TECHNICAL REPORT NO. 134

July 1971

BOUNDS AND OPTIMAL STRATEGIES
FOR STOCHASTIC SYSTEMS

by

Bruce William Turnbull

Prepared under Contract
DA-31-124-ARO-D-474, U.S. Army Research Office-Durham

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government
TABLE OF CONTENTS

CHAPTER I: INTRODUCTION
1.1 A Dam Model 1
1.2 The Theory of Gambling 3
1.3 An Application to a Problem in Biology: Life Historical Consequences of Natural Selection 5
1.4 Summary 6

CHAPTER II: THE BASIC THEORY 8

CHAPTER III: A BRANCHING PROCESS 15
3.1 A Description of the Model 15
3.2 The Total Population Size 17
3.3 The Probability of Extinction and the Expected Time to Extinction 21
3.3.1 Upper Bounds 22
3.3.2 Example 1 (Bounded Litter Sizes) 27
3.3.3 Example 2 (Mean and Variance of the Litter Sizes Constrained) 30
3.3.4 Lower Bounds 37
3.3.5 Example 3 40
3.4 The Probability that a Generation Size Will Exceed a Given Number and the Expected Maximum Generation Size 44

CHAPTER IV: AN ASSOCIATED MULTIPLICATIVE PROCESS 51
4.1 A Description of the Model 51
4.2 The Total Population Size 52
4.3 The Probability of Extinction and the Expected Time to Extinction 54
4.3.1 Upper Bounds 55
4.3.2 Example 1 58
4.3.3 Example 2 60
4.3.4 Lower Bounds 64
4.3.5 Example 3 66
4.4 The Probability that a Generation Size Will Exceed a Given Number and the Expected Maximum Generation Size 68
4.5 The Analogy Between the Branching Process and the Multiplicative Process Models 72
CHAPTER V: A MULTITYPE BRANCHING PROCESS

5.1 The Multitype Galton-Watson Process 74
5.2 The Probability of Extinction (Upper Bounds) and the Expected Time to Extinction (Lower Bounds) 77
5.3 The Probability of Extinction (Lower Bounds) and the Expected Time to Extinction (Upper Bounds) 81

CHAPTER VI: A PROBLEM IN BIOLOGY: SELECTION FOR COLONIZING ABILITY 84

6.1 Introduction 84
6.2 The Model 87
6.3 E-selection: Minimizing the Probability of Extinction 91
6.4 S-selection: Maximizing the Expected Total Population Size 100
6.5 An Illustration: E-selection versus S-selection 108
  6.5.1 A Deterministic Model 108
  6.5.2 A Simple Stochastic Model 110
6.6 Concluding Remarks 112

CHAPTER VII: A DAM MODEL 113

7.1 A Description of the Model 113
7.2 An Exponential House 115
  7.2.1 The Basic Theorem 116
  7.2.2 The Special Case of Normal Inputs; a Release Rule 119
  7.2.3 The Probability of Emptiness Before Overflow 122
7.3 Bounded Inputs 126
  7.3.1 Upper Bounds on the Mean Time to First Emptiness 127
  7.3.2 Lower Bounds on the Mean Time to First Emptiness 130
7.4 An Application to the Theory of Collective Risk 134

REFERENCES 135
CHAPTER I
INTRODUCTION

1.1 A Dam Model

In Moran's finite dam model (1954), the water level $Z_{t+1}$ at time $t+1$ is given by:

$$Z_{t+1} = \min\{(Z_t + I_{t+1} - O_{t+1})^+, b\},$$

where $I_{t+1}$ is the random inflow in the period $(t,t+1)$, $O_{t+1}$ is the demand or desired outflow in $(t,t+1)$, and $b$ is the capacity of the dam. In the equation above we have used the notation "$x^+$" to denote "$\max(x,0)$."

The demands $\{O_t\}$ and the capacity $b$ are controllable parameters. For various values of these parameters, the performance of the dam may be compared using such criteria as the stationary distribution of $\{Z_t\}$ or the distribution of the time to first emptiness $T$. Moran calculated these distributions under the assumptions that $O_t = d$ for all $t$ where $d$ is some known constant, and that the inflows $\{I_t\}$ are independent and identically distributed random variables with a known distribution.

Moran's dam process is a model for many inventory and storage problems. However, it can only be viewed as a first approximation to the reservoir problem, as the assumptions about the inflows $\{I_t\}$ are unrealistic. Firstly, one expects rainfalls to vary with the season and thus if the time periods considered are less than one year, the inflows $\{I_t\}$ would not be identically distributed. This problem has been studied in Lloyd and Odom (1964). Secondly, due to upstream reservoirs, a heavy rainfall in one period may
affect inflows to the dam in several succeeding time periods and thus the inflows will not be independent (see Bhat and Gani 1959). A more realistic assumption is to take into account possible serial correlation among the inputs. For instance Lloyd (1963) suggests a model in which the inflows form a Markov process with known stationary transitions.

Finally, the most impractical assumption is that the probability distributions of the inflows \( \{I_t\} \) are known. These distributions must be estimated from past data of river flows and, in many cases, little or no data exist, say ten or fifteen years at most. Moreover, in situations where a considerable amount of information does exist, records of river flows even thirty years ago may bear little relevance to present day conditions due to topographical changes that occur in the watershed—both manmade and natural.

The problem is how to replace these assumptions with ones that are realistic, yet are not so general that meaningful results cannot be obtained. We will assume only that it is known that the inflow distribution, conditional on the past history of the process, belongs to a given class \( \mathcal{M} \) of probability laws. For instance, \( \mathcal{M} \) might consist of all distributions with mean greater than some real number \( m \). Alternatively, it might be known that the variance or the range of the input random variables are constrained. We will then, for example, find lower bounds on the expected time, \( T \), to first reach a certain critical level (perhaps zero) for a given initial dam content. These bounds will represent the "worst that can happen" and a conservative planner would make his decisions accordingly. This is equivalent to a "minimax strategy" in game theory terminology.

To sum up, previously in the literature, dam models have been studied assuming that the nature and distribution of the input random variables are
fixed and known. (A review of dam models is given in Prabhu 1964, 1965 Chapters 6, 7.) Using such "exact" assumptions, quantities such as the mean time to first emptiness are computed exactly. Here we replace "exact assumptions" with a family of assumptions, and replace "exact answers" with bounds on the quantities under study. The contention is that "far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise." -- Tukey (1962, page 13).

1.2 The Theory of Gambling

The problem posed in §1.1 is similar to the one studied in the theory of optimal gambling, first considered by Blackwell (1954, 1964) and later extensively developed by Dubins and Savage (1965). This theory considers a gambler who, having a fortune $z_n$ at stage $n$, may play any gamble $\gamma$ selected from a specified set $\mathcal{N}(z_n)$ of gambles. Each gamble is a probability measure on the set $C$ of all possible fortunes and $z_{n+1}$, the gambler's fortune at the next stage, is determined following the probabilities in $\gamma$. The gambler's objective is to select a betting strategy, or sequence of gambles, that will maximize his probability of attaining a goal before being ruined. Analogously in the dam problem, $z_n$ represents the water level at time $n$, $C$ is the interval $[0, b]$ of possible water levels, the set of gambles corresponds to the set of possible inflow distributions, and, typically, the objective is to minimize the expected time to first emptiness. This suggests that techniques similar to those used in gambling theory can be applied to the dam problem of §1.1. In fact, as will be shown later, these techniques can also be applied to a wide variety of interesting problems.

Dubins and Savage (1965) treat at length many problems associated with
optimal gambling systems. They proceed under the very general assumption of finite additively gamble defined on all subsets of \( C \). (They give details in Section 2.3.) In this dissertation we shall stay with the traditional assumption of countably additive measures defined on the Borel sets. In addition, our theory will include a general one-stage reward function describing rewards which are accumulated every successive time period until the process is stopped. Dubins and Savage consider only a terminal reward function describing rewards which depend solely on the state of the system at termination \( T \).

Dubins and Savage concentrate mainly on unfavorable games, i.e. those systems where the gambler is ruined with probability one if he plays sufficiently long. Freedman (1967) and Molenaar and Van der Velde (1967) treat the problem of how to gamble so as to delay this fate as long as possible. In favorable games the objective is usually to maximize the probability of avoiding ruination. Results for favorable games have appeared in Breiman (1960a), and the special case of coin-tossing is considered in Ferguson (1965) and Truelove (1970).

Blackwell has obtained results for random sums with the model

\[
Z_n = Z_{n-1} + X_n \quad (n = 1, 2, \ldots),
\]

where \( X_n \) is a random variable with distribution belonging to some set \( \mathcal{M} \) of probability laws. For any positive number \( t \), Blackwell (1954) finds an upper bound for \( \Pr[\sup_{n} (X_1 + \ldots + X_n) > t] \) for a particular set \( \mathcal{M} \), and finds the strategy that achieves the bound. In a later article (1964), he finds sharp upper bounds on the expected time for the random sum \( \{Z_n\} \) to leave a region \( R \) for various sets \( \mathcal{M} \) and regions \( R \). Blackwell's paper does allow a one-stage reward but only in the special case of random sums when this reward is unity.
With a similar setup, Taylor (1971) finds sharp bounds on the expected value of the positive part of the stopped random sum $E[Z^+_1]$, for various sets $\mathcal{N}$. He points out that the bounds have an interpretation in a stock market timing problem. Other related papers include MacQueen (1961), Breiman (1960b), and Dubins and Freedman (1965).

However, none of the papers mentioned in this section really indicates the far-ranging potential applications of this theory of gambling in applied probability. Some of these applications are developed in this dissertation.

1.3 An Application to a Problem in Biology: Life Historical Consequences of Natural Selection

In addition to the dam problem of §1.1, there are other potential applications of the theory of optimal gambling. One of the most interesting of these is the biological problem of optimal life history strategies and modes of natural selection. A full description of this problem appears in Chapter VI but we will give a brief introduction to it here.

Animals from the same species can exhibit a wide variation of their life histories. Gadgil and Bossert (1970) illustrate this with the example of two ecological forms of trout (Salmo trutta). The river form of this species (s. t. fario) lives in the poorer environment, grows to a smaller size, and matures at an age of 3-5 years; whereas the lake form (s. t. lacustris), living in more favorable conditions, grows to a larger size and matures at 5-7 years. An explanation for this phenomenon is that each form is "fittest" in its environment in some sense, and, by natural selection, the "optimum" genotype is the one to predominate. Two measures of fitness that we will consider are the probability of extinction and the expected total population
size. The use of the probability of extinction as a measure of fitness was first proposed by Holgate (1967). Maximizing mean total population size will be shown to be equivalent to maximizing total fecundity (i.e. the expected total number of offspring produced by an individual during its lifetime) and this is related to the growth rate of the population. Many researchers, among them Fisher (1930) and Cole (1954), have considered this growth rate as a measure of fitness.

In Chapter VI, we will examine some mathematical models for life histories, and we will use the techniques of gambling theory to identify those genotypes associated with the optimum life history strategies.

1.4 Summary

In Chapter II, we shall set up the general theory and prove four basic theorems which will be used in the remaining chapters.

In Chapter III, a branching process is considered. Bounds are obtained for (i) the expected total population size, (ii) the probability of extinction, (both for finite and infinite horizons), (iii) the mean time to extinction, (iv) the probability that a generation size exceeds a given number, and (v) the expected maximum generation size. In (i), (ii), and (iii), the optimal strategies which attain the bounds are identified. Some of the results of this chapter were first obtained by Freedman and Purves (1967) and by Goodman (1968). However, here their results are presented as part of a unified theory.

In Chapter IV, a multiplicative process is defined and results analogous to those of Chapter III are obtained.

In Chapter V, a model for a multitype branching process is considered, and bounds are obtained on the probability of extinction with both finite and infinite horizons.
In Chapter VI, a "life history" process is defined as a special case of the multitype branching process considered in the previous chapter, and the problem presented in §1.3 is discussed. Bounds are obtained for the probability of extinction of the species and on the expected total population size, and those genotypes for which the bounds are attained are identified.

In Chapter VII, a dam process is defined and the problem of §1.1 is discussed. Bounds are obtained on the mean time to first emptiness, on the mean time to reach any given critical level, and on the probability of emptiness before overflow under various sets of assumptions on the input distributions. A release rule is discussed.
CHAPTER II
THE BASIC THEORY

Here we will develop the general theory which will be applied to various specific models in the remaining chapters. The concepts of an $\mathcal{M}$-sequence and a stopping time are defined, and so is the reward function which assigns a real number to each stopped $\mathcal{M}$-sequence. This reward function consists of the sum of one-stage rewards and a terminal reward. Four basic theorems give upper and lower bounds for the expected rewards both in the finite and infinite horizon case. The proofs involve the theory of martingales and the results will be familiar to those acquainted with the ideas of dynamic programming.

For random variables $X, Y$ we use the notation $P_X$ or $P(X)$ for the probability distribution of $X$, and $P_X|Y$ or $P(X|Y)$ for the conditional distribution of $X$ given $Y$. Also let $C$ be a Borel subset of some Polish space. In the applications considered in the remaining chapters, this space will be the space of reals or of real-valued $n$-vectors (for some positive integer $n$).

**Definition.** Let $X_1, X_2, \ldots$ be random variables defined on some probability space $(\Omega, \mathcal{B}, \mathbb{P})$ and taking values in $C$. Let $\mathcal{M}$ be a non-empty set of probability distributions on $C$. Then we call $X_1, X_2, \ldots$ an $\mathcal{M}$-sequence if the distributions $P(X_1)$ and $P(X_n|X_1, X_2, \ldots, X_{n-1})$ for all $n = 2, 3, \ldots$ are always in $\mathcal{M}$.

For instance, if $\mathcal{M}$ contains only one element then $X_1, X_2, \ldots$ are independent and identically distributed. However, if $\mathcal{M}$ contains more than one element, in general the sequence is not even Markov.
Definition. Let $Z_0, Z_1, Z_2, \ldots$ be random variables defined on a probability space $(\Omega, \mathcal{B}, P)$ and taking values in $C$. For each $z$ in $C$, let $\mathcal{M}(z)$ be a non-empty set of probability distributions on $C$. Then we call $Z_0, Z_1, Z_2, \ldots$ an $\mathcal{M}$-sequence starting at $z$ if

(a) $Z_0 = z$, and

(b) $P(Z_{n+1} \mid Z_0, Z_1, \ldots, Z_n) \in \mathcal{M}(Z_n)$ for every $n = 0, 1, 2, \ldots$.

For example, if $\{X_n\}$ $(n = 1, 2, \ldots)$ is an $\mathcal{M}$-sequence of real-valued random variables, $\mathcal{M}(z) = \{P(z+X) : P(X) \in \mathcal{M}\}$, and $Z_n = z + X_1 + \ldots + X_n$, then $\{Z_n\}$ is an $\mathcal{M}$-sequence starting at $z$.

Definition. A stopping time relative to $\{Z_n\}$ is a random variable, $T$, taking values $\{0, 1, 2, \ldots, \infty\}$, for which, for every $n$, the event $\{T = n\}$ is in the $\sigma$-algebra generated by $Z_0, Z_1, \ldots, Z_n$.

(If, in any particular context, we fail to define a stopping time on part of the sample space, then we take its value to be infinite on that part.)

Let $r$ be a real-valued Baire function on $C$, called the one-stage reward, and define $T(N) = \min[T, N]$ for $N = 0, 1, 2, \ldots$.

Theorem 2.1

Let $N$ be some positive integer and let $r, \{f_k\}$ $(k = 0, 1, 2, \ldots, N)$ be real-valued non-negative Baire functions on $C$ such that:

\begin{equation}
(2.1) \quad f_{k+1}(z) \geq r(z) + Ef_k(Z)
\end{equation}

whenever $P(Z) \in \mathcal{M}(z)$, for all $z \in C$, and for every $k = 0, 1, 2, \ldots, N-1$.

Then, for $Z_0, Z_1, \ldots$, an $\mathcal{M}$-sequence starting at $z$, we have:
(2.2) \[ f_N(z) \geq E\left[ \sum_{k=0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)}) \right], \]

for all \( z \in C \), and for all stopping times \( T \).

**Proof**

We first show that \( \{f_{N-k}(Z_k) + \sum_{i=0}^{k-1} r(Z_i)\} \) (\( k = 1,2,\ldots,N \)) is a non-negative supermartingale. Note that:

\[
E[f_{N-k-1}(Z_{k+1}) + \sum_{i=0}^{k} r(Z_i) | Z_0, Z_1, \ldots, Z_k] \\
= \sum_{i=0}^{k-1} r(Z_i) + r(Z_k) + E[f_{N-k-1}(Z_{k+1}) | Z_0, Z_1, \ldots, Z_k] \\
\leq \sum_{i=0}^{k-1} r(Z_i) + f_{N-k}(Z_k),
\]

where the inequality follows by condition (2.1). This verifies the supermartingale property and the result (2.2) follows by applying a theorem due to Doob (1953; page 300, Theorem 2.1), with the sign changed.

**Theorem 2.2**

Let \( N \) be some positive integer and let \( r, \{f_k\} \) (\( k = 0,1,2,\ldots,N \)) be real-valued non-negative Baire functions on \( C \) such that:

(2.3) \[ f_{k+1}(z) \leq r(z) + Ef_k(Z), \]

whenever \( P(Z) \in \mathcal{M}(z) \), for all \( z \in C \), and for every \( k = 0,1,2,\ldots,N-1 \).

Then, whenever \( Z_0, Z_1, \ldots \) form an \( \mathcal{M} \)-sequence starting at \( z \) and satisfying:
(2.4) \[ E[\frac{r_{N-k}(Z_k)}{\xi_k} + \sum_{i=0}^{k-1} r(Z_i)] < \infty, \]
for all \( k = 1, 2, \ldots, N \),
we have:

(2.5) \[ f_N(z) \leq E[\sum_{k=0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)})], \]
for all \( z \in \mathbb{C} \), and for all stopping times \( T \).

**Proof**

By conditions (2.3) and (2.4) we have that \( \{f_{N-k}(Z_k) + \sum_{i=0}^{k-1} r(Z_i)\} \)
\( (k = 1, 2, \ldots, N) \) is a submartingale. Then (2.5) follows from a martingale
theorem which can be found in Neveu, (1965; Lemma, page 132).

**Notation**

If \( T = \infty \), and \( a_i \geq 0 \) for all \( i \), define \( \sum_{i=0}^{T} a_i \) to be \( \lim_{N \to \infty} \sum_{i=0}^{N} a_i \),
which may possibly take the value \( +\infty \). Also let \( I_B \) denote the indicator
random variable of the event \( B \).

**Theorem 2.3**

Let \( r, f \) be real-valued non-negative Baire functions on \( \mathbb{C} \) satisfying:

(2.6) \[ f(z) \geq r(z) + Ef(Z), \]

whenever \( P(Z) \in \mathcal{M}(z) \) and for all \( z \in \mathbb{C} \).

Then for \( Z_0, Z_1, \ldots \), an \( \mathcal{M} \)-sequence starting at \( z \), we have:
(2.7) \[ f(z) \geq E\left[ \sum_{k=0}^{T-1} r(Z_k) + f(Z_T) \cdot I_{T<\infty} \right], \]

for all \( z \in C \) and for all stopping times \( T \).

(Note: For the sake of completeness we may define \( f(Z_T) = 0 \) if \( T = \infty \); i.e. there is no terminal reward if the process never stops.)

Proof

Using Theorem 2.1 with \( f_k = f \) for all \( k \), we have, for each \( N = 0, 1, 2, \ldots \):

\[ f(z) \geq E\left[ \sum_{k=0}^{T(N)-1} r(Z_k) + f(Z_{T(N)}) \right] \]

\[ \geq E\left[ \sum_{0 \leq k < T(N)} r(Z_k) + f(Z_{T(N)}) \cdot I_{T<N} \right] \]

\[ = E\left[ \sum_{0 \leq k < T(N)} r(Z_k) + f(Z_T) \cdot I_{T<N} \right] \]

\[ \Rightarrow E\left[ \sum_{0 \leq k < T} r(Z_k) + f(Z_T) \cdot I_{T<\infty} \right] \text{ as } N \to \infty, \]

by monotone convergence, since \( r, f \) are non-negative. Hence the result (2.7) is proved. \( \square \)

Theorem 2.4

Let \( r, f \) be non-negative Baire functions on \( C \) such that

(2.8) \[ f(z) \leq r(z) + Ef(Z), \]
whenever $P(Z) \in \mathcal{M}(z)$ and for all $z \in C$. Then, whenever $Z_0, Z_1, \ldots$ form an $\mathcal{M}$-sequence starting at $z$, and $T$ is any stopping time satisfying

\begin{enumerate}
\item[(a)] $E[f(Z_k)] + \sum_{i=0}^{k-1} r(Z_i) < \infty$ for $k = 1, 2, \ldots$
\item[(b)] $\lim \inf_{N \to \infty} E[f(Z_N) \cdot I_{T>N}] = 0$,
\end{enumerate}

(2.9)

we have:

\begin{equation}
\tag{2.10}
f(z) \leq E[\sum_{k=0}^{T-1} r(Z_k) \cdot f(Z_T) \cdot I_{T<\infty}],
\end{equation}

for all $z \in C$.

**Proof**

Using (2.8), (2.9a), we may apply Theorem 2.2 with $f_k = f$ for all $k$ and obtain:

\[ f(z) \leq E[\sum_{k=0}^{T(N)-1} r(Z_k) + f(Z_{T(N)})] \]

\[ = E[\sum_{k=0}^{T(N)-1} r(Z_k) + f(Z_{T(N)}) \cdot I_{T<N}] + E[f(Z_{T(N)}) \cdot I_{T\geq N}] . \]

This is true for all $N$. Letting $N \to \infty$, we have:

\[ f(z) \leq E[\sum_{k=0}^{T-1} r(Z_k) + f(Z_T) \cdot I_{T<\infty}] + \lim \inf_{N \to \infty} E[f(Z_{T(N)}) \cdot I_{T\geq N}] . \]

The result (2.10) now follows by condition (2.9b).
Remark In this chapter we have not concerned ourselves with the question of the existence of functions \( r, f, \{f_k\} \); and we have assumed they do exist. In the applications in the following chapters, the existence of these functions and corresponding \( \mathcal{M} \)-sequences will be demonstrated.
CHAPTER III

A BRANCHING PROCESS

In this chapter we define a branching process, a special case of which is the Galton-Watson process, defined in Harris (1963, Chapter 1) and originally in Galton and Watson (1874). Using the theory developed in Chapter II, we derive results about quantities of interest such as the mean population size, the probability of extinction, the mean time to extinction, the probability that a generation size will exceed a given number, and the expected maximum generation size.

3.1 A Description of the Model

Let $\mathcal{M}$ be a set of probability distributions concentrated on the non-negative integers.

For each non-negative integer $z$, define a sequence of random variables $Z_0, Z_1, \ldots$ as follows:

\begin{align*}
Z_0 &= z \\
Z_n &= X(n,1) + X(n,2) + \ldots + X(n,Z_{n-1}) \quad (n = 0, 1, 2, \ldots),
\end{align*}

where, "conditional on the past," \{X(n,i): i = 1, 2, \ldots, Z_{n-1}\} are independent and with an identical distribution in $\mathcal{M}$ for each $n$.

By "conditional on the past," we mean "conditional on the random variables $\{Z_i, X(i,j): j = 1, 2, \ldots, Z_i\}, i = 1, 2, \ldots, n-1."$ It is understood that if $Z_{n-1} = 0$ then $Z_n = 0$, i.e. the state 0 is absorbing. If $Z_n = 0$, we say the process has become extinct.
More precisely, in terms of $\mathcal{M}$-sequences, we assume that $Z_0, Z_1, \ldots$ is an $\mathcal{M}$-sequence starting at $z$ for

$$(3.2) \quad \mathcal{M}(z) = \{P(X_1 + \ldots + X_z) : \{X_i\} \text{ are i.i.d.}, P(X_i) \in \mathcal{M}\},$$

where "i.i.d." is an abbreviation for "independent and identically distributed."

This represents a simple population growth model, in which $X(n,j)$ is the number of offspring of the $j$'th individual of the $(n-1)$th generation, and $Z_n$ is the size of the $n$'th generation.

