Some generalized age-dependent branching processes
by Kenny Sherman Crump

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY in Mathematics
Montana State University

Abstract:
Several models for branching processes are investigated which extend the age-dependent model of Bellman and Harris (1952). These models are developed primarily to provide reasonable alternatives to the assumptions that individuals live and reproduce independently and that the death of a parent and the birth of offspring must occur simultaneously.

A process in which correlations occur among siblings is obtained by assuming the life-spans (as well as the numbers of offspring) of a group of siblings are exchangeable random variables. Other necessary probabilistic relations are governed by independence. The generating function of z^(k)(t), the number of individuals alive at time t descending from k siblings born at time t = 0 , satisfies a system of integral equations, which, upon differentiation, becomes a simple renewal equation for E[z^(1)(t)z^(l)(t+τ)] which in certain cases leads to the result that Z^(k) (t)/e^α(2t + τ) converges in mean square to a r.v. W^(k) as t→∞ and, in turn, W^(k) converges in mean square to a nondegenerate r.v. W as k→∞. In the binary case, the distribution of W^(2) is continuous and, provided 1 - G1(t) =0(e^εt) , ε > 0 , it is also absolutely continuous.

Dependence between generations is introduced by assuming that if o1,...,oj,... is a sequence of individuals with oj the parent of oj+1 then the life-spans of these individuals form a Markov chain. An integral equation satisfied by the generating function of the process is used to study the probability of extinction and the first moments. A simpler model with dependence between generations is also described and it is indicated how standard techniques may be employed to study this model.

A rather general birth-and-death process is considered in which an individual may give birth at various times during its life. The number of offspring N(t) born to an individual with life-span l during the age-interval [0,t] is defined by N(t) = K(t) if t ≤ l, and N(t) = K(l) if t > l, where K(t) is an arbitrary counting process. Individuals are assumed to live and reproduce independently. If E[K(t)]<∞ for all t and E[K(0)] < 1, then Z(t), the size of the population at time t, is finite a.s. The probability of extinction of Z(t) is the smallest nonnegative root of the equation s = E[exp{N(l)(log s)}]. A renewal-type integral equation is derived for M(t) = E[Z(t)], and this equation is used to investigate the monotonicity and asymptotic properties of M(t). Again, Z(t)/e^αt converges in mean square (a.s. in some instances) to a r.v. W as t→∞. For the special case where K(t) is a compound Poisson process, the generating function of Z(t) satisfies a useful integral equation which is utilized in investigating the distribution of W. and in showing that Markov branching processes form a special case of this process.

In a process where the life-span of an individual depends on the size of his family, the first two moments satisfy systems of non-linear renewal-type integral equations. Asymptotic expressions for these moments are obtained using complex variable techniques.

195

SOME GENERALIZED AGE-DEPENDENT BRANCHING PROCESSES

by

KENNY SHERMAN CRUMP

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Mathematics

Approved:

Louis C. Barrett
Head, Major Department

Charles J. Mode
Chairman, Examining Committee

H. Loering
Graduate Dean

MONTANA STATE UNIVERSITY
Bozeman, Montana

June, 1968

ACKNOWLEDGEMENT

The author wishes to express his deepest gratitude to his advisor Dr. Charles J. Mode for his guidance and encouragement during the preparation of this thesis. The preparation of the final copy and a portion of the research reported herein was supported by the National Institute of Health Contract PH-43-67-1380 and for this timely financial assistance the author is indeed grateful. Dr. Robert H. Rodine deserves a particular word of thanks for reading a rough draft of the entire manuscript and making many helpful suggestions. Thanks, also, are due Miss Carol Jonas and Mrs. Billie Albrecht for their diligence in transforming the nearly illegible first draft into the finished copy. Finally, the author wishes to acknowledge a debt of gratitude due his wife and parents for their encouragement and patience while this work was being completed.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

ABSTRACT

Several models for branching processes are investigated which extend the age-dependent model of Bellman and Harris (1952). These models are developed primarily to provide reasonable alternatives to the assumptions that individuals live and reproduce independently and that the death of a parent and the birth of offspring must occur simultaneously.

A process in which correlations occur among siblings is obtained by assuming the life-spans (as well as the numbers of offspring) of a group of siblings are exchangeable random variables. Other necessary probabilistic relations are governed by independence. The generating function of $Z^{(k)}(t)$, the number of individuals alive at time t descending from k siblings born at time $t = 0$, satisfies a system of integral equations, which, upon differentiation, becomes a simple renewal equation for $E[Z^{(1)}(t)Z^{(1)}(t+\tau)]$ which in certain cases leads to the result that $Z^{(k)}(t)/e^{\alpha(2t+\tau)}$ converges in mean square to a r.v. $W^{(k)}$ as $t \to \infty$ and, in turn, $W^{(k)}$ converges in mean square to a nondegenerate r.v. W as $k \to \infty$. In the binary case, the distribution of $W^{(2)}$ is continuous and, provided $1 - G_1(t) = 0(e^{-\epsilon t})$, $\epsilon > 0$, it is also absolutely continuous.

Dependence between generations is introduced by assuming that if $o_1, \ldots, o_j \ldots$ is a sequence of individuals with o_j the parent of o_{j+1} then the life-spans of these individuals form a Markov chain. An integral equation satisfied by the generating function of the process is used to study the probability of extinction and the first moments. A simpler model with dependence between generations is also described and it is indicated how standard techniques may be employed to study this model.

A rather general birth-and-death process is considered in which an individual may give birth at various times during its life. The number of offspring $N(t)$ born to an individual with life-span ℓ during the age-interval $[0,t]$ is defined by $N(t) = K(t)$ if $t \le \ell$, and $N(t) = K(\ell)$ if $t > \ell$, where $K(t)$ is an arbitrary counting process. Individuals are assumed to live and reproduce independently. If $E[K(t)] < \infty$ for all t and $E[K(0)] < 1$, then $Z(t)$, the size of the population at time t, is finite a.s. The probability of extinction of $Z(t)$ is the smallest non-negative root of the equation $s = E[\exp\{N(\ell)(\log s)\}]$. A renewal-type integral equation is derived for $M(t) = E[Z(t)]$, and this equation is used to investigate the monotonicity and asymptotic properties of $M(t)$. Again, $Z(t)/e^{\alpha t}$ converges in mean square (a.s. in some instances) to a r.v. W as $t \to \infty$. For the special case where $K(t)$ is a compound Poisson process, the generating function of $Z(t)$ satisfies a useful integral equation which is utilized in investigating the distribution of W and in showing that Markov branching processes form a special case of this process.

In a process where the life-span of an individual depends on the size of his family, the first two moments satisfy systems of non-linear renewal-type integral equations. Asymptotic expressions for these moments are obtained using complex variable techniques.

Chapter I

INTRODUCTION

1.1 Statement of Problem

Generally speaking, a branching process is a mathematical model for the development of a population whose members reproduce and die, subject to laws of chance. For convenience, in this thesis we shall usually refer to the members of a population as "individuals", although the mathematical model being discussed may be applicable to bacteria, plants, atomic particles, or many other things. The reader interested in applications, as well as the basic theory of branching processes, may refer to the monograph by T. E. Harris (1963). Kendall (1966) gives an interesting account of the history of branching processes.

Most of the models for branching processes studied previously have incorporated the assumption that members of the population must not interfere with one another. In other words, they must live and reproduce independently. This assumption seems reasonable for some applications; e.g., when the population is composed of atomic particles, and there is on the average, a sizable distance between adjacent particles. However, for the study of complex biological populations, it would be advantageous to be able to introduce certain types of dependence among individuals, particularly among related individuals. Such an assumption might, for example, enable one to incorporate the effects of heredity into the framework of branching processes.

Another often undesirable feature of the more well-known models is

that the death of the parent[1] must occur simultaneously with the birth of offspring.[2] Although this feature has a natural application in populations of bacteria that reproduce by splitting, a model which allows reproduction to occur throughout the life of the parent would certainly be of interest.

The purpose of this thesis, then, is to formulate mathematical models for branching processes which incorporate some of the suggestions in the foregoing paragraphs, and then to investigate the properties of these models. Chapter II concerns a model in which the life-spans of siblings are correlated, as well as the numbers of offspring produced by siblings, but otherwise individuals live and reproduce independently. Next, two models are discussed in which an individual's life-span is influenced by the life-spans of his ancestors. In Chapter IV a fairly general process is studied wherein it is possible for an individual to give birth at various times during his life. Finally, in Chapter V a brief description of two processes is given. The first process deals with a special type of dependence between generations and the second process allows for the possibility of an individual becoming dormant and remaining in the population indefinitely but never having offspring.

Before proceeding to these models, however, it seems appropriate

1 All of the processes discussed in this thesis consider only one sex, so an individual will have only one parent in the population.

2 In age-independent processes such as Markov branching processes (Harris (1963)) it actually makes no difference whether we assume an individual dies at time t and is replaced by $n \geq 1$ offspring or that he continues to live and has $n - 1$ offspring at time t.

that we pause here to describe two of the most well-known branching processes and to state some results for these models which we will find convenient to refer to later. We shall also present some results from renewal theory which will be useful in the sequel.

1.2 The Galton-Watson Process

It was about 100 years ago that Watson and Galton (1874) formulated a model to study the problem of the extinction of family surnames. Because of the disappearance of the surnames of many families that had once occupied conspicuous positions, it had been conjectured that distinguished families are more likely to die out than ordinary ones. In order to explore this hypothesis, Galton recognized that it would be desirable to first know the probability that an ordinary family becomes extinct.

In the mathematical model developed by Galton and Watson for this purpose (called a Galton-Watson process), a man has probabilities p_0, p_1, p_2 $\cdots$ of having $0,1,2,\cdots$ sons and in turn each of these offspring has sons of his own with the same probabilities, and so on. The number of sons sired by any man is assumed independent of the number of male progeny of any other man.

An important role in the investigation of this process is played by the generating function

$$h(s) = \sum_{n=0}^{\infty} p_n s^n , \quad |s| \le 1 .$$

For instance, the smallest non-negative root of the functional equation

$h(s) = s$ turns out to be the probability that the male line originating from a single man will eventually terminate. If $h'(1) \leq 1$ this root is one (assuming $p_1 < 1$), but if $h'(1) > 1$ this root is strictly less than one. (Due to an oversight, Watson concluded erroneously that this root was always equal to one.) Although the model of Galton and Watson was overlooked for many years, numerous papers on this model have appeared in the past two decades.

1.3 The Bellman-Harris Process

Bellman and Harris (1952) studied an extension of the Galton-Watson process in which the life-spans of individuals are also taken into consideration. A single individual born at time $t = 0$ lives for a random length of time with distribution function $G(t)$. At the end of its life it is replaced by n offspring, $n = 0,1,2,\ldots$, with probability p_n. These progeny, in turn, behave in the same way as the progenitor and so the process continues for as long as individuals are present. The life-span and the number of children of each member of the population have the same probability distributions as the corresponding ones for the original individual. The other necessary probabilistic relations are governed by independence, so that this process is completely determined by the distribution $G(t)$ and the generating function $h(s) = \Sigma p_n s^n$. It is usually assumed that these functions satisfy the conditions $G(0) = 0$, $G(\infty) = 1$ and $m \equiv h'(1) < \infty$. All distribution functions that appear in this thesis are taken to be continuous from the right.

To avoid confusion, we shall call the random function $Z(t)$

giving the size of the population at time t a Bellman-Harris process, although the term "age-dependent branching process" seems to be more in vogue in the literature. However, the latter expression also appropriately describes several other models discussed in this thesis.

The generating function $F(s,t)$ of the random function $Z(t)$ is defined by

$$F(s,t) = E\left[s^{Z(t)}\right] = \sum_{k=0}^{\infty} P[Z(t)=k]s^k , \quad t \geq 0 , \quad |s| \leq 1 . \tag{1.3.1}$$

Harris (1963) proves that $F(s,t)$ satisfies the integral equation[1]

$$F(s,t) = s[1-G(t)] + \int_0^t h[F(s,t-u)]dG(u) . \tag{1.3.2}$$

Uppon differentiating (1.3.2) with respect to s and putting $s = 1$, it follows that $M(t) \equiv E[Z(t)]$ satisfies the renewal equation

$$M(t) = 1 - G(t) + m\int_0^t M(t-u)dG(u) . \tag{1.3.3}$$

The following lemma summarizes some other well-known facts about $M(t)$.

---

1 In this thesis the symbol $\int_a^b$ will mean $\int_{[a,b]}$. To denote $\int_{[a,b)}$, $\int_{(a,b]}$, and $\int_{(a,b)}$, we will use $\int_a^{b-}$, $\int_{a+}^{b}$, and $\int_{a+}^{b-}$, respectively.

Lemma 1.3.1: (i) Equation (1.3.3) has a unique solution that is bounded on every finite interval.

(ii) This solution may be written in the form

$$M(t) = 1 + (m-1)\sum_{k=1}^{\infty} m^{k-1} G_k(t) , \tag{1.3.4}$$

where G_k is the kth convolution of G with itself.

(iii) If G is not a lattice distribution[1] and there exists an α such that

$$m\int_0^{\infty} e^{-\alpha u} dG(u) = 1 \tag{1.3.5}$$

and

$$\int_0^{\infty} u e^{-\alpha u} dG(u) < \infty , \tag{1.3.6}$$

then

$$M(t)e^{-\alpha t} \to \frac{\int_0^{\infty} e^{-\alpha u}[1 - G(u)]du}{m\int_0^{\infty} t e^{-\alpha u} dG(u)} \quad \text{as} \quad t \to \infty . \tag{1.3.7}$$

These results concerning the renewal equation (1.3.3) are all given

1 We say that G is a lattice distribution if its only points of increase are integer multiples of some fixed number.

explicitly in Harris (1963) (with the exception of 1.3.4, which is easily deducible from Harris' Lemma 1, page 161). However, we shall prove parts (ii) and (iii) at the end of the next section to illustrate the results from renewal theory presented there.

1.4 Some results from renewal theory

Since the moments of several of the processes to be studied satisfy renewal-type equations, rather extensive use will be made of the results in this section. The treatment of renewal theory given here is patterned after Feller (1966).

Throughout this section we shall suppose that $F(t)$ and $H(t)$ are finite, nondecreasing, continuous from the right and equal to zero when $t \leq 0$. Of considerable interest in renewal theory is the function

$$U(t) = \sum_{k=0}^{\infty} F_k(t) , \tag{1.4.1}$$

where $F_1 = F$, F_k is the kth convolution of F with itself, $k = 2,3,\ldots$, and

$$F_0(t) = \begin{cases} 1 & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 . \end{cases}$$

<u>Lemma 1.4.1</u>: The function $U(t)$ is finite for all t.

<u>Proof</u> This fact is well known but perhaps the following proof (similar to Mode (1968a)) is of interest. Let t be fixed. If $F(t) < 1$, then

$$U(t) = \sum_{k=0}^{\infty} F_k(t) \le \sum_{k=0}^{\infty} F^k(t) = \frac{1}{1 - F(t)} < \infty .$$

If $F(t) \ge 1$, then

$$\int_0^t e^{-ax} dF(x) < 1 ,$$

for some $a > 0$. If we let

$$\overline{F}(u) = \int_0^u e^{-ax} dF(x) \qquad 0 \le u \le t ,$$

it can be shown using Laplace transforms or other techniques that the kth convolution of $\overline{F}(u)$ with itself is given by

$$\overline{F}_k(u) = \int_0^u e^{-ax} dF_k(x) \qquad 0 \le u \le t .$$

It follows from the case $F(t) < 1$ that

$$\sum_{k=0}^{\infty} \overline{F}_k(t) < \infty .$$

But

$$\sum_{k=0}^{\infty} F_k(t) \le e^{at} \sum_{k=0}^{\infty} \overline{F}_k(t)$$

and the lemma is therefore proved. The proof of the next result is similar to that of Lemma 2.5.1 and will be omitted.

Lemma 1.4.2: If $f(t)$ is bounded and $f(t) \to A$ as $t \to \infty$ and $F(\infty) < \infty$, then

$$\int_0^t f(t-u)dF(u) \to A\,F(\infty) \quad \text{as} \quad t \to \infty. \tag{1.4.2}$$

Parts (i) and (ii) of the following lemma are often called Blackwell's theorem and the key renewal theorem, respectively.

Lemma 1.4.3: Suppose F is not a lattice distribution and $F(\infty) = 1$. Then

(i)

$$U(t) - U(t-h) \to \frac{h}{\int_0^\infty y dF(y)} \quad \text{as} \quad t \to \infty \tag{1.4.3}$$

for every fixed $h > 0$,

(ii) if $f(t)$ is the difference of two bounded nondecreasing functions which are both integrable on $[0,\infty]$ and $f(t) = 0$ for $t < 0$, then

$$\int_0^t f(t-y)dU(y) \to \frac{\int_0^\infty f(y)dy}{\int_0^\infty y dF(y)} \quad \text{as} \quad t \to \infty. \tag{1.4.4}$$

The proof of this lemma may be found in Feller (1966).

Corollary: If $H(\infty) < \infty$, $F(\infty) = 1$, and $V(t)$ satisfies the equation

$$V(t) = H(t) + \int_0^t V(t-y)dF(y) ,$$

then

$$\frac{V(t)}{t} \to \frac{H(\infty)}{\int_0^\infty y dF(y)} \quad \text{as} \quad t \to \infty . \tag{1.4.5}$$

Proof: The proof follows from Lemmas 1.4.2 and 1.4.3 (i) by applying the technique set forth in XI.3 of Feller (1966).

As an application of the results of this section and also to illustrate techniques that will be used later, we now give the

Proof of Lemma 1.3.1:

It can be easily verified that $M(t)$ defined by (1.3.4) satisfies equation (1.3.3). If we define $F(t) = mG(t)$, then

$$M(t) = 1 + \frac{(m-1)}{m} \sum_{k=1}^{\infty} F_k(t)$$

and it follows from Lemma 1.4.1 that this last expression is finite, which proves (ii). If we define

$$\overline{G}(t) = m\int_0^t e^{-\alpha u} dG(u) ,$$

$$\overline{f}(t) = (1 - G(t))e^{-\alpha t} ,$$

$$\overline{M}(t) = M(t)e^{-\alpha t} ,$$

and

$$\overline{U}(t) = \sum_{k=0}^{\infty} \overline{F}_k(t) ,$$

then (1.3.3) becomes

$$\overline{M}(t) = \overline{f}(t) + \int_0^t \overline{M}(t - u)d\overline{G}(u) . \qquad (1.4.6)$$

The solution of (1.4.6) is given by

$$\overline{M}(t) = \int_0^t \overline{f}(t - y)d\overline{U}(y) ,$$

and (iii) of Lemma 1.3.1 now follows from (ii) of Lemma 1.4.3 .

The interested reader may consult Feller (1941) for the proof of uniqueness.

Chapter II

CORRELATION AMONG SIBLINGS

2.1 Introduction

In this chapter we shall study an age-dependent branching process in which the independence assumptions of the Bellman-Harris process described in Section 1.3 are relaxed so that interactions may occur among siblings. Individuals that are not siblings continue to live and reproduce independently; in particular, there is no interaction between individuals in different generations.

The process begins with k siblings born at time $t = 0$. The life-spans, as well as the numbers of offspring, of these siblings are correlated. At the end of its life an individual is replaced by its offspring who are correlated in the same way as the siblings in the first generation. As is generally true, the relaxation of independence assumptions makes matters more complicated. In order to have some homogeneity on which to base the analysis of the process, we shall assume that the life-spans $\ell_1,\ldots,\ell_n$ of any group of n siblings have the same joint distribution function $G_n(x_1,\ldots,x_n)$, and this function is invariant under permutations of the x's. In the literature, this property of $G_n(x_1,\ldots,x_n)$ is expressed by saying that the random variables $\ell_1,\ldots,\ell_n$ are "exchangeable" or "interchangeable". Likewise, we shall assume that the numbers of offspring of siblings are exchangeable random variables. Intuitively, these assumptions say that one would expect all groups of n siblings to behave in the same way which seems to be a natural first step in relaxing assumptions of independence.

Harris (1963) discusses briefly, for the case of binary reproduction, how a Bellman-Harris process can be modified to include dependence between sister cells. The model he discusses is a special case of the process described above.

Data of Powell (1955) suggests that in bacterial populations the life-spans of sister cells are correlated while the life-spans of mother and daughter are not. If this is true, the study of bacterial populations could be an important application of the results obtained in this chapter.

2.2 The probability space

Although we shall be primarily concerned with analytic properties of the process, we shall base the development of these properties upon an underlying probability space. The construction of this space follows that of Harris (1963), Chapter 6, and much of the notation used here is borrowed from that source.

Let us denote the collection of all finite sequences $i_1,\ldots,i_k$, where $i_1,\ldots,i_k$ and k are positive integers, by $\mathcal{J}$. For each $I \in \mathcal{J}$, let $\langle I \rangle$ represent a distinct individual whose line of descent is given by I. For example, $\langle 213 \rangle$ is the third child of $\langle 21 \rangle$, who, in turn, is the first child of $\langle 2 \rangle$. Since we are assuming that all children of an individual appear simultaneously, the ordering "first child", "second child", etc., has no real significance. The collection $\langle 1 \rangle, \langle 2 \rangle, \ldots$ is called the first generation and we shall not be concerned with ancestors of these individuals. We say naturally that two individuals of the form $\langle i_1 \ldots i_k n \rangle$ and $\langle i_1 \ldots i_k m \rangle$ are siblings. By fiat, all members of

the first generation are siblings. Similarly, if $I_1 = i_1,\ldots,i_k$ and $I_2 = i_1,\ldots,i_j$ with $j \geq k$, we say that $\langle I_2 \rangle$ is a descendent of $\langle I_1 \rangle$. We note that according to this definition, every individual is a descendent of himself.

We shall be interested in only two facts about each individual; the length of its life and the number of children it has. For each $I \in \mathcal{J}$ let ℓ_I and ν_I denote the life-span and the number of offspring of $\langle I \rangle$, respectively. The life-span ℓ_I may be any non-negative real number and ν_I may be any non-negative integer.

<u>Definition 2.1.1</u>: By a family history we mean a sequence

$$\omega = (\ell_1, \nu_1; \ell_{11}, \nu_{11}; \ell_2, \nu_2; \ldots) \tag{2.2.1}$$

where the subscripts range over all elements of $\mathcal{J}$ in some arbitrary but fixed order. The collection of all such family histories shall be denoted by Ω.

It should be pointed out that many $\omega \in \Omega$ correspond to the same realization of the growth of an actual population. This is because each ω gives information about individuals who are never born. For example $\nu_1 = 3$ means that $\langle 1 \rangle$ has three children, but ℓ_{14} represents the life-span of the <u>fourth</u> child of $\langle 1 \rangle$. It might seem more natural to define Ω so that each $\omega \in \Omega$ would contain only the following information: the life-spans of the k individuals in the first generation and the number of offspring of each, the life-spans of these offspring and the number of offspring they have, and so on. Everett and Ulam (1948) and

Otter (1949) have constructed spaces of this type, although they consider only generation sizes.

In view of these remarks, Definition 2.2.1 may seem rather artificial, but it certainly does no harm to have "finer" events in a probability space than those one is actually concerned with. Moreover, this definition makes Ω a countably infinite product of spaces that are either the non-negative real line or the non-negative integers, which is a great help in defining the σ-algebra $\mathcal{A}$ and probability measure P, and also in deriving integral equations as we shall see.

The problem discussed here is not unlike that of computing probabilities in a certain dice game. Although one is only interested in events such as "the sum of the numbers is seven", it is more convenient to define the basic space as a product space (the set of all ordered pairs of integers from one to six) whose elements correspond to "finer" events. The main difference in this simple example and our situation is that our "finer" events are not actually observable.

As mentioned before, the set Ω may be represented as an infinite product space

$$\Omega = B_1 \times B_2 \times B_3 \times \ldots \tag{2.2.2}$$

where each B_i is either the non-negative real line or the set of non-negative integers. In the former case, let the σ-algebra $\mathcal{A}_i$ on B_i be the Borel sets, and in the latter case, let the σ-algebra $\mathcal{A}_i$ on B_i be the power set of B_i. Then define the σ-algebra $\mathcal{A}$ on Ω as the

minimal σ-algebra over the class of all cylinders of the form

$$\prod_{k=1}^{N} A_k \times \prod_{k=N+1}^{\infty} B_k$$

where N is any positive integer and $A_k \in \mathcal{G}_k$, $k = 1,2,\ldots,N$.

We must now define a probability measure P on the measurable space $(\Omega,\mathcal{G})$. For each $k = 1,2,\ldots$, let $G_k(x_1,\ldots,x_k)$ be a distribution satisfying

(i) if $x_j < 0$ for some j, $1 \le j \le k$, then

$$G_k(x_1,\ldots,x_k) = 0 \tag{2.2.3}$$

(ii) the function $G_k(x_1,\ldots,x_k)$ is symmetric in its arguments; i.e.,

$$G_k(x_{i(1)},\ldots,x_{i(k)}) = G_k(x_1,\ldots,x_k) \tag{2.2.4}$$

where $i(\ )$ is any permutation of $1,\ldots,k$, and

(iii) the family of distributions is consistent[1] in the sense that

$$G_{k-1}(x_1,\ldots,x_{k-1}) = G_k(x_1,\ldots,x_{k-1},\infty), \tag{2.2.5}$$

We shall frequently denote $G_k(x_1,\ldots,x_k)$ simply by $G(x_1,\ldots,x_k)$.

---

1 Condition (iii) is weaker than the usual definition of consistency, but (ii) and (iii) taken together imply consistency as it is usually defined.

Similarly, suppose that for each k and each set of non-negative integers $j_1,\ldots,j_k$, $p_{j_1,\ldots,j_k}$ is a non-negative number satisfying

(i) $$\sum_{j_1=0}^{\infty} \cdots \sum_{j_k=0}^{\infty} p_{j_1 \ldots j_k} = 1 , \tag{2.2.6}$$

(ii) $$p_{j_{i(1)} \ldots j_{i(k)}} = p_{j_1 \ldots j_k} \tag{2.2.7}$$

where $i(\ )$ is any permutation of $1,\ldots,k$, and

(iii) $$\sum_{j_k=0}^{\infty} p_{j_1 \ldots j_k} = p_{j_1 \ldots j_{k-1}} . \tag{2.2.8}$$

Now let $S_1,\ldots,S_k,\ldots$ be the partition of $\mathcal{J}$ into equivalence classes determined by the equivalence relation "I_1 is equivalent to I_2 if and only if $\langle I_1 \rangle$ and $\langle I_2 \rangle$ are siblings". Then Ω may be written as

$$\Omega = L_{S_1} \times V_{S_1} \times L_{S_2} \times V_{S_2} \times \ldots$$

where

$$L_{S_j} = \prod_{I \in S_j} L_I$$

and

$$V_{S_j} = \prod_{I \in S_j} V_I$$

each L_I being the non-negative real line, and each V_I being the set of non-negative integers.

The functions $G_k(x_1,\ldots,x_k)$ $k = 1,2,\ldots$ form a family of consistent distribution functions which determines a probability on each L_{S_j} by Kolmogorov's fundamental theorem (Kolmogorov (1933)). Likewise, the $p_{i_1 \ldots i_k}$ determine a probability on each V_{S_j} by the same theorem. These probabilities then determine a probability P on $(\Omega,\mathcal{G})$ according to the product probability theorem (Loeve (1963), p. 91). The following facts are immediate consequences of the definition of P.

(i) For any k siblings, $G_k(x_1,\ldots,x_k)$ is the joint distribution function for their life-spans, and $p_{i_1 \ldots i_k}$ is the joint probability function for the numbers of offspring they have. It follows that for each $I \in \mathcal{J}$, the collections $\{\ell_{I1},\ldots,\ell_{Ik},\ldots\}$ and $\{\nu_{I1},\ldots,\nu_{Ik},\ldots\}$ are sequences of exchangeable random variables.

(ii) The life-spans ℓ_I are independent of the numbers of offspring ν_I.

(iii) If none of the individuals $\langle I_1'\rangle,\ldots,\langle I_k'\rangle,\ldots$ are siblings of the individuals $\langle I_1''\rangle,\ldots,\langle I_k''\rangle,\ldots$ then the collection of life-spans $\{\ell_{I_1'},\ldots,\ell_{I_k'},\ldots\ldots\}$ is independent of the collection $\{\ell_{I_1''},\ldots,\ell_{I_k''},\ldots\}$, and a corresponding statement is true for the numbers of offspring.

