ON SOME DISCRETE MODELS IN BRANCHING PROCESSES

By U. G. Nadkarni

I.A.R.S., New Delhi

Introduction

A branching stochastic process is a process which corresponds to the phenomenon such as propagation of animal species and nuclear chain reactions. In each of these, the growth of a set of initial units (*i.e.*, organisms or particles) is considered. The initial set of units can produce more units of the same type. These units can themselves generate more individuals and the set gets enlarged. Such processes have been studied by different authors[1-3] under various assumptions. In the discrete models considered, it is assumed that the probabilities of fission of every unit of the set are constants. Also the law of propagation for each individual is assumed to be same and is taken to be independent of the mode of fission of the other. These models generally hold good over a finite interval of generations with infinite natural resources. However, under the natural limitations, the probabilities of fission need not be constants. Some modified models when the probabilities of fission of an unit into r units are dependent upon the existing number of units are considered in this paper. The expressions for the mean, the variance, the probabilities of extinction and maximum likelihood estimates of the parameters involved are obtained.

Modified Branching Process

A sequence of random variables, $z_0, z_1, z_2, \cdots$ is a modified branching process if and only if the following conditions are satisfied:

1. $z_0, z_1, z_2, \cdots$ are non-negative integer-valued random-variables $z_0 = 1$, $P(z_1 = r) = p_1(r)$, $p_1(0) + p_1(1) < 1$.

2. Sequence is Markovian.

3. The variable z_n is a sum of a number independent identically distributed random variables, the number depending upon the value assumed by z_{n-1}. It represents the number of units in the n-th generation.

4. There exists a function $\phi(j, r)$ which is a decreasing function of j uniformly relative to r or otherwise, and is the probability that any unit in a population of size j produces r individuals, all these probabilities being independent of each other and $\phi(1, r) = p_1(r)$.

Let

$$\phi(j, r) = \frac{e^{-m_j} m_j^r}{r!}$$

where

$$m_j = a + \frac{\beta}{j} \tag{1}$$

and $a + \beta = m_1$, a and β being positive constants.

Define the generating functions for the 1st and n-th generation by

$$f_1(s) = \sum_{i=0}^{\infty} p_1(i)\, s^i \text{ and } f_n(s) = \sum_{i=0}^{\infty} p_n(i)\, s^i$$

where $|s| \leqslant 1$ and $p_n(i)$ is the probability of there being i individuals in the n-th generation.

Then, from the definition of $\phi(j, r)$ we have

$$f_{n+1}(s) = \sum_{j=0}^{\infty} p_n(j) \left\{ \sum_{r=0}^{\infty} \frac{e^{-(a+\beta/j)} \left(a + \frac{\beta}{j}\right) s^r}{r!} \right\}^j$$

$$= \sum_{j=0}^{\infty} p_n(j)\, e^{-(aj+\beta)} \quad e^{(aj+\beta)s}$$

$$= e^{\beta(s-1)} f_n\left[e^{a(s-1)}\right]. \tag{2}$$

The mean of the $(n+1)$-th generation is obtained from this recursion relation by differentiating it with respect to s and putting $s = 1$

$$f'_{n+1}(1) = E(z_{n+1}) = \beta + a f_n'(1)$$

$$= \beta \left[\frac{1 - a^n}{1 - a}\right] + a^n m_1. \tag{3}$$

In the limit as n the number of generations becomes very large, we have

$$\underset{n\to\infty}{Lt}\, E(z_{n+1}) = \frac{\beta}{1-a} \qquad 0 < a < 1$$

$$= \infty \qquad a \geqslant 1, \tag{4}$$

And for the variances consider,

$$f''_{n+1}(1) = \beta^2 + 2\alpha\beta f_n'(1) + \alpha^2 \{f_n''(1) + f_n'(1)\}$$

or

$$f_n''(1) = \beta^2 \left\{\frac{1-\alpha^{2n-2}}{1-\alpha^2}\right\} + \beta\left(\frac{\alpha^2+2\alpha\beta}{1-\alpha}\right)\left[\frac{1-\alpha^{2n-4}}{1-\alpha^2}\right]$$