In the classic model of Galton and Watson (1874), the litter size distributions in all generations are i.i.d. This corresponds to our model for the case when $\mathcal{M}$ contains exactly one element. The recent theory of "branching processes in random environments" (see Smith 1968, Smith and Wilkinson 1969, and Athreya and Karlin 1970) treats the case when the litter sizes of individuals in the $n$'th generation are conditionally i.i.d. with distribution function $\xi_n$, where $\{\xi_i; i \geq 0\}$ is a stationary random process and $\xi_n$ is independent of the past history of the process ($n = 0, 1, 2, \ldots$). This becomes the Galton-Watson process for the special case when, for all $n$, $\xi_n = \xi$ with probability one. For these models, exact results have been obtained about limiting distributions and conditions for almost sure convergence. In our model we also permit the litter size distributions to be themselves random but, in addition, they are allowed to depend on the past history of the process. We specify only that they belong to some class $\mathcal{M}$. Using the theorems of Chapter II, we then compute bounds on the mean population size, the probability of extinction, the expected time to extinction, the probability that a generation size will exceed a given number, and the expected maximum generation size.
We will need to assume the well-known result that if $\phi$ is the generating function of a non-negative integer-valued random variable $X$, then, provided $\Pr[X > 1] > 0$, the equation $x = \phi(x)$ has exactly two roots in $[0, \infty)$, namely $x = 1$ and $x = \alpha$ where $\alpha > 1$, $\alpha = 1$, $\alpha < 1$ according as $E(X) < 1$, $E(X) = 1$, or $E(X) > 1$. In the case $\Pr[X > 1] = 0$, then if $\Pr[X = 1] < 1$ the only root of the equation $x = \phi(x)$ is $x = 1$, while if $\Pr[X = 1] = 1$ then $\phi(x) \equiv x$ for all $x \geq 0$. This result follows from the convexity of $\phi$ and is proved for instance in Karlin (1966, Chapter 11.3).

Throughout this chapter, the set $\mathbb{C}$, referred to in Chapter II, will be taken to be the set of all non-negative integers. Also it will be understood that $\mathcal{M}$ is some subset of probability distributions concentrated on $\mathbb{C}$. We will adopt the convention $0^0 = 1$, and define the time to extinction, $T_e$, by:

$$T_e = \min \{n: Z_n = 0\} \text{ if } Z_n = 0 \text{ for some } n,$$

$$= \infty \text{ otherwise.}$$

3.2 The Total Population Size

In this section, we consider families, $\mathcal{M}$, of distributions for which the means are restricted, and find bounds on the expected total population size, $E[\sum_{i=0}^{N} Z_i]$, for $N$ both finite and infinite. It will turn out that these bounds are sharp and all are achieved by the same Galton-Watson process.

Theorem 3.1 (Upper bounds)

Take $m > 0$, let

$$\mathcal{M}_m = \{P(X): EX \leq m\}$$
and set

\[ M(z) = \{P(X_1 + \ldots + X_z): \{X_i\} \text{ i.i.d., } P(X_i) \in M, \ i = 1,2,\ldots,z.\} \]

for \( z = 0,1,2,\ldots \).

Then for \( Z_0, Z_1, \ldots \), an \( M \)-sequence starting at \( z \), we have:

\[
E[\sum_{i=0}^{N} Z_i] \leq \begin{cases} 
  z \cdot \frac{1-m^{N+1}}{1-m} & \text{if } m \neq 1 \\
  z \cdot (N+1) & \text{if } m = 1
\end{cases}
\]

for \( N = 0,1,2,\ldots \),

and

\[
E[\sum_{i=0}^{\infty} Z_i] \leq \begin{cases} 
  z \cdot \frac{1}{1-m} & \text{if } m < 1 \\
  \infty & \text{if } m > 1
\end{cases}
\]

Furthermore, these bounds are achieved when \( Z_0, Z_1, \ldots \) form a Galton-Watson process with Malthusian rate (mean litter size) equal to \( m \).

**Proof**

Consider first \( m \neq 1 \). We apply Theorem 2.1 with \( C \) the set of non-negative integers, \( r(z) = z, T = N \), and \( f_k(z) = z \cdot (1-m^{k+1})/(1-m) \) for \( z \in C \) and \( k = 0,1,\ldots,N \). Then \( r \) and \( f_k \) are non-negative and for \( P(Z) \in M(z) \):
\[ r(z) + Ef_k(Z) = z + E \left( \frac{1-m}{1-m} \cdot Z \right) \]

\[ = z + E \left( \frac{1-m}{1-m} \cdot (X_1 + \ldots + X_z) \right) \]

\[ \leq z + \frac{1-m}{1-m} \cdot m \cdot z \]

\[ = z \cdot \frac{1-m}{1-m} \cdot k+2 \]

\[ = f_{k+1}(z) \]

which verifies the hypotheses of Theorem 2.1. The result (3.3) now follows for \( m \neq 1 \) by noting that:

\[ \sum_{k=0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)}) = \sum_{k=0}^{N-1} r(Z_k) + f_0(Z_N) \]

\[ = \sum_{k=0}^{N} Z_k . \]

For the case \( m = 1 \) the proof is similar but Theorem 2.1 is applied with \( f_k(z) = (k+1) \cdot z \) for \( z \in \mathbb{C} \) and \( 0 \leq k \leq N \). Now, for \( P(Z) \in \mathbb{M}(z) \), we have:

\[ r(z) + Ef_k(Z) = z + E[(k+1) \cdot Z] \]

\[ = z + (k+1) \cdot E[X_1 + \ldots + X_z] \]

\[ \leq z + (k+1) \cdot z \]

\[ = (k+2) \cdot z \]

\[ = f_{k+1}(z) . \]
Hence the result (3.3) is proved for all positive $m$.

For $m < 1$, we have

$$E[\bigcap_{k=0}^{N} Z_k] \leq z \cdot \frac{1-m}{1-m} \leq z \cdot \frac{1}{1-m}.$$  

This is true for all $N$ and so by monotone convergence the result (3.4) follows. For $m \geq 1$, (3.4) is true trivially.

Finally, the proof that these bounds are attained by a Galton-Watson process with Malthusian rate $m$ can be found in any standard text on stochastic processes. (See, for instance, Harris 1963, Chapter 1.5 or Karlin 1966, Chapter 11.2.)

**Theorem 3.2** (Lower bounds)

Take $m > 0$, let

$$\mathcal{M} = \{P(X): EX \geq m\}$$

and set

$$\mathcal{M}(z) = \{P(X_1 + \ldots + X_z): \{X_i\} \text{ are i.i.d., } P(X_i) \in \mathcal{M}\}.$$  

Then for $z_0, z_1, \ldots$, an $\mathcal{M}$-sequence starting at $z$, we have:

$$E[\bigcap_{i=0}^{N} Z_i] \geq \begin{cases} 
  z \cdot \frac{1-m^{N+1}}{1-m} & \text{if } m \neq 1 \\
  z \cdot (N+1) & \text{if } m = 1
\end{cases}$$

for $N = 0, 1, 2, \ldots$.  


and

\[ E[\sum_{i=0}^{\infty} Z_i] \geq z \cdot \frac{1}{1-m} \quad \text{if} \quad m < 1 \]

\[ = \infty \quad \text{if} \quad m \geq 1. \]

(3.6)

These bounds are achieved when \( Z_0, Z_1, \ldots \) is a Galton-Watson process with Malthusian rate \( m \).

**Proof**

To obtain (3.5), we apply Theorem 2.2 in the same way that Theorem 2.1 was applied in the proof of Theorem 3.1. Theorem 2.2 holds only for those \( \mathcal{M} \)-sequences for which

\[ E[\sum_{i=0}^{k} f_k(Z_i) + \sum_{i=0}^{k-1} r(Z_i)] < \infty \quad (k = 1, 2, \ldots, N). \]

(3.7)

However, (3.5) is still true for those \( \mathcal{M} \)-sequences starting at \( z \) which do not obey the above condition (3.7), since then we must have \( E[\sum_{i=0}^{N} Z_i] = \infty \). The result (3.6) is obtained by letting \( N \to \infty \) first in the left hand side of (3.5) and then in the right hand side of (3.5). Attainment of bounds follows as in Theorem 3.1.

3.3 The Probability of Extinction and the Expected Time to Extinction

In this section, we consider certain families \( \mathcal{M} \) of distributions and find bounds on the probability of extinction of the branching process described in §3.1. In §3.3.1, upper bounds are derived for an \( \mathcal{M} \) which may at first glance appear somewhat artificial. However two interesting special cases are considered in §3.3.2 and §3.3.3; in the former the litter sizes
are bounded, in the latter the variances of the litter size distributions are constrained. Lower bounds are derived in §3.3.4 and a special case also studied by Freedman and Purves (1967) and by Goodman (1968) is presented in §3.3.5. (Goodman also treated the example of §3.3.2.) Finally, results about the mean time to extinction are proven. All the bounds obtained are sharp and are achieved by Galton-Watson processes.

3.3.1 Upper Bounds

Theorem 3.3 (Extinction in a finite time N)

For any positive integer N, let $\gamma_0 = 0, \gamma_1, \gamma_2, \ldots, \gamma_N$ be a sequence of non-negative real numbers. Take

$\mathcal{M} = \{P(X): E[\gamma_j^X] \leq \gamma_{j+1}, \text{ for } 0 \leq j \leq N-1\}$

and set

$\mathcal{M}(z) = \{P(X_1 \ldots X_z): \{X_i\} \text{ are i.i.d., } P(X_i) \in \mathcal{M}\}.$

Assume $\mathcal{M}$ is non-empty. Then for $Z_0, Z_1, \ldots$, an $\mathcal{M}$-sequence starting at $z$, we have:

$$(3.8) \quad Pr[Z_N = 0] \leq (\gamma_N)^Z.$$

Proof

We apply Theorem 2.1 with $C =$ the set of non-negative integers, $r(z) = 0$, $T = N$, $f_k(z) = (\gamma_k)^z$ for $z \in C$ and $0 \leq k \leq N$. Then $r$ and $f_k$ are non-negative and for $P(Z) \in \mathcal{M}(z)$:
\[ \begin{align*}
    r(z) + E f_k(Z) &= E \left[ \gamma_k^{X_1 + \ldots + X_z} \right] \\
    &= \{E[\gamma_k^{X_1}]\}^z \quad \text{(since \( \{X_i\} \) are conditionally i.i.d.)} \\
    &\leq (\gamma_{k+1})^z \quad \text{(by definition of \( \mathcal{M}_0 \))} \\
    &= f_{k+1}(z),
\end{align*} \]

which verifies the hypotheses of Theorem 2.1. The result (3.8) now follows by noting that:

\[ E[\sum_{k=0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)})] = E[f_0(Z_N)] \]

\[ = E[\gamma_N^Z] \]

\[ = \Pr[Z_N = 0], \]

since \( \gamma_0 = 0 \) and we have adopted the convention \( 0^0 = 1. \)

Corollary 1 (The Expected Time to Extinction)

Suppose \( \gamma_0 = 0, \gamma_1, \gamma_2, \ldots \) is a sequence of real numbers in \( [0,1] \).

Take

\[ \mathcal{M}_0 = \{P(X): E[\gamma_j^X] \leq \gamma_{j+1}, \quad \text{for all} \quad j \}. \]
Then

\[ E[T_e] \geq \sum_{i=0}^{\infty} [1 - (\gamma_i)^z], \]

where \( T_e \) is the time to extinction as defined in \( \S 3.1 \).

**Note** Since \( 0 < \gamma_i < 1 \) for all \( i \) and \( z > 0 \), all terms in the infinite sum in (3.9) are non-negative. If \( \gamma_i \) does not tend to a limit or tends to some limit other than one then the sum is infinity. If \( \gamma_i \to 1 \) then the sum may be finite or infinite.

**Proof**

The result (3.9) is obtained by noting that

\[ E[T_e] = \sum_{i=0}^{\infty} \text{Pr}[T_e > i] \]

\[ = \sum_{i=0}^{\infty} \text{Pr}[Z_i > 0] \]

\[ = \sum_{i=0}^{\infty} (1 - \text{Pr}[Z_i = 0]) \]

\[ \geq \sum_{i=0}^{\infty} [1 - (\gamma_i)^z] \]

\[ \square \]

**Corollary 2  (Achievement of Bounds)**

If there exists a Galton-Watson process with litter sizes \( X \) such that \( E[X_j] = \gamma_{j+1} \), for all \( j \), then this achieves the bounds (3.8), (3.9). This follows by Harris (1963; Chapter 1, Theorem 6.1).
Note on Existence  For general \{\gamma_i\}, \mathcal{M} may be empty and no Galton-Watson process defined as above will exist. However, in each of two special cases studied in §3.3.2 and §3.3.3, a certain interesting family of distributions is defined and the \{\gamma_i\} are chosen in such a way that this family will be a subset of \mathcal{M} and that the Galton-Watson process of Corollary 2 will exist.

Theorem 3.4  (Eventual extinction)

Take \alpha > 0, let

\[ \mathcal{M} = \{P(X): E[\alpha^X] \leq \alpha\} \]

and set

\[ \mathcal{M}(z) = \{P(X_1 + \ldots + X_z): \{X_i\} \text{ are i.i.d., } P(X_1) \in \mathcal{M}\} \]

Then, for \(Z_0, Z_1, \ldots\), an \(\mathcal{M}\)-sequence starting at \(z\), we have:

\[ (3.10) \quad \Pr[Z_N = 0 \text{ for some } N] \leq \alpha^z. \]

Proof

We apply Theorem 2.3 with \(C\) the set of non-negative integers, \(r(z) = 0\), \(f(z) = \alpha^z\) for \(z \in C\) and

\[ T = \begin{cases} \min\{n: Z_n = 0\} & \text{if } Z_n = 0 \text{ for some } n \\ \infty & \text{otherwise} \end{cases} \]
Then \( r \) and \( f \) are non-negative, and for \( P(Z) \in \mathcal{M}(z) \):

\[
    r(z) + Ef(Z) = E[\alpha^X_1 + \ldots + X_z] \\
    = \{E[\alpha^X_1]\}^z \\
    \leq \alpha^z \\
    = f(z),
\]

which verifies the conditions of Theorem 2.3. Thus \( f(z) \geq Ef(Z_T) \).

Let \( v(z) = \begin{cases} 1 & \text{if } z = 0 \\ 0 & \text{otherwise} \end{cases} \), then since \( f(z) \geq v(z) \) for all \( z \in C \), we have:

\[
    f(z) \geq Ef(Z_T) \\
    \geq Ev(Z_T) \\
    = \int_{T<\infty} v(Z_T) dP \\
    = Pr[T < \infty] \\
    = Pr[Z_N = 0 \text{ for some } N].
\]

This proves (3.10). \( \Box \)
Corollary 1  (Achievement of Bounds)

If there exists a Galton-Watson process with litter sizes \( X \) such that \( E[\alpha^X] = \alpha \), then this process achieves the bound (3.10). This follows by Harris (1963; Chapter 1, Theorem 6.1).

3.3.2 Example 1  (Bounded Litter Sizes)

Let \( k \) be some positive integer and \( m \) a real number such that \( 0 \leq m \leq k \). We will derive an upper bound on the probability of extinction and a lower bound on the expected time to extinction for the branching process described in §3.1 for which the litter sizes \( X \), conditional on the past, are constrained to be at most \( k \) and with mean at least \( m \). It will turn out that this is a special case of the \( M \)-sequences studied in §3.3.1 if \( \alpha \) and \( \{\gamma_k\} \) are chosen appropriately.

Lemma 3.1

Let \( M = \{P(X): EX \geq m, \ 0 \leq X \leq k\} \). Then for \( 0 \leq \beta \leq 1 \) and \( P(X) \in M \), we have:

\[
E[\beta^X] \leq 1 - m/k + \beta^k \cdot m/k .
\]

Proof

Define \( f(x) = 1 - (1-\beta^k) \cdot x/k \) and \( g(x) = \beta^x \).

If \( \beta = 0 \), then \( f(x) = 1 - x/k \) and \( g(x) = 1 \) for \( x = 0 \), \( g(x) = 0 \) for \( x > 0 \). Hence \( g(x) \leq f(x) \) for \( 0 \leq x \leq k \).

If \( 0 < \beta < 1 \) then since \( f \) is linear, \( g \) is convex and \( f(0) = g(0) = 1, f(k) = g(k) = \beta^k \), we again have that \( g(x) \leq f(x) \) for \( 0 \leq x \leq k \). (See Figure 3.1.)
Figure 3.1

Hence for $P(X) \in M$, we have $Eg(X) \leq Ef(X)$ or:

$$E[b^X] \leq E[1 - (1 - b^k) \cdot X/k]$$

$$\leq 1 - (1 - b^k) \cdot m/k,$$

which completes the proof of Lemma 3.1.

Theorem 3.5

Let $M = \{P(X): EX \geq m, 0 \leq X \leq k\}$ and set $M(z) = \{P(X_1 + \ldots + X_z): \{X_i\}_{i=1}^z\}$ are i.i.d., $P(X_1) \in M$. Then for $Z_0, Z_1, \ldots$ an $M$-sequence starting at $z$, we have:

A) $Pr[Z_N = 0] \leq (\gamma_N)^Z$, $(N = 0, 1, 2, \ldots)$,

where $\gamma_N$ is defined recursively by:
\[ \gamma_0 = 0 \]
\[ \gamma_{j+1} = (1 - m/k) + \gamma_j^k \cdot (m/k) \quad j = 0, 1, \ldots, N - 1; \]

B) \( \Pr[Z_N = 0 \text{ for some } N] \leq \alpha^Z \),

where \( \alpha \) is the smaller root in \([0,1]\) of:
\[ \alpha = (1 - m/k) + \alpha^k \cdot (m/k); \]

C) \[ \mathbb{E}[T_e] \geq \sum_{i=0}^{\infty} [1 - (\gamma_i)^z], \]

where \( \{\gamma_i, i \geq 0\} \) are defined as in (A);

D) The bounds in (A), (B), (C) are sharp and are attained when \( \{Z_i\} \)
form a Galton-Watson process \( \{Z_i^1\} \) with litter size distribution given
by:
\[ X' = 0 \quad \text{with probability } 1 - (m/k) \]
\[ = k \quad \text{with probability } m/k. \]

Proof
(A). Applying Lemma 3.1 with \( \beta = \gamma_j \), we obtain \( \mathbb{E}[\gamma_j^X] \leq \gamma_{j+1} \) for \( P(X) \in \mathcal{M} \) and \( j = 0, 1, 2, \ldots \). The result (A) now follows by Theorem 3.3.

(B). Applying Lemma 3.1 with \( \beta = \alpha \), we obtain \( \mathbb{E}[\alpha^X] \leq \alpha \) for \( P(X) \in \mathcal{M} \). The result (B) now follows by Theorem 3.4. Note that if \( m \leq 1 \), then \( \alpha = 1 \) and if \( m > 1 \) then \( 0 < \alpha < 1 \).

(C). This follows directly from (A) and Corollary 1 of Theorem 3.3.
(D). Note that for $X \sim X'$, $E[\gamma_j^X] = \gamma_{j+1}$ and $E[\alpha_j^X] = \alpha$. Hence the proof of (D) follows from Corollary 2 of Theorem 3.3 and Corollary 1 of Theorem 3.4.

Remarks This example and the results (A), (B), (D) of Theorem 3.5 were first presented by Goodman (1968).

The example is a case of "Bold play is optimal." That is, in order to maximize the probability of extinction or to minimize the mean time to extinction, subject to the restrictions of $EX \geq m$ and $0 < X < k$, the optimal strategy is for a population to have "extreme" litter sizes, namely $0$ or $k$. It is suboptimal to have intermediate litter sizes ("timid play").

3.3.3 Example 2 (Mean and Variance of the Litter Sizes Constrained)

Let $m$, $\sigma$ be positive real numbers. We will derive an upper bound on the probability of extinction and a lower bound on the expected time to extinction for the branching process described in §3.1, for which the litter size distribution, conditional on the past history, has mean $m$ and variance at most $\sigma^2$. It will turn out that, like Example 1, this is a special case of the $M$-sequences studied in §3.3.1 if $\alpha$ and $\{\gamma_k\}$ are chosen appropriately.

The results of this section take on one of two forms depending on whether or not the quantity $(m^2 + \sigma^2)/m$ is an integer. The two cases are:

(i) $(m^2 + \sigma^2)/m = h$, or (ii) $h < (m^2 + \sigma^2)/m < h+1$, where $h$ is a positive integer.

[The case $0 < (m^2 + \sigma^2)/m < 1$ is impossible, since for $P(X) \in M_0^*$, we have:

$$\sigma^2 > \text{Var}[X] = EX^2 - m^2 > EX - m^2 = m - m^2,$$

which implies that $(m^2 + \sigma^2)/m \geq 1$.]
Lemma 3.2

Let \( M = P(Y) \), \( E[Y] = m, \text{Var}[Y] \leq \sigma^2 \). Then for all \( 0 \leq \beta < 1 \) and \( P(X) \in M \), we have:

(i) \[
E[\beta^X] \leq 1 - m/h + \beta^h \cdot m/h, 
\]

if \( h = (m^2 + \sigma^2)/m \) is a positive integer;

or

(ii) \[
E[\beta^X] \leq 1 - [m(1+2h) - (\sigma^2 + m^2)]/h(h+1) \\
+ \beta^h \cdot [m(1+h) - (\sigma^2 + m^2)]/h + \beta^{h+1} \cdot [m^2 + \sigma^2 - hm]/(h+1), 
\]

if \( h < (m^2 + \sigma^2)/m < h+1 \) with \( h \) a positive integer.

Proof

We note first that if \( \psi(x) = 1 - ax + bx^2 - \beta^x \) then the equation \( \psi(x) = 0 \) has at most three roots in \( x \geq 0 \) for \( a, b > 0 \), \( 0 < \beta < 1 \). This follows by examining the roots of \( \psi'(x) = 0 \). These number at most two since they correspond to the points of intersection of the straight line \( y = -a + 2bx \) and the concave function \( y = \beta^x \log \beta \).

We have that \( \psi(0) = 0 \). Suppose \( \psi(x) = 0 \) has two other roots \( \theta_1, \theta_2 \) with \( 0 < \theta_1 < \theta_2 \). Then since \( \psi(x) \to +\infty \) as \( x \to \infty \), we must have:

\[
\psi(x) \geq 0 \quad 0 \leq x \leq \theta_1 \\
\psi(x) \leq 0 \quad \theta_1 \leq x \leq \theta_2 \\
\psi(x) \geq 0 \quad \theta_2 \leq x .
\]
(i) Suppose first that \( h = (m^2 + \sigma^2)/m \) is an integer.

\[ f(x) = \beta^h + (x-h) \cdot \beta^h \log \beta + (x-h)^2 \left[ 1 - (1-h \cdot \log \beta) \beta^h \right]/h^2. \]

Then \( f(\cdot) \) is tangent to \( g(\cdot) \) at \( x = h \) and \( \psi(x) = f(x) - g(x) \) is of the form discussed above, with the two positive roots of \( \psi(x) = 0 \) coincident at \( \theta_1 = \theta_2 = h \). Thus we have \( \psi(x) > 0 \) for all \( x > 0 \) and therefore \( g(x) \leq f(x) \) for all \( x > 0 \). (See Figure 3.2.) This result still holds in the limit \( \beta \to 0 \), for then \( f(x) = (x-h)^2/h^2 \) and \( g(x) = 1 \) for \( x = 0 \) and \( g(x) = 0 \) for \( x > 0 \).
Thus, by construction, \( g(x) \leq f(x) \) for all \( x \geq 0 \) and \( 0 \leq \beta < 1 \).

Hence for \( P(X) \in \mathcal{M} \) we have \( E_g(X) \leq E_f(X) \), or:

\[
E[\beta^X] \leq E[\beta^h + (X-h) \cdot \beta^h \log \beta + (X-h)^2 \cdot \left[1 - (1-h \cdot \log \beta)\beta^h\right]/h^2]
\]

\[
= \beta^h + (EX-h)\beta^h \log \beta + [\text{Var } X + (h-EX)^2] \cdot \left[1 - (1-h \cdot \log \beta)\beta^h\right]/h^2
\]

\[
\leq \beta^h + (m-h)\beta^h \log \beta + [\sigma^2 + (h-m)^2] \cdot \left[1 - (1-h \cdot \log \beta)\beta^h\right]/h^2
\]

\[
= \beta^h + (m-h)\beta^h \log \beta + (h^2-hm)[1 - (1-h \cdot \log \beta)\beta^h]/h^2
\]

\[
= 1 - m/h + \beta^h \cdot m/h ,
\]

which proves part (i) of Lemma 3.2.

(ii) Suppose now that \( h < (m^2 + \sigma^2)/m < h+1 \), where \( h \) is a positive integer.
For $0 < \beta < 1$, let $g(x) = \beta^x$ as before and define $f(x) = 1 - ax + bx^2$, where

$$a = [(1-\beta^h)(1+2h) - h^2\beta^h(1-\beta)]/h(h+1), \quad \text{and}$$

$$b = [1 - \beta^h - h\beta^h(1-\beta)]/h(h+1).$$

It is easy to show that $a, b$ are positive and thus $\psi(x) = f(x) - g(x)$ is of the required form and $\psi(x) = 0$ has two positive roots at $\theta_1 = h, \theta_2 = h + 1$. Thus we have $\psi(x) \geq 0$ for all $0 \leq x < h$, and all $x \geq h + 1$. (See Figure 3.3.) This result still holds in the limit $\beta \to 0$ for then $f(x) = (x-h)(x-h-1)/h(h+1)$.

