TWO PROBLEMS IN THE THEORY OF STOCHASTIC
BRANCHING PROCESSES

by

JOHN WILLIAM BISHIR

Institute of Statistics
Mimeograph Series No. 286
April, 1961
BIOGRAPHY

The author was born in Joplin, Missouri, September 23, 1933 and attended public schools in that state. College work includes study at Joplin (Mo.) Junior College, 1951-53, the University of Missouri, 1953-55, where he received an A.B. in mathematics, and the State University of Iowa, 1955-57, where an M.S. in mathematics was received. He was statistician for the Institute of Agricultural Medicine from February to August, 1957, and joined the staff of North Carolina State College in September, 1957, as instructor in mathematics, a position held since that time.

The author was married in 1954 to Mary Katherine Tims and has three children.
ACKNOWLEDGMENTS

Many persons have, in greater or lesser degree, assisted me in this undertaking. Some are perhaps unaware that a contribution was made and in some instances I, too, may have overlooked some special bit of help. To all these, known and unknown, go my thanks.

Special gratitude must be extended to:

Dr. Walter Smith of the University of North Carolina Department of Mathematical Statistics for suggestion of the second problem and other helpful suggestions along the way.

Dr. John W. Cell and the members of the Department of Mathematics of North Carolina State College, especially Drs. Jack Levine, Walter L. Harrington and H. M. Nahikian, for their interest and encouragement.

Drs. R. L. Anderson and R. J. Monroe as members of the advisory committee.

Dr. D. R. Shreve and Henry Hamilton of the North Carolina State College computation center for their prompt assistance with the computations.

Mary Katherine, Bobby, Carol and Patricia for putting up with the whole thing.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>2.0</td>
<td>LIST OF FIGURES</td>
<td>vi</td>
</tr>
<tr>
<td>3.0</td>
<td>THE CLASSICAL BRANCHING PROCESS</td>
<td>1</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>3.2</td>
<td>Review of Literature</td>
<td>3</td>
</tr>
<tr>
<td>3.3</td>
<td>Generating Functions and the Solution of the Classical Problem</td>
<td>6</td>
</tr>
<tr>
<td>3.4</td>
<td>The Branching Process as a Markov Chain</td>
<td>12</td>
</tr>
<tr>
<td>4.0</td>
<td>THE MAXIMUM POPULATION SIZE IN A BRANCHING PROCESS</td>
<td>14</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>14</td>
</tr>
<tr>
<td>4.2</td>
<td>Distribution of Maximum Population Size</td>
<td>15</td>
</tr>
<tr>
<td>4.3</td>
<td>Bounded Population Size</td>
<td>23</td>
</tr>
<tr>
<td>5.0</td>
<td>DEPENDENCE IN A BRANCHING PROCESS</td>
<td>36</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>36</td>
</tr>
<tr>
<td>5.1.1</td>
<td>General Comments</td>
<td>36</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Description of the Problem</td>
<td>36</td>
</tr>
<tr>
<td>5.2</td>
<td>Variations of the Original Problem</td>
<td>40</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Lower Bounds for $P_{\infty}$</td>
<td>40</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Upper Bounds for $P_{\infty}$</td>
<td>46</td>
</tr>
<tr>
<td>6.0</td>
<td>SUMMARY AND CONCLUSIONS</td>
<td>64</td>
</tr>
<tr>
<td>6.1</td>
<td>The Problem of Maximum Population Size</td>
<td>64</td>
</tr>
<tr>
<td>6.2</td>
<td>The Grid Problem</td>
<td>65</td>
</tr>
<tr>
<td>6.3</td>
<td>Recommendations</td>
<td>66</td>
</tr>
<tr>
<td>7.0</td>
<td>LIST OF REFERENCES</td>
<td>68</td>
</tr>
<tr>
<td>8.0</td>
<td>APPENDIX A</td>
<td>71</td>
</tr>
<tr>
<td>9.0</td>
<td>APPENDIX B</td>
<td>73</td>
</tr>
<tr>
<td>10.0</td>
<td>APPENDIX C</td>
<td>75</td>
</tr>
</tbody>
</table>
1.0 LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Some values of $p_1(k)$ in certain geometric distributions</td>
<td>19</td>
</tr>
<tr>
<td>2.</td>
<td>Some values of $p_5(k)$ in certain geometric distributions</td>
<td>20</td>
</tr>
<tr>
<td>3.</td>
<td>Values of $E_1^<em>(T)$ and $V_1^</em>(T)$ for the geometric distribution with $p = 1/2$, using $k = 10$</td>
<td>27</td>
</tr>
<tr>
<td>4.</td>
<td>Values of $E_1^<em>(G)$ and $V_1^</em>(G)$ for the geometric distribution with $p = 1/2$, using $k = 10$</td>
<td>33</td>
</tr>
<tr>
<td>5.</td>
<td>Values of $p_{nj}$ for $n = 0, 1, 2, 3$</td>
<td>41</td>
</tr>
<tr>
<td>6.</td>
<td>Partial transition matrix</td>
<td>48</td>
</tr>
</tbody>
</table>
2.0 LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Possible shapes of the curve $y = f(s)$</td>
<td>9</td>
</tr>
<tr>
<td>2.</td>
<td>Behavior of $t_n$, the probability of extinction by generation $n$</td>
<td>10</td>
</tr>
<tr>
<td>3.</td>
<td>The infinite grid of points</td>
<td>37</td>
</tr>
<tr>
<td>4.</td>
<td>Graph of $y = 1 - s_0$ when $f(s) = q + ps^2$</td>
<td>39</td>
</tr>
<tr>
<td>5.</td>
<td>Graph of the curve $b_1(p)$</td>
<td>53</td>
</tr>
<tr>
<td>6.</td>
<td>The best upper bound for $P_\infty$</td>
<td>59</td>
</tr>
<tr>
<td>7.</td>
<td>Truncated grid of $n$ rows</td>
<td>60</td>
</tr>
<tr>
<td>8.</td>
<td>Figure 7 redrawn</td>
<td>61</td>
</tr>
</tbody>
</table>
3.0 THE CLASSICAL BRANCHING PROCESS

3.1 Introduction

The stochastic processes known as branching, or multiplicative, processes have drawn the attention of writers in many fields since their introduction by Galton and Watson in 1874 (12). This was apparently the first non-deterministic approach to the problem of population growth. Their work was concerned with the extinction of family surnames. From it came the simplest and probably best known mathematical model of a branching process and the one with which we shall be concerned. Since the mathematical model admits of various physical interpretations we shall first describe it in general and then indicate particular applications.

We consider particles, or objects, which are able to produce other particles of like kind. Every particle has probability $p_k$ ($k = 0, 1, 2, \ldots$) of producing $k$ new particles. The original, or zero, generation is composed of $Z_0$ (usually taken to be one) particles. The direct descendants of the nth generation form the $(n + 1)$st generation. The particles of each generation act independently of each other.

This is the original Galton-Watson model. In their work the role of the particles is played by the male descendants in a particular family line. The number $p_k$ then represents the probability for a newborn boy to become the progenitor of exactly $k$ boys. The classical
problem is concerned with the probabilities of finding exactly \( k \) carriers of the family name in the \( n \)th generation and of ultimate extinction of the line.

A second application may be made to nuclear chain reactions. Here the particles are neutrons which are subject to chance hits by other particles. We take the probability of a particle hitting another (and thus creating \( m \) new particles) to be \( p \), and \( q (= 1 - p) \) to be the probability it remains inactive by being absorbed or removed in some way. Extinction here is equivalent to inactivity while explosion is implied by an indefinitely increasing number of particles.

In spite of the simplicity of the model there are interesting mathematical problems associated with it, many of them still unsolved. One of these concerns the population size in the successive generations. In particular, we may desire the probability that the population size never exceeds a fixed value, \( k \). A solution to this problem, together with related results, is given in the next chapter.

A second problem concerns the fundamental assumption of independence of the particles in a given generation. This is obviously violated in the nuclear reaction as the population size, or number of neutrons, becomes large. Thus we may wish to know what results are obtained under various assumptions of dependence and how these results compare with the classical situation. We shall discuss this
problem for a very simple model in which each particle produces either zero or two progeny.

In the remainder of the present chapter we shall indicate some of the literature on branching processes and related topics and shall discuss those of the classical results which will be needed for later reference and comparison.

3.2 Review of Literature

The first account of what we now call the classical branching process appeared in 1874 in a paper by Francis Galton and the Reverend H. W. Watson (12). The problem apparently was proposed by Galton and the mathematical results are due to Watson. His calculations provide all that is necessary for the solution as he discovered the generating function relation (3.1) governing the distribution of \( Z_n \), the population size in the nth generation. He also recognized the now familiar fact that extinction is inevitable in a population for which average birth and death rates exactly balance. This paper was also reprinted later as Appendix F of Galton's book (11).

Their results drew little attention for a number of years and more than fifty years after the original work appeared, A. K. Erlang (4) proposed the same problem. This time a more complete solution was given by J. F. Steffensen (20, 21), at first unaware of the earlier work. He showed that if \( f(s) \) is the generating function of the probabilities \( p_k \),
then there is always a positive chance that the male line will ultimately become extinct and that this is given by the smallest positive root of the equation $f(s) = s$. Further, extinction is certain to occur if and only if the expected number of male children per father is less than or equal to one. A development of these results is given in the next section.

In 1931, A. J. Lotka (19) applied Steffensen's results to the 1920 data for white males in the United States and found the value for the chance of ultimate extinction of a male line to be approximately 0.88. This, of course, assumes a single "ancestor."

The classical results were generalized by Everett and Ulam (5) to the case of many types of particles, each particle being able to produce particles of all types according to given probability laws. As an application, consider the nuclear reaction. Here it is apparent that neutrons near the center of the material have a greater likelihood of striking other neutrons than those near the outside of the material. For a first approximation, then, we may divide a sphere of active material into two parts, an inner sphere and an outer shell. Neutrons in the two parts will be assumed to generate other neutrons of both kinds according to (possibly different) probability distributions.

In the last twenty years attention has shifted from the simple process we have described to more general ones in which the reproductive properties of the particles may change, varying with time and
age of the particles. Feller (7) was apparently the first to discuss
branching processes in which a continuous time parameter is involved
and since 1939 there has been an extensive literature. The papers of
Harris (15) and Bellman and Harris (2) may be cited in particular.

The 1950 paper of Harris (16) provides the best review available
of the classical branching process. In particular he includes a sum-
mary of Russian efforts along these lines. Mention is also made of
some unsolved problems in the field.

For a comprehensive view of stochastic processes in relation to
study of population growth the 1948 paper of Kendall (18) is especially
recommended. Also included here is an interesting account of the
history of the subject beginning with the work of Galton and Watson in
1874.

In the next two sections we shall develop those classical results
to which we shall need to refer in the sequel. The first deals with the
solution of the classical problem by the usual generating function
approach. In the second we discuss the relation between branching
processes and Markov chains having a denumerable number of states.
This relation is conceptually quite convenient for the presentation of
later results.
3.3 Generating Functions and the Solution of the Classical Problem

In this section we shall develop some of the classical results. In order to facilitate the discussion we first introduce the concept of a probability generating function. Theorems will be stated and used without proof as they may be found in any book on probability in which generating functions are discussed (for example, Feller (8)). We then proceed to a mathematical description of the simple branching process. Reference may be made to Feller's book (8) or that of Bharucha-Reid (3) for further details. A generalization to more than one type of particle may be found in the paper of Everett and Ulam (5).

We begin with the definition of a generating function.

Definition 3.1. Let \( p_0, p_1, p_2, \ldots \) be a sequence of numbers. Then the function

\[
f(s) = \sum_{k=0}^{\infty} p_k s^k
\]

is called the generating function of the sequence.

In our work the \( p_k \) will represent probabilities of mutually exclusive events. Thus \( f(1) = 1 \), and \( f(s) \) must converge for \( |s| \leq 1 \).

Definition 3.2. Let \( \{a_k\} \) and \( \{b_k\} \) be any two number sequences. The sequence \( \{c_r\} \) defined by

\[
c_r = \sum_{k=0}^{r} a_k b_{r-k}
\]

is called the convolution of \( \{a_k\} \) and \( \{b_k\} \).
Theorem 3.1. If \( \{a_k\} \) and \( \{b_k\} \) are sequences with generating functions \( A(s) \) and \( B(s) \), then the generating function of their convolution \( \{c_r\} \) is
\[
C(s) = A(s) B(s).
\]
If \( X \) and \( Y \) are non-negative, integral valued, independent random variables with generating functions \( A(s) \) and \( B(s) \), then the generating function of the sum \( X + Y \) is \( A(s) B(s) \).

This result may be extended by induction to any number of mutually independent random variables. In particular, we find that if \( S_n = X_1 + \ldots + X_n \) where the \( X_j \) have the common generating function \( A(s) \), then the generating function of \( S_n \) is \( A^n(s) \).

Now suppose \( \{X_n\} \) is a sequence of mutually independent random variables with common distribution \( \text{Pr}(X_n = k) = p_k \) and generating function \( f(s) = \sum_{k = 0}^{\infty} p_k s^k \). Let \( S_N = X_1 + \ldots + X_N \) where the number, \( N \), of terms is itself a random variable independent of the \( X_n \).

