ON THE TRANSIENT BEHAVIOUR
OF A POISSON BRANCHING PROCESS

E. SENETA

Summary

In general, information concerning the distribution of the time to absorption, T, of a simple branching (Galton-Watson) process for which extinction in finite mean time is certain, is difficult to obtain. The process of greatest biological interest is that for which the offspring distribution is Poisson, having p.g.f. $F(s) = e^{m(s-1)}$, $m < 1$.

In the present paper, a method which consists of bounding the offspring distribution by two simpler distributions, in this case gives reasonable bounds for ET and for the mean of the asymptotic conditional distribution. The bounds are sufficiently good to permit asymptotic expressions as $m \to 1-$ of the above means, and also for the corresponding variances. Some comparison with the relevent diffusion approximation is then possible. The main results occur in § 5.

1. Introduction

Suppose a single ancestor, at the zeroth generation produces j offspring with probability given by the coefficient p_j of s^j in the probability generating function

$$F(s) = \sum_{j=0}^{\infty} p_j s^j, \qquad\qquad 0 < F(0) < 1$$

where $0 \leq s \leq 1$, and all succeeding offspring reproduce independently in the same way. Then if we denote by Z_n the number of individuals in the nth generation, we have

$$F_n(s) \equiv \sum_{j=0}^{\infty} P[Z_n = j]s^j = F_{n-1}(F(s))$$

(Harris [3], Chapter 1) i.e. $F_n(s)$ is the nth functional iterate of $F(s)$. The process $\{Z_n\}$ is then known as the simple branching process.

It is well known that if $F'(1) = m < 1$, extinction is certain to occur in finite mean time. If $m = 1$, then extinction is bound to occur although

465

the mean time to extinction is infinite, whereas if $m > 1$, there is a positive probability of survival.

Two particular results which we shall need, due to Kolmogorov and Yaglom (Harris [3], Chapter 1) state that if $m < 1$, $F''(1) < \infty$

$$(1.1) \qquad \lim_{n \to \infty} \frac{m^n}{1 - F_n(0)} = \mu \geqq 1$$

and

$$(1.2) \qquad \lim_{n \to \infty} G_n(s) = G(s), \qquad\qquad s \in [0, 1]$$

exists, where

$$G_n(s) = \sum_{j=1}^{\infty} s^j P[Z_n = j / Z_n > 0] = \frac{F_n(s) - F_n(0)}{1 - F_n(0)},$$

and $G(s)$ is a proper probability generating function with $G'(1) = \mu$. It is therefore clear, under these conditions — since

$$\operatorname{Var} Z_n = \frac{v^2 m^n (m^n - 1)}{m^2 - m}, \qquad\qquad m < 1$$

where $v^2 = \operatorname{Var} Z_1$ — that the variance of the distribution defined by $G_n(s)$, is

$$(1.3) \qquad \frac{v^2 m^n (m^n - 1)}{(1 - F_n(0))(m^2 - m)} + \frac{m^{2n}}{1 - F_n(0)} - \frac{m^{2n}}{(1 - F_n(0))^2}$$

which is bounded uniformly in n from (1.1). It can easily be shown (Harris [3], § 9) that in fact taking limits

$$(1.4) \qquad \sigma^2 = \frac{v^2 \mu}{m(1 - m)} - \mu^2$$

is the variance of the distribution defined by $G(s)$. Moreover, if we denote by T the time to absorption, then under the above conditions, since $P[T > n] = 1 - F_n(0)$, we have from (1.1), as $n \to \infty$

$$(1.5) \qquad P[T > n] \sim \mu^{-1} m^n.$$

Generally speaking, it appears difficult to obtain information about either the distribution of T, or that defined by $G(s)$. Some investigation of the general problem has been carried out by Heathcote and Seneta [5] who have obtained bounds for the means of both distributions, in terms of the first two derivatives of the offspring distribution $F(s)$. Although the bounds appear to be quite good for moderately small m in cases of interest, they become progressively worse as m approaches unity (see tables I and II of that paper). The bounds are best possible in the sense that they are

attained for the case when $F(s) = p+qs$. However, one unfortunate defect which makes them of limited practical use, is that e.g. for the Poisson case, the lower bound for ET approaches a finite limit as $m \to 1-$, whereas in view of the fact that

$$ET = \sum_{n=0}^{\infty} P[T > n]$$

and (1.5), we may expect in general (if μ^{-1} behaves sufficiently well as $m \to 1$) $ET \to \infty$ as $m \to 1-$. In biological application however, the cases of relevence occur with $m \approx 1$, and thus one of our aims is to obtain a more adequate bound to the behaviour in this vicinity.

