*The Annals of Probability*
1973, Vol. 1, No. 3, 457–474

# INEQUALITIES FOR BRANCHING PROCESSES[1]

BY BRUCE W. TURNBULL[2]

*Cornell University*

A branching process is considered for which the conditional distributions of the litter sizes, given the past, are allowed to vary from period to period and are required only to belong to some set $\mathscr{M}$. The process is non-Markovian in general. For various interesting $\mathscr{M}$, bounds are derived on (i) the probability of extinction, (ii) the mean time to extinction, (iii) the probability that a generation size exceeds a given number, (iv) the expected maximum generation size, and (v) the mean total population size. In (i), (ii) and (v), the optimal strategies which achieve the bounds are identified.

The techniques used are similar to those used in the theory of gambling as developed by Dubins and Savage (*How to Gamble if You Must*, McGraw-Hill (1965)).

**1. Introduction.** We consider the following population growth model in which the successive generation sizes are given by:

$$Z_0 = z$$
$$Z_n = X(n, 1) + \cdots + X(n, Z_{n-1}), \qquad (n = 1, 2, \cdots);$$

where $z$ is a nonnegative integer representing the initial generation size and $X(n, j)$ is the random number of offspring of the $j$th individual of the $(n-1)$th generation. It is understood that if $Z_{n-1} = 0$ then $Z_n = 0$ and we say that the process has become extinct. We further assume that within each generation, conditional on the past, the individuals reproduce independently of each other and, for each $n$, the random litter sizes $\{X(n, i); i = 1, 2, \cdots, Z_{n-1}\}$ are independent and with distributions (not necessarily identical) in some set $\mathscr{M}$ of probability distributions. By "conditional on the past," we mean "conditional on the random variables $\{Z_i, X(i, j) : j = 1, 2, \cdots, Z_i\}, i = 1, 2, \cdots, n-1$."

In the classic model of Galton and Watson [8], the litter size distributions in all generations are i.i.d. This corresponds to our model for the case when $\mathscr{M}$ contains exactly one element. The recent theory of "branching processes in random environments" (see Smith [13], Smith and Wilkinson [14], and Athreya and Karlin [1]) 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 $\{\xi_i; i \geqq 0\}$

Received October 15, 1971; revised June 14, 1972.

[1] Part of author's doctoral dissertation submitted in the Department of Operations Research, Cornell University. The research was supported in part by the U. S. Army Research Office under Grant DA-31-124-ARO-D-474, and by the National Science Foundation under Grant GK-21460.

[2] Now at Mathematical Institute, University of Oxford.

*AMS* 1970 *subject classifications.* Primary 60J80; Secondary 60G40, 60G45.

*Key words and phrases.* Branching processes; Chebyshev-like inequalities, gambling theory, non-Markovian processes, martingales, stopping times, dynamic programming.

is a stationary random process and $\xi_n$ is independent of the past history of the process ($n = 0, 1, 2, \cdots$). 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 $\mathscr{M}$.

Freedman and Purves [7] treat this model for the special case when $\mathscr{M}$ consists of all distributions concentrated on $\{0, 2, 3, 4, \cdots\}$ and with mean less than 2. They then find sharp lower bounds on the probability of extinction. Goodman [9] extends these results and finds both upper and lower bounds on the probability of extinction when $\mathscr{M}$ consists of distributions concentrated on some arbitrary subset of the nonnegative integers and with mean restricted to be within a certain interval.

In Section 2 we will develop a general theory which will be applied in later sections to obtain Chebyshev-like bounds on not only the probability of extinction but also the mean time to extinction, the expected total population size, the probability that a generation size exceeds a certain number and the expected maximum generation size—all for various $\mathscr{M}$. It will be seen that the results of Freedman and Purves [7] and of Goodman [9] are part of a unified theory.

Goodman points out that the study of such processes are useful for describing populations when the demographic data are subject to error. Also the optimal strategies whereby the bounds are achieved are of interest in the study of natural selection. (For further discussion of this application see Turnbull [18] Chapter 6.)

**2. The basic theory.** The problem posed in Section 1 is similar to the one studied in the theory of optimal gambling, first considered by Blackwell [2], [3] and later extensively developed by Dubins and Savage [6]. This theory considers a gambler who, having a fortune $z_n$ at stage $n$, may play any gamble $\gamma$ selected from a specified set $\mathscr{M}(z_n)$ 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 our branching process model, $z_n$ represents the size of the $n$th generation, $C$ is the set of possible population sizes (the nonnegative integers), the set of gambles corresponds to the set of possible litter size distributions, and, typically, the objective is to minimize the probability of extinction. This suggests that techniques similar to those used in gambling theory can be applied to our problem. In fact, as will be shown later, these techniques can also be applied to a wide variety of interesting problems.

Dubins and Savage [6] treat at length many problems associated with optimal gambling systems. They proceed under the very general assumption of finitely additive gambles defined on all subsets of $C$. (They give details in Section 2.3.) We shall stay with the traditional assumption of countably additive measures defined on the Borel sets, although the bounds derived in Section 3, 4, and 5 may be shown to hold when $\mathscr{M}$ is extended to include finitely additive gambles. The measurable gambling setup has been treated in Blackwell [4], Strauch [15] and by Sudderth [16], [17].

For random variables $X$, $Y$ we use the notation $P(X)$ for the probability distribution of $X$, and $P(X \mid Y)$ for the conditional distribution of $X$ given $Y$. Also let $C$ be a Borel subset of some Polish space. In the applications considered in this paper, this space will be the space of reals.

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

For instance, if $\mathscr{M}$ contains only one element then $X_1, X_2, \cdots$ are independent and identically distributed. However, if $\mathscr{M}$ contains more than one element, in general the sequence is not even Markov.

Definition. Let $Z_0, Z_1, Z_2, \cdots$ be random variables defined on a probability space $(\Omega, B, P)$ and taking values in $C$. For each $z$ in $C$, let $\mathscr{M}(z)$ be a nonempty set of probability distributions on $C$. Then we call $Z_0, Z_1, Z_2, \cdots$ an $\mathscr{M}$-sequence starting at $z$ if

(a) $Z_0 \equiv z$, and

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

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

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

(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.)

If $T = \infty$, and $a_1 \geqq 0$ for all $i$, define $\sum_{i=0}^{T} a_i$ to be $\lim_{N\to\infty} \sum_{i=1}^{N} a_i$, which may possibly take the value $+\infty$. Also let $I_B$ denote the indicator random variable of the event $B$.

