1969

# Birth and death processes

Ronald Keith Wade
*The University of Montana*

129

# BIRTH AND DEATH PROCESSES

By

Ronald Wade

B.A. University of Montana 1966

Presented in partial fulfillment of the requirements

for the degree of

Master of Arts

UNIVERSITY OF MONTANA

1969

Approved by:

H E Reinhardt

Chairman, Board of Examiners

John M. Stewart

Dean, Graduate School

March 14, 1969

Date

ACKNOWLEDGEMENTS

I wish to express my appreciation to Professor Howard E. Reinhardt for his guidance and instruction throughout the preparation of this thesis. Also, I wish to thank Professor David Arterburn, Professor Charles Bryan, and Professor R. S. Hoffman for reading the thesis.

R. W.

# TABLE OF CONTENTS

# INTRODUCTION

Ecologists have studied populations of organisms in the field for a number of years. Many of these studies have been concerned with rather small populations in which random fluctuations may be expected to be of more importance than in very large populations. Therefore, it should be helpful to study an arbitrary population which increases and decreases according to certain probability laws. In this paper, two different formulations of the problem will be considered.

CHAPTER I

# BIRTH AND DEATH PROCESSES AND RELATED DISTRIBUTIONS

## 1. Poisson and Pure Birth and Death Processes

Consider first a system subject to instantaneous changes due to the occurrence of random events. All changes are assumed to be of the same kind. The only concern is the total number of changes. Each change occurs at some point on the time axis. Arrival of telephone calls and breakage of a chromosome under harmful irradiation are examples. These two physical examples have two main properties in common; forces which determine the process do not change (homogeneity in time), and future changes are independent of past changes. The number of changes is recorded by a counting function $N_t$ which enumerates the number of changes in the interval $[0, t)$. The mathematical formulation of such physical processes is called the Poisson process. The postulates are:

(a) Whatever the number of changes during $[0, t)$, the probability that during $[t, t + h)$ a change occurs is $\lambda h + o(h)$, (b) The probability that more than one change occurs is $o(h)$. The term $\lambda h + o(h)$ means $\lambda h$ + "something of smaller order than h". The "something" is a function $f(h)$ satisfying $\lim_{h \to 0} f(h)/h = 0$ and one writes $f(h) = o(h)$. From these postulates alone one can prove that the probability of exactly r occurrences in the time interval $[T, T + t)$ is $\frac{e^{-\lambda t}(\lambda t)^r}{r!}$. The reader is referred to Feller [3], pp. 400-402, for details. A stochastic process is defined as a family of random variables, $(X_t, t \varepsilon T)$; it is a random phenomenon arising through a process which is developing in time according to probabilistic laws. The Poisson process is an example.

In the Poisson process, the probability of a change during $[t, t + h)$ is independent of the number of changes during $[0, t)$. The pure birth process is a modification of the Poisson process. For the pure birth process, postulate (a) of the Poisson process becomes

(a') When n changes occur during $[0, t)$, the probability of a change during $[t, t + h)$ is $\lambda_n h + o(h)$.

In the Poisson process, past and future are independent. In the pure birth process, past and future are not independent, since the probability of a subsequent change, called a birth, depends on the cumulative size of the population through $\lambda_n$, called the birth rate. However, there is no effect of aging; the birth rate is not "age-specific". For a discussion of the pure birth process the reader is referred to Feller [3], pp. 402-407.

The above mentioned processes provide models for systems, such as arrival of telephone calls, which can only undergo one type of change. Neither process can, however, serve as a realistic model for changes in population size where individuals are not only born, but can also die. The pure birth and death process provides a model for such a system. One additional postulate is needed along with a rewording so that the postulates for the pure birth and death process become:

(i) Suppose that the system is in state n at time t; the probability that the system changes to state n + 1 during [t, t + h) is $\lambda_n h + o(h)$.

(ii) Suppose that the system is in state n at time t; the probability that the system changes to state n - 1 during [t, t + h) is $\mu_n h + o(h)$.

(iii) The probability of any other kind of change is o(h).

The functions $\lambda_n$ and $\mu_n$ are called the birth and death rates respectively. The word "pure" refers to the fact that neither $\lambda_n$ nor $\mu_n$ is a function of time. Further discussion of the pure birth and death process may be found in Feller [3], pp. 407-411. In the rest of this paper, the birth and death functions will be allowed to be time dependent; such processes are sometimes called generalized birth and death processes.

## 2. A Generalized Birth and Death Process

Population growth can be considered as a birth and death process in which the birth and death rates are arbitrary functions of time. This consideration leads to the non-homogeneous, integer valued process $(N_t,\ t \geq 0)$ satisfying the following postulates:

(a) $P(N_{t+h} = N_t + 1) = \lambda(t)N_t h + o(h) \qquad N_t \geq 0$

(b) $P(N_{t+h} = N_t - 1) = \mu(t)N_t h + o(h) \qquad N_t \geq 1$

(c) $P(N_{t+h} = N_t) = 1 - (\lambda(t) + \mu(t))N_t h + o(h) \qquad N_t \geq 0$

(d) The probability of any other kind of change is $o(h)$

(e) The functions $\lambda(t)$, $\mu(t)$, and the first derivative of $\lambda(t)$ are assumed to be continuous.