2.3 The branching stochastic process

We shall be primarily interested in the random function $Z^{(k)}(t)$ giving the number of individuals alive at time t in a process that begins

with k siblings born at time $t = 0$.

<u>Definition 2.3.1</u>: For each $t \geq 0$, $\omega \in \Omega$, and $I = i_1,\ldots,i_k \in \mathcal{J}$, let

$$X_I(t,\omega) = \begin{cases} 1 & \text{if } i_2 \leq \nu_{i_1,\ldots,i_k} \leq \nu_{i_1,\ldots,i_{k-1}}, \\ & \ell_{i_1} + \cdots\cdots + \ell_{i_1,\ldots,i_{k-1}} \leq t, \\ & \text{and } \ell_{i_1} + \cdots\cdots + \ell_{i_1,\ldots,i_k} > t \\ 0 & \text{otherwise} \end{cases} \tag{2.3.1}$$

and

$$Z_j(t,\omega) = \Sigma X_I(t,\omega) \tag{2.3.2}$$

where the sum is taken over all descendents of $\langle j \rangle$, and

$$Z^{(k)}(t,\omega) = \sum_{j=1}^{k} Z_j(t,\omega) \tag{2.3.3}$$

We shall often, for convenience, suppress the ω in these functions. Expression (2.3.1) simply says that $X_I(t) = 1$ if $\langle I \rangle$ is ever born and is alive at time t, and $X_I(t) = 0$, otherwise. Therefore, $Z_j(t)$ is the number of descendents of $\langle j \rangle$ alive at time t and $Z^{(k)}(t)$ is the number of individuals alive at time t who descend from $\langle 1 \rangle,\ldots,\langle k \rangle$.

<u>Theorem 2.3.1</u>: The functions $X_I(t,\omega)$, $Z_j(t,\omega)$, and $Z^{(k)}(t,\omega)$ are all measurable in t for each fixed ω and measurable in ω for each fixed

t .

Proof: The proof follows easily from Definition 2.3.1 and the fact that all of the coordinate functions ℓ_I and ν_I are measurable.

The representation of $Z^{(k)}(t)$ to be given in the following lemma will be crucial when we derive an integral equation for the generating function of $Z^{(k)}(t)$. For each $\omega = (\ell_1, \nu_1; \ell_{11}, \nu_{11}; \ell_2, \nu_2; \ldots)$ define $\omega_i = (\ell_{i1}, \nu_{i1}; \ell_{i11}, \nu_{i11}; \ell_{i2}, \nu_{i2}; \ldots)$. Let Ω_i be the set of all such ω_i and let Ω_0 be the set of all sequences of the form $(\ell_1, \nu_1; \ell_2, \nu_2; \ell_3, \nu_3; \ldots)$. Then the space Ω may be represented in the form

$$\Omega = \Omega_0 \times \Omega_1 \times \Omega_2 \times \ldots . \tag{2.3.4}$$

Lemma 2.3.1: If $\ell_1 \le t, \ldots, \ell_i \le t$, $\ell_{i+1} > t, \ldots, \ell_k > t$ and $\nu_1 > 0, \ldots, \nu_i > 0$ for some i, $1 \le i \le k$, then

$$Z^{(k)}(t,\omega) = k - i + \sum_{j=1}^{i} Z^{(\nu_j)}(t - \ell_j, \omega_j) . \tag{2.3.5}$$

Remark: The proof of this lemma is similar to that of Theorem 6.1, page 129 of Harris (1963) and will be omitted. The content of the lemma is illustrated by Figure 1 which represents the first stages of the evolution of a family that begins with three siblings in the first generation. For this particular family history, we have $k = 3$, $i = 2$, $\nu_1 = 2$, $\nu_2 = 3$, $\nu_3 = 3$, $Z^{(\nu_1)}(t - \ell_1, \omega_1) = 2$, $Z^{(\nu_2)}(t - \ell_2, \omega_2) = 4$, and $Z^{(k)}(t,\omega) = 7$. The reader may verify that these values are in agreement with (2.3.5).

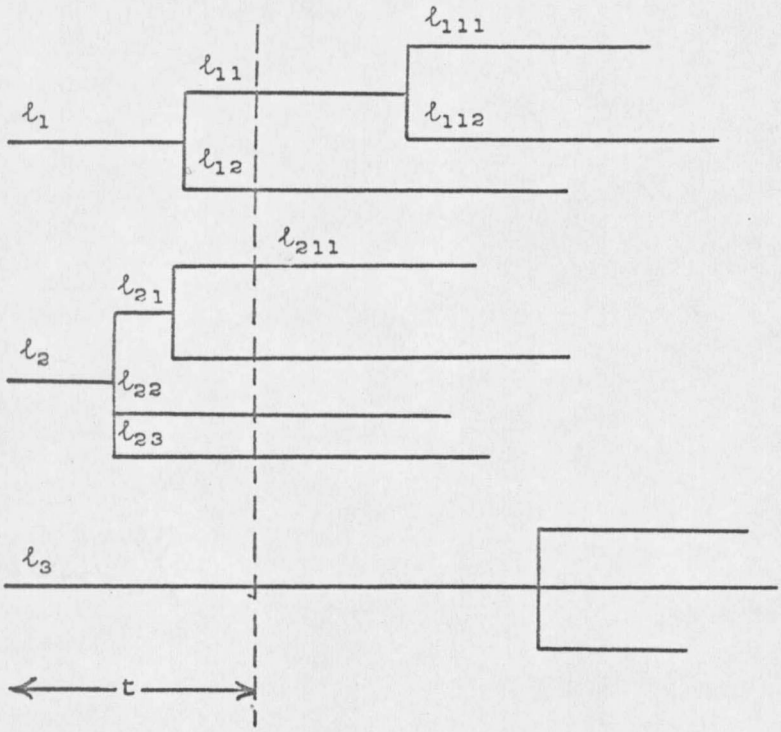

Figure 1

Early stages in the development of a family history

2.4 System of integral equations for the generating functions

<u>Definition 2.4.1</u>: For $t \geq 0$, $|s| \leq 1$, $k = 1,2,\ldots$, let

$$F(s,t,k) = \sum_{r=0}^{\infty} P[Z^{(k)}(t) = r]s^r = E\left[s^{Z^{(k)}(t)}\right]. \qquad (2.4.1)$$

Rather than derive an integral equation for $F(s,t,k)$ for each fixed k similar to (1.3.2) we shall instead derive a system of integral equations in which each equation involves all of the functions $F(s,t,k)$ $k = 1,2,\ldots$.

It follows from Definition 2.4.1 that

$$F(s,t,k) = \int_{\Omega} s^{Z^{(k)}(t)} dP = \sum \int_{A} s^{Z^{(k)}(t)} dP \tag{2.4.2}$$

where the sum is taken over all sets of the form $A = A_1 \times A_2 \times \cdots \times A_k$, and each A_i, $1 \le i \le k$, is either $C_i \equiv [\ell_i \le t]$ or C_i'. (The symbol C_i' stands for the complement of C_i.) If we let $B_i = C_1 \times \cdots \times C_i \times C_{i+1}' \times \cdots \times C_k'$, $1 \le i \le k$, and $B_0 = C_1' \times \cdots \times C_k'$, then

$$\binom{k}{i} \int_{B_i} s^{Z^{(k)}(t)} dP = \sum{}' \int_{A} s^{Z^{(k)}(t)} dP \tag{2.4.3}$$

where Σ' means the sum over all $A = \prod_{j=1}^{k} A_j$ such that exactly i of the A_j's equal C_j and the rest equal C_j'. Expression (2.4.3) is a consequence of the fact that both $G_k(x_1,\ldots,x_k)$ and $p_{i_1 \ldots i_k}$ are symmetric in their arguments. As a result of (2.4.2) and (2.4.3) we see that

$$F(s,t,k) = \sum_{i=0}^{k} \binom{k}{i} \int_{B_i} s^{Z^{(k)}(t)} dP . \tag{2.4.4}$$

and

$$\int_{B_0} s^{Z^{(k)}(t)} dP = s^k P[\ell_1 > t, \ldots, \ell_k > t] \tag{2.4.5}$$

and for $1 \le i \le k$,

$$\int_{B_i} s^{Z^{(k)}(t)} dP = \sum_{n_1=0}^{\infty} \cdots \sum_{n_i=0}^{\infty} \int_D s^{Z^{(k)}(t)} dP \tag{2.4.6}$$

where, for convenience, we have set $D = B_i \cap [\nu_1 = n_1] \cap \cdots \cap [\nu_i = n_i]$. Let us now focus our attention on just one of these summands for which $\nu_j > 0$, $1 \le j \le i$. Applying Lemma 2.3.1 we have

$$\int s^{Z^{(k)}(t)} dP = s^{k-i} \int_D s^{\sum_{j=1}^{i} Z^{(n_j)}(t - \ell_j, \omega_j)} dP . \tag{2.4.7}$$

We now interrupt the proof briefly to introduce some new notation. If $\Omega' \times \Omega'' = \Omega$ for some Ω' and Ω'' let us denote by $P_{\Omega'}$ the probability function on Ω' defined by

$$P_{\Omega'}(A) = P(A \times \Omega'') ,$$

for A any measurable subset of Ω'.

With this notation, the right side of (2.4.7) becomes

$$s^{k-i} \int_D dP_{\Omega_0} \int s^{\sum_{j=1}^{i} Z^{(n_j)}(t - \ell_j, \omega_j)} dP_{\Omega_1 \times \cdots \times \Omega_i} \tag{2.4.8}$$

by an application of Fubini's theorem (Loeve (1963), p. 136). (The

spaces Ω_i , $i = 0,1,\ldots$, are those defined just prior to Lemma 2.3.1.) Upon noting that

$$P_{\Omega_1 \times \cdots \times \Omega_i} = P_{\Omega_1} \times \cdots \times P_{\Omega_i}$$

(i.e., the probability $P_{\Omega_1 \times \cdots \times \Omega_i}$ is the same as the one determined by $P_{\Omega_1},\cdots,P_{\Omega_i}$ by the product probability theorem), another application of Fubini's theorem transforms (2.4.8) into

$$s^{k-i}\int_D dP_{\Omega_0} \sum_{j=1}^{i} \int s^{Z^{(n_j)}(t-\ell_j,\omega_j)} dP_{\Omega_j} , \tag{2.4.9}$$

and since, by the definition of P and P_{Ω_j} , the probability P_{Ω_j} on Ω is identical to the probability P on Ω, (2.4.9) is equal to

$$s^{k-i}\int_D dP_{\Omega_0} \prod_{j=1}^{i} F(s,t-\ell_j,n_j) . \tag{2.4.10}$$

Upon applying the definitions of P_{Ω_0} and the set D , (2.4.10) becomes

$$s^{k-i} p_{n_1,\ldots,n_i} \int_{u_1=0}^{t} \cdots \int_{u_i=0}^{t} \int_{u_{i+1}=t}^{\infty} \cdots \int_{u_k=t}^{\infty} \prod_{j=1}^{i} F(s,t-u_j,n_j)\, dG(u_1,\ldots,u_k) . \tag{2.4.11}$$

All of this work is also valid if $\nu_j = 0$ for some j , $1 \le j \le i$,

provided we set

$$F(s,t,0) = 1 .$$

From (2.4.2) through (2.4.11), we have proved

<u>Theorem 2.4.1</u>: The generating functions $F(s,t,k)$ satisfy the system of integral equations

$$F(s,t,k) = s^k P[\ell_i > t;\ i = 1,2,\ldots,k] + \sum_{i=1}^{k} s^{k-i} \sum_{n_1=0}^{\infty} \cdots \sum_{n_i=0}^{\infty} p_{n_1 \cdots n_i} \tag{2.4.12}$$

$$\cdot \int_{u_1=0}^{t} \cdots \int_{u_i=0}^{t} \int_{u_{i+1}=t}^{\infty} \cdots \int_{u_k=t}^{\infty} \prod_{j=1}^{i} F(s,t-u_j,n_j)\,dG(u_1,\ldots,u_k) ,$$

$k = 1,2,\ldots .$

2.5 First moments

For the remainder of this chapter we shall assume that $G_1(0) = 0$ and $m = h'(1) < \infty$ where $h(s) = \sum_j p_j s^j$.

<u>Definition 2.5.1</u>: Let

$$M(t,k) = E[Z^{(k)}(t)] .$$

For convenience, we shall denote $M(t,1)$ simply by $M(t)$. Since $E[Z_j(t)] = E[Z_i(t)]$ for all positive integers i and j , it follows that

$$M(t,k) = E\left[\sum_{j=1}^{k} Z_j(t)\right] = kE[Z_1(t)] = kM(t) . \tag{2.5.1}$$

Therefore, the properties of the mean functions $M(t,k)$ may be obtained by studying only $M(t)$.

By an argument almost identical to that of Harris (1963), page 139, the assumptions for G and m made at the beginning of this section imply that $M(t) < \infty$ for all t. Consequently, $M(t,k) < \infty$ and $Z^{(k)}(t) < \infty$ almost surely for each t and k.

<u>Theorem 2.5.1</u>: $M(t)$ satisfies the integral equation

$$M(t) = 1 - G(t) + m\int_0^t M(t-u)dG(u) . \tag{2.5.2}$$

<u>Proof</u>: If we let $k = 1$ in equation (2.4.12) we get

$$F(s,t,1) = s[1 - G(t)] + \int_0^t \sum_{n=0}^{\infty} p_n F(s,t-u,n)dG(u), \quad 0 \le s \le 1 , \tag{2.5.3}$$

the interchanging of integration and summation being permissible by the dominated convergence theorem. Since $M(t,k) = F'(1,t,k)$ where

$$F'(s,t,k) = \frac{\partial}{\partial s}F(s,t,k)$$

the proof will follow by differentiating both sides of (2.5.3) with respect to s, setting $s = 1$, and applying (2.5.1). We note that

$$p_n F'(s,t-u,n) \le p_n M(t-u,n) = np_n M(t-u) , \quad 0 \le s \le 1 .$$

Therefore, the differentiated series $\Sigma np_n F'(s,t-u,n)$ converges uniformly $0 \le s \le 1$, and the operations of summation and differentiation may be interchanged. The operations of differentiation and integration may be interchanged by a corollary of the dominated convergence theorem (Loeve (1963), page 126).

The only other observation that really needs to be made about the mean functions of the process is that equation (2.5.2) is identical to (1.3.3) given in Chapter 1 for the mean of the Bellman-Harris process which implies that the particular type of dependence assumed in this chapter has no effect upon the mean. In particular, Lemma 1.3.1 stated for the mean of the Bellman-Harris process holds for the function $M(t)$ being studied in this section also.

For future reference, we now state a result that follows in a trivial fashion from the preceeding remarks, Lemma 1.3.1, and expression (2.5.1).

<u>Corollary</u>: If $m > 1$, and G is not a lattice distribution, then

$$M(t,k)e^{-\alpha t} \to bk, \quad k = 1,2,\ldots, \quad \text{as} \quad t \to \infty, \tag{2.5.4}$$

where

$$b = \frac{\int_0^\infty e^{-\alpha u}[1 - G(u)]du}{m\int_0^\infty ue^{-\alpha u}dG(u)}$$

and α is defined in Lemma 1.3.1.

Before concluding our discussion of first moments, it should be pointed out that Mode (1967) studied a binary branching process with correlation among siblings and obtained a rate of growth of the form $t^{-\frac{1}{2}}e^{\alpha t}$. However, he based his results on an integral equation that is vastly different from (2.4.12) and it is not clear how his process fits in with the one being studied here.

2.6 Second moments

<u>Definition 2.6.1</u>: Let

$$M_2(t,\tau,k) = E[Z^{(k)}(t)Z^{(k)}(t+\tau)] . \tag{2.6.1}$$

At this juncture, we could (most likely) derive an integral equation for the joint generating function of $Z^{(k)}(t)$ and $Z^{(k)}(t+\tau)$ similar to (2.4.12) and then differentiate this expression to obtain an integral equation for $M_2(t,\tau,k)$. This approach is taken by Harris (1963), Ney (1964b), and Mode (1968b). However, because of the complexity of the resulting expressions, in the present situation it seems easier to derive an equation for $M(t,\tau,k)$ directly. This procedure also circumvents delicate arguments concerning interchanging the order of differentiation and other limiting operations as were needed in the proof of Theorem 2.5.1.

First of all, we have

$$M_2(t,\tau,k) = E\left[\sum_{j=1}^{k} Z_j(t) \sum_{j=1}^{k} Z_j(t+\tau)\right]$$

$$= E\left[\sum_{j=1}^{k} Z_j(t)Z_j(t+\tau)\right] + E\left[\sum_{i\neq j} Z_i(t)Z_j(t+\tau)\right] \quad (2.6.2)$$

$$= kE[Z_1(t)Z_1(t+\tau)] + k(k-1)E[Z_1(t)Z_2(t+\tau)] .$$

The last equality is a result of both $G(x_1,\ldots,x_k)$ and $p_{i_1 \cdots i_k}$ being symmetric in their arguments. If we let

$$M(t,\tau) = E[Z_1(t)Z_1(t+\tau)] \quad (2.6.3)$$

and

$$M^*(t,\tau) = E[Z_1(t)Z_2(t+\tau)] , \quad (2.6.4)$$

equation (2.6.2) becomes

$$M_2(t,\tau,k) = kM(t,\tau) + k(k-1)M^*(t,\tau) . \quad (2.6.5)$$

In view of (2.6.5), the next step is to derive integral representations for $M(t,\tau)$ and $M^*(t,\tau)$. Most of the ideas involved have already been applied in deriving (2.4.12). Therefore, without presenting any of the details, the end results are

$$M(t,\tau) = 1 - G(t+\tau) + m\int_{t+}^{t+\tau} M(t+\tau-u)dG(u)$$

$$+ \sum_{n=1}^{\infty} p_n \int_0^t M_2(t-u,\tau,n)dG(u) , \tag{2.6.6}$$

and

$$M^*(t,\tau) = 1 + G(t,t+\tau) - G(t) - G(t+\tau) + m\int_{u=t+}^{\infty}\int_{v=0}^{t+\tau} M(t+\tau-v)dG(u,v)$$

$$+ m\int_{u=0}^{t}\int_{v=t+\tau+}^{\infty} M(t-u)dG(u,v) + m_2\int_{u=0}^{t}\int_{v=0}^{t+\tau} M(t-u)M(t+\tau-v)dG(u,v) \tag{2.6.7}$$

where

$$m_2 = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} j\,k\,p_{jk} . \tag{2.6.8}$$

The reader should observe that (2.6.7) expresses $M^*(t,\tau)$ in terms of $M(t)$ and known quantities, but the right side of (2.6.6) contains all of the functions $M_2(t,\tau,k)$, $k = 1,2,\ldots$. However, if we substitute (2.6.5) into (2.6.6), we get

$$M(t,\tau) = 1 - G(t+\tau) + m\int_{t+}^{t+\tau} M(t+\tau-u)dG(u)$$

$$+ m\int_0^t M(t-u,\tau)dG(u) + h''(1)\int_0^t M^*(t-u,\tau)dG(u) , \tag{2.6.9}$$

which is the desired integral equation for $M(t,\tau)$. It now appears that we may obtain an asymptotic formula for $M(t,\tau,k)$, by first using (2.6.7) to get an asymptotic formula for $M^*(t,\tau)$, then using this result together with (2.6.9) to obtain an asymptotic formula for $M(t,\tau)$, and finally applying these results to (2.6.5). This we shall now do.

<u>Lemma 2.6.1</u>: If $m > 1$ and $h''(1) < \infty$ and G is not a lattice distribution, then

$$M^*(t,\tau)e^{-\alpha(2t+\tau)} \to m_2 b^2 \int_0^\infty\int_0^\infty e^{-\alpha(u+v)}dG(u,v) < \infty \quad \text{as } t \to \infty \qquad (2.6.10)$$

and this limit is uniform in τ.

<u>Proof</u>: Let us first observe that by the Schwartz inequality

$$m_2 \le \sum_{j=0}^{\infty} j^2 p_j \sum_{k=0}^{\infty} k^2 p_k = h''(1) + m < \infty . \qquad (2.6.11)$$

It is fairly evident that when the right side of (2.6.7) is multiplied by $e^{-\alpha(2t+\tau)}$ and the limit as t approaches infinity is taken, all terms except the last approach zero uniformly in τ. Therefore, the proof will be complete upon demonstrating that for any given $\epsilon > 0$ there exists an x such that for any $t \ge x$ and any τ, the expression

$$\left|\int_0^t\int_0^{t+\tau} M(t-u)M(t+\tau-v)e^{-\alpha(2t+\tau)}dG(u,v) - b^2\int_0^\infty\int_0^\infty e^{-\alpha(u+v)}dG(u,v)\right| \qquad (2.6.12)$$

is less than ϵ. Since $M(t)$ is bounded on every finite interval by Lemma 1.3.1, and $M(t)e^{-\alpha t}$ approaches a finite constant as t becomes large, it follows that $M(t)e^{-\alpha t}$ is bounded on the entire positive real line, say by B. From (2.5.4) there exists an x_1 such that

$$|M(t)e^{-\alpha t}M(t+\tau)e^{-\alpha(t+\tau)} - b^2| < \epsilon/3 \tag{2.6.13}$$

for all $t > x_1$. Furthermore, there exists an x_2 such that

$$1 - G(t-x_1, t-x_1) < \epsilon/3(B^2+2b^2) \tag{2.6.14}$$

for $t > x_2$. On the other hand, for all t and τ expression (2.6.12) is less than or equal to

$$\left|\int_0^t\int_0^{t+\tau}\left[M(t-u)M(t+\tau-v)e^{-\alpha(t-u)}e^{-\alpha(t+t-v)} - b^2\right]e^{-\alpha(u+v)}dG(u,v)\right|$$

$$+ b^2\left|\int_0^\infty\int_0^\infty e^{-\alpha(u+v)}dG(u,v) - \int_0^t\int_0^{t+\tau} e^{-\alpha(u+v)}dG(u,v)\right|$$

$$\leq \int_0^{t-x_1}\int_0^{t+\tau-x_1}|M(t-u)e^{-\alpha(t-u)}M(t+\tau-v)e^{-\alpha(t+\tau-v)} - b^2|e^{-\alpha(u+v)}dG(u,v)$$

$$+ \int_A [B^2+2b^2]dG(u,v)$$

where $A = \{[0,\infty) \times [0,\infty)\} - [0,t-x_1] \times [0,t+\tau-x_1]\}$. The proof is now completed by letting $x = \max(x_1,x_2)$ and applying the estimates

(2.6.13) and (2.6.14) to this last inequality.

Lemma 2.6.2: If $m > 1$, $h''(1) < \infty$, and G is not a lattice distribution, then

$$M(t,\tau)e^{-\alpha(2t+\tau)} \to \frac{h''(1)m_2 b^2\left[\int_0^\infty\int_0^\infty e^{-\alpha(u+v)}dG(u,v)\right]\left[\int_0^\infty e^{-2\alpha u}dG(u)\right]}{1 - m\int_0^\infty e^{-2\alpha u}dG(u)} \tag{2.6.15}$$

as $t \to \infty$ and this limit is uniform in τ.

Proof: If we multiply both sides of (2.6.9) by $e^{-\alpha(2t+\tau)}$ and put

$$\overline{M}(t,\tau) = M(t,\tau)e^{-\alpha(2t+\tau)},$$

$$\overline{m} = m\int_0^\infty e^{-2\alpha u}dG(u), \text{ and}$$

$$\overline{G}(t) = \frac{\int_0^t e^{-2\alpha u}dG(u)}{\int_0^\infty e^{-2\alpha u}dG(u)},$$

then (2.6.9) may be written in the form

$$\overline{M}(t,\tau) = \overline{f}(t,\tau) + \overline{m}\int_0^t \overline{M}(t-u,\tau)d\overline{G}(u).$$

Moreover, using Lemma 2.6.1 and an argument similar to the proof of

Lemma 2.6.1, it can be shown that

$$\lim_{t\to\infty} \bar{f}(t,\tau) = b^2 h''(1) m_2 \int_0^\infty \int_0^\infty e^{-\alpha(u+v)} dG(u,v) \cdot \int_0^\infty e^{-2\alpha u} dG(u) ,$$

and this limit is uniform in τ. The proof now follows from Lemma 4, page 163 of Harris (1963) in the same way as Theorem 18.1 of Harris.

We are now ready to obtain the result towards which we have been working, namely, an asymptotic formula for $M_2(t,\tau,k)$. For the sake of brevity, let us set

$$c = \int_0^\infty e^{-2\alpha u} dG(u) \tag{2.6.16}$$

and

$$c' = \int_0^\infty \int_0^\infty e^{-\alpha(u+v)} dG(u,v) . \tag{2.6.17}$$

<u>Theorem 2.6.1</u>: If $m > 1$, $h''(1) < \infty$, and G is not a lattice distribution, then for each $k = 1,2,\ldots,$

$$M(t,\tau,k)e^{-\alpha(2t+\tau)} \to k m_2 b^2 c' \left[\frac{h''(1)c}{1-mc} + k - 1\right] \quad \text{as} \quad t \to \infty \tag{2.6.18}$$

and this limit is uniform in τ.

<u>Proof</u>: The proof follows immediately upon applying Lemmas 2.6.1 and 2.6.2 to equation (2.6.5).

We now get as a special case the corresponding statement for the Bellman-Harris process in which individuals live and reproduce independently.

Corollary: If $m_2 = m^2$ and $G(u,v) = G(u)G(v)$, and conditions of Theorem 2.6.1 hold, then

$$M(t,\tau)e^{-\alpha(2t+\tau)} \to \frac{b^2h''(1)c}{1-mc} \quad \text{as} \quad t \to \infty. \tag{2.6.19}$$

Remark: It should be pointed out that the only difference in the limiting expression for $M(t,\tau)$ in Lemma 2.6.2 and the one obtained for the Bellman-Harris process is the factor m_2c' appearing in (2.6.15). As we shall see in the proof of Theorem 2.7.1, under the hypothesis of Theorem 2.6.1, $m_2c' \geq 1$. It follows that under this hypothesis the limiting variance is greater for the process with correlation among siblings than for the Bellman-Harris process. This observation agrees with a remark made by Harris (1963) in his discussion of correlation among sister cells.

2.7 Limit random variables

In this section we shall investigate the convergence of suitably normed versions of $Z^{(k)}(t)$ to random variables $W^{(k)}$ as $t \to \infty$. We shall also study the convergence of the sequence $W^{(k)}$ as $k \to \infty$.

Definition 2.7.1: Let

$$W^{(k)}(t) = Z^{(k)}(t)/kbe^{\alpha t} \tag{2.7.1}$$

and

$$W_k(t) = Z_k(t)/be^{\alpha t}. \tag{2.7.2}$$

<u>Theorem 2.7.1</u>: If $m > 1$, $h''(1) < \infty$ and $G(t)$ is not a lattice distribution, then $W^{(k)}(t)$ converges in mean square to a random variable $W^{(k)}$ as $t \to \infty$. Furthermore,

$$E(W^{(k)}) = 1 \tag{2.7.3}$$

and

$$\mathrm{Var}[W^{(k)}] = \frac{m_2 c'}{k}\left[\frac{h''(1)c}{1-mc} + k - 1\right] - 1 > 0. \tag{2.7.4}$$

<u>Proof</u>: Consider the simple expansion

$$E[W^{(k)}(t) - W^{(k)}(t+\tau)]^2 = E[W^{(k)}(t)]^2 + E[W^{(k)}(t+\tau)]^2 - 2E[W^{(k)}(t)W^{(k)}(t+\tau)].$$

If we let $t \to \infty$ and apply Theorem 2.6.1 we get

$$\lim_{t\to\infty} E[W^{(k)}(t) - W^{(k)}(t+\tau)]^2 = 0$$

and this limit is uniform for $\tau \geq 0$. Therefore, $W^{(k)}(t)$ is a Cauchy sequence in mean square and it follows that there exists a random variable $W^{(k)}$ such that $W^{(k)}(t) \to W^{(k)}$ in mean square as $t \to \infty$ by the L_r convergence theorem (Loeve (1963), page 161). The values for the mean and

variance follow from the facts that

$$E[W^{(k)}] = \lim_{t\to\infty} E[W^{(k)}(t)]$$

(Loeve (1963), page 161) and

$$E[(W^{(k)})^2] = \lim_{t\to\infty} E[(W^{(k)}(t))^2]$$

(Loeve (1963), page 157). To see that the variance is positive, note that

$$\lim_{k\to\infty} \mathrm{Var}[W^{(k)}] = m_2 c' - 1 \geq 0$$

since variances are non-negative. Therefore

$$\mathrm{Var}[W^{(k)}] \geq \frac{1}{k}\left[\frac{h''(1)c}{1-mc} + k - 1\right] - 1 = \frac{1}{k}\left[\frac{h''(1)c + mc - 1}{1 - mc}\right].$$

This last expression is non-negative according to Harris (1963). Moreover, equality occurs only when G is a step function with one step, which cannot be since G is not a lattice distribution. This completes the proof.