Suppose \( \text{Pr}(N = r) = g_r \) is the distribution of \( N \) and the associated generating function is \( g(s) = \sum_{r = 0}^{\infty} g_r s^r \). Then we have the following theorem concerning the sum \( S_N \):

**Theorem 3.2.** The generating function of \( S_N = X_1 + \ldots + X_N \) is the compound function \( g(f(s)) \).

We are now able to give a mathematical description of a branching process. We write \( Z_n \) for the size of the \( n \)th generation.

Assume that each particle, independently of the other particles, has
probability $p_k$ of producing exactly $k$ new particles; ($k = 0, 1, 2, \ldots$).

Supposing $Z_0 = 1$, which means we begin with one particle, then $Z_1$ has probability distribution $p_k = \Pr(Z_1 = k)$ and generating function $f(s) = \sum_{k=0}^{\infty} p_k s^k$. To avoid trivial cases we assume that $p_0 \neq 0$ and $p_0 + p_1 < 1$. Now the second generation consists of the direct descendants of the $Z_1$ members of the first generation. Thus $Z_2$ is the sum of $Z_1$ mutually independent random variables each having generating function $f(s)$. From Theorem 3.2 the generating function of $Z_2$ must be $f_2(s) = f(f(s))$. Continuing inductively, we see that in general the generating function $f_{n+1}(s)$ of $Z_{n+1}$ is given recursively by

$$f_1(s) = f(s),$$
$$f_{n+1}(s) = f(f_n(s)). \quad (3.1)$$

Thus we have the generating function relations governing the size of the $n$th generation. From these it may be shown (1) that $E(Z_n) = m^n$ where $m = E(Z_1)$. If $m > 1$, we see that the population size tends to increase exponentially.

Since by assumption $0 < p_0 < 1$ and all $p_k$ are non-negative, it is evident that the generating function $f(s) = \sum_{k=0}^{\infty} p_k s^k$, as well as its derivative, is an increasing function in the interval $(0, 1)$. This means that $f(a) < f(b)$ when $a < b$. The curve is thus convex and can intersect the line $y = s$ in at most two points, one of which must be $(1, 1)$. The two possibilities for the curve are shown in Figure 1.
Now let $t_n$ be the probability that the process terminates at or before the $n$th generation. This is the probability that there are no particles in the $n$th generation and hence is given by $f_n(0) = f[f_{n-1}(0)]$. Now $t_1 = f(0) = p_0 > 0$ and $t_2 = f(p_0) > f(0) = t_1$. By induction then, assuming $t_n > t_{n-1}$, we have

$$t_{n+1} = f(t_n) > f(t_{n-1}) = t_n$$

so the sequence is increasing. Moreover, it is bounded above by one and therefore must approach a limit, $s_0$, which is not greater than one. From (3.2) $s_0$ is obviously a solution of the equation $f(s) = s$.

Further, $s_0$ is the smallest positive root of this equation. For, let $s_1$ be any root greater than zero. Then $t_1 = f(0) < f(s_1) = s_1$. By induction $t_{n+1} = f(t_n) < f(s_1) = s_1$ for all $n$. Hence $s_0 \leq s_1$. 
Figure 2 below illustrates the behavior of the $t_n$. In most cases $t_n$ converges to $s_0$ rapidly, indicating that the process will usually terminate during the first few generations, if at all. Also, an argument similar to that employed above for $s = 0$ shows that $f_n(s)$ approaches $s_0$ for all $s < 1$. This means that the coefficients of $s$, $s^2$, $s^3$, ... in the expansion of $f_n(s)$ all approach zero so that after a few generations it is unlikely that we will find a moderate number of particles. That is, either the process terminates quickly or the population size becomes very large.

Figure 2. Behavior of $t_n$, the probability of extinction by generation $n$. 

$y = f(s)$

$y = s$

$f(0) = p_0$

$s_0$

$t_1$

$t_2$

$t_3$

$t_4$

$s$

1
In the interval \((0, 1)\), \(f'(s)\) takes on its maximum value at \(s = 1\). Thus, if \(f'(1) \leq 1\), \(f'(s) < 1\) for all \(s \neq 1\) in the interval. In this case we have, by the mean value theorem

\[
\frac{1 - f(s)}{1 - s} = f'(s_1) < 1
\]

where \(s_1\) is some number between \(s\) and one. This means \(f(s) > s\) for \(s < 1\) and the only intersection occurs at \((1, 1)\). On the other hand, if \(f'(1) > 1\) the curve \(y = f(s)\) must be below the line \(y = s\) for values close to one. Since \(p_0 = f(0) > 0\), the two curves must intersect at some value of \(s\) less than one. Since \(f'(1) = \sum_{k=0}^{\infty} kp_k = m\), the expected number of descendants from a single particle, we have the following theorem:

**Theorem 3.3.** If \(m\), the expected number of descendants from one particle, is less than or equal to one, then \(s_0 = 1\) and as \(n\) increases the probability tends to one that the process will terminate by the \(n\)th generation. If \(m\) is greater than one this probability is given uniquely by the smallest positive root, \(s_0 < 1\), of the equation \(f(s) = s\).

To illustrate the above results, suppose we have a process in which a single particle has probability \(q\) (\(0 < q < 1\)) of producing no new particles and probability \(p = 1 - q\) of producing two. Then \(f(s) = q + ps^2\). Solving the equation \(f(s) = s\) we find that \(s_0\) is the smaller of the
numbers 1, q/p. Hence \( s_0 = 1 \) if \( p \leq 1/2 \) and \( s_0 = q/p \) if \( p > 1/2 \). Here \( m = f'(1) = 2p \).

As a second example consider the geometric distribution for which \( p_k = qp^k \). Here \( f(s) = q/(1 - ps) \) which yields the same values for \( s_0 \) as in the previous example. However, in this case \( m = p/q \) so that \( s_0 = 1/m \) if \( m > 1 \).

All of the above discussion applies when we begin with one particle. However, extension to more than one particle is quite simple. If \( Z_0 = i \) the probability of ultimate extinction is \( s_0^i \) since the particles act independently. The other results follow as before.

3.4 The Branching Process as a Markov Chain

A branching process can be considered as a special case of a Markov chain with a denumerable number of states. To see this, simply choose as states of the process states \( E_j \) such that the process is said to be in state \( E_j \) at time (generation) \( n \) if \( Z_n = j \); i.e., if there are \( j \) particles in the \( n \)th generation. If \( Z_0 = i \), the process starts in state \( E_i \). From the previous section we see that the transition probability, \( p_{ij} \), of going from state \( E_i \) to state \( E_j \) in one step (generation) is given by the coefficient of \( s^j \) in the expansion of \( [f(s)]^i \), where \( f(s) \) is the generating function of progeny from a single particle. Thus the transition matrix \( P = (p_{ij}) \) is obtained by using in the \( i \)th row the successive coefficients in \( [f(s)]^i = \sum_{j=0}^{\infty} p_{ij} s^j \).
Example. For the geometric distribution \( p_k = qp^k \) (\( p + q = 1 \)) and
\[
f(s) = \frac{q}{1 - ps} = q \sum_{k=0}^{\infty} p^k s^k.
\]
It is easily verified by induction that for
\[
i \geq 1, \quad [f(s)]^i = \left[ \frac{q}{1 - ps} \right]^i = q^i \sum_{k=0}^{\infty} \binom{i+k-1}{i-1} p^k s^k.
\]
When
\[
i = 0, \quad [f(s)]^i = 1.
\]
Hence in this case, \( P_{0j} = 1 \) if \( j = 0 \), and zero otherwise, while for \( i > 0 \), \( p_{ij} \) is given by
\[
p_{ij} = \binom{i+j-1}{i-1} q^i p^j.
\]
4.0 THE MAXIMUM POPULATION SIZE IN A BRANCHING PROCESS

4.1 Introduction

In studies of population growth one is often interested in determining bounds, if any exist, for the population size. For example, consider a genetical situation in which each gene of a given organism has a chance of appearing in 0, 1, 2, \ldots direct descendants. If a mutation occurs we can use the branching process to describe the spread of the new gene. The original mutant plays the role of the zero-generation particle. In the initial stages the number of organisms having the mutant gene is small and any bounds on the population are due to chance. As the population size increases external limits may be imposed by the physical environment in which the organisms exist. The branching process scheme does not take these into account since reproductive properties are assumed not to change.

In this chapter two aspects of the problem of upper bounds on population size in a branching process are considered. First, the probability distribution of the maximum population size is obtained. The theory is illustrated using the geometric distribution and some related results are developed. Second, some properties of the process under the conditional hypothesis that the number of particles never exceeds a given fixed value are developed. It is shown that the theory of finite absorbing Markov chains applies here and results are obtained
from this correspondence. Other results which arise only in the branching process are then indicated.

4.2 Distribution of Maximum Population Size

Consider the usual branching process in which each particle has, independently of the other particles present, probabilities $p_k (k = 0, 1, 2, \ldots)$ of giving rise to $k$ like particles. We assume $p_0 \neq 0$ and $p_0 + p_1 < 1$. Let $f(s) = \sum_{k=0}^{\infty} p_k s^k$ be the associated generating function and let $Z_n$ be the size of the $n$th generation. It is convenient to think of the process as a Markov chain (see section 3.4).

The transition matrix $P = (p_{ij})$ is given by $p_{ij} = \Pr \left[ (Z_n = j) / (Z_{n-1} = i) \right]$ = the coefficient of $s^j$ in $\left[ f(s) \right]^i$. The process will be said to be in state $E_j$ at time (generation) $n$ if $Z_n = j$. In the following discussion we will be interested in certain parts of the matrix $P$. For this reason we introduce the following definitions of the matrices $P_k$, $R_k$ and $Q_k$:

$$
P_k = \begin{bmatrix}
p_{00} & p_{01} & p_{02} & \cdots & p_{0k} \\
p_{10} & p_{11} & p_{12} & \cdots & p_{1k} \\
p_{20} & p_{21} & p_{22} & \cdots & p_{2k} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
p_{k0} & p_{k1} & p_{k2} & \cdots & p_{kk}
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 & \cdots & 0 \\
p_0 & p_1 & p_2 & \cdots & p_k \\
p_{20} & p_{21} & p_{22} & \cdots & p_{2k} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
p_{k0} & p_{k1} & p_{k2} & \cdots & p_{kk}
\end{bmatrix}
$$

$$
\equiv \begin{bmatrix}
1 & 0 \\
R_k & Q_k
\end{bmatrix}
$$
The vector \( R_k \) is \( k \times 1 \) while the matrix \( Q_k \) is \( k \times k \). It is easily verified that \( p_{10} = p_0^i \) so that the vector \( R_k \) is given by
\[
R_k = \text{col} \left( p_0^0, p_0^1, \ldots, p_0^k \right).
\]
We see that we have an absorbing chain with a single absorbing state \( E_0 \).

Now consider the probability \( p_1(k) \) that the population size \( Z_n \) never exceeds the value \( k \) if the initial population size is \( Z_0 = i \). That is to say \( p_1(k) = \Pr \left[ Z_n \leq k, \text{ all } n / (Z_0 = i) \right] \). Obviously \( p_1(k) = 0 \) for \( i > k \) and \( p_1(k) = 1 \). For \( 1 \leq i \leq k \) we have
\[
p_1(k) = p_0^i + \sum_{j=1}^{k} p_{ij} p_j(k).
\]
Writing \( P(k) = \text{col} \left[ p_1(k), p_2(k), \ldots, p_k(k) \right] \) we can combine these equations into the single matrix equation
\[
P(k) = R_k + Q_k P(k).
\]
Solving for \( P(k) \) yields
\[
P(k) = (I - Q_k)^{-1} R_k
\]
as the required vector of probabilities.

That \( (I - Q_k)^{-1} \) exists follows from two theorems due essentially to Frobenius (9, 10) although he proved them only for matrices with positive entries. Extension to the case of non-negative entries can be found in the paper of Fan (6).

---

\( \upuparrows \) See the List of Notation for explanation of this and subsequent notation.
Theorem 4.1. If the elements of a square matrix $A$ are all real and non-negative, then its characteristic equation has a root $r$ which is real, positive and larger in absolute value than any other root.

Theorem 4.2. If the elements of an $n$-square matrix $A$ are real and non-negative, then its largest characteristic root lies between the largest and smallest of the row sums $r_i = \sum_{j=1}^{n} a_{ij}$ ($i=1, 2, \ldots, n$) of $A$.

Since the row sums of the transition matrix $P$ are all one, it follows that no row sum of $Q_k$ can exceed $1 - p_0^k$. Hence no root of $Q_k$ can be greater in absolute value than this number and, in particular, since $p_0 \neq 0$, no root of $Q_k$ has the value one. This implies that $I - Q_k$ can have no zero root and the existence of the inverse is guaranteed.

