (including the original i) then each daughter cell has probability $\frac{1}{2}$ of receiving any one of these, independently for each plasmid. If the parent cell has 0 plasmids, it is said to be cured, and each daughter cell receives 0 plasmids. Thus if Y_{t+1} is the number of plasmids in a daughter cell, and Y_t the number in a parent cell, it follows that when $Y_t > 0$

$$Y_{t+1} = U_1 + U_2 + \cdots + U_{N_t} \tag{2}$$

where $N_t = Y_t + Z_t$, with $Z_t \sim \text{Bin}(n, p)$, and $U_i \sim \text{Bin}(1, \frac{1}{2})$. Under the natural independence assumptions implying the Markov property it is readily checked that the model we have proposed holds and in particular (1) obtains with

$$F(s) = \tfrac{1}{2} + \tfrac{1}{2}s \quad \text{and} \quad B(s) = \left(1 - \frac{p}{2} + \frac{p}{2}s\right)^n.$$

Equation (2) may be written in the more conventional form: if $Y_t > 0$

$$Y_{t+1} = U_1 + U_2 + \cdots + U_{Y_t} + I \tag{3}$$

where I is a random variable with p.g.f. $B(s)$. In our special case $m \equiv F'(1-) = \frac{1}{2}$, so for purposes of application we need consider only a subcritical process for the general model, and this is our basic presentation. We shall additionally, and only very briefly give some basic details on the case $m > 1$, since this can be done without extensive analysis.

Iterating (1) we obtain

$$G_{t+1}(s) = P_{t+1}(s)G_0(F_{t+1}(s)) + \sum_{r=0}^{t} G_{t-r}(0)(P_r(s) - P_{r+1}(s))$$

where

$$P_r(s) = 1, \quad \text{if } r = 0; \qquad = \prod_{l=0}^{r-1} B(F_l(s)), \quad \text{if } r \geqq 1.$$

Here $F_l(s)$ is the lth functional iterate of F ($F_0(s) = s$), and $P_r(s)$ is recognizable as the p.g.f. of generation size at time r of an *ordinary* BPI $\{X_t\}$ with offspring and immigration p.g.f.'s $F(s)$ and $B(s)$, starting with 0 individuals ($X_0 = 0$). Thus for $s \in [0,1]$

$$G_t(s) = P_t(s)G_0(F_t(s)) + \sum_{r=1}^{t} G_{t-r}(0)(P_{r-1}(s) - P_r(s)) \tag{4}$$

and consequently

$$1 - G_t(s) = P_t(s)(1 - G_0(F_t(s))) + \sum_{r=1}^{t} (1 - G_{t-r}(0))(P_{r-1}(s) - P_r(s)). \tag{5}$$

Notice that $G_t(0) = P(T \leqq t)$ where T is the time to extinction for the process $\{Y_t\}$, or the first-passage time to 0 for a corresponding ordinary BPI. Putting $s = 0$ in (5) and writing $a_r = P_{r-1}(0) - P_r(0)$, $r \geqq 1$, $u_t = 1 - G_t(0)$, we obtain

$$u_t = c_t + \sum_{r=1}^{t} u_{t-r}a_r \tag{6}$$

where $c_t = P_t(0)(1 - G_0(F_t(0)))$. (6) has the form of a renewal equation for $\{u_t\}$ with $\{a_r\}$ satisfying $a_r > 0$ and $\Sigma_{r=1}^{t} a_r = 1 - P_t(0)$. (The positivity of each a_t, $t \geqq 1$, follows from $P_t(0) = P_{t-1}(0)B(F_{t-1}(0))$ and the basic assumptions $0 < F(0) < 1$, $0 < B(0) < 1$.)

It is known that for $\{X_t\}$ the state space $S = \{0, 1, 2, \cdots\}$ is the union of two disjoint sets J^* and $S - J^*$ such that J^* is irreducible and aperiodic and contains the state 0. Moreover if $j \notin J^*$, $p_{ij}^{(h)} = 0$ for each integer $i \geqq 0$, $h > 0$. The set $S - J^*$ may be empty ([7], §5.2). For the stopped process $\{Y_t\}$ of present interest, similar reasoning shows that the states $\{1, 2, \cdots\}$ may be subdivided into (different) sets J^* (non-empty) and $\{1, 2, \cdots\} - J^*$ with analogous properties.