For $P(X) \in M$, $X$ is concentrated on the non-negative integers and so for $0 \leq \beta < 1$ we have $Eg(X) \leq Ef(X)$, or:

$$E[\beta^X] \leq E[1 - ax + bx^2]$$

$$\leq 1 - am + b(\sigma^2 + m^2)$$

$$= 1 - [m(1+2h) - (\sigma^2 + m^2)]/h(h+1)$$

$$+ \beta^h \cdot [m(1+h) - (\sigma^2 + m^2)]/h + \beta^{h+1} \cdot [m^2 + \sigma^2 - hm]/(h+1),$$

which completes the proof of Lemma 3.2.
Theorem 3.6

Let $M = \{P(X): EX = m, \ Var X \leq \sigma^2\}$ and set $M(z) = \{P(X_1 + \ldots + X_z): \{X_i\}\}$ are i.i.d., $P(X_i) \in M\}$. For $0 \leq \beta \leq 1$, define $\phi(\beta)$ by:

(i) 
$$\phi(\beta) = 1 - \frac{h}{m/h + \beta^h \cdot m/h},$$

if $(m^2 + \sigma^2)/m = h$ is an integer;

or

(ii) 
$$\phi(\beta) = 1 - \frac{[m(1+2h) - (\sigma^2 + m^2)]/h(h+1)}{\beta^h \cdot [m(1+h) - (\sigma^2 + m^2)]/h + \beta^{h+1} \cdot \beta^{2-hm}/(h+1)},$$

if $h < (m^2 + \sigma^2)/m < h+1$ where $h > 1$ is an integer.

Then for $Z_0, Z_1, \ldots$, an $M$-sequence starting at $z$, we have:

A) $Pr[Z_N = 0] \leq (\gamma_N)^z$,

where $\gamma_N$ is defined recursively by:

$\gamma_0 = 0$,

$\gamma_{j+1} = \phi(\gamma_j)$ for $j = 0, 1, 2, \ldots, N-1$;

B) $Pr[Z_N = 0 \ for \ some \ N] \leq \alpha^z$,

where $\alpha$ is the smaller root in $[0, 1]$ of

$\alpha = \phi(\alpha)$.
C) \( \mathbb{E}[T_e] \geq \sum_{i=0}^{\infty} \left[ 1 - (\gamma_i)^2 \right] \),

where \( \{\gamma_i, i \geq 0\} \) are defined as in (A).

D) The bounds in (A), (B), (C) are sharp and are all attained when \( \{Z_i\} \)
form a Galton-Watson process \( \{Z_i^t\} \) with litter size distribution given
by:

in case (i) \( X^t = 0 \) with probability \( 1 - \frac{m}{h} \),
= \( h \) with probability \( \frac{m}{h} \);

or in case (ii) \( X^t = 0 \) with probability \( 1 - \frac{[m(1+2h) - (\sigma^2 + m^2)]}{h(h+1)} \),
= \( h \) with probability \( \frac{[m(1+h) - (\sigma^2 + m^2)]}{h} \),
= \( h+1 \) with probability \( \frac{[\sigma^2 + m^2 - hm]}{(h+1)} \).

**Proof** The proof parallels that of Theorem 3.5.

Applying Lemma 3.2 with \( \beta = \gamma_j \), we obtain \( \mathbb{E}[\gamma_j X] \leq \gamma_{j+1} \) for \( P(X) \in \mathcal{M} \)
and \( j = 0, 1, 2, \ldots \). The result (A) now follows by Theorem 3.3.

Applying Lemma 3.2 with \( \beta = \alpha \), we get \( \mathbb{E}[\alpha X] \leq \alpha \) for \( P(X) \in \mathcal{M} \). The
result (B) now follows by Theorem 3.4. Note that if \( m \leq 1 \) then \( \alpha = 1 \),
whereas if \( m > 1 \) then \( 0 < \alpha < 1 \).

The result (C) follows directly from (A) and Corollary 1 of Theorem 3.3.

Since for \( X \sim X' \), \( \mathbb{E}[\gamma_j X] = \gamma_{j+1} \) and \( \mathbb{E}a X = a \), the result (D) follows
from Corollary 2 of Theorem 3.3 and Corollary 1 of Theorem 3.4. \( \square \)

**Remark** As in Example 1, we see that "Bold play is optimal."
3.3.4 Lower Bounds

Theorem 3.7 (Extinction in a finite time N)

For any positive integer N, let \( g_0 = 0, g_1, g_2, \ldots, g_N \) be a sequence of real numbers \( 0 \leq g_i \leq 1; \ i = 1, 2, \ldots, N \),

\[
\mathcal{M} = \{ P(X): \ E[g_j^X] \geq g_{j+1}, \ j = 1, 2, \ldots, N-1 \},
\]

and set

\[
\mathcal{M}(z) = \{ P(X_1 + \ldots + X_z): \ X_i \text{ are i.i.d., } P(X_i) \in \mathcal{M} \}
\]

for \( z = 0, 1, 2, \ldots \).

Then for \( Z_0, Z_1, \ldots \), an \( \mathcal{M} \)-sequence starting at \( z \), we have:

\[
(3.11) \quad \Pr[Z_N = 0] \geq (g_N)^z.
\]

Proof

We apply Theorem 2.2 with \( C \) = the set of non-negative integers, \( r(z) = 0 \), \( T = N \), \( f_k(z) = (g_k)^z \) for \( z = 0, 1, 2, \ldots \) and \( k = 1, 2, \ldots, N \). Then \( r \) and \( f_k \) are non-negative and for \( P(Z) \in \mathcal{M}(z) \):

\[
r(z) + Ef_k(Z) = E[g_k^{X_1 + \ldots + X_z}]
\]

\[
= (E[g_k^X])^z
\]

\[
\geq (g_k+1)^z
\]

\[
= f_{k+1}(z).
\]
Also the integrability condition, $E\left[f_{N-k}(Z_k) + \sum_{i=0}^{k-1} r(Z_i)\right] < \infty$, $k = 0, 1, 2, \ldots, N$, is satisfied since $r \equiv 0$ and $f$ is bounded by 0 and 1. Thus the conditions of Theorem 2.2 are satisfied, and the result (3.11) follows by noting that (as in Theorem 3.3):

$$E\left[\sum_{k=0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)})\right] = \Pr[Z_N = 0].$$

\[ \square \]

**Corollary 1 (The Expected Time to Extinction)**

Suppose $g_0 = 0, g_1, g_2, \ldots$ is a sequence of real numbers in $[0, 1]$. Take

$$\mathcal{N} = \{P(X): E[g_j^X] \geq g_{j+1} \text{ for all } j\}.$$

Then

$$(3.12) \quad E[T_e] \leq \sum_{i=0}^{\infty} \left[1 - (g_i)^2\right],$$

where $T_e$ is the time to extinction as defined in §3.1.

**Note** As in the Note on page 24, the sum on the right hand side of the inequality (3.12) may be finite or infinite.

**Proof**

Proceeding as in the proof of Corollary 1 of Theorem 3.3 we have that:
$$E[T_e] = \sum_{i=0}^{\infty} \Pr[T_e > i]$$

$$= \sum_{i=0}^{\infty} \Pr[Z_i > 0]$$

$$= \sum_{i=0}^{\infty} (1 - \Pr[Z_i = 0])$$

$$\leq \sum_{i=0}^{\infty} [1 - (g_i)^{Z_i}] ,$$

which gives us the required result (3.12).

Corollary 2 (Achievement of Bounds)

If there exists a Galton-Watson process with litter sizes $X$ such that $E[g_j^X] = g_{j+1}$, for all $j$, then this achieves the bounds (3.11), (3.12).

This follows by Harris (1963; Chapter 1, Theorem 6.1).

Corollary 3 (Eventual Extinction)

Suppose $g_N \to \alpha$ as $N \to \infty$, where $0 \leq \alpha \leq 1$. Then, since

$\Pr[Z_N \geq 0 \text{ for some } N] \geq \Pr[Z_N = 0]$ and the left hand side is independent of $N$, we have:

$$\Pr[Z_N = 0 \text{ for some } N] \geq \alpha^Z .$$

Furthermore, if there exists a Galton-Watson process with litter sizes $X$ such that $E[\alpha^X] = \alpha$, then this process achieves the bound (3.13). This follows by Harris (1963, Chapter 1, Theorem 6.1).

In the next section, we shall give an example in which an interesting family of offspring distributions is defined, and the $g_i$ are chosen in such a way that this family is a subset of $\mathcal{N}$ as defined in Theorem 3.7. Also
the Galton-Watson processes, as described in Corollaries 2 and 3, will be shown to exist.

3.3.5 Example 3

In this example, we shall use similar notation to that of Goodman (1968). Let $H \subseteq C$ be a set of non-negative integers and $m$ be some positive real number. We shall derive a lower bound for the probability of extinction and an upper bound on the mean time to extinction of the branching process described in §3.1 for which, conditional on the past, the allowable litter sizes are in $H$, with mean at most $m$. It will turn out that this is a special case of the $\mathcal{M}$-sequences studied in §3.3.4 if the $\{g_k\}$ are chosen appropriately.

Lemma 3.3

Suppose $m$ is not in $H$, and $m^*$ is the smallest integer in $H$ greater than $m$, $m'$ the largest integer in $H$ smaller than $m$ and $d = m^* - m' > 0$. (We assume $m'$, $m^*$ exist.)

Let $\mathcal{M} = \{P(X): X$ concentrated on $H, EX \leq m\}$. Then for all $0 \leq \beta \leq 1$ and $P(X) \in \mathcal{M}$, we have:

$$E[\beta^X] \geq [(m^*-m)\beta^{m'} + (m-m')\beta^{m^*}] / d.$$

Proof

Let $g(x) = \beta^X$ and $f(x) = [(m^*-x)\beta^{m'} + (x-m')\beta^{m^*}] / d$. We will show that $g(x) \geq f(x)$ for all $x \in H$ and for $0 \leq \beta \leq 1$.

Suppose first that $\beta = 0$, then $g(x) = 1$ for $x = 0$, and $g(x) = 0$ otherwise. If $m' \neq 0$, $f(x) \equiv 0$ and so $g(x) \geq f(x)$ for all real $x$. If $m' = 0$, then for $x > m'$, $g(x) = \beta^x$, and for $x \leq m'$, $g(x) = 1$. Then $g(x)$ and $f(x)$ are both constant functions, and $g(0) = 1$ and $f(0) = 1$, so $g(x) \geq f(x)$ for all $x$. If $m' = 0$, the proof is straightforward.
If \( m' = 0 \), \( f(x) = (m^*-x)/m^* \) and because there are no members of \( H \) between \( m' = 0 \) and \( m^* \) we have \( g(x) \geq f(x) \) for all \( x \) in \( H \). (See Figure 3.4.)

**Figure 3.4**

Now suppose \( 0 < \beta < 1 \). Then since \( f \) is linear, \( g \) is concave, \( f(m') = g(m') \), and \( f(m^*) = g(m^*) \), we have \( f(x) \leq g(x) \) for all real \( x \leq m' \) and \( x \geq m^* \). (See Figure 3.5.)

**Figure 3.5**
Thus by construction \( g(x) \geq f(x) \) for all \( x \in H \) and \( 0 \leq \beta \leq 1 \), and hence, for \( P(X) \in M \), we have \( E g(X) \geq E f(X) \), or:

\[
E[\beta^X] \geq E[(m^*-X)\beta^{m'} + (X-m')\beta^{m*}] / d
\]

\[
= [m^* \beta^{m'} - m' \beta^{m*} - (\beta^{m'} - \beta^{m*}) \cdot EX] / d
\]

\[
\geq [m^* \beta^{m'} - m' \beta^{m*} - (\beta^{m'} - \beta^{m*}) \cdot m] / d
\]

\[
= [m^* - m] \beta^{m'} + (m - m') \beta^{m*} / d
\]

which completes the proof of Lemma 3.3.

Theorem 3.8

Let \( M = \{ P(X) : X \) concentrated on \( H \), \( E X \leq m \) \} and set \( M(z) = \{ P(X_1 + \ldots + X_z) : \{ X_i \} \) are i.i.d., \( P(X_i) \in M \} \). Then for \( Z_0, Z_1, Z_2, \ldots \), an \( M \)-sequence starting at \( z \), we have:

A) \( \Pr[Z_N = 0] \geq (g_N)^z \), \( (N = 0, 1, 2, \ldots) \);

where (i) if \( m \in H \), \( g_N = 0 \) for all \( N \),

or (ii) if \( m \notin H \), \( g_N \) is defined recursively by

\( g_0 = 0, \ g_{j+1} = [(m^*-m)g_j^{m'} + (m-m')g_j^{m*}] / d \)

\( (j = 0, 1, 2, \ldots, N-1) \), where \( m^*, m', d \) are defined as in Lemma 3.3;

B) \( \Pr[Z_N = 0 \ for \ some \ N] \geq \alpha^z \),

where (i) if \( m \in H \), \( \alpha = 0 \),
or (ii) if \( m \notin H \), \( \alpha \) is the smaller root in \([0,1]\) of:

\[
\alpha = \frac{[(m^*-m)m' + (m-m')m^*]}{d};
\]

C) \( E[T_e] \leq \sum_{i=0}^{\infty} \left[ 1 - (g^*_i)^z \right] \),

where \( \{g^*_i, i \geq 0\} \) are defined as in (A).

D) The bounds (A), (B), (C) are sharp and are all attained when \( \{Z_i\} \) form a Galton-Watson process \( \{Z'_i\} \) with litter size distribution given by:

(i) if \( m \in H \), \( X' = m \) with probability one;

or (ii) if \( m \notin H \), \( X' = m' \) with probability \( (m^*-m)/d \)

\( = m^* \) with probability \( (m-m')/d \).

Proof

We consider first the case \( m \in H \). Since \( m > 0 \), this means that \( m \geq 1 \).

Then for the \( M^* \)-sequence \( \{Z'_i\} \) as defined in D(i) we have \( Z'_N = z \cdot m^N \) with probability one and hence \( \Pr[Z'_N = 0] = 0 \) and \( T_e = \infty \) with probability one. The results A(i), B(i), C(i), D(i) follow.

Now consider \( m \notin H \). Applying Lemma 3.3 with \( b = g^*_j \), we obtain

\( E[g^*_j] \geq g^*_j+1 \) for \( P(X) \in M^*_\alpha \), and \( j = 0, 1, 2, \ldots \). The result A(ii) now follows by Theorem 3.7. By Harris (1963, Chapter 1, Theorem 6.1) we have \( g^*_N \to \alpha \) as \( N \to \infty \) and hence B(ii) follows by Corollary 3 of Theorem 3.7.

The result C(ii) follows directly from A(ii) and Corollary 1 of Theorem 3.7.

Since for \( X \sim X' \), \( E[g^*_j] = g^*_j+1 \) and \( E[\alpha^*_j] = \alpha \), the result D(ii) follows by Corollaries 2 and 3 of Theorem 3.7. This completes the proof of Theorem 3.8.
Remarks:

1) If \( m' > 0 \), then \( a = g_0 = g_1 = g_2 = \ldots = 0 \). In fact, with probability one, \( Z_N^t \geq z \cdot (m')^N \).

2) The result agrees with that of Freedman and Purves (1967), who considered the special case with \( H = 0, 2, 3, 4, 5, \ldots \) and \( 0 < m < 2 \). Then the optimal strategy is for the litter sizes to have the distribution:

\[
X' = 0 \quad \text{with probability} \quad 1 - \frac{m}{2}, \quad \text{and} \\
X' = 2 \quad \text{with probability} \quad \frac{m}{2}.
\]

3) This example was discussed also in Goodman (1968) in the more general case where \( m \) and \( H \) (he uses \( \tilde{H} \)) are allowed to depend on the time period \( n \).

4) The example is a case of "Timid play is optimal." That is in order to minimize the probability of extinction or to maximize the expected time to extinction, subject to the restrictions \( \Pr[X \in H] = 1 \) and \( EX \leq m \), the optimal strategy is for the population to have litters with the "smallest variability" in size.

3.4 The Probability that a Generation Size Will Exceed a Given Number and the Expected Maximum Generation Size

Theorem 3.9

Take \( \varepsilon > 0 \), \( a > 1 \), \( M = \{ P(X), E[a^X] \leq a \} \) and set \( M(z) = \{ P(X_1 + \ldots + X_\infty) : \{ X_i \} \} \) are i.i.d., \( P(X_1) \in M \).

Then for \( z_0, z_1, \ldots \), an \( M \)-sequence starting at \( z \), we have:

\[
\Pr[Z_N \geq \varepsilon \text{ for some } N] \leq \frac{(a^{z_0} - 1)}{(a^{\varepsilon} - 1)}, \quad (0 \leq z \leq \varepsilon);
\]

\[
(3.14) \quad = 1, \quad (z \geq \varepsilon).
\]
Proof

Trivially if \( Z_0 = z \geq \ell \), then \( \Pr[Z_N \geq \ell \text{ for some } N] = 1 \). Suppose \( z \) is some non-negative integer less than or equal to \( \ell \). We apply Theorem 2.3 with \( C = \) the set of non-negative integers;

\[
T = T_{\ell} = \min\{n: Z_n \geq \ell\}, \text{ if } Z_n \geq \ell \text{ for some } n,
\]

\[= \infty \quad \text{otherwise}; \]

\( r(z) \equiv 0 \) and \( f(z) = (\alpha^z - 1)/(\alpha^\ell - 1) \) if \( 0 \leq z \leq \ell \), \( f(z) = 1 \) if \( z \geq \ell \).

Then \( r \) and \( f \) are non-negative, and since \( f(z) \leq (\alpha^z - 1)/(\alpha^\ell - 1) \) for all \( z \geq 0 \) (see Figure 3.6), we have for \( P(Z) \in \mathcal{M}(z) \):

\[
r(z) + Ef(Z) \leq E[(\alpha^z - 1)/(\alpha^\ell - 1)]
\]

\[
= E[\frac{X_1 + \ldots + X_z - 1}{(\alpha^\ell - 1)}]
\]

\[
= \frac{[(E\alpha^z)^z - 1]/(\alpha^\ell - 1)}{(\alpha^z - 1)/((\alpha^\ell - 1))}
\]

\[
= f(z).
\]
Thus the conditions for Theorem 2.3 are satisfied. Define \( v(z) = 1 \) if \( z \geq \ell \), \( v(z) = 0 \) otherwise. Hence by Theorem 2.3, and since \( f(z) \geq v(z) \) for all \( z \geq 0 \) we have:

\[
\frac{(a^z - 1)}{\alpha - 1} \leq E[f(Z_T) \cdot I_{T < \infty}]
\]

\[
\geq E[v(Z_T) \cdot I_{T < \infty}]
\]

\[
= \Pr[T < \infty]
\]

\[
= \Pr[Z_N \geq \ell \text{ for some } N],
\]

which completes the proof.
Corollary 1

For any \( \alpha \)-sequence starting at \( z \), we have:

\[
E[\sup_n Z_n] \leq z + (\alpha^Z - 1) \cdot \sum_{i=z+1}^{\infty} (\alpha^i - 1)^{-1}.
\]

Proof

Let \( M = \sup_n Z_n \). Then

\[
E[M|Z_0 = z] = z + \sum_{i=z+1}^{\infty} \Pr[M \geq \xi|Z_0 = z]
\]

\[
= z + \sum_{i=z+1}^{\infty} \Pr[T_{\xi} < \infty|Z_0 = z]
\]

\[
\leq z + \sum_{i=z+1}^{\infty} (\alpha^Z - 1)/(\alpha^i - 1),
\]

which proves the result (3.15).

Example 4

Let \( k \) be some positive integer and \( 0 < m < 1 \). We shall derive an upper bound on the probability that a generation ever equals or exceeds \( \xi \) in number for the branching process described in §3.1 for which the litter sizes \( X \) are constrained to be at most \( k \) and with mean, conditional on the past, to be at most \( m \).

Lemma 3.4

Let \( \mathcal{M} = \{ P(X) : 0 \leq X \leq k, \ EX \leq m \} \). Then for \( \alpha > 1 \) and \( P(X) \in \mathcal{M} \), we have:

\[
E[\alpha^X] \leq 1 - m/k + \alpha^k \cdot m/k.
\]
Proof

Define $g(x) = \alpha^x$ and $f(x) = 1 - x/k + \alpha^k \cdot x/k$. Then $f(x) \geq g(x)$ for $0 \leq x \leq k$ (see Figure 3.7), and so $Eg(X) \leq Ef(X)$ for $P(X) \in \mathcal{M}$. Substituting for $f$ and $g$ we obtain:

$$E[\alpha^X] \leq 1 + (EX) \cdot (\alpha^k - 1)/k$$

$$\leq 1 + m \cdot (\alpha^k - 1)/k$$

$$= 1 - m/k + \alpha^k \cdot m/k$$

which proves the lemma.
Theorem 3.10

Take \( 0 < m < 1 \), \( \mathbb{M} = \{ P(X): \ EX \leq m, \ 0 \leq X \leq k \} \) and set

\( \mathbb{M}(z) = \{ P(X_1 + \ldots + X_z): \{X_i\} \text{ are i.i.d., } P(X_1) \in \mathbb{M} \} \). Then for \( z_0, z_1, \ldots \), an \( \mathbb{M} \)-sequence starting at \( z \), we have:

\[
\Pr[Z_N \geq \ell \text{ for some } N] \leq \frac{\alpha^z - 1}{\alpha^\ell - 1},
\]

(for \( 0 \leq z \leq \ell \));

\[
E[\sup_n Z_n] \leq z + \sum_{i=1}^{\infty} \frac{\alpha^z - 1}{\alpha^i - 1},
\]

where \( \alpha (>1) \) is the larger root of the equation: \( \alpha = 1 - m/k + \alpha^k \cdot m/k \).

Proof

Since \( \alpha = 1 - m/k + \alpha^k \cdot m/k \) and \( 0 < m < 1 \), we have \( \alpha > 1 \) and, by Lemma 3.4, \( E[\alpha^X] \leq \alpha \) for \( P(X) \in \mathbb{M} \). Theorem 3.10 now follows by Theorem 3.9 and Corollary 1.

Remark 1: Attainment of Bounds

The Galton-Watson process \( \{Z_i\} \), with litter size distribution:

- \( X = 0 \) with probability \( 1 - m/k \)
- \( = k \) with probability \( m/k \)

is an \( \mathbb{M} \)-sequence with \( E[\alpha^X] = \alpha \). However the bounds (3.16), (3.17) are not attained in general by this process. This is due to the end effect, i.e. the undesirability of overshooting \( \ell \). We conjecture that in fact slightly more
timid strategies should be used by succeeding generations in order to maximize the probability that one generation size hits \( \ell \) exactly.

**Remark 2** If \( m = 1 \), then taking the limit, \( \alpha \to 1 \) we obtain:

\[
\Pr[Z_N \geq \ell \text{ for some } \ell] < \frac{z}{\ell},
\]

and

\[
E[\sup_n Z_n] < \infty.
\]

This follows because \( \lim_{\alpha \to 1} \frac{\alpha^{z-1}}{\alpha^{z-1}} = \frac{z}{\ell} \) and \( z + \sum_{\ell = z+1}^{\infty} \frac{z}{\ell} = \infty \).

**Remark 3** If we consider the same \( \mathbb{N} \)-sequences as in Example 4 but with \( m > 1 \), then clearly a sharp upper bound on \( \Pr[Z_N \geq \ell \text{ for some } \ell] \) is one, and \( E[\sup_n Z_n] \) is unbounded. These bounds are attained by the Galton-Watson process \( \{Z'_i\} \) with litter sizes \( X \) such that \( \Pr[X = 0] = 0 \) and \( EX = m \). This is clearly possible if \( m > 1 \), and the population will grow unboundedly and can never die out.
CHAPTER IV
AN ASSOCIATED MULTIPLICATIVE PROCESS

In this chapter we will define a multiplicative process and derive results which will parallel exactly those obtained for the branching process in Chapter III. To keep the analogy precise, we will again take $C$ to be the set of non-negative integers and $\mathcal{M}$ will be always a set of probability distributions concentrated on $C$. In fact, the results of this chapter may be extended to more general $C$ and $\mathcal{M}$, and this will be discussed in §4.5.

