

https://dspace.org/

Modelling and control of birth and death processes

Getz, Wayne Marcus

2015-01-29
http://hdl.handle.net/10539/16803

Downloaded from a DSpace repository. https://dspace.org/

MODELLING AND CONTROL OF BIRTH AND DEATH PROCESSES

by

Wayne Marcus Getz

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy

February 1976

A B S T R A C T

This thesis treats systems of ordinary differential equations that are extracted from the Kolmogorov forward equations of a class of Markov processes, known generally as birth and death processes. In particular we extract and analyze systems of equations which describe the dynamic behaviour of the second-order moments of the probability distribution of population governed by birth and death processes. We show that these systems form an important class of stochastic population models and conclude that they are superior to those stochastic models derived by adding a noise term to a deterministic population model. We also show that these systems are readily used in population control studies, in which the cost of uncertainty in the population mean size is taken into account.

The first chapter formulates the univariate linear birth and death process in its most general form. The probability distribution for the constant parameter case is obtained exactly, which allows one to state, as special cases, results on the simple birth and death, Poisson, Pascal, Polya, Palm and Arley processes. Control of a population, modelled by the linear birth and death process, is considered next. Particular attention is paid to system performance indeces which take into account the cost associated with non-zero variance and the cost of improving initial estimates of the size of the population under control.

Chapter 2 is devoted to a broad class of multivariate birth and death processes restricted to the instantaneous occurrence of single events only. The approach taken in deriving ordinary differential

equations is generalized. In particular two birth and death processes are examined in depth. A system of linear ordinary differential equations is derived as a stochastic model for populations governed by the general multivariate linear birth and death process, while consideration of a non-linear process leads to a stochastic model, directly related to the quadratic Lotka-Volterra model often used in the study of the dynamics of interacting populations.

A non-linear multiple migration process is formulated in chapter 3. This study essentially extends the methods discussed in chapter 2 to multivariate birth and death processes in which the instantaneous occurrence of multiple events, such as multiple births and migration of families, is allowed.

In chapter 4 we study the linear system of ordinary differential equations, derived in chapter 2, as a population model, and investigate its use in optimal control studies. In particular, we discuss the application of such systems to harvesting populations, which are divided into distinct age or size classes, taking into account the use of control variables that enter the system in a linear and bilinear way. In this discussion we invoke Pontryagin's maximum principle and certain results from linear-quadratic regulator theory. As a demonstration, the solution to, a theoretical harvesting problem is solved numerically and the optimal control turns out to be bang-bang.

In chapter 4 the non-linear systems of ordinary differential equations derived in chapters 2 and 3 are discussed more fully. The stability and equilibrium properties of solutions of these systems are compared with results obtained for directly related deterministic systems. In this analysis it emerges that these

deterministic system models give biased estimates of the population means as predicted by the stochastic system models. This result directly questions the validity of the deterministic system model.

The thesis concludes with a brief résumé of the contribution made in this work to modelling and control studies in population dynamics. Two facts which emerge are:

(i) very little progress has, up to now, been made in analyzing multivariate birth and death processes, so that the systems of equations derived in chapters 2 and 3 provide an important method for solving for the first two moments of the joint multivariate probability distributions of these processes;

(ii) these systems are *superior in a number of respects* to stochastic systems generated by adding a noise term to a deterministic system: firstly because they contain more information on the nature of the noise of the population process; *secondly because* they can be utilized more easily in control studies, and thirdly because the bias in the estimate of the population means, as given by the solution to the deterministic system, is avoided.

A C K N O W L E D G E M E N T S

I wish to thank:

- Professor D.H. Jacobson for teaching me optimal control theory, guiding me through my formative research years and giving up his time to supervise this thesis;

- Messrs. African Explosives and Chemical Industries for their generous financial support during 1974;

- the Council for Scientific and Industrial Research for supporting my work in full during 1975;

- Mr. F.R. Baudert for his careful editing;

- Miss L. Rolfes for writing the computer program used in chapter 4;

- Mrs. Marie de Villiers for her excellent typing of the manuscript;

- my wife Jennifer who had more confidence in me than anyone else.

C O N T E N T S

ABSTRACT ..... (i)

ACKNOWLEDGEMENTS ..... (iv)

GENERAL INTRODUCTION ..... (vii)

CHAPTER 1 FORMULATION OF A GENERAL UNIVARIATE LINEAR BIRTH AND DEATH PROCESS

1.1 Introduction 1

1.2 Formulation of the process - the development of a general model 2

1.3 General solution distribution 4

1.4 Constant parameter solution distribution 9

1.5 Finding the mean and variance 11

1.6 Direct derivation of mean and variance 15

1.7 Unrestricted constant control 17

1.8 Maintaining a constant mean 19

1.9 Variance for constant parameters and control 20

1.10 Deriving a cost performance index 22

1.11 Minimizing the cost performance index 31

1.12 Piecewise constant parameters and the minimization problem 34

CHAPTER 2 FORMULATION OF A GENERAL MULTIVARIATE BIRTH AND DEATH PROCESS

2.1 Introduction 39

2.2 Definitions 41

2.3 General formulation 43

2.4 Building specific models 48

2.5 Method of solution 51

CHAPTER 3 MULTIPLE-EVENT MULTIVARIATE BIRTH AND DEATH PROCESS

3.1 Introduction 65

3.2 Formulation of a non-linear multiple-event migration process 67

3.3 Differential equations for the moments of the distribution 74

CHAPTER 4 THE LINEAR MULTIVARIATE BIRTH AND DEATH PROCESS POPULATION MODEL

4.1 The general linear stochastic population model 81

4.2 Direct derivation and theoretical validation of the stochastic model using Kronecker products 85

4.3 A continuous time stochastic population model 90

4.4 A bilinear scalar control harvesting model 95

4.5 A linear-quadratic Leslie matrix harvesting model 103

4.6 Numerical evaluation of a harvesting programme using linear scalar control 113

CHAPTER 5 NON-LINEAR MULTIVARIATE BIRTH AND DEATH PROCESS POPULATION MODELS

5.1 Introduction 125

5.2 The stochastic Lotka-Volterra and multiple-migration population models 127

5.3 The existence of explosive solutions in quadratic models 130

5.4 A comparison between solutions to the deterministic and stochastic models 136

CHAPTER 6 C O N C L U S I O N 145

R E F E R E N C E S 148

GENERAL INTRODUCTION

> "..... the power of population is indefinitely greater than the power in the earth to produce subsistence for man.
>
> Population, when unchecked, increases in a geometrical ratio. Subsistence increases only in an arithmetical ratio. A slight acquaintance with numbers will shew the immensity of the first power in comparison of the second."†

The efficient management of the world's renewable resources is a problem for which we have only naïve solutions, despite dramatic technological advances during this quarter-century. Fundamental to an understanding of the influence we have on these resources, is a knowledge of the dynamic behaviour of their constitutive populations.

This thesis is concerned with the development of techniques for modelling the dynamic behaviour of populations, and the use of population models as management tools for formulating optimal exploitation policies. A population is considered to be a homogeneous group of individuals: homogeneous in the sense that the behaviour of each individual group in the population is governed by the same probability distributions. The modelling techniques employed are based on a class of stochastic processes known as birth and death processes.

† T. Malthus (1798). An Essay on the Principle of Population p. 13 (First edition). Printed for J. Johnson in St. Pauls Churchyard London.

The first chapter is concerned with the dynamics of populations whose behaviour is assumed to be modelled by a simple univariate linear birth and death process. The linear birth and death process has been extensively used to model the behaviour of a single population subject to birth, death and migration processes. The simple univariate linear birth and death process in its most general form is a four-parameter process. The four parameters govern the increase and decrease of the population at rates proportional to and independent of the size of the population. Many special forms of this process have been studied; from the one-parameter Poisson processes, to the three-parameter simple birth, death and immigration processes. A comprehensive review of these simple linear processes can be found in Goel and Richter-Dyn (1974).

In the first chapter of this thesis we derive the probability distribution for the general four-parameter process. This generalizes all previously known results on the probability distributions of the various simple linear processes. A new interpretation is also given to the two parameters describing the rates of increase and decrease that are independent of population size, viz. in terms of positive and negative control, respectively. The actual physical interpretation of these control parameters will vary, depending on the nature of the population and the type of control that can be applied, but these parameters may, for example, be related to controlled migration policies. It is possible that the remaining two parameters of the model can also be manipulated, but as no essentially new ideas are involved, this extension is not pursued. Control of a constant parameter and piecewise constant parameter process are studied. Particular attention is paid to deriving a cost performance index where a cost is associated

with the non-zero variance, and estimation, of the population size. The results contained in this chapter have appeared in Getz (1975a).

Although non-linear considerations are a natural extension to the model formulated in chapter 1, a more general extension is made in chapter 2 to linear and non-linear single-event multivariate processes. Multivariate processes are important in modelling a system in which a number of basically different groups of individuals (species, colonies, classes within a given population) interact to influence each other's behaviour. This sort of situation arises, for example, in predator-prey systems, systems of populations competing for the same food source and single populations whose life cycle can be divided into distinct developmental stages.

Although systems of interacting populations are very common, multivariate processes have been subjected to comparatively little analysis. A number of results have, however, been derived for special types of bivariate systems. Billard (1974) has obtained results for bivariate processes in which deaths, but not births, are allowed to occur. Puri (1975) has obtained results for a bivariate process in which one population is influenced by the other but not vice versa. Tsokos and Hinkley (1973) obtained expressions for the conditional distributions of two interacting populations, but these results, as pointed out by Beyer and Keiding (1975) and Cormack and Newton (1975), are incorrect. In fact Cormack and Newton (1975) go as far as to state that it is questionable whether the approach used by Tsokos and Hinkley can ever be successfully used, although Tsokos (1975) has argued to the contrary.

The complete solution to describing the behaviour of a multivariate

population process lies in finding its joint multivariate distribution. This is a difficult problem and little success has been achieved. Notably, however, Boswell and Patil (1972) have obtained the distributions for a class of multivariate linear birth and death processes but only under the initial conditions of zero members in each population. In this second chapter we develop a method, applicable to the general multivariate linear process and a class of non-linear processes, of obtaining the most important (usually first and second-order) moments of the joint multivariate distribution of the process. This is achieved by deriving a system of ordinary differential equations in the first- and second-order moments from the Kolmogorov forward equations of the processes. The procedure is demonstrated by deriving the system of equations for the linear process and for a specific non-linear process. In the non-linear process, the system of equations obtained turns out to be directly related to the deterministic Lotka-Volterra population model. The results presented in this chapter are also in Getz (1976).

The third chapter extends the scope of the formulation discussed in the second chapter, to include the occurrence of instantaneous multiple events such as multiple births and the migration of families. The Kolmogorov forward equation for a non-linear multiple migration process is derived and as in the previous chapter a system of ordinary differential equations is obtained, describing the first- and second-order moments of the joint multivariate distribution of the process. The system discussed in this third chapter has application to modelling migration processes between a system of colonies (populations). Although a number of papers have appeared dealing with migration processes between colonies, the approach taken in this chapter allows for generalizacion that have not been dealt with before: simultaneous

migration of a group of individuals and a migration rate dependent upon the size of the colony to which the group is migrating. The results obtained in this section may be found in Getz (1975c).

In the fourth chapter the system of linear ordinary differential equations arising from the Kolmogorov forward equation of the multivariate linear birth and death process is presented as a matrix population model. First we show that the equations describing the second-order moments of the probability distribution can, using the Kronecker matrix product, be generated from the equations describing the first-order moments. The method of generating the equations describing the second-order moments is shown to be equivalent to the method developed by Pollard (1973) for discrete time systems. In addition we show that this method is valid for a larger class of processes than the class of processes considered by Pollard.

We then formulate a particular specialization of the linear matrix model, that has applications in the management of populations structured into classes (age, size, etc.). The use of this model in deriving management policies for the optimal harvesting of fish or control of pest populations is discussed in detail. In particular both the case where control enters the state equations of the system (model) linearly and where it enters bilinearly are discussed and a number of results from optimal control theory invoked. Finally the ideas discussed in this chapter on the optimal control of population, using the system derived in chapter 2, are demonstrated through the numerical solution of a simple problem based on fictitious data.

In chapter 5 the non-linear systems of equations derived in chapters 2 and 3 are analyzed and compared with deterministic systems to which they are directly related. These non-linear systems possess

quadratic terms and under certain conditions their solutions explode in a finite time. Using a theorem derived by Getz and Jacobson (1975) we show that an explosive solution of the deterministic system is *actually a lower* bound (in the sense of its norm) to the solution, with some initial condition, of the stochastic system. The stability properties of the deterministic system are then *discussed and its* equilibrium points compared with those of the stochastic system.

Chapter 6 concludes the thesis with a brief résumé of the material covered and the results obtained. The advantages are discussed of using the modelling formulations developed in this thesis but the need is also stressed for evaluating these techniques in future research in this area.

Finally the principal contributions of this thesis to modelling and control studies in population dynamics can be summarized as follows.

(i) (a) The probability distribution for the general univariate simple birth and death process is derived;

(b) strategies are formulated for controlling populations modelled by this process; and

(c) cost performance criteria are developed which penalize *uncertainty associated* with the mean size of the population.

(ii) A methodology is developed for modelling stochastic systems of interacting populations whose behaviour can be described by a Markov jump process. In particular the approach allows linear and non-linear interactions and the instantaneous occurrence of multiple events.

(iii) A continuous time matrix model for structured stochastic populations is derived. This model is shown to be suitable for

population harvesting studies and further allows the stochastic elements of the population to be included in the formulation of optimal control strategies.

(iv) Pollard's results on generating the second-order moments of a *multitype Galton-Watson process* (Pollard 1973) are rederived for continuous time systems and are shown to hold as well for multivariate linear birth, death and migration processes.

(v) A stochastic analogue of the deterministic Lotka-Volterra model is developed. Further it is demonstrated that the deterministic model gives a biased estimate of the size of the population.

(vi) Stability results for quadratic-type population models are obtained, particularly results related to the existence of explosive solutions.

CHAPTER I

FORMULATION OF A GENERAL UNIVARIATE LINEAR BIRTH AND DEATH PROCESS

1.1 INTRODUCTION

The formulation of a birth and death type process as pioneered by Yule, Feller and Kendall among others, has been successfully used to model the behaviour of stochastic populations. Recent examples of the application of this technique to finding the probability distribution of the size of the population being modelled, are many and varied. They include modelling multiple births (Doubleday 1973), competing species (Tsokos and Hinkley 1973) and birth death and migration processes (Renshaw 1972) as well as models in which differentiation between the sexes is incorporated (Keyfitz 1968).

In general the object of work done in this field is to find an expression to describe how the probability distribution of the size of the population will vary with time. It appears, however, that very little analysis has been done on the control of a population in order to steer its distribution towards a desired target and at the same time minimize some cost performance criterion. An isolated example, however, is Becker (1970). It should also be noted, though, that there are a number of papers in the ecological field that consider optimization problems (Sancho 1973 and Shoemaker 1973).

In this chapter a birth and death population process is formulated to include positive and negative control parameters. The general solution for the distribution of the size of the population at any instant in time, is obtained in the form of a probability generating function. The forms of the mean and variance are derived for constant birth and death parameters, and the values of the control parameters which steer the mean towards a target value are obtained. Optimal control of this process is discussed with respect to minimizing a cost performance index. The cost of variance and the cost of determining the initial distribution of the population are taken into account. The analysis is extended to include piecewise constant parameters.

1.2 FORMULATION OF THE PROCESS - THE DEVELOPMENT OF A GENERAL MODEL

Let n, the number of members in a population, be a stochastic variable taking on the values 0,1,2 ... at P(n;t) denote the probability that the population is c at time t. Let $\Delta t > 0$ denote a small increment in time.

If the population is of size n at time t, then during the interval $[t,t+\Delta t]$ it is assumed that any of the following five events may occur:

(i) Each individual present at time t may give rise to an additional individual with probability $\lambda(t)\Delta t + o(\Delta t)$.

(ii) Each individual present at time t may leave the population (die or emigrate etc.) with probability $\mu(t)\Delta t + o(\Delta t)$.

(iii) Positive control is applied to the population, in which case an individual may be added to the population (immigrate) with probability $\nu(t)\Delta t + o(\Delta t)$.

(iv) Negative control is applied to the population, in which case an individual may be removed (culling, forced emigration) from the population with probability $\alpha(t)\Delta t + o(\Delta t)$.

(v) None of (i) - (iv) occur, and the population level remains the same.

We also assume that the probability of more than one of the events (i) - (iv) occurring is $o(\Delta t)$. Under these conditions we derive the following relation ship:

$$\begin{aligned} P(n;t+\Delta t) = {} & \{(n-1)[\lambda(t)\Delta t+o(\Delta t)] + \nu(t)\Delta t + o(\Delta t)\}P(n-1;t) \\ & + \{1-n[\lambda(t)\Delta t+\mu(t)\Delta t+o(\Delta t)] - [\nu(t)\Delta t+\alpha(t)\Delta t+o(\Delta t)]\}P(n;t) \\ & + \{(n+1)[\mu(t)\Delta t+o(\Delta t)] + \alpha(t)\Delta t+o(\Delta t)\}P(n+1;t) + o(\Delta t), \end{aligned} \tag{1.2.1}$$

which on letting $\Delta t \to 0$ becomes the following differential equation:

$$\begin{aligned} P'(n;t) = {} & [(n+1)\mu(t) + \alpha(t)]P(n+1;t) - [n(\mu(t)+\lambda(t))+\alpha(t)+\nu(t)]P(n;t) \\ & + [(n-1)\lambda(t)+\nu(t)]P(n-1;t) \\ & n=1,2,3,\ldots, \end{aligned} \tag{1.2.2.1}$$

where ' denotes differentiation with respect to t. Since negative control cannot exist when n=0, we have

$$\frac{dP}{dt}(0;t) = [\mu(t) + \alpha(t)]\,P(1;t) - \nu(t)P(0;t). \qquad (1.2.2.2)$$

If at time t=0 we know that the population has a distribution

$$P(i;0) = \gamma_i \qquad i=0,1,2,\ldots \qquad (1.2.3.1)$$

$$\sum_{i=0}^{\infty} \gamma_i = 1 \qquad (1.2.3.2)$$

(because P(i;0) is a distribution),

$$\text{mean} = n_0, \qquad (1.2.3.3)$$

$$\text{variance} = \sigma_0^2, \qquad (1.2.3.4)$$

we can solve (1.2.2.') subject to (1.2.3.').

Taking one or more of the parameters $\lambda(t)$, $\mu(t)$, $\nu(t)$ and $\alpha(t)$ equal to zero, we obtain the equations for the Poisson, Pascal, Polya, pure birth and death, Palm and Arley processes as special cases. The most general of these is Arley's birth and death with immigration model (Arley 1967) in which only $\alpha(t)$ of the four parameters is zero.

1.3 GENERAL SOLUTION DISTRIBUTION

Define a probability generating function

$$G(s;t) = \sum_{n=0}^{\infty} P(n;t)s^n \qquad (1.3.1)$$

Then

$$\frac{\partial G(s;t)}{\partial t} = \sum_{n=0}^{\infty} P'(n;t)s^n \qquad (1.3.2)$$

$$\frac{\partial G(s;t)}{\partial s} = \sum_{n=0}^{\infty} n P(n;t)s^{n-1} \qquad (1.3.3)$$

$$\frac{\partial^2 G(s;t)}{\partial s^2} = \sum_{n=0}^{\infty} n(n-1)P(n;t)s^{n-2}. \qquad (1.3.4)$$

Obviously (1.3.1) converges for $|s| \leq 1$, since $\sum_{n=0}^{\infty} P(n;t) = 1$. In addition (1.3.3) and (1.3.4) exist for $|s| \leq 1$ provided the mean and variance of the distribution are finite for all finite t. No attempt will be made to be mathematically rigorous, as the physics of the problem justifies most mathematical steps. The reader can refer to Takacs (1972) for a rigorous mathematical account of general queueing processes.

Multiplying equation (1.2.2.1) by s^n and summing over all n, we have, using (1.3.1), (1.3.2) and (1.3.3),

$$\frac{\partial G}{\partial t} = (\lambda(t)s - \mu(t))(s-1)\frac{\partial G}{\partial s} + \left(\nu(t) - \frac{\alpha(t)}{s}\right)(s-1)G. \qquad (1.3.5)$$

where for tractability the arguments of G(s;t) have been dropped. The method of solution to (1.3.5) is sketched below for completeness. The full theory can be found in Piaggio (1956).

The auxilliary equations corresponding to (1.3.5) are

$$\frac{ds}{(\lambda(t)s-\mu(t))(s-1)} = \frac{dt}{-1} = \frac{dG}{(a(t)/s-v(t))(s-1)G}. \qquad (1.3.6)$$

We first find a function

$$U(s;t) = \text{constant}$$

by solving

$$\frac{ds}{dt} = -(s-1)(\lambda(t)s-\mu(t)). \qquad (1.3.7)$$

Making the substitution $s = 1+1/z$ we obtain

$$\frac{dz}{dt} - (\lambda(t)-\mu(t))z = \lambda(t).$$

The integrating factor for this equation is

$$\rho(t) = \exp\{\int_0^t [\mu(\tau)-\lambda(\tau)]\,d\tau\}, \qquad (1.3.8)$$

so that

$$z\rho(t) = \int_0^t \lambda(\tau)\rho(\tau)d\tau + k, \qquad (1.3.9)$$

where k is a constant of integration. Finally, since $z = 1/(s-1)$, we have

$$U(s;t) = \frac{\rho(t)}{s-1} - \int_0^t \lambda(\tau)\rho(\tau)d\tau = k \qquad (1.3.10)$$

as required.

Manipulating (1.3.10), we find that

$$s-1 = \frac{\rho(t)}{k + \int_0^t \lambda(\tau)\rho(\tau)d\tau}, \tag{1.3.11}$$

or

$$s = \frac{k + \rho(t) + \int_0^t \lambda(\tau)\rho(\tau)d\tau}{k + \int_0^t \lambda(\tau)\rho(\tau)d\tau}. \tag{1.3.12}$$

We now find a function

$$V(G,t,k) = \text{constant}$$

by solving

$$\frac{dG}{dt} = \left(\nu(t) - \frac{\alpha(t)}{s}\right)(s-1)G \tag{1.3.13}$$

and substituting for s and s-1 in (1.3.13) from (1.3.11) and (1.3.12). Since

$$\int \frac{dG}{G} = \log G(s;t) - K,$$

where K is the constant of integration, we have

$$\log G(s;t) - \int_0^t \left[\frac{\nu(\tau)\rho(\tau)}{k + \int_0^\tau \lambda(\omega)\rho(\omega)\,d\omega} - \frac{a(\tau)\rho(\tau)}{k + \rho(\tau) + \int_0^\tau \lambda(\omega)\rho(\omega)\,d\omega} \right] d\tau = K \qquad (1.3.14.1)$$

- i.e., we have found

$$V(G,t,k) = K, \qquad (1.3.14.2)$$

where k given by (1.3.10) is replaced after the integration has been performed.

The general solution to (1.3.5) is given by

$$V(G,t,k) = f(U(s;t)), \qquad (1.3.15)$$

where f is an arbitrary function determined from the initial conditions. From (1.2.3.1) and (1.3.1),

$$G(s;0) = \sum_{n=0}^{\infty} \gamma_n s^n,$$

and from (1.3.14.1) and (1.3.14.2) for t = 0 we have

$$V(G,0,k) = \log G(s;0)$$

$$= \log \sum_{n=0}^{\infty} \gamma_n s^n. \qquad (1.3.16)$$

From (1.3.8) and (1.3.10), since $\rho(0) = 1$, we have

$$(s;0) = \frac{1}{s-1} , \tag{1.3.17}$$

which leads from equations (1.3.15) and (1.3.16) to

$$f\left(\frac{1}{s-1}\right) = \log \sum_{n=0}^{\infty} \gamma_n s^n ,$$

giving the general form of f as

$$f(\omega) = \log \sum_{n=0}^{\infty} \gamma_n \left(1 + \frac{1}{\omega}\right)^n , \tag{1.3.18}$$

$$\omega = \frac{1}{s-1} .$$

So finally, from (1.3.15) for all t, using (1.3.10), (1.3.14) and (1.3.18), we have that

$$G(s;t) = \exp \int_0^t \left[\frac{\nu(\tau)\rho(\tau)}{k + \int_0^\tau \lambda(u)\rho(u)du} - \frac{\alpha(\tau)\rho(\tau)}{k + \rho(\tau) + \int_0^\tau \lambda(u)\rho(u)du} \right] d\tau$$

$$\times \sum_{n=0}^{\infty} \gamma_n \left(1 + \frac{1}{k}\right)^n , \tag{1.3.19}$$

where k given by (1.3.10) is replaced after the integration has been performed.

1.4 CONSTANT PARAMETER SOLUTION DISTRIBUTION

Let the parameters be constant for all t so that

$$\lambda(t) = \lambda$$

$$\mu(t) = \mu$$

$$\nu(t) = \nu$$

$$\alpha(t) = \alpha.$$

Then from (1.3.8),

$$\rho(t) = e^{(\mu-\lambda)t}, \tag{1.4.1}$$

and

$$\int_0^t \lambda(\tau)\rho(\tau)d\tau = \frac{\lambda}{\mu-\lambda}\left(e^{(\mu-\lambda)t} - 1\right). \tag{1.4.2}$$

From (1.3.10), we have that

$$k = \left[\frac{1}{s-1} - \frac{\lambda}{\mu-\lambda}\right] e^{(\mu-\lambda)t} + \frac{\lambda}{\mu-\lambda} \quad , \tag{1.4.3}$$

which after adding unity to both sides, becomes

$$k+1 = \left[\frac{1}{s-1} - \frac{\lambda}{\mu-\lambda}\right] e^{(\mu-\lambda)t} + \frac{\mu}{\mu-\lambda} \quad . \tag{1.4.4}$$

Thus (1.3.19) becomes

$$G(s;t) = \exp\left[\nu \int_0^t \frac{e^{(\mu-\lambda)\tau}}{k + [\lambda/(\mu-\lambda)]\left(e^{(\mu-\lambda)\tau} - 1\right)} d\tau \right.$$
$$\left. - \alpha \int_0^t \frac{e^{(\mu-\lambda)\tau}}{k + e^{(\mu-\lambda)\tau} + [\lambda/(\mu-\lambda)]\left(e^{(\mu-\lambda)\tau} - 1\right)} d\tau\right] \left[\sum_{n=0}^{\infty} \gamma_n \left[\frac{k}{k+1}\right]^n\right]. \tag{1.4.5}$$

Using the identity

$$\int \frac{dt}{c+be^{at}} = \frac{1}{ac} \log \frac{be^{at}}{c+be^{at}} + \text{const},$$

we can integrate (1.4.5), which after simplification and substitution for k becomes

$$G(s;t) = \left[\frac{(\mu-\lambda s)+\lambda(s-1)e^{(\lambda-\mu)t}}{\mu-\lambda}\right]^{-\nu/\lambda} \times \left[\frac{(\mu-\lambda s)+\mu(s-1)e^{(\lambda-\mu)t}}{s(\mu-\lambda)}\right]^{\alpha/\mu} \times \sum_{n=0}^{\infty} \gamma_n \left[\frac{\mu-\lambda s+\mu(s-1)e^{(\lambda-\mu)t}}{\mu-\lambda s+\lambda(s-1)e^{(\lambda-\mu)t}}\right]^n \quad . \qquad (1.4.6)$$

Thus we have found the form of the generating function for the distribution of the size of the population at any time t, which here, stated in its most general form for constant parameters, appears to be a new result. However, the task of finding a particular P(n;t), i.e., finding the coefficient of s^n in (1.4.6), poses a difficult problem.