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

Theorem 2.1. (Compare [6] Theorem 2.15.2.) *Let $N$ be some positive integer and let $r$, $\{f_k\}$ $(k = 0, 1, 2, \cdots, N)$ be real-valued nonnegative Baire functions on $C$ such that:*

$$f_{k+1}(z) \geqq r(z) + Ef_k(Z) \tag{2.1}$$

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

*Then, for* $Z_0, Z_1, \cdots$, *an* $\mathscr{M}$*-sequence starting at* $z$, *we have*:

$$f_N(z) \geqq E[\textstyle\sum_{k=0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)})] , \tag{2.2}$$

*for all* $z \in C$, *and for all stopping times* $T$.

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

$$\begin{aligned} E[f_{N-k-1}(Z_{k+1}) + \textstyle\sum_{i=0}^{k} r(Z_i) \mid Z_0, Z_1, \cdots, Z_k] & \\ = \textstyle\sum_{i=0}^{k-1} r(Z_i) + r(Z_k) + E[f_{N-k-1}(Z_{k+1}) \mid Z_0, Z_1, \cdots, Z_k] & \\ \leqq \textstyle\sum_{i=0}^{k-1} r(Z_i) + f_{N-k}(Z_k) , & \end{aligned}$$

where the inequality follows by condition (2.1). This verifies the supermartingale property. Then, with a sign change, we may apply a submartingale theorem due to Doob [5] (see also Neveu [12] Section 4.5) and the result (2.2) follows. □

COROLLARY 1. *Note that equality in* (2.1) *implies equality in* (2.2).

THEOREM 2.2. *Let* $N$ *be some positive integer and let* $r$, $\{f_k\}$ $(k = 0, 1, 2, \cdots, N)$ *be real-valued nonnegative Baire functions on* $C$ *such that*:

$$f_{k+1}(z) \leqq r(z) + Ef_k(Z) , \tag{2.3}$$

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

*Then, whenever* $Z_0, Z_1, \cdots$ *form an* $\mathscr{M}$*-sequence starting at* $z$ *and satisfying*:

$$E[f_{N-k}(Z_k) + \textstyle\sum_{i=0}^{k-1} r(Z_i)] < \infty \tag{2.4}$$

*for all* $k = 1, 2, \cdots, N$,
*we have*:

$$f_N(z) \leqq E[\textstyle\sum_{k=0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)})] , \tag{2.5}$$

*for all* $z \in C$, *and for all stopping times* $T$.

PROOF. By conditions (2.3) and (2.4) we have that $\{f_{N-k}(Z_k) + \sum_{i=0}^{k-1} r(Z_i)\}$ $(k = 1, 2, \cdots, N)$ is a submartingale. Then (2.5) follows from a martingale theorem which can be found in Neveu, [12] Section 4.5. □

COROLLARY 2. *Note that equality in* (2.3) *implies equality in* (2.5).

THEOREM 2.3. (Compare [4] Theorem 2 and [6] Theorem 2.12.1.) *Let* $r, f$ *be real-valued nonnegative Baire functions on* $C$ *satisfying*:

$$f(z) \geqq r(z) + Ef(Z) , \tag{2.6}$$

*whenever* $P(Z) \in \mathscr{M}(z)$ *and for all* $z \in C$.

*Then for* $Z_0, Z_1, \cdots$, *an* $\mathscr{M}$*-sequence starting at* $z$, *we have*:

$$f(z) \geqq E[\textstyle\sum_{k=0}^{T-1} r(Z_k) + f(Z_T) \cdot I_{T<\infty}] , \tag{2.7}$$

*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 $N = 0, 1, 2, \cdots$:

$$\begin{aligned} f(z) &\geqq E[\textstyle\sum_{k=0}^{T(N)-1} r(Z_k) + f(Z_{T(N)})] \\ &\geqq E[\textstyle\sum_{0 \leqq k < T(N)} r(Z_k) + f(Z_{T(N)}) \cdot I_{T<N}] \\ &= E[\textstyle\sum_{0 \leqq k < T(N)} r(Z_k) + f(Z_T) \cdot I_{T<N}] \\ &\nearrow E[\textstyle\sum_{0 \leqq k < T} r(Z_k) + f(Z_T) \cdot I_{T<\infty}] \qquad \text{as} \quad N \to \infty , \end{aligned}$$

by monotone convergence, since $r, f$ are nonnegative. Hence the result (2.7) is proved. □

THEOREM 2.4. *Let $r, f$ be nonnegative Baire functions on $C$ such that*

$$f(z) \leqq r(z) + Ef(Z) , \tag{2.8}$$

*whenever $P(Z) \in \mathscr{M}(z)$ and for all $z \in C$. Then, whenever $Z_0, Z_1, \cdots$ form an $\mathscr{M}$-sequence starting at $z$, and $T$ is any stopping time satisfying*

$$E[f(Z_k) + \textstyle\sum_{i=0}^{k-1} r(Z_i)] < \infty \qquad \textit{for} \quad k = 1, 2, \cdots \tag{2.9a}$$

$$\liminf_{N\to\infty} E[f(Z_N) \cdot I_{T \geqq N}] = 0 , \tag{2.9b}$$

*we have:*

$$f(z) \leqq E[\textstyle\sum_{k=0}^{T-1} r(Z_k) + f(Z_T) \cdot I_{T<\infty}] , \tag{2.10}$$

*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:

$$\begin{aligned} f(z) &\leqq E[\textstyle\sum_{k=0}^{T(N)-1} r(Z_k) + f(Z_{T(N)})] \\ &= E[\textstyle\sum_{k=0}^{T(N)-1} r(Z_k) + f(Z_{T(N)}) \cdot I_{T<N}] + E[f(Z_{T(N)}) \cdot I_{T \geqq N}] . \end{aligned}$$

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

$$f(z) \leqq E[\textstyle\sum_{k=0}^{T-1} r(Z_k) + f(Z_T) \cdot I_{T<\infty}] + \liminf_{N\to\infty} E[f(Z_{T(N)}) \cdot I_{T \geqq N}] .$$

The result (2.10) now follows by condition (2.9b). □

We now apply the theorems of this section to the branching process of Section 1. In terms of $\mathscr{M}$-sequences, we assume that $Z_0, Z_1, \cdots$ is an $\mathscr{M}$-sequence starting at $z$ for

$$\mathscr{M}(z) = \{P(X_1 + \cdots + X_z) : \{X_i\} \text{ are indep.}, \quad P(X_i) \in \mathscr{M}\} . \tag{2.11}$$

We will need to assume the well-known result that if $\phi$ is the generating function of a nonnegative 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 $x \geqq 0$. This result follows from the convexity of $\phi$ and is proved for instance in Karlin [11] Chapter 11.3.