We shall henceforth write $N_k$ for $(I - Q_k)^{-1}$. Equation (4.3) then becomes

$$P(k) = N_k R_k$$

Writing $n_{ij}(k)$ for the $i, j$ term in $N_k$ we have

$$p_i(k) = \sum_{j=1}^{k} n_{ij}(k) p_0^j$$

(4.5)

From standard results (cf. section 3.3) we know that the probability of ultimate extinction of the process is, if $Z_0 = 1$, the smallest positive root, $s_0$, of the equation $f(s) = s$, where $f(s)$ is the probability generating function of the number of progeny from a single individual. If $Z_0 = i$, this probability is $s_0^i$. In each case this can be
interpreted as the probability that population size is always finite, which is to say that all \( Z_n < \infty \). Thus as \( k \) increases \( p_i(k) \) should approach \( s_o^i \) for each \( i \). The approach is from below since \( p_i(k) \) is a monotone increasing function of \( k \).

We know also that in most cases the process either dies out quickly or not at all, so the limiting value should be approached closely for small \( k \). In case \( s_o < 1 \), we can be more precise. Suppose \( Z_o = i \) and take \( k \geq i \). Using an \( i \) subscript to denote dependence on \( Z_o = i \), we have

\[
Pr_i (\text{some } Z_n \geq k+1 \text{ and the process dies out})
\]

\[
= Pr_i (\text{some } Z_n \geq k+1) Pr_i (\text{the process dies out} / \text{some } Z_n \geq k+1)
\]

\[
\leq Pr (\text{some } Z_n \geq k+1) s_o^{k+1} \leq s_o^k.
\]

But \( Pr_i (\text{some } Z_n \geq k+1 \text{ and the process dies out}) \) is also equal to

\[
Pr_i (\text{the process dies out}) - Pr_i (\text{all } Z_n \leq k)
\]

\[
= s_o^i - p_i(k).
\]

Thus for all \( i \leq k \), \( s_o^i - p_i(k) \leq s_o^{k+1} \) or, equivalently,

\[
p_i(k) \geq s_o^i - s_o^{k+1}.
\]

As an example consider the geometric distribution \( (p_k = q p^k) \) with \( p = 2/3 \). Here \( s_o = 1/2 \) and if \( Z_o = 1 \) we have \( p_1(k) \geq s_o - s_o^{k+1} \).

For \( k = 9 \) this becomes \( p_1(9) \geq 1/2 - (1/2)^{10} > 0.499 \). The actual value of \( p_1(9) \) correct to four decimal places is 0.4999.
In Tables 1 and 2 are listed values of \( p_1(k) \) and \( p_5(k) \) for the geometric distributions with \( p \) values \( 1/3 \), \( 1/2 \) and \( 2/3 \), respectively. Various values of \( k \) between one and twenty are used. The corresponding values of \( s_o^i \) are given for comparison. Convergence of \( p_i(k) \) to \( s_o^i \) is less rapid for \( p = 1/2 \) than for the other values. This is to be expected since \( p = 1/2 \) is the critical value for the distribution in the sense that this is the largest value of \( p \) for which extinction is certain to occur. That is to say, \( s_o = 1 \) for \( p \leq 1/2 \) but \( s_o = q/p \) for \( p > 1/2 \).

Table 1. Some values of \( p_1(k) \) in certain geometric distributions

<table>
<thead>
<tr>
<th>( k )</th>
<th>( p = 1/3 )</th>
<th>( p = 1/2 )</th>
<th>( p = 2/3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8571</td>
<td>0.6667</td>
<td>0.4286</td>
</tr>
<tr>
<td>2</td>
<td>0.9379</td>
<td>0.7568</td>
<td>0.4690</td>
</tr>
<tr>
<td>3</td>
<td>0.9728</td>
<td>0.8105</td>
<td>0.4864</td>
</tr>
<tr>
<td>5</td>
<td>0.9946</td>
<td>0.8691</td>
<td>0.4973</td>
</tr>
<tr>
<td>8</td>
<td>0.9995</td>
<td>0.9100</td>
<td>0.4997</td>
</tr>
<tr>
<td>12</td>
<td>1.0000</td>
<td>0.9359</td>
<td>0.5000</td>
</tr>
<tr>
<td>20</td>
<td>1.0000</td>
<td>0.9591</td>
<td>0.5000</td>
</tr>
<tr>
<td>( \infty ) (( s_o ))</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.5000</td>
</tr>
</tbody>
</table>
Table 2. Some values of $p_5(k)$ in certain geometric distributions

<table>
<thead>
<tr>
<th>k</th>
<th>$\frac{1}{3}$</th>
<th>$\frac{1}{2}$</th>
<th>$\frac{2}{3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5$^a$</td>
<td>0.8990</td>
<td>0.3835</td>
<td>0.0281</td>
</tr>
<tr>
<td>6</td>
<td>0.9460</td>
<td>0.4482</td>
<td>0.0296</td>
</tr>
<tr>
<td>8</td>
<td>0.9860</td>
<td>0.5516</td>
<td>0.0308</td>
</tr>
<tr>
<td>10</td>
<td>0.9967</td>
<td>0.6258</td>
<td>0.0311</td>
</tr>
<tr>
<td>12</td>
<td>0.9992</td>
<td>0.6795</td>
<td>0.0312</td>
</tr>
<tr>
<td>20</td>
<td>1.0000</td>
<td>0.7953</td>
<td>0.03125</td>
</tr>
<tr>
<td>$\infty$ ($s_0^5$)</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.03125</td>
</tr>
</tbody>
</table>

$^a p_5(k) = 0$ for $k < 5$.

If $p = 1/3$ it is virtually certain that beginning with one initial particle the population size will never exceed eight. In fact it is quite likely that the initial population is the largest in size. For $p = 1/2$ this is less certain but still to be expected in two-thirds of the cases. With either of these values of $p$ the population will eventually become extinct. When $p = 2/3$ the probability is 1/2 that the population becomes infinite. If population size stays finite then it remains small in most cases. The probability is 0.4997 that there are never more than eight particles in any generation.
When the initial generation contains five members these probabilities are decreased. But even here for \( k \) as small as twelve the values of \( p_5(k) \) are quite close to the limiting values for both \( p = 1/3 \) and \( p = 2/3 \). If \( p = 1/2 \) the population size will exceed twenty in about one of every five cases.

These results are useful in giving an indication of just how the population behaves in the early stages. It is apparent that a branching process will usually either die out after having produced only a few particles or will become extremely large.

From the values of \( p_i(k) \) we may obtain the probability distribution of the maximum population size. For, if \( q_i(k) \) is the probability that the maximum population size is \( k \) if we begin with \( i \) particles, then

\[
q_i(k) = p_i(k) - p_i(k-1)
\]

(4.7)

This, however, is of less value than \( p_i(k) \) and will not be pursued further here. We merely note that in the next section we arrive at another representation (4.34) of the probability \( q_i(k) \) which provides this value directly and thus presents another method of calculating \( p_i(k) \).

There is an interesting probabilistic interpretation for the term \( n_{ij}(k) \ p_j^j \) appearing in equation (4.5), \( p_i(k) = \sum_{j=1}^{k} n_{ij}(k) \ p_j^j \). In fact, this is the probability that in the generation just prior to extinction there were \( j \) particles and the population size never exceeded the value
k, given that \( Z_0 = i \). For, let \( s_{ij}(k) \) be the probability in question.

Then

\[
s_{ij}(k) = \sum_{t=1}^{k} p_{it} s_{tj}(k) \quad \text{if } j \neq i
\]

\[
= p_0^i + \sum_{t=1}^{k} p_{it} s_{tj}(k) \quad \text{if } j = i.
\]

Hence if \( S_k = (s_{ij}(k)) \) and \( R_k = \text{diag}(p_0, p_0^2, \ldots, p_0^k) \), we have

\[
S_k = R_k + Q_k S_k \quad \text{or, solving for } S_k,
\]

\[
S_k = N_k R_k.
\] (4.8)

as asserted.

Now let us denote the last non-zero \( Z \) by \( Z_L \), and write

\[
E \left[ Z_L / (Z_0 = i \text{ and all } Z_n \leq k) \right] = E_i^*(Z_L) \text{ and}
\]

\[
E^*(Z_L) = \text{col} \left[ E_1^*(Z_L), \ldots, E_k^*(Z_L) \right].
\]

Then

\[
E_i^*(Z_L) = \sum_{j=1}^{k} j \Pr_i(Z_L = j / \text{all } Z_n \leq k)
\]

\[
= \frac{\sum_{j=1}^{k} j \Pr_i(Z_L = j \text{ and all } Z_n \leq k)}{\Pr_i(\text{all } Z_n \leq k)}
\]

\[
= \frac{\sum_{j=1}^{k} j n_{ij}(k) p_0^j}{p_i(k)}
\] (4.9)

the \( i \) subscript being used to indicate dependence on \( Z_0 = i \).
But from the original probability distribution we have
\[ j p_{0}^{j-1} p_1 = p_{j1} = \Pr(Z_n = 1 / Z_{n-1} = j) \] the j1 element in the matrix \( Q_k \). Thus the expression (4.9) may be written
\[ E_i^*(Z_L) = \frac{p_0}{p_1 p_i(k)} \sum_{j=1}^{k} n_{ij}(k) p_{j1} \] \( (4.10) \)

Since \( N_k Q_k = N_k - I \), this becomes
\[ E_i^*(Z_L) = \frac{p_0}{p_1 p_i(k)} \left[ n_{i1}(k) - S_{i1} \right] \] \( (4.11) \)

Finally, if we write \( \mathcal{P}(k) = \text{diag} \left[ p_1(k), \ldots, p_k(k) \right] \), we have
\[ E^*(Z_L) = \frac{p_0}{p_1} \left[ \mathcal{P}(k) \right]^{-1} \col \left[ n_{11}(k) - 1, n_{21}(k), \ldots, n_{k1}(k) \right] \] \( (4.12) \)
as the vector of expected values.

In the next section we shall make use of the fact that \( p_i(k) \) is known in discussing the behavior of a branching process subject to the condition that the population size never exceeds a given fixed value.

4.3 Bounded Population Size

In this section we shall consider various quantities associated with a branching process under the conditional hypothesis that the population size never exceeds a fixed value, say \( k \). Since, as is shown below, this situation can be interpreted as an absorbing Markov
chain with a finite number of states, all of that theory can be applied here. Those results which apply here will be indicated and we shall develop other results of interest only in the branching process.

Let $S_k$ be the statement -- all $Z_n$ are less than or equal to $k$. Then we can compute transition probabilities conditional on $S_k$ in the following way. Assume the process is in state $E_i$ so that all probabilities are conditional on this. (An $i$ subscript will be used to denote this dependence.) Let $p_{ij}^*(k)$ be the probability of a one step transition from state $E_i$ to state $E_j$ conditional on $S_k$: i.e., if $Z_{n-1} = i$, then $p_{ij}^*(k)$ is the probability that $Z_n = j$, given that the population size never exceeds the value $k$. Thus

$$p_{ij}^*(k) = Pr_i [(Z_n = j) / S_k] = \frac{Pr_i [(Z_n = j) \land S_k]}{Pr_i [S_k]}$$

$$= \frac{Pr_i [S_k / (Z_n = j)] Pr_i [Z_n = j]}{Pr_i [S_k]}$$

$$= \frac{Pr_j [S_k]}{Pr_i [S_k]} \cdot \frac{p_{ij}}{Pr_i [S_k]} = \frac{p_{ij}p_j(k)}{p_i(k)} \quad (4.13)$$

where $p_i(k)$ and $p_j(k)$ are given by $(4.5)$.

The formula obviously holds when $i = 0$ since $p_{0j} = \delta_{0j}$ (the Kronecker delta) and $p_0(k) = 1$. Also, from equations $(4.1)$ and $(4.13)$ we have
\[
\sum_{j=0}^{k} p^*_t(k) = \frac{1}{r_t(k)} \sum_{j=0}^{k} p_{ij} p_j^*(k) = \frac{1}{r_t(k)} \sum_{j=0}^{k} p_{ij} p_j(k) = 1.
\]

Thus the new situation is indeed a finite Markov chain with one absorbing state, \(E_0\).

Writing \(P^*_k = \begin{bmatrix} p^*_t(k) \end{bmatrix}\), \((i, j = 1, 2, \ldots, k)\), and \(\mathcal{P}(k) = \text{diag} \left[ p_1(k), p_2(k), \ldots, p_k(k) \right]\), we see that

\[
P^*_k = \begin{bmatrix} 1 & 0 \\ 0 & \mathcal{P}(k) \end{bmatrix}^{-1} P_k \begin{bmatrix} 1 & 0 \\ 0 & \mathcal{P}(k) \end{bmatrix}
\]

which gives a simple representation of the new transition matrix in terms of the original matrix \(P_k\). Moreover if we partition as before, writing

\[
P^*_k = \begin{bmatrix} 1 & 0 \\ R_k^* & Q_k^* \end{bmatrix},
\]

then

\[
Q_k^* = \left[ \mathcal{P}(k) \right]^{-1} Q_k \mathcal{P}(k).
\]

The inverse \(N_k^* = (I - Q_k^*)^{-1}\) is given by

\[
N_k^* = (I - Q_k^*)^{-1} = \left[ \mathcal{P}(k)^{-1} \mathcal{P}(k) - \left[ \mathcal{P}(k)^{-1} Q_k \mathcal{P}(k) \right] \right]^{-1}
= \left[ \mathcal{P}(k)^{-1} (I - Q_k) \mathcal{P}(k) \right]^{-1} = \left[ \mathcal{P}(k)^{-1} N_k \mathcal{P}(k) \right].
\]

We are now in a position to apply results from the theory of finite absorbing chains to our problem. Some of these results are
given below, statements being made in the terminology of branching processes. Proofs of standard results are omitted. These may be found, for example, in the book of Kemeny and Snell (17). Proofs will be given for results which are peculiar to the branching process.