1.5 FINDING THE MEAN AND VARIANCE

The two most important moments of a distribution are the mean and variance, and these can be found quite easily.

If we put s = 1 in (1.3.3) and (1.3.4) we obtain

$$\frac{\partial G}{\partial s}(1;t) = \sum_{n=0}^{\infty} nP(n;t) \triangleq E(n), \tag{1.5.1}$$

$$\frac{\partial^2 G}{\partial s^2}(1;t) = \sum_{n=0}^{\infty} n(n-1)P(n;t) \triangleq E[n(n-1)], \tag{1.5.2}$$

which are the mean and second factorial moment of the distribution, respectively. Let $\bar{n}(t)$ denote the mean at time t and $\sigma^2(t)$ the variance at time t. Since the variance is the second moment of the distribution about the mean, we have

$$\sigma^2(t) = E[n(n-1)] + \bar{n}(t) - \bar{n}^2(t). \tag{1.5.3}$$

To facilitate the manipulations involved in differentiating (1.4.6) once and then twice and setting $s = 1$ after each differentiation, we can approach the problem step by step as follows: Let

$$f(s;t) = (\mu-\lambda s) + \lambda(s-1)e^{(\lambda-\mu)t}, \tag{1.5.4}$$

$$h(s;t) = (\mu-\lambda s) + \mu(s-1)e^{(\lambda-\mu)t}. \tag{1.5.5}$$

Then

$$f(1;t) = \mu-\lambda,$$

$$f'(1;t) = -\lambda+\lambda e^{(\lambda-\mu)t},$$

$$f''(1;t) = 0,$$

$$h(1;t) = \mu-\lambda,$$

$$h'(1;t) = -\lambda+\mu e^{(\lambda-\mu)t},$$

$$h''(1;t) = 0.$$

Let

$$F(s;t) = \frac{f(s;t)}{u-\lambda}, \tag{1.5.6}$$

$$H(s;t) = \frac{h(s;t)}{s(\mu-\lambda)}, \tag{1.5.7}$$

$$R(s;t) = \frac{h(s;t)}{f(s;t)}. \tag{1.5.8}$$

Then

$$F(1;t) = 1,$$

$$F'(1;t) = \frac{\lambda}{\lambda-\mu}\left[e^{(\lambda-\mu)t} - 1\right],$$

$$F''(1;t) = 0,$$

$$H(1;t) = 1,$$

$$H'(1;t) = \frac{\mu}{\mu-\lambda}\left[e^{(\lambda-\mu)t} - 1\right],$$

$$H''(1;t) = -\frac{2\mu}{\mu-\lambda}\left[e^{(\lambda-\mu)t} - 1\right],$$

$$R(1;t) = 1,$$

$$R'(1;t) = e^{(\lambda-\mu)t},$$

$$R''(1;t) = -\frac{2\lambda}{\mu-\lambda}\left[e^{(\lambda-\mu)t} - 1\right].$$

Rewriting (1.4.6) in terms of (1.5.6) and (1.5.8), using (1.5.4) and (1.5.5), we have

$$G(s;t) = F(s;t)^{-\nu/\lambda} H(s;t)^{\alpha/\mu} \{ \sum_{r=0}^{\infty} \gamma_n R(s;t)^n \}. \qquad (1.5.9)$$

Recalling equations (1.2.3.·) and using (1.5.3) we have that

$$\sum_{n=0}^{\infty} n\gamma_n = n_0, \qquad (1.5.10)$$

$$\sum_{n=0}^{\infty} n(n-1)\gamma_n = \sigma_0^2 - n_0 + n_0^2. \qquad (1.5.11)$$

We can now proceed to find $\bar{n}(t)$ and $\sigma^2(t)$ by differentiating (1.5.6) once and then twice and setting $s = 1$. Using the preceding analysis we finally obtain

$$\bar{n}(t) = n_0 e^{(\lambda-\mu)t} + \frac{\alpha-\nu}{\mu-\lambda} (e^{(\lambda-\mu)t} - 1). \qquad (1.5.12)$$

$$\sigma^2(t) = \sigma_0^2 e^{2(\lambda-\mu)t} + \frac{\nu\lambda - \alpha\mu}{(\mu-\lambda)^2} (e^{\lambda-\mu)t} - 1)^2$$

$$+ \frac{\lambda+\mu}{\lambda-\mu} n_0 e^{(\lambda-\mu)t} (e^{(\lambda-\mu)t} - 1). \qquad (1.5.13)$$

These results, giving the explicit time-varying trajectory of the mean and variance, will be crucial to the ensuing analysis, as the mean and variance of the population are the two most important state variables of the model.

1.6 DIRECT DERIVATION OF MEAN AND VARIANCE

It is possible to find the mean and variance directly from (1.3.5) without first solving for the distribution. Differentiating (1.3.5) with respect to s and (1.5.1) with respect to t, we have when $s = 1$ that

$$\frac{\partial^2 G}{\partial s \partial t}(1;t) = \frac{d\bar{n}}{dt}(t). \tag{1.6.1}$$

Similarly using (1.5.2)

$$\frac{\partial^3 G}{\partial s^2 \partial t}(1;t) = \frac{d}{dt} E(n(n-1))$$

$$= \frac{d}{dt} E(n^2) - \frac{d}{dt} \bar{n}(t). \tag{1.6.2}$$

Differentiating (1.3.5) with respect to s and setting $s=1$ leads directly to the set of equations

$$\frac{d\bar{n}}{dt}(t) - [\lambda(t) - \mu(t)]\bar{n}(t) = \nu(t) - \alpha(t), \tag{1.6.3}$$

$$\frac{d}{dt} E(n^2) - 2[\lambda(t) - \mu(t)] E(n^2) = [2\nu(t) - 2\alpha(t) + \lambda(t) + \mu(t)]\bar{n}(t)$$

$$+ \nu(t) + \alpha(t). \tag{1.6.4}$$

$E(n^2)$ is the second moment of the distribution around zero, so that

$$\sigma^2(t) = E(n^2) - \bar{n}(t)^2. \tag{1.6.5}$$

Equations (1.6.3) and (1.6.4) are thus subject to the initial conditions

$$\bar{n}(0) = n_0 \tag{1.6.6a}$$

$$E(n^2)\big|_{t=0} = \sigma_0^2 + n_0^2. \tag{1.6.6b}$$

It is easy to solve these equations subject to the initial conditions (1.6.6) if the integral of $\lambda(t) - \mu(t)$ can be evaluated, since the equations are linear, first order, and only coupled one way. Let

$$\int_0^t [\lambda(\tau) - \mu(\tau)]\, d\tau = \eta(t). \tag{1.6.7}$$

The integrating factor for (1.6.3) is then $e^{-\eta(t)}$, so that

$$\bar{n}(t) = e^{\eta(t)} \left[\int_0^t [\nu(\tau) - \alpha(\tau)]\, e^{-\eta(\tau)} d\tau + C \right]. \tag{1.6.8}$$

Thus from the initial conditions the constant of integration C is given by

$$C = n_0 e^{-\eta(0)}. \tag{1.6.9}$$

To evaluate (1.6.8), the form of the controls $\nu(t)$ and $\alpha(t)$ must be given. This freedom may allow us to choose them in some optimal fashion. If $\lambda(t)$ and $\mu(t)$ are constant and we choose $\nu(t)$ and $\alpha(t)$ constant, we have

$$\eta(t) = (\lambda - \mu)t,$$

so that

$$\bar{n}(t) = n_0 e^{(\lambda-\mu)t} + \frac{\nu-\alpha}{\mu-\lambda} e^{(\lambda-\mu)t}(e^{-(\lambda-\mu)t} - 1), \qquad (1.6.10)$$

which is identical to (1.5.12). Similarly, for constant parameters, solving (1.6.4) subject to (1.6.6b) and using the relationship (1.6.5), we will obtain an expression for $\sigma^2(t)$ which will be identical to (1.5.13). When only the mean and variance of the distribution are required, it significantly reduces the amount of work involved to use (1.6.1) and (1.6.2) directly, rather than solving (1.3.5) first and then finding the mean and variance afterwards.

1.7 UNRESTRICTED CONSTANT CONTROL

Suppose that we are interested in steering the population distribution mean towards a desired value (N, say) at a given time t_f, all parameters are constant and there are no restrictions on the controls - i.e., we want

$$\bar{n}(t_f) = N,$$

or by (1.5.12)

$$N = n_0 e^{(\lambda-\mu)t_f} + \frac{\alpha-\nu}{\mu-\lambda} (e^{(\lambda-\mu)t_f} - 1). \qquad (1.7.1)$$

Rearranging (1.7.1), we get the form of the control as

$$\alpha - \nu = \frac{(\mu - \lambda)(N - n_0 e^{(\lambda-\mu)t_f})}{e^{(\lambda-\mu)t_f} - 1} \quad (1.7.2)$$

For all $t_f > 0$ the inequality

$$\frac{\mu - \lambda}{e^{(\lambda-\mu)t_f} - 1} < 0 \quad (1.7.3)$$

holds for all λ and μ, since λ and μ are both non-negative.

If no control is applied, i.e. $\alpha - \nu = 0$, then

$$\bar{n}(t_f) = n_0 e^{(\lambda-\mu)t_f},$$

so that

$$N - n_0 e^{(\lambda-\mu)t_f} \quad (1.7.4)$$

is the difference between the means when control and zero control respectively are applied.

(A) If $N > n_0 e^{(\lambda-\mu)t_f}$, then from (1.7.3) and (1.7.2) we have $\alpha - \nu < 0$. Since α and ν are both non-negative this implies that $\alpha < \nu$. Thus positive control exceeds negative control, giving a net positive control level of $\nu - \alpha$. In order to minimize the amount of control used, choose $\alpha = 0$. Intuitively the above analysis says that if the desired mean N exceeds the natural mean $n_0 e^{(\lambda-\mu)t_f}$, we must apply positive control. From (1.7.2) the level of constant positive control required to achieve the mean N must be

$$\nu = \frac{(\lambda-\mu)(N - n_0 e^{(\lambda-\mu)t_f})}{e^{(\lambda-\mu)t_f} - 1} \qquad (1.7.5)$$

(B) If $N < n_0 e^{(\lambda-\mu)t_f}$, we need to apply a net negative control, and similarly to A we find that

$$\alpha = \frac{(\mu-\lambda)(N - n_0 e^{(\lambda-\mu)t_f})}{e^{(\lambda-\mu)t_f} - 1} \qquad (1.7.6)$$

and

$$\nu = 0.$$

1.8 MAINTAINING A CONSTANT MEAN

Once we have reached the desired mean N at time t_f, we may want to maintain the mean at this level for all time in the future. Thus consider t_f to be the initial time, so that $\bar{n}(t_f) = N$ is our initial condition, and find the form of the control so that

$$\bar{n}(t) = N \qquad \text{for all} \quad t > t_f.$$

Adopting equation (1.7.2), we have

$$\alpha - \nu = \frac{(\mu-\lambda)(N - Ne^{(\lambda-\mu)(t-t_f)})}{e^{(\lambda-\mu)(t-t_f)} - 1} \qquad (1.8.1)$$

for all $t > t_f$, i.e.

$$\alpha - \nu = (\lambda-\mu)N. \qquad (1.8.2)$$

If $N > Ne^{(\lambda-\mu)(t-t_E)}$, which implies $\lambda < \mu$, then (1.7.5) becomes

$$\nu = (\mu - \lambda)N. \tag{1.8.3}$$

If $N < Ne^{(\lambda-\mu)(t-t_E)}$, which implies $\lambda > \mu$, then (1.7.6) becomes

$$\alpha = (\lambda - \mu)N. \tag{1.8.4}$$

The results (1.8.3) and (1.8.4) appeal intuitively, since clearly the effect of the control is to cancel the net decrease in the total population by applying positive control at the same level as in (1.8.3), or to cancel the net increase in the total population by applying negative control at the same level, as in (1.8.4).

1.9 VARIANCE FOR CONSTANT PARAMETERS AND CONTROL

The variance at any time t can be easily found using equation (1.5.13) and the results obtained in sections 1.7 and 1.8 for the control levels. Since

$$\frac{e^{(\lambda-\mu)t}}{\lambda-\mu} > 0 \tag{1.9.1}$$

for all λ,μ non-negative, all the terms in (1.5.13) are non-negative except

$$-\alpha\mu\left(\frac{e^{(\lambda-\mu)t} - 1}{\mu - \lambda}\right)^2 ,$$

which is non-positive. So $\sigma^2(t)$ is always greater than or equal to zero when $\alpha = 0$, as it must be to be mathematically consistent. When α is non-zero, however, the greatest value α can have over the interval $[t_0, t_f]$ is such as to make $N = 0$; i.e. from (1.7.6)

$$\alpha_{max} = n_0 e^{(\lambda-\mu)t_f}\left(\frac{\lambda-\mu}{e^{(\lambda-\mu)t_f}-1}\right) . \qquad (1.9.2)$$

Considering the term

$$-\alpha\mu\left(\frac{e^{(\lambda-\mu)t_f}-1}{\lambda-\mu}\right)^2 + \frac{\lambda+\mu}{\lambda-\mu} n_0 e^{(\lambda-\mu)t_f}(e^{(\lambda-\mu)t_f}-1) \qquad (1.9.3)$$

in $\sigma^2(t_f)$ we find that, on substituting (1.9.2) in (1.9.3), we are left with

$$\lambda n_0 e^{(\lambda-\mu)t_f}\left(\frac{e^{(\lambda-\mu)t_f}-1}{\lambda-\mu}\right) ,$$

which by (1.9.1) is non-negative for all λ,μ non-negative, so that $\sigma^2(t)$ is positive for all t, for all possible constant parameters and no doubt for all feasible controls of the system. Also from (1.5.13), $\sigma^2(t)$ is not necessarily monotonic, because of the appearance of the factor $e^{(\lambda-\mu)t}-1$. If the mean is decreasing (i.e., $\lambda < \mu$ and/or $\nu < \alpha$), however, the variance will tend to decrease after a certain time, and vice-versa.

The variance at any time t for $N < n_0 e^{(\lambda-\mu)t_f}$, using (1.5.13), (1.7.6) and (1.8.4), is as follows:

(i) $0 \leq t < t_f$:

$$\sigma^2(t) = \sigma_0^2 e^{2(\lambda-\mu)t} + \frac{\mu}{\lambda-\mu} \frac{N - n_0 e^{(\lambda-\mu)t_f}}{e^{(\lambda-\mu)t_f} - 1} (e^{(\lambda-\mu)t} - 1)^2$$

$$+ \frac{\lambda+\mu}{\lambda-\mu} n_0 e^{(\lambda-\mu)t} (e^{(\lambda-\mu)t} - 1). \qquad (1.9.4)$$

(ii) $t = t_f$:

$$\sigma^2(t_f) = \sigma_0^2 e^{2(\lambda-\mu)t_f} + \frac{\lambda}{\lambda-\mu} n_0 e^{(\lambda-\mu)t_f} (e^{(\lambda-\mu)t_f} - 1)$$

$$+ \frac{\mu}{\lambda-\mu} N(e^{(\lambda-\mu)t_f} - 1). \qquad (1.9.5)$$

(iii) $t > t_f$:

$$\sigma^2(t) = \sigma^2(t_f) e^{2(\lambda-\mu)(t-t_f)}$$

$$+ \frac{\lambda+\mu}{\lambda-\mu} n_0 e^{(\lambda-\mu)(t-t_f)} (e^{(\lambda-\mu)(t-t_f)} - 1)$$

$$- \frac{\mu}{\lambda-\mu} N(e^{(\lambda-\mu)(t-t_f)} - 1)^2. \qquad (1.9.6)$$

Similarly, we can obtain a set of expressions for $\sigma^2(t)$ when $N > n_0 e^{(\lambda-\mu)t_f}$ using (1.5.13), (1.7.5) and (1.8.3).

1.10 DERIVING A COST PERFORMANCE INDEX

Although in many engineering systems the cost of control is easily evaluated, in most ecosystems there are subtle costs that are just as easily overlooked. For example, culling an overpopulated group

of animals in a tourist game park may cause the population to become xenophobic to the extent where tourism in the park is affected. Since each system has its own inherent cost subtleties, it is impossible to formulate a general cost performance index. In a large class of problems, however, we may be interested in the cost of steering a system towards a *desired mean, and the cost of not having attained* this goal at some given time $t = t_f$. We can expect the cost to increase as the following factors increase:

(i) The difference between the desired mean N and the actual mean $\bar{n}(t_f)$ at the final time.

(ii) the size of the variance – i.e., we "pay" for being uncertain of the true size of the population.

(iii) the magnitude of the control variables.

In the particular model described in this paper there are two modes of control, α *and* ν. *Since* ν, the positive control, increases the mean and α, the negative control, decreases the mean, it is physically incongruous for them to operate simultaneously. Their mathematical dependence can be seen in equation (1.7.2), which in fact shows that there is effectively only one control, namely $\alpha - \nu$. An interesting point to this model is, however, that we may have some control in determining our *initial distribution.* In human population studies the mean of the population can be determined to a high degree of accuracy because full population censuses are conducted periodically, but in many populations, especially certain animal populations, it is enormously difficult and costly to obtain an initial distribution with a small variance, to the point where ar

exact knowledge of the size of the population is "infinitely costly". Thus in some systems we can consider the initial variance σ_0^2 (or equivalently the initial standard deviation σ_0, the positive square root of the variance) as a control variable, and for a given initial mean n_0 we can with increasing cost reduce the magnitude of σ_0.

Suppose at time $t = t_f$ we desire the mean to be N, where $N < n_0 e^{(\lambda-\mu)t_f}$. As in section 1.7(B), we take $\nu = 0$. In optimizing a cost performance index, however, we may find that since there is a cost involved in applying α, the optimal α will in general differ from the α given by (1.7.6), and the actual mean obtained will in general not be equal to the desired mean N. Let x_1 be the difference between the actual and desired means at time t_f. Then from (1.5.12) we have

$$x_1(\alpha) = n_0 e^{(\lambda-\mu)t_f} + \frac{e^{(\lambda-\mu)t_f} - 1}{\mu - \lambda} - N. \tag{1.10.1}$$

Let x_2 denote the variance at time t_f; then from (1.5.13) we have

$$\sigma^2(t_f) \triangleq x_2(\alpha,\sigma_0) = \sigma_0^2 e^{2(\lambda-\mu)t_f} - \alpha\mu\left[\frac{e^{(\lambda-\mu)t_f} - 1}{\mu - \lambda}\right]^2 + \frac{\lambda+\mu}{\lambda-\mu} n_0 (e^{(\lambda-\mu)t_f} - 1) e^{(\lambda-\mu)t_f}. \tag{1.10.2}$$

If $J(\alpha,\sigma_0)$ is our cost performance index, then from (i), (ii) and (iii) we must have

$$J(\alpha,\sigma_0) = f(x_1(\alpha), x_2(\alpha,\sigma_0), \alpha, \sigma_0). \tag{1.10.3}$$

In order to find a stationary minimum of J with respect to the control parameters α and σ_0, f must be non-linear in these parameters.

I. Cost of deviating from the mean

When x_1, is zero, we are at the *desired mean*, so that x_1's contribution to the cost at this value must be by (i) also zero. In addition f must be a monotonically increasing function of x_1. If we can estimate the cost $f(x_1^i,0,0,0)$ at two or more points x_1^i, excluding $x_1 = 0$, we can fit the graph

$$f(x_1,0,0,0) = \varepsilon_1 |x_1(\alpha)|^{k_1} \qquad (1.10.4)$$

exactly or by a least squares method, respectively, to obtain the form of the cost contribution of x_1.

Intuitively, a good guess at k_1 is 2, because mean values close to the desired mean should relatively be less heavily penalized than mean values further away from the desired mean.

II. Cost of variance

From general statistical theory the standard deviation $\sigma(t_f)$ is a linear measure of the confidence interval around the mean $\bar{n}(t_f)$. At a given level of confidence, there is a $\beta > 0$, whose value depends on the level of confidence, and the number of samples used to estimate the size of the population, such that the mean will be found at that level of confidence within the interval

$$[\bar{n}(t_f) - \beta\sigma(t_f), \bar{n}(t_f) + \beta\sigma(t_f)] .$$

From (1.10.2) we have $\sigma(t_f) = x_2(\alpha,\sigma_0)^{\frac{1}{2}}$. Since both $x_1(\alpha)$ and $\sigma(t_f)$ measure the deviation of the mean from a desired value, they should have the same form in the cost function, i.e., $\sigma(t_f)$ should also be raised to the power k_1, although the actual weighting constants will in general differ. Thus

$$f(0,x_2,0,0) = a_2 x_2(\alpha,0)^{k_1/2}. \tag{1.10.5}$$

again, a_2 may be *calculated by estimating* the cost corresponding to a number of different x_2 values. Since the maximum deviation of the mean from $\bar{n}(t_f)$ at the desired level of confidence is $\beta\sigma(t_f)$, using the idea that $\beta\sigma(t_f)$ and $x_1(\alpha)$ should push up the cost at the same rate, i.e.,

$$x_1(\alpha) = \beta x_2(\alpha,\sigma_0)^{\frac{1}{2}},$$

we have, *raising both sides to the power* k_1 and multiplying by a_1, that

$$a_1 x_1(\alpha)^{k_1} = a_1 \beta^{k_1} x_2(\alpha,\sigma_0)^{k_1/2}. \tag{1.10.6}$$

This implies from (1.10.5) that

$$a_2 = a_1 \beta^{k_1}. \tag{1.10.7}$$

Either a_2 is known, so that the necessary level of confidence can be calculated through β, or the level of confidence is known and a_2 can be calculated from (1.10.7). The greater the confidence needed in the mean, the larger β must be and (since $k_1 > 0$ and a_1 is constant) the larger a_2 will be. Thus the greater the confidence with which we need to know our mean, the larger the weighting factor a_1 will be, and the more effect the variance will have on the cost function.

III. Cost of control

The cost involved in applying α will very often be linear, although in general it will be of the form

$$f(0,0,\alpha,0) = a_3\alpha^{k_2}, \tag{1.10.8}$$

where $a_3 > 0$ and $k_2 > 0$ can be estimated as in I.

IV. Cost of improving initial variance

$f(0,0,0,\sigma_0)$ must be such that the cost will increase as the magnitude of σ_0 decreases. It must also model the alternatives of finite or infinite cost in finding an n_0 with $\sigma_0 = 0$.

Consider

$$f(0,0,0,\sigma_0) = \frac{a_4}{(\sigma_0^{k_3} + a_5)^{k_4}} \tag{1.10.9}$$

where $a_4 > 0$, $a_5 \geq 0$, $k_3k_4 > 0$. Since $k_3k_4 > 0$, (1.10.9) is a monotonically decreasing function of σ_0. If $a_5 = 0$ it is

"infinitely costly" to make $\sigma_0 = 0$; otherwise, for $\varepsilon_5 > 0$ it will cost $a_4 \varepsilon_5^{-k_4}$.

Suppose that the size of a biological population is estimated from sampling, using a capture-tag-recapture technique, where the statistics are derived from the proportion of animals captured more than once. Let $\bar{n}$ be a best estimate of the population size. The variance associated with $\bar{n}$, denoted $\text{var}(\bar{n})$, *has the property* that the smaller $\text{var}(\bar{n})$ is the more we known about $\bar{n}$. Let us define information I on $\bar{n}$ as

$$I = \frac{1}{\text{var}(\bar{n})} \quad . \tag{1.10.10}$$

Clearly, *zero information implies infinite* variance and vice versa.

Consider a multiple recapture census comprising a sequence of samples $S_1, S_2, \ldots, S_r$, where the members of $S_1, \ldots, S_{r-1}$ are all tagged before being returned to the population, while the members $S_2, \ldots, S_r$ are classified according to when, if at all, they have been captured before. It is assumed that the method of capture does not kill or affect the future behaviour of a population member. Let p be the probability of an individual being caught in a sample, and let $q = 1 - p$. To *improve our estimate of n* by reducing $\text{var}(\bar{n})$, we can either increase the number of samples taken or decrease the value of q. Let e be the amount of effort expended in obtaining pn members in each sample; then if $e = 0$, pn is zero, so $q = 1$. If it is "infinitely difficult" to count the whole population exactly, then $pn = n$ (which implies $q = 0$), only when e is infinitely large. Thus q is related to e by

$$q = \exp(-\alpha e) \qquad \alpha > 0. \tag{1.10.11}$$

Darroch (1958) proves that if e is increased to ke, $k > 1$, the information I is enlarged to more than $k^2 I$, while if e is held constant and the number of samples is increased from r to r+1, the information is increased by more than $(r+1)/(r-1)$, i.e.

$$I_{2e} > 4I_e,$$

and

$$I_{2r} > \frac{(2r)(2r-1)}{(r)(r-1)} I_r$$

$$\approx 4I_r,$$

where I_e denotes the amount of information obtained with affort e for fixed r, and I_r is the amount of information obtained with r samples for fixed e.

If $\alpha e r \ll 1$, then the first order approximation

$$I_{2e} \approx 4I_e \tag{1.10.12}$$

$$I_{2r} \approx 4I_r \tag{1.10.13}$$

will be good. This condition will invariably hold since pa/n will usually be of the order of a fraction of a percent. This implies that q will be very close to 1, and from (1.10.11) ae will be very small indeed.