Throughout this paper, the set $C$, referred to in Section 2, will be taken to be the set of all nonnegative integers. Also it will be understood that $\mathscr{M}$ is some subset of probability distributions concentrated on $C$. We will adopt the convention $0^0 = 1$, and define the time to extinction, $T_e$, by:

$$\begin{aligned} T_e &= \min[n : Z_n = 0] \quad &&\text{if} \quad Z_n = 0 \quad \text{for some} \quad n, \\ &= \infty &&\text{otherwise.} \end{aligned}$$

**3. The probability of extinction and the expected time to extinction.** In this section, we consider certain families $\mathscr{M}$ of distributions and find bounds on the probability of extinction of the branching process described in Section 1. In Sub-section 3.1, upper bounds are derived for an $\mathscr{M}$ which may at first glance appear somewhat artificial. However two interesting special cases are considered in Sub-section 3.2 and Sub-section 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 Sub-section 3.4 and a special case also studied by Freedman and Purves [7] and by Goodman [9] is presented in Sub-section 3.5. (Goodman also treated the example of Sub-section 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.

3.1. *Upper bounds.*

THEOREM 3.1. (*Extinction in a finite time $N$*). *For any positive integer $N$, let $\gamma_0 = 0, \gamma_1, \gamma_2, \cdots, \gamma_N$ be a sequence of nonnegative real numbers. Take*

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

*and set*

$$\mathscr{M}(z) = \{P(X_1 + \cdots + X_z) : \{X_i\} \ \text{are indep.}, \ P(X_i) \in \mathscr{M}\}.$$

*Assume $\mathscr{M}$ is nonempty. Then for $Z_0, Z_1, \cdots$, an $\mathscr{M}$-sequence starting at $z$, we have*:

$$\Pr[Z_N = 0] \leqq (\gamma_N)^z. \tag{3.1}$$

PROOF. We apply Theorem 2.1 with $C =$ the set of nonnegative integers, $r(z) \equiv 0$, $T = N$, $f_k(z) = (\gamma_k)^z$ for $z \in C$ and $0 \leqq k \leqq N$. Then $r$ and $f_k$ are nonnegative and for $P(Z) \in \mathscr{M}(z)$:

$$\begin{aligned} r(z) + Ef_k(Z) &= E[\gamma_k^{X_1 + \cdots + X_z}] \\ &= \textstyle\prod_{i=1}^z E[\gamma_k^{X_i}] \qquad &&\text{(since } \{X_i\} \text{ are conditionally indep.)} \\ &= (\gamma_{k+1})^z &&\text{(by definition of } \mathscr{M}) \\ &= f_{k+1}(z), \end{aligned}$$

which verifies the hypotheses of Theorem 2.1. The result (3.1) now follows by noting that:

$$\begin{aligned} E[\textstyle\sum_{k=0}^{T(N)-1} r(Z_k) + f_{N-T(N)}(Z_{T(N)})] &= E[f_0(Z_N)] \\ &= E[\gamma_0^{Z_N}] \\ &= \Pr[Z_N = 0]\,, \end{aligned}$$

since $\gamma_0 = 0$ and we have adopted the convention $0^0 = 1$. □

COROLLARY 1 (*The expected time to extinction*). *Suppose* $\gamma_0 = 0, \gamma_1, \gamma_2, \cdots$ *is a sequence of real numbers in* $[0, 1]$. *Take*

$$\mathscr{M} = \{P(X) : E[\gamma_j{}^X] \leqq \gamma_{j+1}\,,\ \textit{for all}\ \ j\}\,.$$

*Then*

$$E[T_e] \geqq \textstyle\sum_{i=0}^{\infty}[1 - (\gamma_i)^z]\,, \tag{3.2}$$

*where* $T_e$ *is the time to extinction.*

NOTE. Since $0 \leqq \gamma_i \leqq 1$ for all $i$ and $z \geqq 0$, all terms in the infinite sum in (3.2) are nonnegative. 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.

PROOF. The result (3.2) is obtained by noting that

$$\begin{aligned} E[T_e] &= \textstyle\sum_{i=0}^{\infty} \Pr[T_e > i] \\ &= \textstyle\sum_{i=0}^{\infty} \Pr[Z_i > 0] \\ &= \textstyle\sum_{i=0}^{\infty} (1 - \Pr[Z_i = 0]) \\ &\geqq \textstyle\sum_{i=0}^{\infty} [1 - (\gamma_i)^z]\,. \end{aligned}$$ □

COROLLARY 2 (*Achievement of bounds*). *If there exists a Galton–Watson process with litter sizes X such that* $E[\gamma_j{}^X] = \gamma_{j+1}$, *for all j, then this achieves the bounds* (3.1), (3.2).

This follows by Harris [10] Chapter 1, Theorem 6.1, or by applying Corollary 1 of Section 2 with $\mathscr{M}$ containing only $P(X)$.

Note on existence. For general $\{\gamma_i\}$, $\mathscr{M}$ may be empty and no Galton–Watson process defined as above will exist. However, in each of two special cases studied in Sub-section 3.2 and Sub-section 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 $\mathscr{M}$ and that the Galton–Watson process of Corollary 2 will exist.

THEOREM 3.2 (*Eventual extinction*). *Take* $\alpha > 0$, *let* $\mathscr{M} = \{P(X) : E[\alpha^X] \leqq \alpha\}$ *and set* $\mathscr{M}(z) = \{P(X_1 + \cdots + X_z) : \{X_i\}$ *are indep.* $P(X_i) \in \mathscr{M}\}$. *Then, for* $Z_0, Z_1, \cdots$, *an* $\mathscr{M}$*-sequence starting at z, we have*:

$$\Pr[Z_N = 0\ \ \textit{for some}\ \ N] \leqq \alpha^z\,. \tag{3.3}$$

PROOF. We apply Theorem 2.3 with $C =$ the set of nonnegative integers,

$r(z) \equiv 0$, $f(z) = \alpha^z$ for $z \in C$ and

$$\begin{aligned} T &= \min[n : Z_n = 0] \quad && \text{if} \quad Z_n = 0 \quad \text{for some} \quad n \\ &= \infty && \text{otherwise.} \end{aligned}$$

Then $r$ and $f$ are nonnegative, and for $P(Z) \in \mathscr{M}(z)$:

$$\begin{aligned} r(z) + Ef(Z) &= E[\alpha^{X_1 + \cdots + X_z}] \\ &= \textstyle\prod_{i=1}^{z} E[\alpha^{X_i}] \\ &\leqq \alpha^z \\ &= f(z), \end{aligned}$$

which verifies the conditions of Theorem 2.3. Thus $f(z) \geqq E[f(Z_T) \cdot I_{T<\infty}]$.