Let us turn for the moment to the genetic aspects of the problem. In this context, the process $\{Z_n\}$ has been used by Fisher [2], Chapter IV, to describe the history of the descendants of a single mutant gene. The particular offspring distribution used by him is that of greatest biological significance, the Poisson, with $F(s) = e^{m(s-1)}$. More recently, Moran [6] Chapter V, has given an account of the theory confining himself also almost completely to this offspring distribution. Moreover, it is interesting to note that both Fisher and Moran have been almost entirely concerned with the cases $m > 1$, $m = 1$, these cases representing the situations when the mutant gene has a selective advantage and no selective advantage (or disadvantage) respectively. The case $m < 1$, although possibly less important, is the object of study in the present paper. In particular, since little interest attaches to transient behaviour unless $m \approx 1$ (for extinction occurs rapidly otherwise), we focus our attention on the asymptotic behaviour in this Poisson case, of ET, μ and the corresponding variances as $m \to 1$.

The reason for restricting oneself just to the Poisson distribution, is a consequence of the method used, which is not a general one, and must be applied separately to any specific distribution. The technique may, however, be useful more generally, and in the next section at least, the approach is not confined to a particular function $F(s)$.

2. Theoretical basis

The following lemma[1] forms the crux of our approach. We suppose that $F^{(1)}(s)$, $F^{(2)}(s)$ are two probability generating functions, defined for $s \in [0, 1]$.

LEMMA A. *If* $F^{(1)}(s) \geqq F^{(2)}(s)$ *for* $s \in [0, 1]$, *then* $F_n^{(1)}(s) \geqq F_n^{(2)}(s)$ *for all positive integral* n, *for* $s \in [0, 1]$.

---

[1] Which arose out of a remark by Professor P. A. P. Moran.

PROOF.
$$F^{(1)}_{n+1}(s) = F^{(1)}_n\big(F^{(1)}(s)\big)$$
$$\geq F^{(2)}_n\big(F^{(1)}(s)\big) \quad \text{by induction;}$$
$$\geq F^{(2)}_n\big(F^{(2)}(s)\big)$$
$$= F^{(2)}_{n+1}(s).$$

It is in the sense of inequalities between generating functions that we shall carry out our bounding procedure, with the aid of this lemma (whose validity is also obvious diagrammatically).

Before proceeding, it is useful to obtain some insight, into the intuitive *probabilistic* meaning of the statement $F^{(1)}(s) \geq F^{(2)}(s)$ by giving a sufficient condition for it to hold. Suppose $X^{(1)}$ and $X^{(2)}$ are two random variables, and we define a relation $\succ$, so that

$$X^{(1)} \succ X^{(2)} \quad \text{iff} \quad P[X^{(1)} \leq x] \geq P[X^{(2)} \leq x]$$

for all x.

LEMMA B. *Let* $\{Z^{(1)}_n\}$ *and* $\{Z^{(2)}_n\}$ *be two discrete branching processes such that*

$$Z^{(1)}_1 \succ Z^{(2)}_1.$$

Then

$$F^{(1)}(s) \geq F^{(2)}(s) \qquad\qquad s \in [0, 1].$$

PROOF.
$$\sum_{j=0}^{\infty} P[Z^{(1)}_1 > j]s^j = \frac{1 - F^{(1)}(s)}{1-s}$$
$$\leq \sum_{j=0}^{\infty} P[Z^{(2)}_1 > j]s^j = \frac{1 - F^{(2)}(s)}{1-s}$$

for $s \in [0, 1]$. Hence
$$F^{(1)}(s) \geq F^{(2)}(s), \qquad\qquad s \in [0, 1].$$

We do not pursue the development of this topic further, and return to the practical problem with which we are at the moment concerned.