<u>Remark</u>: The mean square convergence proved in Theorem 2.7.1 can be strengthened to almost sure convergence is one assumes in addition that G has a density g and

$$\int_0^\infty (g(t))^p dt < \infty$$

for some $p > 1$. The proof of this result is almost identical to the proof of the corollary on page 147 in Harris (1963).

The following lemma will allow us to prove that $W^{(k)}$ converges in mean square to a random variable W. It should be noted that the lemma is applicable to sequences of exchangeable random variables.

<u>Lemma 2.7.1</u>: If $\{X_k\}$, $k = 1,2,\ldots$, is a sequence of random variables with common means and variances (finite) and also having the property that $E(X_i X_j) = E(X_1 X_2)$ for $i \neq j$, then $1/k \sum_{j=1}^{k} X_j$ converges in mean square to some random variable X and $Var(X) = Cov(X_1, X_2)$.

<u>Proof</u>: For $k' < k$ we have

$$E\left[\frac{1}{k}\sum_{j=1}^{k} X_j - \frac{1}{k'}\sum_{j=1}^{k'} X_j\right]^2 = \frac{k-k'}{kk'}[E(X_1^2) - E(X_1 X_2)]$$

$$\leq \frac{1}{k'}[E(X_1^2) - E(X_1 X_2)].$$

It is now obvious that $\frac{1}{k}\sum_{j=1}^{k} X_j$ is a Cauchy sequence in mean square and therefore converges to a random variable X by the L_r convergence theorem. Furthermore,

$$E(X^2) = \lim_{k\to\infty} E\left[\frac{1}{k}\sum_{j=1}^{k} X_j\right]^2$$

$$= \lim_{k\to\infty}\left[\frac{1}{k}E(X_1^2) + \frac{k(k-1)}{k^2}E(X_1X_2)\right] = E(X_1X_2) .$$

Therefore

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = E(X_1X_2) - E(X_1)E(X_2) = \mathrm{Cov}(X_1,X_2) .$$

<u>Corollary</u>: If X_k is a sequence of random variables satisfying the conditions of Lemma 2.7.1, then it is impossible for $\mathrm{Cov}(X_1,X_2)$ to be negative.

<u>Proof</u>: From Lemma 2.7.1 we have $\mathrm{Cov}(X_1,X_2) = \mathrm{Var}(X)$ and variances are non-negative.

<u>Remark</u>: The result of this corollary is well-known, but the proof usually given uses the fact that the matrix of covariances of $X_1,\cdots,X_k$ must be positive-definite.

<u>Theorem 2.7.2</u>: Under the conditions of Theorem 2.7.1 the random variables $W^{(k)}$ converge in mean square to a random variable W as $k \to \infty$ and

$$\mathrm{Var}(W) = m_2c' - 1 .$$

<u>Proof</u>: As in the proof of Theorem 2.7.1 it can be shown that the random variables $W_k(t)$ converge in mean square to a random variable W_k as $t \to \infty$ and, furthermore

$$W^{(k)} = \frac{1}{k}\sum_{j=1}^{k} W_j \quad \text{a.s.}.$$

Lemma 2.7.1 is applicable to the sequence W_k and this proves the first part of the theorem. The formula for the variance of W follows from the fact that

$$\mathrm{Var}(W) = \lim_{k\to\infty}\mathrm{Var}(W^{(k)}) .$$

In the remaining sections of this chapter we shall study the distribution of $W^{(k)}$.

2.8 Integral equations for the characteristic functions

<u>Definition 2.8.1</u>: Let $\varphi(s,t,k)$ be the characteristic function of $kW^{(k)}(t)$ and $\varphi(s,k)$ the characteristic function of $kW^{(k)}$, i.e., let

$$\varphi(s,t,k) = E[\exp(iskW^{(k)}(t))] \tag{2.8.1}$$

and

$$\varphi(s,k) = E[\exp(iskW^{(k)})] . \tag{2.8.2}$$

From this definition and the definition of $F(s,t,k)$, it follows that

$$\varphi(s,t,k) = F(\exp(is/be^{\alpha t}),t,k)$$

and, therefore, we can use equation (2.4.12) to obtain a system of integral

equations for $\varphi(s,t,k)$. Making the necessary substitutions, we obtain

$$\varphi(s,t,k) = \exp(is/be^{\alpha t})p[\ell_j > t; j=1,2,\ldots,k] + \sum_{j=1}^{k}\binom{k}{j}\exp[is(k-j)/be^{\alpha t}]$$

$$\cdot \sum_{n_1=0}^{\infty}\cdots\sum_{n_j=0}^{\infty} p_{n_1\cdots n_j}\int_{u_1=0}^{t}\cdots\int_{u_j=0}^{t}\int_{u_{j+1}=t}^{\infty}\cdots\int_{u_k=t}^{\infty} \tag{2.8.3}$$

$$\cdot \prod_{r=1}^{j}\varphi(se^{-\alpha u_r}, t-u_r, n_r)dG(u_1,\ldots,u_k).$$

<u>Theorem 2.8.1</u>: If $m > 1$, $h''(1) < \infty$, and G is not a lattice distribution, then the functions $\varphi(s,k)$, $k = 1,2,\ldots$, satisfy the equation

$$\varphi(s,k) = \sum_{n_1=0}^{\infty}\cdots\sum_{n_k=0}^{\infty} p_{n_1\cdots n_k}\int_0^{\infty}\cdots\int_0^{\infty}\prod_{j=1}^{k}\varphi(se^{-\alpha u_j}, n_j)dG(u_1,\ldots,u_k). \tag{2.8.4}$$

<u>Proof</u>: By Theorem 2.7.1 $W^{(k)}(t) \to W^{(k)}$ in mean square as $t \to \infty$ which implies $\varphi(s,t,k) \to \varphi(s,k)$ as $t \to \infty$. The proof will be complete therefore when it is shown that the right side of (2.8.3) converges to the right side of (2.8.4) as $t \to \infty$. For $j \neq k$ the j-th term in the sum $\sum_{i=1}^{k}$ appearing in (2.8.3) is less than or equal to $\binom{k}{i}[1-G(t)]$ which approaches zero as t approaches infinity. Likewise, the first term on the right side of (2.8.3) approaches zero as $t \to \infty$. The theorem will be proved then if we can show that

$$\sum_{n_1=0}^{\infty} \cdots \sum_{n_k=0}^{\infty} p_{n_1 \cdots n_k} \int_0^t \cdots \int_0^t \prod_{j=1}^{k} \varphi(se^{-\alpha u_j}, t-u_j, n_j) dG(u_1, \cdots, u_k) \to$$

$$\sum_{n_1=0}^{\infty} \cdots \sum_{n_k=0}^{\infty} p_{n_1 \cdots n_k} \int_0^{\infty} \cdots \int_0^{\infty} \prod_{j=1}^{k} \varphi(se^{-\alpha u_j}, n_j) dG(u_1, \cdots, u_k) \tag{2.8.5}$$

as $t \to \infty$. Let us choose an integer q and a number T_1 such that

$$1 - \Sigma' p_{n_1 \cdots n_k} < \epsilon/8 ,$$

where Σ' denotes $\sum_{n_1=0}^{q} \cdots \sum_{n_k=0}^{q}$, and

$$\int_{A(t)} dG(u_1, \cdots, u_k) < \epsilon/16$$

for any $t \geq T_1$, where $A(t) = \{[0,t] \times \cdots \times [0,t]\}'$. (There are k terms in the cross product.) Since $\varphi(s,t,n_j) \to \varphi(s,n_j)$ uniformly on any finite s-interval as $t \to \infty$ (Loeve (1963), page 191), for an arbitrary $\delta > 0$ and positive integer n_j, there exists a number $T(n_j)$ such that

$$\sup_{\substack{t \geq T \\ |r| \leq |s|}} \{|\varphi(r,t,n_j) - \varphi(r,n_j)|\} < \delta$$

for any $T \geq T(n_j)$. Therefore, it follows that for each set of positive

integers $n_1,\cdots,n_k$, there exists a number $T(n_1,\cdots,n_k)$ such that

$$\sup_{\substack{t\geq T\\ |r|\leq|s|}} \{|\prod_{j=1}^{k}\varphi(r,t,n_j) - \prod_{j=1}^{k}\varphi(r,n_j)|\} < \epsilon/2 .$$

Now let $T_2 = \max\{T(n_1,\cdots,n_k);\ 1\leq n_i \leq q,\ i=1,2,\cdots k\}$ and $T_3 = \max\{T_1,T_2\}$. Then for any $t \geq 2T_3$ the absolute value of the difference of the two terms in (2.8.5) is less than or equal to

$$\Sigma' p_{n_1\cdots n_k}\int_0^{t-T_3}\cdots\int_0^{t-T_3}|\prod_{j=1}^{k}\varphi(se^{-\alpha u_j},t-u_j,n_j) - \prod_{j=1}^{k}\varphi(se^{-\alpha u_j},n_j)|dG(u_1,\cdots,u_k)$$

$$+ 4\Sigma' p_{n_1\cdots n_k}\int\cdots\int_{A(t-T_3)} dG(u_1,\cdots,u_k) + 2[1-\Sigma' p_{n_1\cdots n_k}]$$

$$\leq \epsilon/2 + 4(\epsilon/16) + 2(\epsilon/8) = \epsilon .$$

This completes the proof.

2.9 Distribution of W for the binary case

In this section we will consider the special case where $1 = p_2 = p_{22} = p_{222} = \cdots$, which means that each individual has exactly two offspring. This case is important in its own right since it is applicable to a colony of bacteria which reproduce by splitting. Bellman and Harris (1952) studied the binary case of the Bellman-Harris process and our goal is to extend their results to the present model. When

possible we shall reduce the problems encountered in this section to ones already dealt with by Bellman and Harris.

Throughout this section we shall assume that $m > 1$, $h''(1) < \infty$, and that G is not a lattice distribution.

For convenience, let us denote the characteristic functions $\varphi(s,t;2)$ and $\varphi(s,2)$ by simply $\varphi(s,t)$ and $\varphi(s)$ respectively. For the binary situation expression (2.8.3) then becomes, for $k = 2$,

$$\varphi(s,t) = [\exp(is/be^{\alpha t})][1 + G(t,t) - 2G(t)]$$

$$+ 2[\exp(is/be^{\alpha t})]\int_{u=0}^{t}\int_{v=t}^{\infty}\varphi(se^{-\alpha u}, t-u)dG(u,v) \tag{2.9.1}$$

$$+ \int_0^t\int_0^t \varphi(se^{-\alpha u}, t-u)\varphi(se^{-\alpha v}, t-v)dG(u,v)$$

and Theorem 2.8.1 becomes

<u>Corollary</u>: The function $\varphi(s)$ satisfies the integral equation

$$\varphi(s) = \int_0^\infty\int_0^\infty \varphi(se^{-\alpha u})\varphi(se^{-\alpha v})dG(u,v) . \tag{2.9.2}$$

For this binary case let us denote the limit random variable $W^{(2)}$ by simply W and let $K(x)$ denote its distribution function. The main result of this section is that if $1 - G(t)$ is of exponential order as $t \to \infty$, then $K(x)$ is absolutely continuous. We shall first prove three

lemmas giving information about $\varphi(s)$ for large s.

<u>Lemma 2.9.1</u>: We have $|\varphi(s)| \to 0$ as $s \to \pm\infty$.

<u>Proof</u>: From Loeve (1963), page 199,

$$\varphi(s) = 1 + 2isW - \frac{s^2}{2}E(4W^2) + o(|s|^2) \tag{2.9.3}$$

where $o(|s|^2)$ is a term such that $o(|s|^2)/|s|^2 \to 0$ as $s \to 0$. If we square both sides and take absolute values, (2.9.3) becomes

$$|\varphi^2(s)| = 1 + 4s^2[1 - E(W^2)] + o(|s|)^2 \tag{2.9.4}$$

Since the variance of W is strictly positive from Theorem 2.7.1, it follows that $|\varphi(s)| < 1$ for $|s|$ small enough $(s \neq 0)$. We shall now show that $\operatorname{Limsup}_{s\to\infty}|\varphi(s)| < 1$. Suppose otherwise. Let s_3 be a positive number such that $|\varphi(s)| < 1$ for $0 < s \leq s_3$ and $|\varphi(s_3)| < 1 - d$ for some d, $0 < d < \frac{1}{2}$. By the continuity of $|\varphi(s)|$, let s_1 and s_2 be the first points to the left and right, respectively, of s_3 such that $|\varphi(s_1)| = |\varphi(s_2)| = 1 - d$, and let $B = (1/\alpha)\operatorname{Log}(s_2/s_1)$. Then, using (2.9.2),

$$\begin{aligned}\varphi(s_2) &= \int_0^B \int_0^B \varphi(s_2 e^{-\alpha u})\varphi(s_2 e^{-\alpha v})dG(u,v) \\ &+ \int\int_{A(B)} \varphi(s_2 e^{-\alpha u})\varphi(s_2 e^{-\alpha v})dG(u,v)\end{aligned} \tag{2.9.5}$$

where $A(B) = \{[0,B] \times [0,B]\}'$. Taking absolute values and remembering that $|\varphi(s)| \le 1$, we get

$$1 - d = |\varphi(s_2)| \le (1-d)^2 G(B,B) + 1 - G(B,B)$$

and, therefore,

$$(2-d)G(B,B) \le 1 . \tag{2.9.6}$$

Now if $d \to 0$ then $s_1 \to 0$ while s_2 increases and therefore $B \to \infty$, so that (2.9.6) cannot continue to hold. Therefore, $|\varphi(s)| < 1 - d$ for all s larger than some s_0.

For an arbitrary $\epsilon > 0$ choose C so large that $1 - G(C,C) < \epsilon$ and s so large that $se^{-\alpha C} \ge s_0$. Then

$$|\varphi(s)| \le \int_0^C \int_0^C |\varphi(se^{-\alpha u})\varphi(se^{-\alpha v})| dG(u,v) + \epsilon$$

$$\le (1-d)\int_0^C \int_0^C |\varphi(se^{-\alpha u})| dG(u,v) + \epsilon$$

and, letting

$$\psi(s) = \sup_{t \ge s} |\varphi(t)| ,$$

$$\psi(s) \le (1-d)\psi(se^{-\alpha C})G(C,C) + \epsilon ,$$

or

$$\psi(se^{aC}) \le (1-d)\psi(s) + \epsilon . \tag{2.9.7}$$

Using (2.9.7) we see by induction that

$$\psi(se^{k\alpha C}) \le (1-d)^k\psi(s) + \epsilon/(1-d)$$

for $k = 1,2,\ldots$, and, therefore, $\psi(s) \to 0$ as $s \to \infty$. A similar argument shows that $|\varphi(s)| \to 0$ as $s \to -\infty$.

<u>Lemma 2.9.2</u>: If $1 - G(x) = 0(e^{-cx})$ for some $c > 0$ (see Lukacs (1960), page 207), then $|\varphi(s)| = 0(|s|^{-d})$ for some $d > 0$.

<u>Proof</u>: Since $G(x,y)$ is symmetric in its arguments,

$$1 - G(x) \le 1 - G(x,x) \le 2[1 - G(x)]$$

and therefore $1 - G(x,x) = 0(e^{-cx})$, also. If we let $B = (\frac{1}{2}\alpha)\mathrm{Log}(s)$, then

$$|\varphi(s)| \le \int_0^B \int_0^B |\varphi(se^{-\alpha u})||\varphi(se^{-\alpha v})|dG(u,v) .$$

Again letting

$$\psi(s) = \sup_{t\ge s}|\varphi(t)|$$

Filmed as received

without page(s) 48.

UNIVERSITY MICROFILMS.

$$|\varphi'(s)| \le 2\int_0^T \int_0^T \frac{A}{(1+|s|e^{-\alpha u})} d\,|\varphi'(se^{-\alpha v})|\,e^{-\alpha v} dG(u,v)$$

$$\le 2\int_0^T \int_0^T \frac{A}{(1+|s|e^{-\alpha T})} d\,|\varphi'(se^{-\alpha v})|\,e^{-\alpha v} dG(u,v) .$$

Therefore

$$B(T) \equiv \int_0^T |\varphi'(s)|\,ds$$

$$\le 2\int_0^T \int_0^T \left[\int_0^T \frac{A}{(1+se^{-\alpha T})} d\,|\varphi'(se^{-\alpha v})|\,ds\right] e^{-\alpha v} dG(u,v) .$$

Making the substitution $t = se^{-\alpha v}$, we get

$$B(T) \le 2\int_0^T \int_0^T \left[\int_0^{Te^{-\alpha v}} \frac{A}{(1+se^{-\alpha(T-v)})} d\,|\varphi'(t)|\,dt\right] dG(u,v)$$

$$\le 2\int_0^T \int_0^T \left[\int_0^T \frac{A}{(1+te^{-\alpha T})} d\,|\varphi'(t)|\,dt\right] dG(u,v)$$

$$\le 2A\int_0^T \frac{|\varphi'(t)|}{(1+Dt)} d\,dt ,$$

where $D = e^{-\alpha T}$.

Integrating this last expression by parts, we get

$$2A(1+DT)^{-d}B(T) + 2dAD\int_0^T \frac{B(t)}{(1+Dt)^{d+1}}\,dt$$

$$\leq \frac{B(T)}{2} + 2dAD\int_0^T \frac{B(t)}{(1+Dt)^{d+1}}\,dt\,,$$

provided T is so large that $2A(1+DT)^{-d} \leq \frac{1}{2}$. Therefore

$$\frac{B(T)}{(1+DT)^{d+1}} \leq \frac{4dAD}{(1+DT)^{d+1}}\int_0^T \frac{B(t)}{(1+Dt)^{d+1}}\,dt. \tag{2.9.9}$$

If

$$V(T) = \int_0^T \frac{B(t)}{(1+Dt)^{d+1}}\,dt\,,$$

it follows from (2.9.9) that

$$V'(T) \leq \frac{4dADV(T)}{(1+DT)^{d+1}} \tag{2.9.10}$$

for all T greater than or equal to some constant C. Dividing both sides of (2.9.10) by $V(t)$, we get

$$\int_C^T \frac{V'(t)}{V(t)}\,dt \leq \int_C^T \frac{4dAD}{(1+Dt)^{d+1}}\,dt$$

and therefore

$$\log V(T) \le \log C + \frac{4A}{(1+DC)^d}$$

which implies that $\lim_{t\to\infty} \log[V(T)] < \infty$. Then from (2.9.8) it follows that $\lim_{T\to\infty} B(T) < \infty$ and the proof is complete.

We are now ready to prove the principle result of this section.

<u>Theorem 2.9.1</u>: The distribution function $K(x)$ of the random variable W is continuous. If, in addition, $1 - G(x) = 0(e^{-cx})$ for some $c > 0$, then $K(x)$ is absolutely continuous.

<u>Proof</u>: The first part of the theorem follows directly from the fact that $|\varphi(s)| \to 0$ as $s \to \pm\infty$ (Lukacs (1960), page 27). The second part follows from the facts that $|\varphi(s)| \to 0$ as $s \to \pm\infty$ and $|\varphi'(s)|$ is integrable, by a fairly well-known argument. (See, for example, Ney (1961) and Stigum (1966).) However, for the sake of completeness, we shall include it here as well.

Let us define the function

$$g_T(x) = \frac{1}{2\pi}\int_{-T}^{T} e^{-itx}\varphi(t)dt\,, \quad x \ge 0\,, \quad T = 1,2,\cdots.$$

Integrating $g_T(x)$ by parts, we get

$$g_T(x) = \frac{1}{2\pi ix}\left[\varphi(-T)e^{iTx} - \varphi(T)e^{-iTx} + \int_{-T}^{T} e^{-iTx}\varphi'(t)dt\right].$$

It follows that for $T_1 < T_2$ and $0 < x_1 \le x < \infty$,

$$|g_{T_1}(x) - g_{T_2}(x)| \le \frac{1}{2\pi x_1}\Big[|\varphi(T_1)| + |\varphi(T_2)| + |\varphi(-T_1)| + |\varphi(-T_2)| + \int_{-T_2}^{-T_1}|\varphi'(t)|dt + \int_{T_1}^{T_2}|\varphi'(t)|dt\Big]. \tag{2.9.11}$$

Using the facts about $\varphi(t)$ and $\varphi'(t)$ given in Lemmas 2.9.1 and 2.9.3, expression (2.9.11) implies that for each $x > 0$ $g_T(x)$ is a Cauchy sequence in T and therefore converges to some function $g(x)$. Moreover, since the right side of (2.9.10) does not involve x, the convergence is uniform on $[x_1,\infty)$. Now Levy's inversion formula for characteristic functions states that

$$K(x) - K(x_1) = \lim_{T\to\infty}\int_{-T}^{T}\frac{e^{-itx_2}-e^{-itx_1}}{-2\pi it}\varphi(t)dt .$$

But

$$\int_{x_1}^{x} g_T(y)dy = \frac{1}{2\pi}\int_{x_1}^{x}\Big[\int_{-T}^{T} e^{-ity}\varphi(t)dt\Big]dy$$

$$= -\frac{1}{2\pi it}\int_{-T}^{T}(e^{-itx}-e^{-itx_1})\varphi(t)dt .$$

Therefore

$$K(x) - K(x_1) = \lim_{T\to\infty}\int_{x_1}^{x} g_T(y)dy = \int_{x_1}^{x} g(y)dy . \tag{2.9.12}$$

the last equality holding because of uniform convergence. For x_1 fixed, expression (2.9.12) implies that $K(x)$ is differentiable for all $x > x_1$ and the proof is complete.

Chapter III

CORRELATION BETWEEN GENERATIONS

3.1 Introduction

In this chapter models are discussed in which a parent's life-span has some effect upon the life-span of its children. In practice this would be the case if the life-spans of individuals were influenced by heredity.

The first such model begins with a single individual born at time $t = 0$. This individual has a random life-span and at the end of its life it is replaced by a random number of offspring, and so the process continues. As in the model of Bellman and Harris, the numbers of offspring of individuals in the population form a collection of independent and identically distributed random variables with generating function $h(s)$, and these random variables are also independent of the life-spans of any members of the population. It is always assumed that $m \equiv h'(1) < \infty$.

The dependence is injected into the model by assuming if $\ell_1,\ldots,\ell_j$ are the life-spans of individuals $o_1,\ldots,o_j$, where o_k is the parent of o_{k+1}, $k = 1,2,\ldots,j-1$, then the random variables $\ell_1,\ldots,\ell_j$ form a Markov chain. The transition function of the Markov chain is determined by $G(t,x)$, which represents the conditional distribution of the life-span of an individual given that the individual's parent had a life-span of x. The life-spans of a group of siblings are taken to be conditionally independent given the life-span of the

parent.

The analysis of the process may proceed in two different ways. One may carefully define a probability space and then use this space to derive an integral equation satisfied by the generating function of the process as was done in Chapter II, or one may take the integral equation as formally given, prove the existence and uniqueness of a solution, and begin the analysis at this point. It is the latter route that will be followed here. First of all, however, a discussion of $G(t,x)$ must be given.

3.2 The function $G(t,x)$.

We shall assume that for fixed x, $G(t,x)$ is a distribution function concentrated on the non-negative real line, and that for fixed t, $G(t,x)$ is a Lebesgue measurable function of x.

<u>Definition 3.2.1</u>: Let

$$G_0(t,x) = \begin{cases} 1 & \text{if } t \geq 0 \text{ and } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.2.1}$$

and

$$G_{k+1}(t,x) = \int_0^t G_k(t-u,u)dG(u,x), \quad k = 1,2,\ldots . \tag{3.2.2}$$

It follows from this definition that $G_1(t,x) \equiv G(t,x)$. Although the definition of $G_k(t,x)$ resembles the definition of a k-th order convolution, one must not assume that $G_k(t,x)$ has all the properties of a convolution. For example, in general

$$G_{k+1}(t,x) \neq \int_0^t G(t-u,u)dG_k(u,x) .$$

In the sequel the following regularity conditions for $G(t,x)$ are assumed.

Conditions for $G(t,x)$:

(i) There exists a distribution function $H(t)$ concentrated on the non-negative real line such that $G(t,x) \leq H(t)$ for all x, and $\lim_{t\to 0} H(t) < 1/m$ (where $m = h'(1)$).

(ii) $G(t,x) \to 1$ as $t \to \infty$ uniformly for x on a bounded interval, i.e., for each $\epsilon > 0$ and $B > 0$, there exists a T such that $G(t,x) > 1 - \epsilon$ for all $t \geq T$ and all x, $0 \leq x \leq B$.

Example: If

$$G(t,x) = 1 - e^{-t/(x+1)}$$

then the above conditions are satisfied. One may take $H(t) = 1 - e^{-t}$, $t \geq 0$.

Theorem 3.2.1: For each fixed x and k, $G_k(t,x)$ is a distribution function in t. Moreover, for fixed t and x, the sequence $G_k(t,x)$ is nonincreasing.

Proof: We shall first prove by induction that condition (ii) for $G(t,x)$ holds also for $G_k(t,x)$, $k = 0,1,2,\ldots$. Let ϵ and B be given numbers, $B > 0$ and $0 < \epsilon < 1$. By condition (ii) for $G(t,x)$ there exists a T_1 such that $G(t,x) \geq (1-\epsilon)^{\frac{1}{2}}$ for all $t \geq T_1$, and x, $0 \leq x \leq B$. Given T_1, by the induction hypothesis there exists a $T_2 > T_1$ such that $G_k(t,x) > (1-\epsilon)^{\frac{1}{2}}$ for all $t \geq T_2 - T_1$.

and x, $0 \le x \le B$. Therefore, for $t > T_2$,

$$G_{k+1}(t,x) = \int_0^t G_k(t-u,u)dG(u,x) \ge \int_0^{T_1} G_k(T_2-u,u)dG(u,x)$$

$$\ge (1-\epsilon)^{\frac{1}{2}} G(T_1,x) > 1-\epsilon$$

for all x, $0 \le x \le B$.

The proof of the remaining properties of a distribution function will be omitted.

It follows from the definition of $G_0(t,x)$ and the first part of this theorem that $G_0(t,x) \ge G_1(t,x)$. Assuming that $G_k(t,x) \ge G_{k+1}(t,x)$, we see that

$$G_{k+1}(t,x) = \int_0^t G_k(t-u)dG(u,x) \ge \int_0^t G_{k+1}(t-u,u)dG(u,x) = G_{k+2}(t,x) \quad (3.2.4)$$

which completes the proof of the theorem.

We now interrupt our discussion of $G(t,x)$ to prove two lemmas which will be useful later. The facts in both of these lemmas appear in various places (see, for example, Loeve (1963) page 167), but the proofs are often omitted. The technique employed in proving the second lemma was used by Tukey (1958). This same tehcnique will be used again in Theorems 3.2.2 and 3.5.2.

<u>Lemma 3.2.1</u>: Suppose $G(t)$ and $F(t)$ are distribution functions such that $G(t) \le F(t)$ and U is a random variable with the uniform distribution on $[0,1]$. Then there exists functions $Z_G(u)$ and $Z_F(u)$ defined on $(0,1)$ such that

(1) the distribution of $Z_G(U)$ is $G(t)$,

(2) the distribution of $Z_F(U)$ is $F(t)$, and

(3) $Z_F(u) \le Z_G(u)$ for all $u \in (0,1)$.