Let

$$\begin{aligned} v(z) &= 1 \quad && \text{if} \quad z = 0 \\ &= 0 && \text{otherwise}, \end{aligned}$$

then since $f(z) \geqq v(z)$ for all $z \in C$, we have:

$$\begin{aligned} f(z) &\geqq E[f(Z_T) \cdot I_{T<\infty}] \\ &\geqq E[v(Z_T) \cdot I_{T<\infty}] \\ &= \textstyle\int_{T<\infty} v(Z_T)\, dP \\ &= \Pr[T < \infty] \\ &= \Pr[Z_N = 0 \quad \text{for some} \quad N]. \end{aligned}$$

This proves (3.3). □

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* (3.3).

This follows by Harris [10] Chapter 1, Theorem 6.1.

3.2. *Example* 1 (*Bounded litter sizes*). Let $k$ be some positive integer and $m$ a real number such that $0 \leqq m \leqq k$. 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 Section 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 $\mathscr{M}$-sequences studied in Sub-section 3.1 if $\alpha$ and $\{\gamma_k\}$ are chosen appropriately.

LEMMA 1. *Let $\mathscr{M} = \{P(X) : EX \geqq m,\ 0 \leqq X \leqq k\}$. Then for $0 \leqq \beta \leqq 1$ and $P(X) \in \mathscr{M}$, we have*:

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

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 $g(x) \leqq f(x)$ for $0 \leqq x \leqq k$.

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

Hence for $P(X) \in \mathscr{M}$, we have $Eg(X) \leqq Ef(X)$ or:

$$E[\beta^X] \leqq E[1 - (1 - \beta^k) \cdot X/k] \leqq 1 - (1 - \beta^k) \cdot m/k\ ,$$

which completes the proof of Lemma 1. □

THEOREM 3.3. *Let* $\mathscr{M} = \{P(X): EX \geqq m,\ 0 \leqq X \leqq k\}$ *and set* $\mathscr{M}(z) = \{P(X_1 + \cdots + X_z): \{X_i\}$ *are indep.,* $P(X_i) \in \mathscr{M}\}$. *Then for* $Z_0, Z_1, \cdots$ *an* $\mathscr{M}$-*sequence starting at* $z$, *we have*:

(A) $\Pr[Z_N = 0] \leqq (\gamma_N)^z$, $(N = 0, 1, 2, \cdots)$, *where* $\gamma_N$ *is defined recursively by*:

$$\gamma_0 = 0$$
$$\gamma_{j+1} = (1 - m/k) + \gamma_j^{\,k} \cdot (m/k) \qquad j = 0, 1, \cdots, N - 1\ ;$$

(B) $\Pr[Z_N = 0$ *for some* $N] \leqq \alpha^z$, *where* $\alpha$ *is the smaller root in* $[0, 1]$ *of*:

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

(C) $E[T_e] \geqq \sum_{i=0}^{\infty} [1 - (\gamma_i)^z]$, *where* $\{\gamma_i, i \geqq 0\}$ *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* $\{Z_i'\}$ *with litter size distribution given by*:

$$X' = 0 \quad \text{with probability} \quad 1 - (m/k)$$
$$\phantom{X'} = k \quad \text{with probability} \quad m/k\ .$$

PROOF. (A). Applying Lemma 1 with $\beta = \gamma_j$, we obtain $E[\gamma_j^{\,X}] \leqq \gamma_{j+1}$ for $P(X) \in \mathscr{M}$ and $j = 0, 1, 2, \cdots$. The result (A) now follows by Theorem 3.1.

(B). Applying Lemma 1 with $\beta = \alpha$, we obtain $E[\alpha^X] \leqq \alpha$ for $P(X) \in \mathscr{M}$. The result (B) now follows by Theorem 3.2. Note that if $m \leqq 1$, then $\alpha = 1$ and if $m > 1$ then $0 < \alpha < 1$.

(C). This follows directly from (A) and Corollary 1 of Theorem 3.1.

(D). Note that for $X \sim X'$, $E[\gamma_j^{\,X}] = \gamma_{j+1}$ and $E[\alpha^X] = \alpha$. Hence the proof of (D) follows from Corollary 2 of Theorem 3.1 and Corollary 1 of Theorem 3.2. □

REMARKS. This example and the results (A), (B), (D) of Theorem 3.3 were first presented by Goodman [9].

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 $EX \geqq m$ and $0 \leqq X \leqq k$, 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. *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 Section 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 $\mathscr{M}$-sequences studied in Sub-section 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 $0 \leqq (m^2 + \sigma^2)/m < 1$ is impossible, since for $P(X) \in \mathscr{M}$, we have:

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

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

LEMMA 2. *Let* $\mathscr{M} = \{P(X) : E[X] = m, \operatorname{Var}[X] \leqq \sigma^2\}$. *Then for all* $0 \leqq \beta < 1$ *and* $P(X) \in \mathscr{M}$, *we have* $E[\beta^X] \leqq \psi(\beta)$, *where either*

(i) $\psi(\beta) = 1 - m/h + \beta^h \cdot m/h$, *if* $h = (m^2 + \sigma^2)/m$ *is a positive integer or*

(ii) $\psi(\beta) = 1 - [m(1 + 2h) - (\sigma^2 + m^2)]/h(h + 1) + \beta^h \cdot [m(1 + h) - (\sigma^2 + m^2)]/h + \beta^{h+1} \cdot [m^2 + \sigma^2 - hm]/(h + 1)$, *if* $h < (m^2 + \sigma^2)/m < h + 1$ *with* $h$ *a positive integer.*

PROOF. (i) Suppose first that $h = (m^2 + \sigma^2)/m$ is an integer. For $0 < \beta < 1$. define

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

It is easy to show $f(x)$ is tangent to $\beta^x$ at $x = h$ and $\beta^x \leqq f(x)$ for all $x \geqq 0$ with equality at $x = 0$ and $x = h$. This still holds in the limit as $\beta \to 0$. Since $X$ is concentrated on the nonnegative integers, it follows that $E[\beta^X] \leqq Ef(X) \leqq \psi(\beta)$, where to obtain the latter inequality we have used the fact that $E[X] = m$ and $\operatorname{Var}[X] \leqq \sigma^2$.