Result 4.1. Let $n_{ij}^*$ be the $i,j$ element of the matrix $N_k^*$ and $n_j$ be the number of generations (including the zero generation) in which the population size is $j$. Then $n_{ij}^*$ is the expected value of $n_j$, given the initial population size is $i$ and all $Z_n \leq k$. Writing $E^*(n)$ for the matrix of expected values of the $n_j$ we have

$$E^*(n) = N_k^*$$ (4.17)

Since the total number of generations to extinction is simply the number of generations for which the population size is non-zero, we immediately have the second result.

Result 4.2. Let $T$ be the total number of generations to extinction and write $E_i^*(T)$ for the expected value of $T$ when $Z_0 = i$ and all $Z_n \leq k$. Then

$$E_i^*(T) = \sum_{j=1}^{k} n_{ij}^*.$$ (4.18)

If we define $E^*(T) = \text{col} \left[ E_1^*(T), E_2^*(T), \ldots, E_k^*(T) \right]$.

---

$^2/$ Here and elsewhere inclusion of the number $k$ in an expression such as $E_i^*(T)$ would unduly confuse the notation. In such cases the asterisk (*) will always be taken to imply that the quantities in question are conditional on the statement $S_k -- all Z_n \leq k$. 
and \( l_k = \text{col}(1, 1, \ldots, 1) \) (k ones) then the vector of expected values of \( T \) is given by
\[
E^*\!(T) = N_k^* l_k.
\] (4.19)

Result 4.3. Let \( V^*\!(T) \) be the vector of variances of \( T \) conditional on \( S_k \). Then \( V^*\!(T) \) is given by
\[
V^*\!(T) = \left[ 2N_k^* - I \right] E^*\!(T) - E_{sq}^*\!(T),
\] (4.20)
where \( E_{sq}^*\!(T) \) is a vector whose entries are the squares of the corresponding entries in \( E^*\!(T) \). For example, if \( k = 2 \) and \( E^*\!(T) = \text{col}(2, 3) \), then \( E_{sq}^*\!(T) = \text{col}(4, 9) \).

To illustrate these results we give in Table 3 the vectors \( E^*\!(T) \) and \( V^*\!(T) \) for the geometric distribution with \( p = 1/2 \), using \( k = 10 \). Even if \( Z_0 = 10 \) it will take, on the average, less than seven generations for the population to become extinct when population size never exceeds ten.

<table>
<thead>
<tr>
<th>( i )</th>
<th>( 1 )</th>
<th>( 2 )</th>
<th>( 3 )</th>
<th>( 4 )</th>
<th>( 5 )</th>
<th>( 6 )</th>
<th>( 7 )</th>
<th>( 8 )</th>
<th>( 9 )</th>
<th>( 10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_1^*!(T) )</td>
<td>2.47</td>
<td>3.53</td>
<td>4.31</td>
<td>4.92</td>
<td>5.38</td>
<td>5.75</td>
<td>6.04</td>
<td>6.27</td>
<td>6.46</td>
<td>6.61</td>
</tr>
</tbody>
</table>
Result 4.4. Let $G$ represent the total number of particles in all generations, $E^*_i(G)$ the expected value of $G$, given $Z_0 = i$ and all $Z_n \leq k$, and $E^*(G)$ the vector of the $E^*_i(G)$. Then

$$E^*(G) = N^*_k J_k$$

(4.21)

where $J_k = \text{col}(1, 2, \ldots, k)$.

Proof. The total number of individuals can be obtained by multiplying the number of generations in which a given population size is attained by that population size and summing over values from one to $k$. Thus, using Result 4.1,

$$E^*_i(G) = E^*_i \left[ \sum_{j=1}^{k} j n_j \right] = \sum_{j=1}^{k} j E^*_i(n_j) = \sum_{j=1}^{k} j n^*_i,$$

and the result follows immediately.

We next turn to a discussion of the variances of the quantities $n_j$ and $G$. Since in each case terms of the type $n_j n_r$ $(j, r = 1, 2, \ldots, k)$ appear we begin by deriving an expression for the expected value of this product.

Lemma. The expected value of the product $n_j n_r$ given that $Z_0 = i$ and all $Z_n \leq k$ is

$$E^*_i(n_j n_r) = n^*_i n^*_r + n^*_i n^*_j - \delta_{jr} n^*_i.$$  

(4.22)

Proof. Let $u_j(n)$ be a random variable which has the value one when $Z_n = j$ and is zero otherwise. Then $n_j = \sum_{n=0}^{\infty} u_j(n)$ and
\[ E_i^*(n, n_r) = E_i^* \left[ \left( \sum_{n=0}^{\infty} u_j(n) \right) \left( \sum_{m=0}^{\infty} u_{r}(m) \right) \right] \]

\[ = E_i^* \left[ \sum_{n=0}^{\infty} \sum_{m=0}^{n} u_j(n) u_{r}(m) + \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} u_j(n) u_{r}(m) \right. \]

\[ - \sum_{n=0}^{\infty} u_j(n) u_{r}(n) \left. \right] \]

\[ = E_i^* \left[ \sum_{n=0}^{\infty} \sum_{m=0}^{n} u_j(n) u_{r}(m) \right] + E_i^* \left[ \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} u_j(n) u_{r}(m) \right] \]

\[ - E_i^* \left[ \sum_{n=0}^{\infty} u_j(n) u_{r}(n) \right] \]

\[ = \sum_{n=0}^{\infty} \sum_{m=0}^{n} Pr \left[ u_{r}(m) = 1 \land u_j(n) = 1 \land (Z_o = i) \land S_k \right] \]

\[ + \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} Pr \left[ u_j(n) = 1 \land u_{r}(m) = 1 \land (Z_o = i) \land S_k \right] \]

\[ - \sum_{n=0}^{\infty} Pr \left[ u_j(n) = 1 \land u_{r}(n) = 1 \land (Z_o = i) \land S_k \right] \]

\[ = \frac{1}{p_i(k)} \sum_{n=0}^{\infty} \sum_{m=0}^{n} Pr \left[ u_{r}(m) = 1 \land u_j(n) = 1 \land S_k \land (Z_o = i) \right] \]

\[ + \frac{1}{p_i(k)} \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} Pr \left[ u_j(n) = 1 \land u_{r}(m) = 1 \land S_k \land (Z_o = i) \right] \]

\[ - \frac{1}{p_i(k)} \sum_{n=0}^{\infty} Pr \left[ u_j(n) = 1 \land u_{r}(n) = 1 \land S_k \land (Z_o = i) \right] \].
The probability of going from a population size of \( i \) to a population size of \( j \) in \( n \) generations while all intervening generations have population sizes not greater than \( k \) is given by the \( i, j \) term in \( Q^n_k \) which we shall write as \( Q^n_k(i, j) \). Also, the event \( u_j(n) = 1 \) if \( u_x(n) = 1 \) can happen only if \( j = r \) since otherwise population size cannot at once be both \( j \) and \( r \). Hence

\[
E_1^*(n_jn_r) = \frac{1}{p_i(k)} \sum_{n=0}^{\infty} \sum_{m=0}^{n} Q^n_k(i, r) Q^{n-m}_k(r, j) p_j(k)
\]

\[
+ \frac{1}{p_i(k)} \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} Q^n_k(i, j) Q^{m-n}_k(j, r) p_r(k)
\]

\[
- \frac{1}{p_i(k)} \delta_{jr} \sum_{n=0}^{\infty} Q^n_k(i, j) p_j(k)
\]

The last expression is obtained by noting that \( [u_j(n)]^2 = u_j(n) \).

We now change the variables of summation by writing \( s = m \) and \( t = n - m \) in the first term and \( s = n \) and \( t = m - n \) in the second term.

We then have

\[
E_1^*(n_jn_r) = \frac{1}{p_i(k)} \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} Q^s_k(i, r) Q^t_k(r, j) p_j(k)
\]

\[
+ \frac{1}{p_i(k)} \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} Q^s_k(i, j) Q^t_k(j, r) p_r(k)
\]

\[
- \frac{1}{p_i(k)} \delta_{jr} \sum_{n=0}^{\infty} Q^n_k(i, j) p_j(k)
\]
Since $\sum_{n=0}^{\infty} Q^p_k(i,j)$ is simply $n_{ij}$, the $i,j$ term in $N_k$, we obtain

$$E_i^*(n_{jr} n_{r}) = \frac{1}{p_i(k)} n_{ir} n_{rj} p_j(k) + \frac{1}{p_i(k)} n_{ij} n_{jr} p_r(k)$$

$$- \frac{1}{p_i(k)} \delta_{jr} n_{ij} p_j(k)$$

$$= n_{ir}^* n_{jr}^* + n_{ij}^* n_{jr}^* - \delta_{jr} n_{ij}^* \quad (4.23)$$

as is easily seen from the relation between $N_k$ and $N_k^*$. Thus the lemma is proved.

Result 4.5. Setting $j = r$ in the expression (4.23) gives

$$E_i^*(n_j^2) = 2 n_{ij}^* n_{jj}^* - n_{ij}^* \quad (4.24)$$

Thus for the matrix of expected values of the squares of the $n_j$ we have

$$E^*(n^2) = N_k^* \left[ 2 N_k^*(d) - I \right] \quad (4.25)$$

where $N_k^*(d)$ is the matrix obtained from $N_k^*$ by setting the off diagonal elements equal to zero. Hence, since $E^*(n) = N_k^*$ from (4.17), the matrix of variances of the $n_j$ is

$$V^*(n) = E^*(n^2) - E_{sq}^*(n)$$

$$= N_k^* \left[ 2 N_k^*(d) - I \right] - \left[ N_k^* \right]_{sq} \quad (4.26)$$

Result 4.6. We proceed in a somewhat different way to determine the variance of $G$. Since $G = \sum_{j=1}^{k} n_j$, then

$$G^2 = \sum_{j=1}^{k} \sum_{r=1}^{k} j r n_j n_r \quad \text{and thus}$$
\[ E_i^*(G^2) = \sum_{j=1}^{k} \sum_{r=1}^{k} j \ n_i^* \ n_r^* + \sum_{j=1}^{k} \sum_{r=1}^{k} j \ n_{ij}^* \ n_{jr}^* - \sum_{j=1}^{k} j^2 \ n_{ij}^* \]

If we write \( \mathcal{G}_k = \text{diag}(1, 2, \ldots, k) \) we have

\[ E_i^*(G^2) = \sum_{j=1}^{k} j \left( N_k^* \mathcal{G}_k N_k^* \right)_{ij} + \sum_{r=1}^{k} r \left( N_k^* \mathcal{G}_k N_k^* \right)_{ir} - \sum_{j=1}^{k} j^2 n_{ij}^* \]

\[ = 2 \left( N_k^* \mathcal{G}_k N_k^* \mathcal{G}_k \right)_i - \left( N_k^* \mathcal{G}_k^2 \right)_i \quad (4.27) \]

Since \( \mathcal{G}_k \mathcal{I}_k = \mathcal{I}_k \) we may write

\[ E^*(G^2) = N_k^* \mathcal{G}_k \left[ 2 \ N_k^* \mathcal{I}_k - \mathcal{I}_k \right] \quad (4.28) \]

Recalling (4.21) that \( N_k^* \mathcal{I}_k = E^*(G) \), we have for the vector of variances

\[ V^*(G) = E^*(G^2) - E^*_p(G) \]

\[ = N_k^* \mathcal{G}_k \left[ 2 \ E^*(G) - \mathcal{I}_k \right] - E^*_p(G) \quad (4.29) \]

To illustrate these results we have listed in Table 4 values in the vectors \( E^*(G) \) and \( V^*(G) \) for the geometric distribution with \( p = 1/2 \), using the value ten for \( k \). The expected total number of particles is small in each case although the standard deviations are large relative to the means.
Table 4. Values of $E^*_1(G)$ and $V^*_1(G)$ for the geometric distribution with $p = 1/2$, using $k = 10$

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E^*_1(G)$</td>
<td>4.80</td>
<td>9.27</td>
<td>13.41</td>
<td>17.21</td>
<td>20.69</td>
<td>23.85</td>
<td>26.73</td>
<td>29.35</td>
<td>31.73</td>
<td>33.91</td>
</tr>
<tr>
<td>$V^*_1(G)$</td>
<td>94.7</td>
<td>169.1</td>
<td>226.3</td>
<td>269.1</td>
<td>300.5</td>
<td>322.7</td>
<td>338.2</td>
<td>348.7</td>
<td>355.6</td>
<td>360.1</td>
</tr>
</tbody>
</table>

We now return to results which are deducible directly from theorems on finite Markov chains. (Again see Kemeny-Snell (17).) These in turn will lead directly to an alternative derivation of the distribution of the maximum population size.