Proof: Define $Z_G(u) = \inf\{x;\ G(x) \ge u\}$. From the fact that $Z_G(u)$ is non-decreasing it follows that

$$P[U \le G(t)] \le P[Z_G(U) \le Z_G(G(t))] = P[Z_G(U) \le t] \qquad (3.2.5)$$

On the other hand, $G(Z_G(u)) \le G(t)$ for each u such $Z_G(u) \le t$, since $G(t)$ is nondecreasing. But $u \le G(Z_G(u))$ from the definition of $Z_G(u)$. Therefore, $\{u;\ Z_G(u) \le t\} \subseteq \{u;\ u \le G(t)\}$ which implies that

$$P[Z_G(U) \le t] \le P[U \le G(t)]\,. \qquad (3.2.6)$$

Expressions (3.2.5) and (3.2.6) together imply that

$$P[Z_G(U) \le t] = P[U \le G(t)] = G(t)\,. \qquad (3.2.7)$$

If we define $Z_F(u) = \inf\{x;\ F(x) \ge u\}$ the proof of the second assertion is completely analogous to that of the first.

To prove the third assertion, note that $G(t) \le F(t)$ implies that for each $u \in (0,1)$, $\{x;\ G(x) \ge u\} \subseteq \{x;\ F(x) \ge u\}$ and, therefore,

$$Z_G(u) = \inf\{x;\ G(x) \ge u\} \ge \inf\{x;\ F(x) \ge u\} = Z_F(u)\,. \qquad (3.2.8)$$

which completes the proof of the lemma.

The following lemma, which will be used in Section 3.4, illustrates

the utility of Lemma 3.2.1 .

Lemma 3.2.2: Suppose $G(t)$ and $F(t)$ are distribution functions such that $G(t) \le F(t)$ for all t and f is a measurable function such that $\int_{-\infty}^{\infty} f(t)dG(t)$ and $\int_{-\infty}^{\infty} f(t)dF(t)$ are finite. If $f(t)$ is nonincreasing then $\int_{-\infty}^{\infty} f(t)dG(t) \le \int_{-\infty}^{\infty} f(t)dF(t)$, while if $f(t)$ is nondecreasing, the inequality is reversed.

Proof: From Lemma 3.2.1, if $Y = Z_G(U)$, we have, for $f(t)$ nonincreasing,

$$\begin{aligned}\int_{-\infty}^{\infty} f(t)dG(t) &= E[f(Y)] = E\{f[Z_G(U)]\} = \int_0^1 f[Z_G(u)]du \\ &\le \int_0^1 f[Z_F(u)]du = \int_{-\infty}^{\infty} f(t)dF(t) .\end{aligned} \tag{3.2.8}$$

If $f(t)$ is nondecreasing a similar statement holds.

Let $H^{k*}(t)$ denote the k-th convolution of $H(t)$ with itself.

Theorem 3.2.2: For each $x \ge 0$ and $t \ge 0$ we have

$$G_k(t,x) \le H^{k*}(t) \tag{3.2.9}$$

and

$$\sum_{k=0}^{\infty} m^k G_k(t,x) \le \sum_{k=0}^{\infty} m^k H^{k*}(t) < \infty \tag{3.2.10}$$

Proof: The first inequality in (3.2.10) follows from (3.2.9) and the second follows from Lemma 1.4.1 .

To prove the first assertion, let $Z_x(u)$ and $Z_H(u)$ be the functions

associated with $G(t,x)$ and $H(t)$, respectively, by Lemma 3.2.1. Let U be the uniformly distributed variable of Lemma 3.2.1 and let Y be defined by $Y = Z_x(U)$. By the conditions on $G(t,x)$ we know that $G_1(t,x) \le H(t)$ for all x. Assume, for a given positive integer k, that $G_k(t,x) \le H^{k*}(t)$ for all x. It follows that

$$\begin{aligned} G_{k+1}(t,x) &= \int_0^t G_k(t-u,u)dG(u,x) = E[G_k(t-Y,Y)] \\ &= \int_0^1 G_k(t-Z_x(u),Z_x(u))du . \end{aligned} \tag{3.2.11}$$

But by Lemma 3.2.1, $Z_x(u) \ge Z_H(u)$. Therefore,

$$\begin{aligned} G_{k+1}(t,x) &\le \int_0^1 G_k(t-Z_H(u),Z_x(u)du \\ &\le \int_0^1 H^{k*}(t-Z_H(u))du \\ &= \int_0^t H^{k*}(t-u)dH(u) = H^{(k+1)*}(t) \end{aligned} \tag{3.2.12}$$

which completes the proof of the theorem.

3.3 An integral equation.

For the process described in section 3.1, the law of total probability suggests that $F(s,t,x)$, the generating function of the number of individuals alive at time t, given the parent of the initial individual had lifespan x, satisfies the integral equation

$$F(x,t,x) = s[1 - G(t,x)] + \int_0^t h[F(x,t-u,u)]dG(u,x) . \qquad (3.3.1)$$

Theorem 3.3.1: [1]Under the conditions given previously for $h(s)$ and $G(t,x)$, equation (3.3.1) has a unique solution bounded in absolute value by one in the region $-1 \le s \le 1$, $t \ge 0$, and $x \ge 0$. Moreover, for fixed t and x, $F(s,t,x)$ is a generating function in s and for fixed s, $F(s,t,x)$ is a Lebesgue measurable function of t and x.

Proof: The proof follows fairly closely that of Theorem I of Ney (1964a). Let $F_0(s,t,x) \equiv 0$. Then define inductively

$$F_{k+1}(s,t,x) = s[1 - G(t,x)] + \int_0^t h[F_k(s,t-u,u)]dG(u,x) . \qquad (3.3.2)$$

By induction it follows that $|F_k(s,t,x)| \le 1$ for $k = 1,2,\ldots$.

Furthermore,

$$\begin{aligned} |F_2(s,t,x) - F_1(s,t,x)| &\le \int_0^t |h[F_1(s,t-u,u)] - h[F_0(s,t-u,u)]|dG(u,x) \\ &\le m\int_0^t |F_1(s,t-u,u) - F_0(s,t-u,u)|dG(u,x) \qquad (3.3.3) \\ &\le mG(t,x) , \end{aligned}$$

where we have used the mean value theorem and the fact that $m = h'(1) \ge h'(s)$

1. A less general form of Theorem 3.3.1 can be proved using the fixed-point theorem for contraction mappings on a complete metric space. See Sevast'yanov (1964).

for all s, $-1 \le s \le 1$. It now follows by induction that

$$|F_{k+1}(s,t,x) - F_k(s,t,x)| \le m^k G_k(t,x), \quad k = 1,2,\ldots . \tag{3.3.4}$$

Therefore, applying Theorem 3.2.2,

$$\begin{aligned} |F_n(s,t,x) - F_k(s,t,x)| &\le F_{k+1}(s,t,x) - F_k(s,t,x)| + \cdots \\ &\quad + |F_n(s,t,x) - F_{n-1}(s,t,x)| \qquad (3.3.5) \\ &\le \sum_{j=k}^{n-1} m^j G_j(t,x) \le \sum_{j=k}^{\infty} m^j H^{j*}(t), \quad n > k . \end{aligned}$$

The last term in (3.3.5) can be made as small as we please by choosing k large enough. Therefore, $F_k(s,t,x)$ is a Cauchy sequence for each s, t, and x, and converges to a function which we shall denote by $F(s,t,x)$. Upon noting that the last term in (3.3.5) is independent of s and x, and that $H^{k*}(t)$ is increasing in t, it follows that $F_k(s,t,x) \to F(s,t,x)$ as $k \to \infty$ uniformly for all s, x, and t in the region $-1 \le s \le 1$, $x \ge 0$, and $0 \le t \le T$ where T is arbitrary. Since $|F_k(s,t,x)| \le 1$ for all k, we have $|F(s,t,x)| \le 1$. Taking the limit of both sides of (3.3.2) as $k \to \infty$, we have, using the bounded convergence theorem, that $F(s,t,x)$ satisfies equation (3.3.1).

To prove the uniqueness of the solution, suppose $Q(s,t,x)$ is another solution of (3.3.1) satisfying $|Q(s,t,x)| \le 1$, and let T be a positive number such that $mH(T) < 1$. (Such a number exists since

$\lim_{t\to 0} H(t) < 1/m$.) Then

$$|F(s,t,x) - Q(s,t,x)| \le \int_0^t |h[F(s,t-u,u)] - h[Q(s,t-u,u)]|dG(u,x)$$

$$\le m\int_0^t |F(s,t-u,u) - Q(s,t-u,u)|dG(u,x) . \tag{3.3.6}$$

Taking the symbol "Sup" to mean the sup over the set $x \ge 0$ and $0 \le t \le T$ and assuming $F(s,t,x)$ is not identical to $Q(s,t,x)$ on this set, it follows that

$$\begin{aligned} \mathrm{Sup}\{|F(s,t,x) - Q(s,t,x)|\} &\le m[\mathrm{Sup}\{|F(s,t,x) - Q(s,t,x)|\}] \\ &\quad \cdot [\mathrm{Sup}\int_0^t dG(u,x)] \\ &\le m[\mathrm{Sup}\{|F(s,t,x) - Q(s,t,x)|\}]H(T) \\ &< \mathrm{Sup}\{|F(s,t,x) - Q(s,t,x)|\} . \end{aligned} \tag{3.3.7}$$

This is a contradiction and, therefore, $F(s,t,x) = Q(s,t,x)$ for all s, t, and x, $-1 \le s \le 1$, $x \ge 0$, and $0 \le t \le T$.

Now assume, as an induction hypothesis, that $F(s,t,x) = Q(s,t,x)$ for all s, t, and x, $-1 \le s \le 1$, $x \ge 0$, and $0 \le t \le nT$. Supposing now the symbol "Sup" stands for the sup over all x and t, $x \ge 0$ and $0 \le t \le (n+1)T$, we obtain

$$\operatorname{Sup}\{|F(s,t,x) - Q(s,t,x)|\}$$

$$\leq m\left[\sup\left\{\int_0^{t-nT} |F(s,t-u,u) - Q(s,t-u,u)|\,dG(u,x)\right\}\right] \qquad (3.3.8)$$

$$\leq m \sup\{|F(s,t,x) - Q(s,t,x)|\}H(T)$$

which, again, is a contradiction unless $F(s,t,x) = Q(s,t,x)$, $0 \leq t \leq (n+1)T$. This completes the proof of the uniqueness of the solution.

It is easily seen that $F(s,t,x) = 1$ is a solution of (3.3.1) for $s = 1$, so it follows from uniqueness that $F(1,t,x) = 1$.

By induction, for fixed t and x, $F_k(s,t,x)$ is analytic in s, $-1 \leq s \leq 1$, for $k = 1,2,\ldots$. Since the limit of a uniformly convergent sequence of analytic functions is again analytic, $F(s,t,x)$ is also analytic in s, so that $F_k(s,t,x)$ and $F(s,t,x)$ can be expressed as power series expansions in s. Moreover, the coefficients in the power series for $F(s,t,x)$ are non-negative since this is true for $F_k(s,t,x)$, $k = 1,2,\ldots$. Therefore, for fixed t and x, $F(s,t,x)$ is a generating function in s. Since the limit of a sequence of measurable functions is measurable, $F(s,t,x)$ is measurable in t and x for fixed s. This completes the proof of Theorem 3.3.1.

3.4 Extinction probability

The probability $Q(t,x)$ that the process has terminated by time t is given in terms of the generating function by $F(0,t,x)$.

Lemma 3.4.1: $Q(t,x)$ is nondecreasing in t.

Proof: If we define $Q_k(t,x) = F_k(0,t,x)$ then (3.3.2) becomes

$$Q_{k+1}(t,x) = \int_0^t h[Q_k(t-u,u)]dG(u,x) . \tag{3.4.1}$$

It follows immediately by induction that $Q_k(t,x)$ is nondecreasing in t, using the fact that $h(s)$ is nondecreasing. Since $Q_k(t,x) \to Q(t,x)$ as $k \to \infty$, $Q(t,x)$ is nondecreasing in t, also.

It is obvious that $Q(t,x)$ is nondecreasing in x from a probabilistic viewpoint since if the process has terminated by time t, it has certainly terminated by time $t' > t$. However, it was necessary that this be demonstrated using only an analytic argument.

The probability of extinction is defined as

$$Q = \lim_{t\to\infty} Q(t,x) .$$

This limit always exists since $Q(t,x)$ is nondecreasing in t according to Lemma 3.4.1 and bounded above by 1 according to Theorem 3.3.1. As we shall see, Q is independent of x.

The functions $h_n(s)$ appearing in the following theorem are the iterates of $h(s)$ defined by $h_1(s) = h(s)$ and $h_{n+1}(s) = h[h_n(s)]$ for $n = 2,3,\ldots$.

Theorem 3.4.1: The extinction probability Q is given by

$$Q = \lim_{k\to\infty} h_k(0) .$$

Remark: By this theorem Q is the probability of extinction of a Galton-Watson process with generating function $h(s)$ and, therefore, Q is the smallest non-negative solution of $h(s)$. Excluding the trivial case $h(s) \equiv s$, $Q = 1$ if $m \le 1$ and $Q < 1$ if $m > 1$.

Proof: If $Q_k(t,x)$ is defined as in Lemma 3.4.1, then $Q_k(t,x) \to Q(t,x)$ as $k \to \infty$, so the proof will be complete upon demonstrating that $Q_k(t,x) \to h_k(0)$ as $t \to \infty$. Since from (3.3.2), $Q_1(t,x) = h(0)G(t,x)$, it follows that $Q_1(t,x) \to h_1(0)$ as $t \to \infty$ uniformly for x in any finite interval. Assume that the corresponding property for $G_k(t,x)$ holds for some integer k. For $t > T_2 > T_1$, it can easily be shown using (3.4.1) that

$$\begin{aligned} |Q_{k+1}(t,x) - h_{k+1}(0)| &\le m\int_0^{T_1} |Q_k(t-u,u) - h_k(0)|\,dG(u,x) + [1 - G(T_1,x)] \\ &\le m \sup_{\substack{y \ge T_2 - T_1 \\ 0 \le z \le T_1}} \{ |Q_k(y,z) - h_k(0)| \} + [1 - G(T_1,x)] \end{aligned} \tag{3.4.2}$$

For given $\epsilon > 0$ and $B > 0$, T_1 can be chosen so large that $1 - G(T_1,x) < \epsilon/2$ for all x, $0 \le x \le B$. Then for fixed T_1, T_2 can be chosen so large that the first term in the last line of (3.4.2) is less than $\epsilon/2$ also, which completes the proof.

3.5 First moment

In this section some of the properties of

$$M(t,x) \equiv \frac{\partial}{\partial s} F(s,t,x)\Big|_{s=1}$$

will be investigated.

Theorem 3.4.1: $M(t,x)$ satisfies the integral equation

$$M(t,x) = 1 - G(t,x) + m\int_0^t M(t-u,u)dG(u,x) \tag{3.5.1}$$

and $M(t,x)$ is given by the infinite series

$$M(t,x) = \sum_{k=0}^{\infty} m^k[G_k(t,x) - G_{k+1}(t,x)] \tag{3.5.2}$$

or equivalently,

$$M(t,x) = 1 + (m-1) \sum_{k=1}^{\infty} m^{k-1} G_k(t,x) . \tag{3.5.3}$$

Furthermore, there exists positive constants A_1 and α_1, both independent of x, such that

$$M(t,x) \le A_1 e^{\alpha_1 t} \tag{3.5.4}$$

for all $x \ge 0$ and $M(t,x)$ is the unique solution of (3.5.1) having this property.

Proof: Let

$$M_k(t,x) = \frac{\partial}{\partial s} F_k(s,t,x)\Big|_{s=1} \tag{3.5.5}$$

One can see by induction using (3.3.2) that $M_k(t,x) < \infty$, $k = 1,2,\ldots$. Differentiating both sides of (3.3.2) with respect to s and letting $s \to 1$ yields the relation

$$M_{k+1}(t,x) = 1 - G(t,x) + m\int_0^t M_k(t-u,u)dG(u,x). \tag{3.5.6}$$

By induction, it follows that

$$M_k(t,x) = \sum_{j=0}^{k-1} m^j[G_j(t,x) - G_{j+1}(t,x)]. \tag{3.5.7}$$

We see from Theorem 3.2.1 that $M_k(t,x)$ is a nondecreasing sequence. With the notation $L(t,x) = \lim_{k\to\infty} M_k(t,x)$ it follows that

$$L(t,x) = \sum_{j=0}^{\infty} m^j[G_j(t,x) - G_{j+1}(t,x)] \tag{3.5.8}$$

and, upon rearranging terms, we get the alternate expression

$$L(t,x) = 1 + (m-1)\sum_{j=1}^{\infty} m^{j-1}G_j(t,x). \tag{3.5.9}$$

Now consider a Bellman-Harris age-dependent branching process for which m is the mean number of offspring per individual and H(t) is the distribution of the life-span of each individual where H(t) is the bounding function for G(t,x). From Lemma 1.3.1 $\overline{M}(t)$, the mean of this process, is given by

$$\overline{M}(t) = 1 + (m-1)\sum_{k=1}^{\infty} m^{k-1}H^{k*}(t) . \tag{3.5.10}$$

Moreover, $\overline{M}(t)$ is bounded on every finite interval and $M(t) \sim A_1 e^{\alpha_1 t}$ as $t \to \infty$.[1] Using (3.3.6), (3.3.7), and Theorem 3.2.2, we see that if $m > 1$, $0 \le L(t,x) \le \overline{M}(t)$, and if $m \le 1$, $0 \le L(t,x) \le 1$. In either case, $L(t,x) \le A_1 e^{\alpha_1 t}$ for some positive constants A_1 and α_1.

Since for fixed x and t $F_k(s,t,x)$ is analytic in x $(-1 \le s \le 1)$, and converges to $F(s,t,x)$ as $k \to \infty$ uniformly in s, it follows that

$$\frac{\partial}{\partial s}F_k(s,t,x) \to \frac{\partial}{\partial s}F(s,t,x) \quad \text{as} \quad k \to \infty . \tag{3.5.11}$$

Letting $s \to 1$ in (3.5.11) yields $M(t,x) = L(t,x)$.

If $k \to \infty$ on both sides of (3.5.4), using the monotome convergence theorem, we get equation (3.5.1).

To prove uniqueness, suppose $N(t,x)$ is another solution of (3.5.1) with the given properties. Define

$$B(t,x) = e^{-rt}\left|M(t,x) - N(t,x)\right| \tag{3.5.12}$$

where r is a constant chosen so large that $B(t,x)$ is bounded above and

$$m\int_0^{\infty} e^{-rt}dH(t) < 1 .$$

---

1 See Feller (1957), page 50.

It can easily be shown using (3.5.1) that

$$B(t,x) \leq \int_0^t B(t-u,u)e^{-ru}dG(u,x). \tag{3.5.13}$$

If T is an arbitrary positive number and the symbol "Sup" stands for the sup over all $x \geq 0$ and all t, $0 \leq t \leq T$, it follows that

$$\mathrm{Sup}\{B(t,x)\} \leq m\,\mathrm{Sup}\{B(t,x)\}\,\mathrm{Sup}\left\{\int_0^t e^{-ru}dG(u,x)\right\}. \tag{3.5.14}$$

Using Lemma 3.2.2 we get

$$\mathrm{Sup}\{B(t,x)\} \leq m\,\mathrm{Sup}\{B(t,x)\}\int_0^\infty e^{-ru}dH(u),$$

which is a contradiction to the choice of r unless $\mathrm{Sup}\{B(t,x)\} = 0$. This completes the proof.

Several properties of $M(t,x)$ can be deduced immediately from (3.5.3). For a fixed x

i) if $m > 1$, $M(t,x)$ is nondecreasing in t and $M(t,x) \to \infty$ as $t \to \infty$,

ii) if $m = 1$, then $M(t,x) \equiv 1$, and

iii) if $m < 1$, then $M(t,x)$ is nonincreasing in t and $M(t,x) \to 0$ as $t \to \infty$.

Each of these properties is identical to the corresponding property of the Bellman-Harris process. However, in the model presently being studied, it is not always true, as it is for the Bellman-Harris process,

that for $m > 1$ $M(t,x) \sim Ae^{\alpha t}$ as $t \to \infty$ for some numbers A and α. In fact, it may be the case that $M(t,x)e^{-\alpha t} \to 0$ as $t \to \infty$ for all $\alpha > 0$. The following example will illustrate this point.

Example: Suppose $m > 1$ and the life-span of an individual is twice that of his parent with probability one. Then

$$G(t,\tfrac{1}{2}) = \begin{cases} 0 & \text{if } t < 1 \\ 1 & \text{if } t \geq 1, \end{cases}$$

$$G_2(t,\tfrac{1}{2}) = \begin{cases} 0 & \text{if } t < 3 \\ 1 & \text{if } t \geq 3, \end{cases}$$

and, in general, $G_k(t,\tfrac{1}{2}) = \delta(t+1-2^k)$

where

$$\delta(t) = \begin{cases} 1 & \text{if } t \geq 0 \\ 0 & \text{if } t < 0. \end{cases}$$

Thus

$$M(t,\tfrac{1}{2}) = 1 + (m-1)\sum_{k=1}^{\infty} m^{k-1}\delta(t+1-2^k) = 1 + (m-1)\sum_{k=1}^{N(t)} m^{k-1} \qquad (3.5.15)$$

where $N(t)$ is the greatest integer less than or equal to $\mathrm{Log}(t+1)/\mathrm{Log}(2)$. From (3.5.15) and the fact that $m > 1$, it follows that

$$M(t,\tfrac{1}{2}) \le 1 + (m-1)m^{[\mathrm{Log}(t+1)/\mathrm{Log}(2)]}\mathrm{Log}(t+1)/\mathrm{Log}(2) .$$

It can now be easily shown that if $r > [\mathrm{Log}(m)/\mathrm{Log}(2)] + 1$ then

$$\lim_{t\to\infty} M(t,\tfrac{1}{2})/t^r = 0 \tag{3.5.16}$$

which is a stronger conclusion than the one we set out to illustrate.

By imposing further regularity conditions on $G(t,x)$, it is possible (when $m > 1$) to have

$$M(t,x) \ge A_2 e^{\alpha_2 t} \tag{3.5.17}$$

for some positive constants A_2 and α_2, so that the phenomenon illustrated by the last example cannot occur. Specifically, let us suppose that $G(t,x) \to 1$ as $t \to \infty$ uniformly in x. This implies the existence of a distribution function $k(t)$ with the property that $G(t,x) \ge k(t)$ for all $x \ge 0$. Let $\underline{M}(t)$ be the mean of a Bellman-Harris process for which m is the expected number of offspring per individual and $k(t)$ is the distribution of the life-span of each individual. As in the proof of Theorem 3.5.1, $\underline{M}(t) \le M(t,x)$, and, since $\underline{M}(t)$ is asymptotically exponential as $t \to \infty$, (3.5.17) follows. Combining this result with Theorem 3.5.1, we see that for $m > 1$, under the additional assumption that $G(t,x) \to 1$ as $t \to \infty$ uniformly in x, there exists positive constants A_1, A_2, α_1, α_2 independent of x such that

$$A_2 e^{\alpha_2 t} \leq M(t,x) \leq A_1 e^{\alpha_1 t}. \tag{3.5.18}$$

It would be desirable to prove under certain conditions (other than complete independence) that $M(t,x) \sim Ae^{\alpha t}$ as $t \to \infty$, and to be able to evaluate A and α. At present the author knows of no technique that would enable one to accomplish this. It does seem that if such a result were proved, then α would have to be the unique positive number such that $M^*(\lambda,x)$, the Laplace transform of $M(t,x)$, exists for all $\lambda > \alpha$ and fails to exist for all $\lambda < \alpha$.

As a final result for this process, we show that a reasonable monotonicity assumption on $G(t,x)$ implies a similar result for $M(t,x)$.

<u>Definition 3.5.1</u>: The distributions $G(t,x)$ will be called monotone if for any $x_2 < x_1$ and any t, $G(t,x_1) \geq G(t,x_2)$.

This assumption says essentially that individuals with long life spans tend to have offspring with the same characteristic and, therefore, seems to be a very natural condition to impose on the process.

<u>Theorem 3.5.2</u>: Suppose that the distributions $G(t,x)$ are monotone. Then if $m > 1$, $M(t,x)$ is a nonincreasing function of x, and if $m < 1$, $M(t,x)$ is a nondecreasing function of x.

<u>Proof</u>: The theorem will follow from (3.5.3) if we can establish that $G_k(t,x)$ is monotone for $k = 1,2,\ldots$. Since $G_1(t,x) = G(t,x)$ it follows that G_1 is monotone by hypothesis. Suppose that $G_n(t,x)$ is monotone for some positive integer n, and let $Z_x(u)$ be the function associated with $G(t,x)$ by Lemma 3.2.1. Using this lemma

$$G_{k+1}(t,x_1) = \int_0^t G_k(t-u,u)dG(u,x_1) = \int_0^1 G_k(t-Z_{x_1}(u),Z_{x_1}(u))du .$$

But Lemma 3.2.1 implies that $Z_x(u)$ is nondecreasing in x and, hence, if $x_2 > x_1$,

$$G_{k+1}(t,x_1) \geq \int_0^1 G_k(t - Z_{x_1}(u),Z_{x_2}(u))du$$

$$\geq \int_0^1 G_k(t - Z_{x_2}(u),Z_{x_2}(u))du$$

$$= \int_0^t G_k(t - u,u)dG(u,x_2) = G_{k+1}(t,x_2) .$$

3.6 Another model

We now discuss, briefly, a model with dependence between generations for which it is possible to obtain many results analogous to those obtained for the Bellman-Harris process, although the underlying assumptions for the process perhaps seem quite artificial.

Suppose an individual's life is the sum of a "first life" $\ell^{(1)}$ and a "second life" $\ell^{(2)}$ where $\ell^{(1)}$ and $\ell^{(2)}$ are independent. We shall make no attempt to relate these two lives to some physical phenomenon. If $\ell_0^{(2)}$ is the second life of a parent and $\ell_1^{(1)}$ is the first life of its offspring (it is assumed that all offspring of a single individual have identical first lives), then $\ell_0^{(2)}$ and $\ell_1^{(1)}$ are dependent.

All other necessary relations are governed by independence. It should be pointed out that the dependence in this model lasts only through two generations, i.e., the life-span of an individual and that of his grandparent are independent.

The distribution of the lengths of the first and second lives of an individual will be denoted by $G_1(t)$ and $G_2(t)$, respectively, and the distribution of $\ell_0^{(2)} + \ell_1^{(1)}$, the length of the second life of a parent plus the length of the first life of its offspring, will be denoted by $L(t)$.

If one wishes to study a process in which the life-span of a single individual has mean μ, variance σ^2, and the correlation between the life-spans of a parent and its offspring is ρ, $-\frac{1}{2} \le \rho \le \frac{1}{2}$, one can obtain these parameter values with this model. For example, if $\ell_0 = \ell_0^{(1)} + \ell_0^{(2)}$ and $\ell_1^{(1)} + \ell_1^{(2)}$ are the life-spans of a parent and its child, with $\sigma_1^2 = \mathrm{Var}(\ell_0^{(1)}) = \mathrm{Var}(\ell_1^{(1)})$, $\sigma_2^2 = \mathrm{Var}(\ell_0^{(2)}) = \mathrm{Var}(\ell_1^{(2)})$, and $\rho' = \mathrm{Cor}(\ell_0^{(2)}, \ell_1^{(1)})$, then one may take $E(\ell_0^{(1)} + E(\ell_0^{(2)}) = \mu$,

$$\rho' = \begin{cases} 1 & \text{if} \quad \rho \ge 0 \\ -1 & \text{if} \quad \rho < 0, \end{cases}$$

$$\sigma_1^2 = \sigma^2 \frac{1 + (1 - 4\rho^2)^{\frac{1}{2}}}{2},$$

and $\sigma_2^2 = \sigma^2 - \sigma_1^2$. This follows at once upon noting that $\rho = \rho' \sigma_1 \sigma_2 / (\sigma_1^2 + \sigma_2^2)$.