Suppose we can find offspring distribution p.g.f.'s $F^{(L)}(s)$ and $F^{(U)}(s)$ whose characteristics are simple to determine, such that

(2.1) $$F^{(L)}(s) \leq F(s) \leq F^{(U)}(s).$$

Then if $m < 1$, and $m_L, m_U < 1$,

$$\sum_{n=0}^{\infty} n^k\big(1 - F^{(U)}_n(0)\big) \leq \sum_{n=0}^{\infty} n^k\big(1 - F_n(0)\big) \leq \sum_{n=0}^{\infty} n^k\big(1 - F^{(L)}_n(0)\big)$$

for $k \geq 0$ by Lemma A, so that

(2.2) $$\sum_{n=0}^{\infty} n^k\big(1 - F^{(U)}_n(0)\big) \leq \sum_{n=0}^{\infty} n^k P[T > n] \leq \sum_{n=0}^{\infty} n^k\big(1 - F^{(L)}_n(0)\big).$$

(N.B. If $m < 1$, $1 - F_n(0) \leqq m^n$ whether $F''(1)$ is finite or not.) Thus we may use (2.2) to get bounds for

(2.3) $$ET = \sum_{n=0}^{\infty} P[T > n]$$

(2.4) $$\tfrac{1}{2}\{\mathrm{Var}\, T + (ET)^2 - (ET)\} = \sum_{n=0}^{\infty} n P[T > n]$$

etc. if the bounding sums can be found. This is obviously the drawback of the method. There are only a few offspring generating functions whose iteration characteristics are manipulable, for example the two point offspring distribution

(2.5) $$F(s) = p + qs \qquad\qquad 0 < p < 1$$

for which $1 - F_n(0) = q^n$; and the bilinear fractional form

(2.6) $$F^{(G)}(s) = 1 - \frac{b}{1-c} + \frac{bs}{1-cs} ; \qquad \begin{matrix} b,\, c > 0; \\ b + c < 1 \end{matrix}$$

which arises out of the modified geometric distribution

$$p_0 = \frac{1-b-c}{1-c}, \quad p_k = bc^{k-1} \qquad k = 1, 2, \cdots$$

(Harris [3], page 9), although even the latter requires considerable manipulation as regards its extinction time distribution when $m_G < 1$.

The two-point distribution is unfortunately of little use, although it does provide bounds for any $F(s)$ for which $m < 1$. Clearly for $s \in [0, 1]$

$$F(0) + (1 - F(0))s \geqq F(s) \geqq 1 - m + ms$$

so that

$$[1 - F(0)]^n \leqq 1 - F_n(0) \leqq m^n.$$

Hence

$$\sum_{n=0}^{\infty} n^k [1 - F(0)]^n \leqq \sum_{n=0}^{\infty} n^k P[T > n]$$

$$\leqq \sum_{n=0}^{\infty} n^k m^n.$$

In particular

(2.7) $$\frac{1}{F(0)} \leqq ET \leqq \frac{1}{1-m},$$

(2.8) $$\frac{1 - F(0)}{[F(0)]^2} \leqq \sum_{n=0}^{\infty} n P[T > n] \leqq \frac{m}{(1-m)^2}.$$

Both these bounds are obvious for intuitive reasons also. Since the ower bound of (2.7) does not increase indefinitely as $m \to 1-$ in the

Poisson case, for instance, it is of little use for values close to unity. However, it does somewhat surprisingly provide a slightly higher bound in this case for values of $m \geq .8$ than that of Heathcote and Seneta [5], whose lower bound also approaches a finite limit as $m \to 1$, although not monotonically. Moreover, the two-point offspring distribution may be used to give slightly better bounds than (2.7) and (2.8) by fairly obvious refinements.

The modified geometric distribution bounds give much more satisfactory results for the Poisson distribution. This is in part due to the fact that the geometric and Poisson distributions somewhat resemble each other.

To obtain bounds for the asymptotic conditional mean and variance

$$\mu, \sigma^2 = \frac{v^2 \mu}{m(1-m)} - \mu^2$$

we require in addition to the condition (2.1) with $m_L, m_U, m < 1$, the existence of all second derivatives of the bounding functions at unity [2], and $m_L = m_U = m$, for then

$$\frac{m^n}{1-F_n^{(U)}(0)} \geqq \frac{m^n}{1-F_n(0)} \geqq \frac{m^n}{1-F_n^{(L)}(0)}.$$

We can in fact achieve this for the Poisson distribution with a suitable choice of parameters a, b, c in (2.6), consistent with all other requirements.