More detailed results than the foregoing basics require further assumptions. We assume that EY_0, $m \equiv F'(1-)$ and $\lambda \equiv B'(1-)$ are all finite, and $EY_0 > 0$. It is well known (e.g. [8]) that

$$E(X_t) = P_t'(1-) = \lambda(1 - m^t)/(1 - m), \qquad m \neq 1 \tag{7}$$

and since we can say in general from (3) that

$$Y_{t+1} \leqq U_1 + U_2 + \cdots + U_{Y_t} + I$$

where U_i's, Y_t and I are all independent,

$$E(Y_{t+1} \mid Y_t) \leqq Y_t m + \lambda$$

whence

$$E(Y_{t+1}) \leqq mE(Y_t) + \lambda.$$

Iteration then yields

$$EY_t \leqq m^t E(Y_0) + P_t'(1-) \tag{8}$$

where $P_t'(1-)$ is given by (7).

2. The subcritical case

In the case $m < 1$ ergodicity (i.e. positive-recurrence of the index set J^*) is known to obtain iff $\Sigma_{j=1}^{\infty} b_j \log j < \infty$, and hence certainly obtains when $\lambda < \infty$, for the process $\{X_t\}$ (see [7] for references). This implies in particular that as $t \to \infty$, $P_t(0) \to \alpha > 0$ so that in (6), $\Sigma_{r=1}^{\infty} a_r = 1 - \alpha < 1$.

Now in fact

$$a_r = P_{r-1}(0) - P_r(0) = (1 - B(F_{r-1}(0)))P_{r-1}(0) \sim \lambda(1 - F_{r-1}(0))\alpha$$

and since $(1 - F_r(0))/(1 - F_{r-1}(0)) \to m$ as $r \to \infty$, the series $\Sigma_{r=1}^{\infty} a_r z^r$ has convergence radius m^{-1}. *If we also assume* $\Sigma_{j=1}^{\infty} j \log j f_j < \infty$ as we shall do henceforth in this section, then from [4] $1 - F_r(0) \sim cm^r$ $(0 < c \leqq 1)$, so $P_{r-1}(0) - P_r(0) \sim c\lambda\alpha m^{r-1}$, and we see that the series diverges to ∞ at $z = m^{-1}$. It follows that we can find a positive number s_0, $1 < s_0 < m^{-1}$ such that

$$\sum_{r=1}^{\infty} a_r s_0^r = 1$$

and putting $\tilde{a}_t = a_t s_0^t$, $\tilde{u}_t = u_t s_0^t$, $\tilde{c}_t = c_t s_0^t$ in (6) we find

$$\tilde{u}_t = \tilde{c}_t + \sum_{r=1}^{t} \tilde{u}_{t-r} \tilde{a}_r$$

so that $\{\tilde{a}_t\}$ is a probability distribution all of whose terms are positive and all of whose moments are finite; and $\tilde{c}_t = s_0^t P_t(0)(1 - G_0(F_t(0))) \sim c\alpha E(Y_0)(s_0 m)^t$ is a sequence of positive terms such that $\Sigma_{t=0}^{\infty} \tilde{c}_t < \infty$. Applying [2], p. 330, Theorem 1, all of whose conditions are satisfied, we obtain

$$\tilde{u}_t \to \sum_{t=0}^{\infty} \tilde{c}_t \Big/ \sum_{k=1}^{\infty} k\tilde{a}_k, \quad = K, \text{ say, where } 0 < K < \infty$$

so that

$$s_0^t(1 - G_t(0)) \to K, \qquad 0 < K < \infty. \tag{9}$$

This provides information on the extinction-time distribution which completely parallels that for the ordinary subcritical branching process, with s_0 playing the role of m^{-1} in that situation. We return to the further significance of this point shortly, but are now able to pass on to another feature of similarity: the analogue of the Yaglom conditional limit theorem for the subcritical process.