Result 4.7. Let $s^*_{ij}$ be the probability that the population size will ever be $j$ after time zero given $Z_0 = i$ and all $Z_n \leq k$. If $S^*_k$ is the matrix of the $s^*_{ij}$, then

$$S^*_k = \left[ N^*_k - I \right] \left[ N^*_k(dg) \right]^{-1} \quad (4.30)$$

Proof. Recalling that $n^*_ij$ represents the expected number of generations in which the population size will be $j$, given $Z_0 = i$ and all $Z_n \leq k$, we have

$$n^*_ij = \delta^*_{ij} + s^*_{ij} n^*_jj$$

Hence

$$N^*_k = I + S^*_k N^*_k(dg)$$

from which the result follows immediately.
Result 4.8. Let \( h^*_{ij} \) be the probability that the population size \( j \) is ever achieved (including the initial generation). Obviously \( h^*_{ii} = 1 \) and \( h^*_{ij} = s^*_{ij} \) for \( i \neq j \). Writing \( H^*_k \) for the matrix of the \( h^*_{ij} \), we find

\[
H^*_k = N^*_k \left[ N^*_k(dg) \right]^{-1}.
\]  

(4.31)

From this we see that the quantity \( \frac{1}{n^*_i} \) represents the probability of leaving state \( E_i (Z_0 = i) \) and never returning, while \( 1 - \frac{1}{n^*_i} \) is the probability of eventually returning.

Finally we consider the probability \( g^*_{ij} \) that some \( Z_n \) (including \( n = 0 \)) has the value \( j \) and at the same time all \( Z_n \leq k \), given \( Z_0 = i \). Obviously

\[
g^*_{ij} = p_i(k) h^*_{ij}
\]  

(4.32)

so that writing \( G^*_k \) for the matrix of the \( g^*_{ij} \), we have

Result 4.9

\[
G^*_k = \mathcal{O}(k) N^*_k \left[ N^*_k(dg) \right]^{-1}
\]

\[
= \mathcal{O}(k) \left[ \mathcal{O}(k) \right]^{-1} N_k \mathcal{O}(k) \left[ N_k(dg) \right]^{-1}
\]

\[
= N_k \left[ N_k(dg) \right]^{-1} \mathcal{O}(k)
\]  

(4.33)

since \( N^*_k(dg) = N_k(dg) \) and the two diagonal matrices commute.

Setting \( j = k \) in the expression (4.32), we see that the last column of the \( G^*_k \) matrix is the vector of probabilities that the maximum population size will be exactly \( k \), for values of \( Z_0 \) running from one to \( k \). This gives (compare 4.7)
\[
\begin{bmatrix}
q_1(k) \\
q_2(k) \\
\vdots \\
q_k(k)
\end{bmatrix} = \frac{p_k(k)}{n_{kk}^*} \begin{bmatrix}
n_{1k}^* \\
n_{2k}^* \\
\vdots \\
n_{kk}^*
\end{bmatrix},
\]

thus showing that the \( q_i(k) \) depend in a simple way on \( p_k(k) \).
5.0 DEPENDENCE IN A BRANCHING PROCESS

5.1 Introduction

5.1.1 General Comments. The assumption that particles reproduce independently is a fundamental one in the formulation of the classical branching process. In many applications, however, this assumption may only approximate the actual situation and, indeed, be seriously in error as population size increases. In the nuclear reaction, for example, independence will not hold for very large numbers of particles (neutrons in this case).

Situations of this kind point to the need for a systematic study of dependence among particles in a branching process. In this chapter we shall consider and give a partial solution to this problem for a simple branching process. Ways will then be indicated in which the methods used here might be extended to the solution of the general problem.

5.1.2 Description of the Problem. The problem may best be stated in geometric form. This has the added advantage of making clear the connection between this and some other seemingly unrelated problems.

Suppose we have a grid of connected points of the type indicated in Figure 3. It is of infinite extent in the downward direction. The nth row contains n points which we shall label (n,1), (n,2), ..., (n,n).
Point \((n, k)\) in row \(n\) is connected to points \((n+1, k)\) and \((n+1, k+1)\) in row \(n+1\).

![Figure 3. The infinite grid of points](image)

We assume that each point, independently of all other points, has common probability \(p\) of being open, or passable, and probability \(q = 1 - p\) of being closed, or blocked. If a given point is open we may pass through to the next row, while if it is closed no passage is possible through that point. Let \(P_n\) be the probability that, beginning at point \((1, 1)\), we shall be able to pass through the grid to row \(n\).

That is to say, \(P_n\) is the probability that there is an open path from point \((1, 1)\) to some point in row \(n\). We wish to determine

\[ P_\infty = \lim_{{n \to \infty}} P_n, \]

the probability that there exists some open path of infinite length. This will naturally depend in some way on the probability \(p\) of an individual point being open. We wish to determine this
relationship and to find the critical value of p, this being defined as
the largest value of p for which \( P_\infty \) is zero.

This problem, in a slightly different form, is due to
Hammersley and Broadbent (14). It is a special case of a general type
of stochastic process known as a percolation process. In
Hammersley's terminology we may describe the problem as that of a
liquid "percolating" through some porous material. For example, we
may think of oil passing through sandstone. Here if a point is open the
liquid continues, passing into the two channels leading from the point,
while if the point is closed the liquid cannot pass. The liquid enters
the system at point \((1, 1)\) and \( P_n \) represents the probability that at
least one point in row \( n \) is wetted while \( P_\infty \) is the probability of
wetting an infinite number of points.

Our interest in the problem here lies in the fact that it may be
thought of as a branching process in which the particles do not
reproduce independently. For, consider the simplest non-trivial
branching process in which with probability \( q \) a particle fails to
reproduce and with probability \( p = 1 - q \) reproduces two new particles.
Let each point in the grid represent a particle, the point \((1, 1)\) being
the initial particle. If a point is reached we shall say that the
associated particle has been produced. Otherwise, that particle never
exists. A particle reproduces or fails to reproduce according as its
point is open or closed. Reproduction is not performed independently
since a single particle may have more than one progenitor. Thus there is a positional dependence among the particles.

In general our scheme will tend to produce fewer offspring than when particles reproduce independently. For here, if two adjacent particles reproduce, we obtain only three new ones as opposed to four in the independent case. Otherwise, reproductive properties are identical.

In case particles do act independently we have the generating function \( f(s) = q + ps^2 \). When \( Z_0 = 1 \), \( m = E(Z_1) = 2p \) and \( \sigma^2 = \text{Var}(Z_1) = 4pq \). Also \( s_o = 1 \) for \( p \leq 1/2 \) and \( s_o = q/p \) for \( p > 1/2 \). The critical probability here is \( 1/2 \).

For the grid scheme let \( t_o \) be the probability that extinction occurs, or that only a finite number of rows can be reached. It is apparent that \( t_o \geq s_o \). Since \( P_\infty = 1 - t_o \), we have \( P_\infty \leq 1 - s_o \) and it follows that the critical probability for our problem is greater than or equal to one-half. Figure 4 shows the curve \( y = 1 - s_o \) which we shall call the first upper bound to \( P_\infty \).

![Figure 4. Graph of \( y = 1 - s_o \) when \( f(s) = q + ps^2 \)]
Hammersley's results pertain only to the critical probability which he shows (13) to be at least 0.518. Our work will be concerned with improving this lower bound and lowering the curve representing an upper bound for $P_\infty$. The original problem has not been completely solved. We have, however, been able to show that the critical probability is not less than two-thirds and to solve some variations of the problem. These will be discussed in the next section.

5.2 Variations of the Original Problem

5.2.1 **Lower Bounds for $P_\infty$.** As proposed, our problem is difficult to solve since not only the number of open points in a given row is important but also the position of these points relative to each other. Reproductive properties differ depending on whether the points are contiguous or are separated. For this reason various simplifying assumptions have been made. Each assumption provides a variation of the problem. Several of these will be presented. Each provides certain information about the basic problem and some also have interesting interpretations of their own. We shall deal first with situations in which the chance of passing through the grid is less than in the original problem. It is convenient to have both the terminology of the branching process and of the grid scheme at our disposal and we shall use these interchangeably.
5.2.1.1 Contiguous Open Points. Rather than allow the points to retain their respective positions on the grid, suppose at each stage (row) we take all points which can be reached from point (1, 1) and arrange them consecutively, ignoring all others. Effects of position in a row are thus eliminated and the chance of passing through the grid is, if anything, less than that in the original problem.

Recalling from theorem 3.3 that the critical point in the standard (independent particles) process occurs when the expected number of progeny from one particle is one (and thus the expected number from n particles is n), we wish to consider here the expected size of a succeeding generation given that we have n particles. In the terminology of the grid scheme this is equivalent to the expected number of points reached in row n+1 when we begin with the n points in row n. Specifically we wish to determine the value of p for which \( E_n \), the expected number of children from n parents, is equal to n.

Let us write \( p_{nj} \) for the probability of \( j \) offspring from n parents. Table 5 gives these values for \( n = 0, 1, 2, 3 \).

<table>
<thead>
<tr>
<th>Number of Parents (n)</th>
<th>Number of Children (j)</th>
<th>( 0 )</th>
<th>( 1 )</th>
<th>( 2 )</th>
<th>( 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>( q )</td>
<td>( p )</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>( q^2 )</td>
<td>( 2pq )</td>
<td>( p^2 )</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>( q^3 )</td>
<td>( 3pq^2 )</td>
<td>( 2p^2q )</td>
<td>( p^2 )</td>
<td></td>
</tr>
</tbody>
</table>
From this we see that

\[ E_0 = 0 \text{ regardless of } p; \]
\[ E_1 = 2p \text{ which equals one when } p = 1/2; \]
\[ E_2 = 4p - p^2 \text{ which is 2 when } p = \frac{2 - \sqrt{2}}{1}; \]
\[ E_3 = 6p - 2p^2 \text{ which is 3 when } p = \frac{3 - \sqrt{3}}{2}; \]

When \( n = 2 \) or \( n = 3 \), the required value of \( p \) is given by the equation

\[ 2np - (n - 1)p^2 = n \text{ which has the solution } p = \frac{n - \sqrt{n}}{n - 1}. \]

Using induction, we may establish this formula for all values of \( n \geq 2 \). Let

\( G_1 \) be the event that the last of \( n \) consecutive points (point \( (n, n) \)) fails to reproduce; i.e., that it is closed. This has probability \( q \).

\( G_2 \) be the event that point \( (n, n) \) is open while point \( (n, n-1) \) is closed. This has probability \( pq \).

\ldots

\( G_{n+1} \) be the event that all points are open. This has probability \( p^n \).

The events \( G_1, \ldots, G_{n+1} \) are mutually exclusive and exhaustive since the sum of their probabilities is

\[ \sum_{k=0}^{n-1} qp^k + p^n = 1. \]

Now let \( m_i \) be the expected number of points reached given \( G_i \) has occurred. It is easily seen that \( m_i = i + E_{n-i} \) for \( 2 \leq i \leq n-1 \),
while \( m_1 = E_{n-1} \), \( m_n = n \) and \( m_{n+1} = n+1 \). Hence

\[
E_n = qE_{n-1} + pq(2 + E_{n-2}) + p^2q(3 + E_{n-3}) + \ldots
\]

\[
\ldots + p^{n-2}q(n-1 + E_1) + p^{n-1}q(n) + p^n(n+1).
\]

This reduces to

\[
E_n = q \sum_{j=0}^{n-2} p^jE_{n-j-1} + 2p + \frac{p^2 - p^{n+1}}{q}.
\]

Now using as an inductive assumption \( E_k = 2kp - (k-1)p^2 \) for \( k < n \), we find

\[
E_n = 2np - (n-1)p^2,
\]

thus proving our conjecture.

As \( n \) increases \( \frac{n - \sqrt{n}}{n - 1} \) approaches one. Thus in order to insure that the expected number of progeny is as great as the number of parents, \( p \) must approach one for large \( n \). Conversely, if we ask (for fixed \( p \)) the value of \( n \) for which \( \frac{n - \sqrt{n}}{n - 1} = p \), we find \( n = (p/q)^2 \).

Thus for \( n > (p/q)^2 \), \( E_n < n \), while for \( n < (p/q)^2 \), \( E_n > n \). The process, then, will tend to oscillate about the value \( (p/q)^2 \) but will die out with probability one.

This type of action has certain physical analogs. For example, a nuclear reactor of the simple type having a core, a reflector and an absorber which may be inserted or withdrawn in order to change the level of the reaction, will exhibit the same behavior. That is, the number of neutrons present will remain close to a "critical" value.
which depends on the position of the absorber. Whether the present model can be of value in discussing such processes is a question which should be investigated.

5.2.1.2 Restricted Branching. Suppose, as in the original problem, a point is open with probability $p$. Let us impose the restriction that if the point immediately to the left branches, the given point may branch only to the right, while if the left point does not branch, the given one may branch both ways. Points on the extreme left may always branch both ways. This scheme has the effect of not allowing open paths to cross and thus allows computation of probabilities of reaching each point. If we are at row $n$ we can reach row $n+1$ exactly as in the original problem. However, a point, whether it can be reached from above or not, may affect the behavior of another point. Thus this version is not equivalent to the original problem and it is apparent that the probability of passing through the grid has been reduced.