In this formulation, $N_t$ is a random variable which measures the size of the population at time t. The functions $\lambda(t)$ and $\mu(t)$ are called the birth and death rates respectively. We will be interested in small values of h; the first four postulates say that with overwhelming probability, the population size changes by at most one in a small time interval. The fifth postulate is a regularity condition which will permit certain formal manipulations. For the moment, suppose that the population descends from only one individual so that $N_o = 1$. We seek the distribution of $N_t$. Our development follows closely that of D. G. Kendall [6]. The following result about the so-called Ricatti differential equation $\frac{dz}{dx} = Rz^2 + Qz + P$ turns out to be essential.

THEOREM 1. Suppose that P, Q, R, and the first derivative of R are continuous functions of x and that $R(x) \neq 0$. Then the equation $\frac{dz}{dx} = Rz^2 + Qz + P$ has a solution; the general solution may be expressed in the form $z = \frac{f_1(x) + Af_2(x)}{f_3(x) + Af_4(x)}$ where A is an arbitrary constant. The Ricatti differential equation is discussed in many standard treatises including the one of Ince [5], pp. 73-75, 293-295.

THEOREM 2. Under assumptions (a)-(e) with $N_0 = 1$, $N_t$ has a geometric distribution with a modified zero term. A geometric distribution with a modified zero term is defined by:

$$P_0 = r$$

$$P_n = (1-r)(1-s)s^{n-1} \quad 0 < r \leq 1,\ 0 \leq s \leq 1,\ n = 1,2,\ldots;$$

if r = 0, the preceding defines the geometric distribution.

Proof: Define $P_n(t) = P(N_t = n)$ for $n \geq 0$, $t \geq 0$. We have $P_1(0) = 1$ and $P_n(0) = 0$ for $n \neq 1$. Consider (for $n \geq 1$) $P_n(t + h) = P(N_{t+h} = n)$; the probability is o(h) that $N_t$ is other than n - 1, n, or n + 1. Hence, if $n \geq 1$,

(1) $$P_n(t+h) = (n-1)\lambda(t)hP_{n-1}(t) + (1 - (\lambda(t) + \mu(t))nh)P_n(t) + (n+1)\mu(t)hP_{n+1}(t) + o(h).$$

Thus, $$P_n(t+h) - P_n(t) = (n-1)\lambda(t)hP_{n-1}(t) + (n+1)\mu(t)hP_{n+1}(t) - n(\lambda(t) + \mu(t))hP_n(t) + o(h).$$

Dividing by h and letting $h \rightarrow 0$, we obtain $$\frac{\partial P_n(t)}{\partial t} = (n-1)\lambda(t)P_{n-1}(t) + (n+1)\mu(t)P_{n+1}(t) - n(\lambda(t) + \mu(t))P_n(t).$$

If we define $P_{-1}(t) = 0$, the above is also true for n = 0.

The above is the right hand derivative of $P_n(t)$. Replacing $t$ by $t - h$ in (1) gives the same result for the left hand derivative since $\lambda(t)$ and $\mu(t)$ were assumed continuous and $o(h)$ is independent of $t$. We wish to find the probability generating function of $N_t$; the reader is referred to Feller [3], pp. 248-267, for a discussion of generating functions. Define

(2) $\psi(z, t) = \sum_{k=0}^{\infty} P_k(t)z^k$ which converges for $|z| < 1$.

Now, $\frac{1}{h}(\psi(z, t + h) - \psi(z, t)) = \sum_{k=0}^{\infty} \frac{1}{h}(P_k(t + h) - P_k(t))z^k$.

The fact that equality holds as $h \to 0$ follows from an application of the dominated convergence theorem to the series $\sum_{k=0}^{\infty} \frac{1}{h}(P_k(t + h) - P_k(t))z^k$. To see that this theorem applies, notice that

$$\frac{1}{h}(P_k(t + h) - P_k(t)) = (k - 1)\lambda(t)P_{k-1}(t) + (k + 1)\mu(t)P_{k+1}(t) - k(\lambda(t) + \mu(t))P_k(t) + o(h)/h$$

from (1). Since $\lambda(t)$ and $\mu(t)$ are bounded, the above can be replaced by

$$\frac{1}{h}(P_k(t+h) - P_k(t)) \leq M((k-1)P_{k-1}(t) + (k+1)P_{k+1}(t) + 2kP_k(t) + 1)$$

for some constant $M = \max(|\lambda(t)|, |\mu(t)|, |\frac{o(h)}{h}|)$. Furthermore, the above expression in parentheses is less than or equal to $4k + 1$. Therefore,

$$\sum_{k=0}^{\infty} \frac{1}{h}(P_k(t+h) - P_k(t))z^k \leq M\sum_{k=0}^{\infty}(4k+1)z^k$$

and for $|z| < 1$, $M \sum_{k=0}^{\infty}(4k+1)z^k = \frac{M(3z+1)}{(1-z)^2} < \infty$. Thus, the hypotheses of the dominated convergence theorem are satisfied and the interchange of limits is justified.