3. The bilinear fractional generating function

Before we may proceed with our investigation of the Poissonian offspring distribution, we need to establish certain results for the offspring distribution defined by (2.6). In this case

$$m_G \equiv \frac{b}{(1-c)^2}$$

and we shall be concerned only with the case $m_G < 1$. We shall also need to utilize the constant

$$s_0 = \frac{1-b-c}{c(1-c)} = \frac{1}{c} \{1 - m_G(1-c)\}$$

which clearly exceeds unity when $m_G < 1$. Moreover we have

$$1 - F_n^{(G)}(0) = \frac{(s_0 - 1)m_G^n}{s_0 - m_G^n}, \qquad n \geqq 0$$

[2] This condition can be weakened to $\sum_{j=1}^{\infty} j(\log j) p_j < \infty$. See Heathcote, Seneta and Vere-Jones (1967).

so that, denoting by T_G the time to absorption in this case

(3.1) $$\sum_{n=0}^{\infty} n^k P[T_G > n] = (s_0 - 1) \sum_{n=0}^{\infty} \frac{n^k (m_G^n/s_0)}{(1 - m_G^n/s_0)}.$$

The series on the right hand side is not tractable, even for $k = 0$. We can however obtain sufficient information for our ultimate needs by use of the so called "integral test" technique.

For the series (Heathcote and Seneta [5])

$$\sum_{j=M}^{\infty} \frac{\tau s^j}{1 - \tau s^j} \qquad M > 0, \quad 0 < \tau, \quad s < 1,$$

since the terms are monotone decreasing with respect to j,

$$\int_M^{\infty} \frac{\tau s^x}{1 - \tau s^x} dx \leq \sum_{j=M}^{\infty} \frac{\tau s^j}{1 - \tau s^j} \leq \frac{\tau s^M}{1 - \tau s^M} + \int_M^{\infty} \frac{\tau s^x}{1 - \tau s^x} dx$$

i.e.

(3.2) $$\frac{\log (1 - \tau s^M)}{\log s} \leq \sum_{j=M}^{\infty} \frac{\tau s^j}{1 - \tau s^j} \leq \frac{\tau s^M}{1 - \tau s^M} + \frac{\log (1 - \tau s^M)}{\log s}.$$

This can clearly be applied to the tail of (3.1) with $k = 0$. On the other hand the terms of

$$\sum_{j=0}^{\infty} \frac{j \tau s^j}{1 - \tau s^j}$$

only become monotone decreasing for j exceeding some value N^*, depending on τ and s, and are monotone increasing up to this point. The number N^* can be approximately located by equating to zero the derivative of the function of a continuous variable x

$$\alpha(x) = \frac{x \tau s^x}{1 - \tau s^x}$$

the resulting equation finally being

$$1 + x \log s - \tau s^x = 0.$$

As it happens we shall need to know N^* only for the case when, as $s \to 1-$,

(3.3) $$1 - \tau \sim K(1 - s),$$ $(0 < K = \text{const.})$

in which case it is easy to see that the unique solution of the equation is asymptotically:

(3.4) $$N^* \gtrsim \sqrt{2K}/\sqrt{-\log s} \quad \text{and} \quad N^* \leq -\frac{1}{\log s}.$$

We can now apply the integral test twice, up to the point (N^*-1) and from

the point (N^*+1) onwards to obtain, by adding the results, and compensating

(3.5)
$$\int_0^\infty \alpha(x)dx + \alpha(N^*) - \int_{N^*-1}^{N^*+1} \alpha(x)dx \leq \sum_{j=0}^\infty \alpha(j)$$

$$\leq \int_0^\infty \alpha(x)dx + \alpha(N^*-1) + \alpha(N^*) + \alpha(N^*+1) - \int_{N^*-1}^{N^*+1} \alpha(x)dx.$$

Now

$$\int_0^\infty \frac{x\tau s^x}{1-\tau s^x} dx = -\frac{1}{(\log s)^2}\int_0^\tau \left\{\frac{\log y - \log \tau}{1-y}\right\} dy$$

$$\sim \frac{\pi^2}{6(\log s)^2}, \quad \text{if} \quad (1-\tau) \sim K(1-s)$$

as $s \to 1$, which can be seen by considering the integral in two parts, and neglecting the component of lower order. Moreover

$$\alpha(N^*) = \frac{N^*\tau s^{N^*}}{1-\tau s^{N^*}} \approx \tfrac{1}{2}\int_{N^*-1}^{N^*+1} \alpha(x)dx = O\left(\frac{-1}{(\log s)^{\frac{3}{2}}}\right)$$

as $s \to 1$ if $(1-\tau) \sim K(1-s)$, so that under this condition, finally from (3.5)

(3.6)
$$\sum_{j=0}^\infty \frac{j\tau s^j}{1-\tau s^j} \sim \frac{\pi^2}{6(\log s)^2}.$$