Let $h(s)$ be the generating function for the distribution of the number of offspring of each individual, with $h'(1) = m < \infty$, and let $F_1(s,t)$ be the generating function for a process that begins with one individual that at time $t = 0$ has only its second life remaining. Using the methods of Chapter II it can be shown that

$$F_1(s,t) = s(1 - G_2(t)) + h(s)[G_2(t) - L(t)] + \int_0^t h[F_1(s,t-u)]dL(u) . \tag{3.6.1}$$

The mean

$$M_1(t) = \frac{\partial}{\partial s}F_1(s,t)\Big|_{s=1}$$

satisfies

$$M_1(t) = 1 + G_2(t) + m[G_2(t) - L(t)] + m\int_0^t M_1(t-u)dL(u) . \tag{3.6.2}$$

Applying Lemma 1.4.3, if $G_2(0) = 0$, $L(\infty) = 1$, and $m > 1$, then

$$M_1(t) \sim Ae^{\alpha t} \quad \text{as} \quad t \to \infty , \tag{3.6.3}$$

where α is the solution of

$$m\int_0^\infty e^{-\alpha t}dL(t) = 1 , \tag{3.6.4}$$

and

$$A = \frac{\int_0^\infty \{1 - G_2(t) + m[G_2(t) - L(t)]\}e^{-\alpha t}dt}{m\int_0^\infty te^{-\alpha t}dL(t)} . \tag{3.6.5}$$

Now let $F(s,t)$ be the generating function of a process beginning with one individual born at time $t = 0$ (having both its first life and second life remaining). It can be shown by the methods used in Chapter II that

$$F(s,t) = s(1 - G_1(t)) + \int_0^t F_1(s,t-u)dG_1(u) . \tag{3.6.6}$$

The mean

$$M(t) = \frac{\partial}{\partial s}F(s,t)\Big|_{s = 1}$$

is given by

$$M(t) = 1 - G_1(t) + \int_0^t M_1(t-u)dG_1(u) . \tag{3.6.7}$$

Using Lemma 1.4.2, it can be shown that if $m > 1$, $G_2(0) = 0$ and $L(\infty) = 1$, then

$$M(t) \sim e^{\alpha t} A \int_0^\infty e^{-\alpha u} dG_1(u) \quad \text{as} \quad t \to \infty . \tag{3.6.8}$$

Similar techniques may be used to investigate the higher moments.

Chapter IV

AN AGE-DEPENDENT BIRTH-AND-DEATH PROCESS

4.1 Introduction

As pointed out in Chapter I, a feature of the Bellman-Harris process which limits its range of application is that the birth of a new individual occurs only at the instant of death of the parent. The process investigated in this chapter, however, allows births to occur throughout the life of the parent.

We assume that a single individual is born at time $t = 0$, lives for a random length of time ℓ with distribution $G(t)$. At random times during its lifetime this individual gives birth to offspring that behave in a similar manner, the behaviour of each individual being independent of all others. The random function $N(t)$ which counts the number of offspring the initial individual has in the age-interval $[0,t]$ is determined by randomly stopping an arbitrary counting process $K(t)$ at ℓ; i.e.,

$$N(t) = \begin{cases} K(t) & \text{if } t \le \ell \\ K(\ell) & \text{if } t > \ell , \end{cases}$$

and we assume that $K(t)$ and ℓ are independent.

The random function $Z(t)$ giving the size of the population at

time t is a generalization of the process discussed briefly by Harris (1963) which he also calls an "age-dependent birth-and-death process". In fact if

$$G(t) = 1 - \exp\left[\int_0^t \mu(x)dx\right]$$

and $K(t)$ is a Poisson process with mean function

$$E[K(t)] = \int_0^t \lambda(t)dt ,$$

we have Harris' process.

If

$$G(t) = 1 - e^{-\mu t}$$

and $K(t)$ is a compound homogeneous Poisson process, $Z(t)$ becomes a Markov branching process which is a special case of a Bellman-Harris process. However, in general $Z(t)$ is neither a special case nor a generalization of the Bellman-Harris process.

It has recently been brought to the attention of the author that Ryan (1967) has studied a more general process in which the probability of an individual dying in the interval $(t, t+dt)$ depends upon the times in

$[0,t]$ at which the individual has given birth. However, in a private communication, Ryan has indicated that his results have been mainly for the case in which the mean number of offspring per individual is less than one, so the only overlap in his results and those in this chapter seems to be in Theorems 4.7.2 and 4.7.4 (iii). Apparently we can obtain the level of generality of Ryan's model by dropping the assumption that $K(t)$ and ℓ are independent. This assumption does not play a major role until Section 4.8 and, in particular, the results in Sections 4.4 - 4.7 are not dependent upon it.

It should also be mentioned that most of the results of Goodman (1967) on the probability of extinction for several birth-and-death processes are easily deducible as special cases of Theorem 4.6.1.

4.2 The probability space

As in Section 2.2, for each positive integer n and each n-tuple $I = i_1 \ldots i_n$ of positive integers, $\langle i_1 \ldots i_n \rangle$ denotes a member of the population and the sequence $i_1 \ldots i_k$ denotes its line of descent. In addition, the individual $\langle 0 \rangle$ comprises the 0-th generation and $\langle 1 \rangle, \langle 2 \rangle, \langle 3 \rangle, \ldots$ are his offspring. Thus, for example $\langle 32 \rangle$ is the second child of $\langle 3 \rangle$ who in turn is the third child of $\langle 0 \rangle$.

By a family history we mean a sequence

$$\omega = \left\{ \ell_0, T_0; \ell_1, T_1; \ell_{11}, T_{11}; \ldots \right\} \tag{4.2.1}$$

where the subscripts run over all elements of $\mathcal{J}$ in some specified order. (The set $\mathcal{J}$ is defined as in Section 2.2, except it now contains the additional element 0 . As before, an arbitrary element of $\mathcal{J}$ is denoted by I.) The set of all such family histories is denoted by Ω. Each ℓ_I is a non-negative real number and represents the life-span of individual $\langle I \rangle$. Each T_I is itself a sequence of the form

$$T_I = (t_I^{(1)}, t_I^{(2)}, t_I^{(3)}, \ldots) , \tag{4.2.2}$$

where each $t_I^{(j)}$ is an extended real number and

$$0 \leq t_I^{(1)} \leq t_I^{(2)} \leq \ldots \leq \infty , \tag{4.2.3}$$

and represents the times starting from birth at which the individual $\langle I \rangle$ gives birth. For example, if $t_{123}^{(2)} = x$, then $\langle 1232 \rangle$ is born x time units after the birth of $\langle 123 \rangle$. The equality $t_I^{(j)} = t_I^{(j+1)}$ means that $\langle I \rangle$ gives birth to two offspring at the common time. If $t_I^{(j)} < \infty$ and $t_I^{(j+1)} = t_I^{(j+2)} = \ldots = \infty$, this simply means that $\langle I \rangle$ has no more than j offspring, no matter how long he lives. As was also true for the family histories defined in Section 2.2, each ω contains nonessential information since we simply ignore "births" that occur after the death of the parent.

The space Ω may also be expressed as the product space

$$\Omega_{\ell_o} \times \Omega_{T_o} \times \Omega_{\ell_1} \times \Omega_{T_1} \times \cdots \tag{4.2.4}$$

where each Ω_ℓ is a non-negative real line and each Ω_T is the subspace determined by the inequalities (4.2.3) of a countable product of extended real lines. As in Section 2.2 the σ-algebra $\mathcal{G}$ is the product σ-algebra determined by the representation (4.2.4).

To define a probability measure P on the measurable space $(\Omega, \mathcal{G})$, let P_{ℓ_I} be the probability on Ω_{ℓ_I} determined by a distribution $G(t)$, with $G(0) = 0$, and let P_{T_I} be a probability defined on Ω_{T_I}, with each P_{T_I}, $I \in \mathcal{J}$, having a common law. The probability P is then defined to be the unique probability determined by the probabilities on the coordinate spaces and the product probability theorem (Loeve (1963), page 91). This application of the product probability theorem implies that for any individual $\langle I \rangle$, its life span ℓ_I and the times T_I at which it gives birth are independent and these quanties are also independent of the corresponding quantities of the other individuals.

4.3 Properties of the counting functions $K(t)$ and $N(t)$.

<u>Definition 4.3.1</u>: For each $I \in \mathcal{J}$, let

$$K_I(t) = \begin{cases} 0 \text{ if } t_I^{(1)} > t \\ \sup\{i; t_I^{(i)} \leq t\} \quad \text{otherwise,} \end{cases} \tag{4.3.1}$$

$$N_I(t) = \begin{cases} K_I(t) & \text{if } \ell_I > t \\ K_I(\ell_I) & \text{if } \ell_I \leq t \end{cases} \tag{4.3.2}$$

and

$$N_I = K_I(\ell_I) \tag{4.3.3}$$

The random function $K_I(t)$ simply counts the points $t_I^{(j)}$ in the interval $[0,t]$. Each realization of $K_I(t)$ assumes non-negative integer values only(including possibly infinity) and is nondecreasing and continuous from the right. The function $N_I(t)$ represents the number of offspring born to $\langle I \rangle$ within t units after its birth and N_I represents the total number of offspring born to $\langle I \rangle$. We may also write

$$N_I = \lim_{t \to \infty} N_I(t) \tag{4.3.4}$$

Frequently, we shall denote $K_o(t), N_o(t), N_o, t_o^{(j)}$, and ℓ_o by $K(t), N(t), N, t^{(j)}$, and ℓ, respectivly.

Quite a few of the results we shall obtain are expressible in terms of distribution $G(t) = P[\ell \leq t]$ and the mean function

$$\Delta(t) \equiv E[K(t)] . \tag{4.3.5}$$

For the remainder of this chapter we shall assume the following conditions are met by these functions.

<u>Conditions for $G(t)$ and $\Delta(t)$</u>:

(i) The function $G(t)$ is nondecreasing, continuous from the right, $G(0-) = 0$ and $G(\infty) = 1$.

(ii) The function $\Delta(t)$ is finite, nondecreasing, continuous, equal to zero when $t = 0$, but not identically zero.

It should be noted that for any function $\Delta(t)$ with the properties given in (ii), there is a counting function $K(t)$ such that $E[K(t)] = \Delta(t)$. For example, one may take $K(t)$ to be a nonhomogeneous Poisson process with mean $\Delta(t)$. The requirement that $\Delta(t)$ be continuous is not essential and for most results given in the sequel, a corresponding result may be proved for the case $\Delta(t)$ is only continuous from the right.

Let us denote the generating functions of $K(t)$, $N(t)$, and N by $h_K(s,t)$, $h_N(s,t)$, and $h(s)$, respectively.

<u>Theorem 4.3.1</u>: (i) The generating functions $h_N(s,t)$ and $h(s)$ are given in terms of $h_K(s,t)$ by

$$h_N(s,t) = \int_0^t h_K(s,u)dG(u) + [1 - G(t)]h_K(s,t) \qquad (4.3.6)$$

and

$$h(s) = \int_0^\infty h_K(s,u)dG(u) \qquad (4.3.7)$$

for $t \geq 0$ and $|s| \leq 1$.

(ii) The means of $N(t)$ and N satisfy

$$E[N(t)] = \int_0^t \Delta(u)dG(u) + \Delta(t)[1 - G(t)] \qquad (4.3.8)$$

$$= \int_0^t [1 - G(u)]d\Delta(u)$$

and

$$E(N) = \int_0^\infty \Delta(u)dG(u) = \int_0^\infty [1 - G(u)]d\Delta(u) \qquad (4.3.9)$$

in the sense that if one integral converges then so does the other and they are equal.

Proof: Using conditional expectations and Definition (4.3.1), we get

$$h_N(s,t) = E[s^{N(t)}] = E\{E[s^{N(t)}|\ell]\}$$

$$=\int_0^t E[s^{N(t)}|\ell]dP_\ell + \int_{t+}^\infty E[s^{N(t)}|\ell]dP_\ell$$

$$=\int_0^t E[s^{K(\ell)}]dP_\ell + \int_{t+}^\infty E[s^{K(t)}]dP_\ell$$

$$=\int_0^t h_K(s,u)dG(u) + [1 - G(t)]h_K(s,t)$$

which proves (4.3.6). Expression (4.3.7) follows from (4.3.4) and (4.3.6) by applying the bounded convergence theorem.

Since $\Delta(t)$ is finite on every finite interval, we can differentiate (4.3.6) with respect to s, and interchange operations of integration and differentiation by a corollary of the dominated convergence theorem (Loeve (1963), page 126). If we perform these operations and put $s = 1$ in the resulting expression, we get

$$E[N(t)] = \int_0^t \Delta(u)dG(u) + \Delta(t)[1 - G(t)] \tag{4.3.10}$$

which is the first equality in (4.3.8). The second equality follows by

applying the integration by parts theorem for Riemann-Stieltjes integrals to (4.3.10).

The proof of (4.3.9) is quite similar to that of Lemma 6.1, page 148, of Feller (1966). First note that by the monotone convergence theorem

$$E(N) = \lim_{t\to\infty} E[N(t)] .$$

Therefore, letting $t\to\infty$ in the two extreme terms of (4.3.8) we obtain

$$E(N) = \int_0^\infty [1 - G(u)]d\Delta(u) \tag{4.3.11}$$

whether or not this integral is finite. If it is finite then

$$\int_0^\infty \Delta(u)dG(u) \tag{4.3.12}$$

is finite also, since by (4.3.8), this integral is no larger than the integral in (4.3.11). On the other hand, if (4.3.12) is finite, then

$$\lim_{t\to\infty} \int_t^\infty \Delta(u)dG(u) = 0$$

and since

$$\Delta(t)[1 - G(t)] \leq \int_t^\infty \Delta(u)dG(u)$$

we have

$$\lim_{t\to\infty} \Delta(t)[1 - G(t)] = 0$$

also, Therefore, letting $t\to\infty$ in (4.3.8), we obtain

$$\int_0^\infty \Delta(u)dG(u) = \int_0^\infty [1 - G(u)]d\Delta(u) < \infty$$

which completes the proof of Theorem 4.3.1.

4.4 An imbedded Galton-Watson process

As one might expect, the sizes of the successive generations form a Galton-Watson process. The k-th generation $\mathcal{J}_k$, $k = 0 , 1 , 2 ,\ldots$ is defined inductively by $\mathcal{J}_0 = \{\langle 0\rangle\}$, $\mathcal{J}_1 = \{\langle i\rangle ; 1\leq i\leq N_0\}$ and, assuming $\mathcal{J}_k$ is defined, $\mathcal{J}_{k+1}$ is the set of individuals $\langle i_1\ldots i_{k+1}\rangle$ such that $\langle i_1\ldots i_k\rangle \in \mathcal{J}_k$ and $i_{k+1}\leq N_{i_1\ldots i_k}$. The number of individuals in the family I_k is denoted by ξ_k .

<u>Theorem 4.4.1</u>: The random variables ξ_k are a Galton-Watson process with generating function $h(s)$.

Proof: The characteristic property of a Galton-Watson process is

$$E[s^{\xi_n} | \xi_1, \ldots, \xi_{n-1}] = [h(s)]^{\xi_{n-1}}, \quad n = 1, 2, \ldots \tag{4.4.1}$$

Upon noting that $\{N_I, I \in \mathcal{J}\}$ is a collection of independent random variables, each with generating function $h(s)$, the proof reduces to that of Theorem 5.1 of Harris (1963), page 127.

4.5 The branching processes $Z(t)$, $B(t)$, and $D(t)$

In this section, which is analogous to section 2.3, we shall define random functions denoting the number of individuals that are alive at time t, have been born by time t, and have died by time t. We shall also give a result similar to Lemma 2.4 which will be useful in deriving integral equations.

<u>Definition 4.5.1</u>: For each $I = i_1 \cdots i_k$ let

$$Z_I(t,\omega) = \begin{cases} 1 \text{ if } t_0^{(i_1)} \le \ell_0, \; t_{i_1}^{(i_2)} \le \ell_{i_1}, \ldots, t_{i_1}^{(i_k)} \cdots {}_{i_{k-1}} \le \ell_{i_1} \cdots {}_{i_{k-1}}, \\ \quad t_0^{(i_1)} + t_{i_1}^{(i_2)} + \ldots + t_{i_1}^{(i_k)} \cdots {}_{i_{k-1}} \le t, \\ \quad t_0^{(i_1)} + \ldots + t_{i_1}^{(i_k)} \cdots {}_{i_{k-1}} + \ell_{i_1} \cdots {}_{i_k} > t \\ 0 \quad \text{otherwise}, \end{cases} \tag{4.5.1}$$

$$0 \le t < \infty \ , \ \omega \in \Omega \ .$$

(An obvious modification must be made if $I = 0$).

Then let

$$Z(t,\omega) = \sum_{I \in \mathcal{J}} Z_I(t,\omega) \ . \tag{4.5.2}$$

The function $Z(t)$ is simply the number of individuals alive at time t . For a discussion of similar definitions the reader is referred to section 2.3 of this thesis and to page 125 of Harris (1963).

In an analogous fashion, we may define $B(t)$, the number of individuals born by time t , and $D(t)$, the number which have died by time t . The definition of $B(t)$ essentially amounts to recopying Definition 4.5.1 and deleting the third line of inequalities in (4.5.1). Then $D(t)$ may be defined as

$$D(t) = B(t) - Z(t) \tag{4.5.3}$$

whenever the difference is defined. As we shall see, $D(t)$ is defined almost surely. Since there are analogous results for the functions

$Z(t)$, $B(t)$, and $D(t)$ in many instances, we shall often prove a result for $Z(t)$ and then state the corresponding result for $B(t)$ and $D(t)$. Since the sample functions of $B(t)$ and $D(t)$ are nondecreasing, there are certain instances in which these functions are easier to work with than $Z(t)$.

For each $\omega = (\ell_o , T_o ; \ell_1 , T_1 ; \ldots)$ in the space Ω , let $\omega_i = (\ell_i , T_i ; \ell_{i1} , T_{i1} ; \ldots)$. Let Ω_i , $i = 1 , 2 , \ldots ,$ be the set of all such ω_i and let Ω_o be the set of all ordered pairs (ℓ_o , T_o) . Then Ω may be represented in the form

$$\Omega = \Omega_o \times \Omega_1 \times \Omega_2 \times \cdots$$

Let us denote the probability induced on Ω_i by P_i , $i = 0 , 1 \ldots$ Each space Ω_i ; $i = 1 , 2 , \ldots ,$ is a probabilistic replica of Ω so the symbol $Z(t,\omega_i)$, $i = 1 , 2 , \ldots ,$ has an obvious meaning.

<u>Lemma 4.5.1</u>: The function $Z(t,\omega)$ satisfies the equation

$$Z(t,\omega) = \sum_{j=1}^{N(t)} Z(t-t_o^{(j)},\omega_j) + 1 - \delta(t - \ell_o) \tag{4.5.4}$$

with the convention that if $N(t) = 0$ the sum vanishes. The function δ is defined by

$$\delta(t) = \begin{cases} 1 & \text{if} \quad t \geq 0 \\ 0 & \text{if} \quad t < 0 \end{cases} \tag{4.5.5}$$

Remark: The proof of the lemma will be omitted. The reader may refer to the discussion at the end of section 2.3 and to page 129 in Harris (1963) for a proof of a similar result.

4.6 Probability of extinction

The probability q that the population eventually becomes extinct is defined by

$$q = P[Z(t) = 0 \quad \text{for some} \quad t] \, . \tag{4.6.1}$$

One would expect that q is the same as the extinction probability of the imbedded Galton-Watson process $\{\xi_k\}$. Although this does turn out to be the case, a proof is required since, in fact, neither of the sets $[Z(t) = 0 \quad \text{for some} \quad t]$ or $[\xi_k = 0 \quad \text{for some} \quad k]$ is a subset of the other. To demonstrate this fact we offer the following two examples.

Example 1: Suppose ω is a point in Ω such that every generation is non-empty and each member of the k-th generation has life-span $1/2^k$. Then $Z(3,\omega) = 0$, but $\xi_k > 0$ for all k.

Example 2: Suppose ω is a point in Ω such that $\ell_o = 1$, $1/2 = t_o^{(1)} = t_o^{(2)} = \cdots$, $\ell_k = k$, and $t_k^{(1)} = k + 1$ for $k = 1, 2 \ldots$ It can easily be verified that $\xi_2(\omega) = \xi_3(\omega) = \cdots = 0$ but $Z(t;\omega) = \infty$ for all $t > 1/2$.

By the following theorem, however, the events illustrated in these examples occur with probability zero.

Theorem 4.6.1: Let A be the event $[\xi_k = 0$ for some $k]$ and let B be the event $[Z(t) = 0$ for some $t]$. Then $P(A) = P(B)$.

Proof: For a fixed $k > 0$ and a fixed t, we have

$$P[\xi_k > 0 \; ; \; Z(t) = 0] = P\{\bigcup_{i_1 \cdots i_k} [\langle i_1 \cdots i_k \rangle \in \mathcal{J}_k \; ; \; Z(t) = 0]\}$$

$$\leq \sum_{i_1=1}^{\infty} \cdots \sum_{i_k=1}^{\infty} P[\langle i_1 \ldots i_k \rangle \in \mathcal{J}_k : Z(t) = 0] \tag{4.6.2}$$

$$\leq \sum_{i_1=1}^{\infty} \cdots \sum_{i_k=1}^{\infty} P[t_o^{(i_1)} + t_{i_1}^{(i_2)} + \ldots + t_{i_1 \cdots i_{k-1}}^{(i_k)} + \ell_{i_1 \cdots i_k} \leq t]$$

The second inequality simply expresses the fact that if an individual $\langle I \rangle$ is actually born and the process is extinct by time t, then $\langle I \rangle$ must have died by this time. If we set

$$F_i(t) = P[t_I^{(i)} \le t]$$

and note that $t_I^{(i)} \le t$ if and only if $K_I^{(t)} \geqq i$, we see that (Feller (1957), page 249)

$$\sum_{j=1}^{\infty} F_j(t) = \sum_{j=1}^{\infty} P[K(t) \ge j] = E[K(t)] = \Delta(t) \ . \tag{4.6.3}$$

Using this fact, the last line in (4.6.2) becomes

$$\sum_{i_1=1}^{\infty} \cdots \sum_{i_k=1}^{\infty} F_{i_1} * F_{i_2} * \cdots * F_{i_k} * G(t)$$

$$= G * \sum_{i_1=1}^{\infty} F_{i_1} * \cdots * \sum_{i_k=1}^{\infty} F_{i_k}(t)$$

$$= G * \Delta^{k*}(t) \ .$$

But by Lemma 1.4.1 the last line of (4.6.4) approaches zero for fixed t as $k \to \infty$. Therefore,

$$\lim_{k \to \infty} P[\xi_k > 0 \ ; \ Z(t) = 0] = 0 \ .$$

It now follows easily that

$$P(A'B) = P[\xi_k > 0 \text{ for all } k \,;\, Z(t) = 0 \text{ for some } t]$$

$$= P\{\bigcup_{t=1}^{\infty} \bigcap_{k=1}^{\infty} [\xi_k > 0 \,;\, Z(t) = 0]\}$$

$$= \lim_{t\to\infty} \lim_{k\to\infty} P[\xi_k > 0 \,;\, Z(t) = 0]$$

$$= 0$$

and therefore

$$P(B) = P(AB) + P(A'B) = P(AB) \le P(A) .$$

The proof will be complete upon demonstrating the reverse inequality. If, for some k, $\xi_k = 0$ and $\xi_j < \infty$, $j = 1, 2 \ldots k-1$, then it is obvious that $Z(t) = 0$ for some t. Consequently, if $\omega \in AB'$, then $\xi_j = \infty$ for some j, and this, in turn, implies that $K_I(t) = \infty$ for some t and I. Therefore,

$$P(AB') \le P[K_I(t) = \infty \text{ for some } t \text{ and } I]$$

$$= P\{\bigcup_{I \in \mathcal{J}} \bigcup_{t=1}^{\infty} [K_I(t) = \infty]\}$$

$$\le \sum\sum P[K_I(t) = \infty] = 0 ,$$

since $\Delta(t) < \infty$ for all t implies $P[K(t) = \infty] = 0$ for all t. As before, we have

$$P(A) = P(AB) + P(AB') = P(AB) \le P(B)$$

and the proof is complete.

<u>Corollary</u>: The probability of extinction q is the smallest non-negative root of the equation

$$s = \int_0^{\infty} h_K(s,u)dG(u) .$$

If

$$h'(1) = \int_0^\infty [1 - G(u)]d\Delta(u) \leq 1$$

and $h(s) \neq s$, then $q = 1$. Otherwise, q is less than one.

Proof: The proof follows directly from Theorems 4.3.1 and 4.4.1 and remarks made in Section 1.2.

Example: (the birth and death process)

Suppose $K(t)$ is a Poisson process with $\Delta(t) = \lambda t$ and $G(t) = 1 - e^{-\mu}$. By Theorem 4.3.1

$$h(s) = \mu\int_0^\infty e^{\lambda t(s-1)}e^{-\mu t}dt = \frac{\mu}{\mu - \lambda(s-1)} .$$

The fixed point of $h(s) = s$ can easily be seen to be μ/λ and $h'(1) = \lambda/\mu$. Therefore, by the corollary, $q = 1$ if $\lambda \leq \mu$ and $q = \mu/\lambda$ is $\lambda > \mu$. This result was obtained in another way by Kendall (1949).

4.7 First moments

We shall demonstrate, first of all, that the conditions on $\Delta(t)$ and $G(t)$ given in section 4.3 are sufficient to insure the finiteness of $M(t)$.

Lemma 4.7.1: For each t , $E[B(t)] < \infty$.

<u>Proof</u>: For each $I = (i_1 \ldots i_k) \in \mathcal{J}$ let

$$b_I(t) = \begin{cases} 1 & \text{if } t_o^{(i_1)} + \ldots + t_{i_1 \ldots i_{k-1}}^{(i_k)} \leq t \\ 0 & \text{otherwise} \end{cases}$$

(We assume that $b_o(t) = 1$.) The event "$\langle i_1 \ldots i_k \rangle$ has been born by time t " is a subset of the event $[b_I(t) = 1]$ and therefore

$$B(t) \leq \sum_{I \in \mathcal{J}} b_I(t) \ .$$

Taking expectations and again using (4.6.3) we get

$$E[B(t)] \leq \sum_{I \in \mathcal{J}} E[b_I(t)]$$

$$= 1 + \sum_{k=1}^{\infty} \sum_{i_1=1}^{\infty} \ldots \sum_{i_k=1}^{\infty} F_{i_1} * \ldots * F_{i_k}(t)$$

$$= 1 + \sum_{k=1}^{\infty} \Delta^{k*}(t)$$

and this last expression is finite by Lemma 1.4.1.

<u>Theorem 4.7.1</u>: The function $M(t)$ is bounded on every finite interval and $P[Z(t) = \infty] = 0$.

Proof: The second assertion follows from the first, and the first assertion follows from Lemma 4.7.1 since $Z(t) \leq B(t)$ for all t and $B(t)$ is nondecreasing.

Remark: The conditions given here which are sufficient to insure the finiteness of $M(t)$ are quite different and, in a sense, less stringent than corresponding conditions for the Bellman-Harris process. A necessary condition for the finiteness of the mean of the Bellman-Harris process is $G(0)h'(1) \leq 1$ (Sevast'yanov (1967)). The conditions used in Theorem 4.7.1 to show $M(t) < \infty$ were essentially $\Delta(0) = 0$ and $\Delta(t) < \infty$, which are completely independent of G and include the possibility that $h'(1) = \infty$.

The standard technique (used by Harris (1963), Ney (1964a), and Mode (1968a) for example) for studying the moments of a branching process is to first obtain an integral equation for the generating function of the process and then to differentiate it successively to obtain integral equations for the moments. However, it does not seem possible to derive a useful integral equation for the generating function in the present case without sacrificing a good deal of generality. (Such an equation is derived in Section 4.10 for a special case.) One's ability to derive integral equations of this type seems to depend upon the existence of a "point of regeneration" in the process under investigation, which in the words of Harris (1963), is

"roughly speaking, . . . an event in the history of a process with the following property; if we know that the event has just occured and know the state of the process just after the occurrence, then knowledge of the

history of the process prior to the event is of no further help in predicting the future.

In an age-dependent branching process, the death of the ancestor seems to be a point of regeneration..."

It is fairly clear that the death of the ancestor is not a point of regeneration for the present process and no other event seems to meet the requirements. We may, however, derive a useful integral equation for $M(t)$, without working with the generating function.