$$+ \alpha^{2n-3}(\alpha+2\beta)\, m_1 + \alpha^{n-1}(\alpha+2\beta)\left\{\frac{1-\alpha^{n-2}}{1-\alpha}\right\}$$

$$\times \left\{m_1 - \frac{\beta}{1-\alpha}\right\} + \alpha^{2n-2} f_1''(1).$$

The variance of z_n is therefore given by

$$\text{Var.}(z_n) = \beta^2\left(\frac{1-\alpha^{2n-2}}{1-\alpha^2}\right) + \beta\left(\frac{\alpha^2+2\alpha\beta}{1-\alpha}\right)\left(\frac{1-\alpha^{2n-4}}{1-\alpha^2}\right)$$

$$+ \alpha^{2n-3}(\alpha+2\beta)\, m_1 + \alpha^{n-1}(\alpha+2\beta)\left(\frac{1-\alpha^{n-2}}{1-\alpha}\right)$$

$$\times \left(m_1 - \frac{\beta}{1-\alpha}\right) + \alpha^{2n-2} m_1$$

$$- \left\{\frac{\beta(1-\alpha^{n-1})}{(1-\alpha)} + \alpha^{n-1} m_1\right\}^2$$

$$+ \left\{\frac{\beta(1-\alpha^{n-1})}{(1-\alpha)} + \alpha^{n-1} m_1\right\}.$$

As n tends to infinity, the variance reduces to

$$\text{Var.}_{n\to\infty}(z_n) = \frac{\beta^2}{1-\alpha} + \frac{\beta(\alpha^2+2\alpha\beta)}{(1-\alpha)(1-\alpha^2)} - \frac{\beta^2}{(1-\alpha)^2} + \frac{\beta}{(1-\alpha)}$$

$$= \frac{\beta}{(1-\alpha)(1-\alpha^2)}. \tag{5}$$

The mean and variance of z_n, the number of units in the n-th generation as n becomes large attain constant values for $\alpha < 1$. Th mean gradually decreases to a value $\beta/(1-\alpha)$.

Probability of Extinction

From the recursion relation (2) we have for the probability of extinction:

$$f_{n+1}(0) = e^{-\beta} f_n(e^{-\alpha}) = e^{-\beta} f_n(s_1)$$
$$= e^{\beta(s_1+s_2+\cdots s_{n-1}-n)} \cdot f_1(s_n) = e^{\beta S_{n-1}-1} \cdot f_1(s_n)$$

where

$$s_n = e^{\alpha(s_{n-1}-1)} \text{ and } S_{n-1} = \sum_{i=1}^{n-1}(s_i - 1).$$

Now,

$s_n/s_{n-1} = e^{\alpha(s_{n-1}-s_{n-2})}$ and for $\alpha > 0$ it is seen that $s_n > s_{n-1} > s_{n-2} \cdots > s_3$ if $s_2 > s_1$. But, since $s_2/s_1 = e^{\alpha e^{-\alpha}} > 1$ s_n is monotonically increasing sequence such that $s_n < 1$ as $s_1 < 1$. Monotonic and bounded sequence s_n attains its limit l, which satisfies the equation $l = e^{\alpha(l-1)}$. Consider, then the sum $S_n = \sum_{i=1}^{n}(s_i - 1)$ which is monotonically decreasing with n, as $S_n - S_{n-1} = s_n - 1$ is negative. Further, $(s_i - 1) > -\alpha^i$ as $e^{-\alpha} > 1 - \alpha$ and for $0 < \alpha < 1$

$$S_n > -\{\alpha + \alpha^2 + \alpha^3 + \cdots + \alpha^n\}$$
$$> -\alpha\left\{\frac{1-\alpha^{n-1}}{1-\alpha}\right\}$$
$$> -\frac{\alpha}{(1-\alpha)}.$$