Hence,

(3) $\frac{\partial \Psi(z,t)}{\partial t} =$

$$\sum_{k=0}^{\infty}\{(k-1)\lambda(t)P_{k-1}(t) + (k+1)\mu(t)P_{k+1}(t) - k(\lambda(t)+\mu(t))P_k(t)\}z^k =$$

$$\{\lambda(t)z^2 - (\lambda(t) + \mu(t))z + \mu(t)\} \sum_{k=1}^{\infty} kP_k(t)z^{k-1} =$$

$$\{\lambda(t)z^2 - (\lambda(t) + \mu(t))z + \mu(t)\} \frac{\partial \Psi(z,t)}{\partial z} .$$

Since $P_1(0) = 1$, $\Psi(z, 0) = z$. We wish to look at the subsidiary equations of (3) and obtain two independent integrals of the former; suppose these are $u = a$, $v = b$. Then an integral of (3) is given by $u = \Phi(v)$ where $\Phi$ is an arbitrary function. For further details, the reader may consult Forsyth [4], pp. 392-394. The subsidiary equations are

(4) $\frac{dt}{1} = \frac{dz}{-(z-1)(\lambda(t)z-\mu(t))} = \frac{d\Psi}{0}$. Consider

(5) $\frac{dz}{dt} = -\mu(t) + (\lambda(t) + \mu(t))z - \lambda(t)z^2$; we wish to find an integral of this equation. Now, (5) is a Ricatti equation so the general solution is given by theorem 1, i.e.

$z = \frac{f_1(t) + Af_2(t)}{f_3(t) + Af_4(t)}$. Therefore,

$A = \frac{f_1(t) - zf_3(t)}{af_4(t) - f_2(t)}$. From $\frac{dt}{1} = \frac{d\Psi}{0}$, $\Psi = A'$. Thus,

(6) $\Psi = A' = \Phi\left(\frac{f_1(t) - zf_3(t)}{zf_4(t) - f_2(t)}\right)$ which is the general solution of (3). We now want to determine the precise form of $\Phi$.

From $\Psi(z, 0) = z$,

$z = \Psi(z, 0) = \Phi\left(\frac{f_1(0) - zf_3(0)}{zf_4(0) - f_2(0)}\right)$. Now put

$$u = \frac{f_1(0) - zf_3(0)}{zf_4(0) - f_2(0)}$$

or

$$z = \frac{f_1(0) + uf_2(0)}{f_3(0) + uf_4(0)}$$

where $u = u(z, t, \Psi)$ = constant is an integral of (5). Hence

(7) $\Phi(u) = z = \frac{f_1(0) + uf_2(0)}{f_3(0) + uf_4(0)}$. Combining (6) and (7) gives

$$\Psi(z, t) = \frac{f_1(0) + v(z,t)f_2(0)}{f_3(0) + v(z,t)f_4(0)}$$

where

$$v(z, t) = \frac{f_1(t) - zf_3(t)}{zf_4(t) - f_2(t)}.$$

So,

$$\Psi(z, t) = \frac{bf_1(t) - af_2(t) + z(af_4(t) - f_3(t))}{df_1(t) - cf_2(t) + z(cf_4(t) - df_3(t))}$$

where

$a = f_1(0)$, $b = f_2(0)$, $c = f_3(0)$, $d = f_4(0)$. Expansion shows that

(8) $P_0(t) = r(t)$ and $P_n(t) = (1 - P_0(t))(1 - s(t))s(t)^{n-1} (n \geq 1)$

where

$$r(t) = \frac{bf_1(t) - af_2(t)}{df_1(t) - cf_2(t)}$$

$$s(t) = \frac{cf_4(t) - df_3(t)}{df_1(t) - cf_2(t)}.$$

But this is exactly the conclusion of the theorem.

Determination of the functions $r(t)$ and $s(t)$ will complete the solution. From (8),

$$(9)\quad \psi(z,t) = r(t) + (1-r(t))(1-s(t))z + (1-r(t))(1-s(t))s(t)z^2 + \ldots$$

$$= r(t) + z(1-r(t))(1-s(t)) \sum_{k=0}^{\infty}(s(t)z)^k$$

$$= \frac{r(t) + (1-r(t) - s(t))z}{1 - s(t)z}$$

Thus,

$$\frac{\partial\psi}{\partial t} = \frac{(1-s(t)z)(r'(t)-r'(t)z-s'(t)z)+(r(t)+(1-r(t)-s(t))z)s'(t)z}{(1-s(t)z)^2}$$