Theorem 4.7.2: The function $M(t)$ satisfies the renewal equation

$$M(t) = 1 - G(t) + \int_0^t M(t-u)[1 - G(u)]d\Delta(u) \tag{4.7.1}$$

and $M(t)$ is the only solution of (4.7.1) that is bounded on every finite interval. Moreover, if we let

$$F(t) = \int_0^t [1 - G(u)]d\Delta(u) , \tag{4.7.2}$$

then $M(t)$ may be written as

$$M = 1 + F - G + \sum_{k=1}^{\infty} F^{k*} * [F-G] . \tag{4.7.3}$$

Proof: By applying Lemma 4.5.1 and using the notation introduced just before the lemma, we get

$$M(t) = \int Z(t,\omega)dP$$

$$= \int \sum_{j=1}^{N(t)} Z(t-t^{(j)}, \omega_j)dP + \int [1 - \delta(t-\ell_o)]dP .$$

An application of Fubini's theorem yields

$$M(t) = \int dP_o \{ \sum_{j=1}^{N(t)} \int Z(t-t^{(j)}, \omega_j)dP_o \} + 1 - G(t)$$

$$= \int dP_o \{ \sum_{j=1}^{N(t)} M(t-t^{(j)}) \} + 1 - G(t)$$

$$= E\{ \int_o^t M(t-u)dN(u) \} + 1 - G(t)$$

The last step follows immediately from the meaning of integration with respect to a step function. Lemma 4.7.2 (which appears directly following this proof) tells us that the "E" may be moved inside and we get

$$M(t) = 1 - G(t) + \int_o^t M(t-u)dE(N(u))$$

$$= 1 - G(t) + \int_o^t M(t-u)[1 - G(u)]d\Delta(u)$$

which proves (4.7.1). We already know that $M(t)$ is bounded on every finite interval and the uniqueness of such a solution follows from a result in Feller (1941). By Lemma 1.4.1 expression (4.7.3) is also bounded on every finite interval and one may verify that it satisfies (4.7.1) by direct substitution.

<u>Lemma 4.7.2</u>: If f is a non-negative Borel function, then

$$E\int_{a+}^{b} f(t)dN(t) = \int_{a+}^{b} f(t)dE[N(t)] .$$

This equation is also true if $N(t)$ is replaced by $K(t)$.

<u>Proof</u>: First suppose that f is an indicator function of the interval $(x,y]$ and $a \le x < y \le b$. Then

$$E\int_{a+}^{b} f(t)dN(t) = E\int_{x+}^{y} dN(t) = E[N(y)-N(x)]$$

and

$$\int_{a+}^{b} f(t)dE[N(t)]=\int_{x+}^{y} dE[N(t)] = E[N(y)]-E[N(x)] .$$

By considering the other possible arrangements of x and y, we conclude that the lemma is true whenever f is an indicator function of an interval $(x,y]$. The proof may obviously be extended to the

case where f is an indicator function of a set in the algebra formed by the class of all finite disjoint unions of intervals of the type $(x,y]$, along with complements of sets of this type. Let M be the class of all sets for which the lemma is true when f is an indicator function of a set from this class. Suppose $A_j, j = 1, 2 \ldots,$ is a monotone sequence of sets in M and $A_j \to A$ as $j \to \infty$. By repeated use of the bounded convergence theorem (I_A is the indicator function of the set A),

$$E\int_{a+}^{b} I_A dN(t) = E[\operatorname*{Lim}_j \int_{a+}^{b} I_{A_j} dN(t)]$$

$$= \operatorname*{Lim}_j [E\int_{a+}^{b} I_{A_j} dN(t)]$$

$$= \operatorname*{Lim}_j \int_{a+}^{b} I_{A_j} dE[N(t)]$$

$$= \int_{a+}^{b} I_A dE(N(t)) .$$

This proves that M is a monotone class and therefore M contains the Borel sets (Loeve(1963), page 60). By the usual procedure, the proof may now be extended to the case f is any non-negative Borel measurable function.

Several properties of $M(t)$ may be obtained directly from expression (4.7.3). Since $F(\infty) = h'(1)$, using Lemma 1.4.2, we see that if $h'(1) > 1$, then

$$\lim_{t\to\infty} M(t) = \infty \tag{4.7.4}$$

and if $h'(1) < 1$, then

$$\lim_{t\to\infty} M(t) = 0 \tag{4.7.5}$$

We shall see later from renewal theory that if $h'(1) = 1$ and the mean life-span is finite, then $M(t)$ converges to a finite constant as $t\to\infty$.

Expression (4.7.3) may also be used to study the monotone properties of $M(t)$.

<u>Theorem 4.7.3</u>: If

$$Y(t) \equiv \int_0^t [1 - G(u)]d\Delta(u) - G(t) \tag{4.7.6}$$

is nondecreasing (nonincreasing) in the interval $[0,t]$, then $M(t)$ is nondecreasing (nonincreasing) in this interval also.

<u>Proof</u>: Since $Y(t) = F(t) - G(t)$, the proof follows directly from

(4.7.3).

This condition for monotonicity of $M(t)$ seems to be applicable in quite a few situations. For example, if $G(t) = 1-e^{-\mu t}$ and $\Delta(t) = \lambda t$, then $Y(t) = e^{-\mu t}(\lambda-\mu)$ and by Theorem 4.7.3, $\lambda > \mu$ implies $M(t)$ is nondecreasing, $\lambda < \mu$ implies $M(t)$ is nonincreasing, and $\lambda = \mu$ implies $M(t) \equiv 1$.

Harris (1963) shows that in the Bellman-Harris process $M(t)$ is nondecreasing, equal to one, or nonincreasing depending on whether $h'(1)$ is greater than, equal to, or less than one. In the present situation, however, having $h'(1) \geq 1$ is a necessary, but not sufficient, condition for $M(t)$ to be nondecreasing. (The corresponding statement for $h'(1) \leq 1$ is true, also.) The necessity is implied by (4.7.5), since if $M(t)$ goes to zero, it cannot be increasing. The following example will illustrate the fact that this condition is not sufficient.

<u>Example</u>: Let $G(t) = 1-e^{-at}$ and

$$\Delta(t) = \begin{cases} 0 & 0 \leq t \leq r \\ b(t-r) & t > r \end{cases}$$

where a, b, and r are positive numbers and $b > ae^{ar}$. It follows that

$$Y(t) = e^{-at} - 1 \qquad 0 \le t \le r$$

and by Theorem 4.7.3 $M(t)$ is decreasing $0 \le t \le r$. But $h'(1) = b/ae^{ar} > 1$.

We now turn our attention to the asymptotic properties of $M(t)$. As in the Bellman-Harris process, the behaviour of $M(t)$ as $t \to \infty$ may be determined in most cases of interest from known results in renewal theory. However, there are exceptions that do not arise in the Bellman-Harris process.

<u>Definition 4.7.1</u> Let α be the unique solution (provided it exists) of the equation

$$\int_0^\infty e^{-\alpha t}[1 - G(t)]d\Delta(t) = 1 . \tag{4.7.7}$$

<u>Remark</u>: If $1 < h'(1) < \infty$ there will always be a positive α satisfying (4.7.7), However, (4.7.7) may or may not have a solution if $h'(1) = \infty$. If $h'(1) = 1$, the solution is obviously $\alpha = 0$. There may be no α satisfying (4.7.7) if $h'(1) < 1$, but if such an α exists it must be negative.

<u>Theorem 4.7.4</u>: (i) If $h'(1) > 1$ and (4.7.7) has a solution, then

$$M(t)e^{-\alpha t} \to b \quad \text{as} \quad t\to\infty \tag{4.7.8}$$

where

$$b = \frac{\int_0^\infty [1 - G(t)]e^{-\alpha t}dt}{\int_0^\infty te^{-\alpha t}[1 - G(t)]d\Delta(t)} \quad . \tag{4.7.9}$$

(ii) If $h'(1) = 1$ and

$$\int_0^\infty tdG(t) < \infty \, , \tag{4.7.10}$$

then

$$M(t) \to \frac{\int_0^\infty tdG(t)}{\int_0^\infty t[1 - G(t)]d\Delta(t)} \quad \text{as} \quad t\to\infty \tag{4.7.11}$$

where this limit is interpreted as 0 if the denominator is infinite.

(iii) If $h'(1) < 1$ and there exists an α satisfying (4.7.7) and

$$\int_0^\infty te^{-\alpha t}[1 - G(t)]d\Delta(t) < \infty \tag{4.7.12}$$

then the limit (4.7.8) again holds.

Proof: Part (ii) is a straight-forward application of Lemma 1.4.3 (iii). Parts (i) and (iii) follow from the same lemma in the same way as Lemma 1.3.1.

Remark i: It should be pointed out that there are cases of interest in which $h'(1) = \infty$, and for these cases the rate of growth may be either exponential or super-exponential. By super-exponential we mean that $M(t)e^{-\alpha t} \to \infty$ as $t \to \infty$ for each α, $0 \le \alpha < \infty$. The following two examples illustrate how this may happen.

Example 1: If $1 - G(t) = 1/t^2$ for $t \ge 1$ and $\Delta(t) = \lambda t^2/2$, then $h'(1) = \infty$, but $M(t)$ has an exponential rate of growth according to Theorem 4.7.4 (i).

Example 2: If $1 - G(t) = e^{-\alpha t}$ and $\Delta(t) = e^{t^2} - 1$ then, using (4.7.3), $M(t)e^{-ct} \to \infty$ as $t \to \infty$ for all c since in fact $[F(t)-G(t)]e^{-ct} \to \infty$ as $t \to \infty$ for all c.

Remark ii: Another possible behaviour of $M(t)$ which has no analogue in the Bellman-Harris process is that of having a growth rate which is less than exponential. If for the case $h'(1) = 1$ the numerator in (4.7.11) is infinite (corresponding to an infinite mean life-span) while the denominator is finite, then $M(t) \to \infty$, but the rate is less than

linear (See Theorem 4.7.5 (iii)).

Following R. A. Fisher (1930) we shall call α the Malthusian parameter of population growth. Actually, there seems to be a very close connection between α and the Malthusian parameter m defined by Fisher as the solution of

$$\int_0^\infty e^{-mx}\ell_x b_x dx = 1 \tag{4.7.13}$$

where ℓ_x is the probability of living to age x and b_x is the rate of reproduction at age x. Our $1 - G(x)$ has the same interpretation as ℓ_x and, provided $\Delta(x)$ is differentiable, $\Delta'(x) = \lambda(x)$ measures the rate of reproduction, so that (4.7.13) and (4.7.7) are completely analogous.

The following version of assertion (i) of Theorem 4.7.4 will be needed later.

<u>Lemma 4.7.3</u>: Suppose $h'(1) > 1$ and an α exists satisfying (4.7.7). If $\Delta(t)$ is differentiable with derivative $\lambda(t)$ and

$$\int_0^\infty \{[1 - G(u)]\lambda(u)\}^p du < \infty$$

for some $p > 1$, then

$$M(t) = be^{\alpha t}[1+O(e^{-\epsilon t})] \text{ as } t \to \infty \text{ for some } \epsilon > 0 . \tag{4.7.14}$$

Proof: This assertion follows directly from Lemma 3 of Harris (1963), page 162.

To close this section on first moments we now state some results about the means of $B(t)$ and $D(t)$. The proofs are all similar to others in this section.

Theorem 4.7.5: (i) The functions $M_B(t) = E[B(t)]$ and $M_D(t) = E[D(t)]$ satisfy the equations

$$M_B(t) = 1 + \int_0^t M_B(t-u)[1 - G(u)]d\Delta(u) ,$$

$$M_D(t) = G(t) + \int_0^t M_D(t-u)[1 - G(u)]d\Delta(u) .$$

(ii) If $h'(1) > 1$ and there is an α satisfying (4.7.7) then

$$M_B(t)e^{-\alpha t} \to b_B \quad \text{as} \quad t \to \infty ,$$

$$M_D(t)e^{-\alpha t} \to b_D \quad \text{as} \quad t \to \infty ,$$

where

$$b_B = \frac{1}{\alpha\int_0^\infty te^{-\alpha t}[1 - G(t)]d\Delta(t)}$$

and

$$b_D = b_B - b$$

(iii) If $h'(1) = 1$, then both $M_B(t)/t$ and $M_D(t)/t$ approach

$$\frac{1}{\int_0^\infty t[1 - G(t)]d\Delta(t)} \quad \text{as} \quad t \to \infty ,$$

where the limit is interpreted as 0 if the denominator is infinite.

(iv) Under the hypothesis of Lemma 4.7.3 (4.7.14) again holds with $M(t)$ replaced by $M_B(t)$ $(M_D(t))$ and b replaced by b_B (b_D).

4.8 Second moments

<u>Definition 4.8.1</u>: Let

$$M_2(t,\tau) = E[Z(t)Z(t+\tau)] , \quad t \geq 0 , \quad \tau \geq 0 . \tag{4.8.1}$$

Just as for the first moment, we shall study $M_2(t,\tau)$ by means of an integral equation which we shall derive without the aid of generating functions.

If we apply Theorem 4.5.1 to $M_2(t,\tau)$ we get

$$M_2(t,\tau) = 1 - G(t+\tau)$$

$$+ \int d\mathbb{P}[1 - \delta(t+\tau-\ell)] \sum_{i=1}^{N(t)} Z(t-t^{(i)}, \omega_i)$$

$$+ \int d\mathbb{P}[1 - \delta(t-\ell)] \sum_{i=1}^{N(t+\tau)} Z(t+\tau-t^{(i)}, \omega_i) \tag{4.8.2}$$

$$+ \int d\mathbb{P} \sum_{i=1}^{N(t)} Z(t-t^{(i)}, \omega_i) Z(t+\tau-t^{(i)}, \omega_i)$$

$$+ \int d\mathbb{P} \sum_{\substack{i=1 \\ i\neq j}}^{N(t)} \sum_{j=1}^{N(t+\tau)} Z(t-t^{(i)}\omega_i) Z(t+\tau-t^{(j)}, \omega_j) .$$

Our next task is to simplify this expression. Using the methods employed in the proof of Theorem 4.7.2 it can easily be shown that the second term on the right side of (4.8.2) is equal to

$$\int d\mathbb{P}_o[1 - \delta(t+\tau-\ell)] \int_o^t M(t-u)dN(u) . \tag{4.8.3}$$

Taking into account the relations among ℓ, $K(t)$, and $N(t)$, and applying Lemma 4.7.2, (4.8.3) becomes

$$\int dP_o[1 - \delta(t+\tau-\ell)]\{\delta(t-\ell)\int_o^{\ell} M(t-u)dK(u) + [1 - \delta(t-\ell)]\int_o^{t} M(t-u)dK(u)\}$$

$$= \int dP_o[1 - \delta(t+\tau-\ell)]\int_o^{t} M(t-u)dK(u)$$

$$= [1 - G(t+\tau)]\int_o^{t} M(t-u)d\Delta(u) \ . \qquad (4.8.4)$$

If we use the same types of arguments in simplifying the third term on the right side of (4.8.2) we get

$$\int dP_o[1 - \delta(t-\ell)]\sum_{i=1}^{N(t+\tau)} Z(t+\tau-t^{(i)}, \omega_i)$$

$$= \int dP_o[1 - \delta(t-\ell)]\delta(t+\tau-\ell)\int_o^{\ell} M(t+\tau-u)dK(u)$$

$$+ \int dP_o[1 - \delta(t-\ell)][1 - \delta(t+\tau-\ell)]\int_o^{t+\tau} M(t+\tau-u)dK(u)$$

$$= \int_{t+}^{t+\tau}[\int_o^{u} M(t+\tau-v)d\Delta(v)]dG(u)$$

$$+ [1 - G(t+\tau)]\int_o^{t+\tau} M(t+\tau-u)d\Delta(u) \ . \qquad (4.8.5)$$

There are no new methods used in simplifying the fourth term on the right side of (4.8.2). The final result is

$$\int_0^t M_2(t-u,\tau)[1-G(u)]d\Delta(u) \quad . \tag{4.8.6}$$

The last term on the right side of (4.8.2) is equal to

$$\int dP_0[\sum_{i=1}^{N(t)}\sum_{j=1}^{N(t+\tau)} M(t-t^{(i)})M(t+\tau-t^{(j)}) - \sum_{i=1}^{N(t)} M(t-t^{(i)})M(t+\tau-t^{(i)})]$$

$$= E[\int_{v=0}^{t+\tau}\int_{u=0}^{t} M(t-u)M(t+\tau-v)dN(u)dN(v)] - \int_0^t M(t-u)M(t+\tau-u)(1-G(u)d\Delta(u) \tag{4.8.7}$$

In general, it does not appear that the last expression can be simplified further. Putting all of these various pieces together, we have

<u>Theorem 4.8.1</u>: The function $M_2(t,\tau)$ is a solution of

$$M_2(t,\tau) = \int_0^t M_2(t-u,\tau)[1-G(u)]d\Delta(u)$$

$$+ [1-G(t+\tau)][1 + \int_0^t M(t-u)d\Delta(u) + \int_0^{t+\tau} M(t+\tau-u)d\Delta(u)]$$

$$+ \int_{t+}^{t+\tau}[\int_0^u M(t+\tau-v)d\Delta(v)]dG(u) + E[\int_0^{t+\tau}\int_0^t M(t-u)M(t+\tau-v)dN(u)dN(v)]$$

(over)

$$- \int_0^t M(t-u)M(t+\tau-u)[1 - G(u)]d\Delta(u) \ . \tag{4.8.8}$$

Although this equation appears quite formidable, for fixed τ it is simply a renewal equation for $M_2(t,\tau)$ and therefore results from renewal theory may be used to find an asymptotic formula for $M_2(t,\tau)$.

<u>Theorem 4.8.2</u>: If $1 < h'(1)$ and $h''(1) < \infty$ then

$$M_2(t,\tau)e^{-\alpha(2t+\tau)} \to \frac{b^2 A}{1 - \int_0^\infty e^{-2\alpha u}[1 - G(u)]d\Delta(u)} \quad \text{as} \quad t \to \infty \tag{4.8.9}$$

and this limit is uniform in τ. The number α is defined by (4.7.7), b is defined in Theorem 4.7.4, and

$$A = E\left(\int_0^\infty e^{-\alpha u}dN(u)\right)^2 - \int_0^\infty e^{-2\alpha u}[1 - G(u)]d\Delta(u) \ . \tag{4.8.10}$$

<u>Remark</u>: Since $h''(1) < \infty$ implies $h'(1) < \infty$, in this case there will always be an α satisfying (4.7.7).

<u>Proof</u>: The theorem is proved in the same way as Theorem 18.1 of Harris (1963), page 144. If we multiply both sides of (4.8.8) by $e^{-\alpha(2t+\tau)}$ and put

$$K(t,\tau) = M_2(t,\tau)e^{-\alpha(2t+\tau)}$$

$$\bar{m} = \int_0^\infty e^{-2\alpha t}[1 - G(t)]d\Delta(t) < 1 ,$$

$$\bar{G}(t) = 1/\bar{m} \int_0^t e^{-2\alpha u}[1 - G(u)]d\Delta(u) ,$$

then (4.8.8) takes the form

$$K(t,\tau) = \bar{f}(t,\tau) + \bar{m}\int_0^t K(t-u,\tau)d\bar{G}(u) , \qquad (4.8.11)$$

where $\bar{f}(t,\tau)$ is $e^{-\alpha(2t+\tau)}$ times everything on the right side of (4.8.8) except the first term. Since (4.8.11) is identical to (18.4) on page 145 of Harris (1963), the remainder of the proof is the same as his, provided we can show that

$$\bar{f}(t,\tau) \to A \quad \text{as} \quad t \to \infty$$

uniformly in τ.

It is straight-forward to show that none of the terms on the right side of (4.8.8) make any contribution to $\bar{f}(t,\tau)$ in the limit except

the last two, and that the last term times $e^{-\alpha(2t+\tau)}$ approaches

$$b^2\int_0^\infty e^{-2\alpha u}[1-G(u)]d\Delta(u) \quad \text{as} \quad t \to \infty ,$$

uniformly in τ. Thus we shall only prove that

$$e^{-\alpha(2t+\tau)}E\{\int_0^{t+\tau}\int_0^t M(t-u)M(t+\tau-v)dN(u)dN(v)\} \to$$

$$b^2E[\int_0^\infty e^{-\alpha u}dN(u)]^2 \quad \text{as} \quad t \to \infty ,$$

uniformly in τ. By an argument similar to the proof of Lemma 2.6.1 it can be shown that for each $\omega\epsilon\Omega$ such that $N(\omega) < \infty$,

$$\lim_{t\to\infty}\int_0^{t+\tau}[M(t+\tau-u)e^{-\alpha(t+\tau-u)}]e^{-\alpha u}dN(u,\omega)$$

$$= b\int_0^\infty e^{-\alpha u}dN(u,\omega)$$

and this limit is uniform in τ. Moreover, if C is an upper bound for $M(t)e^{-\alpha t}$, then

$$e^{-\alpha(2t+\tau)}\int_0^{t+\tau}\int_0^{t} M(t-u)M(t+\tau-v)\,dN(u,\omega)\,dN(v,\omega)$$

$$\leq C^2\left[\int_0^{\infty} e^{-\alpha u}\,dN(u,\omega)\right]^2$$

$$\leq C^2 N^2(\omega) ,$$

and this last term is integrable since, by hypothesis

$$h''(1) = E(N^2) - E(N) < \infty .$$

Therefore the dominated convergence theorem applies and the proof is complete.

The following stronger assertion will be useful in the next section.

<u>Lemma 4.8.1</u>: If the conditions of Lemma 4.7.3 are met and $h''(1) < \infty$, then

$$\frac{M_2(t,\tau)}{b^2 e^{\alpha(2t+\tau)}} = \frac{A}{1-\int_0^{\infty} e^{-2\alpha u}[1-G(u)]\,d\Delta(u)} + O(e^{-\varepsilon t})$$

for some $\varepsilon > 0$ independent of τ.

Proof: Using Lemma 4.7.3, the proof follows from Lemmas 4 and 5 on page 163 of Harris (1963), after tedious but straight-forward calculations.

Theorem 4.8.3: If $M_B(t,\tau) = E[B(t)\ B(t+\tau)]$ and $M_D(t,\tau) = E[D(t)D(t+\tau)]$ then both Theorem 4.8.2 and Lemma 4.8.1 are valid when $M_2(t,\tau)$ is replaced by $M_B(t,\tau)$ $(M_D(t,\tau))$ and b is replaced by b_B (b_D).

4.9 Convergence of $Z(t)/be^{\alpha t}$

The results of the last section lead immediately to a result concerning quadratic mean convergence of $Z(t)/be^{\alpha t}$.

Theorem 4.9.1: If $h'(1) > 1$ and $h''(1) < \infty$, then $Z(t)/be^{\alpha t}$ converges in mean square to a random variable W as $t \to \infty$ and

$$E(W) = 1 ,$$

$$\mathrm{Var}(W) = \frac{E(\int_0^\infty e^{-\alpha u} dN(u))^2 - 1}{1 - \int_0^\infty e^{-2\alpha u}[1 - G(u)]d\Delta(u)} .$$

Proof: The proof follows from Theorem 4.8.2 along the same lines as the proof of Theorem 2.7.1.

Remark: In almost all cases of interest the variance of W is positive. A sufficient condition for this to be true is

$$h(0) > 0$$

which says "it is possible for an individual to have zero offspring".

<u>Theorem 4.9.2</u>: With the hypothesis of Theorem 4.9.1, $B(t)/be^{\alpha t}$ $(D(t)/be^{\alpha t})$ converges in mean square to a random variable W_B (W_D) and

$$W = W_B - W_D \quad \text{a.s.} \tag{4.9.1}$$

If $G(t)$ and $\Delta(t)$ satisfy certain additional regularity conditions the convergence to these limit random variables is almost sure. Although this is a fairly deep result, Harris (1963) has already done the work for us. Here we have a situation in which we must first prove the result for $B(t)$ and $D(t)$ in order to obtain if for $Z(t)$.

<u>Theorem 4.9.3</u>: If $h''(1) < \infty$ and the conditions of Lemma 4.7.3 hold, then

$$B(t)/be^{\alpha t} \to W_B \quad \text{a.s.} \quad \text{as} \quad t \to \infty$$

$$D(t)/be^{\alpha t} \to W_D \quad \text{a.s.} \quad \text{as} \quad t \to \infty$$

$$Z(t)/be^{\alpha t} \to W \quad \text{a.s.} \quad \text{as} \quad t \to \infty .$$

Proof: By Theorem 4.8.3

$$E[B(t)/be^{\alpha t} - W_B]^2 = O(e^{-\epsilon t})$$

and therefore

$$\int_0^\infty E[B^{(t)}/be^{\alpha t} - W_B]^2 dt < \infty$$

Moreover, $B(t)$ is almost surely a finite nondecreasing step-function. Since these properties of $B(t)$ are precisly those used by Harris (1963) in proving Theorem 21.1, page 147, the assertion concerning $B(t)$ (and, for $D(t)$) follows from Harris' argument. The last part of the theorem now follows from (4.9.1).

In the next section we shall obtain some results about the distribution of W for an important special case.

4.10 Multiple births; The Poisson (θ,f,G) branching process

Recall from Section 4.2 that the definition of Ω makes provision for the possibility of multiple births. In terms of the counting process $K_I(t)$, a jump in a sample function of height j at time t means $\langle I \rangle$ gives birth to j offspring at this instant. Frequently however, it is desirable to separate the epochs at which births occur from the number of births occuring at each epoch by assuming the epochs of birth are generated by a counting process whose sample functions are all step

functions with unit steps, and the number of births occuring at each epoch is independent of this counting process. We shall indicate briefly how these ideas can be incorporated explicitly into the probability space defined in section 4.2.

For each $I \in \mathcal{J}$, let W_I be a sequence of the form $(w^{(1)}, w^{(2)}, \ldots)$, where each $w_I^{(j)}$ is an extended non-negative real number and $w_I^{(j)} < w_I^{(j+1)}$, $j = 1,2\ldots$ (unless $w_I^{(j)} = \infty$ and then we permit equality), and let V_I be a sequence of the form $(v_I^{(1)}, v_I^{(2)}, \ldots)$ where each $v_I^{(j)}$ is a non-negative integer. Each pair (W_I, V_I) determines uniquely a sequence T_I by the following assignment: Let

$$t_I^{(i)} = w_I^{(1)}, \quad 0 < i \le v_I^{(1)}$$

$$\vdots$$

$$t_I^{(i)} = w_I^{(k)}, \quad v_I^{(1)} + \ldots + v_I^{(k-1)} < i \le v_I^{(1)} + \ldots + v_I^{(k)}.$$

Moreover, each sequence T_I is determined in this way by a pair (W_I, V_I), except for those sequences T_I for which $t_I^{(j)} = t_I^{(j+1)} = \ldots < \infty$ for some integer j. But a sequence of this type corresponds to a sample function of $K(t)$ that is infinite at some finite value of t. Such events have probability zero since we are assuming $\Delta(t)$ is finite and therefore a probability measure on the set of ordered (W_I, V_I) will induce a measure on the space Ω_{T_I}. Let the probability on the set of all sequences V_I be the one determined by assuming the random

variables $v_I^{(j)}$, $j = 1,2...$, are independent with a common generating function $f(s)$. Furthermore, let there be an arbitrary probability measure defined on the set of all sequences W_I (the same for each I) and let the measure on the set of ordered pairs (W_I , V_I) be product measure. Let the counting function of the sequences W_I be denoted by $Q_I(t)$ (See Definition 4.3.1) and its generating function by $h_Q(s,t)$. It can be shown that

$$h_K(s,t) = h_Q(f(s),t) \tag{4.10.1}$$

and therefore

$$\Delta(t) = f'(1)\ \theta\ (t) \tag{4.10.2}$$

where $\theta(t)$ is the mean of $Q(t)$.

Although the procedure just described is applicable in a wide variety of situations, for the remainder of this chapter we shall study the following important special case.

<u>Definition 4.10.1</u>: If $Q(t)$ is a Poisson process with mean $\theta(t)$, then the process $Z(t)$ determined by $\theta(t)$, $f(s)$, and $G(t)$ is called a Poisson (θ,f,G) branching process.