The monotonically decreasing sequence S_n of negative terms is bounded below by $-\alpha/(1-\alpha)$ and so S_n also attains a limit, k. The probability of extinction $f_n(0)$ is, therefore, a monotonically decreasing bounded sequence and attains a limit.

$$\underset{n\to\infty}{Lt}\, f_{n+1}(0) = e^{\beta k-\beta} f_1(l) = e^{\beta(k-1)} \cdot e^{(\beta+\alpha)(l-1)}$$
$$= e^{\beta(k+l-2)+\alpha(l-1)}. \tag{6}$$

It can be further shown that for $\alpha < 1$ the equation $x = e^{\alpha(x-1)}$ has only two roots $l_1 = 1$ and $l_2 < 1$. The probability of extinction, in this case, attains the value $e^{\beta(k-1)}$ as n becomes large, and for $\alpha \geqslant 1$ with zero limiting value of probability of extinction z_n diverges to infinity.

The Table I gives the total number, n, of generations required to attain the limit $l = \cdot 9991$ for different values of $\alpha < 1$ and the probability of extinction in the n-th generation. These probabilities are calculated with no selective advantage in the first generation and also for the case, with 1% initial advantage. In the subsequent generations, the selective advantage is determined by values of α and β. Since, the number n leading s_n to a limit l is determined independently, of the probability of extinction, the probability of extinction in all generations from n onwards remains stationary.

TABLE I

α	n	With no initial advantage	With 1% initial advantage	Difference
0	1	·3679	·3641	·0038
·1	3	·3693	·3654	·0039
·2	5	·3751	·3703	·0048
·3	6	·3844	·3791	·0053
·4	8	·3991	·3931	·0060
·5	10	·4186	·4116	·0070
·6	13	·4401	·4392	·0089
·7	17	·4860	·4748	·0112
·8	26	·5560	·5403	·0157
·9	50	·6661	·6397	·0264

ESTIMATION OF α AND β

With no selective advantage or disadvantage the first generation mean is 1, and therefore subject to the condition $\alpha + \beta = 1$, the maximum likelihood estimates of the two parameters of the process can be obtained. Let the growth of the process be observed up to n-th generation. If $x_0, x_1, x_2, \cdots x_n$ be the number of units in the respective generations

$$P\left[z_{n+1} = x_{n+1} \mid z_n = x_n\right] = \frac{e^{-(\alpha x_n+\beta)} \cdot (\alpha x_n + \beta)^{x_{n+1}}}{x_{n+1}\,!}$$

and

$$L = P(x_0, x_1, x_2, \cdots x_n) = P(x_n \mid x_{n-1}) P(x_{n-1} \mid x_{n-2}) \cdots P(x_2/x_1) P(x_1).$$

$$\text{Log } L = -\sum_{r=1}^{n} (\alpha x_{r-1} + \beta) + \sum_{r=1}^{n} x_r \log (\alpha x_{r-1} + \beta) + \text{const.}$$

Differentiating the likelihood function and equating to zero, we have

$$\left[\sum_{r=1}^{n} \frac{x_r x_{r-1}}{(\alpha x_{r-1} + \beta)} - X_n\right] d\alpha + \left[\sum_{r=1}^{n} \frac{x_r}{\alpha x_{r-1} + \beta} - n\right] d\beta = 0 \tag{7}$$

where

$$X_n = x_0 + x_1 + x_2 + \cdots + x_n.$$

This, together with the equation $d\alpha + d\beta = 0$ gives

$$\sum_{r=1}^{n} \frac{x_r x_{r-1}}{\alpha x_{r-1} + \beta} - X_n = \sum_{r=1}^{n} \frac{x_r}{\alpha x_{r-1} + \beta} - n.$$

Equation (7) together with the relation $\alpha + \beta = 1$ can be solved for α and β by the method of successive approximation to obtain the maximum likelihood estimates of α and β. As a particular case when $\alpha = 0$ the mean value of the number of units in every generation is fixed at β and its estimate is given by $\hat{\beta} = (X_n/n)$.