To summarize, for $k > 1$, points $(n, k)$

(a) fail to branch with probability $q$,
(b) branch only right with probability $p^2$,
(c) branch both ways with probability $pq$,

while points $(n, 1)$ fail to branch with probability $q$ and branch both ways with probability $p$. 
Now let \( P(n, k) \) be the probability of reaching point \((n, k)\). Then

\[
P(n, 1) = p^{n-1} \quad .
\] (5.2)

For points \((n, k)\) where \(k \geq 2\) we find

\[
P(n, k) = Pr[(2, 1) \cap (n, k)] + Pr[(2, 2) \cap (n, k)]
\]

where \((2, j) \cap (n, k)\) means that we reach point \((n, k)\) by passing through point \((2, j)\). The probabilities add since our scheme ensures that a given point cannot be reached from two different points. Thus

\[
P(n, k) = p \cdot P(n-1, k) + p \binom{n-2}{k-2} p^{n-2} q^{n-k}
\]

where we must have \(P(n-1, n) = 0\) since \(k\) can never be greater than \(n\). The second term on the right arises since it takes \(k-2\) steps to the right and \(n-k\) steps left to reach point \((n, k)\) from point \((2, 2)\). Steps right have probability \(p\) and steps left probability \(pq\). Hence

\[
\sum_{k=1}^{n} P(n, k) = p^{n-1} + p \sum_{k=1}^{n-1} P(n-1, k) - p \cdot P(n-1, 1)
\]

\[
+ p^{n-1} \sum_{k=2}^{n} \binom{n-2}{k-2} q^{n-k} \quad .
\]

Writing \(\sigma_n = \sum_{k=1}^{n} P(n, k)\), this reduces to

\[
\sigma_n = p \cdot \sigma_{n-1} + p^{n-1} (1 + q)^{n-2}
\]

which has the solution

\[
\sigma_n = \frac{(1 - q^2)^{n-1} - p^n}{q} \quad .
\] (5.3)
It is easily seen that as $n$ increases $\sigma_n$ approaches zero. Since $P_n$ cannot be greater than $\sigma_n$, then $P_n$ approaches zero. Hence the probability of a path of infinite length is zero in this case.

This approach to the problem may be extended to yield a version which is equivalent to the original formulation. To accomplish this we must ensure that points which are not reached are not allowed to affect future events. Thus we will allow a point to affect its right hand neighbor only if it is open and is reached. As before, points $(n, 1)$ branch both ways with probability $p$ and fail to branch with probability $q = 1 - p$. However, for $2 \leq k \leq n$, points $(n, k)$

(a) fail to branch with probability $q$;
(b) branch only to the right with probability $p^2 P(n, k-1)$;
(c) branch both ways with probability

$$p(1 - P(n, k-1)) + pqP(n, k-1) = p - p^2 P(n, k-1).$$

Unfortunately this restatement of the problem leads to difficulties just as great as those encountered originally. In particular, it is not possible to write a simple expression for the probability of reaching a given point, as we were able to do in the simpler case. However, it is felt that some method such as this of reformulating the problem may well lead to a simple solution.

5.2.2 Upper Bounds for $P_\infty$. In this section we consider three variations of the original problem which yield upper bounds on the probability of passing through the grid.
5.2.2.1 Certain Closed Points Considered Open. We first consider a process in which the condition is imposed that if any two points in a given row are open, then all intervening points are also considered to be open. Representing open points by 1's and closed points by 0's it is apparent that the only possible configurations which may appear in a given row are

```
 0

 1

 1  1

 1  1  1

 1  1  1  1
```

with zeros in all extreme positions. Since the zeros are placed on the ends of the rows they may be ignored for purposes of continuing through the grid. Thus we may think of always reverting to one of the above situations. In this sense we have a Markov chain with one absorbing state (0) and transition probabilities as indicated in Table 6. (The transition matrix is, of course, infinite.) We shall say that an arrangement is of type \( n \) if it contains \( n \) consecutive ones. The transitions geometrically represent movement from one row of the grid to the next and the \( n \)-step transitions are from any row to the one \( n \) rows removed.
Table 6. Partial transition matrix

<table>
<thead>
<tr>
<th>From type</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>$q^2$</td>
<td>2pq</td>
<td>$p^2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>$q^3$</td>
<td>3pq$^2$</td>
<td>2$p^2$q</td>
<td>$p^2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>$q^4$</td>
<td>4pq$^3$</td>
<td>3$p^2$q$^2$</td>
<td>2$p^2$q</td>
<td>$p^2$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>$q^5$</td>
<td>5pq$^4$</td>
<td>4$p^2$q$^3$</td>
<td>3$p^2$q$^2$</td>
<td>2$p^2$q</td>
<td>$p^2$</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>$q^6$</td>
<td>6pq$^5$</td>
<td>5$p^2$q$^4$</td>
<td>4$p^2$q$^3$</td>
<td>3$p^2$q$^2$</td>
<td>2$p^2$q</td>
<td>$p^2$</td>
</tr>
<tr>
<td>6</td>
<td>$q^7$</td>
<td>7pq$^6$</td>
<td>6$p^2$q$^5$</td>
<td>5$p^2$q$^4$</td>
<td>4$p^2$q$^3$</td>
<td>3$p^2$q$^2$</td>
<td>2$p^2$q</td>
</tr>
</tbody>
</table>

We wish to determine the probability that the process never dies out or, equivalently, the probability that we pass through the grid. In order to determine this value we make use of the fact that we have an infinite absorbing Markov chain and compute this as the probability of remaining forever in the transient states. Results from Markov chain theory will be stated and used without proof. Details may be found in Feller (8, page 353, ff).

Let $b_n$ be the probability that the process never dies out if we start with an arrangement of type $n$. Let $B$ be the vector of the $b_n$.

The first entry, $b_1$, corresponds to $P_\infty$ for the original problem and is such that $b_1 \geq P_\infty$ since this variation increases the probability of passing through the grid.
If we write the transition matrix given by Table 6 in the partitioned form

\[ P = \begin{bmatrix} 1 & 0 \\ C & Q \end{bmatrix}, \]

we know that \( B \) must satisfy the relation \( B = QB \) or \( (I - Q)B = 0 \).

This yields the infinite set of equations

\[
(1 - 2pq)b_1 - p^2b_2 = 0 \\
-3pq^2b_1 + (1 - 2p^2q)b_2 - p^2b_3 = 0 \\
-4pq^3b_1 - 3p^2q^2b_2 + (1 - 2p^2q)b_3 - p^2b_4 = 0 \\
\vdots \\
\vdots \\
(5.4)
\]

We can choose an arbitrary value for \( b_1 \) and then determine the other \( b \)'s in terms of this value. The value which we choose is the largest one for which all \( b_n \) are less than or equal to one. Since \( b_n \) is an increasing function of \( n \) we determine \( b_1 \) from the equation

\[
\lim_{n \to \infty} b_n = 1 
\]

To proceed with the solution we first solve for the last \( b \) appearing in each of the equations (5.4). This gives

\[
b_2 = \frac{1 - 2pq}{p^2} b_1 \\
b_3 = \frac{1 - 2p^2q}{p^2} b_2 - \frac{3pq^2}{p^2} b_1
\]
\[ b_4 = \frac{1 - 2p^2q}{p^2} b_3 - \frac{3p^2q^2}{p^2} b_2 - \frac{4pq^3}{p^2} b_1 \]

\[ \ldots \]

and in general for \( n \geq 4 \),

\[ b_n = \frac{1 - 2p^2q}{p^2} b_{n-1} - \sum_{k=2}^{n-2} (n-k+1)q^{n-k} b_k - \frac{nq^{n-1}}{p} b_1 \quad (5.5) \]

We take \( b_0 = 0 \) and \( b_1 \) arbitrary.

Rewriting (5.6) we have

\[ b_n = \frac{1 - 2p^2q}{p^2} b_{n-1} - \sum_{k=0}^{n} (n-k+1)q^{n-k} b_k \]

\[ - \frac{nq^{n-1}}{p} b_1 + nq^{n-1} b_1 + 2q b_{n-1} + b_n \]

which reduces to

\[ 0 = b_{n-1} - p^2 q \sum_{k=0}^{n} t_{n-k} b_k - p^2 \sum_{k=0}^{n} q^{n-k} b_k - b_1 p q t_n \quad (5.7) \]

where \( t_n = nq^{n-1} \).

We now introduce the generating functions \( B(z) = \sum_{n=0}^{\infty} b_n z^n \),

\( T(z) = \sum_{n=0}^{\infty} t_n z^n \) and \( Q(z) = \sum_{n=0}^{\infty} q^n z^n \). Multiplying the expressions (5.5) and (5.7) for \( b_n \) by \( z^n \) and summing on \( n \) yields, after some reduction,

\[ zB(z) = p^2 q T(z) B(z) + p^2 Q(z) B(z) + pq b_1 T(z) - pz b_1 \]
Solving for $B(z)$ and using the expressions

$$T(z) = \frac{z}{(1 - qz)^2} \quad \text{and} \quad S(z) = \frac{1}{1 - qz}$$

we obtain

$$B(z) = \frac{[pq - p(1 - qz)^2]}{z(1 - qz)^2 - p^2} b_1 z,$$

which may be rewritten in the form

$$B(z) = -pb_1 \left[ 1 + \frac{p^2 - qz}{z(1 - qz)^2 - p^2} \right]. \quad (5.8)$$

Considering the denominator of the second term we have

$$z(1 - qz)^2 - p^2 = q^2 \left[ z^3 - \frac{2}{q} z^2 + \frac{1}{q^2} z - \frac{p^2}{q^2} \right]$$

$$= q^2 (z - 1) \left[ z^2 - \left( \frac{2}{q} - 1 \right) z + \frac{p^2}{q^2} \right]$$

$$= q^2 (z - 1) (z - a_1) (z - a_2)$$

where

$$a_1 = \frac{2-q + \sqrt{q(1+3p)}}{2q} \quad \text{and} \quad a_2 = \frac{2-q - \sqrt{q(1+3p)}}{2q}.$$

We find that $a_1 \geq 1$ and $a_1 \geq a_2$ always while $a_2 \geq 1$ when $p \geq 2/3$ and $a_2 < 1$ when $p < 2/3$.

By expanding the second term of (5.8) into partial fractions and noting that $a_1 a_2 = \frac{p^2}{q^2}$ we may write $B(z)$ in the form

$$B(z) = \frac{b_1}{p} \left[ \frac{A_1}{1 - z} + \frac{A_2}{1 - \frac{z}{a_1}} + \frac{A_3}{1 - \frac{z}{a_2}} - p^2 \right]. \quad (5.9)$$

where $A_1$, $A_2$, and $A_3$ are constants given by
\[
A_1 = \frac{p^2(p^2 - q)}{q^2(a_1 - 1)(a_2 - 1)} = \frac{p(p^2 - q)}{1 - 3q},
\]
\[
A_2 = \frac{-p^2\left[q + \sqrt{q(1 + 3p)}\right]}{2q(a_1 - 1)(a_1 - a_2)} \quad \text{and}
\]
\[
A_3 = \frac{p^2\left[\sqrt{q(1 + 3p)} - q\right]}{2q(a_2 - 1)(a_1 - a_2)}.
\]

When \( p > 2/3 \), \( A_1 \) and \( A_3 \) are positive, while \( A_2 \) is always negative.

Expanding each of the fractions in (5.9) into a power series we have

\[
B(z) = \frac{b_1}{p} \left[ A_1 \sum_{n=0}^{\infty} z^n + A_2 \sum_{n=0}^{\infty} \left(\frac{z}{a_1}\right)^n + A_3 \sum_{n=0}^{\infty} \left(\frac{z}{a_2}\right)^n - p^2 \right].
\]

For \( n \geq 1 \) the coefficient of \( z^n \) is \( b_n = \frac{b_1}{p} \left[ A_1 + \frac{A_2}{a_1^n} + \frac{A_3}{a_2^n} \right] \)

which becomes \( \frac{b_1}{p} \left[ A_1 + \frac{A_3}{a_2^n} \right] \) as \( n \) increases since \( a_1 > 1 \) except when \( q = 1 \) (and then \( A_2 = 0 \)). If \( a_2 < 1 \) the series diverges and no bounded solution (other than the zero solution) exists for \( b_n \). Here \( p < 2/3 \) and all \( b_n = 0 \).

If \( a_2 = 1 \), \( A_1 \) does not exist and the same conclusion holds.

Here \( p = 2/3 \).

When \( p > 2/3 \), \( a_2 > 1 \) and \( b_n \) approaches \( \frac{A_1}{p} b_1 = \frac{p^2 - q}{1 - 3q} b_1 \).

Thus a bounded solution exists and we want the maximal solution for which all \( b_n \leq 1 \). To obtain this we set \( \frac{p^2 - q}{1 - 3q} b_1 = 1 \) and find, for values of \( p \) exceeding two-thirds.
\[ b_1 = \frac{1 - 3q}{p^2 - q} \]  \hspace{1cm} (5.10)

A graph of this curve is shown in Figure 5. The curve rises rapidly from the value zero at \( p = 2/3 \) to the value 4/5 when \( p = 3/4 \). Of course \( b_1 = 1 \) when \( p = 1 \).