<u>Remarks</u>: (i) In view of (4.10.2), in order to have the conditions on $\Delta(t)$ that were given in section 4.3, we must impose the same conditions

on $\theta(t)$ and also assume that $f'(1) < \infty$.

(ii) Since $\theta(t)$ is continuous and since $E[Q^2 t] = \theta(t) + \theta^2(t)$, it follows that $Q(t)$ is continuous in mean square. Therefore, by B on page 547 of Loeve (1963), almost all sample functions of $Q(t)$ are step-functions with unit jumps, so that a Poisson law is permissible on our sample space.

(iii) The Poisson (θ,f,G) process is of interest in its own right since by the remarks in section 4.1 it generalizes some well-known processes.

By (4.10.1) the results of Theorem 4.3.1 now become

$$h_N(s,t) = \int_0^t e^{\theta(u)(f(s)-1)} dG(u) + [1 - G(t)] e^{\theta(t)(f(s)-1)} , \qquad (4.10.3)$$

and

$$h(s) = \int_0^\infty e^{\theta(u)(f(s)-1)} dG(u) . \qquad (4.10.4)$$

With the added assumption of the Poisson (θ,f,G) process, several results in the last two sections may be sharpened. For example, the last two terms on the right side of (4.8.8) may be replaced by

$$[f'(1)]^2\int_0^t\int_0^u M(t-u)M(t+\tau-v)[1-G(u)]d\theta(v)d\theta(u)$$

$$+[f'(1)]^2\int_0^{t+\tau}\int_0^u M(t+\tau-u)M(t-v)[1-G(u)]d\theta(v)d\theta(u) \qquad (4.10.5)$$

$$+f''(1)\int_0^t M(t-u)M(t+\tau-u)[1-G(u)]d\theta(u) ,$$

and the constant A defined by (4.8.10) can be written

$$A = 2[f'(1)]^2\int_0^\infty\int_0^u e^{-\alpha(u+v)}[1-G(u)]d\theta(v)d\theta(u) \qquad (4.10.6)$$

$$+f''(1)\int_0^\infty e^{-2\alpha u}[1-G(u)]d\theta(u) .$$

We shall illustrate the method of proof of these results by deriving (4.10.6). By the definition of N, (4.8.10) may be written in the form

$$A = E[\sum_{k=1}^{N}\sum_{j=1}^{N} e^{-\alpha(t^{(i)}+t^{(j)})}] - m\int_0^\infty e^{-2\alpha u}[1-G(u)]d\theta(u) . \qquad (4.10.7)$$

If we define

$$Q_I = Q_I(\ell_I) \qquad (4.10.8)$$

the first term of (4.10.7) may be rewritten as

$$E\{\sum_{k=1}^{Q}\sum_{j=1}^{Q} v^{(i)}v^{(j)}e^{-\alpha(w^{(i)}+w^{(j)})}\}$$

$$= E\{\sum_{k=1}^{Q}\sum_{\substack{j=1\\ i\neq j}}^{Q} v^{(i)}v^{(j)}e^{-\alpha(w^{(i)}+w^{(j)})}\} + E\{\sum_{k=1}^{Q}[v^{(k)}]^2e^{-2\alpha w^{(k)}}\} . \quad (4.10.9)$$

Using Lemma 4.7.2 and straight-forward techniques we can show that

$$E\{\sum_{k=1}^{Q}[v^{(k)}]^2e^{-2\alpha w^{(k)}}\} = E[v^2]\int_0^\infty e^{-2\alpha u}[1-G(u)]d\theta(u) \quad (4.10.10)$$

which combines with the second term on the right side of (4.10.7) to yield the second term on the right side of (4.10.6). Furthermore, the first term on the second line of (4.10.9) may be shown to be equal to

$$2[f'(1)]^2E\sum_{k=2}^{Q}\sum_{j=1}^{k-1} e^{-2(w^{(k)}+w^{(j)})}$$

$$= 2[f'(1)]^2\int_0^\infty E\{\int_0^y\int_0^{u-} e^{-\alpha(u+v)}dQ(v)dQ(u)\}dG(y) .$$

By Lemma 4.10.1 which follows this discussion the last expression becomes

$$2[f'(1)]^2\int_0^\infty\{\int_0^y\int_0^u e^{-\alpha(u+v)}d\theta(v)d\theta(u)\}dG(y) .$$

If we interchange the order of integration of y and u in this integral, we get the first term on the right side of (4.10.6) which completes the derivation of this formula.

<u>Lemma 4.10.1</u>: If $f(u)$ is a non-negative Borel function, then

$$E[\int_0^t\int_0^{u-} f(u)f(v)dQ(u)dQ(v)] = \int_0^t\int_0^u f(u)f(v)d\theta(v)d\theta(u) . \qquad (4.10.11)$$

(We are still assuming, of course, that $Q(t)$ is a Poisson process.)

<u>Proof</u>: First suppose that f is an indicator function of an interval of the form $(a,b]$ and $0 \le a < b \le t$. Since each sample function of $Q(t)$ is almost surely a step function with unit steps and with only a finite number of steps in each finite interval, we may assume for a fixed sample path that

$$w^{(m)} \le a < w^{(m+1)} < \cdots < w^{(m+k)} \le b < w^{(m+k+1)} .$$

Then

$$E\{\int_0^t \int_0^{u-} f(u)f(v)dQ(v)dQ(u)\}$$

$$= E\{\int_{a+}^{b} \int_{a+}^{u-} dQ(v)dQ(u)\}$$

$$= E\{\int_{a+}^{b} [Q(u-) - Q(a)]dQ(u)\}$$

$$= E\{m + m + 1 + \cdots + m + k + 1\} - E\{Q(a)[Q(b) - Q(a)]\}$$

$$= \tfrac{1}{2} E\{m + k - 1)(m + k) - (m - 1)\, m\} - E\{Q(a)[Q(b) - Q(a)]\}$$

$$= \tfrac{1}{2} E\{Q^2(b) - Q(b) - Q^2(a) + Q(a)\} - E\{Q(a)[Q(b) - Q(a)]\} .$$

If we recall that a Poisson process has independent increments and that the variance of a Poisson-distributed random variable is equal to its mean, this last expression becomes

$$\tfrac{1}{2}[\theta^2(b) + \theta^2(a)] - \theta(a)\theta(b) .$$

But, on the other hand,

$$\int_0^t \int_0^u f(u)f(v)d\theta(u)d\theta(v)$$

$$= \int_a^b [\theta(u) - \theta(a)]d\theta(u)$$

$$= \tfrac{1}{2}[\theta^2(b) + \theta^2(a)] - \theta(a)\theta(b) .$$

Therefore, the lemma is true if f is an indicator function of a half-open interval, and the proof can now be extended in exactly the same way as Lemma 4.7.2.

We shall now demonstrate that for a Poisson (θ,f,G) branching process, it is possible to derive a useful integral equation for the generating function (even without the apparent existence of a "point of regeneration"). Let

$$F(s,t) = E[s^{Z(t)}], |s| \leq 1, t \geq 0 . \qquad (4.10.12)$$

<u>Theorem 4.10.1</u>: For a Poisson (θ,f,G) process, the generating function $F(s,t)$ satisfies the integral equation

$$F(s,t) = \int_0^t \exp[\int_0^y \{f[F(s,t-u)] - 1\}d\theta(u)]dG(y) \qquad (4.10.13)$$

$$+ s[1 - G(t)]\exp[\int_0^t \{f[F(s,t-u)] - 1\}d\theta(u)] .$$

<u>Proof</u>: Using Theorem 4.5.1, the fact that

$$s^{[1 - \delta(t-\ell)]} = \delta(t-\ell) + [1 - \delta(t-\ell)]s , \qquad (4.10.14)$$

and arguments similar to those used earlier in this section and in section 2.4 we get

$$F(s,t) = \int dP_o \delta(t-\ell) \prod_{j=1}^{N(t)} F(s,t-t^{(j)})$$

$$+ \int dP_o s[1 - \delta(t-\ell)] \prod_{j=1}^{N(t)} F(s,t-t^{(j)}) . \qquad (4.10.15)$$

From this point we will work only with the first term on the right side of (4.10.15) since the evaluation of the second is similar. The first term is equal to

$$\int_{[\ell \le t]} dP_o \prod_{j=1}^{N(t)} F(s,t-t^{(j)})$$

$$= \int_{[\ell \le t]} dP_o \prod_{j=1}^{K(\ell)} F(s,t-t^{(j)})$$

$$= \int_o^t E\{\prod_{j=1}^{K(y)} F(s,t-t^{(j)})\}dG(y)$$

$$= \int_o^t E\{\prod_{j=1}^{Q(y)} F^{v^{(j)}}(s,t-w^{(j)})\}dG(y)$$

$$= \int_o^t E\{\prod_{j=1}^{Q(y)} f[F(s,t-w^{(j)})]\}dG(y)$$

$$= \int_o^t \sum_{k=0}^{\infty} P[Q(y) = k]E\{\prod_{j=1}^{k} f[F(s,t-w^{(j)})]|Q(y) = k\}dG(y) \tag{4.10.16}$$

But it is a property of non-homogeneous Poisson processes (Parzen (1962), page 143), that, conditioned on the event $Q(y) = k$, the times $w^{(1)},\ldots, w^{(k)}$ are distributed as order statistics corresponding to k independent random variables with common distribution function

$$H(u) = \frac{\theta(u)}{\theta(y)}, \quad 0 \le u \le y . \tag{4.10.17}$$

Using this fact, (4.10.16) becomes

$$\int_0^t \sum_{k=0}^{\infty} \frac{e^{-\theta(y)}[\theta(y)]^k}{k!} \left[\frac{k!}{[\theta(y)]^k} \int_0^y \int_0^{u_k} \cdots \int_0^{u_2} \prod_{j=1}^{k} f[F(s,t-u_j]d\theta(u_1)\cdots d\theta(u_k)\right] dG(y)$$

$$= \int_0^t \sum_{k=0}^{\infty} e^{-\theta(y)} \left[\frac{\int_0^y f[F(s,t-u)]d\theta(u)}{k!}\right]^k dG(y)$$

$$= \int_0^t \exp\left[\int_0^y \{f[F(s,t-u)]d\theta(u)\right] dG(y) .$$

Upon performing similar operations on the other term in (4.10.15), the proof is complete.

As a first application of the integral equation (4.10.13), we shall demonstrate the relationship of Poisson (θ,f,G) branching processes to Markov branching processes.

<u>Theorem 4.10.2</u>: If in a Poisson (θ,f,G) process, $\theta(t) = \lambda t$ and $G(t) = 1 - e^{-\mu t}$, then $F(s,t)$ is the generating function of a temporally homogeneous Markov branching process.

<u>Remark</u>: Presumably, not only the generating functions, but the processes are the same, also.

<u>Proof</u>: In this special case, (4.10.13) can be put in the form

$$F(s,t) = \mu e^{-t(\lambda+\mu)} \int_0^t \exp\left\{\lambda \int_x^t f[F(s,y)]dy + x(\lambda+\mu)\right\} dx \tag{4.10.18}$$

$$+ s \exp\left\{\lambda \int_0^t f[F(s,y)]dy - t(\lambda+\mu)\right\}$$

If we differentiate (4.10.18) with respect to t and compare the result with the undifferentiated equation, we obtain

$$\frac{\partial F(s,t)}{\partial t} = \mu + \lambda f[F(s,t)]F(s,t) - (\lambda+\mu)F(s,t) \ . \tag{4.10.19}$$

Upon putting $b = \lambda + \mu$ and

$$h^*(s) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} s f(s) \ ,$$

(4.10.19) becomes

$$\frac{\partial F(s,t)}{\partial t} = b[h^*[F(s,t)] - F(s,t)] \ . \tag{4.10.20}$$

Moreover, we see from (4.10.13) that $F(s,0) = s$. Since (4.10.20) is the equation satisfied by the generating function of a temporally homogeneous Markov branching process (Harris (1963), page 106) and (4.10.21) is the proper initial condition, the theorem is proved.

Utilizing techniques similar to those employed in section 2.8, we may use the integral equation (4.10.13) for $F(s,t)$ to derive an integral equation for the characteristic function of W .

<u>Theorem 4.10.3</u>: If for the Poisson (θ,f,G) process, $h'(1) > 1$ and $h''(1) < \infty$, then the characteristic function $\varphi(s)$ of W satisfies the integral equation

$$\varphi(s) = \int_0^\infty \exp\left[\int_0^y \{f(\varphi(se^{-\alpha u})) - 1\} d\theta(u)\right] dG(y) \tag{4.10.22}$$

The proof of this theorem is similar to that of Theorem 2.8.1 and will be omitted.

We may use this integral equation to study the distribution of W , as we did with a similar equation in section 2.9. The proof of the next theorem follows the same general pattern as the proof of Lemma 2.9.1.

<u>Theorem 4.10.3</u>: If for a Poisson (θ,f,G) process $h'(1) > 1$, $h''(1) < \infty$, and $f(s) = s$, then the distribution of W is continuous except for a jump of height q at the origin.

<u>Proof</u>: The moment-generating function

$$\phi(s) = E(e^{-sW})$$

satisfies the integral equation (4.10.22), also, so by letting $s \to \infty$ in (4.10.22) we get

$$\phi(\infty) = \int_0^\infty e^{(\phi(\infty)-1)\theta(y)} dG(y) .$$

Since h is convex the only roots of this equation between zero and one are 1 and q. But since $E(W) = 1$, $P[W = 0] < 1$, so therefore

$$P[W = 0] = \phi(\infty) = q .$$

It follows (Lukacs (1960)) that the proof will be complete if we can show that

$$\varphi^*(s) = E[e^{isW} | W > 0]$$

approaches zero as $s \to \pm \infty$.

Since

$$h(0) = \int_0^\infty e^{-\theta(u)} dG(u) > 0 ,$$

it follows from a remark made just after Theorem 4.9.1 that $\mathrm{Var}(W) > 0$. Therefore, as in the proof of Lemma 2.9.1, $|\varphi(s)| < 1$, provided $|s|$ is small enough and $s \neq 0$.

We shall next show that

$$\operatorname*{Lim\,sup}_{s\to\infty} |\varphi(s)| < 1 .$$

If this is not the case, there exists numbers s_3 and $d > 0$ such that

$$|\varphi(s_3)| < 1 - d$$

and

$$|\varphi(s)| < 1 , \quad 0 < s \leq s_3 .$$

Let s_1 and s_2 be the first points to the left and right, respectively, of s_3 such that

$$|\varphi(s_1)| = |\varphi(s_2)| = 1 - d ,$$

and let

$$A = (1/\alpha)\mathrm{Log}(s_2/s_1)$$

Using (4.10.22) we have

$$|\varphi(s_2)| \le \int_0^A \exp\left\{\int_0^y [\varphi(s_2 e^{-\alpha u}) - 1] d\theta(u)\right\} dG(y) + 1 - G(A)$$

and therefore

$$1 - d \le \int_0^A e^{-d\theta(y)} dG(y) + 1 - G(A) ,$$

which can be put in the form

$$1 \ge (-1/d)\left[\int_0^A e^{-d\theta(y)} dG(y) - G(A)\right] .$$

Expanding $e^{-d\theta(y)}$ in a Maclaurin series, we get

$$1 \geq \int_0^A \theta(y)dG(y) - (1/d)\int_0^A \sum_{k=2}^{\infty} \frac{[-d\theta(y)]^k}{k!} dG(y) . \qquad (4.10.23)$$

If $d \to 0$ then $s_1 \to 0$ while s_2 increases so that $A \to \infty$. Therefore if $d \to 0$ in (4.10.23) the last term approaches zero and we get

$$1 \geq \int_0^\infty \theta(y)dG(y) = h'(1) .$$

This is a contradiction since, by hypothesis, $h'(1) > 1$. Since

$$\varphi(s) = q + (1 - q)\varphi^*(s) \qquad (4.10.24)$$

it follows that

$$\operatorname*{Lim\,sup}_{s\to\infty} |\varphi^*(s)| < 1 ,$$

also. Therefore there exists an s_o and a $d > 0$ such that

$$|\varphi^*(s)| < 1 - d$$

for all $s \geq s_o$. For an arbitrary ϵ, choose A so large that

$$\frac{1 - G(A)}{1 - q} < \epsilon$$

and then choose s so large that

$$se^{-\alpha A} > s_o .$$

Since from (4.10.24) and (4.10.22), $\varphi^*(s)$ satisfies the integral equation

$$(1-q)\varphi^*(s) = \int_o^\infty e^{(q-1)\theta(y)} \exp\left[\int_o^y (1-q)\varphi^*(se^{-\alpha u})d\theta(u)\right] dG(y) - q$$

so that

$$\varphi^*(s) = (1/1-q)\int_o^\infty e^{(q-1)\theta(y)} \sum_{k=1}^\infty \left[\int_o^y (1-q)\varphi^*(se^{-\alpha u})d\theta(u)\right]^k dG(y) .$$

If we let

$$\psi(s) = \sup_{t \geq s} |\varphi^*(t)|$$

we get

$$\psi(s) < (1-d)\int_0^A \left\{ e^{(q-1)\theta(y)}\theta(y)\left\{1 + \sum_{k=2}^{\infty} \frac{\left[\int_0^y (1-q)\psi(se^{-\alpha u})d\theta(u)\right]^{k-1}}{k!}\right\} dG(y) + \epsilon$$

$$< (1-d)\int_0^\infty e^{(q-1)\theta(y)}\theta(y)dG(y) + (1-d)\psi(se^{-\alpha A})\int_0^\infty e^{(q-1)\theta(y)}\left[\sum_{k=2}^{\infty}\frac{(1-q)^{k-1}\theta^k(y)}{k!}\right]$$

$$\cdot\, dG(y) + \epsilon\ .$$

If we let

$$B = \int_0^\infty e^{(q-1)\theta(y)}\theta(y)dG(y)$$

it follows that

$$\psi(s) < (1-d)B + \frac{(1-d)\psi(se^{-\alpha A})}{1-q}\int_0^\infty e^{(q-1)\theta(y)}\left[e^{(1-q)\theta(y)} - 1 - (1-q)\theta(y)\right]dG(y) + \epsilon$$

or

$$\psi(s) < (1-d)B + (1-d)(1-B)\psi(se^{-\alpha A}) + \epsilon\ . \qquad (4.10.25)$$

The remainder of the proof may be found in Ney (1964b). By iterating this last inequality, we obtain

$$\psi(se^{n\alpha A}) < \frac{(1-d)B + \epsilon}{1-(1-d)(1-B)} + [(1-d)(1-B)]^n \psi(s) \quad n = 1,2 \ldots \quad (4.10.26)$$

Since

$$\frac{B}{1-(1-d)(1-B)} < 1 ,$$

we may choose ϵ such that

$$\frac{(1-d)B + \epsilon}{1-(1-d)(1-B)} < (1-d)\ C$$

where $C < 1$. Therefore there exists an s_1 such that for $s \geq s_1$,

$$\psi(s) < C(1-d) .$$

If we repeat the earlier argument with $1-d$ replaced by $C(1-d)$, we arrive at (4.10.26) with $1-d$ replaced by $C(1-d)$. Since

$$\frac{B}{1-C(1-d)(1-B)} < \frac{B}{1-(1-d)(1-B)}$$

it follows that ϵ may now be chosen such that

$$\frac{(1-d)CB + \epsilon}{1-C(1-d)(1-B)} < (1-d)C^2 .$$

Therefore, there exists an s_2 such that for $s \geq s_2$

$$\psi(s) \leq (1-d)\ C^2 .$$

We now see by induction that

$$\lim_{s\to\infty} \psi(s) = 0 .$$

Likewise, we can show that

$$\lim_{s\to -\infty} |\varphi(s)| = 0$$

and this completes the proof.

Chapter V

OTHER MODELS

5.1 A model in which family size affects longevity

Several authors (Waugh (1961) and Sevast'yanov (1964), for example) have considered branching processes in which the number of offspring depends upon the length of life of the parent. In this section a process is studied in which this situation is reversed, so that the length of life of an individual depends upon the size of his family. We shall suppose that if an individual is one of k offspring then the distribution of his lifespan is $G_k(t)$. In other respects, the process is assumed to be identical to the Bellman-Harris process.

In order to study the moments of the process we must come to grips with countably infinite systems of integral equations. Since the renewal theorems in Chapter I are not applicable, other techniques must be used to obtain asymptotic results for the moments. For the case $m > 1$, using complex variable techniques we find that the first two moments are again asymptotically exponential, as they are for the Bellman-Harris process, but the definitions of the constants involved are somewhat more complicated.

Perhaps this process can be used to study animal populations in which a child has less chance of surviving if it comes from a large litter. This seems to be a reasonable situation since in larger litters each child receives a proportionately smaller share of the mother's care. Even a better model for such an application might be obtained by modifying the birth-and-death process of Chapter IV just as we are modifying the Bellman-Harris process in this section. The technicalities encountered in such a

process would differ only in details from the one being studied here.

We will require that each $G_k(t)$, $k = 1,2,\cdots$, be a distribution with its mass concentrated on the non-negative real line and that there is a distribution function $H(t)$ with $H(0) = 0$ such that $H(t) \geq G_k(t)$, $k = 1,2,\cdots$. The reader is referred to the example preceeding Theorem 3.2.1 for an example of an H and a set of G_k's that satisfy this condition.

An important role will be played by the functions $G_{jn}(t)$, $j = 1,2,\cdots$, $n = 1,2,\cdots$, defined by

$$G_{j1}(t) = G_j(t), \text{ and } G_{j(n+1)}(t) = \sum_{k=1}^{\infty} kp_k G_{kn} * G_j(t)$$

where $*$ stands for the operation of convolution and p_n is the probability that an individual has n offspring.

<u>Lemma 5.1.1</u>: $G_{jn}(t) \leq m^{n-1}H^{n*}(t)$ for $j = 1,2,\cdots$, and $n = 1,2,\cdots$.

<u>Proof</u>: The proof is by induction. For $n = 1$ it is obvious. Assuming the lemma is true for a fixed $n \geq 1$ and for all j, we have

$$G_{j(n+1)}(t) = \Sigma kp_k \int_0^t G_{kn}(t-u)dG_j(u) \leq m^{n-1}\Sigma kp_k \int_0^t H^{n*}(t-u)dG_j(u)$$

$$= m^{n-1}\Sigma kp_k \int G_j(t-u)dH^{n*}(u)$$

$$\leq m^n H^{(n+1)*}(t),$$

which completes the proof.

Since the discussion given here deals with only the first two moments, it is more convenient not to introduce a generating function for the process. If $Z(t)$ is the number of individuals alive at time t and v denotes the number of offspring of the initial individual's parent, the techniques found in the proof of Theorem 4.7.2 can be used to show that $M(t,k) = E[Z(t)|v=k]$ satisfies

$$M(t,k) = 1 - G_k(t) + \sum_{n=1}^{\infty} np_n \int_0^t M(t-u,n)dG_k(u). \tag{5.1.1}$$

<u>Theorem 5.1.1</u>: Equation (5.1.1) possesses a unique solution among the functions satisfying

$$|M(t,k)| \le Ae^{rt} \tag{5.1.2}$$

where A and r are positive constants independent of k, and this solution is given by

$$M(t,k) = 1 + (m-1) \sum_{j=1}^{\infty} G_{kj}(t). \tag{5.1.3}$$

<u>Proof</u>: One can easily verify that (5.1.3) is a solution of (5.1.1). By Lemma 5.1.1

$$M(t,k) \le 1 + (m-1) \sum_{j=1}^{\infty} m^{j-1} H^{j*}(t),$$

and so the exponential bound follows from Lemma 1.3.1 . The proof of uniqueness is similar to that given in Theorem 3.5.1 and will be omitted.

As mentioned earlier, the results on renewal theory in section 1.4 do not seem to be applicable to equation (5.1.1) and consequently other methods must be used to obtain asymptotic results for $M(t,n)$. We shall follow the method outlined in 7.10 of Bellman and Cooke (1963), (also used by Harris (1963) in proving Lemma 3, page 162), and employ complex variable techniques to obtain such a result. Let λ be a complex number and set

$$M^*(\lambda,k) = \int_0^\infty e^{-\lambda t} M(t,k)dt \tag{5.1.4}$$

and

$$G_k^*(\lambda) = \int_0^\infty e^{-\lambda t} dG_k(t) . \tag{5.1.5}$$

In view of (5.1.2) the integrals (5.1.4) converge for all λ in some half-plane $\mathrm{Re}(\lambda) > \beta$, where β is a constant independent of k . Taking ordinary Laplace transforms of (5.1.) and using the fact that

$$\int_0^\infty e^{-\lambda t} G(t)dt = \frac{1}{\lambda}\int_0^\infty e^{-\lambda t} dG(t) ,$$

we obtain

$$M^*(\lambda,k) = \lambda^{-1}[1 - G_k^*(\lambda)] + \sum_n np_n M^*(\lambda,n)G_k^*(\lambda) . \tag{5.1.6}$$

Let us now introduce the functions

$$\overline{M}(\lambda,k) = \frac{1 + (m-1)G_k^*(\lambda) - \varphi(\lambda)}{\lambda[1 - \varphi(\lambda)]} \tag{5.1.7}$$

where

$$\varphi(\lambda) = \Sigma np_n G_n^*(\lambda) . \tag{5.1.8}$$

By substituting $\overline{M}(\lambda,k)$ for $M^*(\lambda,t)$ in (5.1.6) we see that $\overline{M}(\lambda,k)$ is a solution of (5.1.6). (In the special case $p_n = 0$ for $n > 3$, (5.1.6) becomes a system of three equations and three unknowns. If one solves this special case, he is led to predict the general solution (5.1.7).) Since the solution of (5.1.7) may not be unique, we do not know, a priori, that $\overline{M}(\lambda,k) = M^*(\lambda,k)$ or even that $\overline{M}(\lambda,k)$ is a Laplace transform. Before proving these facts, let us first note that the function $\varphi(\sigma)$ defined in (5.1.8) is, for σ real, a continuous, decreasing function and that $\lim_{\sigma\to\infty}\varphi(\sigma) = 0$. If $\varphi(0) = m > 1$, there exists a unique $\alpha > 0$ such that

$$\varphi(\alpha) = 1 . \tag{5.1.9}$$

As we shall see, this α will be the Malthusian parameter of population

growth.

Lemma 5.1.2: $\overline{M}(\lambda,k)$ is a Laplace transform, $k = 1,2,\cdots$.

Proof: By a result on page 31 of Ditkin and Prudnikov (1965), it will suffice to show that $\overline{M}(\sigma+i\tau)$ is analytic in some half-plane $\sigma \geq \sigma_0$, and that

$$\sup_{\sigma\geq\sigma_0}\int_{-\infty}^{\infty}|\overline{M}(\sigma+i\tau,k)|^2 d\tau < \infty .$$

Let us choose $\sigma_0 > \alpha$ where α is defined by (5.1.9). (If $m \leq 1$, set $\alpha = 0$.) Then for any $\lambda = \sigma + i\tau$ such that $\sigma \geq \sigma_0$, $|\varphi(\lambda)| \leq \varphi(\sigma_0) < 1$, so that $1 - \varphi(\sigma+i\tau)$ has no zeros in the half-plane $\sigma \geq \sigma_0$. Since the series $\Sigma np_n G_n^*(\sigma+i\tau)$ converges uniformly for $\sigma \geq 0$, $\varphi(\sigma+i\tau)$ is analytic in the half-plane, also. Moreover,

$$\sup_{\sigma\geq\sigma_0}\int_{-\infty}^{\infty}\left[\frac{(m-1)\overline{G}_k(\sigma+i\tau)}{(\sigma+i\tau)[1-\varphi(\sigma+i\tau)]}\right]^2 d\tau$$

$$\leq \left[\frac{m-1}{1-\varphi(\sigma_0)}\right]^2 \int_{-\infty}^{\infty}\left[\frac{1}{\sigma_0+i\tau}\right]^2 d\tau < \infty ,$$

which completes the proof, since

$$\overline{M}(\lambda,k) = \lambda^{-1} + \frac{(m-1)\overline{G}_k(\lambda)}{\lambda[1-\varphi(\lambda)]} .$$

In Theorem 5.1.2 it will be shown that under appropriate conditions the inverse Laplace transform of $\overline{M}(t,n)$ (the existence of which is guaranteed by Lemma 5.1.2) is of exponential order. Since this inverse Laplace transform also must satisfy (5.1.1), and since (5.1.1) has a unique solution of exponential order, it follows that, under the conditions of Theorem 5.1.2,

$$\overline{M}(t,k) = M^*(t,k) . \tag{5.1.10}$$

We are now ready to obtain an asymptotic formula for $M(t,k)$.