Multiplying (5) by s_0^t we notice in the expression

$$\begin{aligned} s_0^t(1-G_t(s)) &= s_0^t P_t(s)(1-G_0(F_t(s))) \\ &\quad + \sum_{r=1}^{t} s_0^{t-r}(1-G_{t-r}(0))s_0^r(P_{r-1}(s)-P_r(s)) \end{aligned} \tag{10}$$

that the first summand on the right is $\leqq s_0^t(1-G_0(F_t(s)))\leqq s_0^t(1-G_0(F_t(0)))\sim cE(Y_0)(s_0m)^t \to 0$ as $t\to\infty$, since $s_0m<1$. Also

$$\begin{aligned} P_{r-1}(s)-P_r(s) &= (1-B(F_{r-1}(s)))P_{r-1}(s)\leqq 1-B(F_{r-1}(0)) \\ &\leqq \lambda(1-F_{r-1}(0))\leqq \lambda m^{r-1} \end{aligned} \tag{11}$$

by the mean value theorem, so that $\Sigma_{r=1}^{\infty} s_0^r(P_{r-1}(s)-P_r(s))<\infty$ for each $s\in[0,1]$. Noting (9) we find, letting $t\to\infty$ in (10), that

$$\lim_{t\to\infty} s_0^t(1-G_t(s)) = K\sum_{r=1}^{\infty} s_0^r(P_{r-1}(s)-P_r(s)). \tag{12}$$

From (9) and (12)

$$\begin{aligned} \frac{G_t(s)-G_t(0)}{1-G_t(0)} = 1-\frac{(1-G_t(s))}{1-G_t(0)} &\xrightarrow{t\to\infty} 1-\sum_{r=1}^{\infty} s_0^r(P_{r-1}(s)-P_r(s)) \\ &= V(s), \qquad \text{say,} \end{aligned}$$

$s\in[0,1]$, which is clearly independent of the initial distribution vector $\boldsymbol{\pi}$. Also by dominated convergence on account of (11)

$$\lim_{s\to 1-} V(s) = 1-\lim_{s\to 1-}\sum_{r=1}^{\infty} s_0^r(P_{r-1}(s)-P_r(s)) = 1,$$

so we have a proper limit distribution for $P(Y_t=j\mid Y_t>0)$, $j=1,2,\cdots$. Using analogous reasoning

$$\begin{aligned} \lim_{\delta\to 0+}(1-V(1-\delta))/\delta &= -\sum_{r=1}^{\infty} s_0^r(d(P_{r-1}(s)-P_r(s))/ds)_{s=1} \\ &= \sum_{r=1}^{\infty} s_0^r\lambda m^{r-1} = s_0\lambda/(1-s_0m) \end{aligned} \tag{13}$$

from (7).

If we denote by $p_{ij}^{(t)}$ the t-step transition probabilities of the Markov chain underlying $\{Y_t\}$, we have as a consequence of the above limit theorem that for $j\in J^*$ in particular

$$\sum_{i=1}^{\infty} \pi_i p_{ij}^{(t)}/\{1 - G_t(0)\} \to v_j, \qquad \sum_{j \in J^*} v_j = 1$$

where $\Sigma_{j=1}^{\infty} v_j s^j = V(s)$. Specializing to an initial distribution concentrated at $i \in J^*$, we have from (9)

$$s_0^t p_{ij}^{(t)} \to K(i) v_j, \qquad i, j \in J^* \tag{14}$$

where $K(i)$ can be expected to depend on i. For fixed $i \in J^*$ the limit is positive for at least one $j \in J^*$. Now the irreducible aperiodic set J^* must be either R-positive, R-null or R-transient, and the behaviour exhibited in (14) can occur only in the *R-positive case* with $s_0 = R$ ([9], Chapter 6). The R-positivity theory now shows that $v_j > 0$ for each $j \in J^*$ (since the limit in (14) must be positive for each i, j), that $\{K(i)\}$, $i \in J^*$ is the unique (to constant multiples) R-invariant vector of the set J^*; and $\{v_j\}$, $j \in J^*$ the unique (to constant multiples) invariant measure, scaled (as is possible in this particular situation) to have sum 1. These results are analogous to those of Seneta and Vere-Jones [10] for the ordinary subcritical process.