and

$$\frac{\partial\psi}{\partial z} = \frac{(1-s(t)z)(1-s(t)-r(t))+(r(t)+(1-r(t)-s(t))z)s(t)}{(1-s(t)z)^2}$$

where

$$r'(t) = \frac{dr(t)}{dt} \text{ and } s'(t) = \frac{ds(t)}{dt}.$$

Substitution of the expressions for $\frac{\partial\psi}{\partial t}$ and $\frac{\partial\psi}{\partial z}$ in (3) gives

$$(1-s(t)z)(r'(t) - r'(t)z - s'(t)z) + (r(t) + (1 - r(t) - s(t))z)s'(t)z =$$
$$(\lambda(t)z^2 - (\lambda(t) + \mu(t))z + \mu(t))((1-s(t)z)(1-r(t) - s(t)) + (r(t) + (1-r(t) - s(t))z)s(t))$$ or,

$$(r'(t)s(t)-r(t)s'(t)+s'(t))z^2 + (r(t)s'(t) - r'(t) - s'(t) - r'(t)s(t))z + r'(t) =$$
$$\lambda(t)(1-r(t))(1-s(t))z^2 - (\lambda(t) + \mu(t))(1-r(t))(1-s(t))z + \mu(t)(1-r(t))(1-s(t)).$$

Thus,

$$(10)\quad r'(t)s(t) - r(t)s'(t) + s'(t) = \lambda(t)(1-r(t))(1-s(t))$$

and

(11) $r'(t) = \mu(t)(1 - r(t))(1 - s(t))$

are obtained by equating coefficients of $z^2$ and constants. Now let $U = 1 - r(t)$, $V = 1 - s(t)$. Then (11) becomes $U' = -\mu(t)UV$ and (10) becomes $U'(V - 1) - V'U = \lambda(t)UV$. Substitution of the former into the latter gives $V' = (\mu(t) - \lambda(t))V - \mu(t)V^2$ which is a Bernoulli equation and which can be solved by writing $W = \frac{1}{V}$ so that $W' + (\mu(t) - \lambda(t))W = \mu(t)$. At $t = 0$, $P_0(0) = 0 = r(0)$; for $n = 1$, $P_1(0) = 1 = (1 - r(0))(1 - s(0)) = 1 - s(0)$ so $s(0) = 0$. So, $U = V = W = 1$ at $t = 0$. Thus the solution is

(12) $$W = e^{-A(t)}\int_0^t e^{A(x)}\mu(x)dx + e^{-A(t)}$$

where $A(t) = \int_0^t \mu(x) - \lambda(x)dx$.

From the above, $\frac{U'}{U} = -\mu(t)V = -\frac{\mu(t)}{W} = -\frac{W'}{W} - A'(t)$. Therefore, $\frac{U'}{U} = -\frac{W'}{W} - A'(t)$ so that

$$\int_0^t \frac{dU(s)}{U(s)} + \int_0^t \frac{dW(s)}{W(s)} = -\int_0^t A'(s)ds$$

or $\log(UW) = -A(t)$. Hence, $1 - r(t) = U = \frac{e^{-A(t)}}{W}$ or

(13) $r(t) = 1 - \frac{e^{-A(t)}}{W}$. Since $1 - s(t) = V = \frac{1}{W}$, $s(t) = 1 - \frac{1}{W}$. Combining (13) with (8) and (12) determines the $P_n(t)$ as functions of $t$.

It is now possible to find expressions for the mean and variance of $N_t$. $E(N_t) = \sum_{n=0}^{\infty} nP_n(t) =$

$$(1 - r(t))(1 - s(t)) \sum_{n=1}^{\infty} ns(t)^{n-1}$$
$$= (1 - r(t))(1 - s(t))\frac{d}{ds(t)}\left(\frac{1}{1-s(t)}\right)$$
$$= (1 - r(t))(1 - s(t)) \frac{1}{(1-s(t))^2} = e^{-A(t)} \text{ since } |s(t)| < 1.$$

Similarly, an expression for the variance is

$$\mathrm{Var}(N_t) = \sum_{n=0}^{\infty} n^2 P_n(t) - E(N_t)^2.$$

Now,

$$\sum_{n=0}^{\infty} n^2 P_n(t) = (1 - r(t))(1 - s(t)) \sum_{n=1}^{\infty} n^2 s(t)^{n-1}$$
$$= (1 - r(t))(1 - s(t))s(t) \sum_{n=1}^{\infty} n^2 s(t)^{n-2}$$
$$= (1 - r(t))(1 - s(t))s(t) \sum_{n=1}^{\infty} n^2 s(t)^{n-2} -$$
$$(1 - r(t))(1 - s(t))s(t) \sum_{n=1}^{\infty} ns(t)^{n-2} +$$
$$(1 - r(t))(1 - s(t))s(t) \sum_{n=1}^{\infty} ns(t)^{n-2}$$
$$= (1 - r(t))(1 - s(t))s(t) \sum_{n=1}^{\infty} n(n-1)s(t)^{n-2} + E(N_t)$$
$$= (1 - r(t))(1 - s(t))s(t) \frac{d^2}{ds(t)^2}\left(\frac{1}{1-s(t)}\right) + E(N_t)$$
$$= \frac{2s(t)(1-r(t))}{(1-s(t))^2} + \frac{1 - r(t)}{1 - s(t)}$$