Applying the expressions (3.2) and (3.6) to (3.1) we have

(3.7)
$$1 + (s_0-1)\frac{\log(1-m_G/s_0)}{\log m_G} \leqq ET_G$$

$$\leqq 1 + \frac{(s_0-1)m_G}{s_0-m_G} + (s_0-1)\frac{\log(1-m_G/s_0)}{\log m_G}$$

and if, as $m_G \to 1-$, $(1-1/s_0) \sim K(1-m_G)$, then also

(3.8) $$\tfrac{1}{2}[\text{Var } T_G + (ET_G)^2 - (ET_G)] \sim \frac{(s_0-1)\pi^2}{6(\log m_G)^2} \sim \frac{K\pi^2}{6(1-m_G)}.$$

The expressions (3.7) and (3.8) are of great importance in the following, apart from some independent interest.

4. Bounding the Poisson generating function

Our next step is to bound the Poisson generating function, which we shall henceforth denote by $F^{(P)}(s)$, by two generating functions of the bilinear fractional form, which we shall denote by $F^{(U)}(s)$ (the upper bound) and $F^{(L)}(s)$ (the lower).

By the mean value theorem

$$F^{(P)}(s) - F^{(G)}(s) = \left\{ me^{m(s^*-1)} - \frac{b}{(1-cs^*)^2} \right\}(s-1)$$

where $s^* \in (s, 1)$. It follows that,

$$me^{m(s-1)} - \frac{b}{(1-cs)^2} \begin{matrix} \geq 0 \\ \leq 0 \end{matrix}$$

for all $s \in (0, 1)$ imply respectively that

$$F^{(P)}(s) \leq F^{(G)}(s)$$
$$\geq F^{(G)}(s).$$

Thus if the parameters can be chosen to respectively fulfil these two conditions, we shall be in a position to apply the consequences of lemma A. In addition we shall need to fulfil the following restrictions on the parameters

(4.1) $\qquad \begin{cases} b, c \geq 0, \quad b+c \leq 1 \\ \dfrac{b}{(1-c)^2} < 1. \end{cases}$

We proceed by investigating the function

$$\phi(s) = e^{ms}(1-cs)^2 - bm^{-1}e^m,$$

so that

$$\frac{d}{ds}\phi(s) = 0$$

implies

$$\{1-cs\}\{m(1-cs) - 2c\} = 0$$

so that the two roots are

(4.2) $\qquad s_1 = 1/c, \quad s_2 = (1-2c/m)/c.$

The smaller root corresponds to a maximum of $\phi(s)$ and the larger to a minimum. Consideration reveals that if we choose

$$s_2 = 0 \quad \text{i.e.} \quad c = m/2$$

then

$$\phi(1) = e^m(1-c)^2 - bm^{-1}e^m \geq 0$$

(and so

$$\phi(s) \geq 0, \ s \in [0, 1] \ )$$

implies

$$m \geq \frac{b}{(1-c)^2}.$$

Hence a suitable choice of b and c which in fact equates the means of the Poisson and geometric distributions and for which

(4.3)
$$
\begin{cases}
F^{(P)}(s) \leq F^{(G)}(s) \equiv F^{(U)}(s), \quad s \in [0, 1] \\[2mm]
\text{is} \\[2mm]
c = \dfrac{m}{2}, \quad b = m\left(1 - \dfrac{m}{2}\right)^2 \\[2mm]
\text{where} \\[2mm]
s_0 = \dfrac{(1-m)^2+1}{m}, \quad \dfrac{b}{(1-c)^2} = m.
\end{cases}
$$

NOTE: $b+c < 1$, so that all the conditions (4.1) are satisfied.

To obtain the reverse inequality

$$
F^{(P)}(s) \geq F^{(L)}(s) \quad \text{for} \quad s \in [0, 1]
$$

we proceed as follows. Choose c so that

$$
s_2 = \frac{1}{c}\left(1 - \frac{2c}{m}\right) = 1
$$

i.e.
$$
c = \frac{m}{m+2}.
$$

Then, since $s_1 = 1/c > 1$, we shall get the required (i.e. $\phi(s) \leq 0$, $s \in [0, 1]$) if

$$
\phi(1) \leq 0
$$

i.e.
$$
\frac{b}{(1-c)^2} \geq m.
$$

Hence to make this an equality we have the choices

$$
c = \frac{m}{m+2}, \quad b = m\left(\frac{2}{m+2}\right)^2, \quad (b+c < 1).
$$