COVARIANCE FUNCTION

We consider the covariance function of the process $z_m, z_{m+1} \cdots z_n$, after a sufficiently large number of generations, m_0, such that

$$E(z_m) = \frac{\beta}{1-\alpha} + \epsilon$$

where ϵ is a small positive constant.

For, $m > m_0$, neglecting terms in ϵ

$$E\left(z_m - \frac{\beta}{1-\alpha}\right)\left(z_n - \frac{\beta}{1-\alpha}\right)$$

$$= \sum_{r=0}^{\infty} p_m(r) E\left[\left(z_n - \frac{\beta}{1-\alpha}\right)\left(r - \frac{\beta}{1-\alpha}\right)\Big|_{z_m = r}\right]$$

$$= \sum_{r=0}^{\infty} p_m(r) \left(r - \frac{\beta}{1-\alpha}\right) \left(r\alpha^{n-m} - \frac{\beta\alpha^{n-m}}{1-\alpha}\right)$$

$$= \sum_{r=0}^{\infty} p_m(r) \left(r - \frac{\beta}{1-\alpha}\right)^2 \alpha^{n-m}$$

$$= \alpha^{n-m} E\left(z_m - \frac{\beta}{1-\alpha}\right)^2$$

$$= \frac{\alpha^{n-m}\beta}{(1-\alpha)(1-\alpha^2)}.$$

The covariance function, therefore, takes the form

$$E\left(z_m - \frac{\beta}{1-\alpha}\right)\left(z_{m+p} - \frac{\beta}{1-\alpha}\right) = \frac{\beta\alpha^p}{(1-\alpha)(1-\alpha^2)},$$

which shows that the Morkovian process is nearly stationary, after a sufficiently large number of generations.

Some Other Forms of $\phi(j, r)$

In the above, Poisson form for $\phi(j, r)$ is adopted as it satisfies the required conditions. We shall consider some alternative forms for $\phi(j, r)$ which satisfy these.

Let

$$\phi(j, r) = e^{-\beta/j} \sum_{r_1+r_2=r}^{\infty} \frac{1}{r!}\left(\frac{\beta}{j}\right)^{r_1}\left(\frac{\alpha}{1+\alpha}\right)^{r_2}$$

which is the coefficient s^r in $[1 - \alpha(s-1)]^{-1} e^{\beta(s-1)/j}$. This leads to the recursion relation $f_{n+1}(s) = e^{\beta(s-1)} f_n [1 - \alpha(s-1)]^{-1}$. The mean number of units in the n-th generation in this case and in the other two cases to be considered is

$$m_n = \beta\left(\frac{1-\alpha^{n-1}}{1-\alpha}\right) + \alpha^{n-1} m_1,$$

the same as in Poisson form. The variance however, differs in each case. In the present case, it is

$$\beta^2\left(\frac{1-\alpha^{2n-2}}{1-\alpha}\right)+\frac{2\alpha\beta^2}{(1-\alpha)}\left(\frac{1-\alpha^{2n-4}}{1-\alpha^2}\right)+2\alpha^{n-3}\beta m_1$$

$$+\,2\beta\alpha^{n-1}\left(m_1-\frac{\beta}{1-\alpha}\right)\left(\frac{1-\alpha^{n-2}}{1-\alpha}\right)$$

$$+\,\alpha^{2n-2}\,(2\alpha^2+\beta^2+2\alpha\beta)-m_n^{\,2}+m_n.$$

The probability of extinction in the $(n+1)$-th generation is given by $f_{n+1}(0)=e^{\beta s_n-\beta}f_1(s_n)$ where

$$s_n=\frac{1}{1-\alpha\,(s_{n-1}-1)}\quad\text{and}\quad S_n=\sum_{i=1}^{n}(s_i-1).$$

Now, $s_n>s_{n-1}>s_{n-3}\cdots>s_3$ if $s_2>s_1$ and as

$$s_2-s_1=\frac{\alpha}{1+\alpha+\alpha^2}>0,$$ s_n is an increasing monotonic sequence bounded by 1 and for $0<\alpha<1$ attains its bound. As before since $(s_i-1)>-\alpha^i$, S_n is monotinically decreasing sequence of negative terms bounded by $-\alpha/(1-\alpha)$, it tends to a limit k_1. The limiting values of the probabilities of extinction for different values of α are given by $e^{\beta(k_1-1)}$.