![Figure 5. Graph of the curve \( b_1(p) \)](image)

It may be seen from the graph that the value two-thirds is the critical probability for the process. Since \( b_1(p) \) provides an upper bound for \( P_\infty \) it is clear that the critical probability for the original problem is greater than or equal to this value.

Comparing the curve \( b_1(p) \) with the first upper bound \( 1 - s_0 = 1 - q/p \) we find that the curves intersect at \( p = 1/\sqrt{2} = 0.707 \). Thus while \( b_1(p) \) provides the better estimate of the critical probability, it is a poorer bound throughout most of the interval \((2/3, 1)\).
The next variation will provide a bound which is still better than 
1 - q/p throughout the interval.

5.2.2.2 Arrangements of Zeros and Ones. While the critical 
value 2/3 is the best we have been able to obtain, it is possible to 
improve the upper bound b_1(p) throughout most of the range (2/3, 1) by 
using a different approach to the problem.

Let P_n (any arrangement A) be the probability of reaching row 
n and beginning with the given arrangement, A. For example 
P_n \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} would represent the probability of reaching row n and having 
points (1, 1) and (2, 2) open while point (2, 1) is closed. Let P_n^i(A) be 
the conditional probability of reaching row n given that we started 
with the arrangement A. Then P_n is given by

\[ P_n = \sum_A P(A) \cdot P_n^i(A) \]

where the summation is taken over all k rowed arrangements (for any 
fixed k). P(A) is the probability of the arrangement A.

For arrangements involving only one row we have \( P_n = pP_n^i(1) \) 
so that

\[ P_n^i(1) = \frac{P_n}{p} \] \hspace{1cm} (5.11)

For two rows

\[ P_n = P_n \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + P_n \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + P_n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \]

\[ = 2P_n \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + P_n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \]
\[ = 2P_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} P_n \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + P_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} P_n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \]
\[ = 2p^2q P_{n-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + p^3 P_n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \]
\[ = 2pq P_{n-1} + p^3 P_n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{5.12} \]

From this we obtain two useful results. First,
\[ P_n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{p^3} (P_n - 2pq P_{n-1}) \tag{5.13} \]

Second, if we write \( S_n(2, 1) \), \( S_n(2, 2) \) and \( S_n \begin{bmatrix} (2, 1), (2, 2) \end{bmatrix} \) for, respectively, the statements, row \( n \) is reached from point \( (2, 1) \), row \( n \) is reached from point \( (2, 2) \), and row \( n \) is reached from both points \( (2, 1) \) and \( (2, 2) \), all conditional on the given arrangement \( \begin{pmatrix} 1 \\ 1 \end{pmatrix} \), we have
\[ P_n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = P_1 \left[ S_n(2, 1) \right] + P_1 \left[ S_n(2, 2) \right] - P_1 \left[ S_n \begin{bmatrix} (2, 1), (2, 2) \end{bmatrix} \right]. \]

Now
\[ P_1 \left[ S_n \begin{bmatrix} (2, 1), (2, 2) \end{bmatrix} \right] = P_1 \left[ S_n(2, 1) \right] P_1 \left[ S_n(2, 2) / S_n(2, 1) \right], \]
and since it must hold that
\[ P_1 \left[ S_n(2, 2) / S_n(2, 1) \right] \geq P_1 \left[ S_n(2, 2) \right] = P_1 \left[ S_n(2, 1) \right], \]
we have the inequality
\[ P_n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} 2P_1 \left[ S_n(2, 1) \right] - \left[ P_1 \left[ S_n(2, 1) \right] \right]^2. \]
But \( P_1 \left[ S_n(2, 1) \right] \) is the same as \( P_{n-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \) and thus we may write
\[
\begin{align*}
\mathbf{P}_n = & \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mathbf{2P}_n^l \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \left[ \mathbf{P}_n^l \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right]^2 \\
= & \frac{2}{p} \mathbf{P}_n - \frac{\mathbf{P}_n^2}{p^2} \\
\end{align*}
\]

(5.14)

using (5.11). Hence from (5.12) we have

\[
\begin{align*}
\mathbf{P}_n = & 2pq\mathbf{P}_n^l + p^3\mathbf{P}_n^l \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
= & 2pq\mathbf{P}_n^l + p^3 \left[ \frac{2}{p} \mathbf{P}_n - \frac{\mathbf{P}_n^2}{p^2} \right] \\
\end{align*}
\]

or

\[
\mathbf{P}_n \leq 2p\mathbf{P}_n^l - p\mathbf{P}_n^2 \\
\]

(5.15)

Since \( \mathbf{P}_n \) is a monotone decreasing function of \( n \), bounded below by zero, it must approach a limit \( \mathbf{P} \) as \( n \) becomes infinite.

Thus in the limit we have

\[
\mathbf{P} \leq 2p\mathbf{P} - p\mathbf{P}^2 \\
\]

(5.16)

If \( \mathbf{P} = 0 \) our problem is solved. If \( \mathbf{P} \neq 0 \) division of both sides of (5.16) by \( \mathbf{P} \) gives

\[
1 \leq 2p - p\mathbf{P} \quad \text{or} \quad \mathbf{P} \leq \frac{2p - 1}{p} \\
\]

(5.17)

Now \( \mathbf{P} \) corresponds to \( b_1 \) in the previous discussion and is actually equal to the first upper bound, \( 1 - q/p \), for \( \mathbf{P}_\infty \). Thus we already know that when \( p \geq 1/\sqrt{2} \) this is a better upper bound than \( b_1 \).

Let us pursue the same considerations for three rowed arrangements. The discussion may be simplified by noting that the conditional
probabilities depend only on the arrangement of ones which can be reached in row three. For instance, for purposes of continuing, the arrangement

\[
\begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 0 \\
1 & 0 & 1
\end{pmatrix}
\]

is essentially \((0 0 1)\) since the left one cannot be reached while the one on the right is reached. (To be reached a point must be connected to point \((1, 1)\) by an unbroken string of ones, or open points.) Also we note that for conditional purposes the arrangements \((1 0 1)\) and \((1 1 1)\) are equivalent since all points in row four can be reached from either. Thus we will now write \(P'(1 0)\) instead of \(P'\begin{pmatrix}1 \\ 1 \end{pmatrix}\), \(P'(1 1)\) in place of \(P'\begin{pmatrix}1 \\ 1 \end{pmatrix}\), \(P'(1 0 0)\) for \(P'\begin{pmatrix}1 \\ 1 \\ 0 \end{pmatrix}\), etc. Then for three rows we have

\[
P_n = P_n(1 0 0) + P_n(1 1 0) + P_n(1 0 1)
\]

\[
= P(1 0 0) P_n'(1 0 0) + P(1 1 0) P_n'(1 1 0)
\]

\[
+ P(1 0 1) P_n'(1 0 1)
\]

\[
= p^3 q^2 (4 + 3p) P_{n-2}'(1) + 2p^4 q(1 + p) P_{n-1}'(1 1)
\]

\[
+ p^5 P_n'(1 0 1)
\]

(5.18)

The probabilities \(P(1 0 0)\), etc., are computed by simply writing out all arrangements equivalent for purposes of continuing and adding their probabilities.

Using (5.11) and (5.13), (5.18) becomes

\[
P_n = -p^3 q^2 P_{n-2} + 2pq(1 + p) P_{n-1} + p^5 P_n'(1 0 1)
\]

(5.19)
Also, arguing as in the case of two rows we find

\[ P'(101) \equiv 2P'_{n-2}(1) - \left[P'_{n-2}(1)\right]^2 \]

\[ \equiv \frac{2}{p} P_{n-2} - \frac{P_{n-2}^2}{p^2} . \]

Putting this in (5.19) and passing to the limit as before gives

\[ P \leq T(p) \equiv \frac{2p - 1}{p^3} - (3 - p)(1 - p), \quad P \neq 0 \quad (5.20) \]

The critical probability obtained by setting the right side of (5.20) equal to zero is just slightly greater than 0.6. The function \( T(p) \) is smaller in magnitude than \( \frac{2p - 1}{p} = 1 - q/p \) for all values of \( p \) in the interval \( (2/3, 1) \) and hence provides a better upper bound for \( P_\infty \). We still use \( b_1 \) for values of \( p \) close to 2/3.

For \( k > 3 \), the algebra involved in this approach becomes prohibitive and improvement comes slowly. For example, for \( k = 4 \) the bound at \( p = 2/3 \) is approximately 0.31 which is only slightly better than the corresponding value 25/72 for \( k = 3 \). The calculations are quite extensive and no pattern seems to emerge.

By using both the curves \( b_1(p) \) and \( T(p) \) the best upper bound for the probability \( P_\infty \) is obtained. These curves intersect at the point \( p = p^* \), say, where \( p^* \) is approximately 0.692. Thus the best upper bound, \( P, \) is given by

\[ P = 0 \quad \text{for} \ P \leq 2/3 , \]

\[ P = b_1(p) = \frac{1 - 3q}{p^2 - q} \quad \text{for} \ 2/3 < p < p^* , \]
\[ P = T(p) = \frac{2p - 1}{p^3} - (3 - p)(1 - p) \quad \text{for } p \geq p^*. \quad (5.21) \]

Figure 6 is a graph of this curve.

Figure 6. The best upper bound for \( P_\infty \)

5.2.2.3 A Combinatorial Approach. The following is a third method for obtaining an upper bound on the probability \( P_\infty \) of passing through the grid. It yields nothing new in terms of results but is included for its own interesting features.

Let us define \( P'(n, k) \) to be the probability of reaching point \((n, k)\) but not \((n, j)\) for any \( j < k \); i.e., we reach no point in row \( n \) to the left of point \((n, k)\). In order to reach a point here we shall require a string of ones to and including the point. Then defining \( P'(n, 1) = P(n, 1) \) we have
\[ P_{n+1} = \sum_{k=1}^{n} P'(n, k) \]  \hspace{1cm} (5.22)

where \( P_n \), as usual, represents the probability of reaching row \( n \).

In order to compute these probabilities we try to count all independent arrangements which allow us to reach \((n, k)\) but no point to the left in row \( n \). We are not able to make this count exactly but can obtain an upper bound. Consider the truncated grid shown in Figure 7.

\begin{center}
\begin{tabular}{c}
. (1, 1) \\
. . .
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{c}
. . .
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{c}
. . . .
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{c}
. . .
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{c}
. . .
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{c}
. .
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{c}
(1, 1) . . . (n, k)
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{c}
. . .
\end{tabular}
\end{center}

Figure 7. Truncated grid of \( n \) rows

In order to reach point \((n, k)\) and no point \((n, j)\) to the left we must reach the \( k \)th (positive slope) diagonal and block all lower numbered diagonals. We must also have a continuous string of \( n \) ones from \((1, 1)\) to \((n, k)\). To block the first \( k-1 \) diagonals requires at least \( k-1 \) zeros, one in each diagonal, arranged in such a way that the zero in diagonal \( j-1 \) is not below the zero in diagonal \( j \). The
probability of \( k-1 \) zeros and \( n \) ones in given positions (in each case we place the ones as closely above the zeros as possible so that each arrangement of zeros determines the arrangement of the ones) is \( q^{k-1} p^n \). The other positions are arbitrary. Hence arrangements of zeros may be duplicated, and we may obtain an upper bound by multiplying the probability \( q^{k-1} p^n \) by the total number of arrangements of \( k-1 \) zeros subject to the specified conditions. This number is given by the sum

\[
\sum_{i_{k-1}=1}^{n-(k-1)} \sum_{i_{k-2}=i_{k-1}}^{n-(k-2)} \ldots \sum_{i_1=i_2}^{n-1} (1) \equiv J_n, k \quad (5.23)
\]

Now let us redraw the figure as indicated in Figure 8.

```
n, k
   . . . . . . .
   . . . . . . .
   . . . . . . .
   . . . . . . .
(1, 1) . . . . . . . (n, 1)
```

Figure 8. Figure 7 redrawn

The original figure has simply been rotated until the left diagonal appears on the bottom and distorted slightly so that diagonals and rows are perpendicular. In this position it is apparent that the total number of arrangements of \( k-1 \) zeros subject to the conditions specified above is the same as the number of distinct paths (strings of ones) from \( (1, 1) \) to \( (n, k) \) subject to the restriction that the horizontal coordinate
of any point on the path is not less than the vertical component. This is closely associated with the ballot problem described by Feller (8) and an exact expression is given by Whitworth (22). As is shown below the only J value we require is \( J_{n, n} \). In this case we have

\[
J_{n, n} = \frac{1}{n} \binom{2n-2}{n-1}
\]

Now we have shown that \( P'(n, k) \equiv J_{n, k} p^n q^{k-1} \) and thus

\[
P_{n+1} = \sum_{k=1}^{n} P'(n, k) \equiv \sum_{k=1}^{n} J_{n, k} p^n q^{k-1} \equiv Q_{n+1}, \text{ say.}
\]

It may be verified that the \( J_{n, k} \) satisfy the recursion relation

\[
J_{n+1, k+1} = J_{n, k+1} + J_{n+1, k}.
\]