Suppose increase in cost is directly proportional to increase in effort or increase in the number of samples. Then doubling the effort or number of samples will increase the information by a factor of 4, or by (1.10.10) will decrease the variance by a factor of 4. This is equivalent to decreasing the standard deviation by a factor of 2 - i.e., the increase in cost is proportional to $1/\sigma_0$, and thus (1.10.9) becomes

$$f(0,0,0,\sigma_0) = \frac{a_4}{\sigma_0} . \tag{1.10.14}$$

Clearly (1.10.14) is only applicable to populations where it is "infinitely difficult" to determine the size of the population exactly. $a_4 > 0$ can be calculated as before. In general if the cost can be determined empirically for a number of points $(\sigma_0^i, f(0,0,0,\sigma_0^i))$, then a_4, a_5, k_3 and k_4 in (1.10.8) can be determined (as previously discussed) by fitting a curve through these points.

The reader interested in capture-recapture statistics should consult Darroch (1959) and Jolly (1965).

Finally, using (1.10.3), (1.10.4), (1.10.5), (1.10.7), (1.10.8) and (1.10.9), (1.10.3) can be written as

$$J(\alpha,\sigma_0) = a_1 |x_1(\alpha)|^{k_1} + a_1 \beta^{k_1} x_2(\alpha,\sigma_0)^{k_1/2} + a_3 \alpha^{k_2} + \frac{a_4}{(\sigma_0^{k_3} + a_5)^{k_4}} . \tag{1.10.15}$$

A similar analysis can be done for $N > n_0 e^{(\lambda-\mu)t_f}$, in which case $\alpha = 0$, $\nu > 0$ and $J = J(\nu,\sigma_0)$.

1.11 MINIMIZING THE COST PERFORMANCE INDEX

Suppose that $k_1 = 2$, $k_2 = 1$, and we are analysing a large animal population whose size is "infinitely costly" to determine exactly. Then (1.10.15) becomes

$$J(\alpha,\sigma_0) = a_1 x_1(\alpha)^2 + a_1\beta^2 x_2(\alpha,\sigma_0) + a_3\alpha + \frac{a_4}{\sigma_0}. \qquad (1.11.1)$$

We can minimize (1.11.1) by finding the stationary points of $J(\alpha,\sigma_0)$ with respect to the parameters α and σ_0 and checking to see if the second derivative matrix of J is positive semi-definite.

Recalling (1.10.1) and (1.10.2), and letting

$$\frac{e^{(\lambda-\mu)t_f} - 1}{\mu - \lambda} = r(t_f), \qquad (1.11.2)$$

we have

$$\frac{\partial x_1(\alpha)}{\partial\alpha} = r(t_f),$$

$$\frac{\partial^2 x_1(\alpha)}{\partial\alpha^2} = 0,$$

$$\frac{\partial x_2(\alpha,\sigma_0)}{\partial\alpha} = -\mu r(t_f)^2,$$

$$\frac{\partial^2 x_2(\alpha,\sigma_0)}{\partial\alpha^2} = 0,$$

$$\frac{\partial x_2(\alpha,\sigma_0)}{\partial\sigma_0} = 2\sigma_0 e^{2(\lambda-\mu)t_f},$$

$$\frac{\partial^2 x_2(\alpha,\sigma_0)}{\partial\sigma_0^2} = 2e^{2(\lambda-\mu)t_f}$$

$$\frac{\partial^2 x_2(\alpha,\sigma_0)}{\partial\alpha\partial\sigma_0} = 0.$$

Thus

$$\frac{\partial J}{\partial\alpha} = 2a_1 x_1(\alpha)\frac{\partial x_1}{\partial\alpha} + a_1\beta^2\frac{\partial x_2}{\partial\alpha} + a_3$$

$$= 2a_1[n_0 e^{(\lambda-\mu)t_f} - N + \alpha r(t_f)]\, r(t_f)$$

$$- a_1\beta^2\mu r(t_f) + a_3.$$

Solving for α in $\partial J/\partial\alpha = 0$ and substituting for $r(t_f)$ gives

$$\alpha = \frac{(\mu-\lambda)(N - n_0 e^{(\lambda-\mu)t_f}}{(e^{(\lambda-\mu)t_f} - 1)} + \frac{\mu\beta^2}{2}$$

$$- \frac{a_3}{2a_1}\left(\frac{\mu-\lambda}{e^{(\lambda-\mu)t_f} - 1}\right), \qquad (1.11.3)$$

$$\frac{\partial J}{\partial\sigma_0} = a_1\beta^2\frac{\partial x_2}{\partial\sigma_0} - \frac{a_4}{\sigma_0^2}$$

$$= 2a_1\beta^2\sigma_0 e^{2(\lambda-\mu)t_f} - \frac{a_4}{\sigma_0^2}.$$

Solving for σ_0 in $\frac{\partial J}{\partial \sigma_0} = 0$ gives

$$\sigma_0 = \left(\frac{a_4}{2a_1\beta^2 e^{2(\lambda-\mu)t_f}}\right)^{1/3} \qquad (1.11.4)$$

Since

$$\frac{\partial^2 J}{\partial \alpha^2} = 0,$$

$$\frac{\partial^2 J}{\partial \sigma_0^2} = 2a_1\beta^2 e^{2(\lambda-\mu)t_f} + \frac{2a_4}{\sigma_0^3} > 0$$

(because $a_1, \beta, a_4, \sigma_0$ are all positive) and

$$\frac{\partial^2 J}{\partial \alpha \partial \sigma_0} = 0,$$

it follows that the matrix

$$\begin{bmatrix} \frac{\partial^2 J}{\partial \alpha^2} & \frac{\partial^2 J}{\partial \sigma_0 \partial \alpha} \\ \frac{\partial^2 J}{\partial \alpha \partial \sigma_0} & \frac{\partial^2 J}{\partial \sigma_0^2} \end{bmatrix}$$

is positive semi-definite, so that the values of α and σ_0 given by (1.11.3) and (1.11.4) minimize $J(\alpha,\sigma_0)$. If $\beta = 0$ and $a_3 = 0$, in which case we are not interested in cost of variance or control, (1.11.3) is identical to (1.7.6). Thus in (1.11.3) we have found the optimal level at which to apply constant negative control on the system, and in (1.11.4) we have found how accurately we must deter-

mine our initial mean by indicating the size of the initial variance, so as to minimize the cost performance index (1.11.1) at the final time t_f.

1.12 PIECEWISE CONSTANT PARAMETERS AND THE MINIMIZATION PROBLEM

So far we have only considered systems with constant parameters λ, μ, ν and α over the interval $[t_0 = 0, t_f]$. The extension of the analysis to piecewise constant parameters over this interval is worth while, since this may be a good approximation to complicated time varying parameters, especially birth and death parameters that vary seasonally.

Suppose that we divide the interval $[t_0, t_f]$ into m subintervals. Denote the i-th subinterval $[t_{i-1}, t_i]$ by Δt_i, and let the parameters over this interval have the constant values λ_i, μ_i, α_i and ν_i. If the mean at time t_{i-1} is n_{i-1} and the variance σ_{i-1}, then adapting (1.5.12) and (1.5.13) to the new notation, we have

$$n_i = n_{i-1} e^{(\lambda_i - \nu_i)\Delta t_i} + \frac{\alpha_i - \nu_i}{\mu_i - \lambda_i} \left(e^{(\lambda_i - \mu_i)\Delta t_i} - 1\right), \qquad (1.12.1)$$

$$\sigma_i = \sigma_{i-1}^2 e^{2(\lambda_i - \mu_i)\Delta t_i} + \frac{\lambda_i + \mu_i}{\lambda_i - \mu_i} e^{(\lambda_i - \mu_i)\Delta t_i} \left(e^{(\lambda_i - \nu_i)\Delta t_i} - 1\right)$$

$$+ \frac{\nu_i \lambda_i - \alpha_i \mu_i}{(\mu_i - \lambda_i)^2} \left(e^{(\lambda_i - \mu_i)\Delta t_i} - 1\right). \qquad (1.12.2)$$

Thus (1.10.1) and (1.10.2) become

$$x_1(n_{i-1},\alpha) = n_i - N \tag{1.12.3}$$

$$x_2(n_{i-1},\sigma_{i-1},\alpha_i) = \sigma_i^2 \ . \tag{1.12.4}$$

The cost of running the system over the i-th interval will be, using (1.10.15),

$$J(n_{i-1},\sigma_{i-1},\alpha_i) = a_1|x_1(n_{i-1},\alpha_i)|^{k_1} + a_1\beta^{k_1}x_2(n_{i-1},\sigma_{i-1},\alpha_i)^{k_1/2}$$

$$+ a_3\alpha_i^{k_1} + \frac{a_4}{(\sigma_{i-1}^{k_3} + a_5)^{k_4}} \ . \tag{1.12.5}$$

Hence the total cost of running the system over $[t_0,t_f]$ will be given by

$$F(n_0,\sigma_0,\alpha_1,\alpha_2,\ldots,\alpha_m) \sum_{i=1}^{m} J(n_{i-1},\sigma_{i-1},\alpha_i). \tag{1.12.6}$$

As before, n_0 is fixed, but we can choose the m+1 parameters $\alpha_0,\alpha_1,\ldots,\alpha_m$ optimally so as to minimize (1.12.6). It is worth noting that; choosing each α_i to minimize $J(n_{i-1},\sigma_{i-1},\alpha_i)$ and σ_0 to minimize $J(n_0,\sigma_0,\alpha_i)$ will not in general minimize F, since σ_i and σ_{i-1} are dependent of $\sigma_0,\alpha_1,\ldots,\alpha_{i-1}$.

There are *two important computational* procedures to find a

$$u \triangleq (\sigma_0,\alpha_1,\alpha_2,\ldots,\alpha_m) \tag{1.12.7}$$

that will minimize (1.12.6).

I. Non-linear programming

There are a number of algorithms, including the various gradient methods (Himmelblau 1972 and Hadley 1964), that are designed to calculate

$$u^* = (\alpha_1^*, \alpha_j^*, \ldots, \alpha_m^*)$$

such that

$$F(n_0, u^*) \leq F(n_0, u)$$

for all admissible u defined by (1.12.7). A more detailed analysis of this type of approach would only be worth tackling if a specific problem were being considered.

II. Dynamic programming

Define

$$F_k(n_{k-1}, \sigma_{k-1}) \triangleq \min_{\alpha_k, \ldots, \alpha_m} \sum_{i=k}^{m} J(n_{i-1}, \sigma_{i-1}, \alpha_i).$$

Then

$$F_k(n_{k-1}, \sigma_{k-1}) = \min_{\alpha_k} J(n_{k-1}, \sigma_{k-1}, \alpha_k)$$

$$+ \min_{\alpha_{k+1}, \ldots, \alpha_m} \sum_{i=k+1}^{m} J(n_{i-1}, \sigma_{i-1}, \alpha_i),$$

since $n_{i-1}, \sigma_{i-1}, \alpha_k$ are all independent of $\alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_m$. Thus

$$F_k(n_{k-1}, \sigma_{k-1}) = \min_{\alpha_k} [J(n_{k-1}, \sigma_{k-1}, \alpha_k) + F_{k+1}(n_k, \sigma_k)],$$
$$k = m-1, m-2, \ldots, 1, \quad (1.12.8)$$

forms a recurrence relation to solve for

$$F_0(n_0) = \min_{\sigma_0} F_1(n_0, \sigma_0)$$
$$+ \min_{u} F(n_0, u), \quad (1.12.9)$$

as required, with the given initial condition

$$F_m(n_{m-1}, \sigma_{m-1}) = \min_{\sigma_m} J(n_{m-1}, \sigma_{m-1}, \alpha_m). \quad (1.12.10)$$

The problem of solving for (1.12.9) is complicated by the fact that (1.12.1), (1.12.2) and (1.12.8) are coupled and that (1.12.1) and (1.12.2) are solved forwards in time while (1.12.8) is solved backwards in time.

Some optimization problems lend themselves readily to a dynamic programming approach, whenever it is possible to obtain a general form for F_k in terms of a reproducing function, e.g., the optimization of a general stochastic linear system with a quadratic performance criterion (Meditch 1969). Is it not possible to obtain a

reproducing form for F_k in the above problem, as F_m contains terms in σ^2_{m-1} and $1/\sigma_{m-1}$; F_{m-1} contains terms in σ^4_{m-2}, $1/\sigma^2_{m-2}$, $\sigma^2_{m-2}\alpha_{m-2}$ and $\sigma^2_{m-2}\alpha_{m-1}$; and F_{m-2} contains even more non-linear and cross product terms. Hence the application of dynamic programming to the cost function in its general form as given in (1.12.5), is not pursued further at this point.

For a particular problem, however, a suitable algorithm may possibly be found in Hadley (1964), Bellman and Dreyfus (1962) or Jacobson and Mayne (1970).

CHAPTER 2

FORMULATION OF A GENERAL MULTIVARIATE BIRTH AND DEATH PROCESS

2.1 INTRODUCTION

The previous chapter deals with the simple linear univariate birth and death process in its most general form. One extension of this work is to consider more complex univariate processes allowing non-linear rates and multiple events.

We will, however, go straight on to a second possible extension of the previous work i.e., multivariate birth and death process models. Both linear and non-linear single event multivariate processes will be dealt with in this chapter, while the extension to multiple events is dealt with in the next chapter.

General multivariate populations have proved difficult to analyse, partly owing to the notational complexities involved. As was pointed out in the introduction, not many results have been obtained for bivariate birth and death processes. For the more general multivariate process joint multivariate probability distributions have been derived only for a few special cases. Notable among these results are the distributions obtained for a class of multivariate linear birth and death processes under the initial conditions of zero members in each population (Boswell and Patil 1972), and migration processes between colonies of the same species where the birth, death and migration parameters are linearly dependent on population size (Renshaw 1972).

Usually it is sufficient for most modelling purposes if the means, variances and covariances of the individual populations are the only parameters of the joint multivariate distribution known. These results have been derived for the bivariate two-sex problem (*Prabhu 1965, chapter 16*) and bivariate Lotka-Volterra predator-prey system (Goel et al. 1971).

A general procedure has been derived to generate the variance, covariance and higher-order terms for any *Galton-Watson multitype* process (Pollard 1973). Pollard's stochastic version of the Leslie model of population age structure is an important application of this theory. No methodology exists, however, for generating the desired statistics for general Markov jump processes. In addition, most analyses of systems in this area have been primarily concerned with the properties of the stochastic matrix of a particular Markov process. Inglehart (1964), following Reuter's method of analysing the transition probability matrix of a Markov process, gives conditions of uniqueness recurrence and properties of absorption states of a class of multivariate Markov jump processes known as competition processes.

In this chapter we are primarily concerned with deriving ordinary differential equations describing the first and second order moments of the joint multivariate distribution for single event Markov jump processes. Although the formulation is general, two specific processes are dealt with in depth. The first is the multivariate version of the linear control process discussed in the previous chapter and the second is a non-linear process that leads to a stochastic form of the deterministic Lotka-Volterra population model.

For a general survey of the application of stochastic processes especially of the Markov type, the reader is referred to two volumes by Iosifescu and Tautu (1973a, 1973b).

2.2 *DEFINITIONS*

In view of the lack of a standardized notation and the notational complexities involved when a multivariate system is analysed, the following definitions are introduced, which facilitate formulation of the problem.

2.2.1 Let n_i denote the i-th element of a vector $\underline{n} \in R^m$.

Then

$$\Omega \triangleq \{\underline{n}^T = (n_1, n_2, \ldots, n_m) \mid n_i \in Z^+ \qquad i=1,\ldots,m\}$$

where

$Z^+ \triangleq$ {non negative integers}.

2.2.2 For notational purposes, we define the following vectors in Ω

$$\underline{0}^T \triangleq (0,0,\ldots,0)$$

$$\underline{1}^T \triangleq (1,1,\ldots,1)$$

$$\underline{e}_i^T \triangleq (0,\ldots,1,\ldots,0) \quad (1 \text{ in the i-th position}).$$

2.2.3 We define the ℓ_1^m norm on the elements $\underline{n}, \underline{n}^* \in \Omega$, i.e.

(i) $$\|\underline{n}\| = \sum_{i=1}^{m} n_i$$

(ii) $\|\underline{n}-\underline{n}^*\| = \sum_{i=1}^{m} |n_i - n_i^*|$.

2.2.4 An m-species population process is an *m-dimensional stochastic process* $X(t)$ *defined on* Ω *for* $t\in[0,\infty)$ by the probability *density function*

$$P(\underline{n};t) = \Pr\{X(t) = \underline{n}\} \quad \text{for all } \underline{n}\in\Omega$$

$$= 0 \quad \text{for all } \underline{n}\notin\Omega .$$

2.2.5 *We define the transition probability distribution of the process in terms of any* $\underline{n}, \underline{n}^*\in\Omega$ and $t,s\in[0,\infty)$ as:

$$P(\underline{n},\underline{n}^*;s,t) = \Pr\{X(t+s) = \underline{n}^* | X(s) = \underline{n}\}.$$

Also define

$$P(\underline{n},\underline{n}^*;s,o) = \delta(\underline{n},\underline{n}^*) = \begin{cases} 0 \text{ if } \underline{n} \neq \underline{n}^* \\ 1 \text{ if } \underline{n} = \underline{n}^* \end{cases}$$

so that equation (2.3.1) *considered in the next section holds when* $\Delta t = 0$.

Further since $P(\underline{n},\cdot;s,t)$ is a probability distribution we must have

$$P(\underline{n},\underline{n}^*;s,t) \geq 0 \quad \text{and} \quad \sum_{\underline{n}^*\in\Omega} P(\underline{n},\underline{n}^*;s,t) = 1.$$

In this chapter we *shall* be primarily concerned with formulating, in *detail*, models of the following process.

2.2.6 An m-dimensional multivariate birth-and-death process is a population process $X(t)$ defined on Ω with a transition probability distribution satisfying the following conditions:

(i) $P(\underline{n},\underline{n}^*;t,\Delta t) = o(\Delta t)$ for all $\underline{n},\underline{n}^* \in \Omega$ such that $\|\underline{n}-\underline{n}^*\| \geq 2$

(ii) $P(\underline{n},\underline{n}^*;t,\Delta t) = 0$ for all $\underline{n},\underline{n}^* \in \Omega$ such that $n_i = 0$ and $n_i^* > 0$.

Condition (i) allows only single events to occur at any instant in time. This assumption is standard for most birth and death models, a notable exception being models which provide for the occurrence of multiple births (Doubleday 1973).

Condition (ii) states that no spontaneous generation (i.e. births) can take place in a species that has zero members. This assumption provides one of the essential differences between birth and death and queueing processes.

2.2.7 We define vectors $\underline{u}(t) \in R^p$, $\underline{y}(t) \in R^q$ for all $t \in [0,\infty)$. Using systems terminology we sometimes refer to $X(t)$ as the state vector of the system (process) and $\underline{u}(t)$, $\underline{y}(t)$ as the external or control vectors of the system.

2.3 GENERAL FORMULATION

In this section we shall derive the forward Kolmogorov differential equation for $P(\underline{n};t)$ to be solved under a given initial condition, $P(\underline{n};0)$ say, for the process defined in 2.2.6. This process will be shown to be a denumerable Markov jump process.

Since condition (i) in 2.2.6 restricts state transitions to single events, at any instant in time, with probability one, we can write down the following equation

$$P(\underline{n};t+\Delta t) = P(\underline{n};t)P(\underline{n},\underline{n};t,\Delta t) + \sum_{i=1}^{m} P(\underline{n}+e_i;t)P(\underline{n}+e_i,\underline{n};t,\Delta t) + P(\underline{n}-e_i;t)P(\underline{n}-e_i,\underline{n};t,\Delta t) + o(\Delta t) \tag{2.3.1}$$

which holds for all $\underline{n}\in\Omega$.

If the limits as defined in the following equation exist we have, from equation (2.3.1), the forward Kolomogorov differential equations of the process:

$$P'(\underline{n};t) = \lim_{\Delta t\to o}\{[P(\underline{n},\underline{n};t,\Delta t) - 1]/\Delta t\}P(\underline{n};t) + \lim_{\Delta t\to o}\{\sum_{i=1}^{m}[P(\underline{n}+e_i,\underline{n};t,\Delta t)/\Delta t]P(\underline{n}+e_i;t) + \sum_{i=1}^{m}[P(\underline{n}-e_i,\underline{n};t,\Delta t)/\Delta t]P(\underline{n}-e_i;t)\} \tag{2.3.2}$$

for all $\underline{n}\in\Omega$, which can be solved subject to a given initial condition $P(\underline{n};0)$.

The existence and uniqueness of a solution $P(\underline{n};t)$ $t\in[0,\infty)$ will depend on the functional form of the limits as defined in the above equation. In addition, a limiting distribution

$$P(\underline{n}) = \lim_{t\to\infty} P(\underline{n};t) \tag{2.3.3}$$

will exist for a time homogeneous system under certain conditions.

Since Ω is denumerable, the process defined by 2.2.6 turns out to be a denumerable Markov jump process under the following conditions (Prabhu 1965, p. 125):

We define

$$C(\underline{n};t) \triangleq \lim_{\Delta t \to 0} [1 - P(\underline{n},\underline{n};t,\Delta t)]/\Delta t \qquad (2.3.4)$$

and

$$C(\underline{n}^*;t)\Pi(\underline{n}^*,\underline{n};t) \triangleq \lim_{\Delta t \to 0} P(\underline{n}^*,\underline{n};t,\Delta t)/\Delta t. \qquad (2.3.5)$$

Then $C(\underline{n};t)$, $C(\underline{n}^*;t)$, $\Pi(\underline{n}^*,\underline{n};t)$ must be continuous in t for all pairs $(\underline{n}^*,\underline{n})$, $\underline{n}^*$, $\underline{n} \in \Omega$ with

$$C(\underline{n};t) \geq 0, \qquad \Pi(\underline{n},\underline{n};t) = 0 \qquad \text{and}$$

$$\sum_{\underline{n} \in \Omega} \Pi(\underline{n}^*,\underline{n};t) = 1 \quad \text{for all} \quad t \in [0,\infty).$$

$C(\underline{n};t)$ is referred to as the "jump rate" from state $\underline{n}$ and $\Pi(\underline{n}^*,\underline{n};t)$ is the probability of a jump from $\underline{n}^*$ to state $\underline{n}$. The reader should refer to Prabhu (1965, chapter 4) and Iosifescu and Tautu (1973a, p. 201) for an account of the theory on which the Kolmogorov forward and backward equations of a Markov jump process are based.

Suppose that the transitional probabilities of the process are functions of the state and external vectors of the process, as well as t explicitly. If we define these probabilities as:

$$P(\underline{n}-\underline{e}_i,\underline{n};t,\Delta t) \triangleq \lambda_i(\underline{n}-\underline{e}_i,\underline{u}(t),\underline{v}(t),t)\Delta t + o(\Delta t)$$
$$i=1,\ldots,m$$
$$P(\underline{n}+\underline{e}_i,\underline{n};t,\Delta t) \triangleq \mu_i(\underline{n}+\underline{e}_i,\underline{u}(t),\underline{v}(t),t)\Delta t + o(\Delta t) \qquad (2.3.6)$$
$$i=1,\ldots,m .$$

Since $P(\underline{n},\underline{n}^*;t,\ t)$ is a probability distribution satisfying the conditions of an m-dimensional multivariate birth-and-death process (definition 2.2.6), we have, using (2.3.6) above

$$P(\underline{n},\underline{n};t,\Delta t) = 1 - \sum_{i=}^{m} \{\lambda_i(\underline{n},\underline{u}(t),\underline{v}(t),t) + \mu_i(\underline{n},\underline{u}(t),\underline{v}(t),t)\} + o(\Delta t). \qquad (2.3.7)$$

In order that condition (ii) in definition 2.2.6 be satisfied we must have

$$\lambda_i(\underline{0},\cdot,\cdot) \equiv 0.$$

Further, since the probabilities are only defined on Ω we must have

$$\mu_i(\underline{n}+\underline{e}_i,\cdot,\cdot,\cdot) \equiv 0 \qquad \text{if } n_i - 1 < 0.$$

Substituting (2.3.6) and (2.3.7) in (2.3.2) we see that the limits in equation (2.3.2) are defined by (2.3.6) and (2.3.7). So finally we obtain the equation

$$P'(\underline{n};t) = \sum_{i=1}^{m} \{\lambda_i(\underline{n}-\underline{e}_i,\underline{u}(t),\underline{v}(t),t)P(\underline{n}-\underline{e}_i;t) - [\lambda_i(\underline{n},\underline{u}(t),\underline{v}(t),t) +$$
$$\mu_i(\underline{n},\underline{u}(t),\underline{v}(t),t)]P(\underline{n};t) + \mu_i(\underline{n}+\underline{e}_i,\underline{u}(t),\underline{v}(t),t)P(\underline{n}+\underline{e}_i;t)\} \qquad (2.3.8)$$

which describes the trajectory of the probability density function $P(\underline{n};t)$ for the m-dimensional multivariate birth-and-death process $X(t)$ defined in 2.2.6.

From (2.3.4) we have

$$C(\underline{n};t) = \sum_{i=1}^{m} [\lambda_i(\underline{n},\underline{u}(t),\underline{v}(t),t) + \mu_i(\underline{n},\underline{u}(t),\underline{v}(t),t)] \qquad (2.3.9)$$

Clearly $C(\underline{n};t) \geq 0$ since it is the sum of non-negative functions (a consequence of (2.3.6)). Using (2.3.5), (2.3.6) and (2.3.9) we see that for any $\underline{n} \in \Omega$ $\Pi(\underline{n},\underline{n}^*;t)$ will have two forms:

$$\begin{aligned} \Pi(\underline{n},\underline{n}+\underline{e}_i;t) &= \lambda_i(\underline{n},\underline{u}(t),\underline{v}(t),t)/C(\underline{n};t) \\ &\qquad i=1,\ldots,m \\ \Pi(\underline{n},\underline{n}-\underline{e}_i;t) &= \mu_i(\underline{n},\underline{u}(t),\underline{v}(t),t)/C(\underline{n};t) \\ &\qquad i=1,\ldots,m \ . \end{aligned} \qquad (2.3.10)$$

Further, $\Pi(\underline{n},\underline{n}^*;\cdot) \equiv 0$ for all other $\underline{n}^* \in \Omega$ including $\underline{n}^* = \underline{n}$ so that clearly from (2.3.9) and (2.3.10)

$$\sum_{\underline{n}^* \in \Omega} \Pi(\underline{n},\underline{n}^*;t) = 1$$

for all $t \in [0,\infty)$. Therefore, under the stated conditions, if expressions (2.3.9) and (2.3.10) are *continuous* in t for $\underline{n} \in \Omega$, the process defined in 2.2.6 is a denumerable Markov jump process and equation (2.3.8) is the well-defined forward Kolmogorov equation of this process.