Working from the expression

$$V(s) = 1 - \sum_{r=1}^{\infty} s_0^r (P_{r-1}(s) - P_r(s))$$

and putting $\rho = 1/s_0$ $(m < \rho < 1)$, we find

$$B(s) V(F(s)) = \rho V(s) + (1 - \rho), \qquad s \in [0, 1] \tag{15}$$

which resembles the functional equation $B(s) P(F(s)) = P(s)$ for the p.g.f. $P(s)$ of the stationary-limiting distribution of the corresponding ordinary BPI, $\{X_t\}$. This suggests writing $V(s) = P(s) H(s)$, where $H(s)$ satisfies $H(0) = 0$ (since $P(0) > 0$), $H(1) = 1$, on $s \in [0, 1]$ and subsequently using the usual fixed-point shift by writing $f(x) = 1 - F(1 - x)$, $b(x) = 1 - B(1 - x)$, $h(x) = 1 - H(1 - x)$, $p(x) = 1 - P(1 - x)$, to obtain finally

$$(1 - \rho) g(x) = \rho h(x) - h(f(x)), \qquad x \in [0, 1] \tag{16}$$

where $g(x) = p(x)/\{1 - p(x)\}$.

If in (16) we regard h as an unknown function, with all other functions as specified previously, and ρ as any finite number satisfying $\rho > m$, on iteration we obtain

$$h(x) = \frac{h(f_n(x))}{\rho^n} + \frac{(1 - \rho)}{\rho} \left\{ \sum_{i=0}^{n-1} g(f_i(x))/\rho^i \right\}$$

where $f_i(x) = 1 - F_i(1 - x)$ is the ith functional iterate of $f(x)$. If we impose the constraints that $h(0) = 0$, $0 < h'(0+) < \infty$, since $h(f_n(x)) \sim h'(0+) f_n(x) \leqq h'(0+) f_n(1) \sim h'(0+) c m^n$ it follows that

(17)
$$h(x)=\frac{(1-\rho)}{\rho}\sum_{i=0}^{\infty} g(f_i(x))/\rho^i$$

(which evidently converges uniformly in $x \in [0,1]$). It is easily checked that this is indeed a solution of (16) which, under these conditions, is unique. If we seek $h(x)$ such that $h(1)=1$, then ρ must satisfy

(18)
$$\frac{\rho}{1-\rho}=\sum_{i=0}^{\infty}\frac{g(f_i(1))}{\rho^i}.$$

It is readily seen that there is just one such ρ in the interval $m<\rho$, and that it must satisfy $m<\rho<1$.

Thus returning to s_0 and the p.g.f. $V(s)$ of the Yaglom-type limit distribution, (17) and (18) indicate how these relate to the p.g.f. $P(s)$. Equation (15) provides the simplest way to compute moments of $\{v_j\}$; $\Sigma_{j=1}^{\infty} jv_j = \lambda/(\rho - m)$, in agreement with (13).

3. The supercritical case

Putting $W_t = Y_t/m^t$ in the case $m>1$ we notice from (3) that

$$\begin{aligned} E(Y_{t+1} \mid Y_t) &= mY_t + \lambda \geqq mY_t, \qquad \text{if } Y_t>0 \\ &=0 \qquad\qquad\quad \geqq mY_t \qquad \text{if } Y_t=0 \end{aligned}$$

so that in general $E(W_{t+1} \mid Y_t) \geqq W_t$. The sequence $\{Y_t\}$ is Markovian, and from (7) and (8) $EW_t \leqq EY_0 + \lambda/(m-1)$, so $\{W_t\}$ is a submartingale with bounded mean and hence by the submartingale convergence theorem (as in [8] for $\{X_t\}$)

$$W_t \xrightarrow{\text{a.s.}} W \quad \text{with } E(W) \leqq EY_0 + \lambda/(m-1)$$

where the proper random variable W may, however, be degenerate. To investigate its distribution, notice from (4) that for $s \geqq 0$