4.1 A Description of the Model

Let $\mathcal{M}$ be some set of probability distributions concentrated on $C$, the set of all non-negative integers. Then, for each $z$ in $C$, define a sequence of random variables $Z_0, Z_1, \ldots$ as follows:

$Z_0 = z,$

$Z_{n+1} = Z_n \cdot X_{n+1},$ where $P(X_{n+1} \mid Z_0, Z_1, \ldots, Z_n) \in \mathcal{M},$

$n = 0, 1, 2, \ldots.$

More precisely, $Z_0, Z_1, Z_2, \ldots$ is an $\mathcal{M}$-sequence starting at $z$ with:

$\mathcal{M}(z) = \{P(zX) : P(X) \in \mathcal{M}\}.$

This process can be regarded as a branching process with deterministic offspring in a random environment (see also the discussion in §3.1). The litter sizes of all individuals in a given generation are deterministic and
equal but these common litter sizes vary randomly from generation to generation.

As before define $T_e$, the time to extinction, by:

$$T_e = \min\{N: Z_N = 0\} \text{ if } Z_N = 0 \text{ for some } N,$$
$$= \infty \text{ otherwise}.$$

4.2 The Total Population Size

Theorem 4.1 (Upper bounds)

Take $m > 0$, let $\mathcal{M} = \{P(X), EX \leq m\}$ and set $\mathcal{M}(z) = \{P(zX): P(X) \in \mathcal{M}\}$ for $z \in C$.

Then for $Z_0, Z_1, \ldots$, an $\mathcal{M}$-sequence starting at $z$, we have:

$$E[\sum_{i=0}^{N} Z_i] \leq \begin{cases} 
z \cdot \frac{1-m^{N+1}}{1-m} & \text{if } m \neq 1 \\
(N+1) \cdot z & \text{if } m = 1 
\end{cases}$$

for $N = 0, 1, 2, \ldots$;

and

$$E[\sum_{i=0}^{\infty} Z_i] \leq \begin{cases} 
z \cdot \frac{1}{1-m} & \text{if } m < 1 \\
\infty & \text{if } m = 1 
\end{cases}.$$

Furthermore these bounds are attained for the $\mathcal{M}$-sequence $\{Z'_1\}$ for which the multiplicands $X_1, X_2, \ldots$ are i.i.d. with mean $m$. 

Proof

The proof is completely analogous to that of Theorem 3.1 with \( r(\cdot) \), \( f_k(\cdot) \), and \( T \) as defined there. In this case

\[
\begin{align*}
   r(z) + Ef_k(Z) &= \begin{cases} 
   z + E[\frac{1-m}{1-m} zX] & \text{if } m \neq 1 \\
   z + E[(k+1)zX] & \text{if } m = 1 
   \end{cases} 
\end{align*}
\]

\[< f_{k+1}(z), \text{ for } P(Z) \in \mathcal{N}(z). \]

If \( X_1, X_2, \ldots \) are i.i.d. random variables with mean \( m \) then

\[
E[Z_1] = z \cdot E[\sum_{j=1}^{\infty} X_j] = z \cdot E[X_1] = z \cdot m. \]

Since \( E[\sum_{i=0}^{N} Z_i] = \sum_{i=0}^{N} E[Z_i] \) and

\[
E[\sum_{i=0}^{\infty} Z_i] = \sum_{i=0}^{\infty} E[Z_i] \text{ (by monotone convergence theorem; the } \{Z_i\} \text{ are all non-negative)},
\]

it follows that the bounds (4.1), (4.2) are attained when the \( \{X_i\} \) form such a sequence. \( \square \)

Theorem 4.2 (Lower bounds)

Take \( m > 0 \), let \( \mathcal{N} = \{P(X), EX \geq m\} \) and set \( \mathcal{M}(z) = \{P(zX): P(X) \in \mathcal{N}\} \) for \( z \in \mathbb{C} \). Then for \( Z_0, Z_1, \ldots \), an \( \mathcal{N} \)-sequence starting at \( z \), we have:

\[
(4.3) \quad E[\sum_{i=0}^{N} Z_i] \geq \begin{cases} 
   z \cdot \frac{1-m}{1-m}^{N+1} & \text{if } m \neq 1 \\
   (N+1) \cdot z & \text{if } m = 1 
   \end{cases},
\]

for \( N = 0,1,2,\ldots \); and
\[(4.4) \quad E[\sum_{i=0}^{\infty} Z_i] \geq z \cdot \frac{1}{1-m} \quad \text{if } m < 1\]

\[= \infty \quad \text{if } m = 1.\]

Furthermore, these bounds are attained for the \( M \)-sequence \( \{Z_i\} \) for which the multiplicands \( X_1, X_2, \ldots \) are i.i.d. with mean \( m \).

**Proof**

The proof is completely analogous to that of Theorem 3.2 with \( r(\cdot) \), \( f_k(\cdot) \), and \( T \) as defined there. In this case

\[r(z) + Ef_k(Z) = \begin{cases} 
  z + E[\frac{1-m}{1-m} \cdot zX] & \text{if } m \neq 1 \\
  z + E[(k+1)zX] & \text{if } m = 1
\end{cases}\]

\[\geq f_{k+1}(z), \quad \text{for } P(Z) \in M(z).\]

The attainment of the bounds in (4.3), (4.4) follows as in Theorem 4.1.

\[\square\]

**4.3 The Probability of Extinction and the Expected Time to Extinction**

In this section, we consider certain families \( M \) of distributions and find bounds on the probability of extinction (or absorption at zero) and on the mean time to extinction for the multiplicative process described in §4.1. The examples and results will be analogous to those found in Chapter III. All bounds that are derived will be sharp and will be attained by \( M \)-sequences for
which the multiplicands \( \{X_i\} \) form an i.i.d. sequence and are distributed as the litter sizes of the optimal Galton-Watson process in corresponding branching process.

### 4.3.1 Upper Bounds

**Theorem 4.3**

Take \( 0 \leq q \leq 1 \), let \( \mathcal{M} = \{P(X): \Pr[X = 0] \leq q\} \) and set \( \mathcal{M}_z = \{P(zX): P(X) \in \mathcal{M}\} \) for all \( z \in \mathbb{C} \). Then for \( Z_0, Z_1, \ldots \), an \( \mathcal{M}_z \)-sequence starting at \( z \), we have:

\[
\Pr[Z_N = 0] \leq \begin{cases} 
1 - (1-q)^N & \text{if } z > 0 \\
1 & \text{if } z = 0 
\end{cases}
\]

\[(4.5)\]

for \( N = 0, 1, 2, \ldots \).

**Proof**

The proof is similar to that of Theorem 3.3. We apply Theorem 2.1 with \( r(z) \equiv 0 \), \( T = N \), and

\[
f_k(z) = \begin{cases} 
1 - (1-q)^k & \text{if } z > 0 \\
1 & \text{if } z = 0 
\end{cases}
\]

for \( k = 0, 1, 2, \ldots, N \), and for \( z \) in \( \mathbb{C} \).

Then \( r \) and \( f_k \) are non-negative and if \( z = 0 \) then:

\[
r(z) + Ef_k(z) = Ef_k(zX) = 1 = f_{k+1}(z), \quad \text{for } P(Z) \in \mathcal{M}(0). \quad \text{If } z > 0, \text{ then}
\]
\[ r(z) + E f_k(Z) = E f_k(zX) \]
\[ = 1 \cdot \Pr[X = 0] + [1 - (1-q)^k] \cdot \Pr[X > 0] \]
\[ = (1-q)^k \cdot \Pr[X = 0] + 1 - (1-q)^k \]
\[ \leq (1-q)^k q + 1 - (1-q)^k \]
\[ = 1 - (1-q)^{k+1} \]
\[ = f_{k+1}(z) , \]

for \( P(Z) \in \mathcal{M}(z) \), \( k = 0,1,2,... \).

Thus the conditions of Theorem 2.1 are verified and the result (4.5) follows by noting that:

\[ E[L_{0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)})] = E[f_0(Z_N)] \]
\[ = \Pr[Z_N = 0] , \]

since \( f_0(z) = \begin{cases} 
0 & \text{if } z > 0 \\
1 & \text{if } z = 0 
\end{cases} . \)

\[ \square \]

**Corollary 1**

If \( 0 < q \leq 1 \) and \( z > 0 \) then

\[ E[T_e] \geq 1/q . \]
Proof

As in Corollary 1 of Theorem 3.3 we have:

\[ E[T_e] = \sum_{i=0}^{\infty} \Pr[T_e > i] \]

\[ = \sum_{i=0}^{\infty} \Pr[Z_i > 0] \]

\[ = \sum_{i=0}^{\infty} (1 - \Pr[Z_i = 0]) \]

\[ > \sum_{i=0}^{\infty} (1-q)^i \]

\[ = \frac{1}{q} , \quad \text{which proves the result.} \]

\[ \square \]

Corollary 2

The bounds (4.5), (4.6) are sharp and are attained by the \( M \)-sequence for which the multiplicands \( \{X_i\} \) are i.i.d. with \( \Pr[X_i = 0] = q \).

Proof

Suppose \( \{X_i; i = 1,2,\ldots,N\} \) are i.i.d. with \( \Pr[X_i = 0] = q \) then \( \Pr[Z_N = 0] = 1 \) if \( z = 0 \), while for \( z > 0 \) we have:

\[ \Pr[Z_N = 0] = 1 - \Pr[Z_N \neq 0] \]

\[ = 1 - \prod_{i=1}^{N} \Pr[X_i \neq 0] \]

\[ = 1 - (1-q)^N , \]

which proves Corollary 2. \[ \square \]
4.3.2 Example 1

Let \( k \) be some positive integer and \( m \) a real number such that \( 0 \leq m \leq k \). We will derive an upper bound on the probability of extinction and a lower bound on the mean time to extinction for the multiplicative process defined in §4.1, for which the multiplicands \( X_i \) \((i = 1, 2, \ldots)\) are bounded by 0 and \( k \) and have mean, conditional on the past, \( Z_0, Z_1, \ldots, Z_{i-1} \), of at least \( m \). It will turn out that this is a special case of the \( M \)-sequences studied in §4.3.1 with \( q = 1 - m/k \).

Lemma 4.1

Let \( M = \{ P(X): 0 \leq X \leq k, \ E[X] > m \} \), then for all \( P(X) \in M \) we have:

\[
\Pr[X = 0] \leq 1 - \frac{m}{k}.
\]

Proof

Let \( f(x) = 1 - x/k \) and \( g(x) = 1 \) if \( x = 0 \) and \( g(x) = 0 \) otherwise.

![Figure 4.1](image-url)
Now \( g(x) \leq f(x) \) for \( 0 \leq x \leq k \) (see Figure 4.1) and so, for \( P(X) \in \mathcal{M} \), we have \( E_g(X) \leq E_f(X) \), or

\[
Pr[X = 0] \leq 1 - E[X/k] \leq 1 - m/k
\]

which completes the proof of the lemma.

\( \square \)

**Theorem 4.4**

Let \( \mathcal{M} = \{ P(X): 0 \leq X \leq k, \ EX \geq m \} \) and set \( \mathcal{M}(z) = \{ P(zX): P(X) \in \mathcal{M} \} \) for each \( z \in C \). Then for \( z_0, z_1, \ldots, \) an \( \mathcal{M} \)-sequence starting at \( z > 0 \), we have:

\[
Pr[z_N = 0] \leq 1 - (m/k)^N, \quad \text{and}
\]

\[
E[T_N] \geq k/(k-m).
\]

These bounds are sharp and are attained by the \( \mathcal{M} \)-sequence \( \{ z_i' \} \) for which the multiplicands \( \{ X_i' \} \) are i.i.d. with common distribution given by:

\[
X_i' = 0 \quad \text{with probability } \ 1 - m/k \\
= k \quad \text{with probability } \ m/k.
\]

**Proof**

The theorem follows directly from Lemma 4.1 and Theorem 4.3 and its corollaries using \( q = 1 - m/k \). \( \square \)
4.3.3 Example 2

Let \( m, \sigma \) be positive real numbers. We will derive an upper bound on the probability of extinction and a lower bound on the expected time to extinction for the multiplicative process for which the random multiplicands \( X_i \) \((i = 1, 2, \ldots)\), conditional on \( Z_0, Z_1, \ldots, Z_{i-1} \), have mean \( m \) and variance at most \( \sigma^2 \).

There are two cases: (i) \((m^2+\sigma^2)/m = h\), or (ii) \(h < (m^2+\sigma^2)/m < h+1\), where \( h \) is a positive integer \( \geq 1 \). (As in §3.3, the case \(0 \leq (m^2+\sigma^2)/m < 1\) is impossible if \(X\) is to take on non-negative integer values only.)

It will turn out that this is a special case of the \( M \)-sequences studied in §4.3.1 with \( q \) chosen appropriately.

Lemma 4.2

Let \( M = \{P(X): EX = m, \ Var X \leq \sigma^2\} \), then for \( P(X) \in M \), we have

\[
Pr[X = 0] \leq q,
\]

where

(i) \( q = 1 - m/h \) if \((m^2+\sigma^2)/m = h\) is a positive integer, or

(ii) \( q = 1 - [m(1+2h) - (\sigma^2+m^2)]/h(h+1) \) if \( h < (m^2+\sigma^2)/m < h+1 \) with \( h \) a positive integer.

Proof

(i) Suppose \( h = (m^2+\sigma^2)/m \) is an integer. Let

\[
f(x) = (x-h)^2/h^2, \ g(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}
\]
Then \( g(x) \leq f(x) \) for all \( x \) (see Figure 4.2) and thus for \( P(X) \in \mathcal{M} \), we have \( E g(X) \leq E f(X) \), or:

\[
Pr[X = 0] \leq E[(X-h)^2/h^2]
\]

\[
= E[(X-m)^2 + (m-h)^2]/h^2
\]

\[
\leq (\sigma^2 + (m-h)^2)/h
\]

\[
= 1 - m/h, \quad \text{since } h = (m^2 + \sigma^2)/m .
\]

(ii) Suppose \( h < (m^2 + \sigma^2)/m < h+1 \) with \( h \) an integer. Let

\[
f(x) = (x-h)(x-h+1)/h(h+1), \quad g(x) = \begin{cases} 
1 & \text{if } x = 0 \\
0 & \text{otherwise}.
\end{cases}
\]
Now \( g(x) \leq f(x) \) for all \( x \notin (h, h+1) \) (see Figure 4.3), and thus for \( P(X) \in \mathcal{M}_0 \), we have \( E_g(X) \leq E_f(X) \), or:

\[
Pr[X = 0] \leq E[(X-h)(X - \bar{h+1})]/h(h+1)
\]

\[
= [\text{Var}(X) + m^2 - (2h+1)m + h(h+1)]/h(h+1)
\]

\[
\leq 1 - [m(1+2h) - (\sigma^2 + m^2)]/h(h+1)
\]

which completes the proof of Lemma 4.2.

**Theorem 4.5**

Let \( \mathcal{M}_0 = \{P(X): EX = m, \text{Var} X \leq \sigma^2\} \) and set \( \mathcal{M}(z) = \{P(zX): P(X) \in \mathcal{M}_0\} \) for \( z \in \mathbb{C} \). Then for \( Z_0, Z_1, \ldots \), an \( \mathcal{M} \)-sequence starting at \( z > 0 \), we have:
\[ \Pr[Z_N = 0] \leq 1 - p^N, \quad (N = 0,1,2,\ldots) , \]

and

\[ E[T_e] \geq (1-p)^{-1} , \]

where (i) if \( h = (\sigma^2 + m^2) / m \) is an integer then \( p = m / h \), or

(ii) if \( h < (\sigma^2 + m^2) / m < h+1 \) with \( h \) a positive integer, then

\[ p = \frac{m(1+2h) - (\sigma^2 + m^2)}{h(h+1)} . \]

These bounds are sharp and are attained by the \( N \)-sequence \( \{Z'_i\} \) for which the multiplicands \( \{X'_i\} \) are i.i.d. with a common distribution given by:

in case (i) \( X = 0 \) with probability \( 1 - m / h \),

\[ = h \text{ with probability } m / h ; \]

in case (ii) \( X = 0 \) with probability \( 1 - \frac{m(1+2h) - (\sigma^2 + m^2)}{h(h+1)} \),

\[ = h \text{ with probability } \frac{m(h+1) - (\sigma^2 + m^2)}{h} , \]

\[ = h+1 \text{ with probability } \frac{\sigma^2 + m^2 - hm}{(h+1)} . \]

**Proof**

The theorem follows directly from Lemma 4.2 and Theorem 4.3 and its corollaries using \( q = 1 - p = 1 - m / h \) in case (i), and

\[ q = 1 - p = 1 - \frac{m(1+2h) - (\sigma^2 + m^2)}{h(h+1)} \text{ in case (ii).} \]
4.3.4 Lower Bounds

Theorem 4.6

Take $0 \leq q \leq 1$, $\mathcal{N} = \{P(X); \ Pr[X = 0] \geq q\}$, and set

$\mathcal{N}(z) = \{P(zX); \ P(X) \in \mathcal{N}\}$ for each $z \in \mathcal{C}$. Then for $Z_0, Z_1, \ldots$, an $\mathcal{N}$-sequence starting at $z$, we have:

$$
Pr[Z_N = 0] \geq 1 - (1-q)^N \quad \text{if } z > 0
$$

$$
= 1 \quad \text{if } z = 0
$$

\hspace{1cm} (N = 0, 1, 2, \ldots).

Proof

We apply Theorem 2.2 with $r(z) \equiv 0$, $T = N$ and

$$
f_k(z) = \begin{cases} 
1 - (1-q)^k & \text{if } z > 0 \\
1 & \text{if } z = 0 
\end{cases}
$$

(k = 0, 1, 2, \ldots).

Then $r$ and $f$ are non-negative, and if $z = 0$ then $r(z) + Ef_k(Z) = Ef_k(zX) = 1 = f_{k+1}(z)$ for $P(Z) \in \mathcal{M}(0)$. If $z > 0$, then
\[ r(z) + E f_k(Z) = Ef_k(zX) \]

\[ = 1 \cdot \Pr[X = 0] + [1 - (1-q)^k] \cdot \Pr[X \neq 0] \]

\[ = (1-q)^k \Pr[X = 0] + 1 - (1-q)^k \]

\[ \geq (1-q)^k \cdot q + 1 - (1-q)^k \]

\[ = 1 - (1-q)^{k+1} \]

\[ = f_{k+1}(z), \]

for \( P(Z) \in \mathcal{M}(z), \ k = 0, 1, 2, \ldots \).

Also the integrability condition, \( E[\mathbf{f}_{N-k}(Z_k) + \sum_{i=0}^{k-1} r(Z_i)] < \infty \)

\( (k = 1, 2, \ldots, N) \), is satisfied since \( r \equiv 0 \) and \( f_k \) is bounded by 0 and 1. Thus the conditions of Theorem 2.2 are satisfied and the result (4.7) follows by noting, as in Theorem 4.3, that:

\[ E[\sum_{k=0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)})] = \Pr[Z_N = 0]. \]

\[ \square \]

**Corollary 1**

If \( 0 < q < 1 \) and \( z > 0 \) then

\[ (4.8) \quad E[T_e] \leq 1/q. \]

**Proof**

The proof goes through as in the proof of Corollary 1 of Theorem 4.3 but with all inequalities reversed.
Corollary 2

The bounds (4.7), (4.8) are sharp and are attained by the $M$-sequence for which the multiplicands $\{X_i\}$ are i.i.d. with $Pr[X_i = 0] = q$.

Proof

The proof is the same as that of Corollary 2 of Theorem 4.3.

4.3.5 Example 3

Let $H \subseteq C$ and $m$ be some positive real number. We shall derive a lower bound for the probability of extinction and an upper bound on the mean time to extinction for the multiplicative process for which, conditional on the past, $Z_0, Z_1, \ldots, Z_{i-1}$, the distribution of the multiplicand $X_i$ is concentrated on $H$ with mean at most $m$, for each $i = 1, 2, \ldots$.

Define $m^*$ to be the smallest number $\geq m$ in $H$ and $m'$ the largest number $\leq m$ in $H$. We assume $m^*$ and $m'$ exist. Of course if $m \in H$ then $m = m^* = m'$.

Lemma 4.3

Let $M = \{P(X): X$ concentrated on $H$, $EX \leq m\}$ then for $P(X) \in M$, we have:

\[
Pr[X = 0] \geq \begin{cases} 
1 - m/m^* & \text{if } m' = 0 \\
0 & \text{if } m' \neq 0
\end{cases}
\]

Proof

Trivially $Pr[X = 0] \geq 0$.

Suppose $m' = 0$, then since $m > 0$ we must have $0 < m < m^*$. Let $f(x) = 1 - x/m^*$ and $g(x) = 1$ if $x = 0$ and $g(x) = 0$ otherwise.
Thus \( g(x) \geq f(x) \) for \( x = m' = 0 \) and for all \( x \geq m^* \), and so \( g(x) \geq f(x) \) for all \( x \) in \( H \). (See Figure 4.4.) Hence for \( P(X) \in M \), \( E_g(X) \geq E_f(X) \), or:

\[
Pr[X = 0] \geq 1 - E[X/m^*] \geq 1 - m/m^* ,
\]

which completes the proof of Lemma 4.3.

\[\square\]

**Theorem 4.7**

Let \( M = \{P(X): X \text{ concentrated on } H, EX \leq m\} \) and set \( M(z) = \{P(zX): P(X) \in M\} \) for all \( z \in C \). Then for \( z_0, z_1, \ldots \), an \( M \)-sequence starting at \( z > 0 \), we have:
\[ \Pr[Z_N = 0] \geq \begin{cases} 
0 & \text{if } m' \neq 0 \\
1 - (m/m^*)^N & \text{if } m' = 0 
\end{cases} 
\]

\[ N = 0, 1, 2, \ldots \] ;

and

\[ E[T_e] \leq \begin{cases} 
\infty & \text{if } m' \neq 0 \\
m^*/(m^*-m) & \text{if } m' = 0 
\end{cases} .
\]

These bounds are sharp and are attained by the \( n \)-sequence for which the multiplicands \( \{X_i\} \) are i.i.d. with common distribution given by:

(i) If \( m = m' = m^* \), \( X = m \) with probability 1,

or

(ii) If \( 0 \leq m' < m < m^* \), \( X = m' \) with probability \((m^*-m)/d\)

\[ = m^* \] with probability \((m-m')/d\),

where \( d = m^* - m' \).

Proof

The proof follows directly by applying Lemma 4.3 and Theorem 4.6 and its corollaries with \( q = 0 \) if \( m' \neq 0 \), and \( q = 1 - (m/m^*) \) if \( m' = 0 \).

4.4 The Probability that a Generation Size Will Exceed a Given Number and the Expected Maximum Generation Size

In this section we will consider the multiplicative process for which, conditional on the past, \( Z_0, Z_1, \ldots, Z_{i-1} \), the random multiplicands \( X_i \) have a mean of at most one, for each \( i \geq 1 \).
Theorem 4.8

Take $\ell > 0$, $\alpha > 1$, $M = \{P(X), E[X] \leq 1\}$ and set
$M(z) = \{P(zX): P(X) \in M\}$. Then for $Z_0, Z_1, \ldots$, an $M$-sequence starting at $z$, we have:

$$\Pr[Z_N \geq \ell \text{ for some } N] \leq \frac{z}{\ell} \quad (0 \leq z \leq \ell),$$

(4.9)

$$= 1 \quad (z > \ell).$$

Proof

Trivially if $Z_0 = z \geq \ell$ then $\Pr[Z_N \geq \ell \text{ for some } N] = 1$. Suppose $z \in C$, $0 < z < \ell$. We apply Theorem 2.3 with

$$T = T_\ell = \min[n: Z_n > \ell] \quad \text{if } Z_n > \ell \text{ for some } n$$

$$= \infty \quad \text{otherwise},$$

$r(z) = 0$ and $f(z) = z/\ell$ if $0 \leq z \leq \ell$, and $f(z) = 1$ for $z > \ell$. Then $r$ and $f$ are non-negative and, since $f(z) \leq z/\ell$ for all $z \geq 0$, we have for $P(Z) \in M(z)$:

$$r(z) + E[f(Z)] \leq E[Z/\ell]$$

$$= E[zX/\ell]$$

$$\leq z/\ell$$

$$= f(z), \quad \text{since } z \leq \ell.$$
Thus the conditions for Theorem 2.3 are satisfied. Define $v(z) = 1$ if $z \leq \ell$ and $v(z) = 0$ otherwise. Note that $f(z) \geq v(z)$ for all $z \geq 0$. Hence, by Theorem 2.3, we have:

$$z/\ell \geq E[f(Z_{\ell}) \cdot I_{T<\infty}] \geq E[v(Z_{\ell}) \cdot I_{T<\infty}] = 1 \cdot \Pr[T < \infty]$$

$$= \Pr[Z_N \geq \ell \text{ for some } N],$$

which completes the proof.