To sum up: If we choose $\lambda_i(\cdot,\cdot,\cdot,\cdot)$ and $\mu_i(\cdot,\cdot,\cdot,\cdot)$ as continuous non-negative functions of time for all $\underline{n} \in \Omega$, the continuity

conditions are satisfied, the various probability distributions are well-defined and equation (2.3.8) is the forward Kolmogorov differential equation of the stochastic process $X(t)$ defined in 2.2.6 which turns out to be a denumerable Markov jump process.

2.4 BUILDING SPECIFIC MODELS

In this section we shall choose specific forms for the functions $\mu_i(\cdot,\cdot,\cdot,\cdot)$ and $\lambda_i(\cdot,\cdot,\cdot,\cdot)$ introduced in the previous section and state the physical interpretation that can be given to these forms. The systems considered turn out to be the stochastic analogues of the linear and Lotka-Volterra deterministic systems that have been used widely in population dynamics (Goel et al. 1971, Usher 1972 and Smith 1973). The building of controls into the system will also be considered, allowing the models to be used in optimal control studies (Dreyfus 1972).

A. Linear control system

For $i=1,\ldots,m$ we define

$$\lambda_i(\underline{n},\underline{u}(t),t) \triangleq \sum_{j=1}^{m} \lambda_{ij}(t)n_j + \sum_{r=1}^{p} \alpha_{ir}(t)u_r(t) \quad \text{for } n_i > 0 \qquad (2.4.1)$$

$$= 0 \quad \text{for } n_i = 0$$

$$\mu_i(\underline{n},\underline{\mu}(t),\underline{\nu}(t),t) \triangleq \sum_{j=1}^{m} \mu_{ij}(t)n_j + \sum_{r=1}^{p} \beta_{ir}(t)u_r(t)$$

$$\text{for } n_i > 0 \qquad (2.4.2)$$

$$= 0 \qquad \text{for } n_i = 0$$

where all time-dependent functions are continuous in t, $t\in[0,\infty)$.

Using the above definitions, for all $\underline{n}\in\Omega$ equation (2.3.8) becomes

$$P'(\underline{n};t) = \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m} \{\lambda_{ij}(t)n_jP(\underline{n}-\underline{e}_i;t) - [\lambda_{ij}(t)+\mu_{ij}(t)]n_jP(\underline{n};t)$$

$$+\mu_{ij}(t)n_jP(\underline{n}+\underline{e}_i;t)\} + \sum_{i=1}^{m} \{\lambda_{ii}(t)(n_i-1)P(\underline{n}-\underline{e}_i;t)$$

$$-[\lambda_{ii}(t)+\mu_{ii}(t)]n_iP(\underline{n};t) + \mu_{ii}(t)(n_i+1)P(\underline{n}+\underline{e}_i;t)\}$$

$$+ \sum_{i=1}^{m}\sum_{r=1}^{p} \{\alpha_{ir}(t)u_r(t)P(\underline{n}-\underline{e}_i;t) - [\alpha_{ir}(t)+\beta_{ir}(t)]u_r(t)P(\underline{n};t)$$

$$+ \beta_{ir}(t)u_r(t)P(\underline{n}+\underline{e}_i;t)\}. \qquad (2.4.3)$$

The physical interpretation of (2.4.1) and (2.4.2) is that each i-th species is influenced to increase and decrease at

(i) *a linear time-dependent rate with respect to its own size and the size of every other species;*

(ii) *a linear time-dependent rate with respect to the size of the elements of a non-negative external vector* $\underline{\mu}$.

B. Lotka-Volterra competitive systems

For $i=1,\ldots,m$ and $\underline{n}\in\Omega$ we define

$$\lambda_i(\underline{n},\underline{\mu}(t),\underline{\nu}(t),t) \triangleq \lambda_i(\underline{\nu}(t),t)n_i + \sum_{j=1}^{m}\lambda_{ij}(t)n_in_j \qquad (2.4.4)$$

$$\mu_i(\underline{n},\underline{\mu}(t),\underline{\nu}(t),t) \triangleq \mu_i(\underline{\nu}(t),t)n_i + \sum_{j=1}^{m}\mu_{ij}(t)n_in_j \qquad (2.4.5)$$

where all time-dependent functions are continuous in t, $t\in[0,\infty)$.

It is clear from definitions (2.4.4) and (2.4.5) that $\lambda_i(\cdot,\cdot,\cdot,\cdot)$ and $\mu_i(\cdot,\cdot,\cdot,\cdot)$ are zero when $n_i = 0$.

Using the above definitions, for all $\underline{n}\in\Omega$ equation (2.3.8) becomes

$$\begin{aligned}P'(\underline{n};t) = \sum_{i=1}^{m}\{&[\lambda_i(\underline{\nu}(t),t)(n_i-1)+\lambda_{ii}(t)(n_i-1)^2]P(\underline{n}-e_i;t)\\ &-[(\lambda_i(\underline{\nu}(t),t)+\mu_i(\underline{\nu}(t),t))n_i+(\lambda_{ii}(t)+\mu_{ii}(t))n_i^2]P(\underline{n};t)\\ &+[\mu_i(\underline{\nu}(t),t)(n_i+1)+\mu_{ii}(t)(n_i+1)^2]P(\underline{n}+e_i;t)\}\\ +\sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m}&\{\lambda_{ij}(t)(n_i-1)n_jP(\underline{n}-e_i;t)\\ &-[\lambda_{ij}(t)+\mu_{ij}(t)]n_in_jP(\underline{n};t)+\mu_{ij}(t)(n_i+1)n_jP(\underline{n}+e_i;t)\}.\end{aligned} \qquad (2.4.6)$$

The physical interpretation of (2.4.4) and (2.4.5) is that each i-th species is influenced to increase and decrease at

(i) a linear time-dependent rate with respect to its own size, where the rate may depend on an external vector $\underline{y}(t)$;

(ii) a quadratic time-dependent rate with respect to its own size; the inclusion of this term is equivalent to building a Pearl-Verhulst type limit term into the model (Goel et al. 1971, p. 237);

(iii) a time-dependent rate directly proportional to size of the product of the i-th and j-th species, for $j=1,\ldots,m$ and $j\neq i$; this rate is usually termed the interaction rate between the i-th and j-th populations.

Note:

Linear controls can be added to model B as was done in A. More general models can be built using combinations of A and B as well as by introducing various non-linearities, e.g. bilinear terms in $\underline{n}$ and $\underline{u}(t)$ (Mohler, 1973).

2.5 METHOD OF SOLUTION

We are now faced with the problem of solving equations of the type given by (2.4.3) and (2.4.6) for the probability density function $P(\underline{n};t)$, which will then define our process $X(t)$ (see definition 2.2.4), subject to a given initial condition $P(\underline{n};0)$. The method outlined in this section is versatile in that it can be applied to general population processes of the type discussed in section 2.4, although only the statistics of the distribution $P(\underline{n};t)$ can be derived, e.g. means, variance and covariance terms and not $P(\underline{n};t)$ itself.

We define a generating function

$$G(\underline{s};t) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_m=0}^{\infty} P(\underline{n};t)\, s_1^{n_1} s_2^{n_2},\ldots,s_m^{n_m}$$

$$\triangleq \sum_{\underline{n}\in\Omega} P(\underline{n};t)\underline{s}^{\underline{n}}. \qquad (2.5.1.1)$$

The summation in (2.5.1.1) is well-defined since Ω is denumerable, and uniformly convergent for $|s_i| \leq 1 \quad i=1,\ldots,m$ as $\sum_{\underline{n}\in\Omega} P(\underline{n};t) = 1$. For a well-behaved multivariate distribution $P(\underline{n};t)$, i.e. finite second and third order moments around the mean, we can derive the following relationships:

$$\frac{\partial G(\underline{s};t)}{\partial s_i} = \sum_{\underline{n}\in\Omega} n_i\, P(\underline{n};t)\underline{s}^{\underline{n}-\underline{e}_i} \qquad (2.5.1.2)$$

or multiplying (2.5.1.2) by s_i:

$$s_i \frac{\partial G(\underline{s};t)}{\partial s_i} = \sum_{\underline{n}\in\Omega} n_i\, P(\underline{n};t)\underline{s}^{\underline{n}} \qquad (2.5.1.3)$$

Clearly (2.5.1.2) can be written as

$$\frac{\partial G(\underline{s};t)}{\partial s_i} = \sum_{\underline{n}\in\Omega_i^-} (n_i+1)\, P(\underline{n}+\underline{e}_i;t)\underline{s}^{\underline{n}}$$

where we have changed the index of summation from running over the space Ω to running over the space $\Omega_i^- \triangleq \{\underline{n}-\underline{e}_i \,|\, \underline{n}\in\Omega\}$. It is easily seen that the only element in Ω_i^- that is not in Ω, is the element

$\underline{n}-\underline{e}_i$ and that

$$(n_i+1)P(\underline{n}+\underline{e}_i;t)\underline{g}^{\underline{n}}\Big|_{\underline{n}-\underline{e}_i} = 0$$

since the i-th element of the vector $\underline{n}-\underline{e}_i$ is -1 i.e. $(n_i+1) = 0$.

Thus the specific index $\underline{n}-\underline{e}_i$ can be dropped from the summation index set Ω_i^- leaving the index set Ω itself. i.e.

$$\frac{\partial G(\underline{g};t)}{\partial s_i} = \sum_{\underline{n}\in\Omega} (n_i+1)P(\underline{n}+\underline{e}_i;t)\underline{g}^{\underline{n}}. \qquad (2.5.1.4)$$

Similarly by changing the summation index set from Ω to $\Omega_i^+ \triangleq \{\underline{n}+\underline{e}_i | \underline{n}\in\Omega\}$ and multiplying by s_i, (2.5.1.3) can be written as

$$s_i^2 \frac{\partial G(\underline{g};t)}{\partial s_i} = \sum_{\underline{n}\in\Omega_i^+} (n_i-1)P(\underline{n}-\underline{e}_i;t)\underline{g}^{\underline{n}}.$$

Clearly $\{\underline{0}, \underline{n}+\underline{e}_i | \underline{n}\in\Omega\} = \Omega$ and by definition 2.2.4 $P(\underline{0}-\underline{e}_i;t) = 0$ since $\underline{0}-\underline{e}_i \notin \Omega$, so that adding the vector $\underline{0}$ to the index set Ω_i^+ will not alter the summation and the index set can be taken as Ω i.e.

$$s_i^2 \frac{\partial G(\underline{g};t)}{\partial s_i} = \sum_{\underline{n}\in\Omega} (n_i-1)P(\underline{n}-\underline{e}_i;t)\underline{g}^{\underline{n}}. \qquad (2.5.1.5)$$

As in (2.5.1.2) we can easily show for $i\neq j$ that

$$\frac{\partial^2 G(\underline{g};t)}{\partial s_i\, \partial s_j} = \sum_{\underline{n}\in\Omega} n_i n_j P(\underline{n};t)\underline{g}^{\underline{n}-\underline{e}_i-\underline{e}_j}.$$

Using this identity, identities (2.5.1.1) and (2.5.1.2) and where applicable the summation index set arguments discussed above, we can derive the following identities:

$$\frac{1}{s_i} G(\underline{s};t) = \sum_{\underline{n}\in\Omega_i^-} P(\underline{n}+\underline{e}_i;t)\underline{s}^{\underline{n}} \qquad (2.5.1.6)$$

$$s_i G(\underline{s};t) = \sum_{\underline{n}\in\Omega} P(\underline{n}-\underline{e}_i;t)\underline{s}^{\underline{n}} \qquad (2.5.1.7)$$

$$\frac{s_i}{s_i}\frac{\partial G(\underline{s};t)}{\partial s_i} = \sum_{\underline{n}\in\Omega_i^-} n_j P(\underline{n}+\underline{e}_i;t)\underline{s}^{\underline{n}} \qquad (2.5.1.8)$$

$$s_i s_j \frac{\partial G(\underline{s};t)}{\partial s_j} = \sum_{\underline{n}\in\Omega} n_j P(\underline{n}-\underline{e}_i;t)\underline{s}^{\underline{n}} \qquad (2.5.1.9)$$

$$s_j \frac{\partial^2 G(\underline{s};t)}{\partial s_i\, \partial s_j} = \sum_{\underline{n}\in\Omega} (n_i+1)n_j P(\underline{n}+\underline{e}_i;t)\underline{s}^{\underline{n}} \qquad (2.5.1.10)$$

$$s_i s_j \frac{\partial^2 G(\underline{s};t)}{\partial s_i\, \partial s_j} = \sum_{\underline{n}\in\Omega} n_i n_j P(\underline{n};t)\underline{s}^{\underline{n}} \qquad (2.5.1.11)$$

$$s_i^2 s_j \frac{\partial^2 G(\underline{s};t)}{\partial s_i\, \partial s_j} = \sum_{\underline{n}\in\Omega} (n_i-1)n_j P(\underline{n}-\underline{e}_i;t)\underline{s}^{\underline{n}} \qquad (2.5.1.12)$$

Further, since $s_i \frac{\partial}{\partial s_i} \sum_{\underline{n}\in\Omega} n_i P(\underline{n};t)\underline{s}^{\underline{n}} = \sum_{\underline{n}\in\Omega} n_i^2 P(\underline{n};t)$ we have

$$s_i \frac{\partial^2 G(\underset{\sim}{s};t)}{\partial s_i^2} + \left\{\frac{\partial G(\underset{\sim}{s};t)}{\partial s_i}\right\} = \sum_{\underset{\sim}{n}\in\Omega} (n_i+1)^2 P(\underset{\sim}{n}+\underset{\sim}{e}_i;t)\underset{\sim}{s}^{\underset{\sim}{n}} \qquad (2.5.1.13)$$

$$s_i^2 \frac{\partial^2 G(\underset{\sim}{s};t)}{\partial s_i^2} + s_i\left\{\frac{\partial G(\underset{\sim}{s};t)}{\partial s_i}\right\} = \sum_{\underset{\sim}{n}\in\Omega} n_i^2 P(\underset{\sim}{n};t)\underset{\sim}{s}^{\underset{\sim}{n}} \qquad (2.5.1.14)$$

$$s_i^3 \frac{\partial^2 G(\underset{\sim}{s};t)}{\partial s_i^2} + s_i^2\left\{\frac{\partial G(\underset{\sim}{s};t)}{\partial s_i}\right\} = \sum_{\underset{\sim}{n}\in\Omega} (n_i-1)^2 P(\underset{\sim}{n}-\underset{\sim}{e}_i;t)\underset{\sim}{s}^{\underset{\sim}{n}} \qquad (2.5.1.15)$$

Note:

Since (2.5.1.6) and (2.5.1.8) do not contain a term in (n_i+1), they are the only expressions where summation cannot be taken over Ω, as was the case in (2.5.1.4). We shall see that this problem arises in models of the type 2.4A, but can be overcome because the birth and death parameters $\lambda_i(\cdot,\cdot,\cdot,\cdot)$ and $\mu_i(\cdot,\cdot,\cdot,\cdot)$ are defined to be zero when $n_i = 0$.

Differentiating (2.5.1.1) with respect to t we obtain

$$\frac{\partial G(\underset{\sim}{s};t)}{\partial t} = \sum_{\underset{\sim}{n}\in\Omega} P'(\underset{\sim}{n};t)\underset{\sim}{s}^{\underset{\sim}{n}}. \qquad (2.5.2)$$

Consider equation (2.4.3). Multiplying it by $\underset{\sim}{s}^{\underset{\sim}{n}}$ and summing over all $\underset{\sim}{n}\in\Omega$ we obtain:

$$\sum_{\underline{n}\in\Omega} P'(\underline{n};t)s^{\underline{n}} = \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m}\{\lambda_{ij}(t)\sum_{\underline{n}\in\Omega} n_j\, P(\underline{n}-\underline{e}_i;t)s^{\underline{n}}$$

$$- [\lambda_{ij}(t)+\mu_{ij}(t)]\sum_{\underline{n}\in\Omega} n_j\, P(\underline{n};t)s^{\underline{n}}$$

$$+ \mu_{ij}(t)\sum_{\underline{n}\in\Omega} n_j\, P(\underline{n}+\underline{e}_i;t)s^{\underline{n}}\}$$

$$+ \sum_{i=1}^{m}\{\lambda_{ii}(t)\sum_{\underline{n}\in\Omega}(n_i-1)P(\underline{n}-\underline{e}_i;t)s^{\underline{n}}$$

$$- [\lambda_{ii}(t)+\mu_{ii}(t)]\sum_{\underline{n}\in\Omega} n_i\, P(\underline{n};t)s^{\underline{n}}$$

$$+ \mu_{ii}(t)\sum_{\underline{n}\in\Omega}(n_i+1)P(\underline{n}+\underline{e}_i;t)s^{\underline{n}}\}$$

$$+ \sum_{i=1}^{m}\sum_{r=1}^{p}\{\alpha_{ir}(t)u_r(t)\sum_{\underline{n}\in\Omega} P(\underline{n}-\underline{e}_i;t)s^{\underline{n}}$$

$$- [\alpha_{ir}(t)+\beta_{ir}(t)]u_r(t)\sum_{\underline{n}\in\Omega} P(\underline{n};t)s^{\underline{n}}$$

$$+ \beta_{ir}(t)u_r(t)\sum_{\underline{n}\in\Omega} P(\underline{n}+\underline{e}_i;t)s^{\underline{n}}\} \qquad (2.5.3)$$

The process of taking the summation over $\underline{n}\in\Omega$ inside the finite summations over i, j and r is valid (as previously remarked) for *well-behaved distributions* which lead to uniformly convergent sums for $|s_i| \leq$ $i=1,\ldots,m$.

In (2.5.3), since the birth and death parameters of section 2.4A (defined in (2.4.1) and (2.4.2)) are zero when the size of the i-th species is zero, we can replace

$$\mu_{ij}(t) \sum_{\underline{n}\in\Omega} n_j P(\underline{n}+\underline{e}_i;t)\underline{g}^{\underline{n}} \quad \text{and} \quad \beta_{ij}(t)u_r(t) \sum_{\underline{n}\in\Omega} P(\underline{n}+\underline{e}_i;t)\underline{g}^{\underline{n}}$$

by the same expressions, except that summation can be taken over Ω_i^- instead of Ω.

Thus using expressions of the type (2.5.1.1) to (2.5.1.9) equation (2.5.3) can be written as

$$\frac{\partial G(\underline{g};t)}{\partial t} = \sum_{i=1}^{m}\sum_{j=1}^{m} (\lambda_{ij}(t) - \mu_{ij}(t)/s_i)s_j(s_i-1)\frac{\partial G(\underline{g};t)}{\partial s_i}$$

$$+ \sum_{i=1}^{m}\sum_{r=1}^{p} (\alpha_{ir}(t) - \beta_{ir}(t)/s_i)(s_i-1)G(\underline{g};t)u_r(t). \quad (2.5.4)$$

Similarly, equation (2.4.6) can be multiplied by $\underline{g}^{\underline{n}}$ summed over $\underline{n}\in\Omega$ and using identities (2.5.1.1) to (2.5.1.15) we obtain

$$\frac{\partial G(\underline{g};t)}{\partial t} = \sum_{i=1}^{m} \{[\lambda_i(\underline{\chi}(t),t)s_i - \mu_i(\underline{\chi}(t),t)] \ (s_i-1)\frac{\partial G(\underline{g};t)}{\partial s_i}$$

$$+ (\lambda_{ii}(t)s_i - \mu_{ii}(t))(s_i-1)\left[s_i\frac{\partial^2 G(\underline{g};t)}{\partial s_i^2} + \frac{\partial G(\underline{g};t)}{\partial s_i}\right]\}$$

$$+ \sum_{\substack{i=1\\ j\neq i}}^{m}\sum_{j=1}^{m} (\lambda_{ij}(t)s_i - \mu_{ij}(t))s_j(s_i-1)\frac{\partial^2 G(\underline{g};t)}{\partial s_i \partial s_j}. \quad (2.5.5)$$

Solving (2.5.4) and (2.5.5) for $G(\underline{g};t)$ subject to the initial condition

$$G(\underline{g};0) = \sum_{\underline{n}\in\Omega} P(\underline{n};0)\underline{g}^{\underline{n}}$$

we can, from (2.5.1.1) find $P(\underline{n};t)$ for each $\underline{n}\in\Omega$ as the coefficient of $\underline{s}^{\underline{n}}$ in $G(\underline{s};t)$. However, solving the above equations for $G(\underline{s};t)$ proves to be extremely difficult and even then it may be impossible to find a closed form expression for the coefficients of $\underline{s}^{\underline{n}}$ in $G(\underline{s};t)$. In most circumstances it may be sufficient to know only the first few moments of the multivariate distribution, e.g. means, variance and covariance terms. These can be found as follows:

Put $\underline{s}=\underline{1}$ (see definition 2.2.2) in expressions (5.1.1) to (5.1.15) and (5.2). Then we have

$$G(\underline{1};t) = \sum_{\underline{n}\in\Omega} P(\underline{n};t) = 1 \qquad (2.5.6.1)$$

since $P(\underline{n};t)$ is a probability distribution.

$$\frac{\partial G(\underline{1};t)}{\partial s_i} = \sum_{\underline{n}\in\Omega} n_i P(\underline{n};t) \triangleq E(n_i) \qquad (2.5.6.2)$$

$$\frac{\partial^2 G(\underline{1};t)}{\partial s_i^2} = \sum_{\underline{n}\in\Omega} n_i(n_i-1)P(\underline{n};t) \triangleq E(n_i^2) - E(n_i) \qquad (2.5.6.3)$$

$$\frac{\partial^2 G(\underline{1};t)}{\partial s_i\,\partial s_j} = \sum_{\underline{n}\in\Omega} n_i n_j P(\underline{n};t) \triangleq E(n_i n_j) \qquad i\neq j. \qquad (2.5.6.4)$$

Clearly since $G(\underline{s};t)$ is well-behaved for well-defined probability distributions

$$E(n_i n_j) = E(n_j n_i)$$

$$\frac{\partial^2 G(\underline{1};t)}{\partial s_i^2} + \frac{\partial G(\underline{1};t)}{\partial s_i} = \sum_{\underline{n}\in\Omega} n_i^2 P(\underline{n};t) \triangleq E(n_i^2). \qquad (2.5.6.5)$$

Further for a well-defined probability distribution

$$\frac{\partial}{\partial s_i}\frac{\partial G(\underline{s};t)}{\partial t} = \frac{\partial}{\partial t}\frac{\partial G(\underline{s};t)}{\partial s_i}$$

so that

$$\frac{\partial}{\partial t}\frac{\partial G(\underline{1};t)}{\partial s_i} = \frac{d}{dt}E(n_i). \qquad (2.5.6.6)$$

Similarly, we can easily show that

$$\frac{\partial^3 G(\underline{1};t)}{\partial s_i \partial s_j \partial s_k} = \sum_{\underline{n}\in\Omega} n_i n_j n_k P(\underline{n};t) \triangleq E(n_i, n_j, n_k) \quad \ldots \; i\neq j\neq k \qquad (2.5.6.7)$$

$$\frac{\partial^3 G(\underline{1};t)}{\partial s_i^2 \partial s_j} = \sum_{\underline{n}\in\Omega} n_i(n_i-1)n_j P(\underline{n};t) \triangleq E(n_i^2 n_j) - E(n_i n_j) \quad i\neq j \qquad (2.5.6.8)$$

$$\frac{\partial}{\partial s_i}\left[\frac{\partial^2 G(\underline{1};t)}{\partial s_i^2} + \frac{\partial G(\underline{1};t)}{\partial s_i} - \right] = \sum_{\underline{n}\in\Omega} n_i^2(n_i-1)P(\underline{n};t)$$

$$\triangleq E(n_i^3) - E(n_i^2) \qquad (2.5.6.9)$$

$$\frac{\partial}{\partial s_j}\left[\frac{\partial^2 G(\underline{1};t)}{\partial s_i^2} + \frac{\partial G(\underline{1};t)}{\partial s_i} - \right] = \sum_{\underline{n}\in\Omega} n_i^2 n_j P(\underline{n};t)$$

$$\triangleq E(n_i^2 n_j), \qquad (2.5.6.10)$$

If we denote the mean of the i-th population by $\bar{n}_i$, its variance by σ_{ii}, its covariance with the j-th population by σ_{ij} and

the general third order moment of the i-th, j-th and k-th populations around the mean, where possibly j=i or k=j=i etc., by σ_{ijk}, then we have the following definitions and relationships:

$$\bar{n}_i \triangleq E(n_i) \qquad i=1,\ldots,m. \tag{2.5.7.1}$$

Since $\sigma_{ij} \triangleq E(n_i-\bar{n}_i)(n_j-\bar{n}_i)$ we have

$$E(n_in_j) = \sigma_{ij} + \bar{n}_i\bar{n}_j \qquad i,j=1,\ldots,m. \tag{2.5.7.2}$$

Since $\sigma_{ijk} \triangleq E(n_i-\bar{n}_i)(n_j-\bar{n}_j)(n_k-\bar{n}_k)$ we have

$$E(n_in_jn_k) = \sigma_{ijk} + \bar{n}_i\sigma_{jk} + \bar{n}_j\sigma_{ik} + \bar{n}_k\sigma_{ij} + \bar{n}_i\bar{n}_j\bar{n}_k \qquad i,j,k=1,\ldots,m. \tag{2.5.7.3}$$

Clearly (2.5.7.3) is symmetrical with respect to the indeces i, j and k. As in (2.5.6.6) we can derive the following expressions using (2.5.7.2)

$$\frac{\partial^3 G(1;t)}{\partial s_i \partial s_j \partial t} = \frac{d}{dt} E(n_in_j)$$

$$= \frac{d\sigma_{ij}}{dt} + \bar{n}_i \frac{d\bar{n}_j}{dt} + \bar{n}_j \frac{d\bar{n}_i}{dt} \tag{2.5.8.1}$$

$$\frac{\partial}{\partial t}\left[\frac{\partial^2 G(1;t)}{\partial s_i^2} + \frac{\partial G(1;t)}{\partial s_i}\right] = \frac{d}{dt} E(n_i^2) \tag{2.5.8.2}$$

$$= \frac{d\sigma_{ii}}{dt} + 2\bar{n}_i \frac{d\bar{n}_i}{dt}$$

I[illegible]entiate (2.5.4) or (2.5.5) with respect to s_i for some i [illegible]1,...,m] and then put $s=1$ we obtain, using expressions (2.5.6.·) and (2.5.7.·), an ordinary differential equation in t for $\bar{n}_i(t)$ which is then the mean of the distribution of the process corresponding to the particular set of equations that have been differentiated. Similarly differentiating (2.5.4) or (2.5.5) twice, first with respect to s_i and then with respect to s_j (where possibly j=i), using expressions (2.5.6.·), (2.5.7.·) and (2.5.8.·) we obtain an ordinary differential in t for $\sigma_{ij}(t)$. Ordinary differential equations in t for higher-order moments can be similarly derived.