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

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

It is easy to check that $\beta^x \leqq f(x)$ for all $0 \leqq x \leqq h$ and $x \geqq h + 1$ with equality at $x = 0, h, h + 1$. This result still holds in the limit $\beta \to 0$ for then $f(x) = (x - h)(x - h + 1)/h(h - 1)$. Since $X$ is concentrated on the nonnegative integers we have that $E[\beta^X] \leqq Ef(X) \leqq \psi(\beta)$, where to derive the second inequality we have used $E[X] = m$ and $\operatorname{Var}[X] \leqq \sigma^2$. This completes the proof of Lemma 2. □

THEOREM 3.4. *Let* $\mathscr{M} = \{P(X) : EX = m, \operatorname{Var} X \leqq \sigma^2\}$ *and set* $\mathscr{M}(z) = \{P(X_1 + \cdots + X_z) : \{X_i\}$ *are indep.*, $P(X_i) \in \mathscr{M}\}$. *For* $0 \leqq \beta \leqq 1$, *define* $\psi(\beta)$ *as in Lemma* 2. *Then for* $Z_0, Z_1, \cdots$, *an* $\mathscr{M}$-*sequence starting at* $z$, *we have*:

(A) $\Pr[Z_N = 0] \leqq (\gamma_N)^z$, *where* $\gamma_N$ *is defined recursively by*:

$$\gamma_0 = 0,$$
$$\gamma_{j+1} = \psi(\gamma_j) \qquad \textit{for} \quad j = 0, 1, 2, \cdots, N - 1;$$

(B) $\Pr[Z_N = 0$ *for some* $N] \leqq \alpha^z$, *where* $\alpha$ *is the smaller root in* $[0, 1]$ *of*

$$\alpha = \psi(\alpha) .$$

(C) $E[T_e] \geqq \sum_{i=0}^{\infty} [1 - (\gamma_i)^z]$, *where* $\{\gamma_i, i \geqq 0\}$ *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* $\{Z_i'\}$ *with litter size distribution given by*: *in case* (i)

$$\begin{aligned} X' &= 0 \quad \text{with probability} \quad 1 - m/h , \\ &= h \quad \text{with probability} \quad m/h ; \end{aligned}$$

*or in case* (ii)

$$\begin{aligned} X' &= 0 \quad \text{with probability} \quad 1 - [m(1 + 2h) - (\sigma^2 + m^2)]/h(h + 1) , \\ &= h \quad \text{with probability} \quad [m(1 + h) - (\sigma^2 + m^2)]/h , \\ &= h + 1 \quad \text{with probability} \quad [\sigma^2 + m^2 - hm]/(h + 1) . \end{aligned}$$

PROOF. The proof parallels that of Theorem 3.3.

Applying Lemma 2 with $\beta = \gamma_j$, we obtain $E[\gamma_j^X] \leqq \gamma_{j+1}$ for $P(X) \in \mathscr{M}$ and $j = 0, 1, 2, \cdots$. The result (A) now follows by Theorem 3.1.

Applying Lemma 2 with $\beta = \alpha$, we get $E[\alpha^X] \leqq \alpha$ for $P(X) \in \mathscr{M}$. The result (B) now follows by Theorem 3.2. Note that if $m \leqq 1$ then $\alpha = 1$, whereas if $m > 1$ then $0 < \alpha < 1$.

The result (C) follows directly from (A) and Corollary 1 of Theorem 3.1.

Since for $X \sim X'$, $E[\gamma_j^X] = \gamma_{j+1}$ and $E\alpha^X = \alpha$, the result (D) follows from Corollary 2 of Theorem 3.1 and Corollary 1 of Theorem 3.2. □

REMARK. As in Example 1, we see that "Bold play is optimal."

3.4. *Lower bounds.*

THEOREM 3.5. (*Extinction in a finite time* $N$). *For any positive integer* $N$, *let* $g_0 = 0, g_1, g_2, \cdots, g_N$ *be a sequence of real numbers* $(0 \leqq g_i \leqq 1; i = 1, 2, \cdots, N)$,

$$\mathscr{M} = \{P(X) : E[g_j^X] \geqq g_{j+1}, j = 1, 2, \cdots, N - 1\} ,$$

*and set*

$$\mathscr{M}(z) = \{P(X_1 + \cdots + X_z) : X_i \text{ are indep.}, \ P(X_i) \in \mathscr{M}\}$$

*for* $z = 0, 1, 2, \cdots$.

*Then for* $Z_0, Z_1, \cdots$, *an* $\mathscr{M}$-*sequence starting at* $z$, *we have*:

$$\Pr[Z_N = 0] \geqq (g_N)^z . \tag{3.4}$$

PROOF. We apply Theorem 2.2 with $C =$ the set of nonnegative integers, $r(z) \equiv 0$, $T = N$, $f_k(z) = (g_k)^z$ for $z = 0, 1, 2, \cdots$ and $k = 1, 2, \cdots, N$. Then $r$ and $f_k$ are nonnegative and for $P(Z) \in \mathscr{M}(z)$:

$$\begin{aligned} r(z) + Ef_k(Z) &= E[g_k^{X_1 + \cdots + X_z}] \\ &= \textstyle\prod_{i=1}^{z} E[g_k^{X_i}] \\ &\geqq (g_{k+1})^z \\ &= f_{k+1}(z) . \end{aligned}$$

Also the integrability condition, $E[f_{N-k}(Z_k) + \sum_{i=0}^{k-1} r(Z_i)] < \infty$, $k = 0, 1, 2, \cdots, N$, 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 (3.4) follows by noting that (as in Theorem 3.1):

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

COROLLARY 1 (*The expected time to Extinction*). *Suppose* $g_0 = 0, g_1, g_2, \cdots$ *is a sequence of real numbers in* $[0, 1]$. *Take*

$$\mathscr{M} = \{P(X) : E[g_j{}^X] \geqq g_{j+1} \text{ for all } j\}\,.$$

*Then*

$$E[T_e] \leqq \textstyle\sum_{i=0}^{\infty} [1 - (g_i)^z]\,, \tag{3.5}$$

*where* $T_e$ *is the time to extinction.*

PROOF. The proof follows as in the proof of Corollary 1 of Theorem 3.1 but with the inequalities reversed. $\Box$

COROLLARY 2 (*Achievement of bounds*). *If there exists a Galton–Watson process with sizes* $X$ *such that* $E[g_j{}^X] = g_{j+1}$, *for all* $j$, *then this achieves the bounds* (3.4), (3.5).