To summarize

(4.4)
$$
\begin{cases}
F^{(P)}(s) \geq F^{(G)}(s) \equiv F^{(L)}(s), \quad s \in [0, 1] \\[2mm]
c = \dfrac{m}{m+2}, \quad b = m\left(\dfrac{2}{m+2}\right)^2 \\[2mm]
s_0 = \dfrac{2-m}{m}, \quad \dfrac{b}{(1-c)^2} = m.
\end{cases}
$$

NOTE. There is, clearly, a certain arbitrariness in the choice of the parameters b, c. The above choices of (4.3) and (4.4) are justified in that

they not only equate the means of the two bounding offspring distributions with that of the one under investigation, but (possibly as a result) give good bounds, and asymptotes (as $m \to 1-$).

5. Bounds and asymptotes

As a consequence of the results established in all the previous sections it is a simple matter to proceed to bounds and asymptotes.

5.1. MEANS. From lemma A, and the double inequality (3.7) we have for the Poisson offspring distribution $F(s) = e^{m(s-1)}$, using (4.3) and (4.4):

(5.1)

$$1 + \frac{(1-m)(2-m) \log \left\{\frac{2(1-m)}{1+(1-m)^2}\right\}}{m \log m}$$
$$\leqq ET$$
$$\leqq 1 + \frac{2m}{2+m} + \frac{2(1-m) \log \left\{\frac{(1-m)(2+m)}{2-m^2}\right\}}{m \log m}.$$

Moreover,

$$\frac{2-m-m^{n+1}}{2(1-m)} \leqq \frac{m^n}{1-F_n(0)} \leqq \frac{1+(1-m)^2-m^{n+1}}{(1-m)(2-m)}$$

yields in the limit as $n \to \infty$

(5.2)
$$\frac{2-m}{2(1-m)} \leqq \mu \leqq \frac{1+(1-m)^2}{(1-m)(2-m)}.$$

From (5.1)

(5.3)
$$ET \sim -\theta_m \log(1-m)$$

as $m \to 1-$, where $1 \lesssapprox \theta_m \lesssapprox 2$. From (5.2)

(5.4)
$$\mu \sim \frac{1}{\psi_m(1-m)}$$

as $m \to 1-$, where $1 \lesssapprox \psi_m \lesssapprox 2$.

Tables I and II compare the true values of ET and μ with the bounds (5.1), (5.2) (c.f. Heathcote and Seneta [5], tables[3] I and II).

Precise computation of the true values of ET and μ is lengthy, even on a high-speed computer for higher values of m; the bounds however are simple to calculate.

[3] The upper values of these tables are incorrect: see correction to [5].

TABLE I

$ET; \ F(s) = e^{m(s-1)}$

m	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
Upper Bound	1.4983	1.7168	1.9232	2.1359	2.3712	2.6504	3.0153	3.5597	4.5984	5.7627
True Value	1.1057	1.2257	1.3657	1.5332	1.7405	2.0083	2.3762	2.9368	3.9974	5.1637
Lower Bound	1.0411	1.1105	1.2053	1.3278	1.4829	1.6708	1.9325	2.2846	2.8786	3.5912

TABLE II

$\mu; \ F(s) = e^{m(s-1)}$

m	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
Upper Bound	1.0585	1.1389	1.2521	1.4167	1.6667	2.0714	2.7949	4.3333	9.1818	19.10
True Value	1.0564	1.1284	1.2327	1.3493	1.5261	1.7932	2.2332	3.1065	5.7148	10.77
Lower Bound	1.0556	1.1125	1.2143	1.3333	1.5000	1.7500	2.1667	3.0000	5.5000	10.50

Thus for $m = .99, \ 4.97 \leqq ET \leqq 7.76; \quad 50.5 \leqq \mu \leqq 99.0$

           $m = \cdot 999, \ 7.23 \leqq ET \leqq 13.3; \quad 500 \quad \leqq \mu \leqq 909.$

5.2 VARIANCES. Since for both bounding distributions, as $m \to 1-$

$$\left(1 - \frac{1}{s_0}\right) \sim K(1-m),$$

it follows from (3.8) that

$$\tfrac{1}{2}[\operatorname{Var} T + (ET)^2 - (ET)] \sim -\rho_m \frac{\pi^2}{6(\log m)}$$

(where $1 \gtrless \rho_m \gtrless 2$) as $m \to 1$, so that utilizing (5.3)

(5.5) $\operatorname{Var} T \sim -\rho_m \dfrac{\pi^2}{3(\log m)} \sim \rho_m \dfrac{\pi^2}{3(1-m)}.$

The following table gives the true values of Var T for comparison.