Proceeding as described above, we can from equation (2.5.4), derive the equations describing the dynamic behaviour of the means, variance terms and covariance terms of the linear control system defined in 2.4A. These equations turn out to be the following:

$$\frac{d\bar{n}_i}{dt} = \sum_{j=1}^{m} [\lambda_{ij}(t) - \mu_{ij}(t)]\bar{n}_j + \sum_{r=1}^{p} [\alpha_{ir}(t) - \beta_{ir}(t)]u_r(t) \qquad i=1,\ldots,m \qquad (2.5.9)$$

$$\frac{d\sigma_{ii}}{dt} = \sum_{j=1}^{m} 2[\lambda_{ij}(t) - \mu_{ij}(t)]\sigma_{ij} + [\lambda_{ij}(t) + \mu_{ij}(t)]\bar{n}_j + \sum_{r=1}^{p} [\alpha_{ir}(t) + \beta_{ir}(t)]u_r(t) \qquad i=1,\ldots,m \qquad (2.5.10)$$

$$\frac{d\sigma_{i\ell}}{dt} = \sum_{j=1}^{m} \{[\lambda_{\ell j}(t) - \mu_{\ell j}(t)]\sigma_{ij} + [\lambda_{ij}(t) - \mu_{ij}(t)]\sigma_{\ell j}\}$$

$$i,\ell=1,\ldots,m \quad i\neq\ell \qquad (2.5.11)$$

Similarly, from equation (2.5.5) the following equations arise:

$$\frac{d\bar{n}_i}{dt} = [\lambda_i(x(t),t) - \mu_i(x(t),t)]\bar{n}_i + \sum_{j=1}^{m}[\lambda_{ij}(t) - \mu_{ij}(t)](\bar{n}_i\bar{n}_j + \sigma_{ij})$$

$$i=1,\ldots,m \qquad (2.5.12)$$

$$\frac{d\sigma_{ii}}{dt} = [\lambda_i(x(t),t) + \mu_i(x(t),t)]\bar{n}_i + 2[\lambda_i(x(t),t) - \mu_i(x(t),t)]\sigma_{ii}$$

$$+ \sum_{j=1}^{m}\{2[\lambda_{ij}(t) - \mu_{ij}(t)](\bar{n}_i\sigma_{ij} + \bar{n}_j\sigma_{ii} + \sigma_{ii}\bar{n}_j) +$$

$$+ [\lambda_{ij}(t) + \mu_{ij}(t)](\bar{n}_i\bar{n}_j + \sigma_{ij})\}$$

$$i=1,\ldots,m \qquad (2.5.13)$$

$$\frac{d\sigma_{i\ell}}{dt} = [\lambda_i(x(t),t) - \mu_i(x(t),t) + \lambda_\ell(x(t),t) - \mu_\ell(x(t),t)]\sigma_{i\ell}$$

$$+ \sum_{j=1}^{m}\{[\lambda_{\ell j}(t) - \mu_{\ell j}(t)](\bar{n}_i\sigma_{ij} + \bar{n}_j\sigma_{i\ell}) +$$

$$+ [\lambda_{ij}(t) - \mu_{ij}(t)](\bar{n}_i\sigma_{\ell j} + \bar{n}_j\sigma_{i\ell} + \sigma_{i\ell j})\}$$

$$i,\ell=1,\ldots,m \quad \ell\neq i \qquad (2.5.14)$$

Equations (2.5.12) describe the dynamic behaviour of the mean size of the m different species in a Lotka-Volterra competitive system as defined in section 2.4B. Equations (2.5.13) and (2.5.14)

describe the dynamic behaviour of the variance and covariance terms of this system.

It should be noted that equations (2.5.12), (2.5.13) and (2.5.14) cannot be solved directly since (2.5.13) and (2.5.14) are dependent of the third order moments $\sigma_{ij,j}$. A method for dealing with this problem is discussed in more detail after equation (3.3.10) in the next chapter.

As previously mentioned, under certain conditions there will exist, for a time homogeneous system, a limiting distribution (Gray 1971). The generating function for this distribution will, by (2.3.3) and (2.5.1.1), be

$$G(\underline{s}) = \sum_{\underline{x}\in\Omega} P(\underline{x})\underline{s}^{\underline{x}}$$

since the probabilities will no longer be a function of t. Thus $\frac{\partial G(\underline{s})}{\partial t} = 0$, so that replacing $G(\underline{s};t)$ by $G(\underline{s})$ in equations (2.5.4) and (2.5.5) and putting the left-hand sides of these equations equal to zero we can solve each of these two systems for $G(\underline{s})$. Further the means, variance and covariance equations of each system will satisfy $\frac{d\bar{m}_k}{dt} = 0$, $\frac{d\sigma_{kk}}{dt} = 0$ and $\frac{d\sigma_{kk}}{dt} = 0$ respectively and can be found by solving the corresponding set of algebraic equations that will arise by setting the left-hand sides of equations (2.5.9) to (2.5.14) equal to zero.

Systems {(2.5.9), (2.5.10), (2.5.11)} and {(2.5.12), (2.5.13), (2.5.14)} can now be used as stochastic linear and Lotka-Volterra

type population models. The linear system is discussed more fully in chapter four, especially with regard to population harvesting models. In chapter five an analysis of the stability properties of the Lotka-Volterra system is pursued further.

CHAPTER 3

MULTIPLE EVENT MULTIVARIATE BIRTH AND DEATH PROCESSES

3.1 INTRODUCTION

In this chapter we are concerned with relaxing condition (i) as given in the definition (see 2.2.6) of an m-dimensional multivariate birth and death process, to allow $P(\underline{n},\underline{n}^*;t,\Delta t)/\Delta t$ to be non-zero in the limit as Δt tends to zero, for certain $\underline{n},\underline{n}^* \in \Omega$ satisfying $\|\underline{n}-\underline{n}^*\| \geq 2$. The procedure for relaxing this condition is best demonstrated by an example. In this respect we will, in this chapter, formulate a non-linear migration process allowing multiple births and group migration, to occur.

Recalling definition 2.2.3, a migration process allowing a group of k individuals to migrate from one colony (species) to another is characterized by $P(\underline{n},\underline{n}^*;t,\Delta t)/\Delta t$ being non-zero in the limit as Δt tends to zero, for certain $\underline{n},\underline{n}^* \in \Omega$ satisfying $\|\underline{n}-\underline{n}^*\| = 2k$ and $\|\underline{n}\| = \|\underline{n}^*\|$. A more explicit definition is given in the next section, where the transition probabilities for the non-linear multiple migration process are defined.

Models of migration processes between colonies subject to birth and death processes have been used to analyze spatially distributed populations (Bailey 1968), and have important applications in the life sciences (Iosifescu and Tautu 1973) and social sciences.

The early work of Bailey (1968), Puri (1968) and Whittle (1967, 1968) on migration processes has been followed by the more recent

work of Renshaw (1972) and Aksland (1975). The following underlying assumptions apply to the models analysed in the research papers mentioned above: the migration rate is assumed to be either linearly (Puri 1968, Renshaw 1972 and Aksland 1975) or non-linearly (Whittle 1967 and 1968) dependent on the size of the colony from which an individual is emigrating; the migration rate is not dependent on the size of the colony to which an individual is immigrating; and only single individuals are permitted to migrate at any instant in time.

The above restrictions prove unrealistic in modelling certain systems, especially in the social sciences, for the following reasons. Firstly, it may be more reasonable, in these systems, to assume that the rate of migration between two colonies is dependent on the size of both colonies, and secondly multiple events can realistically occur as in the case of a "family" of individuals in a colony simultaneously migrating to another colony.

In this chapter we derive the Kolmogorov forward equations for migration processes between colonies subject to non-linear birth and death processes. The migration rates will dependent non-linearly on the size of the two colonies concerned and the event of a multiple birth in any colony is allowed to occur with non-zero probability. The Kolmogorov equations are then used to derive a system of ordinary differential equations in the means, variance terms and covariance terms of the probability distributions describing the whole process.

3.2 FORMULATION OF A NON-LINEAR MULTIPLE EVENT MIGRATION PROCESS

Recalling the definitions of section 2.2, we restrict ourselves to processes for which

$$P(\underline{n},\underline{n}';t,\Delta t)P(\underline{n}',\underline{n}^*;t+\alpha\Delta t,(1-\alpha)\Delta t) = o(\Delta t)$$

for all α satisfying $0 < \alpha < 1$.

This allows us to write down the following equation

$$P(\underline{n};t+\Delta t) = \sum_{\underline{n}^*\in\Omega} P(\underline{n}^*;t)P(\underline{n}^*,\underline{n};t,\Delta t) + o(\Delta t). \qquad (3.2.1)$$

The migration process considered in this paper is formulated under the following assumptions:

$$P(\underline{n},\underline{n}+r\underline{e}_i;t,\Delta t) = \lambda_{ir}(t)n_i(k_{ii}-n_i)\Delta t + o(\Delta t) \quad r=1,\ldots,q_i \quad k_{ii}\in Z^+, \quad i=1,\ldots,m \qquad (3.2.2)$$

$$P(\underline{n},\underline{n}-\underline{e}_i;t,\Delta t) = \mu_i(t)n_i\Delta t + o(\Delta t) \qquad i=1,\ldots,m \qquad (3.2.3)$$

$$P(\underline{n},\underline{n}-r\underline{e}_i+r\underline{e}_j;t,\Delta t) = \gamma_{ijr}(t)n_i(k_{ij}-n_j)\Delta t + o(\Delta t) \quad n_i-r \geq 0 \quad r=1,\ldots,q_{ij}, \; k_{ij}\in Z^+, \; i\neq j \; i,j=1,\ldots,m$$
$$= 0 \quad i=j \text{ or } n_i-r < 0 \qquad i,j=1,\ldots,m \qquad (3.2.4)$$

$$P(\underline{n},\underline{n}^*;t,\Delta t) = o(\Delta t) \text{ for all } \underline{n}^*\in\Omega \text{ other than } \underline{n}+r\underline{e}_i, \; \underline{n}-\underline{e}_i \text{ and } \underline{n}-r\underline{e}_i+r\underline{e}_j \quad i\neq j \quad i,j=1,\ldots,m. \qquad (3.2.5)$$

A physical interpretation of (3.2.2) - (3.2.5) can be made as follows:

(i) The i-th colony, i=1,...,n, is subject to a non-linear, multiple birth and death process defined by (3.2.2) and (3.2.3). More specifically: for $n_i \leq k_{ii}$ the birth rate for the i-th colony is

$$\sum_{r=1}^{q_i} \lambda_{ir}(t) n_i (k_{ii} - n_i),$$

where r denotes the litter size of a birth event, and the death rate for the i-th colony is $\mu_i(t) n_i$: for $n_i > k_{ii}$, the birth rate of the i-th colony is zero and the death rate is given by

$$\mu_i(t) n_i + \sum_{r=1}^{q_i} \lambda_{ir}(t) n_i (n_i - k_{ii}).$$

$\sum_{r=1}^{q_i} \lambda_{ir}(t) n_i (k_{ii} - n_i)$ switches from a birth rate to a death rate when the i-th colony reaches size k_{ii}, since the sign of this expression switches from positive to negative at this point.

Clearly k_{ii} is a saturation population level for the i-th colony, beyond which no births occur. The concept of a saturation level in population model, has been frequently used (Goel et. al. 1971).

As an illustration, Diagrams 1 and 2 graph the birth and death rates respectively as a function of colony size, for parameters λ_{ir}, $r=1,\ldots,q_i$ and μ_i independent of time.

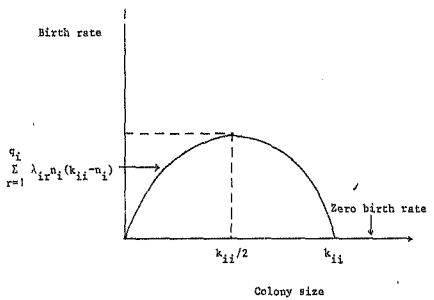


DIAGRAM 1. Birth rate in the i-th colony

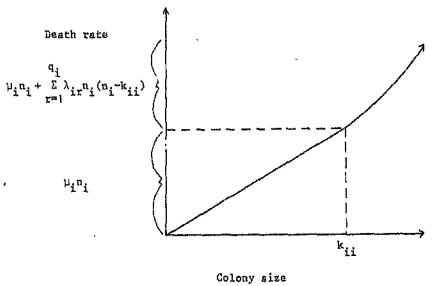


DIAGRAM 2. Death rate in the i-th colony

(ii) The migration process between the i-th and j-th colonies is defined by (3.2.4). The index r in this case refers to the number of individuals in a particular "family" that are migrating from the i-th colony to the j-th colony at the rate

$$\gamma_{ij_r}(t) n_i (k_{ij} - n_j).$$

Clearly if the size of the i-th colony is large and the size of the j-th colony is small (compared with k_{ij}) the migration rate is comparatively high. This rate drops as the size of the recipient colony (j-th in this case) approaches a saturation

level given by k_{jj}. Although the saturation level is denoted here as being dependent on the donor colony (j-th in this case) in most real situations, the assumption

$$k_{ij} = k_{jj} \qquad i=1,\ldots,m$$

will hold, i.e. the saturation level does not distinguish between individuals from different colonies, but depends only on the numbers of individuals concerned.

As in (1)

$$\sum_{r=1}^{q_{ij}} \gamma_{ijr}(t) n_i (k_{ij} - n_j)$$

changes sign at $n_j = k_{ij}$ so that for $n_j > k_{ij}$ the actual direction of migration of individuals from the i-th colony to the j-th colony reverses.

It should be noted that if a given colony is above its saturation level and all other colonies are below, the probability of the given colony increasing in a time interval Δt is $o(\Delta t)$. Further, if all colonies are above their respective saturation levels, the probability of a new individual being born in the total system in a time interval Δt is also $o(\Delta t)$. Hence the saturation levels are probabilistic upper bounds on the size of the colonies within the system.

In the two-colony case, letting $k_{11} = k_{21} = k_1$ and $k_{12} = k_{22} = k_2$ the actual birth and death rates for the four cases that arise are given in the tables below.

Two-colony birth, death, and migration model with saturation levels k_1 and k_2 in colonies 1 and 2 respectively

	Colony 1	Colony 2
Case 1	$n_1 \leq k_1$	$n_2 \leq k_2$
Birth rate	$\sum_{r=1}^{q_1} \lambda_{1r} n_1 (k_1 - n_1)$	$\sum_{r=2}^{q_2} \lambda_{1r} n_2 (k_2 - n_2)$
Death rate	$\mu_1 n_1$	$\mu_2 n_2$
Migration rate	$1 \rightarrow 2$ $\sum_{r=1}^{q_{12}} \gamma_{12r}(t) n_1 (k_2 - n_2)$	
	$2 \rightarrow 1$ $\sum_{r=1}^{q_{21}} \gamma_{21r}(t) n_2 (k_1 - n_1)$	
Case 2	$n_1 > k_1$	$n_2 \leq k_2$
Birth rate	nil	$\sum_{r=2}^{q_2} \lambda_{2r} n_2 (k_2 - n_2)$
Death rate	$\mu_1 n_1 + \sum_{r=1}^{q_1} \lambda_{1r} n_1 (n_1 - k_1)$	$\mu_2 n_2$
Migration rate	$1 \rightarrow 2$ $\sum_{r=1}^{q_{12}} \gamma_{12r}(t) n_1 (k_2 - n_2) + \sum_{r=1}^{q_{21}} \gamma_{21r}(t) n_2 (n_1 - k_1)$	
	$2 \rightarrow 1$ nil	

Table continued 73/ .....

Table continued

	Colony 1	Colony 2
Case 3	$n_1 \leq k_1$	$n_2 > k_2$
Birth rate	$\sum_{r=1}^{q_1} \lambda_{1r} n_1 (k_1 - n_1)$	nil
Death rate	$\mu_1 n_1$	$\mu_2 n_2 + \sum_{r=2}^{q_2} \lambda_{2r} n_2 (n_2 - k_2)$
Migration rate	$1 \to 2$: nil $2 \to 1$: $\sum_{r=1}^{q_{12}} \gamma_{12r}(t) n_1 (n_2 - k_2) + \sum_{r=1}^{q_{21}} \gamma_{21r}(t) n_2 (k_1 - n_1)$	
Case 4	$n_1 > k_1$	$n_2 > k_2$
Birth rate	nil	nil
Death rate	$\mu_1 n_1 + \sum_{r=1}^{q_1} \lambda_{1r} n_1 (n_1 - k_1)$	$\mu_2 n_2 + \sum_{r=1}^{q_2} \lambda_{2r} n_2 (n_2 - k_2)$
Migration rate	$1 \to 2$: $\sum_{r=1}^{q_{21}} \gamma_{21r}(t) n_2 (n_1 - k_1)$ $2 \to 1$: $\sum_{r=1}^{q_{12}} \gamma_{12r}(t) n_1 (n_2 - k_2)$	

3.3 DIFFERENTIAL EQUATIONS FOR THE MOMENTS OF THE DISTRIBUTION

Using expressions (3.2.1) – (3.2.5) we derive the following differential equation:

$$\frac{d}{dt}P(\underline{n};t) = \sum_{i=1}^{m}\{\sum_{r=1}^{q_i}\lambda_{ir}(t)[k_{ii}(n_i-r)-(n_i-r)^2]P(\underline{n}-r\underline{e}_i;t) -$$

$$-[\sum_{r=1}^{q_i}\lambda_{ir}(t)(k_{ii}n_i-n_i^2)+\mu_i(t)n_i]P(\underline{n};t) +$$

$$+\mu_i(t)n_i+1)P(\underline{n}+\underline{e}_i;t)\} +$$

$$+\sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m}\{\sum_{r=1}^{q_{ij}}\gamma_{ijr}(t)[k_{ij}(n_i+r) -$$

$$-(n_i+r)(n_j-r)]P(\underline{n}+r\underline{e}_i-r\underline{e}_j;t) -$$

$$-\gamma_{ijr}(t)(k_{ij}n_i-n_in_j)P(\underline{n};t)\}. \qquad (3.3.1)$$

Equation (3.3.1) is thus a difference-differential equation governing the dynamics of the probability distribution $P(\underline{n};t)$ of the migration process $X(t)$ defined in the previous section.

To derive a partial differential equation in the probability generating function

$$G(\underline{s};t) = \sum_{\underline{n}\in\Omega} P(\underline{n};t)\underline{s}^{\underline{n}} \qquad (3.3.2)$$

from (3.3.1) for the migration process defined in the previous section, as was done for the two processes discussed in the second chapter (see equations (2.5.4) and (2.5.5)) we need to generalize

expressions (2.5.1.').

It can be shown that for any integer r, the following expressions hold for i,j=1,...,n:

$$s_i^{1-r}\frac{\partial G(\underline{s};t)}{\partial s_i} = \Sigma(n_i+r)P(\underline{n}+re_i;t)\underline{s}^{\underline{n}} \tag{3.3.3}$$

$$s_i^{1-r}s_j^{r}\frac{\partial G(\underline{s};t)}{\partial s_i} = \Sigma(n_i+r)P(\underline{n}+re_i-re_j;t)\underline{s}^{\underline{n}} \qquad i\neq j \tag{3.3.4}$$

$$s_i^{1-r}s_i^{r+1}\frac{\partial^2 G(\underline{s};t)}{\partial s_i\partial s_j} = \Sigma(n_i+r)(n_j-r)P(\underline{n}+re_i-re_j;t)\underline{s}^{\underline{n}} \qquad i\neq j \tag{3.3.5}$$

$$s_i^{r+1}\left[s_i\frac{\partial^2 G(\underline{s};t)}{\partial s_i^2}+\frac{\partial G(\underline{s};t)}{\partial s_i}\right] = \Sigma(n_i-r)^2P(\underline{n}-re_i;t)\underline{s}^{\underline{n}}. \tag{3.3.6}$$

The indices of summation in (3.3.4) - (3.3.6) are not over Ω. When multiplying (3.3.1) by $\underline{s}^{\underline{n}}$ and summing over $\underline{n}\in\Omega$, however, expressions (3.3.3) - (3.3.6) can be used since the fact that $P(\underline{n};t) = 0$ if $\underline{n}\notin\Omega$ and $\gamma_{ijr}(t) = 0$ if $n_i-r<0$ allows the summation indices to range over *all* $\underline{n}\in\Omega$ (see the more detailed argument given in section 2.5).

Using (3.3.3) - (3.3.6) and (2.5.2) equation (3.3.1) can, after rearranging terms, be written as:

$$\frac{\partial G(\underline{s};t)}{\partial t} = \sum_{i=1}^{m} \left[\sum_{r=1}^{q_i} s_i(s_i^r-1)k_{ii}\lambda_{ir}(t) - (s_i-1)\mu_i(t) + \right.$$

$$\left. + \sum_{\substack{j=1 \\ j\neq i}}^{m} \sum_{r=1}^{q_{ij}} s_i(s_i^{-r}s_j^r-1)k_{ij}\gamma_{ijr}(t) \right] \frac{\partial G(\underline{s};t)}{\partial s_i} -$$

$$- \sum_{i=1}^{m} \sum_{r=1}^{q_i} s_i(s_i^r-1)\lambda_{ir} \left[s_i \frac{\partial^2 G(\underline{s};t)}{\partial s_i^2} + \frac{\partial G(\underline{s};t)}{\partial s_i} \right] +$$

$$+ \sum_{\substack{i=1 \\ j\neq i}}^{m} \sum_{r=1}^{q_{ij}} s_i s_j(1-s_i^{-r}s_j^r)\gamma_{ijr}(t) \frac{\partial^2 G(\underline{s};t)}{\partial s_i \partial s_j} . \qquad (3.3.7)$$

Solving (3.3.7) subject to the initial condition

$$G(\underline{s};0) = \sum_{\underline{n}\in\Omega} P(\underline{n};0)\underline{s}^{\underline{n}}$$

we can from (3.3.2) find $P(\underline{n};t)$ for each $\underline{n}\in\Omega$ as the coefficient of $\underline{s}^{\underline{n}}$ in $G(\underline{s};t)$. It is not easy, however, to find a solution to (3.3.7). As in the previous section, however, we can find ordinary differential equations describing the means, variance terms and covariance terms of the distribution $P(\underline{n};t)$. These ordinary differential equations can then be solved numerically to yield all the first and second order moments. These are the only moments of the joint multivariate probability distribution $P(\underline{n};t)$ that are usually required.

Differentiating (3.3.7) an appropriate number of times with respect to g and using expressions (2.5.6.·), (2.5.7.·) and (2.5.8.·) the differential equations are:

$$\frac{d\bar{n}_i}{dt} = \left[\sum_{r=1}^{q_i} rk_{ii}\lambda_{ir}(t) - \mu_i(t) \right]\bar{n}_i + \sum_{\substack{j=1 \\ j\neq i}}^{m} \left[\sum_{r=1}^{q_{ji}} rk_{ji}\gamma_{ji}(t)\bar{n}_j - \sum_{r=1}^{q_{ij}} rk_{ij}\gamma_{ij}(t)\bar{n}_i \right] - \sum_{j=1}^{q_i} r\lambda_{ir}(t)(\bar{n}_i^2+\sigma_{ii}) + \sum_{\substack{j=1 \\ j\neq i}}^{m} \left[\sum_{r=1}^{q_{ij}} r\gamma_{ijr}(t) - \sum_{r=1}^{q_{ji}} r\gamma_{jir}(t) \right](\bar{n}_i\bar{n}_j+\sigma_{ij})$$

$$i=1,\ldots,m \qquad (3.3.8)$$

$$\frac{d\sigma_{ii}}{dt} = \left[\sum_{r=1}^{q_i} r^2k_{ii}\lambda_{ir}(t) + \mu_i \right]\bar{n}_i + \left[\sum_{r=1}^{q_i} (2rk_{ii}-r^2)\lambda_{ir}(t) - \mu_i(t) \right]\sigma_{ii} +$$

$$+ \sum_{\substack{j=1 \\ j\neq i}}^{m} \sum_{r=1}^{q_{ij}} \{(-2rk_{ij}\sigma_{ii}-r^2\sigma_{ij}-r^2\bar{n}_i\bar{n}_j+(r^2-2r)k_{ij}\bar{n}_i\}\gamma_{ijr}(t) +$$

$$+ \sum_{\substack{j=1 \\ j\neq i}}^{m} \sum_{r=1}^{q_{ji}} \{(2rk_{ji}-r^2)\sigma_{ij}-r^2\bar{n}_i\bar{n}_j + r^2k_{ji}\bar{n}_j\}\gamma_{jir}(t)$$

$$- \sum_{r=1}^{q_i} \{r^2\bar{n}_i^2 + 2r(2\bar{n}_i\sigma_{ii}+\sigma_{iii})\}\lambda_{ir}(t)$$

$$+ 2 \sum_{\substack{j=1 \\ j\neq i}}^{m} \{ \sum_{r=1}^{q_{ij}} r\gamma_{ij}(t) - \sum_{r=1}^{q_{ji}} r\gamma_{ji}(t)\}(\bar{n}_i\sigma_{ij}+\bar{n}_j\sigma_{ii}+\sigma_{iij})$$

$$i=1,\ldots,m \qquad (3.3.9)$$

$$\frac{d\sigma_{i\ell}}{dt} = \left[\sum_{r=1}^{q_i} rk_{ii}\lambda_{ir}(t) - \mu_i + \sum_{r=1}^{q_\ell} rk_{\ell\ell}\lambda_{\ell r}(t) - \mu_\ell \right] \sigma_{i\ell} +$$

$$+ \sum_{\substack{j=1 \\ j\neq i}}^{m} \{ \sum_{r=1}^{q_{j\ell}} rk_{j\ell}\gamma_{j\ell}(t)\sigma_{ij} + \sum_{r=1}^{q_{ji}} rk_{ji}\gamma_{ji}(t)\sigma_{\ell j} - \sum_{r=1}^{q_{\ell j}} rk_{\ell j}\gamma_{\ell jr}(t)\sigma_{i\ell} -$$

$$- \sum_{r=1}^{q_{ij}} rk_{ij}\gamma_{ij}(t)\sigma_{i\ell} \} +$$

$$+ \sum_{r=1}^{q_{i\ell}} \{ [(r-r^2)k_{i\ell} - r^2\bar{n}_\ell]\gamma_{i\ell r}(t)\bar{n}_i - r^2\gamma_{i\ell r}(t)\sigma_{i\ell} \} +$$

$$+ \sum_{r=1}^{q_{\ell i}} \{ [(r-r^2)k_{\ell i} - r^2\bar{n}_i]\gamma_{\ell ir}(t)\bar{n}_\ell - r^2\gamma_{\ell ir}(t)\sigma_{\ell i} \} -$$

$$- \sum_{r=1}^{q_i} r\lambda_{ir}(t)(\sigma_{ii\ell} + 2\bar{n}_i\sigma_{i\ell}) - \sum_{r=1}^{q_\ell} r\lambda_{\ell r}(t)(\sigma_{\ell\ell i} + 2\bar{n}_\ell\sigma_{\ell i}) +$$

$$+ \sum_{\substack{j=1 \\ j\neq i}}^{m} \left[\sum_{r=1}^{q_{ij}} r\gamma_{ijr}(t) - \sum_{r=1}^{q_{ji}} r\gamma_{jir}(t) \right] \left[\bar{n}_i\sigma_{\ell j} + \bar{n}_j\sigma_{i\ell} + \sigma_{i\ell j} \right] +$$

$$+ \left[\sum_{r=1}^{q_{\ell j}} r\gamma_{\ell jr}(t) - \sum_{r=1}^{q_{i\ell}} r\gamma_{j\ell r}(t) \right] \left[\bar{n}_\ell\sigma_{ij} + \bar{n}_j\sigma_{i\ell} + \sigma_{i\ell j} \right]$$

$$i, = 1, \ldots, m. \qquad (3.3.10)$$

The above equations are very similar to the Lotka-Volterra system {(2.5.12), (2.5.13), (2.5.14)} in that they contain the same type of non-linearities. The following discussion on the above system applies equally well to the Lotka-Volterra system derived in chapter two.