Thus,

$$\mathrm{Var}(N_t) = \frac{(1-r(t))(r(t) + s(t))}{(1-s(t))^2}$$
$$= \frac{e^{-A(t)}}{W}\left(2 - \frac{1 + e^{-A(t)}}{W}\right) \div \frac{1}{W^2}$$
$$= e^{-2A(t)} - e^{-A(t)} + 2e^{-2A(t)}\int_0^t e^{A(x)}\mu(x)dx.$$

These formulas hold provided $N_0 = 1$. If $N_0 > 1$, the usual assumption is that sub-populations of coexisting individuals are independent. The above formulas for the mean and variance are then multiplied by the number of ancestors to give the mean and variance of $N_t$.

It is also of interest to find the probability of extinction of the population.

THEOREM 3. When $N_0 = N \geq 1$, the probability of ultimate extinction of the population is $\left(\frac{I}{1 + I}\right)^N$ where $I = \int_0^\infty e^{A(s)}\mu(s)ds$.

Proof: When $N_0 = 1$, $P_0(t) = r(t) = 1 - \frac{e^{-A(t)}}{W} =$

$$\frac{e^{-A(t)}\int_0^t e^{A(s)}\mu(s)ds}{e^{-A(t)} + e^{-A(t)}\int_0^t e^{A(s)}\mu(s)ds} = \frac{\int_0^t e^{A(s)}\mu(s)ds}{1 + \int_0^t e^{A(s)}\mu(s)ds}$$

so that the probability of ultimate extinction is $\frac{I}{1 + I}$. Since coexisting populations were assumed independent, this becomes $\left(\frac{I}{1 + I}\right)^N$ if $N_0 = N \geq 1$.

From the above, it is also seen that a necessary and sufficient condition for ultimate extinction of the population (in the sense of being an event of probability 1) is that $\int_0^\infty e^{A(s)}\mu(s)ds$ diverge to $+\infty$.

$N_t$ may assume various integer values in $[0, t_1)$. Suppose that $N_{t_1} = n$; call $n$ the state of the process $(N_t, t \geq 0)$ at time $t_1$. We may ask how long it will be until the process first leaves state $n$. Thus, we wish to find the distribution of the time spent in state n by the process until it first leaves state n. To this end, define $T_n$ to be the time until the process first leaves state n, given that $N_{t_1} = n$. The next theorem gives the distribution of $T_n$.

THEOREM 4. Under the assumptions about the process $(N_t,\ t \geq 0)$, $P(T_n < t) = 1 - e^{-n\int_0^t \lambda(t_1+s)+\mu(t_1+s)ds}$.

Proof: Define $F_n(t) = P(T_n \geq t)$. Now for $h > 0$, $F_n(t) - F_n(t+h) = P(t \leq T_n < t+h)$ which is the probability that the process first leaves state n in the time interval $[t_1 + t,\ t_1 + t + h)$, given that $N_{t_1} = n$. The process may leave state n by means of a birth or a death. Since $F_n(t)$ is the probability that the process has not left state n in $[t_1,\ t_1 + t)$, we have

$$P(t \leq T_n < t + h) = ((\lambda(t_1 + t) + \mu(t_1 + t))hn + o(h))F_n(t).$$

So,

$$\lim_{h \to 0} \frac{F_n(t+h) - F_n(t)}{h} = -(\lambda(t_1 + t) + \mu(t_1 + t))nF_n(t).$$

Replacing t by t - h in the above and using the fact that $\lambda(t)$ and $\mu(t)$ are continuous,

$$\lim_{h \to 0} \frac{F_n(t) - F_n(t-h)}{h} = -(\lambda(t_1 + t) + \mu(t_1 + t)nF_n(t)\ ,\ \text{so}$$

$F_n'(t) = (\lambda(t_1 + t) + \mu(t_1 + t))nF_n(t)$. The solution of this differential equation is $F_n(t) = e^{-n\int_0^t \lambda(t_1+s) + \mu(t_1+s)ds}$.

Now, $P(T_n < t) = 1 - F_n(t) = 1 - e^{-n\int_0^t \lambda(t_1+s) + \mu(s)ds}$.

# CHAPTER II

## AN INTEGRAL EQUATION

Population growth as a birth and death process may also be formulated in the following way. The state of the process at time t will be described by the integer valued function $N(x, t)$ which will specify the age distribution of the population in the sense that $\int_{x_1}^{x_2} dN(x, t)$ is the number of individuals in the age group $(x_1, x_2)$. The Stieltjes integration is with respect to the age variable x. The postulates for this formulation are as follows:

(a) The sub-populations generated by two coexisting individuals develop independently of one another.