TABLE III

$\operatorname{Var} T; \ F(s) = e^{m(s-1)}$

m	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
Var T	0.1179	0.2859	0.5353	0.9218	1.554	2.670	4.883	10.25	30.55	79.47
$\pi^2/3(1-m)$	—	—	—	—	—	—	—	16.45	32.90	65.80

For the variance of the asymptotic conditional distribution, we may readily obtain bounds from (5.2), noting that in the Poisson case $v^2 = m$. Asymptotically as $m \to 1-$

$$(5.6) \qquad \sigma^2 = \frac{\mu}{(1-m)} - \mu^2 \sim \frac{1}{\psi_m} \left(1 - \frac{1}{\psi_m}\right) \frac{1}{(1-m)^2}$$

where $1 \gtrless \psi_m \gtrless 2$. It is therefore interesting to note that in this case, as $m \to 1$

$$(5.7) \qquad \frac{\mu^2}{\sigma^2} \sim \frac{1}{\psi_m - 1}$$

with Table II indicating that this is in fact close to unity.

To conclude this section, several remarks seem appropriate.

Firstly, from (5.3) and (5.5) we notice that the mean of the distribution of T increases much more slowly than the variance as $m \to 1-$: This is also evident from Tables I and III, from which we see that at $m = .95$ the mean is only ≈ 5 whereas the variance is in the vicinity of 80. A consequence of this fact is that in the vicinity of $m = 1$, ET is a poor indication of the time to extinction, a fact which is occasionally overlooked in applications, and emphasized by the virtual uselessness of the Tchebycheff inequalities, e.g.

$$P[|T - ET| > y] < \operatorname{Var} T / y^2$$

in the present situation, as can easily be checked. In contrast, the situation for the asymptotic conditional distribution, as summarized by (5.4), (5.6) and (5.7) is more "stable".

Secondly, it appears likely that the quantities θ_m, ψ_m and ρ_m may be replaced by constants θ, ψ and ρ as $m \to 1$, although it is a weakness of the present method that this may not be deduced from it. In fact one may not unreasonably conjecture that $\theta = \psi = \rho$. Some light is cast on these problems by the following section.

Lastly, one is led to wonder from (5.7) whether a relation of the form $\mu^2/\sigma^2 \sim 1$ as $m \to 1-$ does not hold under fairly general conditions on the offspring distribution. For instance, as will be seen, it is supported by the diffusion approximation theory of the next section with ψ_m replaced by 2, which only requires that $m \le 1$, $m \sim 1$, $F''(1-) < \infty$, $F'''(1-) < \infty$. For the bilinear fractional case defined in § 3, the asymptotic conditional p.g.f. is

$$G(s) = \frac{\left(1 - \dfrac{1}{s_0}\right) s}{1 - \dfrac{s}{s_0}} .$$

Hence

$$\mu = \cfrac{1}{1 - \cfrac{1}{s_0}}$$

$$\sigma^2 = \cfrac{\cfrac{1}{s_0}}{\left(1 - \cfrac{1}{s_0}\right)^2}$$

$$\frac{\mu^2}{\sigma^2} = s_0 \to 1 \qquad\qquad \textit{if } s_0 \to 1 \textit{ as } m \to 1.$$

This point however, needs deeper analysis, because of the general difficulty of what is meant for an offspring distribution not as simply defined as the Poisson by the statement "$m \to 1$". For instance, it is not difficult to see that by choosing

$$c = 1 - \xi, \quad b = \xi^3 \qquad\qquad (0 < \xi < 1)$$

so that

$$m = \frac{b}{(1-c)^2} = \xi$$

in the bilinear fractional case, all the necessary restrictions are satisfied, and yet

$$s_0 = \frac{1}{c}\{1 - m(1-c)\} = 1 + m$$

$$\to 2 \qquad\qquad\qquad\qquad\qquad \text{as } m \to 1-.$$

Thus it appears that care is necessary in how the parameters of an offspring distribution behave as $m \to 1$. We shall not, however, pursue this topic further in the present paper, leaving it to a further note.