Equations (3.3.8) cannot be directly solved for the means $\bar{n}_i, i=1,\ldots,m$ since these equations are dependent on the variance and covariance terms σ_{ij}. Furthermore, the system of equations (3.3.8), (3.3.9) and (3.3.10) is not closed, since (3.3.9) and (3.3.10) are dependent on the third-order moments $\sigma_{ij\ell}$ of the probability distribution. Suitably differentiating (3.3.7) a number of times with respect to the elements of $\underline{s}$, an ordinary differential equation in any moment of the probability distribution of the migration process can be found, but because the differential equation for any n-th order moment contains (n+1)-th order moments, one cannot obtain a finite closed system of equations to solve. Goel et. al. (1971), discussing a non-linear two-species interaction model, suggests solving the two equations for the means of the populations concerned, by considering the variance and covariance terms in these equations as random driving forces (noise). In making such an approximation little use would, however, have been made of the considerable information available about the behaviour of these second-order moments. A far better approximation would be to solve the system equations for the means, variance and covariance terms, and rather assume the form of the third-order moments $\sigma_{ij\ell}$ for the following reasons:

(a) For reasonably large populations we can expect the probability distribution of the system to sufficiently approximate a joint multivariate normal distribution (for which $\sigma_{ij\ell} = 0$, $i,j,\ell=1,\ldots,m$) so that the assumption

$$\bar{n}_i\sigma_{jk} + \bar{n}_j\sigma_{ik} + \sigma_{ijk} \cong \bar{n}_i\sigma_{jk} + \bar{n}_j\sigma_{ik}$$

can be used.

(b) If we either set $\sigma_{ijk} = 0$ or replace σ_{ijk} by an appropriate noise term, we lose little information on the behaviour of the second-order moments of the system.

Using the techniques developed in this and the previous chapters we now have the machinery to formulate a system of equations that will model the behaviour of a number of interacting populations. We are able to deal with the occurrence of single and multiple instantaneous events, as well as linear and non-linear growth and interaction rates. Hence we are now ready to discuss the application of these models to the development of management strategies for controlling populations, as well as to analyse their properties.

CHAPTER 4

THE LINEAR MULTIVARIATE BIRTH AND DEATH PROCESS POPULATION MODEL

4.1 THE GENERAL LINEAR STOCHASTIC POPULATION MODEL

From the linear birth and death parameters defined in section 2.4A we derived the system of ordinary differential equations {(2.5.9), (2.5.10), (2.5.11)} (which is easily seen to be linear) describing the dynamic behaviour of the means, variance terms and covariance terms.

It is interesting to compare equation (2.5.9), i.e., the equation in the mean, with the general linear deterministic model. In order to write system {(2.5.9), (2.5.10), (2.5.11)} in matrix form we define the following matrices for i,j=1...,m and r=1,...,p.

$$(A_\lambda(t))_{ij} = \lambda_{ij}(t) \qquad (4.1.1.1)$$

$$(A_\mu(t))_{ij} = \mu_{ir}(t) \qquad (4.1.1.2)$$

$$(B_\alpha(t))_{ir} = \alpha_{ir}(t) \qquad (4.1.1.3)$$

$$(B_\beta(t))_{ir} = \beta_{ir}(t) \qquad (4.1.1.4)$$

$$A(t) = A_\lambda(t) - A_\mu(t) \qquad (4.1.1.5)$$

$$B(t) = B_\alpha(t) - B_\beta(t). \qquad (4.1.1.6)$$

Using the above definitions, equations (2.5.9) can be written in matrix form as

$$\frac{d\bar{\underline{n}}}{dt} = A(t)\,\bar{\underline{n}} + B(t)\underline{u}(t). \tag{4.1.2}$$

This equation is identical to the deterministic model

$$\frac{d\underline{n}}{dt} = A(t)\underline{n} + B(t)\underline{u}(t) \tag{4.1.3}$$

except for the interpretation of the elements of the matrices A(t) and B(t). In the deterministic system (4.1.3) these elements are instantaneous growth and interaction rates while in (4.1.2) the elements are parameters measuring the instantaneous probability of certain events occurring.

Further, we see that the solutions to (4.1.2) and (4.1.3) must be identical when solved subject to the same initial conditions, i.e.

$$\underline{n}(0) = \bar{\underline{n}}(0) = \begin{bmatrix} \sum_{\underline{n}\in\Omega} n_1\, P(\underline{n};0) \\ \vdots \\ \sum_{\underline{n}\in\Omega} n_m\, P(\underline{n};0) \end{bmatrix}.$$

This solution cannot, however, be realistically interpreted in the context of the deterministic system since we know from differential equation theory that for A(t), B(t), and $\underline{u}(t)$ continuous in t, the solution $\underline{n}(t)$ will be continuous in t. Thus $\underline{n}(t) \notin \Omega$ for all

$t \in [0,\infty)$. However, $\underline{n}(t)$ is not expected to belong to Ω since $\bar{n}$ is an average measure of a population size which need not itself be an integer to be realistically interpreted.

It is common practice to stochasticize a deterministic system by selecting a disturbance vector $\underset{\sim}{w}(t)$ and suitably coupling it to the system through a matrix G(t) (Meditch 1969) so that (4.1.3) becomes

$$\frac{d\underline{n}}{dt} = A(t)\underline{n} + B(t)\underline{u}(t) + G(t)\underset{\sim}{w}(t). \tag{4.1.4}$$

To solve (4.1.4) for $\underline{n}(t)$, which will now represent a probable trajectory of the stochastic process X(t), we are required to know the statistics of the disturbance vector $\underset{\sim}{w}(t)$ (which is often assumed to be Gaussian) and the matrix G(t). The statistics of the noise for the stochastic linear control system {(2.5.9), (2.5.10), (2.5.11)} can, in contrast to the stochastic linear control system given by equation (4.1.4), be directly solved for, utilizing equations (2.5.10) and (2.5.11) subject to given initial conditions, i.e. the stochasticity is implicitly defined in the system given by equations (2.5.9), (2.5.10) and (2.5.11). Further, if $\underline{\sigma}(t)$ is a vector of elements $\sigma_{ij}(t)$ listed in dictionary order, say, the equations (2.5.9), (2.5.10) and (2.5.11) form an extended linear system

$$\frac{d}{dt}\begin{pmatrix}\underline{n}\\ \underline{\sigma}\end{pmatrix} = A(t)\begin{pmatrix}\underline{n}\\ \underline{\sigma}\end{pmatrix} + B(t)\,\underline{u}(t) \tag{4.1.5}$$

where $A(t)$ and $B(t)$ are appropriately defined in terms of the parameters of the three equations. The actual form of $A(t)$ and $B(t)$ is discussed more fully in the next section.

Using the extended system (4.1.5) it is possible to formulate an optimal control policy that takes the cost of variance and covariance into account, by directly utilizing results of the well-known quadratic linear control problem (Lee and Markus 1967), where $\begin{pmatrix} \bar{n} \\ \Sigma \end{pmatrix}$ will now be the system vector under consideration.

A model of the type (4.1.5) avoids some of the mathematical intricacies involved in analyzing continuous stochastic systems of the type (4.1.4) where, if $\underset{\sim}{w}(t)$ is uncorrelated in time, the rigour of Ito calculus is necessary. It is, however, simpler to work with the discrete form of equation (4.1.4), and well-known results exist for optimally controlling such systems (Meditch 1969).

The comparison of the stochastic model given by equation (4.1.4) with the stochastic model developed in this paper, i.e. equations (2.5.9), (2.5.10) and (2.5.11), applies only to the case where the stochasticity in the system is due to the probabilistic nature of the birth, death and interaction parameters of the species being modelled. The comparison does not cover external noise interfering with the system, as this type of noise has clearly not been accounted for in equations (2.5.9), (2.5.10) and (2.5.11). More explicitly, we arrived at equations (2.5.9), (2.5.10) and (2.5.11) using a generalized stochastic birth and death formulation and the resulting model represented by these three equations generates the time-dependent *means, variance and covariance terms of, in general,* the non-stationary, multivariate probability distribution for this birth and death

process. In contrast the model described by equation (4.1.4) is based on the assumption that the population process has an instantaneous deterministic growth rate so that a random noise component is added to compensate for this simplifying assumption.

4.2 DIRECT DERIVATION AND THEORETICAL VALIDATION OF THE STOCHASTIC MODEL USING KRONECKER PRODUCTS

In this section, using the matrix Kronecker product, we discuss in detail the structure of the matrices $A(t)$ and $B(t)$ given in equation (4.1.5). We also show that the covariance matrix satisfies a linear matrix equation and is indeed positive semi-definite.

The Kronecker product of two general matrices $A \in R^{m\times n}$ and $B \in R^{p\times q}$ is defined (Bellman 1970) as a matrix $C \in R^{mp\times nq}$ denoted by

$$C = A \otimes B$$

whose elements in the $(i-1)p$ to ip rows and $(j-1)q$ to jq columns are $(a_{ij}B)$ $i=1,\ldots,m$ $j=1,\ldots,n$ where a_{ij} is the ij-th element of A, i.e.

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{bmatrix} \qquad (4.2.1.)$$

Let p and q be vectors of elements p_{ij} and q_{ij} respectively listed in dictionary order for $i,j=1,\ldots,n$. Then for a given n-square matrix A we can easily, using (4.2.1), show that the expression

$$p_{ij} = \sum_{k=1}^{n} (a_{ik}q_{jk} + a_{jk}q_{ik}) \qquad i,j=1,\ldots,n \qquad (4.2.2.1)$$

can be written in matrix form as

$$\underline{p} = (A \otimes I + I \otimes A)\underline{q} \qquad (4.2.2.2)$$

where I is the n-square identity matrix.

'Examining system {(2.5.9), (2.5.10), (2.5.11)} we see that (4.2.2.·) can be applied. Using definitions (4.1.1.·) and applying (4.2.2.·) the matrices $A(t)$ and $B(t)$ in (4.1.5) can be given in detail so that (4.1.5) can be written (suppressing the argument t) as

$$\frac{d}{dt}\begin{bmatrix} \underline{\bar{a}} \\ \cdots \\ \underline{g} \end{bmatrix} = \begin{bmatrix} A_\lambda - A_\mu & \vdots & 0 \\ \cdots & \vdots & \cdots \\ (A_\lambda + A_\mu)_1 & \vdots & \\ 0 & \vdots & \\ \vdots & \vdots & \\ 0 & \vdots & \\ (A_\lambda + A_\mu)_i & \vdots & (A_\lambda - A_\mu) \otimes I + I \otimes (A_\lambda - A_\mu) \\ 0 & \vdots & \\ \vdots & \vdots & \\ 0 & \vdots & \\ (A_\lambda + A_\mu)_m & \vdots & \end{bmatrix} \begin{bmatrix} \underline{\bar{a}} \\ \cdots \\ \underline{g} \end{bmatrix} + \begin{bmatrix} B_\alpha - B_\beta \\ \cdots \\ (B_\alpha + B_\beta)_1 \\ 0 \\ \vdots \\ 0 \\ (B_\alpha + B_\beta)_i \\ 0 \\ \vdots \\ 0 \\ (B_\alpha + B_\beta)_m \end{bmatrix} \begin{bmatrix} \underline{u} \end{bmatrix} \qquad (4.2.3)$$

where $(A_\lambda + A_\mu)_i$ and $(B_\alpha + B_\beta)_i$, are the i-th rows of $(A_\lambda + A_\mu)$ and $(B_\alpha + B_\beta)$ respectively and are to be found in the $[(i-1)m+i]$-th row $i=1,\ldots,m$ of the $m^2 \times m$ and $m^2 \times p$ matrices in which they respectively

appear. The remaining rows of these two matrices are zero for each column entry.

It should be noted from symmetry that $\sigma_{ij}(t) = \sigma_{ji}(t)$ $i,j=1,\ldots,m$. The redundant $m(m-1)/2$ equations are, however, included in (4.2.3) to facilitate the use of the Kronecker notation.

Extracting the homogeneous equations in $\underline{\mu}$ and $\underline{\sigma}$ from system (4.2.3) we have using (4.1.1.5) that

$$\frac{d\underline{\mu}}{dt} = A\underline{\mu} \tag{4.2.4.1}$$

and

$$\frac{d\underline{\sigma}}{dt} = A\otimes I + I\otimes A. \tag{4.2.4.2}$$

Examining system (4.2.3) we see that the equations for $\underline{\sigma}$ can be generated directly from the matrices A_λ, A_μ, B_α and B_β defined in the equation for $\bar{\mu}$.

Pollard (1973) has shown that the homogeneous equation in $\underline{\sigma}$ for a discrete multitype Galton-Watson process can be generated from the equation for the mean

$$\bar{\mu}(t+1) = L\underline{\mu}(t) \tag{4.2.5.1}$$

using Kronecker products, and that it will have the form

$$\underline{\sigma}(t+1) = L\otimes L\ \underline{\sigma}(t). \tag{4.2.5.2}$$

On comparing systems (4.2.4.2) and (4.2.5.2) it is interesting to note that they are the discrete and continuous time analogues of each other.

Discretizing (4.2.4.1) we have that

$$\underline{\mu}(t+\Delta t) = (A\Delta t + I)\underline{\mu}(t)$$

whence on comparison with (4.2.5.1) we can identify L and $A\Delta t + I$ so that, when $\Delta t = 1$ we have that $L = A+I$. Further for arbitrary Δt (4.2.5.2) can be written as

$$\underline{g}(t+\Delta t) = (A\Delta t + I) \otimes (A\Delta t + I)\underline{g}(t).$$

Multiplying out the brackets and dividing by Δt, the above equation, after rearranging terms, becomes

$$\frac{\underline{g}(t+\Delta t) - \underline{g}(t)}{\Delta t} = (A \otimes I + I \otimes A)\underline{g}(t) + (A \otimes A)\Delta t\, \underline{g}(t)$$

which on letting $\Delta t \to 0$ is seen to be identical to equation (4.2.4.2).

Hence Pollard's results for the multitype *Galton-Watson process* apply equally well to the general multivariate linear birth and death process. As is pointed out in the next section this result also holds when *linear migration* is included in the process.

An important question to answer concerning the solution $\underline{g}(t)$ in equation (4.2.3) is: does the covariance matrix $S(t)$ of elements $\sigma_{ij}(t)$ $t > t_0$, remain positive semi-definite if $S(t_0)$ is positive semi-definite.

Before answering this question we note that the solution to the equation

$$\frac{d\underset{\sim}{n}}{dt} = A(t)\underset{\sim}{n} + B(t)\underset{\sim}{\mu}(t) \tag{4.2.6}$$

is non-negative for all $t > t_0$. This follows from the assumptions explicitly stated in equations (2.4.1) and (2.4.2) that the birth and death parameters for the i-th population are zero when the i-th population is zero, i.e. system (4.2.6) is not strictly linear since A(t) and B(t) are dependent on $\underset{\sim}{n}$ in the sense that at a boundary point ($n_i = 0$ for some i) of the process modelled by (4.2.6), certain of the elements of A(t) and B(t) are switched to zero.

Examining the equation for $\dot{S}$ in (4.2.3) we see that it can be written in matrix form as

$$\dot{S} = A^T(t)S + SA(t) + K(\underset{\sim}{n},\underset{\sim}{\mu},t) \tag{4.2.7}$$

where $K(\underset{\sim}{n},\underset{\sim}{\mu},t)$ is a diagonal matrix whose diagnol elements are given by the vector

$$\text{diag}[K] = (A_\lambda(t) + A_\mu(t))\underset{\sim}{n} + (B_\alpha(t) + B_\beta(t))\underset{\sim}{\mu}. \tag{4.2.8}$$

Since the matrices $A_\lambda(t)$, $A_\mu(t)$, $B_\alpha(t)$ and $B_\beta(t)$ and the vector $\underset{\sim}{\mu}$ by definition contain non-negative elements only (the birth and death parameters must always be non-negative) diag[K] is a non-negative vector and hence $K(\underset{\sim}{n},\underset{\sim}{\mu},t)$ is positive semi-definite.

It is easy to verify (Brockett 1970) that the solution to (4.2.7) is given by

$$S(t) = \Phi^T(t,t_0)S(t_0)\Phi(t,t_0) + \int_{t_0}^{t} \Phi^T(t,\tau)K(\underline{x},\underline{u},\tau)\Phi(t,\tau)d\tau \qquad (4.2.9)$$

where $\Phi(t,t_0)$ is the transition matrix (principal solution) of the vector equation

$$\frac{d\underline{x}}{dt} = A(t)\underline{x} .$$

$K(\underline{x},\underline{u},t)$ is positive semi-definite, hence from (4.2.9) we see that $S(t)$ is indeed positive semi-definite for all $t > t_0$, whenever $S(t_0)$ is positive semi-definite.

We have thus shown (4.2.3) to be a physically realistic model under the assumption that certain parameters in the matrices $A(t)$ and $B(t)$ will switch to zero at the lower boundary of the population mean size (i.e. $n_i = 0$ for some i).

4.3 A CONTINUOUS TIME STOCHASTIC POPULATION MODEL

Matrix population models have been used extensively to simulate the dynamic behaviour of a population that can be separated into a number of distinct classes.

The first matrix population models were introduced by Lewis (1942) and Leslie (1945 and 1948). These models were designed for populations divided into age classes of equal length. Given m age classes each of unit time length (except possibly for the last age

class), let n_i denote the number of individuals in the i-th age class. The models proposed by Lewis and Leslie have the following form:

$$\underline{n}(t+1) = \begin{bmatrix} f_1 & f_2 & \cdots & f_{m-1} & f_m \\ p_1 & 0 & & & 0 \\ 0 & p_2 & \ddots & & \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & 0 & p_{m-1} & 0 \end{bmatrix} \underline{n}(t) \qquad (4.3.1)$$

where $f_i \geq 0$ $i=1,\ldots,m$ and $0 \leq p_i \leq 1$, $i=1,\ldots,m-1$ respectively represent in a unit time interval for an individual in class i, the expected number of offspring and the probability of survival. Note that the lower right diagonal term is zero which results from the assumption that all members in the oldest age group die in a unit time interval. This assumption is easily modified by making the lower right diagonal term non-zero. This and other modifications of the basic model (4.3.1), e.g. differentiating between sexes, have been introduced in different applications, some of which can be found in Usher's survey of the discrete-time Leslie matrix model (Usher 1972).

Clearly the continuous time analog of (4.3.1) is given by (see section 4.2)

$$\frac{d\underline{n}}{dt} = \begin{bmatrix} f_1-1 & f_2 & \cdots & f_{m-1} & f_m \\ p_1 & -1 & 0 & & 0 \\ 0 & p_2 & \ddots & \ddots & \\ & \ddots & \ddots & 0 & \\ 0 & & 0 & p_{m-1} & -1 \end{bmatrix} \underline{n} \,. \qquad (4.3.2)$$

We can generate the stochastic equivalent of (4.3.2) as described in the previous section (see equation (4.2.3)) where A_λ will be the matrix defined in (4.3.1) and A_μ is the identity matrix.

We will, however, derive from first principles a birth and death process population model which is closely related to, but more general than the stochastic model based on (4.3.2). The major difference will be the diagonal terms. In the model considered hereunder, these terms will not necessarily be -1. This generalization allows classes to be used which are not directly related to age structure, so that the progression of an individual from one class to the next is not rigidly controlled by given time intervals. An example of this is the division of a population into size classes when the growth rates of all the individuals in the population differ slightly. In the discrete model (4.3.1) this type of structure would be characterized by positive diagonal entries.

Consider a population divided into m classes. Let $n_i=1,\ldots,m$ denote the number of individuals in the i-th class. Suppose that

the population can be modelled by the following birth and death process.

(i) An individual in the i-th class may give birth to an individual in the first class in the interval $[t,t+\Delta t]$ with probability $f_i(t)\Delta t + o(\Delta t)$ $i=1,\ldots,m$.

(ii) An individual in the i-th class may die in the interval $[t,t+\Delta t]$ with probability $\lambda_i(t)\Delta t + o(\Delta t)$ $i=1,\ldots,m$.

(iii) An individual in the i-th class may enter the (i+1)-th class (e.g. by virtue of increasing age or size for age-class or size-class structures respectively) in the interval $[t,t+\Delta t]$ with probability $p_i(t)\Delta t + o(\Delta t)$ $i=1,\ldots,m-1$.

(iv) The probability of more than one event occurring in the interval $[t,t+\Delta t]$ is $o(\Delta t)$.

Clearly assumptions (i) - (iv) define a process that is a special case of the general linear multivariate birth, death and migration process. Although in chapter 2 we dealt only with the general linear multivariate birth and death process, the addition of migration parameters poses no additional problems and following the methodology of chapter 2, ordinary differential equations in the means, variance terms and covariance terms can be derived to yield

$$\frac{d\bar{n}}{dt} = (L(t) - D(t))\bar{n} \qquad (4.3.3.1)$$

and

$$\frac{d\underline{g}}{dt} = \begin{bmatrix} [L(t)+D(t)]_1 \\ \vdots \\ [L(t)+D(t)]_m \end{bmatrix} \underline{g} + [(L(t)-D(t))\otimes I + I\otimes(L(t)-D(t))]\underline{g} \qquad (4.3.3.2)$$

where

$$L(t) = \begin{bmatrix} f_1(t) & f_2(t) & \cdots & f_{m-1}(t) & f_m(t) \\ p_1(t) & 0 & & & 0 \\ 0 & p_2(t) & & & \\ & & \ddots & & \\ 0 & & & p_{m-1}(t) & 0 \end{bmatrix} \qquad (4.3.4.1)$$

and

$$D(t) = \begin{bmatrix} \ell_1(t)+p_1(t) & 0 & & & 0 \\ 0 & \ell_2(t)+p_2(t) & & & \\ & & \ddots & & \\ & & & \ell_{m-1}(t)+p_{m-1}(t) & 0 \\ 0 & & & 0 & \ell_m(t) \end{bmatrix} \qquad (4.3.4.2)$$

Comparing systems (4.3.3.') and (4.2.3) and equating A_λ with L and A_μ with D, we see that they are identical modulo the introduction of the control vector $\underline{\mu}$, and hence (4.2.3) is valid when interspecies migration is introduced. (Although we have demonstrated the validity of (4.2.3) only for the particular process considered above,

(4.2.3) is valid in general for linear migration between species.) Further, since the elements of D(t) and L(t) in (4.3.4.') are non-negative and D(t) is diagonal, examining (4.3.3.1) we see that the assumption of zero death rate for a population of size zero is naturally satisfied and hence there is no need to artificially switch the parameters to zero as remarked in the discussion after equation (4.2.6). Thus system (4.3.3.') remains a linear system at the boundary points of the population.

4.4 A BILINEAR SCALAR CONTROL HARVESTING MODEL

The matrix population model has been used to develop harvesting policies for the exploitation of various animal populations. Lefkovitch (1967) investigated the effect of harvesting upon the development structure of a four-class insect model (viz. egg, larva, pupa and adult). Doubleday (1975) used the Leslie matrix model to analyze harvesting policies, particularly taking seasonal variation into account. Rorres and Fair (1975) used an age specific Leslie matrix model of the females in a population to devise a harvesting policy to maximize the yield from a population under conditions in which the population is to return to a fixed initial age structure after each harvest.

In this section we analyze a harvesting model based on system (4.3.3.'). Unlike system (4.2.3), however, control is introduced in a bilinear association with $\bar{n}$. As will be seen, in many harvesting models, this type of control arises more naturally than linear control but the non-linearities introduced into the system make the problem more difficult to solve.

Consider a *population modelled by system* (4.3.3.*). Suppose we wish to exploit or control this population in some optimal fashion by applying a scalar control of intensity u(t). In fish resource management problems, for example, u(t) would be the intensity of the fishing activity while in pest control problems for example, u(t) would be the intensity or rate at which pesticides are applied.

Clearly in this type of problem u(t) will be constrained above by a maximum activity rate $\hat{u}$ say, which may be dependent on time (e.g. the capacity of the fleet of fishing boats may be increased) and will be *naturally constrained below by zero.*

Recalling assumption (i) - (iv) in section 4.3, on which system (4.3.3.*) is based, we make the following additional assumption.