Remark 1  Attainment of the Bounds

If $\ell/z$ is an integer, then the bound is attained when $X_1$ is chosen such that

$$X_1 = \ell/z \text{ with probability } z/\ell,$$

$$= 0 \text{ with probability } 1 - z/\ell.$$
Thus \[ Z_1 = \ell \quad \text{with probability} \quad \frac{\ell}{z} \]
\[ = 0 \quad \text{with probability} \quad 1 - \frac{\ell}{z} , \]

and the optimal strategy is to aim for the goal in one step ("bold play").

If \( \ell/z \) is not an integer, then we cannot choose \( \ell/z \) to be a possible value of \( X_1 \), and such a strategy is not open to us. The effect of overshoot must be taken into account (see Remark 1 of §3.4).

**Remark 2** The Expected Maximum Generation Size

For all \( M \)-sequences starting at \( z \), we have:

\[
(4.10) \quad E[\sup_n Z_n] < \infty .
\]

This bound is sharp and is attained by the \( M \)-sequence for which the multiplicands \( X_i \) are i.i.d. with distribution given by:

\[
X = 0 \quad \text{with probability} \quad 1 - \frac{1}{k} \\
= k \quad \text{with probability} \quad \frac{1}{k} ,
\]

where \( k \) is any integer greater than one. This follows from the fact that for this \( M \)-sequence:

\[
E[\sup_n Z_n] = z \cdot \sum_{i=1}^{\infty} \frac{k^i \cdot (1-1/k) \cdot (1/k)^i}{(1-k)^i} \]
\[ = \infty . \]
4.5 The Analogy Between the Branching Process and the Multiplicative Process Models

It can be seen that the results of §4.1 - 4.4 parallel closely those of §3.1 - 3.4. In fact, the distributions of the i.i.d. litter sizes \( \{X_1\} \) in the branching process model which achieve the bounds are the same as those for the random multiplicands \( \{X_i\} \) which achieve the bounds for the corresponding results in the multiplicative process.

In both chapters, in order to make the analogy closer, \( C \) was taken to be the set of non-negative integers. However, the results of Chapter 4 can be extended to more general \( C \). It is easy to show that Theorems 4.1 - 4.7 can be adapted as follows:

**Theorems 4.1, 4.2** These hold for \( C = [0, \infty) \), (we must have \( r(z) > 0 \)).

**Theorem 4.3** This holds for \( C = (-\infty, \infty) \).

**Theorem 4.4** This holds for \( C = [0, \infty) \) and \( k \) no longer needs to be integer. (If \( z < 0 \), Lemma 4.1 does not hold.)

**Theorem 4.5** This holds for \( C = (-\infty, \infty) \), and \( m \) need no longer be positive. Since \( h = (\sigma^2 + m^2)/m \) is always in \( C \), the result (i) holds. (ii) is inapplicable.

**Theorem 4.6** This holds for \( C = (-\infty, \infty) \).

**Theorem 4.7** This holds for \( C = [0, \infty) \), and \( H \) any subset of \( C \). \( m', m^* \) no longer need to be integers. (Lemma 4.3 does not hold for \( z < 0 \).)

**Theorem 4.8** This holds for \( C = [0, \infty) \). Since now \( \ell/z \) is always in \( C \), the bound (4.9) can always be achieved with \( X_1 = \ell/z \) with probability \( z/\ell \) and \( X_1 = 0 \) with probability \( 1 - z/\ell \). (If \( z < 0 \), then
f(z) \not\in v(z).] To achieve the bound (4.10) for the expected maximum
generation size, k may now be any real number strictly greater than
one.
CHAPTER V
A MULTITYPE BRANCHING PROCESS

5.1 The Multitype Galton-Watson Process

The multitype (or vector) Galton-Watson Process is defined in Harris (1963, Chapter 2) and in Bharuchta-Reid (1960, Chapter 1.8). We shall use the notation of the latter. Vectors and matrices will be denoted by underlined letters, and we shall say that a vector or matrix is 'positive,' 'non-negative,' or 'zero' if all its components have those properties. For n some positive integer, define \( \underline{0} \) and \( \underline{1} \) as the n-vectors with all components 0 or 1, respectively. Also let \( \underline{e}_i \), \( i = 1,2,...,n \), denote the n-vector whose i'th component is 1 and whose other components are 0.

We consider a population in which each individual belongs to precisely one of n distinct types. Let \( Z_N = (Z_{1N},Z_{2N},...,Z_{nN}) \) represent the population at time \( N \), where \( Z_{iN} \) is the number of individuals of type \( i \) alive at time \( N \) \( (i = 1,2,...,n; N = 0,1,2,...) \).

Let \( p_i(a_1,a_2,...,a_n) \) be the probability that an individual of type \( i \) produces \( a_1 \) individuals of type 1, \( a_2 \) individuals of type 2,..., and \( a_n \) individuals of type \( n \) in the next generation \( (i = 1,2,...,n) \). All branches evolve independently of each other.

Let \( F_{iN}(s) \) be the generating function of \( Z_N \) when \( Z_0 = \underline{e}_i \), where \( s = (s_1,s_2,...,s_n) \), for \( 1 \leq i \leq n \); \( N = 0,1,2,... \). Hence:

\[
F_{i0}(s) = s_i,
\]

and
\[ F_{i1}(s) = \sum_{a_1=0}^{\infty} \sum_{a_2=0}^{\infty} \ldots \sum_{a_n=0}^{\infty} p_i(a_1, a_2, \ldots, a_n) s_1^{a_1} s_2^{a_2} \ldots s_n^{a_n} \]

= \( F_i(s) \), say.

A fundamental result (Harris 1963; Chapter 2, Theorem 3.1) is that:

\[ F_{i,N+1}(s) = F_i[F_{1N}(s), F_{2N}(s), \ldots, F_{nN}(s)] \]

\( (i = 1, 2, \ldots, n) \).

Differentiating (5.2) we get:

\[ E[Z_N] = Z_0 \cdot M^N, \]

where \( M = (m_{ij}) \) and \( m_{ij} = E[Z_{j1} | Z_0 = e_i] \). We will need the following two definitions:

The process is called positively regular if \( M^N \) is positive for some positive integer \( N \). In this case, \( M \) has a positive real eigen-value, \( \lambda \) say, that is larger in absolute value than any other eigen-value (Harris 1963; Chapter 2, Theorem 5.1).

The process will be called singular if the generating functions \( F_i(s) \) \((1 \leq i \leq n)\) are all linear with no constant term, i.e. each individual has exactly one offspring.

Define \( \rho_i(N) = \Pr[Z_N = 0 | Z_0 = e_i] \), for \( 1 \leq i \leq n \) and \( N = 0, 1, 2, \ldots \). Then if \( \rho(N) = (\rho_1(N), \ldots, \rho_n(N)) \) and \( F(s) = (F_1(s), \ldots, F_n(s)) \) we have, by (5.1):
(5.3) \[ \varrho(N+1) = F(\varrho(N)) . \]

Since the branches are independent we have the further result that:

(5.4) \[ \Pr[Z_N = 0 | Z_0 = z] = [\varrho_1(N)]^{z_1} \cdot [\varrho_2(N)]^{z_2} \ldots [\varrho_n(N)]^{z_n} , \]

where \( z = (z_1, z_2, \ldots , z_n) \).

Harris (1963; Chapter 2, Theorems 7.1, 7.2) shows that if the process is positively regular and not singular then \( \varrho(N) \rightarrow \varrho \) as \( N \rightarrow \infty \), where

\( \varrho = (\varrho_1, \varrho_2, \ldots , \varrho_n) \) and the convergence is componentwise. Also \( \varrho \) satisfies the equation \( \varrho = F(\varrho) \). Furthermore, if \( \lambda < 1 \) then \( \varrho = 1 \) and the population will become extinct with probability one, whereas if \( \lambda > 1 \), then \( 0 < \varrho < 1 \).

In the remaining sections of this chapter we will consider multitype branching processes where the transition laws \( \{ \varrho_i(a_1, \ldots , a_n); 1 \leq i \leq n \} \) are allowed to vary within a certain class from period to period and at time \( N (\geq 1) \) may depend on the past history \( \{ Z_i, 1 \leq i \leq N-1 \} \) of the process. These processes will be defined as \( N \)-sequences. In \$5.2 \), upper bounds for the probability of extinction and lower bounds for the expected time to extinction are derived and in \$5.3 \) bounds in the reverse direction for the same quantities are obtained. The bounds are sharp and are attained by multitype Galton-Watson processes.

Throughout this chapter (and also Chapter VI), \( C \) will be taken to be the set of all \( n \)-vectors whose components are non-negative integers. As in Chapter III, define \( T_\varepsilon \), the time to extinction by:
5.2 The Probability of Extinction (Upper Bounds) and the Expected Time to Extinction (Lower Bounds)

Theorem 5.1 (Extinction in a finite time)

Take $N$ some positive integer and let $\{\rho(k); k = 0,1,2,\ldots,N\}$ be a sequence of real $n$-vectors such that $0 \leq \rho(k) \leq 1$, $(0 \leq k \leq N)$ and $\rho(0) = 0$; and set

$$
M(z) = \{F(Z_1|Z_0 = z): \text{ where } Z_1 \text{ is the population vector of a multitype Galton-Watson branching process such that } f(k) \leq \rho(k+1), k = 0,1,2,\ldots,N-1\},
$$

for all $z \in C$.

Then for $Z_0, Z_1, Z_2, \ldots$, an $M$-sequence starting at $z = (z_1, z_2, \ldots, z_n)$, we have:

$$
(5.5) \quad \Pr[Z_N = 0] \leq [\rho_1(N)]^{z_1} \cdot [\rho_2(N)]^{z_2} \ldots [\rho_n(N)]^{z_n}.
$$

Proof

This is the multitype analogue of Theorem 3.3. We apply Theorem 2.1 with $C$ the set of non-negative integer $n$-vectors, $T = N$, $r(z) \equiv 0$,

$$
f_k(z) = [\rho_1(k)]^{z_1} \cdot [\rho_2(k)]^{z_2} \ldots [\rho_n(k)]^{z_n} \text{ for } z \in C \text{ and } 0 \leq k \leq N.
$$
Then \( r \) and \( f_k \) are non-negative and, for \( P(Z) \in \mathcal{M}(z) \), we have:

\[
\begin{align*}
r(z) + Ef_k(Z) &= E\{[\rho_1(k)]^z_1 \cdot [\rho_2(k)]^z_2 \cdots [\rho_n(k)]^z_n\} \\
&= \left[ F_1(\rho(k)) \right]^{z_1} \cdot \left[ F_2(\rho(k)) \right]^{z_2} \cdots \left[ F_n(\rho(k)) \right]^{z_n} \\
&\leq [\rho_1(k+1)]^{z_1} \cdot [\rho_2(k+1)]^{z_2} \cdots [\rho_n(k+1)]^{z_n} \\
&= f_{k+1}(z),
\end{align*}
\]

which verifies the hypotheses of Theorem 2.1.

The result (5.5) now follows by noting that

\[
E\left[ \sum_{k=0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)}) \right] = Ef_0(Z_N) = \Pr[Z_N = 0],
\]

since \( \rho(0) = 0 \), and we have adopted the convention that \( 0^0 = 1 \). \( \square \)

**Corollary 1 (The Expected Time to Extinction)**

Suppose \( \{\rho(k); k = 0,1,2,\ldots\} \) is a sequence of real \( n \)-vectors such that \( 0 \leq \rho(k) \leq 1 \), for all \( k \), and \( \rho(0) = 0 \). Take

\[
\mathcal{M}(z) = \{P(Z_1 | Z_0 = z) \text{ where } Z_1 \text{ is the population vector of a multitype Galton-Watson process such that } F(\rho(k)) \leq \rho(k+1), \text{ for all } k\},
\]

for all \( z \in C \).
Then, for \( Z_0, Z_1, Z_2, \ldots \), an \( N \)-sequence starting at \( Z = (z_1, z_2, \ldots, z_n) \), we have:

\[
E[T_e] \geq \sum_{k=0}^{\infty} \left( 1 - \left[ \rho_1(k) \right]^{z_1} \cdot \left[ \rho_2(k) \right]^{z_2} \cdots \left[ \rho_n(k) \right]^{z_n} \right).
\]

**Proof**

The proof is analogous to that of Corollary 1 of Theorem 3.3 and follows by noting that:

\[
E[T_e] = \sum_{k=0}^{\infty} \left( 1 - \Pr[Z_k = 0] \right).
\]

\( \square \)

**Corollary 2 (Achievement of Bounds)**

(A) If there exists an \( N \)-sequence such that at stage \( N-k \) (i.e. with \( k \) periods remaining), conditional on the past, the generating function of the transition probabilities, as defined in (5.1), is \( F^{(k)} \) with

\[
\rho(k) = F^{(k)}(\rho(k-1))
\]

for \( 1 \leq k \leq N \), then this \( N \)-sequence achieves the bound (5.5). The proof follows by induction on \( N \) using (5.3).

(B) If it turns out that \( F^{(k)} \) is independent of \( k \) (\( F^{(k)} = F \) say), then the \( N \)-sequence in (A) is a multitype Galton-Watson process and, by (5.3) and (5.4), this process achieves the bound (5.5) simultaneously for all \( N \), and thus the bound (5.6) on the expected time to extinction is also attained.

This is unlike the problems studied in Chapters III and IV, in all of which
the optimal policy was independent of the time horizon \( N \). There always
the strategy which maximized (minimized) the probability of extinction also
minimized (maximized) the mean time to extinction.

**Theorem 5.2 (Eventual extinction)**

Let \( \rho\), \( 0 \leq \rho \leq 1 \) be a real \( n \)-vector and set

\[
M(z) = \{ P(Z_1 | Z_0 = z) : \text{where } Z_1 \text{ is the population vector of a multitype Galton-Watson branching process such that } f(\rho) \leq \rho, \}
\]

for all \( z \in \mathbb{C} \).

Then for \( Z_0, Z_1, \ldots \), an \( M \)-sequence starting at \( z = (z_1, \ldots, z_n) \), we have:

\[
(5.7) \quad \Pr[Z_N = 0 \text{ for some } N] \leq \rho_1^{z_1} \cdot \rho_2^{z_2} \cdots \rho_n^{z_n}.
\]

**Proof**

This is the multitype analogue of Theorem 3.4, and the proof follows similarly, with \( f(z) = \rho_1^{z_1} \cdot \rho_2^{z_2} \cdots \rho_n^{z_n} \). The details are omitted.

**Corollary 1 (Achievement of bounds)**

If there exists a positively regular, non-singular multitype Galton-Watson process with \( f(\rho) = \rho \) then, as explained in §5.1, this process achieves the bound (5.7).
5.3 The Probability of Extinction (Lower Bounds) and the Expected Time to Extinction (Upper Bounds)

Theorem 5.3

Take \( N \) some positive integer and let \( \{\rho(k); k = 0,1,\ldots,N\} \) be a sequence of real \( n \)-vectors such that \( 0 < \rho(k) < 1, (0 \leq k \leq N) \) and \( \rho(0) = 0 \); and set

\[
M(z) = \{P(Z_1|Z_0 = z)\}: \text{ where } Z_1 \text{ is the population vector of a multitype Galton-Watson process such that } \rho(k+1) \geq \rho(k) \quad k = 0,1,2,\ldots,N-1
\]

for \( z \in C \).

Then, for \( Z_0,Z_1,Z_2,\ldots \), an \( M \)-sequence starting at \( z = (z_1,z_2,\ldots,z_n) \), we have:

\[
(5.8) \quad \Pr[Z_N = 0] \geq [\rho_1(N)]^{z_1} \cdot [\rho_2(N)]^{z_2} \ldots [\rho_n(N)]^{z_n}.
\]

Proof

This is the multitype analogue of Theorem 3.7 and the proof follows similarly, with \( f_k(z) = [\rho_1(k)]^{z_1} \cdot [\rho_2(k)]^{z_2} \ldots [\rho_n(k)]^{z_n} \). The details are omitted.

Corollary 1 (The Expected Time to Extinction)

Suppose \( \{\rho(k); k = 0,1,2,\ldots\} \) is a sequence of real \( n \)-vectors \( 0 < \rho(k) < 1, \) for all \( k \), and \( \rho(0) = 0 \). Take
\[ M(z) = \{ P(Z_1 | Z_0 = z) : \text{where } Z_1 \text{ is the population} \]
\[ \text{vector of a multitype Galton-Watson process} \]
\[ \text{such that } P(\rho(k)) > \rho(k+1), \text{ for all } k \]

for all \( z \in C \).

Then, for \( Z_0, Z_1, Z_2, \ldots \), an \( M \)-sequence starting at \( z = (z_0, z_1, \ldots, z_n) \), we have:

\[ E[T_e] \leq \sum_{k=0}^{\infty} \{ 1 - [\rho_1(k)]^{z_1} \cdot [\rho_2(k)]^{z_2} \cdots [\rho_n(k)]^{z_n} \} . \]

\[ \tag{5.9} \]

**Proof**

The proof is the same as that for Corollary 1 of Theorem 5.1 but with the inequalities reversed.

\[ \square \]

**Corollary 2 (Achievement of Bounds)**

(A) If there exists an \( M \)-sequence such that at stage \( N-k \) (i.e. with \( k \) periods remaining), conditional on the past, the generating function of the transition probabilities, as defined in (5.1), is \( F(k) \) with

\[ \rho(k) = F(k)(\rho(k-1)) \]

for \( 1 \leq k \leq N \), then this \( M \)-sequence achieves the bound (5.8).

(B) If it turns out that \( F(k) \) is independent of \( k \) (\( F(k) = F \) say), then the \( M \)-sequence in (A) is a multitype Galton-Watson process and this process achieves the bound (5.8) simultaneously for all \( N \), and thus the bound (5.9) on the expected time to extinction is also attained.

This result follows as in Corollary 2 of Theorem 5.1.
Corollary 3 (Eventual Extinction)

Suppose \( \rho(k) \to \rho \) as \( k \to \infty \) for some \( \rho, 0 < \rho < 1 \), then, since
\[
Pr[Z_N = 0 \text{ for some } N] > Pr[Z_N = 0]
\]
and the left-hand side is independent of \( N \), we have:

\[
(5.9) \quad Pr[Z_N = 0 \text{ for some } N] > \prod_{i=1}^{n} \rho_i.
\]

Furthermore if there exists a multitype Galton-Watson process with \( F(\rho) = \rho \) then, as explained in §5.1, this process achieves the bound (5.9).

Note on Existence

For the theorems in this chapter we have assumed that \( \mathcal{M} \)-sequences, as defined, do exist. In fact, for general \( \{ \rho(k) \} \), \( \mathcal{M}(z) \) may be empty for some or all \( z \in \mathbb{C} \), and no such multitype Galton-Watson process as defined in the corollaries will exist. In Chapter VI, a certain interesting class of multitype branching processes is described, and the \( \{ \rho(k) \} \) are chosen in such a way that this class is a subset of the \( \mathcal{M} \)-sequences defined in this chapter.
6.1 Introduction

"Multaque tum interiisse animantium saecla necesse est, 
Nec potuisse propagando producere prolem. 
Nam quaecomque uides uesci uitalibus aureis, 
Aut dolus, aut uitus aut denique mobilitas est 
Ex ineunte aeuo genus id tutata, reservans."¹

(Lucretius, *De Rerum Natura*)

Thus Lucretius stated the principle of natural selection, in a rather crude form, some two thousand years before Darwin published his book, "The Origin of the Species" (1859). The first efforts to apply mathematics to explain evolutionary phenomena took place in the 1920's. Pearson (1930) accumulated numerous data on the inheritance of various characteristics in man and made some conclusions about human populations. Fisher (1930), Haldane (1932), and Wright (1931) developed theories of selection in Mendelian populations, while Lotka (1925) and Volterra (1926, 1931) applied mathematics to the study of populations and ecology. Much of the recent work in this area has stemmed from a review article by Cole (1954). The present day status of the problem has been summarized by MacArthur and Wilson (1967).

The theory of natural selection implies that, within a given species, the genotype most likely to survive will be the one which is fittest in its

¹ "And many lines of organisms must have perished then, and been unable to propagate their kind. For whatever you see feeding on the vital air, either craft, strength, or finally mobility has been protecting and preserving that race from its earliest times."
environment. In the next few paragraphs we will discuss various measures of fitness.

The most common measure of fitness is \( r \), the intrinsic rate of natural increase. Lotka (1925) showed that, asymptotically, the number of individuals in any age group grows exponentially with time at a common rate \( r \), provided that the age-specific birth and death rates remain constant. In fact, \( r \) is the solution to the equation:

\[
\int_0^\infty e^{-rx} \lambda(x)m(x)dx = 1,
\]

where \( \lambda(x) \) is the probability of an individual living to age \( x \) and \( m(x) \) is the mean number of offspring produced per unit time at age \( x \). This parameter \( r \) has also been called the "Malthusian rate" (by Fisher 1930) and the "biotic potential" (by Chapman 1931). Fisher (1930), Haldane (1932), and Wright (1931) were the first to use \( r \) as a measure of fitness ("r-selection"). Cole (1954) compared the intrinsic rate of increase for semelparous and iteroparous organisms. Lewontin (1965) showed how \( r \) was affected by changes in age of first reproduction, age of last reproduction, age of greatest fecundity, and total fecundity (\( \int_0^\infty \lambda(x)m(x)dx \)). He illustrated his results with life history data for the rice-beetle, *Calandra oryzae*, and the fruit-fly, *Drosophila serrata*. Murphy (1968) considered a model for two competing populations and used \( r \) to compare their success.

MacArthur and Wilson (1967, Chapter 4) developed a model for the population of a colonizing species, for which the birth and death rates are density dependent and for which there is an upper bound \( K \) on the size of the population due to environmental restrictions. (\( K \) is called the "carrying capacity of the environment." ) MacArthur and Wilson considered both the carrying
capacity, \( K \), ("K-selection") and the mean time to extinction, \( T \), as measures of fitness. (Since all populations are bounded in size, they will become extinct with probability one as long as the death rate is always positive.)

Holgate (1967) considered an unbounded stochastic population model and then proposed the probability of extinction as a measure of fitness. He derived the result that, with geometric or Poisson offspring distributions, biennials have a lower probability of extinction than annuals with the same intrinsic rate of growth \( r \). The most recent work is by Gadgil and Bossert (1970). They proposed a general model in which natural selection takes place, and considered general "profit" and "cost" functions as measures of fitness, in order to determine optimal ages for reproduction.

In this chapter we will formulate a model for the life history of a species. It will be a special case of the multitype branching process discussed in Chapter V, with each type corresponding to an age group of the population.

(A different model is proposed by Mode (1968). He develops a theory of multidimensional age-dependent branching processes of the Bellman-Harris type (see Harris 1963, Chapter 6), and, in a brief note at the end of the paper, he suggests that these processes can be applied in the study of natural selection.)

We will consider two different measures of fitness, namely:

(i) the probability of extinction,

(ii) the expected total population size.

In (ii), each individual is only counted once, say in the year it is new-born. It will turn out that the life history strategy that maximizes the expected total population size also maximizes the total fecundity \( R = \int_0^\infty \kappa(x)m(x)dx \).
$R$ is a measure of the growth rate of the population. (For its relation to $r$, the intrinsic rate of increase, see Lotka 1925, page 87.)

$\mathcal{N}$-sequences will be defined, each of which will correspond to the possible life histories of different genotypes. Lower bounds for the objective (i) and upper bounds for the objective (ii) will be derived and those $\mathcal{N}$-sequences which attain the bounds will define the optimal life history strategies.

6.2 The Model

We will use a generalization of the population model proposed by Lewis (1942) and Leslie (1945, 1948). We consider only a population of females (we could just as easily consider only males), and assume that the population is partitioned into $n$ distinct age groups, labelled $1, 2, \ldots, n$. At age $i$ ($1 \leq i \leq n$), with probability $\pi_i$, an individual does not make a reproductive effort. Usually $\pi_i$ will be zero or one but if $0 < \pi_i < 1$ then, for example, it might be regarded as the probability of failure to find a mate. In this case its survival probability to age $i + 1$ is $\theta_i$ (with $\theta_n = 0$). With probability $1 - \pi_i$, the individual does make a reproductive effort and produces $x$ offspring with probability $p_i(x)$, $(x = 0, 1, 2, \ldots)$. In this case its survival probability to age $i + 1$ is $\gamma_i$ (with $\gamma_n = 0$).

In his model, essentially Leslie assumed $\pi_i = 0$ for all $i$ and, since the model was deterministic, he did not need to specify the entire age-specific offspring distributions but only the mean litter sizes, $\bar{m}_i = \sum_{x=0}^{\infty} x p_i(x)$. Typically $\theta_i \geq \gamma_i$ since the hazards of reproduction are added to those of predation and disease. Here we make the assumption that $\gamma_i$ is independent of the litter size $x$. (In fact, Murdoch (1966) studied the flour beetle, Carabidae, and showed that the survival of the adult female from near the end
of one breeding season to the start of the next is approximately inversely proportional to the litter size in that first breeding season.) Also we might expect that, to a certain extent, the mean litter size $m_i$ will increase with size and therefore with age $i$ of the parent. This is true of some fish for example (see Lack 1954, page 180). However Lack (page 96) also showed that for some birds, e.g. the lapwing (*Vanellus vanellus*), $\theta_i$ and $m_i$ are independent of $i$.