This follows by Harris [10] Chapter 1, Theorem 6.1, or by applying Corollary 2 of Section 2 with $\mathscr{M}$ containing only $P(X)$.

COROLLARY 3 (*Eventual extinction*). *Suppose* $g_N \to \alpha$ *as* $N \to \infty$, *where* $0 \leqq \alpha \leqq 1$. *Then, since* $\Pr[Z_N \geqq 0$ *for some* $N] \geqq \Pr[Z_N = 0]$ *and the left-hand side is independent of* $N$, *we have*:

$$\Pr[Z_N = 0 \text{ for some } N] \geqq \alpha^z\,. \tag{3.6}$$

*Furthermore, if there, exists a Galton–Watson process with litter sizes* $X$ *such that* $E[\alpha^X] = \alpha$, *then this process achieves the bound* (3.6).

This follows by Harris [10] Chapter 1, Theorem 6.1.

In the next section, we shall given an example in which an interesting family of offspring distributions is defined, and the $g_i$ are chosen in such a way that this is a subset of $\mathscr{M}$ as defined in Theorem 3.5. Also the Galton–Watson processes, as described in Corollaries 2 and 3, will be shown to exist.

3.5. *Example* 3. In this example, we shall use similar notation to that of Goodman [9]. Let $H \subseteqq C$ be a set of nonnegative 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 Section 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 $\mathscr{M}$-sequences studied in Sub-section 3.4 if the $\{g_k\}$ are chosen appropriately.

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

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

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

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 $g(x) \geqq f(x)$ for all real $x$. If $m' = 0$, $f(x) = (m^* - x)/m^*$ and because there are no members of $H$ between $m' = 0$ and $m^*$ we have $g(x) \geqq f(x)$ for all $x$ in $H$.

Now suppose $0 < \beta \leqq 1$. Then since $f$ is linear, $g$ is convex, $f(m') = g(m')$, and $f(m^*) = g(m^*)$, we have $f(x) \leqq g(x)$ for all real $x \leqq m'$ and $x \geqq m^*$.

Thus, by construction $g(x) \geqq f(x)$ for all $x \in H$ and $0 \leqq \beta \leqq 1$, and hence, for $P(X) \in \mathscr{M}$, we have $Eg(X) \geqq Ef(X)$, or:

$$\begin{aligned} E[\beta^X] &\geqq E[(m^* - X)\beta^{m'} + (X - m')\beta^{m^*}]/d \\ &= [m^* \cdot \beta^{m'} - m' \cdot \beta^{m^*} - (\beta^{m'} - \beta^{m^*}) \cdot EX]/d \\ &\geqq [m^* \cdot \beta^{m'} - m' \cdot \beta^{m^*} - (\beta^{m'} - \beta^{m^*}) \cdot m]/d \\ &= [(m^* - m)\beta^{m'} + (m - m')\beta^{m^*}]/d , \end{aligned}$$

which completes the proof of Lemma 3. □

THEOREM 3.6. *Let $\mathscr{M} = \{P(X): X$ concentrated on $H$, $EX \leqq m\}$ and set $\mathscr{M}(z) = \{P(X_1 + \cdots + X_z): \{X_i\}$ are indep., $P(X_i) \in \mathscr{M}\}$. Then for $Z_0, Z_1, Z_2, \cdots$, an $\mathscr{M}$-sequence starting at $z$, we have:*

(A) $\Pr[Z_N = 0] \geqq (g_N)^z$, $(N = 0, 1, 2, \cdots)$; *where*

(i) *if $m \in H$, $g_N = 0$ for all $N$, or*

(ii) *if $m \notin H$, $g_N$ is defined recursively by $g_0 = 0$, $g_{j+1} = [(m^* - m)g_j^{m'} + (m - m')g_j^{m^*}]/d$ $(j = 0, 1, 2, \cdots, N - 1)$, where $m^*, m'$, $d$ are defined as in Lemma* 3;

(B) $\Pr[Z_N = 0$ *for some* $N] \geqq \alpha^z$, *where*

(i) *if $m \in H$, $\alpha = 0$, or*

(ii) *if $m \notin H$, $\alpha$ is the smaller root in* $[0, 1]$ *of*: $\alpha = [(m^* - m)\alpha^{m'} + (m - m')\alpha^{m^*}]/d$;

(C) $E[T_e] \leqq \sum_{i=0}^{\infty} [1 - (g_i)^z]$, *where $\{g_i, i \geqq 0\}$ 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 $(m^* - m)/d$; $X' = m^*$ with probability $(m - m')/d$.*

PROOF. We consider first the case $m \in H$. Since $m > 0$, this means that $m \geqq 1$. Then for the $\mathscr{M}$-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 with $\beta = g_j$, we obtain $E[g_j{}^X] \geqq g_{j+1}$ for $P(X) \in \mathscr{M}$, and $j = 0, 1, 2, \cdots$. The result A(ii) now follows by Theorem 3.5. By Harris [10] 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 3.5.

The result C(ii) follows directly from A(ii) and Corollary 1 of Theorem 3.5.

Since for $X \sim X'$, $E[g_j{}^X] = g_{j+1}$ and $E[\alpha^X] = \alpha$, the result D(ii) follows by Corollaries 2 and 3 of Theorem 3.5. This completes the proof of Theorem 3.6. □

REMARK 1. If $m' > 0$, then $\alpha = g_0 = g_1 = g_2 = \cdots = 0$. In fact, with probability one, $Z_N' \geqq z \cdot (m')^N$.

REMARK 2. The result agrees with that of Freedman and Purves [7], who considered the special case with $H = 0, 2, 3, 4, 5, \cdots$ and $0 < m < 2$. Then the optimal strategy is for the litter sizes to have the distribution:

$$\begin{aligned} X' &= 0 \quad \text{with probability} \quad 1 - m/2\,, \qquad \text{and} \\ X' &= 2 \quad \text{with probability} \quad m/2\,. \end{aligned}$$

REMARK 3. This example was discussed also in Goodman [9] in the more general case where $m$ and $H$ (he uses $\bar{H}$) are allowed to depend on the time period $n$.