(b) An individual of age x existing at time t has a chance $\lambda(x)h + o(h)$ of producing a new individual of age zero during the time interval $(t, t + h)$. The function $\lambda(x)$, called the birth rate, is not a function of t, and is assumed to be continuous.

(c) An individual of age x existing at time t has a chance $\mu(x)h + o(h)$ of dying during the time interval $(t, t + h)$. The function $\mu(x)$, called the death rate, is not a function of t, and is assumed to be continuous. Also, $N(x, 0)$ will be supposed given.

Consider the random variable $dN(x, t)$ which enumerates the individuals in the age group $(x, x + h)$ where $x < t$. The function $dN(x, t)$ is really a function of x, t, and h. One wants to consider this function for small h. Now, introduce the function $\alpha(x, t)$ by assuming that

$dN(x, t) = 0$ with probability $1 - \alpha(x, t)h + o(h)$

$dN(x, t) = 1$ with probability $\alpha(x, t)h + o(h)$

$dN(x, t) \geq 2$ with probability $o(h)$.

The function $\alpha(x, t)$ is required to be bounded in every finite rectangle. Hence, $E(dN(x, t)) = \alpha(x, t)h + o(h)$. Also, $\beta(x, t) = Var(dN(x, t)) = \alpha(x, t)h - \alpha^2(x, t)h^2 + o(h) = \alpha(x, t)h + o(h)$. Suppose that initially there is just one member of the population of age A so that $N(x, 0) = 0$ if $x < A$ and $N(x, 0) = 1$ if $x \geq A$.

In Kendall's [7] discussion of this model, he assumed $\mu(x)$ to be a constant function. In the following, $\mu(x)$ is only required to satisfy (c).

Define a random variable $z(t)$ by

$z(t) = 1$ if the single ancestor is still alive at time t

$z(t) = 0$ otherwise.

Let $F(t)$ be the probability that the ancestor is dead by time t. Then $F(t + h) - F(t) = (\mu(t)h + o(h))(1 - F(t))$ by postulate (c). Dividing by h and letting $h \to 0$, we obtain $\frac{F'(t)}{1 - F(t)} = \mu(t)$. Actually, this computation only shows differentiability to the left; however, replacing t by t - h in the preceding also gives differentiability to the right since $\mu(t)$ is assumed to be continuous and $o(h)$ is independent of t. The above differential equation has the solution $F(t) = 1 - e^{-\int_0^t \mu(s)ds}$. Hence, $P(z(t) = 1) = e^{-\int_0^t \mu(s)ds}$.

Consider the random variables $X_t$ and $X_{t-s}$ of the stochastic process $(X_t,\ t \geq 0)$ where $t > 0$, $s > 0$, $s < t$. We may ask for $E(X_t | X_{t-s} = x)$ which is the expected value of $X_t$ given that $X_{t-s}$ has assumed a certain value. This type of question involves the idea of conditioning by a random variable which is discussed by Parzen [9], pp. 41-53. Further, we may allow x to assume all possible values and ask for the expected value of $E(X_t | X_{t-s} = x)$.

Lemma. Given the preceding conditions, $E(X_t) = E(E(X_t | X_{t-s} = x))$ where $E(E(X_t | X_{t-s} = x))$ means taking the expected value of $X_t$ given that $X_{t-s} = x$ and then taking the expected value of the latter quantity over all possible x.

Proof: $E(E(X_t | X_{t-s} = x)) = \int E(X_t | X_{t-s} = x)dF(x)$ where F is the distribution function of $X_{t-s}$ and integration is Riemann-Stieltjes. But $\int E(X_t | X_{t-s} = x)dF(x) = E(X_t)$. Thus, the lemma states that $E(X_t)$ can be calculated in two steps; first calculate $E(X_t | X_{t-s} = x)$ and then calculate the expected value of this quantity over all x.

Consider dN(x, t) given the conditions at time t - x where x < t. We form $E(dN(x, t) \mid dN(0, t-x) = n)$ which is np since dN(x, t), given the value of dN(0, t - x), is binomial. In this case, n = dN(0, t - x) and $p = e^{-\int_0^x \mu(s)ds}$. We then find the expectation of $E(dN(x,t) \mid dN(0, t-x) = n) = np$. Now, $dN(0, t-x) = (\lambda(A + t-x)z(t-x) + \int_0^{t-x} \lambda(y)dN(y, t-x))h$ so that

$$E(np) = E(dN(0, t-x)p) = p(\lambda(A + t-x)e^{-\int_0^{t-x}\mu(s)ds} + \int_0^{t-x}\lambda(y)\alpha(y, t-x)dy)h.$$

According to the discussion in the preceding paragraph, the last quantity equals E(dN(x, t)) which we know to be $\alpha(x, t)h + o(h)$. Hence, we obtain

$$(14) \quad \alpha(x, t) = \lambda(A + t-x)e^{-\int_0^{t-x}\mu(s)ds - \int_0^x \mu(s)ds} + e^{-\int_0^x \mu(s)ds}\int_0^{t-x} \lambda(y)\alpha(y, t-x)dy.$$