Similarly,

$$\phi(j,r)=(1+\beta)^{1/j}\,e^{-\alpha}\sum_{r_1+r_2=r}^{\infty}\frac{\alpha^{r_2}}{r_2!}\left(\frac{\beta}{1+\beta}\right)^{r_1}\Gamma\left(\frac{1}{j}+r_1\right)$$

which is the coefficient of s^r in $[1-\beta(s-1)]^{-1/j}\,e^{\alpha(s-1)}$

gives the recursion relation

$$f_{n+1}(s)=[1-\beta(s-1)]^{-1}f_n\,[e^{\alpha(s-1)}]$$

and the variance of z_n as

$$2\beta^2\left(\frac{1-\alpha^{2n-2}}{1-\alpha^2}\right)+\frac{\alpha\beta\,(\alpha+2\beta)}{(1-\alpha)}\left(\frac{1-\alpha^{2n-4}}{1-\alpha^2}\right)+\alpha^{2n-3}\,m_1\,(\alpha+2\beta)$$

$$\alpha^{n-2}\left(m_1-\frac{\beta}{1-\alpha}\right)(\alpha^2+2\alpha\beta)\left(\frac{1-\alpha^{n-2}}{1-\alpha}\right)$$

$$+\,\alpha^{2n-2}\,(\alpha^2+2\beta^2+2\alpha\beta)-m_n^{\,2}+m_n.$$

The probability of extinction in this case for $(n+1)$-th generation is given by

$f_{n+1}(0) = (1+\beta)^{-1} f_1(s_{n-1}) \prod_{i=1}^{n} [1-\beta(s_i - 1)]^{-1}$. After a finite number of generations $|\beta(s_i - 1)| < 1$, therefore, the infinite product $\prod_{i=1}^{n} [1-\beta(s_i-1)]^{-1}$ as n increases is convergent as $\sum_{i=1}^{n}(s_i - 1)$ is convergent, and the product tends to a limit k_2. The limiting value of the probability of extinction as n tends to infinity is $k_2(1+\beta)^{-1}$.

Finally, the form

$$\phi(j, r) = (1+\alpha)^{-1}(1+\beta)^{-1/j} \cdot \sum_{r_1+r_2=r}^{\infty} \left(\frac{\beta}{1+\beta}\right)^{r_1} \left(\frac{\alpha}{1+\alpha}\right)^{r_2} \Gamma\left(\frac{1}{j} + r_2\right)$$

the coefficient of s^r in $[1-\beta(2-1)]^{-1/j}[1-\alpha(s-1)]$ leads to the recursion relation

$$f_{n+1}(s) = [1-\beta(s-1)]^{-1} f_n[1-\alpha(s-1)]$$

and the variance of Z_n, as

$$2\beta^2\left(\frac{1-\alpha^{2n-2}}{1-\alpha^2}\right) + \frac{2\alpha\beta^2}{(1-\alpha)}\left(\frac{1-\alpha^{2n-4}}{1-\alpha^2}\right)$$
$$+2\beta\alpha^{n-1}\left(m_1 - \frac{\beta}{1-\alpha}\right)\left(\frac{1-\alpha^{n-2}}{1-\alpha}\right)$$
$$+ 2\beta m_1 \alpha^{2n-3} + 2\alpha^{2n-2}(\alpha^2+\beta^2+\alpha\beta) - m_n^2 + m_n.$$

The probability of extinction in the $(n+1)$-th generation is

$$(1+\beta)^{-1} \prod_{i=1}^{n} [1-\beta(s_i-1)]^{-1} [(1-\alpha(s_n-1)].$$

As n tends to infinity this attains a limiting value $k_2^{-1}(1+\beta)^{-1}$ where k_3 is the limit of the product $\prod_{i=1}^{n} [1-\beta(s_i-1)]^{-1}$.