REMARK 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 $EX \leqq m$, the optimal strategy is for the population to have litters with the "smallest variability" in size.

**4. The probability that a generation size will exceed a given number and the expected maximum generation size.** We will need to define the following stopping time.

DEFINITION. For each nonnegative integer $l$, let

$$\begin{aligned} T_l &= \min[n : Z_n \geqq l]\,, \qquad \text{if} \quad Z_n \geqq l \quad \text{for some} \quad n\,, \\ &= \infty \qquad\qquad\qquad\quad \text{otherwise}\,; \end{aligned}$$

THEOREM 4.1. *Take $l > 0$, $\alpha > 1$, $\mathscr{M} = \{P(X), E[\alpha^X] \leqq \alpha\}$ and set $\mathscr{M}(z) = \{P(X_1 + \cdots + X_z) : \{X_i\}$ are indep., $P(X_i) \in \mathscr{M}\}$.*

*Then for $Z_0, Z_1, \cdots$, an $\mathscr{M}$-sequence starting at $z$, we have*:

$$\begin{aligned} (4.1) \qquad \Pr[Z_N \geqq l \ \textit{for some} \ N] &\leqq (\alpha^z - 1)/(\alpha^l - 1)\,, \qquad (0 \leqq z \leqq l)\,; \\ &= 1\,, \qquad\qquad\qquad\qquad\quad (z \geqq l)\,. \end{aligned}$$

PROOF. Trivially if $Z_0 = z \geqq l$, then $\Pr[Z_N \geqq l$ for some $N] = 1$. Suppose $z$ is some nonnegative integer less than or equal to $l$. We apply Theorem 2.3 with

$C =$ the set of nonnegative integers, $T = T_l$, $r(z) \equiv 0$ and $f(z) = (\alpha^z - 1)/(\alpha^l - 1)$ if $0 \leqq z \leqq l$, $f(z) = 1$ if $z \geqq l$.

Then $r$ and $f$ are nonnegative, and since $f(z) \leqq (\alpha^z - 1)/(\alpha^l - 1)$ for all $z \geqq 0$, we have for $P(Z) \in \mathscr{M}(z)$:

$$\begin{aligned} r(z) + Ef(Z) &\leqq E[(\alpha^Z - 1)/(\alpha^l - 1)] \\ &= E[\alpha^{X_1 + \cdots + X_z} - 1]/[\alpha^l - 1) \\ &= [\prod_{i=1}^z E(\alpha^{X_i}) - 1]/(\alpha^l - 1) \\ &\leqq (\alpha^z - 1)/(\alpha^l - 1) . \end{aligned}$$

Clearly $r(z) + Ef(Z) \leqq 1$. Hence $r(z) + Ef(Z) \leqq f(z)$.

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

$$\begin{aligned} (\alpha^z - 1)/(\alpha^l - 1) &\geqq E[f(Z_T) \cdot I_{T<\infty}] \\ &\geqq E[v(Z_T) \cdot I_{T<\infty}] \\ &= \Pr[T < \infty] \\ &= \Pr[Z_N \geqq l \text{ for some } N] , \end{aligned}$$

which completes the proof. □

Corollary 1. *For any $\mathscr{M}$-sequence starting at $z$, we have*:

$$E[\sup_n Z_n] \leqq z + (\alpha^z - 1) \cdot \sum_{i=z+1}^\infty (\alpha^i - 1)^{-1} . \tag{4.2}$$

Proof. Let $M = \sup_n Z_n$. Then

$$\begin{aligned} E[M \mid Z_0 = z] &= z + \sum_{i=z+1}^\infty \Pr[M \geqq i \mid Z_0 = z] \\ &= z + \sum_{i=z+1}^\infty \Pr[T_i < \infty \mid Z_0 = z] \\ &\leqq z + \sum_{i=z+1}^\infty (\alpha^z - 1)/(\alpha^i - 1) , \end{aligned}$$

which proves the result (4.2). □

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 $l$ in number for the branching process described in Section 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 4. *Let $\mathscr{M} = \{P(X) : 0 \leqq X \leqq k,\ EX \leqq m\}$. Then for $\alpha > 1$ and $P(X) \in \mathscr{M}$, we have*:

$$E[\alpha^X] \leqq 1 - m/k + \alpha^k \cdot m/k .$$

Proof. Define $g(x) = \alpha^x$ and $f(x) = 1 - x/k + \alpha^k \cdot x/k$. Then $f(x) \geqq g(x)$ for $0 \leqq x \leqq k$, and so $Eg(X) \leqq Ef(X)$ for $P(X) \in \mathscr{M}$. Substituting for $f$ and $g$

we obtain:

$$\begin{aligned} E[\alpha^X] &\leqq 1 + (EX) \cdot (\alpha^k - 1)/k \\ &\leqq 1 + m \cdot (\alpha^k - 1)/k \\ &= 1 - m/k + \alpha^k \cdot m/k , \end{aligned}$$

which proves the lemma. □

THEOREM 4.2. *Take* $0 < m < 1$, $\mathscr{M} = \{P(X) : EX \leqq m,\ 0 \leqq X \leqq k\}$ *and set* $\mathscr{M}(z) = \{P(X_1 + \cdots + X_z) : \{X_i\}$ *are indep.*, $P(X_i) \in \mathscr{M}\}$. *Then for* $Z_0, Z_1, \cdots$, *an* $\mathscr{M}$*-sequence starting at* $z$, *we have*:

$$\Pr[Z_N \geqq l \text{ for some } N] \leqq (\alpha^z - 1)/(\alpha^l - 1) , \qquad (\text{for } 0 \leqq z \leqq l) ; \tag{4.3}$$

$$E[\sup_n Z_n] \leqq z + \textstyle\sum_{i=z+1}^{\infty} (\alpha^z - 1)/(\alpha^i - 1) , \tag{4.4}$$

*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 the Lemma, $E[\alpha^X] \leqq \alpha$ for $P(X) \in \mathscr{M}$. Theorem 4.2 now follows by Theorem 4.1 and Corollary 1. □

REMARK 1. Attainment of bounds. The Galton–Watson process $\{Z_i'\}$, with litter size distribution:

$$\begin{aligned} X &= 0 \quad \text{with probability} \quad 1 - m/k \\ &= k \quad \text{with probability} \quad m/k , \end{aligned}$$

is an $\mathscr{M}$-sequence with $E[\alpha^X] = \alpha$. However the bounds (4.3), (4.4) are not attained in general by this process. This is due to the end effect, i.e. the undesirability of overshooting $l$. 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 $l$ exactly.