In the case of semelparous organisms (or "big-bang" reproducers) where reproduction involves the death of the parent, we would have $\gamma_i = 0$ for all $i$. Examples of such organisms are bamboo (e.g. *Arundinaria*) and the Pacific salmon (*Oncorhynchus*). Also, as in the case of many insects, if an individual can have at most one litter during its lifetime then the quantities (i) the probability of extinction or (ii) the expected total population size are unaffected by taking $\gamma_i = 0$ for all $i$.

In the terminology of Chapter V, $Z_{iN}$ represents the number of individuals of age $i$ alive at time $N$ for $i = 1,2,\ldots,n$; and $N = 0,1,2,\ldots$. Births and deaths are assumed to take place at times $0^+,1^+,2^+,\ldots$ etc. The transition functions are defined as follows:
For $i = 1,2,\ldots,n-1,$

\[
P_i(\alpha_1,\alpha_2,\ldots,\alpha_n) = \begin{cases} 
\begin{aligned}
\pi_i \delta_i & \quad \text{if } \alpha_{i+1} = 1; \alpha_j = 0 \text{ for } j \neq i + 1; \\
(1-\pi_i)\gamma_i p_i(x) & \quad \text{if } \alpha_1 = x, \alpha_{i+1} = 1; \alpha_j = 0 \text{ for } j \neq 1, i + 1; x = 0,1,2,\ldots; \\
(1-\pi_i)(1-\gamma_i)p_i(x) & \quad \text{if } \alpha_1 = x; \alpha_j = 0 \text{ for } j \neq 1; x = 1,2,\ldots; \\
(1-\pi_i)(1-\gamma_i)p_i(0) & \\
+ \pi_i (1-\delta_i) & \quad \text{if } \alpha_i = 0 \text{ for all } i; \\
0 & \quad \text{otherwise}; 
\end{aligned}
\end{cases}
\]

and

\[
P_n(\alpha_1,\alpha_2,\ldots,\alpha_n) = \begin{cases} 
\begin{aligned}
(1-\pi_n)p_n(x) & \quad \text{if } \alpha_1 = x; \alpha_j = 0 \text{ for } j \neq 1; x = 1,2,\ldots; \\
\pi_n + (1-\pi_n)p_n(0) & \quad \text{if } \alpha_j = 0 \text{ for all } j.
\end{aligned}
\end{cases}
\]

Let $G_i, \phi_i$ be the distribution function and the probability generating function of the litter size of a parent of age $i$ ($i = 1,2,\ldots,n$), that is:

\[
G_i(x) = \sum_{y=0}^{x} p_i(y) ; \quad \phi_i(\alpha) = \sum_{y=0}^{\infty} \alpha^y \cdot p_i(y) ,
\]

for $x = 0,1,2,\ldots,$ and $\alpha \geq 0.$
Let \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \), \( \theta = (\theta_1, \theta_2, \ldots, \theta_n) \), \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n) \) and 
\( G = (G_1, G_2, \ldots, G_n) \), where of course \( \theta_n = \gamma_n = 0 \). Then \( \pi, \theta, \gamma, \) and \( G \) define a multitype Galton-Watson process \( Z_0, Z_1, Z_2, \ldots \). We shall call such a process a "Life History process" (or "L. H. process") with parameters \((\pi, \theta, \gamma, G)\).

We let \( \Pi, \Theta, \Gamma \) be sets of real \( n \)-vectors with components lying between 0 and 1 (inclusive). For \( \Theta \) and \( \Gamma \) we further require that the last component of all members be zero. For \( 1 \leq i \leq n \), let \( G_i \) be a set of distribution functions concentrated on the non-negative integers, and define \( G = (G_1, G_2, \ldots, G_n) \). We say that \( G \in G \) if \( G_i \in G_i \) for all \( i \) (1 \( \leq i \leq n \)).

We define \( M = (\Pi, \Theta, \Gamma, G) \) and we say that \((\pi, \theta, \gamma, G) \in M\) if \( \pi \in \Pi, \theta \in \Theta, \gamma \in \Gamma, \) and \( G \in G \). \( M \) can be viewed as indicating the genealogical variation possible within the colonizing species. For instance, if \( M \) consists of only one element then the species has no scope to adapt and evolve.

As in Chapter V we take \( C \) to be the set of all non-negative integer \( n \)-vectors, and we consider \( M \)-sequences starting at \( z \) defined by:

\[(6.1) \quad M(z) = \{ P(Z_1 = z) : \text{where } (Z_n; n = 0, 1) \text{ is a L. H. process} \}
\]

with parameters \((\pi, \theta, \gamma, G) \in M\).

If \( M \) contains exactly one element, then the sequence defined is a L. H. process.

In the following sections, for selected \( M \), we shall find bounds on (i) the probability of extinction (both with finite and infinite horizons), and (ii) the mean total population size. The bounds are sharp and the
\(N\)-sequences that attain them are identified. The mechanisms whereby nature may tend to select the optimum life history parameters \(\pi, \theta, \gamma, G\) with respect to (i; infinite horizon) and (ii) will be called "E-selection" and "S-selection" respectively. These terms are consistent with the terms "r-selection" and "K-selection" of MacArthur and Wilson (1967).

6.3 E-selection: Minimizing the Probability of Extinction

In this section, we find sharp lower bounds for the probability of extinction in some finite time \(N\) for a class of \(N\)-sequences defined in (6.1). In Theorem 6.2, these results are extended to the infinite horizon case and it is the life history strategies which minimize the probability of eventual extinction that will be called fittest with respect to E-selection. Three special cases are considered; more examples will be given in §6.5.

Theorem 6.1 (Extinction in a finite time)

Take \(0 \leq \pi^L \leq \pi^U \leq 1\), \(0 \leq g \leq 1\), \(0 \leq t \leq 1\), where \(\pi^L = (\pi^L_1, \ldots, \pi^L_n)\), \(\pi^U = (\pi^U_1, \ldots, \pi^U_n)\), \(g = (g_1, \ldots, g_n)\), and \(t = (t_1, \ldots, t_n)\) with \(g_n = t_n = 0\). For \(1 \leq i \leq n\), let \(\phi_i^*(\alpha)\) be a continuous increasing non-negative function for \(0 \leq \alpha \leq 1\), with \(\phi_i^*(1) = 1\); and set

\[
\mathcal{U} = \{\pi \in \mathbb{R}^n: \pi^L \leq \pi \leq \pi^U\}
\]

\[
\mathcal{S} = \{\theta \in \mathbb{R}^n: 0 \leq \theta \leq t\}
\]

(6.2)

\[
\mathcal{G} = \{\gamma \in \mathbb{R}^n: 0 \leq \gamma \leq g\}
\]

\[
\mathcal{G}_i = \{P(X): E[\alpha^X] \geq \phi_i^*(\alpha), 0 \leq \alpha \leq 1\}.
\]

Here \(\mathbb{R}^n\) denotes the space of all real \(n\)-vectors.
Define the sequence of n-vectors \( \{ \rho(0), \rho(1), \ldots \} \) by the recursive relations: 
\[
\rho(i+1) = \min_{y=\pi_i^L, \pi_i^U} (y[(1-t_i) + \tau_i \rho_{i+1}(j)] + (1-y)\phi_i^*(\rho_1(j))[(1-g_i) + g_i \rho_{i+1}(j)]) ,
\]
\[
(1 \leq i \leq n-1),
\]
(6.3)

\[
\rho_n(j+1) = \pi_n^L + (1-\pi_n^L)\phi_n^*(\rho_1(j)) ,
\]
for \( j = 0,1,2, \ldots \).

Then for \( Z_0, Z_1, \ldots \), an \( \mu \)-sequence starting at \( z = (z_1, z_2, \ldots, z_n) \) as defined by (6.1), we have:

\[
(6.4) \quad \Pr[Z_N = 0] \geq [\rho_1(N)]^{z_1} \cdot [\rho_2(N)]^{z_2} \ldots [\rho_n(N)]^{z_n} ,
\]
for \( N = 0,1,2, \ldots \).

\textbf{Proof}

We will apply Theorem 5.3. First note that \( 0 \leq \rho(k) \leq 1 \) for all \( k = 0,1,2, \ldots \). Considering the generating function \( F \), we have, for \( 1 \leq i \leq n-1 \):
\[ F_i(\rho(k)) = \pi_i \left[ (1 - \theta_i) + \theta_i \rho_{i+1}(k) \right] + (1 - \pi_i) \phi_i(\rho_i(k)) \gamma_i \rho_{i+1}(k) + 1 - \gamma_i \]
\[ \geq \pi_i \left[ (1 - t_i) + t_i \rho_{i+1}(k) \right] + (1 - \pi_i) \phi_i(\rho_i(k)) \gamma_i \rho_{i+1}(k) + 1 - \gamma_i \]
(since \( \theta_i \leq t_i \), \( \gamma_i \leq g_i \) and \( 0 \leq \rho_{i+1}(k) \leq 1 \))
\[ \geq \pi_i \left[ (1 - t_i) + t_i \rho_{i+1}(k) \right] + (1 - \pi_i) \phi_i^*(\rho_i(k)) \gamma_i \rho_{i+1}(k) + 1 - \gamma_i \]
\[ \geq \rho_i(k+1) \quad \text{(by construction)} \]

Also,
\[ F_n(\rho(k)) = \pi_n + (1 - \pi_n) \phi_n(\rho_n(k)) \]
\[ \geq \pi_n^L + (1 - \pi_n^L) \phi_n^*(\rho_n(k)) \]
\[ = \rho_n(k+1) \]

Thus \( F(\rho(k)) \geq \rho(k+1) \) for all \( k \), and the result now follows by Theorem 5.3.

For \( i = 1, 2, \ldots, n \) and \( 0 \leq \alpha \leq 1 \), suppose there exist random variables \( \{X_i^\prime\} \) (with corresponding distribution functions \( \{G_i^\prime\} \)) such that
\[ E[\alpha X_i^\prime] = \phi_i^*(\alpha) \]. Let \( \mathcal{G}^\prime = (G_1^\prime, \ldots, G_n^\prime) \), then by Corollary 2(A) of Theorem 5.3, the bound (6.4) is achieved by the \( n \)-sequence with, at all stages, \( \mathcal{G} = \mathcal{G}^\prime \), \( \theta = t, \gamma = g \) and, at stage \( N-j \) (\( 0 \leq j \leq N-1 \)), i.e. with \( j \) periods remaining, \( \pi = \pi(j) \), where
\begin{equation}
\pi_i^j = \begin{cases} 
\pi_i^U & \text{if } (1-t_i) + t_i \rho_{i+1}(j-1) < \phi_1^*(\rho_1(j-1))[(1-g_1) + g_1 \rho_{i+1}(j-1)] \\
\pi_i^L & \text{otherwise}.
\end{cases}
\end{equation}

If it happens that \( \pi_i^j \) is independent of \( j \), then the optimal policy is stationary and the above \( \pi \)-sequence is a L. H. process. In this case, the upper bound on the mean time to extinction, given by (5.9) with the \( \{ \rho(k) \} \) defined by (6.3), is attained and the results of Corollary 2(B) of Theorem 5.3 apply.

However, in general, the bound (6.4) cannot be attained by a L. H. process since, proceeding optimally, the underlying transition law \( (\pi, \theta, \gamma, G) \) changes from period to period because \( \pi \) does. Of course (5.9) is still an upper bound on the mean time to extinction although now it is not necessarily sharp.

Minimizing the probability of extinction with a finite time horizon may not be very interesting as a possible criterion for natural selection. Theorem 6.1 does lead to an upper bound on the mean time to extinction, a measure of fitness considered by MacArthur and Wilson (1967, Chapter 4). However, in general, this bound is not sharp and thus cannot be used to find the strategy which maximizes the mean time to extinction.

The next theorem will give us a bound on the probability of eventual extinction, a much more interesting criterion and one which has been proposed by Holgate (1967). The results of Theorem 6.1 will be needed in the proof.
Theorem 6.2 (Eventual extinction)

With the setup of Theorem 6.1, for \( Z_0, Z_1, \ldots \), an \( M \)-sequence starting at \( Z = (z_1, z_2, \ldots, z_n) \), we have:

\[
(6.6) \quad \Pr[Z_N = 0 \text{ for some } N] \geq \rho_1 \cdot \rho_2 \cdots \rho_n,
\]

where \( \rho = (\rho_1, \rho_2, \ldots, \rho_n) \) is the smaller root (using the Euclidean norm, say) in the region \( 0 < \rho < 1 \) of the equations:

\[
\rho_i = \min \{ y[(1-t_i) + t_i \rho_{i+1}] + (1-y)\phi^*(\rho_1)[(1-g_i) + g_i \rho_{i+1}] \} \quad (1 \leq i \leq n-1)
\]

\[
(6.7)
\]

\[
\rho_n = \pi_n^L + (1-\pi_n^L)\phi^*(\rho_1).
\]

For \( i = 1, 2, \ldots, n \) and \( 0 < \alpha < 1 \), suppose there exist random variables \( \{X_i^\prime\} \) (with corresponding distribution functions \( \{G_i^\prime\} \)) such that

\[
E[\alpha X_i^\prime] = \phi^*_i(\alpha).
\]

Set \( G_i^\prime = (G_1^\prime, \ldots, G_n^\prime) \), then the bound (6.6) is sharp and is attained by the L. H. process with parameters \( (\hat{\pi}, t, g, G') \) where:

\[
(6.8) \quad \hat{\pi}_i = \begin{cases} 
\pi_i^U & \text{if } 1-t_i + t_i \rho_{i+1} < \phi^*_i(\rho_1) \cdot [(1-g_i) + g_i \rho_{i+1}] \\
\pi_i^L & \text{otherwise}.
\end{cases}
\]
Proof

For $1 \leq i \leq n$ and $N \geq 1$, take $\rho_i(N)$ to be defined by (6.3). The bound (6.4) of Theorem 6.1 is sharp and hence $\rho_i(N) = \Pr[Z_N = 0 | Z_0 = e_i]$ for some $\mathcal{U}$-sequence $\{Z_i ; i = 0,1,2,\ldots\}$. Since $\{0\}$ is an absorbing state, we have $\rho_i(N) \leq \rho_i(N+1)$. Now $0 \leq \rho_i(N) \leq 1$ for all $i$, $N$, and hence, by monotone convergence, $\rho(N) \to \rho$, with $\{\rho_i, 1 \leq i \leq n\}$ as defined by (6.7). The result now follows by Corollary 3 of Theorem 5.3. The attainment of the bounds follows by noting that for the L. H. process with parameters $(\mu, t, g, G')$, the generating function $F$ satisfies the relation $\rho = F(\rho)$.

Note that here the optimal policy is stationary, and the bound is attained by a L. H. process. We shall now give three examples to illustrate the results of Theorems 6.1, 6.2.

Example 1

We first consider a multitype analogue of the Example 3 of §3.3.5. Suppose that at age $i$ ($1 \leq i \leq n$) an individual, given that it makes a reproductive effort, may produce a litter of size zero or at least $k_i$, but with mean at most $m_i$, where $0 < m_i < k_i$ and $k_i$ is some positive integer.

More precisely, the class of possible offspring distributions for a parent of age $i$ ($1 \leq i \leq n$) is given by:

$$G_{\alpha_i} = \{P(X) : E[X] \leq m_i, X \text{ concentrated on } H_i\},$$

where $H_i = \{0, k_i, k_i + 1, k_i + 2,\ldots\}$.

If we choose $\phi_i^*(\alpha) = 1 - m_i/k_i + \alpha \cdot k_i \cdot m_i/k_i$ ($0 \leq \alpha \leq 1$), then, by applying Lemma 3.3 with $\beta = \alpha$, $m' = 0$, $m^* = k_i$, we have:
\[ E[\alpha^X] \geq \phi^*_i(\alpha) , \]

for \( 0 \leq \alpha \leq 1, \ P(\alpha) \in G_i^c. \)

The \( \{G_i^c\} \) are now of the form specified in Theorems 6.1 and 6.2, and we can apply all the results obtained in this section.

The optimal litter size distributions, for \( 1 \leq i \leq n, \) are given by:

\[
p_i(x) = \begin{cases} 
1 - \frac{m_i}{k_i}, & \text{if } x = 0, \\
\frac{m_i}{k_i}, & \text{if } x = k_i, \\
0, & \text{otherwise}; 
\end{cases}
\]

since for a random variable \( X \) with this distribution (\( G \) say), we have \( E[\alpha^X] = \phi^*_i(\alpha), \) for \( 0 \leq \alpha \leq 1. \) As in Example 3 of §3.3.5 we see that "timid play is optimal."

**Example 2 (Semelparous organisms with certain survival)**

Suppose each individual has complete freedom of choice whether to make a reproductive effort or not. If it does, it dies with probability one, while if it does not, it will live on to the next period with certainty. Then \( p_i^L = 0, \ p_i^U = 1, \ g_i = 0 \ (1 \leq i \leq n), \ t_i = 1 \ (1 \leq i \leq n-1), \) and \( t_n = 0. \)

Then the defining relations (6.3) for the \( \{p(k)\} \) become:

\[
\rho_i(k+1) = \min[\rho_{i+1}(k), \phi^*_i(\rho_1(k))], \ (1 \leq i \leq n-1),
\]

\[
\rho_n(k+1) = \phi^*_n(\rho_1(k)) .
\]
Here the two terms inside the square brackets correspond to the alternative strategies of deciding to live to the next period or of choosing to reproduce and die.

The limiting value \( \rho \) is defined by:

\[
\rho_n = \phi^*_n(\rho_1)
\]

\[
\rho_i = \min[\rho_{i+1}, \phi^*_i(\rho_1)], \text{ i = n-1, n-2, \ldots, 2, 1}.
\]

Hence \( \rho_1 = \min[\phi^*_1(\rho_1), \phi^*_2(\rho_1), \ldots, \phi^*_n(\rho_1)] \). (Note that the function \( \psi(\alpha) = \min[\phi^*_1(\alpha), \ldots, \phi^*_n(\alpha)] \) is continuous and increasing for \( 0 \leq \alpha \leq 1 \) with \( \psi(1) = 1 \).)

The stationary optimal policy is therefore to choose \( G = G' \), (as defined in Theorem 6.2), \( \theta = 1, \gamma = 0, \) and, for \( 1 \leq i \leq n \):

\[
\pi_i = \begin{cases} 
1 & \text{if } \rho_{i+1} < \phi^*_i(\rho_1) \text{ and } 1 \leq i \leq n-1, \\
0 & \text{otherwise.}
\end{cases}
\]

The optimum genotype reproduces at that age \( L \), for which

\[
\phi^*_L(\rho_1) \leq \phi^*_i(\rho_1) \text{ for all } 1 \leq i \leq n.
\]

**Example 3 (Identical litter size distributions)**

This example is a special case of both the two previous examples. Take \( g, t, \pi^L, \pi^U \) as defined in Example 2, and \( G_{\theta i} \) as defined in Example 1 but
with $m_i = m$, $k_i = k$ for all $1 \leq i \leq n$. Thus now $\phi^*_i(a) = 1 - m/k + a^{k} \cdot m/k = \phi^*(a)$ say.

The defining relations (6.3) for $\{\rho(k)\}$, $k = 0, 1, 2, \ldots$ become:

$$\rho_i(k+1) = \begin{cases} \min[\rho_{i+1}(k), \phi^*(\rho_i(k))] & \text{if } 1 \leq i \leq n-1, \\ \phi^*(\rho_i(k)) & \text{if } i = n. \end{cases}$$

It is easy to show by induction that this implies that $\rho_i(k+1) = \rho_i(k+1)$ for $1 \leq i \leq n-1$, and that, if $N = r \cdot n + s$ with $r$, $s$ non-negative integers, then:

$$\rho_i(N) = \begin{cases} u_r & \text{if } 1 \leq i \leq n-s \\ u_{r+1} & \text{if } n-s+1 \leq i \leq n, \end{cases}$$

where $\{u_r; r = 0, 1, 2, \ldots\}$ are defined by $u_0 = 0$, $u_{r+1} = \phi^*(u_r)$.

The optimal policy is to choose $\pi_i = 1$ $(1 \leq i \leq n-1)$ and $\pi_n = 0$, i.e. to live for as long as possible and then reproduce. This result is intuitive since now the litter size distributions are independent of the age of the parent so nothing is lost by delaying reproduction and death (of the parent) for as long as possible.

Since the optimal choice of $\pi$ is now independent of the number of periods remaining, we have that the bounds (6.4) on the probability of extinction by time $N$ for all $N$ and the corresponding bound (5.9) on the mean time to extinction are attained by the L. H. process with parameters.
(\hat{n}, \hat{t}, 0, G), \text{ where } \hat{n} = \hat{t} = (1,1,\ldots,1,0) \text{ and } G = (G,G,\ldots,G) \text{ with } G \text{ as defined in Example 1. (See the discussion following the proof of Theorem 6.1.) This policy is stationary and also minimizes the probability of eventual extinction. In fact, } \rho_i(N) \rightarrow \alpha \text{ for all } i, \text{ where } \alpha \text{ is the smaller root in } [0,1] \text{ of } \alpha = 1 - \frac{m}{k} + \frac{\alpha^k}{m/k}. \text{ Thus by Theorem 6.2 we have:}

\[ \Pr[Z_N = 0 \text{ for some } N] \geq \alpha^{z_1+z_2+\ldots+z_n}, \]

for all \( \mathcal{N} \)-sequences starting at \( z = (z_1,z_2,\ldots,z_n) \).

Remark: Upper bounds

Using Theorems 5.1 and 5.2, upper bounds can be derived for the probability of extinction in a similar manner. It is not planned to do so here but they are of some interest as they could explain how some species become extinct, or indicate the optimal strategies of use in pest population control.

6.4 S-selection: Maximizing the Expected Total Population Size

In this section we find sharp upper bounds on the expected total population size, \( E[S] \), for a class of \( \mathcal{N} \)-sequences as defined in (6.1). For the litter size distributions, \( \{G_i\} \), only the means are restricted. These \( \mathcal{N} \)-sequences are defined precisely in (6.11). It will turn out that the strategy that maximizes \( E[S] \) also maximizes the total fecundity \( R \), or the expected total number of offspring produced by an individual during its lifetime. The life history strategies which maximize these quantities will be called fittest with respect to S-selection. Two special cases are considered; more examples will be given in $6.5$.
Suppose \( Z_0, Z_1, \ldots \) is an L. H. process with parameters \((\pi, \theta, \gamma, \alpha)\), where the mean of the distribution \( G_i \) is \( \mu_i \) \((1 \leq i \leq n)\). Let
\[
Q_i = E[\sum_{k=0}^{\infty} Z_{1k} | Z_0 = e_i].
\]
Then since every individual, born after stage 0, must go through a period of being age 1, we have that the expected total population size \( E[S] \), given \( Z_0 = (z_1, z_2, \ldots, z_n) \), is equal to \(
\sum_{i=2}^{\infty} z_i + \sum_{i=1}^{n} z_i Q_i.
\)
Thus the problem is equivalent to one of maximizing \( \sum_{i=1}^{n} z_i Q_i \).

For \(1 \leq i \leq n\), define \( a_i = \pi_i \theta_i + (1-\pi_i)\gamma_i \) and
\[
A_i = (1-\pi_i)\mu_i + a_{i-1}(1-\pi_{i-1})\mu_{i-1} + \ldots + a_{i-1}a_{i-1} \ldots a_{n-1}(1-\pi_{n-1})\mu_{n-1}.
\]
Thus \( a_i \) represents the conditional probability that an individual reaches age \( i+1 \) given that it reaches age \( i \), and \( A_i \) is the expected number of offspring an individual will have during the rest of its life given that it is presently of age \( i \).

We note that the \( \{A_i\} \) satisfy the (backward) recurrence relations:

\[
A_n = (1-\pi_n)\mu_n
\]
\[(6.9)\]
\[
A_k = \pi_k \theta_k A_{k+1} + (1-\pi_k)(\gamma_k A_{k+1} + \mu_k)
\]
\[= a_k A_{k+1} + (1-\pi_k)\mu_k \]
\((k = n-1, n-2, \ldots, 2, 1)\).