Now, let $\Theta(t-x) = \alpha(x, t)e^{\int_0^x \mu(s)ds}$ so that

$$(15) \quad \Theta(v) = \lambda(A + v)e^{-\int_0^v \mu(s)ds} + \int_0^v \lambda(y)e^{-\int_0^y \mu(s)ds}\Theta(v-y)dy.$$

The above is a Volterra integral equation which is discussed by Feller [2]. The above functions satisfy the conditions of his Theorem 2; the conclusion is that there exists a unique non-negative solution which is bounded in every finite interval. In some cases, one can find the explicit solution of (15); one method of solution is by means of the Laplace transform [2]. However, there are many cases in which an explicit solution of (15), and hence of (14), is difficult to obtain; in these cases, it would be helpful to know if $\alpha(x, t) \to \alpha(x)$ as $t \to \infty$. Theorem 5 adopted from Feller [2] gives such information.

THEOREM 5. Suppose that we have $\Theta(v) = g(v) + \int_0^v \Theta(v-y)f(y)dy$ where f and g are non-negative continuous functions. Suppose that $\int_0^\infty f(x)dx = 1$, $\int_0^\infty g(x)dx = b < \infty$. Suppose further that there exists an integer $n \geq 2$ such that the moments $m_k = \int_0^\infty x^k f(x)dx$, $k = 1, \ldots, n$, are finite, and that the functions $f(x)$, $xf(x)$, ..., $x^{n-2}f(x)$ are of bounded total variation over $(0, \infty)$. Suppose finally that

$\lim_{x \to \infty} x^{n-2}g(x) = 0$ and $\lim_{x \to \infty} x^{n-2}\int_x^\infty g(s)ds = 0$. Then

$$\lim_{v \to \infty} \Theta(v) = b/m_1.$$

In an example to be considered later, we shall use this theorem to help find the limiting value of $\alpha(x, t)$.

In the formulation of the problem in Theorem 2, one is really looking at the population in a broad sense; the birth and death functions are the population birth and death functions. In the prior formulation, a much closer look is taken at the population; the birth and death functions here describe what happens to an individual. The preceding gives a loose idea of the different interpretations of the birth and death rates in the two formulations. In a situation in which the population is the members of a species, the second formulation may be the more accurate one. In many cases of interest in biology, it is possible to write down an expression for the Laplace transform of the solution of the integral equation occurring in the second formulation. However, it seems to be difficult to choose functions for $\lambda(x)$ and $\mu(x)$ which would be considered realistic by a biologist and at the same time allow an explicit solution of the equation. Two examples are given below. The first one is quite unrealistic; the second one is typical of some organisms.

Example 1. Suppose $\lambda(x)$ and $\mu(x)$ are the constant functions $\lambda$ and $\mu$. Equation (15) then becomes

$$\Theta(v) = \lambda e^{-\mu v} + \lambda\int_0^v e^{-\mu y}\Theta(v - y)dy.$$

Thus

$$L(\Theta) = \frac{\lambda}{s + \mu} + \frac{\lambda}{s + \mu}L(\Theta)$$

where L denotes the Laplace transform or

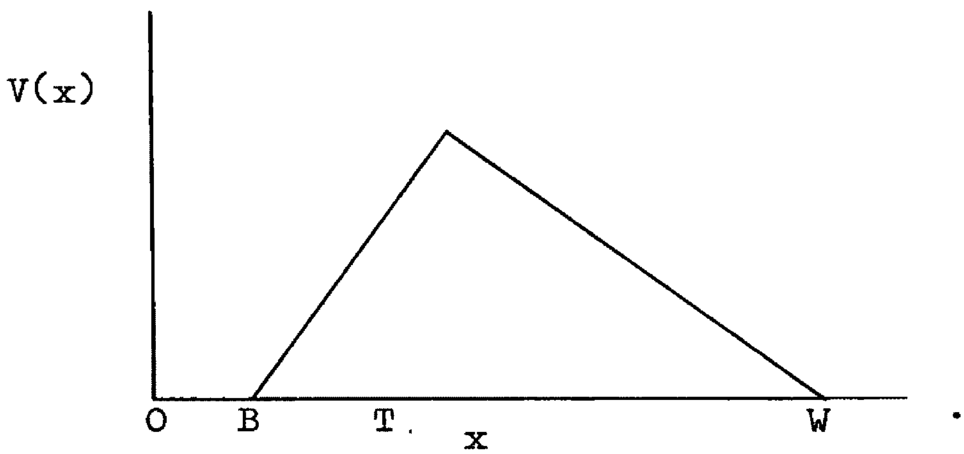
$$L(\Theta) = \frac{\frac{\lambda}{s + \mu}}{1 - \frac{\lambda}{s + \mu}} \quad \text{so that } \Theta(v) = \lambda e^{-(\mu-\lambda)v}.$$

This method of solving for $\Theta$ is standard Laplace transform procedure; the reader may consult Churchill [1], pp. 36-37. Since $v = t - x$ and $e^{-\mu x}\Theta(t-x) = \alpha(x, t)$, then $\beta(x, t) = \alpha(x, t) = \lambda e^{-(\mu-\lambda)t-\lambda x}$ when $x < t$.