(v) The probability that an individual in the i-th class is removed in the interval $[t,t+\Delta t]$ due to the application of u(t) is $b_i(t)u(t)\Delta t + o(\Delta t)$ $i=1,\ldots,m$.

This assumption then leads to a modified form of system (4.3.3.*) viz.

$$\frac{d\bar{x}}{dt} = (L(t) - D(t) - u(t)B(t))\bar{x}(t) \tag{4.4.1}$$

$$\frac{d\bar{x}}{dt} = \begin{bmatrix} [L(t)+D(t)+u(t)B(t)]_1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \\ & & & \\ & [L(t)+D(t)+u(t)B(t)]_i & & \\ & & 0 & \\ & & & \ddots \\ & & & 0 \\ & & & \\ & & & [L(t)+D(t)+u(t)B(t)]_m \end{bmatrix} \bar{x}(t) + [(L(t)-D(t)-u(t)B(t)) \otimes I +$$

$$+ I \otimes (L(t)-D(t)-u(t)B(t))]g(t) \tag{4.4.2}$$

$$B(t) = \begin{bmatrix} b_1(t) & 0 & \cdots & 0 \\ 0 & b_2(t) & & 0 \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & b_n(t) \end{bmatrix}.$$

If the i-th class, say, does not respond to the application of u then obviously we must have $b_i(t)$ in B(t) equal to zero. This will occur in the fisheries management problem if, for example, the mesh size of the fishing nets is so chosen as to capture fish in the k-th class and above, say, or in an insect control problem if, for example, only the larval life stage is affected by the pesticide being applied.

The remarks at the end of section 4.3 on the non-negativity of the solution to (4.3.3.1) still apply to (4.4.1) since B(t) is diagonal.

Consider the following two problems relating to the control of a population modelled by system {(4.4.1), (4.4.2)}.

A. The deterministic problem

Minimize the integral performance criterion

$$J_d(\bar{n},u,t) = \int_{t_0}^{t_f} [-\bar{x}^T(t)B(t)\bar{n}(t)u(t) + k(t)u(t) + f_1(\bar{n}(t)]\,dt + f_2(\bar{n}(t_f)) \tag{4.4.3}$$

over $u(t) \in [0,\bar{u}]$, subject to equations (4.4.1) where $\bar{u}$ is assumed to be constant.

B. The stochastic problem

Minimize the integral performance criterion

$$J_s(\tilde{n},g,u,t) = \int_{t_0}^{t_f} [-\underline{w}^T(t)B(t)\bar{n}(t)u(t) + k(t)u(t) + f_1(\underline{n}(t)) + f_3(\underline{g}(t))]\,dt + f_2(\tilde{n}(t_f)) + f_4(\underline{g}(t_f)) \quad (4.4.4)$$

over $u(t) \in [0,\bar{u}]$ subject to equations (4.4.1) and (4.4.2).

The terms appearing in J_d and J_s are interpreted as follows.

(i) From (4.4.1) we see that $u(t)B(t)\underline{n}(t)$ is the rate of harvesting the population. The value of the harvest is obtained by weighting the harvested individuals classwise, using a non-negative vector $\underline{w}(t) \in R^m$ whose elements are chosen to reflect the economic value of the corresponding class, and integrating this weighted sum over the harvesting period. Hence the total value of the harvest over the period $[t_0,t_f]$ is

$$\int_{t_0}^{t_f} \underline{w}^T(t)B(t)\bar{n}(t)u(t)dt.$$

Since we wish to maximize this quantity, it must appear with a negative sign in J_d and J_s (since these performance integrals are to be minimized).

(ii) The term $k(t)u(t)$ is associated with the cost of applying the control u. Although we have assumed in this case that the cost of control is linear, in many cases non-

linear functions for the cost of control may be more suitable. The function $k(t)$ is assumed to be time-dependent since the cost of applying $u(t)$ at a given level may vary seasonally or escalate with time.

(iii) The terms $f_1(\bar{n}(t))$ and $f_2(\bar{n}(t_f))$ have been included for generality to reflect inherent costs related to the size of the various population classes. For example: in the control of a pest population $f_1(\bar{n}(t))$ would reflect economic losses due to pest damage; in the exploitation of a fish population $f_2(\bar{n}(t_f))$ could be used to reflect the cost of leaving an overexploited population at the end of the harvesting program (i.e. at $t = t_f$). Suitable functions could be of the form discussed in section 1.10.I (see equation (1.10.4)), provided that their first partial derivatives, with respect to $\bar{n}$, exist.

(iv) In J_d we considered costs associated only with $\bar{n}$ and u. In J_s, in addition to $\bar{n}$ and u, we have included costs associated with $\underline{\sigma}$. The function $f_3(\underline{\sigma}(t))$ will typically penalize the variance terms $\sigma_{ii}(t)$ $i=1,\ldots,n$ when they are non-zero (i.e. positive). Given the variance terms for each class of the population, we are able to estimate a confidence interval around $\bar{n}$ within which the real sizes of the classes are likely to be found. Since the control strategy is based on the mean class size a large class interval could lead to under- or over-exploitation of the population. Hence the reason for associating a cost to non-zero variance (see also section 1.10.II). The func-

tion $f_4(\underline{x}(t_f))$ is used to penalize non-zero variance at the final time.

Using Pontryagin's maximum principle (Lee and Markus 1967, Leitmann 1966, Athans and Falb 1966) we can set up a system of equations that will be satisfied by the controls that minimize J_s and J_d. Since the control u appears linearly in J_s, J_d and system {(4.4.1), (4.4.2)} the optimal control will be bang-bang with possible singular arcs depending on the form of $f_1(\bar{\underline{x}}(t))$ and $f_3(\underline{x}(t))$.

Problem B arises when the cost of the variance is a consideration in the performance criterion, otherwise f_3 and f_4 are identically zero and we have problem A. Consider problem A. The Hamiltonian associated with this problem is

$$H(\bar{\underline{x}}(t),\underline{p}(t),u(t),t) = -\underline{x}^T(t)B(t)\bar{\underline{x}}(t) + k(t)u(t) + f_1(\bar{\underline{x}}(t)) \,,$$
$$+ \underline{p}(t)^T[L(t) - D(t)]\bar{\underline{x}}(t) - \underline{p}(t)^T B(t)\bar{\underline{x}}(t)u(t) \tag{4.4.5}$$

where p(t) satisfies the adjoint equation

$$-\dot{\underline{p}}(t) = [L(t) - D(t)]^T\underline{p}(t) + B(t)^T u(t)[\underline{p}(t) + \underline{x}(t)] - \frac{\partial f_1(\bar{\underline{x}}(t))}{\partial \bar{\underline{x}}}$$

$$p(t_f) = \left.\frac{\partial f_2(\bar{\underline{x}}(t))}{\partial \bar{\underline{x}}(t)}\right|_{t=t_f} . \tag{4.4.6}$$

Pontryagin's maximum principle states the following necessary conditions for a control to be optimal (i.e. for J_d to be minimal).

Let $u^*(t)$ be a piecewise continuous function satisfying the constraint

$$0 \leq u^*(t) \leq \hat{u} \qquad t \in [t_0, t_f]. \tag{4.4.7}$$

Define $\underline{\hat{x}}^*(t)$ as the solution to (4.4.1) with $u(t) = u^*(t)$. Then for $u^*(t)$ to be optimal it is necessary that

(i) a solution $\underline{p}^*(t)$ to (4.4.6) (with $\underline{\hat{x}}(t) = \underline{\hat{x}}^*(t)$ and $u(t) = u^*(t)$) exists.

(ii) $H(\underline{\hat{x}}^*(t), \underline{p}^*(t), u^*(t), t) \leq H(\underline{\hat{x}}^*(t), \underline{p}^*(t), u(t), t)$ for all $u(t)$ satisfying (4.4.7),

(iii) $H(\underline{\hat{x}}^*(t)$ $^*(t), t) = H(\underline{\hat{x}}^*(t_f), \underline{p}^*(t_f), u^*(t_f, t_f) -$

$$- \int_{t_f}^{t} \frac{\partial H}{\partial t}(\underline{\hat{x}}^*(\tau), \underline{p}^*(\tau), u^*(\tau), \tau) d\tau.$$

Using (ii) we can immediately deduce that since

$$\frac{\partial H}{\partial u} = -(\underline{x}(t) + \underline{p}(t))^T B(t) + k(t)$$

the optimal value must be given by

$$u^*(t) = 0 \quad \text{whenever} \quad -(\underline{x}(t) + \underline{p}^*(t))^T B(t) + k(t) > 0 \tag{4.4.8.1}$$

and

$$u^*(t) = \hat{u} \quad \text{whenever} \quad -(\underline{x}(t) + \underline{p}^*(t))^T B(t) + k(t) < 0. \tag{4.4.8.2}$$

If

$$H_u^* = -(\underline{q}(t) + \underline{p}^*(t))^T B(t) + k(t) \qquad (4.4.9)$$

changes sign at $t_1 \in [t_0, t_f]$ say, then $u^*(t)$ can be chosen to be 0 or $\hat{u}$ at $t = t_1$. If this function, however, is zero on an interval $[t_1, t_2] \subset [t_0, t_f]$ where $t_1 < t_2$, then the problem is singular. In this case the form of the control (referred to as a singular subarc) can be derived by examining higher-order conditions (Bell and Jacobson 1975). In many problems singular subarcs play an important part in the control of biological populations (Cliff and Vincent 1973, Goh et. al. 1974).

Clearly the choice of $u^*(t)$ depends on $\underline{p}^*(t)$ which itself depends on $u^*(t)$ and the optimal trajectory $\underline{\bar{x}}^*(t)$. Solving for $\underline{\bar{x}}^*(t)$ requires the substitution of $u(t) = u^*(t)$ in (4.4.1). Further from (4.4.6) we see that $\underline{p}^*(t)$ is solved backwards in time from a boundary condition $\underline{p}^*(t_f)$ dependent on $\underline{\bar{x}}^*(t_f)$. Equations (4.4.1) and (4.4.6) must be solved simultaneously, subject to $u^*(t)$ generated by condition (4.4.8..). This type of problem, usually referred to as a two-point boundary value problem with switching condition, cannot be solved using the techniques applied to solving regular two-point boundary value problems. The presence of the switching function H_u^* (equation 4.4.9) complicates the problem of solving for u^*. A differential dynamic programming type algorithm described in Jacobson and Mayne (chapter 3, 1970) can, however, be applied to solving two-point boundary value problems with a switching condition.

Since Pontryagin's principle provides only necessary conditions for $u^*(t)$ to be optimal, $u^*(t)$ may in fact not minimize the integral J_d. We can, however, check to see whether $u(t)$ satisfies any of the known sufficiency conditions. Typically these arise from the existence of a scalar function that satisfies various conditions (Leitmann 19 3, Lee and Markus 1967) usually related to the existence of a solution to the Hamilton-Jacobi-Bellman partial differential equation

$$-\frac{\partial V(\underline{x},t)}{\partial t} = \min_{u(t)\in[0,\bar{u}]} H(\underline{x}(t), \frac{\partial V(\underline{x},t)}{\partial t}, u(t), t)$$

or conditions arising from the requirement that the second variation of J_d should be positive definite.

Equations corresponding to (4.4.5), (4.4.6), (4.4.8.') and (4.4.9) can, in an entirely analogous manner, be derived for problem B. As in problem A we shall have to solve a two-point boundary value problem with a switching condition, although the system of equations for the state (i.e. $\underline{x},\underline{z}$) and costate variable (i.e. $\underline{p}$) will in general be (n+3)/2 times larger than in problem A.

4.5 A LINEAR-QUADRATIC LESLIE MATRIX HARVESTING MODEL

When the harvesting model is linear the minimization of a quadratic cost performance criterion can easily be solved by direct application of the numerous results available on the linear-quadratic optimal control problem (Anderson and Moore 1971, Brockett 1970). Hence it is desirable, when feasible, to model the harvested popula-

tion by *a linear control system.* Although a linear model is not as physically appealing as the bilinear model (equation (4.4.1)) considered in the previous section it can be valid for populations for from a boundary point (i.e. $n_i = 0$ for some i).

As in the previous section we shall consider a population modelled by system (4.3.3.). In addition, suppose that we can manipulate the i-th class of the population by adding or removing individuals from this class at a rate $u_i(t)$ which will be respectively positive or negative. Further suppose that control is only applied to *classes* i_k $k=1,\ldots,p \leq m$ *and where only* control u_k is applied to class i_k. Then equation (4.3.3.1) can be extended to include control and written as

$$\frac{d\underline{n}}{dt} = (L(t) - D(t))\underline{n} + B\,\underline{u}(t) \qquad (4.5.1)$$

where the i_k-th row of the $m \times p$ matrix B contains a 1 in the k-th column $k=1,\ldots,p$ and *the rest of the elements of B are zero.* If $\underline{u}$ represents the application of a deterministic type of control then (4.4.3.2) remains unmodified. Otherwise (as in system (4.2.3)) additional terms in $\underline{u}$ will have to be included.

$u_k(t)$ will, for example, be deterministic if it reflects a deterministic number of fish, say, removed from the i_k-th class within a unit time interval. $u_1(t)$ may, for example, be stochastic *if it represents adding* fertilized eggs to a fish population, and each egg has a given probability of hatching into a member of the first class.

Examining equation (4.5.1) we see that although (as mentioned in section 4.3) the solution to the homogeneous system will always be positive for positive initial conditions, an element in the solution may become negative through the application of $\underline{u}$. Obviously this situation is physically impossible and the model should either be rejected when the class being controlled is small or at least modified to ensure that the results remain valid. In fact the validity of the model should be examined a posteriori with respect to the solutions it has generated. In the application of this model to the optimal control of populations one can ensure that the population classes remain sufficiently large to ensure the validity of the model. This point is discussed in more detail further on (see discussion under (iii)).

As in the previous section we can set up a deterministic and a stochastic minimization problem. Clearly the deterministic problem is just a special case of the more general stochastic problem so that for generality we consider the problem in its stochastic form. Since the addition of the stochastic elements merely extends the dimension of the problem, however, the deterministic and stochastic problems are equivalent from a notational point of view.

Consider the following cost performance index

$$J = \tfrac{1}{2}\int_{t_0}^{t_f}\Bigg[\sum_{k=1}^{p}\{2w_k(t)u_k(t) + r_k(t)u_k^2(t)\} + \sum_{i=1}^{m}\{q_i(t)[\bar{n}_i(t) - \hat{n}_i(t)]^2 + q_{ii}(t)\sigma_{ii}^2(t)\}\Bigg]dt + \tfrac{1}{2}\sum_{i=1}^{m}\{f_i(n_i(t_f) - \hat{n}_i(t_f))^2 + f_{ii}\sigma_{ii}^2(t_f)\}. \quad (4.5.2)$$

The terms appearing in (4.5.2) can be interpreted as follows:

(i) The elements $w_k(t)$ of the vector $\underline{w}(t) \in R^k$ are all assumed to be non-negative. Hence the first term in J, compactly written as $\underline{x}^T(t)\underline{w}(t)$, is a weighted sum of the rates at which individuals are added to and harvested from the various classes when the corresponding control variable is positive or negative. Since $w_k(t)u_k(t)$ is positive when $u_k(t)$ is positive we shall be minimizing the cost associated with artificially increasing the k-th population. If, however, $u_k(t)$ is negative, then $w_k(t)u_k(t)$ is negative and we shall actually be maximizing the quantity $w_k(t)u_k(t)$ which is associated with the economic value of the yield from harvesting the i_k-th class.

(ii) $\frac{1}{2}r_k(t)u_k^2(t)$ is the term associated with the cost of the application of $u_k(t)$ to the i_k-th population. We see that we have assumed that the cost increases quadratically with the elements of $\underline{u}$. The following four terms in J are also quadratic. This form of costing is applicable when small deviations from zero are proportionately better tolerated than large deviations from zero. In matrix notation the rate at which the cost for all controls is accumulating can be written as $\frac{1}{2}\underline{u}^T R \underline{u}$, where R is the $p \times p$ diagonal matrix of positive elements $r_k(t)$ $k=1,\ldots,p$. In general R need not be diagonal but R is always assumed to be positive definite.

(iii) Consider the term $\frac{1}{2}\sum_{i=1}^{m} q_i(t)(\tilde{n}_i(t) - \hat{n}_i(t))^2$ for all $q_i(t) \geq 0$. Clearly this term penalizes deviations of $\tilde{n}_i(t)$ from $\hat{n}_i(t)$ with small deviations being proportionately more tolerated than large deviations. The vector $\hat{\underline{n}}$ can be interpreted as the ideal popu-

lation size for exploitation. Although this vector may be difficult to obtain, the role of this term is to prevent the population from being under- or over-exploited (since both lead to a loss in economic terms). If, for example the optimal control allows a particular class of the population to go too close to zero hence threatening the viability of the population model, the problem must be reformulated by suitably increasing the elements of $\underline{R}$. If, however, the optimal control predicts a low yield with the population level still remaining high, the yield could possibly be improved by suitably reducing the elements of $\underline{R}$, provided of course that it is not a high cost of control that is preventing a reasonable level of exploitation from taking place.

(iv) The term $\frac{1}{2}\sum_{i=1}^{m} q_{ii}(t)\sigma_{ii}^2(t)$, $q_{ii}(t) > 0$ for all i, penalizes, costwise, non-zero variance, again tolerating proportionately small deviations from zero more readily than large ones.

(v) The final two terms play roles very similar to those discussed in (iii) and (iv) the only difference being that they penalize costwise an under- or over-exploited population and large variance at the end of the harvesting program. The vector $\underline{\bar{n}}(t_f)$ can be interpreted as a population level that is viable for immediate exploitation. This follows since after the harvesting program it is desirable to leave a population that can be exploited again. If this constraint is not needed then obviously $\underline{\bar{n}}(t_f)$ can be made close to zero. As mentioned earlier the linear model is only valid for $\underline{n}(t)$ not close to zero.

Hence $\bar{n}(t_f)$ will always be taken to be larger than the point at which the population comes sufficiently close to zero for the model to break down.

Note: In (ii) we have made the implicit assumption that to remove members at a given rate from a small population is as costly as removing them from a large population. Clearly this is not true and the cost will increase as the population decreases. Hence $r_k(t)$ should in fact depend on $\bar{n}(t)$. However, in some cases this assumption can be made as long as the population remains sufficiently far from a boundary point. This is where the bilinear model considered in the last section is superior to the linear model in this section, since in the bilinear model the rate at which members are removed is proportional to the intensity of the control multiplied by the size of the population.

Using matrix notation, our control problem can be stated as follows.

Minimize the integral performance criterion

$$J = \tfrac{1}{2}\int_{t_0}^{t_f} [2\underline{x}^T(t)\underline{u}(t) + \underline{u}^T(t)R(t)\underline{u}(t) + (\underline{x}(t) - \bar{\underline{x}}(t))^T Q(t)(\underline{x}(t) - \bar{\underline{x}}(t))]\,dt$$

$$+ \tfrac{1}{2}(\underline{x}(t_f) - \bar{\underline{x}}(t_f))^T F(\underline{x}(t_f) - \bar{\underline{x}}(t_f)) \qquad (4.5.3)$$

over all piecewise continuous functions $\underline{u}(t)$ subject to the equation

$$\dot{\underline{x}} = A(t)\underline{x} + B(t)\underline{u}(t) \qquad \underline{x}(t_0) = \underline{x}_0 \qquad (4.5.4)$$

where $x^T(t) = (\underline{x}^T(t), \underline{\xi}^T(t)) \in R^{(m^2+m)}$.

$A(t) \in R^{(m^2+m)\times(m^2+m)}$ and $B(t) \in R^{(m^2+m)\times p}$

are the system matrices generated as described in section 4.2 from $L(t)$, $D(t)$ and the elements of $B(t)$ that correspond to stochastic elements of $\underline{u}(t)$ in equation (4.5.1). $R(t)$ and $\underline{w}(t)$ are as defined in remarks (i) and (ii) above while $Q(t)$ and F are the $(m^2+m)\times(m^2+m)$ diagonal matrices of elements $(q_1(t),\ldots,q_m(t),q_{11}(t),q_{12},\ldots,q_{mm})$ and $(f_1,\ldots,f_m,f_{11},f_{12},\ldots,f_{mm})$ respectively. These matrices are assumed to be positive semi-definite which holds in this case, since the elements $q_i(t)$, $q_{ii}(t)$, f_i and f_{ii} are by definition non-negative.

This problem can be reduced to a standard form as follows (for notational convenience we suppress the argument t in all time-varying functions). Completing the square in (4.5.3) we have

$$J = \tfrac{1}{2}\int_{t_0}^{t_f} [(\underline{u}+R^{-1}\underline{w})R(\underline{u}+R^{-1}\underline{w}) - \underline{w}^T R^{-1}\underline{w} + (\underline{x}-\underline{\xi})^T Q(\underline{x}-\underline{\xi})\ dt$$

$$+ \tfrac{1}{2}(\underline{x}(t_f)-\underline{\xi}(t_f))^T F(\underline{x}(t_f)-\underline{\xi}(t_f)).$$

In addition (4.5.4) can be written as

$$\dot{\underline{x}} = A\underline{x} + B(\underline{u}+R^{-1}\underline{w}) - BR^{-1}\underline{w}.$$

Let

$$\underline{v} = \underline{u} + R^{-1}\underline{y}$$

and

$$\underline{z} = \underline{x} - A^{-1}B R^{-1}\underline{y}$$

where we must now assume $A(t)$ is non-singular on $[t_0, t_f]$. Also let $\hat{\underline{z}} = \hat{\underline{x}} - A^{-1}B R^{-1}\underline{y}$.

Then in terms of the new variable $\underline{z}$ and $\underline{v}$ our problem is stated as follows.

Minimize

$$J = \tfrac{1}{2}\int_{t_0}^{t_f} [\underline{v}^T R \underline{v} + (\underline{z}-\hat{\underline{z}})^T Q (\underline{z}-\hat{\underline{z}})^T]\, dt - \tfrac{1}{2}\int_{t_0}^{t_f} \underline{y}^T R^{-1}\underline{y}\, dt$$

$$+ \tfrac{1}{2}(\underline{z}(t_f) - \hat{\underline{z}}(t_f))^T F(\underline{z}(t_f) - \hat{\underline{z}}(t_f)) \qquad (4.5.5)$$

over the set of all piecewise continuous $\underline{v}$, subject to the equation

$$\dot{\underline{z}} = A\underline{z} + B\underline{v} \qquad \underline{z}(t_0) = \underline{x}_0 - A^{-1}B R^{-1}\underline{y}. \qquad (4.5.6)$$

Except for the term $-\tfrac{1}{2}\int_{t_0}^{t_f} \underline{y}^T R^{-1}\underline{y}\, dt$, the above problem has the form of a class of linear quadratic regulator problems often referred to as the tracking problem (since $\underline{z}(t)$ tracks the desired state $\hat{\underline{z}}(t)$ as closely as possible). This additional term is, however, independent of the choice of control $\underline{v}$ and hence does enter into the minimization problem. We can therefore directly apply results that have been obtained (Athans and Falb 1966) to generate the optimal control that minimizes (4.5.5) subject to (4.5.6).

Clearly the minimum value of J must be greater than or equal to $-\frac{1}{2}\int_{t_0}^{t_f} \underset{\sim}{y}^T R^{-1} \underset{\sim}{y}\, dt$ (since R is positive definite, Q and F are positive semi-definite). Thus the numerical value of this term is the minimum cost we can expect to incur, but since it is actually negative it is the greatest profit we can expect to make. If Q and F are zero we see from (4.5.5) that this profit is achieved by setting $\underset{\sim}{y}(t) = 0$ for $t \in [t_0, t_f]$ i.e., $\underset{\sim}{u}(t) = -R^{-1}\underset{\sim}{y}(t)$.

For non-trivial Q and F the solution to our harvesting problem is (Athans and Falb 1966, section 9.9)

$$\underset{\sim}{u}(t) = -R^{-1}(t)B^T(t)[S(t)\underset{\sim}{x}(t) - \underset{\sim}{g}(t)]$$

i.e.

$$\underset{\sim}{u}(t) = -R^{-1}(t)B^T(t)[S(t)\underset{\sim}{x}(t) - \underset{\sim}{g}(t)] + \\ + R^{-1}(t)[B^T(t)A^{-1}B(t)R^{-1}(t) - I]\underset{\sim}{x} \qquad (4.5.7)$$

where S(t) and $\underset{\sim}{g}(t)$ satisfy

$$-\dot{S}(t) = Q(t) + S(t)A(t) + A^T(t)S(t) - S(t)B(t)R^{-1}(t)B^T(t)S(t)$$

$$S(t_f) = F \qquad (4.5.8)$$

$$\dot{\underset{\sim}{g}}(t) = -Q(t)\underset{\sim}{g}(t) - [A(t) - B(t)R^{-1}(t)B^T(t)S(t)]^T\underset{\sim}{g}(t)$$

$$\underset{\sim}{g}(t_f) = F\underset{\sim}{x}(t_f). \qquad (4.5.9)$$

Equation (4.5.8) is the well-known matrix Riccati equation (Reid 1972, Brockett 1970). Various numerical methods are available for solving it. In most cases standard numerical integration routines will be sufficient, although in certain circumstances (when Q and F are not positive semi-definite) the solution may blow up on the interval $[t_0, t_1]$ (Getz and Jacobson 1975) in which case the optimal controller does not exist. Note that this equation is solved backwards in time.

Once a solution to (4.5.8) has been obtained (4.5.9) is a linear system, which presents no problems, and again is solved backwards in time. The optimal controller is then given in closed-loop *feedback form* (i.e. direct dependence on $x(t)$) but can be obtained in open-loop form (i.e. as a function of time only) by solving (4.5.4) with y substituted from (4.5.7) and then substituting the solution of (4.5.4) into (4.5.7) to obtain a function dependent on the solutions to (4.5.8), (4.5.9) and x_0 only.

Finally it may be noted that the elements of $\hat{x}(t)$ will usually be smaller than those of $x(t)$ on $[t_0, t_1]$. This follows since for an exploitable population the classes of $\bar{n}(t)$ that are going to be harvested must initially by definition of $\hat{n}(t)$, be reasonably larger than those of $\hat{n}(t)$. Also $\hat{x}(t)$ can, in most cases, be set to zero - where this is obviously the most desirable value it can attain (i.e. purely deterministic). This means that a particular $u_r(t)$ will only switch sign if the corresponding class $n_{i_r}(t)$ is over-exploited on $[t_o, t_F]$. This may happen if $u_f(t)$ is cheap or $\bar{n}_{i_r}(t)$ is not sufficiently larger than $\hat{n}_{i_r}(t)$ to prevent over-exploitation from taking place. If $u_r(t)$ does switch from a negative to a positive sign the

constant $w_r(t)$ chosen to weight a harvested unit of the n_{i_r}-th classes in terms of its economic value now gives the same weight to stocking a unit of the n_{i_r}-th class in terms of the costs involved. If the cost of stocking the n_{i_r}-th population with a given number of individuals is equal to the economic value of this number of individuals when harvested, no problem arises. If this is not the case, however, we cannot introduce a change in $w_i(t)$ when $u_i(t)$ changes sign without introducing a non-linearity that violates the linear quadratic formulation.