Substituting this in the formula for \( Q_n \) and simplifying we find that

\[
Q_{n+1} = Q_n - p^n q^{n+1} J_{n+1, n+1}
\]

and thus

\[
Q_{n+1} = Q_2 - \sum_{j=2}^{n} p^j q^{j+1} J_{j+1, j+1}
\]

\[
= p^2 (1+q) - \sum_{j=2}^{n} p^j q^{j+1} \frac{1}{j+1} \binom{2j}{j}.
\]

(5.24)

Letting \( n \) approach infinity and using the identity \( \binom{2j}{j} = 2 \binom{2j-1}{j-1} \) this becomes

\[
Q_\infty = p - \frac{2}{p} \sum_{j=1}^{\infty} (pq)^{j+1} \frac{1}{j+1} \binom{2j-1}{j-1}
\]
\[
= p - \frac{2}{p} \int_0^{pq} \left( \sum_{j=1}^{\infty} (pq)^j \binom{2j-1}{j-1} \right) \ d(pq)
\]

The series \( \sum_{j=1}^{\infty} (pq)^j \binom{2j-1}{j-1} \) is simply the expansion of \( \frac{1}{2} \left[ (1-4pq)^{-1/2} - 1 \right] \). Substituting this and integrating we find, after some reduction, that

\[
Q_\infty = 1 + \frac{(1-4pq)^{1/2} - 1}{2p} = 1 + \frac{\left[(1-2p)^2\right]^{1/2} - 1}{2p}.
\]

Thus we see that if \( p \leq 1/2 \)

\[
Q_\infty = 1 + \frac{1-2p-1}{2p} = 0 \tag{5.25}
\]

while if \( p > 1/2 \)

\[
Q_\infty = 1 + \frac{2p-1-1}{2p} = 1 - \frac{q}{p} \tag{5.26}
\]

Hence this bound for \( P_\infty \) is precisely the same as that obtained in the previous discussion using arrangements involving two rows.
6.0 SUMMARY AND CONCLUSIONS

6.1 The Problem of Maximum Population Size

In this thesis we have considered two problems which arise in the theory of stochastic branching processes. The first of these was concerned with the maximum population size in a branching process. It was found that by considering the branching process as a special case of a Markov chain with a denumerable number of states it was possible to derive an exact expression for the probability that the population size did not exceed a given fixed value, say \( k \). This in turn led to an expression for the exact distribution of the maximum population size. Examples were given using the geometric distribution which showed that even for small values of \( k \), the probability that the population size did not exceed \( k \) was quite close to the theoretical limiting value. This was interpreted as implying that in most cases the population size either remains small or becomes indefinitely large.

Using the expressions for the derived probabilities we were then able to discuss properties of a branching process when it is known that population size is bounded. In this case we found that the theory of finite absorbing Markov chains applied and certain results were obtained from this correspondence. Other results of interest only in the branching process were also developed. In particular, expressions were derived for the mean and variance of the number of generations
to extinction and of the total number of individuals in the process. Finally, this approach was used to arrive at another expression for the distribution of maximum population size.

The form of the results obtained is such that extensive computations are required in order to obtain numerical results. For each value of \( k \) it is necessary to invert a \( k \times k \) matrix. However, since convergence to limiting values is rapid, it is not necessary in many cases to go above \( k = 10 \). For applications, then, the solution will probably not be of great value. For theoretical purposes, the solution is quite simple and may be useful in other related investigations.

6.2 The Grid Problem

The second problem was concerned with the probability of passing through an infinite grid of connected points. The points were arranged in rows in such a way that the \( n \)th row contained \( n \) points denoted \((n, 1), (n, 2), \ldots, (n, n)\). Point \((n, k)\) in row \( n \) was connected to points \((n+1, k)\) and \((n+1, k+1)\) in row \( n+1 \). Each point, independently of all other points, had common probability \( p \) of being open and probability \( q = 1 - p \) of being blocked. The probability in question, then, is \( P_\infty \), the probability that, beginning at point \((1, 1)\), we can find a non-terminating open path through the grid. This problem was shown to be a special case of the general type of stochastic process
known as a percolation process. It was also interpreted as a branching process in which there is a dependence among particles.

The original problem was not solved completely but it was shown that the critical value of $p$ is not less than two-thirds; i.e., that $p$ must be greater than two-thirds in order for $P_\infty$ to be non-zero. An upper bound for $P_\infty$ as a function of $p$ was also obtained. A number of variations, providing both upper and lower bounds for $P_\infty$, were considered. One of these described the behavior of a process which was similar to that of a nuclear reactor.

6.3 Recommendations

The percolation process and the problem of dependence in a branching process are both problems of some importance in the theory of stochastic processes. The grid problem provides a link between these two areas and thus may aid in the study of both. Certain aspects of the problem deserve to be pursued further. It is felt that there exists a "right" approach which will provide the complete solution to the problem. This may consist of another way of attacking the stated problem or may provide an equivalent restatement of the problem. One such restatement was indicated. If this solution is obtained it will conceivably be of value in studying the more general problem involving the whole plane rather than just the first quadrant.
The general problem of dependence has only been touched on here. It is necessary to extend these methods to more general branching processes, and to arrive at a systematic way of describing dependence in any process. One such way might consist in assigning a measure of dependence between every pair of particles which would depend on the strength of the family line relationship between them. This might be made to depend on the number of generations back to the first common ancestor.
7.0 LIST OF REFERENCES


8.0 APPENDIX A

THE DETERMINANT OF THE MATRIX $Q_k$

The following theorem shows that the matrix $Q_k$ will be non-singular if and only if the probability $p_1$ of a particle producing exactly one new particle is not zero.

Theorem 8.1. The determinant of $Q_k$ is given by

$$\left| Q_k \right| = p_1^2 \frac{k(k+1)}{2}.$$  

Proof. To facilitate the proof we introduce the following definition. Let $t_{jr}$ be the probability that $j$ particles in any generation produce $r$ particles in the next generation in such a way that no subgroup of less than $j$ particles produces all $r$ particles; i.e., so that each of the $j$ particles produces at least one new particle.

Obviously $t_{1r} = p_{1r} = p_r$,

$$t_{rr} = p_1^r$$

and $t_{jr} = 0$ for $j > r$.

From the definition it is easily seen that for all $j$ and $r$

$$t_{jr} = p_{jr} - t_{1r} p_0^{j-1} \binom{j}{1} - \cdots - t_{j-1,r} p_0 \binom{j}{j-1}. \tag{1}$$

From this equation (taking $j = 2$ and using the fact that $p_{1r} = t_{1r}$) we see that by multiplying row one of $Q_k$ by $2p_0$ and subtracting from row two we can introduce $t_{2r}$ in place of $p_{2r}$ ($r = 1, 2, \ldots, k$). This operation will not change the determinant of $Q_k$. Now assuming the
values \( t_{jr} \) have been introduced by row operations in rows 2, 3, \ldots, \( j - 1 \) (note that this is originally true for row one) we see from the general relation (1) that it is then possible for row \( j \) also. None of the operations involved changes the value of the determinant.

Thus we have shown that

\[
\begin{vmatrix}
  p_{11} & \cdots & p_{1k} \\
  \vdots & \ddots & \vdots \\
  p_{k1} & \cdots & p_{kk}
\end{vmatrix}
= \begin{vmatrix}
  t_{11} & \cdots & t_{1k} \\
  \vdots & \ddots & \vdots \\
  t_{k1} & \cdots & t_{kk}
\end{vmatrix}
\]

\[
\begin{vmatrix}
p_{1} & t_{12} & \cdots & t_{1k} \\
p_{1}^{2} & \cdots & t_{2k} \\
\vdots & \ddots & \vdots \\
o & \cdots & p_{1}^{k}
\end{vmatrix}
= p_{1} p_{1}^{2} \cdots p_{1}^{k} = p_{1}^{k(k+1)/2}
\]

and the theorem is proved.
9.0 APPENDIX B

ALGEBRAIC PROPERTIES OF THE MATRIX $N_k$

Theorem 9.1. \(|I - Q_k| > 0\) and \(|N_k| = |(I - Q_k)^{-1}| > 0\).

This theorem may be proved in various ways, some direct and some using the series representation $N_k = \sum_{n=0}^{\infty} Q_k^n$. The first simple proof was given by Frobenius (9) for the case $Q_k > 0$. The result was later extended to non-negative $Q_k$ (see e.g. Fan (6)).

Theorem 9.2. The $ij$ entry in $N_k$, $n_{ij}(k)$, is such that

$$n_{ii}(k) \equiv \left( \frac{s_o}{p_0} \right)^i$$

while

$$n_{ij}(k) \equiv \frac{s_o^i - p_0^i}{p_0^j} \quad i \neq j .$$

Proof. From result (4.5) and the discussion following,

$$p_i(k) = \sum_{j=1}^{k} n_{ij}(k) p_0^j \leq s_o^i .$$

Since $n_{ij}(k) \geq 0$ for all $i$ and $j$ and $p_0 > 0$, we have $n_{ii}(k) p_0^i \leq s_o^i$ and the first result follows. Moreover, $n_{ii}(k) \equiv 1$ so we may write

$$p_i(k) = \sum_{j \neq i} n_{ij}(k) p_0^j + n_{ii}(k) p_0^i \leq s_o^i$$

or

$$\sum_{j \neq i} n_{ij}(k) p_0^j \leq s_o^i - n_{ii}(k) p_0^i \leq s_o^i - p_0^i$$

and thus obtain the second result.
This theorem actually provides a poor bound for the values $n_{ij}(k)$. However, since an indication is given of possible values of these numbers, it is helpful in setting the problem up for computation.
APPENDIX C

LIST OF NOTATION

Notation is listed only if it appears in more than one context.

Page numbers indicate the first introduction of the symbol.

Standard Notation

\[ \{ a_k \} \quad \text{The sequence of numbers } a_1, a_2, \ldots \]

\[ A \land B \quad \text{The logical conjunction of the statements } A \text{ and } B. \]

\[ \text{col}(x_1, x_2, \ldots, x_k) \quad \text{Column vector with entries } x_1, x_2, \ldots, x_k. \]

\[ \delta_{ij} \quad \text{The Kronecker delta.} \]

\[ E(X) \quad \text{The expected value of the random variable } X. \]

\[ \binom{n}{k} \quad \text{Binomial coefficient.} \]

\[ \Pr(A) \quad \text{The probability of the event } A. \]

\[ \Pr(A/B) \quad \text{The conditional probability of } A \text{ given } B \text{ has occurred.} \]

\[ \text{Var}(X) \quad \text{The variance of the random variable } X. \]

Special Notation

\[ A(dg) \quad \text{Matrix obtained from } A \text{ by setting all off diagonal terms equal to zero}. \]

\[ A_{sq} \quad \text{Matrix whose entries are the squares of the corresponding elements in the matrix } A. \]

\[ b_1 \quad \text{Upper bound on } P_\infty. \]
\( b_n \) Probability of passing through the grid. 45

\( B(z) \) Generating function of the \( b_n \). 46

\( f(s) \) Generating function of the \( p_k \). 3

\( G \) Total number of particles in all generations. 25

\( J_k \) \( \text{col}(1, 2, \ldots, k) \). 25

\( G_k \) \( \text{diag}(1, 2, \ldots, k) \). 28

\( l_k \) Column vector of \( k \) one's. 24

\( 1 - p/q \) First upper bound for \( P_\infty \). 35

\( m \) The expected value of \( Z_1 \). 8

\( N_k \) \( (I - Q_k)^{-1} \). 16

\( n_{ij}(k) \) The \( ij \) entry in \( N_k \). 16

\( n_j \) Number of times the population size is \( j \). 24

\( N^*_k \) \( (I - Q^*_k)^{-1} \). 23

\( n^*_{ij} \) The \( ij \) entry in \( N^*_k \). 24

\( (n, k) \) The \( k \)th point in row \( n \) of the grid. 32

\( P \) Same as \( P_\infty \). 51

\( P_1(k) \) Probability that population size never exceeds \( k \) given \( Z_0 = i \). 15

\( P_k \) Probability that one particle produces \( k \) new particles. 1

\( \mathcal{P}(k) \) \( \text{diag}\left[p_1(k), p_2(k), \ldots, p_k(k)\right] \). 21

\( P_n \) The probability of reaching row \( n \) of the grid. 33
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(n,k)$</td>
<td>The probability of reaching point $(n,k)$.</td>
</tr>
<tr>
<td>$P_\infty$</td>
<td>$\lim_{n \to \infty} P_n$.</td>
</tr>
<tr>
<td>$Q_k$</td>
<td>k$x$k matrix.</td>
</tr>
<tr>
<td>$Q^*_k$</td>
<td>k$x$k matrix.</td>
</tr>
<tr>
<td>$R_k$</td>
<td>k$x$1 vector.</td>
</tr>
<tr>
<td>$S_k$</td>
<td>The statement - all $Z_n$ are less than or equal to $k$.</td>
</tr>
<tr>
<td>$s_o$</td>
<td>The smallest positive root of the equation $f(s) = s$.</td>
</tr>
<tr>
<td>$T(p)$</td>
<td>Upper bound for $P_\infty$.</td>
</tr>
<tr>
<td>$Z_n$</td>
<td>Number of particles in the $n$th generation.</td>
</tr>
<tr>
<td>$Z_0$</td>
<td>Initial number of particles.</td>
</tr>
</tbody>
</table>