A constant death rate corresponds to what is usually called a type III survivorship curve in population study. A type IV survivorship curve, which shows high infant mortality, is considered to be the more usual case. A constant birth rate over all time does not occur in a living population.

Since $\lambda(x)$ is the chance of an individual aged x producing another individual and $e^{-\int_0^x \mu(s)ds}$ is the chance that an individual will live to age x, we see that $\int_0^\infty \lambda(x) e^{-\int_0^x \mu(s)ds} dx = R$ is the contribution of an individual to the population of which he is a member. We infer that the population size is increasing if $R > 1$, decreasing if $R < 1$, and remaining constant if $R = 1$. Biological populations that exist for a very long period of time must show an approximate zero rate of increase in size as a long term average; the condition $R = 1$ is equivalent to a zero rate of increase.

Example 2. Let $V(x) = \lambda(x) e^{-\int_0^x \mu(s)ds}$. Instead of specifying the form of $\lambda(x)$ and $\mu(x)$ which would then determine $V(x)$, we may bypass $\lambda(x)$ and $\mu(x)$ and simply specify the form of $V(x)$. For some organisms, there have been experimental determinations of the function $V(x)$; it is approximately triangular in shape (see Lewontin [8], pp. 77-94). We define $V(x)$ by the following diagram:




B, T, and W are the age of first reproduction, the age of peak reproduction, and the age of final reproduction, respectively. We have that $1 = \int_0^\infty V(x)dx = \frac{(W - B)}{2}V(T)$, so that $V(T) = 2/(W - B)$. From similar triangles,

$$\frac{V(x)}{x - B} = \frac{V(T)}{T - B} \qquad B \leq x < T$$

$$\frac{V(x)}{W - x} = \frac{V(T)}{W - T} \qquad T \leq x \leq W \quad \text{so that}$$

$$V(x) = \frac{2(x - B)}{(T-B)(W-B)} \qquad B \leq x < T$$

$$V(x) = \frac{2(W - x)}{(W-T)(W-B)} \qquad T \leq x \leq W$$

$V(x) = 0$ otherwise.

We wish to use Theorem 5 to find $\Theta(v)$ as $v \to \infty$. Now, $\int_0^\infty xV(x)dx = 1/3(B + T + W)$ which is $m_1$ of Theorem 5. If A is approximately zero and $\lambda(v)$ and $e^{-\int_0^v \mu(s)ds}$ are such that $\lambda(v + A)e^{-\int_0^v \mu(s)ds} = V(v + A)$, then $\int_0^\infty V(v + A)dv = \frac{W - B}{2}V(T - A)$. Thus, Theorem 5 asserts that $\Theta(v) \to \frac{3(W-B)V(T-A)}{2(B+T+W)}$ as $v \to \infty$. But $v = t - x$ and $\Theta(t - x) = \alpha(x, t)e^{\int_0^x \mu(s)ds}$; hence, for fixed x as $t \to \infty$, $\beta(x, t) = \alpha(x, t) \to \frac{3(W-B)V(T-A)}{2(B+T+W)}e^{-\int_0^x \mu(s)ds}$. If there is more than one individual, the means and variances add since sub-populations are assumed independent. In practice, one may be able to assume that t is large enough for the above formula to be approximately correct.

## REFERENCES CITED

(1) Churchill, R. V., "Modern Operational Mathematics in Engineering", McGraw-Hill, 1944.

(2) Feller, W., "On the Integral Equation of Renewal Theory", Annals of Mathematical Statistics, Vol. 12 (1941), pp. 243-267.

(3) Feller, W., "An Introduction to Probability Theory and its Applications," Vol. I, John Wiley and Sons, 1957.

(4) Forsyth, A. R., "A Treatise on Differential Equations", MacMillan and Co., 1921.

(5) Ince, E. L., "Ordinary Differential Equations", Dover Publications, 1944.

(6) Kendall, D. G., "On the Generalized Birth and Death Process," Annals of Mathematical Statistics, Vol. 19 (1948), pp. 1-15.

(7) Kendall, D. G., "Stochastic Processes and Population Growth," Journal of the Royal Statistical Society, B, Vol. 11 (1949), pp. 230-264.

(8) Lewontin, R. C., "Selection for Colonizing Ability", in The Genetics of Colonizing Species, eds. H. G. Baker and G. L. Stebbins, Academic Press, (1965), pp. 77-94.

(9) Parzen, E., "Stochastic Processes", Holden-Day, Inc., 1965.