Then \( \{Q_i : 1 \leq i \leq n\} \) satisfy the following linear equations:
\[ Q_1 = 1 + a_1 Q_2 + (1-\pi_1) \mu_1 Q_1 \]

\[ Q_i = a_i Q_{i+1} + (1-\pi_i) \mu_i Q_1 \quad (2 \leq i \leq n-1) \]

\[ Q_n = (1-\pi_n) \mu_n Q_1. \]

Since \( Q_i > 0 \), the solution to these equations is:

\[ Q_i = A_i Q_1 \quad \text{for } 2 \leq i \leq n, \]

where

\[ Q_1 = \begin{cases} \frac{1}{1-A_1} & \text{if } A_1 < 1, \\ \infty & \text{otherwise.} \end{cases} \]

Hence

\[ E\left[ \sum_{n=0}^{\infty} z_{1N} | Z_0 = z \right] \]

\[ = E\left[ \sum_{i=1}^{n} z_i Q_i \right] \]

\[ = \begin{cases} (z_1 + \sum_{i=2}^{n} A_i z_i)/(1-A_1) & \text{if } A_1 < 1, \\ \infty & \text{if } A_1 \geq 1. \end{cases} \]

\( M \)-sequences

Take \( 0 \leq \pi^L \leq \pi^U \leq 1, \)

\( 0 \leq g < 1, \quad 0 \leq t < 1 \) (with \( g_n = t_n = 0 \)) and \( m_i > 0 \) \( (1 \leq i \leq n) \). We will consider \( M \)-sequences as defined in (6.1) with
\( \mathbb{U} = \{ \pi \in \mathbb{R}^n : \pi^L \leq \pi \leq \pi^U \} \)

\( \Theta = \{ \theta \in \mathbb{R}^n : 0 \leq \theta \leq t \} \)

\( \Gamma = \{ \gamma \in \mathbb{R}^n : 0 \leq \gamma \leq g \} \)

\( \mathcal{G}_i = \{ P(X) : EX \leq m_i \} \quad (1 \leq i \leq n) , \)

i.e. the litter size of a parent of age \( i \ (1 \leq i \leq n) \) is allowed to have any distribution (concentrated on the non-negative integers) but with mean at most \( m_i \). Here \( \mathbb{R}^n \) denotes the space of all real \( n \)-vectors.

We define the quantities \( \{ \hat{\pi}_i \}, \{ \hat{a}_i \}, \{ \hat{A}_i \}, \ (1 \leq i \leq n) \), recursively by:

\[
\hat{\pi}_n = \pi^L_n \\
\hat{a}_n = 0 \\
\hat{A}_n = (1 - \pi^L_n) r_n ;
\]

(6.12)

\[
\hat{\pi}_k = \begin{cases} 
\pi^U_k & \text{if } t_k \hat{A}_k + 1 > g_k \hat{A}_k + 1 + m_k \\
\pi^L_k & \text{otherwise ,}
\end{cases}
\]

\[
\hat{a}_k = \hat{\pi}_k t_k + (1 - \hat{\pi}_k) g_k ,
\]

\[
\hat{A}_k = \hat{a}_k \hat{A}_k + 1 + (1 - \hat{\pi}_k) m_k ,
\]

successively for \( k = n-1, n-2, \ldots, 1 \).
It will turn out that if \( \hat{\pi}_k = \pi_k^L \) then the optimal strategy for the individual is to make a reproductive effort at age \( k \) and that \( \hat{A}_k \) is the expected number of future offspring following an optimal strategy given that the organism has already reached age \( k \) \((1 \leq k \leq n)\).

Solving (6.12) explicitly we obtain:

\[
\hat{A}_k = (1-\hat{\pi}_k)m_k + \hat{a}_k(1-\hat{\pi}_{k+1})m_{k+1} + \ldots + \hat{a}_{k+1}\ldots\hat{a}_{n-1}(1-\hat{\pi}_n)m_n ,
\]

\((1 \leq k \leq n)\).

**Lemma 6.1**

Suppose \( \pi_i^L \leq \pi_i \leq \pi_i^U, \ 0 \leq \theta_i \leq t_i, \ 0 \leq \gamma_i \leq g_i, \ \mu_i \leq m_i \) \((1 \leq i \leq n)\), then:

\[
\hat{A}_i \geq \mu_i(1-\pi_i) + a_i\hat{A}_{i+1} . \quad 1 \leq i \leq n-1
\]

**Proof**

We have:

\[
\mu_i(1-\pi_i) + a_i\hat{A}_{i+1} = \hat{A}_{i+1}\pi_i + (1-\pi_i)[\mu_i + \gamma_i\hat{A}_{i+1}]
\]

\[
\leq \pi_i t_i\hat{A}_{i+1} + (1-\pi_i)[m_i + g_i\hat{A}_{i+1}]
\]

\[
\leq \pi_i t_i\hat{A}_{i+1} + (1-\pi_i)[m_i + g_i\hat{A}_{i+1}]
\]

(by definition of \( \hat{\pi}_i \))

\[
= \hat{A}_i \quad (1 \leq i \leq n-1)
\]

which proves the lemma. \(\square\)
Theorem 6.3

For all $\omega$-sequences $Z_0, Z_1, \ldots$ starting at $z = (z_1, z_2, \ldots, z_n)$ we have:

$$E[\sum_{N=0}^{\infty} Z_{1N}] \leq \begin{cases} (z_1 + \sum_{i=2}^{n} \hat{A}_i z_1)/(1-\hat{A}_1) & \text{if } \hat{A}_1 < 1, \\ \infty & \text{otherwise}. \end{cases}$$

(6.13)

Proof

Trivially $E[\sum_{N=0}^{\infty} Z_{1N}] < \infty$. Suppose $\hat{A}_1 < 1$, then we apply Theorem 2.3, with $C$ the set of non-negative integer $n$-vectors, $T = \min[N: Z_N = 0]$ if $Z_N = 0$ for some $N$ and $T = \infty$ otherwise, $f(z) = (z_1 + \sum_{i=2}^{n} \hat{A}_i z_1)/(1-\hat{A}_1)$, and $r(z) = z_1$ for $z \in C$. Then $r$ and $f$ are non-negative and for $P(z) \in M(z)$:

$$r(z) + Ef(z) = z_1 + E\{[Z_{11} + \sum_{i=2}^{n} \hat{A}_i Z_{11}]/(1-\hat{A}_1)\}$$

$$= z_1 + \frac{1}{1-\hat{A}_1} \{z_1 [\mu_1 (1-\pi_1) + a_1 \hat{A}_2] + z_2 [\mu_2 (1-\pi_2) + a_2 \hat{A}_3]$$

$$\quad + \ldots + z_n [\mu_n (1-\pi_n)]\}$$

$$\leq z_1 + \frac{1}{1-\hat{A}_1} \{z_1 \hat{A}_1 + \sum_{i=2}^{n} \hat{A}_i z_1\}$$

$$= [z_1 + \sum_{i=2}^{n} \hat{A}_i z_1]/(1-\hat{A}_1)$$

$$= f(z),$$

which verifies the conditions of Theorem 2.3.
In the derivation of the inequality we have used Lemma 6.1 and the fact that:

\[ E[Z_{i1}|Z_0 = z] = \sum_{i=1}^{n} z_i\mu_i(1-\pi_i), \]
\[ E[Z_{i1}|Z_0 = z] = a_{i-1}z_{i-1} \quad (2 \leq i \leq n). \]

The proof now follows by Theorem 2.3 by noting that:

\[ E[\sum_{k=0}^{T-1} r(Z_k) + f(Z_T)I_{T<\infty}] = E[\sum_{k=0}^{T-1} Z_{1k} + 0] = E[\sum_{k=0}^{\infty} Z_{1k}] . \]

\[ \square \]

**Corollary 1** (Attainment of Bounds)

From the equations (6.10), we see that the bounds are attained by the L. H. process with parameters \((\hat{\pi}, t, g, G')\), where \(G' = (G'_1, G'_2, \ldots, G'_n)\), and \(G'\) is any distribution concentrated on the non-negative integers with mean \(m_i\), \((1 \leq i \leq n)\).

In words, the optimal strategy is always to pick the largest survival probabilities, namely \(\{t_i\}\), \(\{g_i\}\), and to choose litter size distributions with the largest means \(\{m_i\}\). However, at age \(k\) \((1 \leq k \leq n)\) it is optimal for the organism to make a reproductive effort \((\pi_k = \frac{L}{\pi_k})\) only if the conditional mean number of future offspring (given that he has reached age \(k\)) is greater than that if reproduction were not attempted that period; assuming, in either case, that the organism will follow an optimal strategy for the rest of its life.

**Corollary 2** (The Total Fecundity)

By definition, \(A_1\) represents the total fecundity \((R)\) or the expected total number of offspring produced by an individual during its lifetime.
Using Lemma 6.1, we can show inductively that \( A_k \leq \hat{A}_k \) \( (k = n, n-1, \ldots, 1) \) for \( \pi^L < \pi < \pi^U \), \( 0 < \theta < t \), \( 0 < \gamma < g \), and \( 0 < \nu < m \). In particular, \( A_1 \leq \hat{A}_1 \) and this maximum total fecundity is achieved by the L. H. process defined in Corollary 1. Hence we have the result that the total fecundity is maximized by the same life history strategy that maximizes the expected total population size.

**Remark 1**

The result (3.4) of Theorem 3.1 is a special case of Theorem 6.3 with \( n = 1 \) and \( \pi^L_1 = 0 \).

**Example 1**

If \( t = g \) then \( \pi = \pi^L \) and the optimal strategy is to try to reproduce at every stage.

**Example 2** (Semelparous organisms with certain survival)

Suppose each organism can have at most one litter. We will not affect \( \sum_{k=0}^{\infty} Z_{1k} \) if we suppose \( g = 0 \). Suppose also that at age \( k \) \( (1 \leq k \leq n) \), if it is still alive, an individual has complete freedom to attempt to breed or not. If it does not, it survives to the next period with certainty, unless it is of age \( n \) in which case it dies. Thus \( \pi^L = 0 \), \( \pi^U = 1 \), and \( t = (1, 1, \ldots, 1, 0) \). In this case:

\[
\hat{\pi}_i = \begin{cases} 
0 & \text{if } m_i = \max[m_1, m_{i+1}, \ldots, m_n] \\
1 & \text{otherwise}
\end{cases}
\]

and

\[
\hat{A}_i = \max[m_1, m_{i+1}, \ldots, m_n].
\]
In words, the optimal strategy is to reproduce at that age in the future when the largest expected litter size occurs. For an individual of age $L$, the optimal policy is for it to reproduce at that age $L$ ($1 \leq L \leq n$) satisfying: $m_L = \max\{m_1, m_2, \ldots, m_n\}$.

Remark 2: **Lower Bounds**

In a similar manner lower bounds can be derived for the mean total population size (or, equivalently, the expected total number of births). It is not planned to do so here, although the strategies which achieve these bounds may be of interest for the purpose of population control. The results would form a multitype analogue of Theorem 3.2.

### 6.5 An Illustration: E-selection versus S-selection

In this section we will present two examples which will illustrate how the different measures of fitness considered in the last two sections can lead to different optimum life history strategies.

#### 6.5.1 A Deterministic Model

We consider a special case in which each individual may reproduce at most once (and hence we may take $\gamma = 0$) and there is a constant survivorship probability ($\theta_i = \theta$ for $1 \leq i \leq n-1$, $\theta_n = 0$). We suppose also that the organism may reproduce or not with complete freedom at every age ($\pi^L = 0$, $\pi^U = 1$), and that if it reproduces at age $i$ ($1 \leq i \leq n$), it bears $m_i$ offspring with probability one. The $\{m_i\}$ are all non-negative integers. Thus $G$ consists of exactly one element.

**E-selection**

We take $\phi^*_i(a) = a^{m_i}$ and therefore, by Theorem 6.2, the eventual extinction probabilities $\{p_i: 1 \leq i \leq n\}$ for the optimum $\mathcal{M}$-sequence (E-selection)
are given by:

\[ \rho_n = \phi_n^*(\rho) \]

\[ \rho_i = \min[1 - \theta + \theta \rho_{i+1}, \phi_i^*(\rho)], \quad (1 \leq i \leq n-1). \]

Hence \( \rho_1 \) is the smaller root in \([0,1] \) of:

\[ \rho_1 = \min[\rho_1^m, 1 - \theta + \theta \rho_2^m, \ldots, 1 - \theta^{n-1} + \theta^{n-1} \rho_1^m], \]

and the stationary optimal policy (E-selection) is for the organism to reproduce at that age \( L \) satisfying:

\[ 1 - \theta^{L-1} + \theta^{L-1} \rho_1^m \leq 1 - \theta^{i-1} + \theta^{i-1} \rho_1^m, \]

for all \( i \) \( (1 \leq i \leq n). \)

Of course if \( m_1 \neq 0 \), then trivially \( \rho_1 = 0 \) and the optimal policy is to reproduce at age 1.

S-selection

For this example, the \( \{\hat{A}_i \} \) are given by:

\[ \hat{A}_n = m_n \]

\[ \hat{A}_i = \max[m_i, \theta \hat{A}_{i+1}], \]

\[ = \max[m_i, \theta m_{i+1}, \ldots, \theta^{n-i} m_n], \]

for \( i = n-1, n-2, \ldots, 1. \)
Hence \( \hat{A}_1 = \max[m_1, \theta m_2, \ldots, \theta^{n-1} m_1] \) and the stationary optimal policy (S-selection) is for the organism to reproduce at that age \( L \) satisfying:

\[
\theta^{L-1} m_L > \theta^{i-1} m_i \quad \text{for all } i \ (1 \leq i \leq n).
\]

### 6.5.2 A Simple Stochastic Model

We consider the same model as in §6.5.1, except that now we suppose that the litter size distribution, \( G_i \), of a parent of age \( i \ (1 \leq i \leq n) \) is given by:

\[
p_i(x) = \begin{cases} 
1/2 & \text{if } x = 0, \\
1/2 & \text{if } x = i, \\
0 & \text{otherwise,}
\end{cases}
\]

i.e. litter sizes are either 0 or \( i \) with equal probability. Again \( \alpha \) consists of this one element only.

**E-selection**

We take \( \phi_1^*(\alpha) = (1+\alpha^1)/2 \) and therefore, by Theorem 6.2, the eventual extinction probabilities for the optimum \( \nu \)-sequence (E-selection) are given by:

\[
\rho_n = \phi_n^*(\rho_1),
\]

\[
\rho_i = \min[1-\theta + \theta \rho_{i+1}, \phi_i^*(\rho_1)], \quad (1 \leq i \leq n-1).
\]

Hence \( \rho_1 \) is the smallest root in \([0,1]\) of:
\[ \rho_1 = \min [\phi_1^*(\rho_1), 1 - \theta + \theta\phi_2^*(\rho_1), \ldots, 1 - \theta^{n-1} + \theta^{n-1}\phi_n^*(\rho_1)] \]

\[ = \min_{1 \leq i \leq n} [1 - \theta^{i-1} + \theta^{i-1}(1 + \rho_1^i)/2] . \]

The optimal stationary policy (E-selection) is:

\[ \hat{\pi}_i = \begin{cases} 
1 & \text{if } 1 - \theta + \theta\rho_{i+1} < (1 + \rho_1^i)/2 \\
\text{and } 1 \leq i \leq n - 1 , \\
0 & \text{otherwise} .
\end{cases} \]

The optimum genotype (E-selection) will reproduce at that age \( L \) satisfying:

\[ 1 - \theta^{L-1} + \theta^{L-1}(1 + \rho_1^L)/2 \leq 1 - \theta^{i-1} + \theta^{i-1}(1 + \rho_1^i)/2 , \]

for all \( i \) \((1 \leq i \leq n)\).

**S-selection**

Theorem 6.3 is concerned only with expected values and we may apply the results for S-selection of the previous example (in §6.5.1) with \( m_i = i/2 \) \((1 \leq i \leq n)\). Thus

\[ \hat{A}_1 = \frac{1}{2} \cdot \max[1, 2\theta, 3\theta^2, \ldots, n\theta^{n-1}] , \]

and the stationary optimal strategy (S-selection) is for the organism to reproduce at that age \( L \) satisfying:
\[ L \Theta^{L-1} \geq i \Theta^{i-1} \quad \text{for all} \quad i \quad (1 \leq i \leq n). \]

By differentiating the function \( f(x) = x \Theta^{x-1} \), we obtain \( L \propto |\log \Theta|^{-1} \). Of course \( L \) must be rounded off to be an integer in \([1,n]\).

6.6 Concluding Remarks

In this chapter, the results have been used to show how nature might force selection of genotypes with certain "optimum" life history parameters from among a given class of genotypes (characterized by \( \mathcal{N} \)). Alternatively, these results can be viewed as representing bounds on probability of extinction, mean time to extinction and expected total population size of a colonizing species with fixed life history parameters, but when these parameters are known only to lie within a certain class \( \mathcal{N} \). It will be this latter interpretation of \( \mathcal{N} \)-sequences that will be useful for the topic which is to be discussed in the next chapter.
7.1 A Description of the Model

We now return to the dam model discussed in §1.1.

For \( n = 0, 1, 2, \ldots \), we let \( I_n \) be the amount of water which flows into the dam in period \( n \) and \( O_n \) the demand or desired outflow in that period. We define \( X_n = I_n - O_n \) to be the net input to the dam in period \( n \) and assume \( P(X_n | I_1, \ldots, I_{n-1}; O_1, \ldots, O_{n-1}) \in \mathcal{M} \), where \( \mathcal{M} \) is some set of probability distributions on the real line. Then the amount of water, \( Z_n \), stored
at time $n$ is given by:

$$Z_0 = z$$

(7.1) $$Z_n = \min\left((Z_{n-1} + X_n)^+, b\right) \text{ for } n = 1, 2, \ldots$$

where $b$ is the capacity of the dam, and $z$ its initial content ($0 \leq z \leq b$). Thus $Z_0, Z_1, Z_2, \ldots$ is an $\mathcal{M}$-sequence starting at $z$ with:

(7.2) $$\mathcal{M}(z) = \{P[\min((z+X)^+, b)]: P(X) \in \mathcal{M}\}.$$ 

A typical sample path is shown in Figure 7.1.

For $0 \leq a \leq b$, we define the stopping time, $T_a$, by:

(7.3) $$T_a = \begin{cases} \min\{n: Z_n \leq a\}, & \text{if } Z_n \leq a \text{ for some } n, \\ \infty, & \text{otherwise}. \end{cases}$$

$T_a$ represents the time taken for the water volume in the dam to first reach the given critical level, $a$. Of course, if $a \geq z$ then trivially we have $T_a = 0$. Usually we will be interested in the special case $a = 0$, for then $T_0$ represents the time to first emptiness; however the case $a > 0$ may be of interest when there are certain water levels that are critical, for navigation or recreation, for example.

Moran (1954) studied this model under the assumptions that the inputs were independent and identically distributed and with a known distribution. This is a special case of our model when $\mathcal{M}$ contains only one element.
However, for reasons cited in §1.1, these assumptions are not very realistic and we suggest that our model is more useful as we assume only that the conditional distributions of the net inputs, given the past history of the process, belong to some given class \( \mathcal{M} \). In the following sections, for various selected \( \mathcal{M} \), we will give bounds on \( E[T_0] \), the mean time to first emptiness; on \( E[T_a] \), the mean time to first reach a critical level \( a \); and on the probability of overflow before emptiness.

7.2 An Exponential House

In this section we take \( m \) and \( \lambda \) to be positive real numbers and consider the dam process as defined in (7.1), (7.2), for which the distribution of the net input \( X_n \) (\( n = 1, 2, \ldots \)), conditional on the past history, \( \{I_1, I_2, \ldots, I_{n-1}, 0_1, 0_2, \ldots, 0_{n-1}\} \), has mean no greater than \( m \) and such that the expected value of \( e^{-\lambda X_n} \) is no greater than one. This last condition implies that also the mean of \( X_n \) must be positive (apart from the trivial case \( X_n = 0 \) with probability one, in which case the mean could be zero). In fact, for any input distributions with positive means and their Laplace transforms existing, we can find positive \( m, \lambda \) such that the above conditions hold. Thus many standard probability distributions (e.g. normal, uniform, double exponential etc.) fall into this class, provided that their means are positive.

The title of this section comes from the theory of gambling in a problem where the restrictions placed on the set of permitted gambles are similar to the restrictions we place on the allowable input distributions (see Dubins and Savage 1965, Chapter 8.7).
7.2.1 The Basic Theorem

Theorem 7.1 (Lower Bounds on the Mean Time to First Emptiness)

Let \( m, \lambda \) be positive real numbers and set \( M = \{ P(X): E X \leq m, E e^{-\lambda X} \leq 1 \} \). Then for \( Z_0, Z_1, \ldots \), an \( M \)-sequence starting at \( z \) (\( 0 \leq z \leq b \)) as defined in (7.1), (7.2), we have:

\[
(7.4) \quad E[T_0] \geq -z/m + e^{\lambda b} [1 - e^{-\lambda z}] / m \lambda.
\]

Proof

We apply Theorem 2.4 with \( C \) the real interval \([0, b]\); \( r(z) = 1 \), \( f(z) = -z/m + e^{\lambda b} [1 - e^{-\lambda z}] / m \lambda \) for \( 0 \leq z \leq b \); and \( T = T_0 \), the time to first emptiness.

The random variables \( \{ Z_n: n = 0, 1, 2, \ldots \} \) are all constrained to take values in \([0, b]\) and, since \( f \) and \( r \) are bounded on \([0, b]\), we see that \( E[\sum_{k=0}^{\infty} r(Z_k) + f(Z_k)] < \infty \) for all \( k \geq 0 \). Also we have that either \( E[T] = \infty \) (and trivially (7.4) holds) or else \( \lim_{N \to \infty} \sup E[f(Z_N) I_{T \geq N}] = 0 \). Thus we may suppose that the latter is true.

Now \( f(0) = 0 \), and therefore \( E[\sum_{k=0}^{T-1} r(Z_k) + f(Z_T) \cdot I_{T < \infty}] = E[T_0] \). Thus, since \( r \) and \( f \) are non-negative on \([0, b]\), it only remains to show that: \( 1 + E f(Z) \geq f(z) \) whenever \( P(Z) \in M(z) \).

Extend \( f(z) \) by defining \( f(z) = f(b) \) for \( z \geq b \), and \( f(z) = 0 \) for \( z \leq 0 \). Thus \( f(Z) = f(z + X) \) for \( P(Z) \in M(z) \). Also define \( g(z) = -z/m + e^{\lambda b} [1 - e^{-\lambda z}] / m \lambda \) for all real \( z \).
Figure 7.2

Then \( f(z) \geq g(z) \) for all real \( z \) and \( f(z) = g(z) \) for \( 0 \leq z \leq b \) (see Figure 7.2). Hence, for \( P(Z) \in \mathcal{M}(z) \), we have:

\[
1 + Ef(Z) = 1 + Ef(z+X) ,
\]

\[
\geq 1 + Eg(z+X) ,
\]

\[
= 1 - E(z+X)/m + e^{\lambda b} [1-e^{-\lambda z}E(e^{-\lambda X})]/m\lambda ,
\]

\[
\geq -z/m + e^{\lambda b} [1-e^{-\lambda z}]/m\lambda ,
\]

\[
= f(z) ,
\]

which completes the proof.
Corollary 1 (The Time to Reach a Given Critical Level)

For the setup of Theorem 7.1, we have, for all $\mathcal{M}$-sequences starting at $z$ ($a \leq z \leq b$):

\[(7.5) \quad E[T_a] \geq -(z-a)/m + e^{\lambda b} \frac{1-e^{-\lambda(z-a)}}{m \lambda},\]

where $T_a$ is the time to first reach the level, $a$, as defined in (7.3). (Of course, if $z < a$ then $E[T_a] = 0$.)

**Proof**

Essentially the process is stopped at time $T_a$, and hence it does not matter how we define the process to continue after time $T_a$. Thus we may suppose that the state $z = a$ is absorbing without affecting $E[T_a]$. Now the result (7.5) follows by applying Theorem 7.1 with a simple change of origin, i.e. $z$ is replaced by $z - a$.

Corollary 2 (Attainment of Bounds)

Suppose $\sigma^2 = 2m/\lambda$, then for $X$ normally distributed with mean $m$ and variance $\sigma^2$ we have $EX = m$ and $E[e^{-\lambda X}] = 1$. However the $\mathcal{M}$-sequence for which the $\{X_n\}$ are i.i.d. with this distribution, $N(m, \sigma^2)$, does not achieve the bounds (7.4), (7.5). This is again due to the problem of "overshoot" which was encountered also in Theorems 3.10, 4.8.

The bounds (7.4), (7.5) are attained only in a limiting sense when the time scale becomes continuous and $Z(t)$ is a Brownian motion process with $Z(0) = z$ and with drift $m$ and infinitesimal variance $\sigma^2$. For this process, in time $(t, t+\Delta t)$, there is a net input of $\Delta X$, where $E[\Delta X] = m \cdot \Delta t + o(\Delta t)$ and $\text{Var}[\Delta X] = \sigma^2 \cdot \Delta t + o(\Delta t)$. The proof is given in Cox and Miller (1965, page 233, Example 5.6). In this case there is no problem of "oversooting
the boundary." Note that the same process minimizes $E[T_a]$ for all choices of $z$ and $a$.