REMARK 2. If $m = 1$, then taking the limit, $\alpha \to 1$ we obtain:

$$\Pr[Z_N \geqq l \text{ for some } l] \leqq z/l ,$$

and

$$E[\sup_n Z_n] \leqq \infty .$$

REMARK 3. If we consider the same $\mathscr{M}$-sequences as in Example 4 but with $m > 1$, then clearly a sharp upper bound on $p[Z_N \geqq l$ for some $l]$ 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.

**5. The mean total population size.** In this section, we consider families, $\mathscr{M}$, 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 5.1. (*Upper bounds*). *Take $m > 0$, let $\mathscr{M} = \{P(X): EX \leqq m\}$ and set $\mathscr{M}(z) = \{P(X_1 + \cdots + X_z): \{X_i\}$ are indep. $P(X_i) \in \mathscr{M}\}$. Then for $Z_0, Z_1, \cdots$, an $\mathscr{M}$-sequence starting at $z$, we have*:

$$E[\textstyle\sum_{i=0}^{N} Z_i] \leqq z \cdot \frac{1 - m^{N+1}}{1 - m} \quad \text{if} \quad m \neq 1 \tag{5.1}$$

$$\leqq z \cdot (N + 1) \quad \text{if} \quad m = 1 \quad \text{for } N = 0, 1, 2, \cdots,$$

*and*

$$E[\textstyle\sum_{i=0}^{\infty} Z_i] \leqq z \cdot \frac{1}{1 - m} \quad \text{if} \quad m < 1 \tag{5.2}$$

$$\leqq \infty \quad \text{if} \quad m \geqq 1 .$$

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

PROOF. The proof follows by Theorem 2.1 with $r(z) = z$, $T = N$, and $f_k(z) = z \cdot (1 - m^{k+1})/(1 - m)$ if $m \neq 1$ and $f_k(z) = (k + 1) \cdot z$ if $m = 1$. □

Lower bounds can be derived similarly when $\mathscr{M} = \{P(X): EX \geqq m\}$. The results are the same as in (5.1), (5.2) but with the inequalities reversed.

**6. An associated multiplicative process.** Let $\mathscr{M}$ be some set of probability distributions concentrated on $C$, the set of all nonnegative integers. Then, for each $z$ in $C$, define a sequence of random variables $Z_0, Z_1, \cdots$ as follows:

$$Z_0 = z ,$$

$$Z_{n+1} = Z_n \cdot X_{n+1} ,$$

where $P(X_{n+1} \mid Z_0, Z_1, \cdots, Z_n) \in \mathscr{M}$, $n = 0, 1, 2, \cdots$.

More precisely, $Z_0, Z_1, Z_2, \cdots$ is an $\mathscr{M}$-sequence starting at $z$ with $\mathscr{M}(z) = \{P(zX): P(X) \in \mathscr{M}\}$. This process can be regarded as a branching process with deterministic offspring in a random environment (see also the discussion in Section 1). It is an interesting exercise to derive, for this process, the results analogous to those of Sections 3, 4, 5.

**7. Note.** Goodman [9] allows the set $\mathscr{M}$ to vary from period to period. We can denote by $\mathscr{M}_j$ the set of allowable offspring distributions in the $j$th period. The theorems of Section 2 can be adapted to handle this problem by taking $C$ to be set of all 2-tuples, $\{(z, j)\}$, where $j$ indicates the current period.

The theory of Section 2 has far-ranging potential applications to many problems in applied probability. Some of these are developed in Turnbull [18].

**Acknowledgment.** The author wishes to express his sincere thanks to Professor H. M. Taylor III for his guidance and encouragement during the preparation of this paper.

REFERENCES

[1] ATHREYA, K. B. and KARLIN, S. (1971). Branching processes with random environments I, II. *Ann. Math. Statist.* **42** 1499–1520, 1843–1858.

[2] BLACKWELL, D. (1954). On optimal systems. *Ann. Math. Statist.* **25** 394–397.

[3] BLACKWELL, D. (1964). Probability bounds via dynamic programming. *Proc. Symp. Appl. Math.* **16** 277–280. American Mathematical Society, Providence.

[4] BLACKWELL, D. (1965). Positive dynamic programming. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 415–418. Univ. of California Press.

[5] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[6] DUBINS, L. E. and SAVAGE, L. J. (1965). *How to Gamble if You Must.* McGraw-Hill, New York.

[7] FREEDMAN, D. A. and PURVES, R. (1967). Timid play is optimal II. *Ann. Math. Statist.* **38** 1284–1285.

[8] GALTON, F. and WATSON, H. W. (1874). On the probability of the extinction of families. *J. Anthrop. Inst.* **4** 138–144.

[9] GOODMAN, L. A. (1968). How to minimize or maximize the probabilities of extinction in a Galton-Watson process and in some related multiplicative population processes. *Ann. Math. Statist.* **39** 1700–1710.

[10] HARRIS, T. E. (1963). *The Theory of Branching Processes.* Prentice-Hall, Englewood Cliffs.

[11] KARLIN, S. (1966). *A First Course in Stochastic Processes.* Academic Press, New York.

[12] NEVEU, J. (1965). *Mathematical foundations of the Calculus of Probability.* Holden-Day, San Francisco.

[13] SMITH, W. (1968). Necessary conditions for almost sure extinction of a branching process with random environment. *Ann. Math. Statist.* **39** 2136–2140.

[14] SMITH, W. and WILKINSON, W. (1969). On branching processes in random environments. *Ann. Math. Statist.* **40** 814–827.

[15] STRAUCH, R. E. (1967). Measurable gambling houses. *Trans. Amer. Math. Soc.* **126** 64–72. (Correction in **130**, 184.)

[16] SUDDERTH, W. D. (1969). On the existence of good stationary strategies. *Trans. Amer. Math. Soc.* **135** 399–414.

[17] SUDDERTH, W. D. (1971). On measurable gambling problems. *Ann. Math. Statist.* **42** 260–269.

[18] TURNBULL, B. W. (1971). Bounds and optimal strategies for stochastic systems. Ph. D. dissertation, Cornell Univ.

MATHEMATICAL INSTITUTE
24-29 ST. GILES
OXFORD, ENGLAND