<u>Theorem 5.1.2</u>: Suppose $m > 1$ and $G_k(t)$ has a density $g_k(t)$ and $R_k(t)$ is defined by

$$M(t,k)e^{-\alpha t} = b_k + R_k(t) \tag{5.1.11}$$

where α is determined by (5.1.9) and

$$b_k = \frac{(m-1)G_k^*(\alpha)}{\alpha \Sigma n p_n \int_0^\infty t e^{-\alpha t} g_n(t)dt} \tag{5.1.12}$$

(i) If

$$\int_0^\infty |g_k(t)|^{p_k} dt < \infty \tag{5.1.13}$$

for some $p_k > 1$, $k = 1,2,\cdots$, then

$$|R_k(t)| \le A_k e^{-\epsilon t} \tag{5.1.14}$$

for some positive numbers A_k and ϵ, and ϵ does not depend on k.

(ii) If

$$\int_0^\infty |g_k(t)|^p dt \le B < \infty \tag{5.1.15}$$

for some $p > 1$, then

$$|R_k(t)| \le Ae^{-\epsilon t} \tag{5.1.16}$$

for some positive numbers A and ϵ, neither of which depend upon k.

<u>Proof</u>: Clearly α is the only real root of (5.1.9) and if $\sigma + i\tau$ is a complex root, then $\sigma > \alpha$, since

$$1 \le \Sigma n p_n \left|\int_0^\infty e^{-(\sigma+i\tau)t} g_n(t)dt\right| < \Sigma n p_n \int_0^\infty e^{-\sigma t} g_n(t)dt = \varphi(\sigma) .$$

If $\overline{N}(t,k)$ denotes the inverse Laplace transform of $\overline{M}(t,k)$, then by the inversion formula for Laplace transforms

$$\overline{N}(t,k) = \lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^{T} e^{t(x_2+i\tau)}\overline{M}(x_2+i\tau)d\tau \tag{5.1.17}$$

for any $x_2 > \alpha$. It is easy to see from (5.1.7) that $e^{\lambda t}\overline{M}(\lambda,k)$ has a simple pole at $\lambda = \alpha$ and the residue at this pole is

$$b_k e^{\alpha t} \tag{5.1.18}$$

Now choose x_1 such that α is the only root of $\varphi(\sigma + i\tau) = 1$ in the half-plane $\sigma \geq x_1$ and consider the rectangle pictured in Figure 2 formed by the vertical lines $\sigma = x_1$, $\sigma = x_2$, and the horizontal lines $\tau = T$, $\tau = -T$.

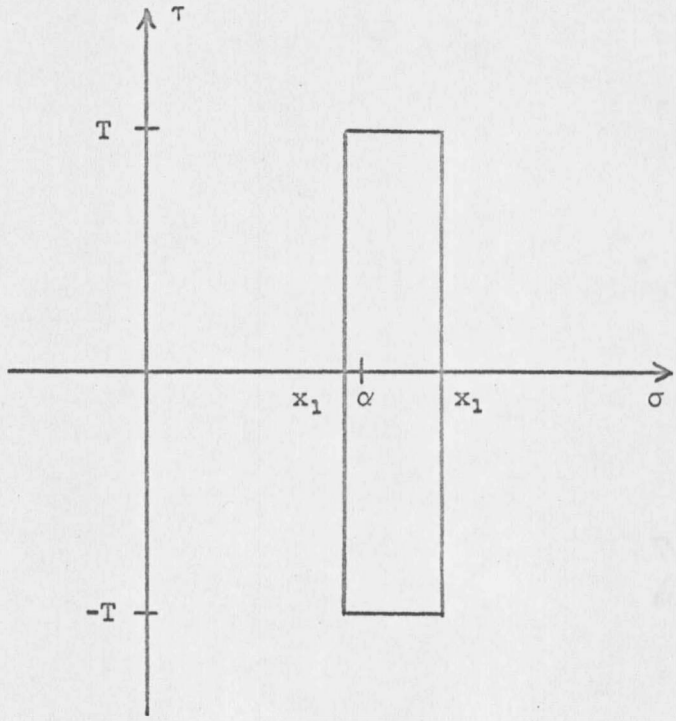

Figure 2

Contour of integration

From (5.1.17), (5.1.18), and the Cauchy integral formula, it follows that

$$\overline{N}(t,k) = b_k e^{\alpha t} + \lim_{T\to\infty}\left\{\frac{1}{2\pi i}\int_{-T}^{T} e^{(x_1+i\tau)t}\overline{M}(x_1+i\tau)d\tau\right.$$

$$+\frac{1}{2\pi i}\int_{x_1}^{x_2} e^{(\sigma+iT)t}\overline{M}(\sigma+iT)d\sigma \qquad (5.1.19)$$

$$\left.+\frac{1}{2\pi i}\int_{x_2}^{x_1} e^{(\sigma-iT)t}\overline{M}(\sigma-iT)d\sigma\right\} .$$

The proof of assertion (i) will be completed by showing that the last two integrals in (5.1.19) approach zero as $T \to \infty$ and that the first integral is $0(e^{x_1 t})$.

First of all, by the Riemann-Lebesgue lemma (Bellman and Cooke (1963)), $\varphi(\sigma+iT) \to 0$ as $T \to \infty$ for $x_1 \le \sigma \le x_2$ so that for T_0 large enough $|1-\varphi(\sigma+iT)| > \epsilon$ for $T > T_0$. Therefore, using (5.1.7), the second integral in (5.1.17) is in absolute value less than or equal to

$$\frac{2me^{x_1 t}}{2\pi\epsilon}\int_{x_1}^{x_2}\frac{1}{|\sigma+iT|}d\sigma \to 0 \quad \text{as} \quad T\to\infty .$$

Likewise, the third integral in (5.1.19) approaches zero as $T \to \infty$.

The first integral in (5.1.19) may be written in the form

$$\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{(x_1+i\tau)t}[1+(m-1)G_k^*(x_1+i\tau)]}{x_1+i\tau}d\tau$$

$$+\frac{m-1}{2\pi}\int_{-T}^{T}\frac{e^{(x_1+i\tau)t}G_k^*(x_1+i\tau)}{(x_1+i\tau)[1-\varphi(x_1+i\tau)]}d\tau \tag{5.1.20}$$

The first term in (5.1.20) approaches $1+(m-1)G_k(t)$ as $T\to\infty$, and the second term is in absolute value

$$\leq\left[\frac{(m-1)e^{x_1t}}{2\pi C}\right]\left[\int_{-\infty}^{\infty}\frac{|G_k^*(x_1+i\tau)||\varphi(x_1+i\tau)|}{|x_1+i\tau|}d\tau\right] \tag{5.1.21}$$

since by another application of the Riemann-Lebesgue lemma,

$$|1-\varphi(x_1+i\tau)|>C\,,\quad -\infty<\tau<\infty$$

for some $C>0$. By the Hoëlder inequality, the second term in (5.1.21) is

$$\leq\left\{\int_{-\infty}^{\infty}|G_k^*(x_1+i\tau)|^{q_k}d\tau\right\}^{1/q_k}\left\{\int_{-\infty}^{\infty}\left|\frac{\varphi(x_1+i\tau)}{x_1+i\tau}\right|^{p_k}d\tau\right\}^{1/p_k} \tag{5.1.22}$$

where q_k is chosen such that $1/q_k+1/p_k=1$. Since $|\varphi(\lambda)|$ is bounded above by m for all $\lambda\geq 0$, the second term in (5.1.22) is finite.

Let us denote the Fourier transform of $e^{-x_1t}g_k(t)$ by $f_k(\tau)$, i.e.,

$$f_k(\tau) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\tau t} e^{-x_1 t} g_k(t)dt .$$

Therefore

$$\frac{G_k^*(x_1 + i\tau)}{\sqrt{2\pi}} = f_k(-\tau) .$$

Then by Theorem 7.4[1] of Titchmarsh (1937) we get

$$\begin{aligned}(2\pi)^{-q_k/2}\int_{-\infty}^{\infty} |G_k^*(x_1 + i\tau)|^{q_k} d\tau &= \int_{-\infty}^{\infty} |f_k(\tau)|^{q_k} d\tau \\ &\le (2\pi)^{1-q_k/2}\left[\int_{-\infty}^{\infty} e^{-p_k x_1 t}(g_k(t))^{p_k} dt\right]^{1/p_k - 1}\end{aligned} \tag{5.1.23}$$

This last term is finite by hypothesis and therefore the first term in (5.1.22) is also finite.

This completes the proof of assertion (i) since, by the remarks just preceeding this theorem, it follows that $\overline{N}(t,k) = M(t,k)$. To prove (ii) simply replace p_k by p and q_k by q in the proof of (1). It then follows that the last term in (5.1.23) is

1 If (5.1.10) is true for some $p_k > 1$, it is also true for some p_k , $1 < p_k \le 2$, so we can assume without loss of generality that $1 < p_k \le 2$.

$$\leq (2\pi)^{1-q/2} B^{1/(p-1)}$$

and the proof of Theorem 5.1.2 is complete.

Remarks (i): Conditions (5.1.13) and (5.1.15) on the functions $g_k(t)$ are met by a wide class of densities. For example (5.1.13) is satisfied by any gamma with $p_k = 2$ and (5.1.15) is satisfied by any countable collection of densities $g_k(t)$ taken from the collection of gammas

$$f(x;\theta,\beta) = \frac{x^{\theta} e^{-x/\beta}}{\Gamma(\theta+1)\beta^{\theta+1}} \quad \theta \geq 0, \quad \beta \geq \epsilon > 0$$

with $p = 2$.

(ii) It is interesting to note that the exponent in the exponential approximation to $M(t,k)$ given in (5.1.11) is the same for all k. This says the rate of growth of the process is independent of where the process "begins".

Turning now to the second moment, it can be shown by using familiar techniques that if $h''(1) < \infty$, then $M_2(t,\tau,k) = E[Z(t)Z(t+\tau)|v=k]$ satisfies the equation

$$M_2(t,\tau,k) = f_k(t,\tau) + \Sigma n p_n \int_0^t M_k(t-u,\tau,n)\,dG_k(u) \tag{5.1.24}$$

where

$$f_k(t,\tau) = \Sigma n(n-1)p_n\int_0^t M(t-u,n)M(t+\tau-u,n)dG_k(u)$$

$$+ \Sigma np_n\int_t^{t+\tau} M(t+\tau-u,n)dG_k(u) + 1 - G_k(t+\tau) . \tag{5.1.25}$$

Moreover, (5.1.24) has a unique solution with the property that

$$|M_2(t,\tau,k)| \le A_2 e^{r_2 t} \tag{5.1.26}$$

for some positive numbers r_2 and A_2 independent of t and τ.

In order to obtain the mean square convergence of a suitably normed version of $Z(t)$, we must investigate the asymptotic properties of $M_2(t,\tau,k)$. Toward this end the following definitions will be useful. Let

$$\begin{aligned}
\overline{M}_2(t,\tau,k) &= M_2(t,\tau,k)e^{-\alpha(2t+\tau)} \\
\overline{M}(t,k) &= M(t,k)e^{-\alpha t} \\
\overline{f}_k(t,\tau) &= f_k(t,\tau)e^{-\alpha(2t+\tau)} \\
\overline{G}_k(t) &= \int_0^t e^{-2\alpha u}dG_k(u) \\
a_k &= \overline{G}_k(\infty)
\end{aligned} \tag{5.1.27}$$

and

$$\overline{H}_k(t) = \sum_{j=1}^{\infty} \overline{G}_{k\,j}(t) ,$$

where $\overline{G}_{k1}(t) = \overline{G}_k(t)$ and

$$\overline{G}_{k(n+1)}(t) = \Sigma j p_j \overline{G}_{jn} * \overline{G}_k(t) .$$

<u>Lemma 5.1.3</u>: If $m > 1$, then

(i) $$\overline{G}_{kn}(t) \to a_k(\Sigma j p_j a_j)^{n-1} \quad \text{as} \quad t \to \infty \tag{5.1.28}$$

and

(ii) $$\overline{H}_k(t) \to \frac{a_k}{1 - \Sigma j p_j a_j} \quad \text{as} \quad t \to \infty \tag{5.1.29}$$

<u>Proof</u>: By the definition of α in (5.1.9), we know that $\Sigma j p_j a_j < 1$. Therefore (ii) follows from (i), using the formula for the sum of a geometric series. Assertion (i) follows by induction on n using Lemma 1.4.2 .

<u>Lemma 5.1.4</u>: If $h''(1) < \infty$ and the conditions of assertion (ii) of Theorem 5.1.2 hold, then

(i) there exists a constant A such that

$$f_k(t,\tau) \le A \quad \text{for all} \quad k,\ t,\ \text{and}\ \tau . \tag{5.1.30}$$

(ii) $$\lim_{t\to\infty} \overline{f}_k(t,\tau) = a_k \Sigma n(n-1) p_n b_n^2 < \infty \tag{5.1.31}$$

and this limit is uniform in τ. (We defined b_n in (5.1.12).)

<u>Proof</u>: If we multiply both sides of (5.1.25) by $e^{-\alpha(2t+\tau)}$ we get

$$\overline{f}_k(t,\tau) = \Sigma n(n-1)p_n\int_0^t \overline{M}(t-u,n)\overline{M}(t+\tau-u,n)d\overline{G}_k(u) \tag{5.1.32}$$

$$+ e^{-\alpha t}\Sigma np_n\int_t^{t+\tau} \overline{M}(t+\tau-u,n)d\overline{G}_k(u) + e^{-\alpha(2t+\tau)}[1-G_k(t+\tau)] .$$

Since by assertion (ii) of Theorem (5.1.2) $\overline{M}(t,k)$ converges uniformly in k as $t \to \infty$, and by Theorem 5.1.1 $\overline{M}(t,k)$ is bounded uniformly in k on any finite t-interval, it follows that there exists a D such that

$$\overline{M}(t,k) \le D \quad \text{for all } t \text{ and } k. \tag{5.1.33}$$

Applying this to (5.1.30) we get

$$\overline{f}_k(t,\tau) \le h''(1)D^2 + mD + 1$$

and this proves (i).

It can be shown using (5.1.30) that for $t > T$

$$|\overline{f}_k(t,\tau) - a_k\Sigma n(n-1)p_n b_n^2|$$

$$\le \Sigma n(n-1)p_n\int_0^T |\overline{M}(t-u,n)\overline{M}(t+\tau-u,n) - b_n^2|dG_k(u)$$

$$+ 3D^2h''(1)[\overline{G}_k(\infty) - \overline{G}_k(T)] + (mD+1)[1-G_k(t)]$$

and the second assertion follows easily from this expression. The sum in (5.1.31) is finite since the sequence of b_n's is bounded above.

We are now ready to prove the principal result for the second moment.

<u>Theorem 5.1.3</u>: If $h''(1) < \infty$ then $\overline{M}_2(t,\tau,k)$ is given by

$$\overline{M}_2(t,\tau,k) = \overline{f}_k(t,\tau) + \Sigma np_n \int_0^t \overline{f}_n(t-u,\tau)d\overline{H}_k(u) \tag{5.1.34}$$

If, in addition, the conditions of assertion (ii) of Theorem 5.1.2 hold, then

$$\frac{M_2(t,\tau,k)}{e^{\alpha(2t+\tau)}} \to \frac{a_k \Sigma n(n-1)p_n b_n^2}{1 - \Sigma np_n a_n} \tag{5.1.35}$$

and this limit is uniform in τ.

<u>Proof</u>: If we multiply both sides of (5.1.24) by $e^{-\alpha(2t+\tau)}$, we get

$$\overline{M}_2(t,\tau,k) = \overline{f}_k(t,\tau) + \Sigma np_n \int_0^t \overline{M}_2(t-u,\tau,n)d\overline{G}_k(u). \tag{5.1.36}$$

By substituting (5.1.32) into the right side of (5.1.34) we get (In each case $*$ indicates a convolution with respect to t with τ fixed.)

$$\begin{aligned} &\overline{f}_k + \Sigma np_n[\overline{f}_n + \Sigma jp_j\overline{f}_j * \overline{H}_n] * \overline{G}_k \\ &= \overline{f}_k + \Sigma np_n\overline{f}_n * \overline{G}_k + \Sigma jp_j\overline{f}_j * (\Sigma np_n\overline{H}_n * \overline{G}_k) . \end{aligned} \tag{5.1.37}$$

But

$$\Sigma np_n \overline{H}_n * \overline{G}_k = \sum_{r=2}^{\infty} \overline{G}_{kr}. \tag{5.1.38}$$

Therefore the last line in (5.1.35) can be written as

$$\begin{aligned} &\overline{f}_k + \Sigma np_n [\overline{f}_n * (\overline{G}_{k1} + \sum_{r=2}^{\infty} \overline{G}_{kr})] \\ &= \overline{f}_k + \Sigma np_n (\overline{f}_n * \overline{H}_k) . \end{aligned} \tag{5.1.39}$$

It follows that (5.1.32) is a solution of (5.1.34) and the first part of the theorem is proved.

Let us denote $a_k \Sigma n(n-1)p_n b_n^2$ by c_k. Since by Lemma 5.1.3, $\overline{f}_k(t,\tau) \le A$, it follows that $c_k \le A$ for all k. For a given $\epsilon > 0$ choose an integer x such that

$$\frac{2a_k A}{1 - \Sigma r p_r a_r} \sum_{n=x+1}^{\infty} np_n < \frac{\epsilon}{4} \tag{5.1.40}$$

Then, using Lemma 5.1.3, choose T such that

$$\sup_{\substack{r \ge T/2 \\ \tau}} |\overline{f}_n(r,\tau) - c_n| \le \frac{\epsilon}{4mH(\infty)} \quad \text{for} \quad n = 1,2,\cdots,x , \tag{5.1.41}$$

and

$$mA[H_k(\infty) - H_k(T/2)] < \frac{\varepsilon}{4} . \tag{5.1.42}$$

Then for any $t \geq T$ and any τ, we have

$$\left|\Sigma np_n \int_0^t \overline{f}_n(t-u,\tau)d\overline{H}_k(u) - \frac{a_k \Sigma np_n c_n}{1 - \Sigma rp_r a_r}\right|$$

$$\leq \Sigma np_n \int_{t/2}^t |\overline{f}_n(t-u,\tau)|d\overline{H}_k(u) + \sum_{n=1}^{x} np_n \int_0^{t/2} |\overline{f}_n(t-u,\tau) - c_n|d\overline{H}_k(u) \tag{5.1.43}$$

$$+ \sum_{n=x+1}^{\infty} np_n \int_0^{t/2} \{|\overline{f}_n(t-u,\tau)| + |c_n|\}d\overline{H}_k(u) + \Sigma np_n c_n[\overline{H}_k(\infty) - \overline{H}_k(t/2)].$$

If the estimate (5.1.42) is applied to the first and fourth terms of the last expression and (5.1.41) and (5.1.40) are applied to the second and third terms, respectively, we see that

$$\Sigma np_n \int_0^t \overline{f}_n(t-u,\tau)d\overline{H}_k(u) \to \frac{a_k \Sigma np_n c_n}{1 - \Sigma rp_r a_r} \quad \text{as} \quad t \to \infty \tag{5.1.44}$$

and the limit is uniform in τ. The proof of the second part of the theorem may now be completed by applying (5.1.44) and (5.1.31) to (5.1.34) and rearranging terms.

If $Z_n(t)$ denotes the size of the population at time t in a process for which $v = n$ a.s. then we have the following

Corollary: Under the conditions of Theorem 5.1.3, $Z_n(t)e^{-\alpha t}$ converges in mean square to a random variable W_n as $t \to \infty$, and $E(W_n) = b_n$

$$\mathrm{Var}(W_n) = \frac{a_n \Sigma j(j-1)p_j b_j^2}{1 - \Sigma j p_j a_j} - 1 .$$

The variance of W_n is positive.

Proof: The proof is similar to that of Theorem 2.7.1 and will be omitted.

Remark: By a deeper analysis one could most likely show that the convergence indicated in (5.1.35) is of exponential order as in Lemma 4.8.1. This would imply that $Z_n(t)e^{-\alpha t} \to W_n$ a.s. as $t \to \infty$.

5.2 Population with dormant members

Consider a process identical to the Bellman-Harris process described in Chapter I except that $G(t)$ is defective; that is to say, $q = G(\infty) < 1$, which means that an individual has probability $1-q$ of living forever. Such a process might be applicable to a colony of bacteria in which each organism can possibly go into a dormant state so that it always remains in the population but never reproduces. If the process is given this interpretation, $1-q$ is simply the probability of an organism becoming dormant.

As in the Bellman-Harris model the generating function of this process satisfies

$$F(s,t) = s[1 - G(t)] + \int_0^t h[F(s,t-u)]dG(u) \qquad (5.2.1)$$

and

$$M(t) = 1 - G(t) + m\int_0^t M(t-u)dG(u) \tag{5.2.2}$$

where $M(t)$ is again the expected number of organisms alive at time t. By Lemma 1.3.1, equation (5.2.2) can be solved to yield

$$M(t) = 1 + (m-1)\sum_{k=1}^{\infty} m^{k-1}G_k(t) . \tag{5.2.3}$$

The asymptotic behavior of $M(t)$ can be determined by considering several different cases.

(i) If $m < 1$, by letting $t \to \infty$ in (5.2.3), noting that $G_k(t) \to q^k$ as $t \to \infty$, it follows that

$$M(t) \to \frac{1-q}{1-mq} < 1 \quad \text{as} \quad t \to \infty .$$

(ii) If $m = 1$, then $M(t) = 1$.

(iii) If $m > 1$ but $mq < 1$, then again from (5.2.3)

$$M(t) \to \frac{1-q}{1-mq} > 1 \quad \text{as} \quad t \to \infty .$$

(iv) If $mq > 1$, there exists a unique $\alpha > 0$ such that

$$m\int_0^\infty e^{-\alpha t}dG(t) = 1$$

and from Lemma 1.4.3 ,

$$\frac{M(t)}{e^{\alpha t}} \to \frac{\int_0^\infty e^{-\alpha u}(1 - G(u))du}{m\int_0^\infty ue^{-\alpha u}dG(u)} \quad \text{as} \quad t \to \infty .$$

(v) If $m > 1$ and $mq = 1$, then from the corollary to Lemma 1.4.3

$$\frac{M(t)}{t} \to \frac{1 - q}{m\int_0^\infty udG(u)} \quad \text{as} \quad t \to \infty .$$

Cases (i), (ii), and (iii) show that $M(t)$ can be made to converge to any non-negative number. Case (v) is interesting in that it furnishes an example of a branching process with a linear rate of growth.

The results for the extinction probability for a Bellman-Harris process must also be modified in order to fit this situation. If we set $Q(t) = F(0,t)$, then we see from (5.2.1) that

$$Q(t) = \int_0^t h[Q(t - u)]dG(u) . \tag{5.2.4}$$

Using techniques similar to those used in section 3.4, it can be shown that Q, the probability of extinction, is the smallest non-negative root of

$$s = qh(s) \tag{5.2.5}$$

and, since $q < 1$, it follows that $Q < 1$ also, whatever the value of m.

Another quantity of interest in this model is the probability that eventually the population either becomes extinct or contains only dormant members. Let us call this quantity the "probability of dormancy" and denote it by D.

If $Z^*(t)$ is the number of non-dormant members of the population at time t and $F^*(s,t)$ is the generating function of $Z^*(t)$, it can be shown that

$$F^*(s,t) = 1 - q + s(q - G(t)) + \int_0^t h[F^*(s,t-u)]dG(u) \ . \tag{5.2.6}$$

Applying the techniques of section 3.4 to this equation it can be shown that $Q^* = P[Z^*(t) = 0 \text{ for some } t]$ is the smallest non-negative root of

$$s = 1 - q + qh(s) \tag{5.2.7}$$

Since the population is dormant at time t if and only if $Z^*(t) = 0$,

it follows that $D = Q^*$ and therefore D is also the smallest non-negative root of (5.2.7). From the results for Galton-Watson processes in Section 1.2, we conclude that if $mq \leq 1$ then $D = 1$, and if $mq > 1$ then $D < 1$.

LITERATURE CITED

Bellman, R. E. and K. L. Cooke (1963). Differential-Difference Equations. Princeton University Press.

Bellman, R. E. and T. E. Harris (1952). On age-dependent binary branching processes. Annals of Mathematics 55, 280-295.

Ditkin, V. A. and A. P. Prudnikov (1965). Integral Transforms and Operational Calculus. New York: Pergamon Press. (translated by D. E. Brown)

Everett, C. J. and S. Ulam (1948). Multiplicative systems in several variables, III. Los Alamos Scientific Laboratory, LA-707.

Goodman, L. A. (1967). The probabilities of extinction for birth-and-death processes that are age-dependent or phase-dependent. Biometrica 54, 579-596.

Feller, William (1941). On the integral equation of renewal theory. Annals of Mathematical Statistics 12, 243-267.

_____ (1957). An Introduction to Probability Theory and Its Applications, Vol. I, second edition. New York: John Wiley and Sons, Inc.

_____ (1966). An Introduction to Probability Theory and Its Applications, Vol. II. New York: John Wiley and Sons, Inc.

Fisher, R. A. (1930). The Genetical Theory of Natural Selection. Oxford University Press. New York: Dover Publications, Inc., 1958.

Harris, T. E. (1963). The Theory of Branching Processes. Englewood Cliffs, N. J.: Prentice Hall, Inc.

Kendall, D. G. (1949). Stochastic processes and population growth. Journal of the Royal Statistical Society B 11, 230-264.

_____ (1966). Branching processes since 1873. Journal London Math. Soc. 41, 385-406.

Kolmogorov, A. (1933). Grundbegriffe der Wahrscheinlichkeitsrechnung. Berlin: Springer-Verlag; translated as Foundations of the Theory of Probability (1956). New York: Chelsea Publishing Co.

Loeve, Michel (1963). Probability Theory, third edition. New York: D. Van Nostrand Company, Inc.

Lukacs, E. (1960). Characteristic Functions. London: Charles Griffin & Company Limited.

Mode, C. J. (1967). A renewal density theorem in the multi-dimensional case. Journal of Applied Probability 4, 62-76.

_____ (1968a). A multi-dimensional age-dependent branching process with applications to natural selection I. Mathematical Biosciences (to appear).

_____ (1968b). A multi-dimensional age-dependent branching process with applications to natural selection II. Mathematical Biosciences (to appear).

Ney, P. E. (1961). Some contributions to the theory of cascades. Ph.D. thesis, Columbia University.

_____ (1964a). Generalized branching processes I. Illinois Journal of Mathematics 8, 316-331.

_____ (1964b). Generalized branching processes II. Illinois Journal of Mathematics 8, 332-350.

Otter, Richard (1949). The multiplicative process. Annals of Mathematical Statistics 20, 206-224.

Parzen, Emanuel (1962). Stochastic Processes. San Francisco: Holden-Day, Inc.

Powell, E. O. (1955). Some features of the generation times of individual bacteria. Biometrica 42, 16-44.

Ryan, T. J., Jr. (1967). On age-dependent branching processes. Preliminary report. Notices of the American Mathematical Society 14, 854.

Sevast'yanov, B. A. (1964). Age-dependent branching processes. Theory of Probability and Its Applications 9, 521-537.

_____ (1967). On the regularity of branching processes. Matematicheskie Zametki I, 53-62. (in Russian)

Stigum, B. (1966). A theorem on the Galton-Watson process. Annals of Mathematical Statistics 37, 695-698.

Titchmarsh, E. C. (1937). Introduction to the Theory of Fourier Integrals. Oxford University Press.

Tukey, John W. (1958). A problem of Berkson and minium variance orderly estimators. Annals of Mathematical Statistics 29, 588-592.

Watson, H. W. and Francis Galton (1874). On the probability of the extinction of families. J. Anthropol. Inst. Great Britain and Ireland 4, 138-144.

[redacted]

Crump, K.S.
Some generalized age-dependent branching processes

[redacted]

NAME AND ADDRESS

[redacted]

[redacted]