7.2.2 The Special Case of Normal Inputs; a Release Rule

Suppose $X$ is a normally distributed random variable with mean $\mu$ and variance $\sigma^2$. Then $E[e^{-\lambda X}] = \exp[-\mu \lambda + \lambda^2 \sigma^2 / 2]$, and the condition $E[e^{-\lambda X}] \leq 1$ implies $\lambda \sigma^2 \leq 2\mu$. We also have the restriction $0 < \mu \leq m$. Thus if the point $(\mu, \sigma^2)$ lies in the shaded area shown in Figure 7.3, the distribution of $X$ belongs to the set $\mathcal{M}_\nu$, as defined in Theorem 7.1.

![Figure 7.3](image)

Now suppose that the net input $X_n$ (n = 1,2,...) is normally distributed with mean $\mu_n$ and variance $\sigma_n^2$, conditional on the past history,
\( \{I_1, \ldots, I_{n-1}; 0_1, \ldots, 0_{n-1}\} \), where \( \mu_n \) and \( \sigma_n^2 \) may be unknown. However suppose that we do know that \( \mu_n \leq m \) and \( \sigma_n^2/2\mu_n \leq 1/\lambda \), for all \( n \geq 1 \), where \( m \) and \( \lambda \) are known positive constants. Thus we may apply Theorem 7.1 to obtain lower bounds on the mean time to first emptiness or to reach any given critical level, \( a \). A class of release rules will now be proposed and these results will be used to evaluate their performance with respect to the lower bounds on \( E[T_0] \) or \( E[T_a] \) for any given \( a \) (\( 0 \leq a \leq b \)).

We suppose that the total inflow \( I_n \) (\( n = 1, 2, \ldots \)) in period \( n \) is normally distributed with mean \( \mu_n \) and variance \( \sigma_n^2 \), where the \( \{I_n\} \) are independent. We consider the release rule:

\[
(7.6) \quad q_n = \alpha_n \cdot I_n + \beta_n(Z_{n-1}),
\]

where, for each \( n = 1, 2, \ldots \), we have that \( \alpha_n \) is some real number and \( \beta_n \) is some real-valued function defined on \([0, b]\). Thus the outputs \( \{q_n\} \) and net inputs \( \{X_n\} \) are no longer independent. In fact, the distribution of \( X_n \) (\( n = 1, 2, \ldots \)) conditional on the past history of the process, is normal with mean \( (1-\alpha_n)\mu_n - \beta_n(Z_{n-1}) \) and variance \( (1-\alpha_n)^2\sigma_n^2 \). Thus if we set:

\[
m = \sup_{0 \leq z \leq b} \sup_{n} \left[(1-\alpha_n)\mu_n + \beta_n(z)\right]
\]

and

\[
\frac{2}{\lambda} = \sup_{0 \leq z \leq b} \sup_{n} \left[ \frac{(1-\alpha_n)^2\sigma_n^2}{(1-\alpha_n)\mu_n - \beta_n(z)} \right],
\]
then, provided \( m \) and \( \lambda \) are positive, a lower bound on the mean time to first emptiness is given by Theorem 7.1, and Corollary 1 provides a lower bound on the mean time to first reach any given critical level, \( a (0 \leq a \leq b) \). A water resources planner could compare these lower bounds for various numbers \( \{\alpha_n\} \) and functions \( \{\beta_n\} \) and use the information to help decide which release rule should be employed. The choice of these \( \{\alpha_n\}, \{\beta_n\} \) is an interesting open question.

As an example, take \( \alpha_n = \alpha \) and \( \beta_n(z) = \beta \cdot z \) for all \( n \), where \( 0 \leq \alpha, \beta \leq 1 \). Then:

\[
m = \sup_{0 < z < b} \{ (1-\alpha)\mu - \beta z \},
\]

\[
= (1-\alpha)\mu;
\]

and

\[
\frac{2}{\lambda} = \frac{(1-\alpha)^2 \sigma^2}{(1-\alpha)\mu - \beta b}.
\]

Suppose \( m \) and \( \lambda \) are positive, then, by Theorem 7.1, we have:

\[
E[T_0 | Z_0 = z] \geq z/\mu(1-\alpha)
\]

\[
+ \frac{(1-\alpha)^2}{2\mu[(1-\alpha)\mu - \beta b]} \cdot (1 - \exp(-\lambda z)) \cdot \exp\left\{ \frac{2[(1-\alpha)\mu - \beta b]b}{(1-\alpha)^2 \sigma^2} \right\}.
\]

For \( z = b \), this becomes:
\[ E[T_0|Z_0=b] = -\frac{b}{(1-\alpha)\mu} + \frac{\exp(2b[(1-\alpha)\mu - \beta\mu]/(1-\alpha)2\sigma^2) - 1}{2\mu[(1-\alpha)\mu - \beta\mu]/(1-\alpha)\sigma^2} \].

7.2.3 The Probability of Emptiness Before Overflow

We now remove the restriction on the means of the net input distributions (i.e. we let \( m \to \infty \)).

**Theorem 7.2**

Take \( \lambda \geq 0 \) and set \( \mathcal{M} = \{ P(X): E[e^{-\lambda X}] \leq 1 \} \). Then for \( Z_0, Z_1, \ldots \), an \( \mathcal{M} \)-sequence starting at \( z \) \( (0 \leq z \leq b) \) as defined in (7.1), (7.2), we have:

\[ \text{Pr}[\text{Dam empties before overflowing}] \leq e^{-\lambda z}. \quad (7.7) \]

(Note: If the dam never empties or overflows then we shall say that "emptiness before overflow" has not occurred.)

**Proof**

We apply Theorem 2.3 with \( C = [0,b] \); \( r(z) \equiv 0 \); \( f(z) = e^{-\lambda z} \) \((0 \leq z < b)\), \( f(b) = 0 \); and

\[ T = \begin{cases} \infty & \text{if } 0 < Z_n < b \text{ for all } n, \\ \min[n: Z_n < 0 \text{ or } Z_n > b] & \text{otherwise}. \end{cases} \]

We extend \( f \) by defining \( f(z) = 1 \) for \( z < 0 \) and \( f(z) = 0 \) for \( z > b \). Thus \( f(Z) = f(z+X) \) for \( P(Z) \in \mathcal{M}(z) \). Also define \( g(z) = e^{-\lambda z} \) for
all real $z$, and $v(z) = 1$ for $z \leq 0$, $v(z) = 0$ for $z > 0$. Thus $g(z) \geq f(z) \geq v(z)$ for all real $z$ (see Figure 7.4).

![Figure 7.4](image)

Certainly $r$ and $f$ are non-negative and for $P(Z) \in \mathbb{M}(z)$ we have (for $0 \leq z < b$):

$$ r(z) + Ef(Z) = Ef(z+X) $$

$$ \leq Eg(z+X) $$

$$ = E[e^{-\lambda(z+X)}] $$

$$ \leq e^{-\lambda z} $$

$$ = f(z) $$
We note that, as in Corollary 1 of Theorem 7.1, it does not matter how we define the process to continue after the stopping time $T$. Hence we may assume that the state $z = b$ is absorbing without affecting the probability of emptiness before overflow. Thus if $z = b$, for $P(Z) \in \mathcal{M}(z)$, we have:

$$r(z) + Ef(Z) = f(b) = f(z).$$

Hence $r(z) + Ef(Z) \leq f(z)$ for $P(Z) \in \mathcal{M}(z)$ and for all $0 \leq z \leq b$. Therefore we may apply Theorem 2.3 to obtain:

$$f(z) \geq E[f(Z_T) \cdot I_{T<\infty}]$$

$$\geq E[v(Z_T) \cdot I_{T<\infty}]$$

$$= \Pr[Z_T = 0 \text{ and } T < \infty]$$

$$= \Pr[\text{Emptiness before overflow}],$$

which completes the proof of Theorem 7.2.

Corollary 1 (Attainment of Bounds)

The bound (7.7) is attained by the $\mathcal{M}$-sequence $Z_0, Z_1$, where the distribution of $X_1$, given $Z_0 = z$, is improper and such that:

$$X_1 = \begin{cases} -z & \text{with probability } e^{-\lambda z}, \\ \infty & \text{with probability } 1 - e^{-\lambda z}. \end{cases}$$

This $\mathcal{M}$-sequence either overflows or empties the dam in one step. Thus the optimal strategy is a "one move, two point policy" as in Dubins and Savage.
(1965, Chapter 7.4, Example 1); however in their gambling model, they do not allow the possibility of "overshooting" the boundaries 0 and b, as is possible in the dam model.

**Corollary 2** *(The Probability of Reaching a Given Critical Level Before Overflow)*

For the setup of Theorem 7.2, and for $0 \leq a \leq z \leq b$, we have:

$$\Pr[\text{Water level reaches } a \text{ before it overflows}] \leq e^{-\lambda(z-a)}.$$ (As before, if $a < Z_n < b$ for all $n$, then we say that the water level has not reached $a$ before overflowing.)

This bound is sharp and is attained by the $\mathbb{M}$-sequence $Z_0, Z_1$, where the distribution of $X_1$, given $Z_0 = z$, is improper and such that:

$$X_1 = \begin{cases} 
-(z-a) & \text{with probability } e^{-\lambda(z-a)} \\
\infty & \text{with probability } 1-e^{-\lambda(z-a)} 
\end{cases}.$$ 

This is the counterpart of Corollary 1 of Theorem 7.1 and the proof follows as for that of Theorem 7.2 but with a simple change of origin.

**Remark 1**

For the Brownian motion process (defined in Corollary 2 of Theorem 7.1) with drift $m$ and infinitesimal variance $\sigma^2 = 2m/\lambda$ we have (see Cox and Miller 1965, Example 5.5, page 233):

$$\Pr[\text{Dam empties before overflowing} | Z(0) = z] = \frac{e^{-\lambda z} - e^{-\lambda b}}{1 - e^{-\lambda b}},$$
which is less than $e^{-\lambda z}$ ($0 \leq z \leq b$). Hence the continuous time process which attains the bound in Theorem 7.1 (minimizes the mean time to first emptiness) is strictly suboptimal for the purpose of maximizing the probability of emptiness before overflow—a result which may appear surprising at first glance.

**Remark 2**

We now calculate the mean time to first emptiness for the strategy of Corollary 1 of Theorem 7.2 which maximizes the probability of emptiness before overflow. We have:

$$E[T_0] = 1 \cdot e^{-\lambda z} + 2 \cdot (1 - e^{-\lambda z}) e^{-\lambda b} + 3 \cdot (1 - e^{-\lambda z})(1 - e^{-\lambda b}) e^{-\lambda b} + \ldots$$

$$= e^{-\lambda z} + (1 - e^{-\lambda z}) e^{-\lambda b} \
\sum_{i=0}^{\infty} (i+2)(1 - e^{-\lambda b})^i$$

$$= e^{-\lambda z} + e^{\lambda b}(1 - e^{-\lambda z})(1 + e^{-\lambda b}) .$$

However, letting $m \to \infty$ in (7.4), we obtain $E[T_0] \geq 0$ for all $\lambda$-sequences starting at $z$ ($0 \leq z \leq b$), for which $M = \{P(X): E[e^{-\lambda X}] \leq 1\}$. Hence the two point policy which is optimal for the goal of Theorem 7.2 (maximizing the probability of emptiness before overflow) is strictly suboptimal for the goal of Theorem 7.1 (minimizing the mean time to first emptiness) when the mean net input is unrestricted ($m \to \infty$).

**7.3 Bounded Inputs**

In this section we take $\tau$ to be some positive real number, and consider the dam process as defined in (7.1), (7.2), for which the distribution of the
net input, $X_n \ (n = 1, 2, \ldots)$, conditional on the past, $\{I_1, \ldots, I_{n-1}; O_1, \ldots, O_{n-1}\}$, takes values in $[-r, r]$ with probability one. By rescaling the units, we can take $r = 1$ without loss of generality.

**7.3.1 Upper Bounds on the Mean Time to First Emptiness**

**Theorem 7.3**

Take $m \ (-1 \leq m \leq 1)$, and set $M = \{P(X): E[X] \leq m, |X| \leq 1\}$. Then for $Z_0, Z_1, \ldots$, an $M$-sequence starting at $z \ (0 < z \leq b)$ as defined in (7.1), (7.2), we have:

$$E[T_0] \leq \begin{cases} \infty & \text{if } m \geq 0 \\ \frac{(1+z)/|m|}{m} & \text{if } m < 0 \end{cases}$$

(7.8)

**Proof**

Clearly $E[T_0] \leq \infty$. Suppose $m < 0$. Then we apply Theorem 2.3 with $C = [0, b]; r(z) \equiv 1, f(z) = (1+z)/|m|$ for $0 \leq z \leq b$; and $T = T_0$, the time to first emptiness.

Extend $f(z)$ by defining $f(z) = f(b)$ for $z \geq b$, and $f(z) = 0$ for $z \leq 0$. Thus $f(Z) = f(z+X)$ for $P(Z) \in M(z)$. Also define $g(z) = (1+z)/|m|$ for all real $z$. 
Then $f(z) \leq g(z)$ for all $z \geq -1$ and $f(z) = g(z)$ for $0 \leq z \leq b$ (see Figure 7.5). Hence, for $P(Z) \in W(z)$, we have:

$$1 + Ef(Z) = 1 + Ef(z+X),$$

$$\leq 1 + Eg(z+X),$$

(since $z + X \geq -1$ w.p. 1)

$$= 1 + E[(1 + z + X)/|m|],$$

$$\leq (1+z)/|m|,$$

$$= f(z).$$
Also $r$ and $f$ are non-negative on $[0,b]$ and thus all the conditions of Theorem 2.3 are satisfied. The proof of Theorem 7.3 now follows by noting that:

$$E[\sum_{k=0}^{T-1} r(Z_k) + f(Z_T) \cdot I_{T<\infty}] = E[T_0] + \Pr[T<\infty]/|m| \geq E[T_0].$$

Remark 1: Attainment of Bounds

If $z = 0$ then $E[T_0] = 0$. Suppose $0 < z < b$. Then if $m > 0$ and the $\{X_i\}$ are i.i.d. and equal to $m$ with probability one, we have $E[T_0] = \infty$ and the bound (7.8) is attained.

If $m < 0$ then the bound is not sharp. This is caused by the problem of "overshooting the boundary" (see also Theorems 3.10, 4.8 and 7.1). Rewritten in our terminology, Blackwell (1964; Example B) showed that if $b = \infty$ and

$$T = \begin{cases} 
\min[n: \text{ } Z_{n-1} + X_n < 0], & \text{if } Z_{n-1} + X_n < 0 \text{ for some } n, \\
\infty, & \text{otherwise},
\end{cases}$$

where the inequality is strict, then $E[T] < (1+z)/|m|$, and for integer $z$ this bound is sharp and is attained when the $\{X_n\}$ are i.i.d. with $X_n = \pm 1$ with probability $1 + m/2$.

Remark 2: The Mean Time to First Reach a Given Critical Level

By a change in origin, a similar result can be obtained for the mean time to first reach a given critical level, $a$ ($0 < a < z < b$). For $m < 0$ we now have:
\[ E[T_a] \leq \frac{(1 + z - a)}{|m|} . \]

(Cf. Corollary 1 of Theorem 7.1, and Corollary 2 of Theorem 7.2.)

### 7.3.2 Lower Bounds on the Mean Time to First Emptiness

The results of this section have not been proved completely and thus they will be stated as a conjecture. However our numerical calculations and certain other considerations all point to the validity of the results.

**Conjecture**

Take \( m \ (0 < m < 1) \) and \( b \) some positive integer, and set

\[ M = \{P(X): \ EX > m, \ |X| < 1\} \]. Suppose \( f(z) \) is defined by:

(7.9)

\[
\begin{align*}
  f(z) &= 0 & (z = 0) \\
  f(z) &= 1 + (m+z) \cdot f(1) & (0 < z \leq c_0) \\
  f(z) &= \frac{[1 + \frac{m+z}{1+z} \cdot f(1)]}{[1 - \frac{m+z}{1+z} \cdot \frac{f(2) - f(1)}{f(1)}]} & (c_0 < z \leq 1) \\
  f(i+z) &= 1 + \frac{1+m}{2-z} \cdot f(i+1) + \frac{1 - (m+z)}{2-z} \cdot f(i-1+z) & (i < i+z \leq i+c_1) \\
  f(i+z) &= f(i) + \frac{f(i+1) - f(i)}{f(1) - f(i-1)} \cdot [f(i-1+z) - f(i-1)] & (i+c_1 < i+z \leq i+1) 
\end{align*}
\]

, for \( i = 1, 2, \ldots, b-1 \);

where

\[
f(i) = -\frac{i}{m} + \frac{1+m}{2m^2} \left(\frac{1+m}{1-m}\right)^b \left[1 - \left(\frac{1}{1+m}\right)^i\right] & (i = 0, 1, 2, \ldots, b),
\]
and the \( \{c_i\} \) are defined such that \( f(z) \) is continuous at \( z = i + c_i \) (\( 0 \leq i \leq b-2 \)), with \( 1-m \geq c_0 \geq c_1 \geq \ldots \geq c_{b-1} = 0 \).

Then for \( Z_0, Z_1, \ldots \), an \( M \)-sequence starting at \( z \) (\( 0 \leq z \leq b \)) as defined in (7.1), (7.2), we have:

\[
(7.10) \quad E[T_0] > f(z).
\]

**Remarks**

The lower bound function \( f(z) \) is shown in Figure 7.6 for the case \( b = 3, m = 1/2 \).
The function is discontinuous at \( z = 0, 1, 2 \) and non-differentiable at
\[ z = c_0, \quad 1 + c_1, \quad 1 + c_0, \quad 2 + c_1, \quad 2 + c_0, \]
where \( c_0 = 3/13 = 0.2308 \) and \( c_1 = 0.1393 \). In general the function is discontinuous at \( 0, 1, 2, \ldots, b-1 \)
and non-differentiable at \( i + c_j \), for \( j = 0, 1, 2, \ldots, i \) and \( i = 0, 1, 2, \ldots, b-1 \).

Various intermediate results need to be shown. Firstly the \( \{c_i\} \) must
be uniquely defined with \( 1-m \geq c_0 \geq c_1 \geq \ldots \geq c_{b-1} = 0 \), or else \( f \) is not
uniquely defined. Theorem 2.4 (with \( C = [0, b] \), \( r(z) = 1 \), \( T = T_0 \), and \( f \)
as defined in (7.9)) would be used in the proof as in the proof of Theorem
7.1. The condition (2.8) reduces to: \( f(z) \leq l + Ef(Z) \) for \( P(Z) \in \mathcal{M}(z), \)
\( 0 \leq z \leq b \). This will be true if:

(a) for \( 0 < z < c_0 \), the chord joining \((0,0)\) to \((1,f(1))\) lies completely below the curve \( y = f(x) \) for \( 0 \leq x \leq z+1 \);

(b) for \( c_0 < z < 1 \), the chord joining \((0,0)\) to \((z+1, f(z+1))\) lies completely below the curve \( y = f(x) \) for \( 0 \leq x \leq z+1 \);

(c) for \( i < z < i+c_i \) \((i = 1, 2, \ldots, b-1)\), the chord joining \((z-1, f(z-1))\)
to \((i+1, f(i+1))\) lies completely below the curve \( y = f(x) \) for \( z-1 \leq x \leq z+1 \);

(d) for \( i+c_i < z < i+1 \) \((i = 1, 2, \ldots, b-1)\), the chord joining
\((z-1, f(z-1))\) to \((z+1, f(z+1))\) lies completely below the curve \( y = f(x) \)
for \( z-1 \leq x \leq z+1 \);

where we have defined \( f(z) = f(b) \) for \( z \geq b \).

**Attainment of the Bound**

The bound (7.10) is achieved by the \( \mathcal{M} \)-sequence defined as follows:

(a) for \( 0 < z \leq c_0 \), \( P(Z) \in \mathcal{M}(z) \) such that:

\[
Z = 0 \quad \text{(i.e. } X = -z) \quad \text{with probability } \frac{1-m}{z} ,
\]

\[
= 1 \quad \text{(i.e. } X = 1-z) \quad \text{with probability } \frac{m+z}{z} .
\]

(This distribution is well-defined since \( c_0 \leq 1-m \).)
(b) for \( c_0 < z \leq 1 \), \( P(Z) \in \mathcal{M}(z) \) such that:

\[
Z = 0 \quad \text{(i.e. } X = -z) \quad \text{with probability } \frac{(1-m)}{(z+1)},
\]
\[
= z+1 \quad \text{(i.e. } X = +1) \quad \text{with probability } \frac{(m+z)}{(z+1)};
\]

(c) for \( i < z \leq i+c_i \) \((i = 1, 2, \ldots, b-1)\), \( P(Z) \in \mathcal{M}(z) \) such that:

\[
Z = z-1 \quad \text{(i.e. } X = -1) \quad \text{with probability } \frac{(1-m-z+i)}{(2-z+i)},
\]
\[
= i+1 \quad \text{(i.e. } X = 1-z+i) \quad \text{with probability } \frac{(1+m)}{(2-z+i)};
\]

(d) for \( i+c_i < z \leq i+1 \) \((i = 1, 2, \ldots, b-1)\), \( P(Z) \in \mathcal{M}(z) \) such that:

\[
Z = z-1 \quad \text{(i.e. } X = -1) \quad \text{with probability } \frac{(1-m)}{2},
\]
\[
= z+1 \quad \text{(i.e. } X = +1) \quad \text{with probability } \frac{(1+m)}{2}.
\]

Using this strategy, the mean time to first emptiness starting at \( z \) is given by \( f(z) \), where \( f(z) \) satisfies the functional equations:

\[
f(z) = 1 + (m+z)f(1) \quad (0 < z \leq c_0),
\]

\[
f(z) = 1 + (m+z)f(1+z)/(1+z) \quad (c_0 < z \leq 1),
\]

\[
f(z) = 1 + [(1+m)f(i+1) + (1-m-z+i)f(z-1)]/(2-z+i)
\]
\[
\quad (i < z \leq i+c_i; \ i = 1, 2, \ldots, b-1),
\]

\[
f(z) = 1 + \frac{1+m}{2} f(z+1) + \frac{1-m}{2} f(z-1)
\]
\[
\quad (i+c_i < z \leq i+1; \ i = 1, 2, \ldots, b-1).
\]

These functional equations are satisfied by the function \( f \) as defined by (7.9) and hence the bound (7.10) is achieved.
Note that for \( z \) integer \((0 \leq z \leq b)\) we have:

\[
(7.11) \quad f(z) = \frac{z}{m} + \frac{1+m}{2m^2} \left( \frac{1+m}{1-m} \right)^b \left[ 1 - \left( \frac{1-m}{1+m} \right)^z \right],
\]

which is of a similar form to the bound (7.4) of Theorem 7.1. For integer \( z \), the bound (7.11) is attained when the net inputs \( \{X_i: i = 1, 2, \ldots\} \) are i.i.d. with common distribution given by \( X_i = \pm 1 \) with probability \((1+m)/2\). Hence the optimal \( \mathcal{M}_n \)-sequence starting at \( z \) always takes on integer values with probability one. (The expression (7.11) was first obtained by Hardin and Sweet (1970) as the mean time to absorption in the nonsymmetric random walk between a semireflecting and an absorbing barrier.)

7.4 An Application to the Theory of Collective Risk

The models studied in this chapter can also be applied to the theory of collective risk, in a case studied by Borch (1967). Here, \( Z_n \) is the total capital of an insurance company at time \( n \), \( I_n \) the premiums collected in the period \((n-1, n)\), \( Q_n \) the total amount in claims paid out in the period \((n-1, n)\), \( b \) the level of capital above which any excess revenue is paid out in dividends or taxes, and \( T_0 \) represents the time until the ruination of the company. For this application the \( \{I_n\} \) are usually considered the controllable variables instead of the \( \{Q_n\} \), as in the case of the dam model.
REFERENCES


Some results from the theory of gambling are extended and applied to some problems in applied probability.

A branching process is considered in which the litter size distributions are known only to lie in some set \( M \) of probability distributions. For various \( M \), Chebyshev-like bounds are derived for various quantities of interest such as the probability of extinction and the mean time to extinction. The conditions under which the bounds are achieved are examined. The results are generalized to multitype branching processes and applied to the problem in mathematical ecology of optimal life history strategies and modes of natural selection.

Finally, a dam model is considered in which the annual water inflows are not necessarily i.i.d. and the exact distributions of the inflows are known only to lie in some class \( M \). Sharp bounds are obtained on such quantities as the mean time to first emptiness. Some release rules are discussed.