Table II gives the limiting values of the mean, variance and the probabilities of extinction in each of the four cases.

TABLE II

$\phi(j, r)$	Mean	Variance	Probability of extinction
1. Poisson	$\frac{\beta}{1-\alpha}$	$\frac{\beta}{(1-\alpha)(1-\alpha^2)}$	$e^{\beta(k-1)}$
2. Poisson Binomial	$\frac{\beta}{1-\alpha}$	$\frac{\beta}{(1-\alpha)}$	$e^{\beta(k_1-1)}$
3. Binomial-Exponential	$\frac{\beta}{1-\alpha}$	$\frac{\beta(1-\alpha\beta+\beta)}{(1-\alpha)(1-\alpha^2)}$	$k_2^{-1}(1+\beta)^{-1}$
4. Binomial-Binomial	$\frac{\beta}{1-\alpha}$	$\frac{\beta(1+\alpha+\beta)}{(1-\alpha^2)}$	$k_3^{-1}(1+\beta)^{-1}$

The expression for covariance function for the process in each of the four cases after a sufficiently large number of generations is $\alpha^{n-m}\,\text{Var.}(Z_m)$, $m > m_0$.

ILLUSTRATION

The modified branching process considered here occurs in the study of population growth of unicellular organisms under controlled conditions. In the branching process evolving, while finding the probability of survival of mutant gene, the probability of producing r offsprings is assumed to be independent of the population size. Consider a single mutant in a population consisting of AA individuals where one of A genes mutates to a new allele a. This single mutant gives offsprings in the population of AA individuals. Let q be the number of offsprings produced by $AA \times Aa$ matings. The probability that the new allele a will be absent among q offsprings is $\frac{1}{2}$. For obtaining total probability that a will be lost in n generations, we make an assumption that the mean number of offsprings per mating in the first generation is $q = 2$, and the number of Aa offsprings are distributed according to Poisson series, which implies that the average number of mutant individuals in the first generation is unity. Further $p_0(1)$, $p_1(1)$, $p_1(2)$, $\cdots$ form the respective terms in the series $e^{-1}\{1, 1, 1/2!, 1/3!, \cdots\}$, *i.e.*, $f(s) = e^{s-1}$ where $f(s)$ is the generating function of the number of mutants in the first generation. If it be further assumed that in the subsequent generations, the probabilities of producing $r\,(r > r_0)$ offsprings are inversely dependent upon the size of the population of the mutant individuals in that generation, this process reduces to the modified branching process with the condition that $\alpha + \beta = 1$. This shows that $f_n'(1) = 1$ for all n and

$$\text{Var.}\,(Z_n) = (1 - \alpha^2)\left\{\frac{1 - \alpha^{2n-2}}{1 - \alpha^2}\right\} + \alpha\,(2 - \alpha)\left\{\frac{1 - \alpha^{2n-4}}{1 - \alpha^2}\right\}$$

$$+ \alpha^{2n-3}\,(2 - \alpha) + \alpha^{n-2}.$$

The constant value of the mean of the n-th generation shows that there is neither advantage nor disadvantage throughout the process though the probability of producing offsprings is inversely dependent on the number of mutants. For different value of α the probabilities of extinction would be same as given in Table I.

ACKNOWLEDGEMENT

I am indebted to Dr. M. N. Ghosh, Indian Council of Agricultural Research, for his guidance and valuable suggestions during the preparation of this paper. I am also thankful to the referee for the valuable improvements suggested by him.

REFERENCES

1. Harris, T. E. .. *Annals of Mathematical Statistics*, 1948, **19**.
2. ———— .. *Second Berkeley Symposium on Mathematical Statistics and Probability*, 1951.
3. Fisher, R. A. .. *Genetical Theory of Natural Selection*, 1958.
4. Li, C. C. .. *Population Genetics*, 1955.