6. The diffusion approximation

It has been shown by Feller [1] that if new units are introduced for measuring time and population in a discrete branching process, a diffusion approximation, with a known estimate of error is possible. When $m = 1 - \delta$ where δ is small and positive (which is precisely the case of interest in the present paper) and an individual in the old counting and the time of one generation are both δ, the density function $\phi(x, t)$ of the modified population variate x at modified time t satisfies a Fokker-Planck equation, where the error in the equation is $0(\delta)$. (For more detail on the present section see § 5 of Seneta [7].)

The asymptotic conditional distribution is then approximated by the density

$$f(x) = \frac{2}{v^2} e^{-(2/v^2)x}, \qquad\qquad x \in (0, \infty)$$

where v^2 is the variance of the original offspring distribution (assumed finite). The mean and variance of the above distribution are well known to be

$$\mu = \frac{v^2}{2}, \quad \sigma^2 = \frac{v^4}{4}$$

so that

$$\frac{\mu^2}{\sigma^2} = 1$$

which agrees with our conjecture at the conclusion of the previous section, since the diffusion approximation used above may only be considered to hold "exactly" in the limit as $\delta \to 0$. Notice also, that in the above ratio, scale factors are irrelevant.

Further, we note that by reintroducing the scaling factor $\delta = 1-m$, we obtain for the unscaled asymptotic conditional distribution, that

$$\mu \simeq \frac{v^2}{2(1-m)} \sim \frac{1}{2(1-m)}$$

as $m \to 1-$ in the Poisson case.

This agrees with (5.4) and suggests that ψ_m may be replaced by 2.

We may also obtain some information about T, by noticing that if T^* is the time to extinction of the modified process with initial point x_0, then

$$(6.1) \qquad P[T^* > t] = 1 - \exp\left\{\frac{2x_0\, e^{-t}}{v^2(e^{-t} - 1)}\right\}$$

so that, in theory, the moments of T^* are obtainable, e.g.

$$ET^* = \int_0^\infty \left[1 - \exp\left\{\frac{2x_0\, e^{-t}}{v^2(e^{-t} - 1)}\right\}\right] dt.$$

By putting $s = 1/e^t - 1$

$$(6.2) \qquad ET^* = \int_0^\infty \frac{1 - e^{-ks}}{s(1+s)}\, ds$$

where

$$k = \frac{2x_0}{v^2}.$$

Unfortunately, the integral of form (6.2) does not appear to be known. It is not worthwhile evaluating it by numerical methods, since the amount of labour is no less than to calculate $\sum_{n=0}^{\infty}(1-F_n(0))$, and even so, only approximate values of ET can be obtained on account of the diffusion approximation involved.

In conclusion, and in view of our purpose of considering transient behaviour when $m \approx 1$, we may obtain the following information from the diffusion approximation, which may be compared with the other asymptotic results of the paper. Denoting by M the maximum number of individuals which the population attains before extinction, starting from a single individual, it can be shown that in the Poisson case

$$(6.3) \qquad \begin{cases} EM \approx -\log(1-m) \\[2mm] \operatorname{Var} M \approx \dfrac{\pi^2}{6(1-m)}. \end{cases}$$

It is thus interesting to note that as $m \to 1-$, EM and ET are — apart from a bounded factor — of the same order of variation (a result that is intuitively satisfying). Similarly, as $m \to 1-$

$$\operatorname{Var} T \sim \rho_m \frac{\pi^2}{3(1-m)} \approx 2\rho_m \operatorname{Var} M.$$

Acknowledgements

References

[1] W. Feller, 'Diffusion processes in genetics'. *Proc. Second Berkeley Symp. Math. Stat. and Prob.* (1951), 227—246.
[2] R. A. Fisher, *The Genetical Theory of Natural Selection* (Dover, New York, 1958).
[3] T. E. Harris, *The Theory of Branching Processes* (Springer-Verlag, Berlin, 1963).
[4] C. R. Heathcote, E. Seneta, and D. Vere-Jones, 'A refinement of two theorems in the theory of branching processes.' *Teor. Veroiat. Primenen.* 12 (1967), 197—201.
[5] C. R. Heathcote, and E. Seneta, 'Inequalities for branching processes'. *J. Appl. Prob.*, 3 (1966), 261—267.
[6] P. A. P. Moran, *The Statistical Processes of Evolutionary Theory* (Clarendon Press, Oxford, 1962).
[7] E. Seneta, Quasi-stationary distribution and time-reversion in genetics, *J.R.S.S. (B)* 28 (1966), 253—277.

Australian National University
Canberra, Australia