(19)
$$\begin{aligned} G_t(e^{-s/m^t}) &= P_t(e^{-s/m^t})G_0(F_t(e^{-s/m^t})) \\ &\quad + \sum_{k=0}^{t-1} G_k(0)\{P_{t-k-1}(e^{-s/m^t}) - P_{t-k}(e^{-s/m^t})\}. \end{aligned}$$

Now if in this supercritical case also we make henceforth *the additional assumption* that $\Sigma_{j=1}^{\infty} j \log j f_j < \infty$, it is known (e.g. [6], [1]) that

$$F_t(e^{-s/m^t}) \to \psi(s), \qquad P_t(e^{-s/m^t}) \to \phi(s),$$

$s \geqq 0$, where $\psi(s)$ and $\phi(s)$ are the Laplace transforms of proper non-degenerate distributions with finite means.

Now

$$P_{t-k-1}(e^{-s/m^t}) - P_{t-k}(e^{-s/m^t}) \leqq 1 - P_{t-k}(e^{-s/m^t})$$
$$= \frac{s}{m^t}\frac{d}{dx}P_{t-k}(e^x)\Big|_{x=\xi}$$

by the mean value theorem, where $-s/m^t < \xi < 0$;

$$\leqq (s/m^t)P'_{t-k}(1-)$$

and from (7)

$$= (s/m^t)\lambda(m^{t-k}-1)/(m-1) \leqq s\lambda m^{-k}/(m-1).$$

Hence by dominated convergence, letting $t \to \infty$ in (19)

$$\theta(s) \equiv E(e^{-sW}) = \phi(s)G_0(\psi(s)) + \sum_{k=0}^{\infty} G_k(0)\{\phi(s/m^{k+1}) - \phi(s/m^k)\}. \tag{20}$$

Since $(-\phi'(0+)) < \infty$, it can similarly be established by dominated convergence that $(-\theta'(0+)) < \infty$, and further information can be deduced from (20).

Note added in proof. The is some overlap between the early part of our results (to Equation (9)) and results of Zubkov [11]; ours are somewhat more general. Results for the critical case have been obtained by Ivanoff and Seneta [12].

References

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

[2] FELLER, W. (1968) *An Introduction to Probability Theory and Its Applications,* Vol. 1, 3rd edn. Wiley, New York.

[3] GARDNER, R. C., METCALF, P. AND BERGQUIST, P. L. (1977) Replication, segregation, and the hereditary stability of bacterial plasmids. Unpublished report, University of Auckland.

[4] HEATHCOTE, C. R., SENETA, E. AND VERE-JONES, D. (1967) A refinement of two theorems in the theory of branching processes. *Theory Prob. Appl.* **12**, 297–301.

[5] NOVICK, R. P. AND HOPPENSTEADT, F. C. (1978) On plasmid incompatibility. *Plasmid* **1**, 421–434.

[6] PAKES, A. G. (1971) A branching process with a state dependent immigration component. *Adv. Appl. Prob.* **3**, 301–314.

[7] SENETA, E. (1969) Functional equations and the Galton–Watson process. *Adv. Appl. Prob.* **1**, 1–42.

[8] SENETA, E. (1970) A note on the supercritical Galton–Watson process with immigration. *Math. Biosci.* **6**, 305–312.

[9] SENETA, E. (1981) *Non-negative Matrices and Markov Chains,* 2nd edn. Springer-Verlag, New York.

[10] SENETA, E. AND VERE-JONES, D. (1966) On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Prob.* **3**, 403–434.

[11] ZUBKOV, A. M. (1972) Life periods of a branching process with immigration. *Theory Prob. Appl.* **17**, 174–183.

[12] IVANOFF, B. G. AND SENETA, E. (1983) The critical branching process with immigration. *J. Appl. Prob.* **20** (4).