4.6 NUMERICAL EVALUATION OF A HARVESTING PROGRAM USING LINEAR SCALAR CONTROL

In this section we shall demonstrate the use of the stochastic Leslie matrix model, developed in section 4.3, by numerically evaluating a scalar control that appears linearly (as in section 4.5) in the population model and minimizes a linear integral performance index. The system matrix and cost performance parameters are, however, based on fictitious data.

Suppose that a population $\underline{n}$ is divided into 3 age classes, young (n_1), juvenile (n_2) and adult (n_3) and under normal (uncontrolled) conditions the population (means $\bar{\underline{n}}$) can be modelled by the Leslie matrix model

$$\frac{d\bar{\underline{n}}}{dt} = A(t)\bar{\underline{n}}(t) \qquad \bar{\underline{n}}(t_0) = \underline{n}_0 \tag{4.6.1}$$

where the actual values of the elements of A(t) are given by

$$A(t) = A_1 = \begin{bmatrix} -0.06 & 0.0 & 0.12 \\ 0.010 & -0.014 & 0.0 \\ 0.0 & 0.004 & -0.004 \end{bmatrix} \qquad t \in [0,15) \qquad (4.6.2.1)$$

$$A(t) = A_2 = \begin{bmatrix} -0.06 & 0.0 & 0.0 \\ 0.013 & -0.014 & 0.0 \\ 0.0 & 0.006 & -0.000 \end{bmatrix} \qquad t \in [15,35) \qquad (4.6.2.2)$$

$$A(t) = A_3 = \begin{bmatrix} -0.04 & 0.0 & 0.0 \\ 0.015 & -0.014 & 0.0 \\ 0.0 & 0.008 & -0.004 \end{bmatrix} \qquad \in [35,50) \qquad (4.6.2.3)$$

and

$$A(t+50) = A(t) \quad \text{for all} \quad t. \qquad (4.6.2.4)$$

Note that we have put

$$A(t) = L(t) + D(t) \qquad (4.6.3)$$

where $D(t)$ is the diagonal part of $A(t)$.

In this example we have chosen the unit time interval to be one-fiftieth of a year so that all rate parameters are in terms of so many units per fiftieth of a year. This unit of time has been chosen as it approximates very closely the time interval of one week which is a useful time unit for designing a program spanned over a number of years.

From (4.6.2.4) we see that A(t) is periodic with period one year. The year begins at the beginning of the breeding season which lasts 15-16 weeks (taken as 15 time units). During this period we see from $a_{13} = 0.12$ that the adults give rise to young at the instantaneous rate of 0.12 juveniles per adult per week. For the rest of the year i.e. $t \in [15,50)$ a_{13} is zero and no breeding takes place. The other parameters are similarly interpreted in terms of migration and death rates, where the migration rates are seen to vary slightly over the year and the death rate for class one is higher during the breeding season than during the rest of the year.

Suppose that the harvesting procedure consists of removing adults from the population. Then (4.6.1) can be written as

$$\frac{d\bar{n}}{dt} = A(t)\bar{n}(t) + \underline{b}\,u(t) \qquad \bar{n}(t_0) = n_0 \tag{4.6.4}$$

where $\underline{b}^T = (0,0,-1)$. $u(t)$ is constrained to a maximum rate of 5 000 individuals per week.

Since in this example $u(t)$ is a deterministic quantity, the system of differential equations for the elements of the covariance matrix will be given by system (4.3.3.2). Thus letting $\underline{x}^T(t) = (\bar{n}^T(t)\underline{c}^T(t))$ we have as in (4.5.4) the system

$$\dot{\underline{x}}(t) = A(t)\underline{x}(t) + \underline{b}\,u \tag{4.6.5}$$

where

$$A(t) = \begin{bmatrix} A(t) & \vdots & 0 \\ \cdots\cdots & \vdots & \cdots\cdots \\ [L(t)+D(t)]_1 & \vdots & \\ 0 & \vdots & \\ 0 & \vdots & \\ 0 & \vdots & \\ [L(t)+D(t)]_2 & \vdots & A(t) \oplus I + I \oplus A(t) \\ 0 & \vdots & \\ 0 & \vdots & \\ 0 & \vdots & \\ [L(t)+D(t)]_3 & \vdots & \end{bmatrix}$$

and

$$\underline{b}^T = (0,0,\cdots 1,0,\ldots,0).$$

As commented in the previous section the solution obtained from (4.6.4) through the application of a particular $\underline{u}$ must be examined to ensure that class three of the population does not become negative or even too close to zero to invalidate the assumption th number of fish removed from this class is proportional to te applied.

Suppose we are interested in finding a harvesting policy u(t), $0 \leq u(t) \leq 5\ 000$ over a ten-year interval, i.e. $t \in [0,10]$, such that the following performance criterion is minimized.

$$J = \int_0^{10} [\underline{g}^T \underline{b}\, u(t) + k\, u(t)]\, dt \quad \sum_{i=1}^{3} [f_i \bar{n}_i(10) + g_i \sigma_{ii}(10)]. \qquad (4.6.6)$$

As in the previous two sections the first two terms are the value and cost of the harvest respectively. The expression outside the integral

attaches a cost to the final state of the system, i.e. $\bar{x}(10)$ and $\bar{y}(10)$. The constants f_i will be non-positive as we wish to maximize the quantities $\bar{n}_i(10)$ $i=1,\ldots,3$ to leave as viable a population as possible at the end of the harvesting program. The constants g_i relating to the variance terms $\sigma_{ii}(10)$ are taken as positive since we should like to minimize the variance at the end of the harvesting program.

Since (4.6.6) is linear in $u(t)$, $\bar{x}(10)$ and $\bar{y}(10)$, Pontrygin's maximal principle (see section 4.4) leads to a set of uncoupled state and costate equations, thus reducing the usual two-point boundary value problem to two initial-value problems. (Note that a final value problem is equivalent to an initial value problem solved backwards in time). Using Pontrygin's principle, the 12 costate variables $p^T(t) = (p_1, p_2, p_3, p_{11}, p_{12}, \ldots, p_{33})$ must satisfy the adjoint system of (4.6.5), i.e.

$$\dot{p} = -A^T(t)p \qquad \begin{matrix} p_i(10) = f_i \\ p_{ii}(10) = g_i \\ p_{ij}(10) = 0 \quad j \neq i \end{matrix} \qquad \begin{matrix} \\ i=1,\ldots,3 \\ j=1,\ldots,3 \end{matrix} \tag{4.6.7}$$

Also $u(t)$ must minimize the Hamiltonian

$$H = p^T(t)A(t)x + p(t)^T b\, u(t) + x^T b\, u(t) + k$$

which as in (4.4.8.') implies the following choice for $u(t)$

$$u^*(t) = 0 \qquad \text{whenever} \quad H_u = (\underline{p}(t) + \underline{g})^T \underline{h} + k > 0$$
$$u^*(t) = 5\,000 \quad \text{whenever} \quad H_u = (\underline{p}(t) + \underline{g})^T \underline{h} + k < 0. \tag{4.6.8}$$

In this problem, singular control does not arise so that at any point t' where H_u crosses the zero axis we can choose $u^*(t')$ as we like. To solve for u^* we need the function $p(t)$. The constants $\underline{g}$, $\underline{h}$ and k have been chosen a priori in setting up (4.6.6). We can, however solve, (4.6.7) immediately as it is not coupled to the state equation (4.6.5) and again the constants f_i and g_i $i=1,\ldots,3$ are chosen a priori. This situation strongly contrasts with the general situation described in section 4.4. Thus when applicable a linear harvesting formulation is a computationally desirable model from which to generate the optimal harvesting program. Although a non-linear formulation will invariably be more physically appealing as a population model, the introduction of a single non-linearity into the linear formulation will immediately couple the systems (4.6.7) and (4.6.5), as system (4.6.7) will then depend on the solution to (4.6.5).

Once we have solved (4.6.7) for $\underline{p}(t)$, $u^*(t)$ is generated via (4.6.8) over the interval $[0,10]$ and then the optimal trajectory $x^*(t)$ can be solved for from system (4.6.5).

Since Pontryagin's principal is a necessary condition for optimality we know that $u^*(t)$ is only a candidate control for minimizing (4.6.6). It is clear, however, from the linearity of the problem that $u(t)$ in fact is the optimal control. (Equation (4.6.7) is totally independent of the rest of the problem and possesses a unique solution which from (4.6.8) generates a unique $u^*(t)$).

In the numerical example (whose solution is discussed below) k was chosen as 0.92 and $\underline{g}$ as a vector of ones, so that the integrand of (4.6.6) reduces to 0.08 $\underline{u}$. The interpretation of this is that the cost of the harvesting program is 92 per cent of the value of the harvest.

The f_i and g_i $i=1,\ldots,n$ are difficult to choose a priori. If for a given f_i and g_i, however, the state of the population at the end of the harvest is unsatisfactory, these parameter can be adjusted until the final state of the population is within a suitable region. For example, suppose we impose the restriction that $\tilde{n}_3(10)$ must be greater than $\bar{n}_3(0)/2$, and $\tilde{n}_3(10)$ in fact turns out to be less than half its value at the start of the program. Then the problem can be rerun by decreasing the value of f_3 (i.e. increasing the modulus of its value since it is negative) which will increase the penalty on $\tilde{n}_3(10)$ for being small. This will cause the amount of control to be reduced so that $\bar{n}_3(10)$ is allowed to increase until, for a sufficient decrease in f_3, $\bar{n}_3(10)$ satisfies the required condition.

An optimal control policy was obtained in the following two cases.

(i) $f_1 = -0.01$, $f_2 = -0.04$, $f_3 = -0.08$, $q_i = 0$ $i=1,\ldots,3$

(ii) f_i *as above and* $q_i = 0.002$ $i=1,\ldots,n$

where the units of f_i and q_i are measured as unit cost per unit individual and unit cost per unit variance respectively. Also $f_3 > f_2 > f_1$ since more stress is placed on having a viable adult population that can be re-harvested immediately after the present harvesting program has been completed than on having a population of

young that takes a number of years to reach maturity, so delaying re-harvesting of the population.

Since the model is based on fictitious data a detailed discussion of the results is pointless. The following remarks are, however, made as they relate directly to a comparison between the performance of a deterministic model ($q_i = 0$ $i=1,2,3$) and the performance of the stochastic model ($q_i > 0$ $i=1,2,3$).

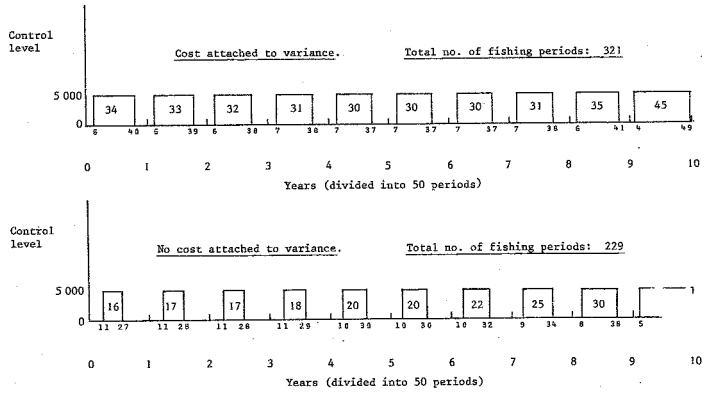
From system (4.6.5) we see that the control u(t) enters into the system only through the equation in $x_3(t) = \bar{n}_3(t)$. Hence the only way of controlling the variance $\sigma_{ii}(t)$ is through the appearance of $\bar{n}_3(t)$ in the equations for $\sigma_{ii}(t)$ $i=1,2,3$, i.e. $x_4(t)$, $x_8(t)$ and $x_{12}(t)$. Further the elements in the lower right-hand partition of A(t) (in (4.6.5)) are all positive so that the only way that the variance can be reduced is by reducing $x_3(t)$, which can be done by applying more control to $x_3(t)$.

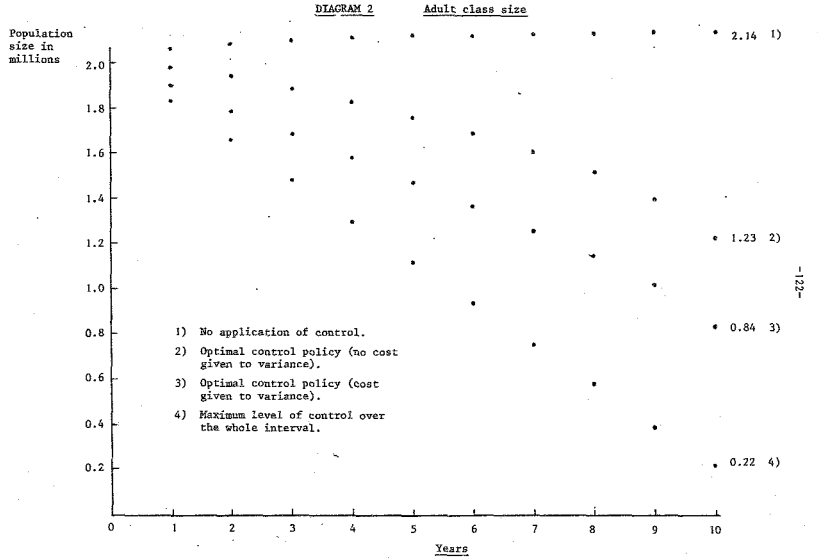
The optimal control policies that result for these two problems are graphed in Diagram 1. In both cases the optimal control policy is to fish for a given number of weeks per year starting at a given week each year with the period of fishing increasing dramatically in the last (i.e. 10-th) year of the harvest. If a cost is attached to non-zero variance, the fishing period, is longer each year than otherwise in an effort to keep the population size, and hence the variance, down).

In Diagram 2 the size of the adult class $n_3(t)$ has been plotted at the end of each yearly interval for the two cases discussed above. The size of the adult class has also been plotted for the two extreme cases: uncontrolled, and controlled at the maximum level, throughout the ten-year period.

DIAGRAM 1

Optimal Controller $u^*(t)$




DIAGRAM 2
Adult class size
Population size in millions
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
1
2
3
4
5
6
7
8
9
10
Years
2.14 1)
1.23 2)
0.84 3)
0.22 4)
1) No application of control.
2) Optimal control policy (no cost given to variance).
3) Optimal control policy (cost given to variance).
4) Maximum level of control over the whole interval.

The initial class sizes for the population were chosen as

$$n_1(0) = 1 \times 10^6 \text{ individuals}$$

$$n_2(0) = 1.5 \times 10^6 \text{ "}$$

$$n_3(0) = 2.0 \times 10^6 \text{ "}$$

and it was assumed at the outset that the population size was completely determined, i.e.

$$\sigma_{ij}(0) = 0 \qquad i,j=1,\ldots,n.$$

The following Table lists the final value for the three classes n_i and the six independent variance and convariance terms σ_{ii} and σ_{ij} $i \neq j$ $i,j=1,\ldots,3$.

	No exploitation 1	Optimal exploitation no variance cost 2	Optimal exploitation variance cost 3	Maximum exploitation 4
$\bar{n}_1$	0.69 millions	0.45 millions	0.32 millions	0.12 millions
$\bar{n}_2$	1.38 "	0.97 "	0.73 "	0.37 "
$\bar{n}_3$	2.14 "	1.23 "	0.84 "	0.22 "
σ_{11}	1.37 "	1.04 "	0.85 "	0.56 "
σ_{22}	4.22 "	3.41 "	2.89 "	2.14 "
σ_{33}	7.30 "	6.22 "	5.47 "	4.45 "
σ_{12}	1.62 "	1.35 "	1.18 "	0.93 "
σ_{13}	2.15 "	1.87 "	1.66 "	1.38 "
σ_{23}	4.44 "	3.83 "	3.40 "	2.81 "

Comparing cases 2 and 3 in the above Table we see that there is only a small improvement in the variance in case 3 over the results for case 2. This, however, is not surprising since the smallest possible variance obtainable is given by case 4, which compared with case 1 shows roughly only a two-fold difference between these values.

As mentioned before, the difficulty of controlling the variance is due to the appearance of the control only in the equation for $\bar{n}_3(t)$. In a model such as (4.4.2) where the control does directly influence the variance equations (since the control is stochastic) the final values of the variance terms will be more sensitive to the control policy applied.

Although this example is based on fictitious data we have demonstrated that the inclusion of the cost of variance in the cost performance index has a marked influence on the choice of the optimal control strategy to be implemented, and hence on the final state of the population.

CHAPTER 5

NON-LINEAR MULTIVARIATE BIRTH AND DEATH PROCESS POPULATION MODELS

5.1 INTRODUCTION

Chapter 4 contains a comprehensive discussion on the use of the linear multivariate birth and death process as a population model and as a tool for developing population control and management policies. The application of non-linear multivariate birth and death processes to population management could similarly be discussed. We are, however, faced with the following dilemma: the more general one attempts to make a discussion on non-linear systems the less one can say about their qualitative properties.

This dilemma was partially resolved for the non-linear modelling studies discussed in chapters 2 and 3 as follows. We formulated the the methodology for *modelling a general class of multivariate* Markov jump processes (see section 2.3), and then presented in depth models of two important but specialized non-linear examples of these processes. For a given problem, however, neither of these models may be applicable, although it may be possible to derive a suitable model by adapting the methods presented in chapters 2 and 3. Clearly these methods can be extended to any multiple-event multivariate Markov jump process whose parameters are polynomial functions of the process variables.

The major problem posed by the non-linear models is that any finite system of differential equations in the moments of the pro-

bability distribution of the process remains open-ended. Recalling the discussion at the end of chapter 3, we see that we were able to close the two systems considered, by assuming their joint multivariate probability distributions to be normal (or at least symmetrically distributed about the population means). This assumption, although intuitively appealing, remains to be validated for general Markov jump processes, or at *least for specific examples.*

As in chapter 4, we can discuss the application of Pontryagin's maximum principle in optimal control studies using these non-linear *models.* This results in setting up systems of non-linear differential equations. In contrast to what was possible in the case of the linear model, however, there is no concise notation to describe these equations and further, from a theoretical point of view, the procedure for solving the problem is the same as that described in section 4.4. Hence *this topic is not pursued* in the context of non-linear models.

Examining the equations derived for the means of the two non-linear *processes* discussed in chapters 2 and 3, we notice that they contain quadratic terms in the population means and are related to the deterministic Lotka-Volterra system of equations often used to model interacting populations. An interesting feature of the deterministic model is its stability chracteristics (Rescigno and Richardson 1973, Getz *1975b). Stability results for* the stochastic analogues of these deterministic models are more difficult to derive. This is due to the appearance of the variance and covariance *terms as forcing* functions in the equations for the means of the populations. In comparing the solutions generated by the stochastic

and deterministic models, some interesting (and even surprising) results do, however, emerge.

5.2 THE STOCHASTIC LOTKA-VOLTERRA AND MULTIPLE MIGRATION POPULATION MODELS

In this and the following sections we discuss the stability of the autonomous form of the models considered and hence all parameter arguments in t are dropped.

The deterministic autonomous Lotka-Volterra model for interacting populations has the following form,

$$\frac{dn_i}{dt} = \varepsilon_i n_i + \sum_{j=i}^{m} a_{ij} n_i n_j \qquad n_i(t_0) = n_{i0} > 0 \qquad i=1,\ldots,m. \quad (5.2.1)$$

In system (2.5.12) if we define

$$\lambda_i - \mu_i = \varepsilon_i \qquad i=1,\ldots,m$$

$$\lambda_{ij} - \mu_{ij} = a_{ij} \qquad i,j=1,\ldots,m$$

we can write this system as

$$\frac{d\bar{n}_i}{dt} = \varepsilon_i \bar{n}_i + \sum_{j=1}^{m} a_{ij}(\bar{n}_i \bar{n}_j + \sigma_{ij}(t)). \qquad i=1,\ldots,m \qquad (5.2.2)$$

The essential difference between systems (5.2.1) and (5.2.2) is the appearance of the terms $\sigma_{ij}(t)$ in (5.2.2), otherwise they are identical. We also note that the linear term in the i-th equation of

(5.2.1) and (5.2.2) involves only $\bar{n}_i$. This is in contrast to the system of equations given in (3.3.8) where the i-th equation is linearly dependent on all the elements of $\underline{\bar{n}}$. Redefining the parameters in (3.3.8), e.g. letting

$$\varepsilon_i = \sum_{r=1}^{q_i} r k_{ii} \lambda_{ir} - \mu_i$$

the system can be rewritten as

$$\frac{d\bar{n}_i}{dt} = \varepsilon_i \bar{n}_i - a_{ii}(\bar{n}_i^2 + \sigma_{ii}(t)) + \sum_{\substack{j=1 \\ j\neq i}}^{m} (b_{ji}\bar{n}_j - b_{ij}\bar{n}_i)$$

$$+ \sum_{\substack{j=1 \\ j\neq i}}^{m} (a_{ij} - a_{ji})(\bar{n}_i\bar{n}_j + \sigma_{ij}(t)) \qquad i=1,\ldots,m \qquad (5.2.3)$$

where we note that a_{ij} and b_{ij} $i,j=1,\ldots,m$ are non-negative. Clearly the deterministic model corresponding to (5.2.3) is

$$\frac{dn_i}{dt} = (\varepsilon_i - a_{ii}n_i)n_i + \sum_{\substack{j=1 \\ j\neq i}}^{m} [(b_{ji} - a_{ji}n_i)n_j - (b_{ij} - a_{ij}n_j)n_i]$$

$$i=1,\ldots,m. \qquad (5.2.4)$$

In system (5.2.4) we see clearly the assumptions underlying the process on which (5.2.3) is based (see section 3.2). Under zero migration between the different colonies the first term of the i-th equation describes pure logistic growth for the i-th colony. The term $(b_{ji} - a_{ji}n_i)n_j$, for example, implies that members of the j-th population migrate at a rate proportional to the size of the j-th

population and enter the i-th population at a rate decreasing linearly with increasing size of the i-th population. The ceiling at which the i-th population will no longer accept members from the j-th population is obviously given by $n_i = b_{ji}/a_{ji}$.

As in the Leslie matrix model, systems (5.2.1) and (5.2.4) are physically appealing since in both cases for $n_{i0} > 0$ i=1,...,n all populations remain positive (i.e. there is no need to switch the parameters to zero as discussed for system (4.2.6)). This follows in (5.2.1) since n_i is a factor of the i-th equation. In (5.2.4) this follows since n_i is a factor of the negative terms of the i-th equation (recall $a_{ij} > 0$, $b_{ij} > 0$ i,j=1,...,m). A similar statement cannot be made about the stochastic systems (5.2.2) and (5.2.3) without analyzing the behaviour of the terms $\sigma_{ij}(t)$ in both cases. We do, however know that the terms $\sigma_{ij}(t)$ come from the covariance matrix S(t) and making use of the fact that S(t) must be positive semi-definite (although we have not proved this as we did in the linear case in section 4.2) some interesting results are deduced in the next two sections of this chapter.

Finally, before going on to discuss qualitatively the stability properties of systems (5.2.1) to (5.2.4) we note that these four systems are all special cases of the general non-autonomous quadratic system of ordinary differential equations

$$\dot{x}_i = \underline{x}^T A_i(t)\underline{x} + \underline{b}_i^T(t)\underline{x} + c_i(t) \qquad i=1,\ldots,m \tag{5.2.5}$$

where $A_i(t)$, $\underline{b}_i(t)$ and $c_i(t)$, i=1,...,m are respectively m-square-matrix, m-vector and scalar functions of time.

The special structure of systems (5.2.1) - (5.2.4) allows us to deduce results about each system that do not apply to (5.2.5). On the other hand, however, results can be derived for (5.2.5), *especially pertaining to the existence of explosive solutions* (discussed in the following section) that can be usefully applied to systems (5.2.1) - (5.2.4).

5.3 THE EXISTENCE OF EXPLOSIVE SOLUTIONS IN QUADRATIC MODELS

It is well-known that certain systems of quadratic differential equations possess solutions that "blow up" in a finite time. The simplest example of this phenomenon is given by the solution to the quadratic scalar equation

$$\frac{dx}{dt} = a x^2 \qquad a > 0 \qquad x(0) = x_0 > 0.$$

The solution to this equation is

$$x(t) = x_0/(1 - a x_0 t) \qquad t \in [0, 1/a x_0).$$

Clearly if $x_0 > 0$ then $x(t) \to \infty$ as $t \to 1/a x_0$. In this case the solution is explosive and is said to have a finite escape time $t_e = 1/a x_0$. Although an explosive solution has no real interpretation for a physical system at t_e, systems with explosive solutions for $t < t_e$ may, over a region of time, model satisfactorily such *phenomena as the outbreak of an epidemic or a population explosion.*

In this section we shall discuss sufficiency conditions for solutions to (5.2.5) to explode, and the application of these results to systems (5.2.1) to (5.2.4). These conditions were obtained by Getz and Jacobson (1975) to which paper we refer the reader for a proof of the following theorem. It should also be mentioned that Frayman (1974) obtained sufficiency conditions similar to those stated below. His results are, however, restricted to systems for which $c_i(t) \equiv 0$ and he does not obtain explicit bounds for the regions in which explosive solutions will occur.

Theorem 5.3.1

Consider the autonomous system

$$\dot{x}_i = \underset{\sim}{x}^T A_i \underset{\sim}{x} + b_i^T \underset{\sim}{x} + c_i \qquad i=1,\ldots,m. \tag{5.3.1}$$

Suppose there exists a vector $\underset{\sim}{w} \in R^m$ of elements w_i such that the matrix $\sum_{i=1}^{m} w_i A_i$ is positive definite. Then the solution to (5.3.1) explodes

(i) for all x_0, whenever $\Delta(\underset{\sim}{w}) < 0$

(ii) for all x_0 satisfying

$$\underset{\sim}{w}^T \underset{\sim}{x}_0 > -\tfrac{1}{2}\underset{\sim}{w}^T\Big(\sum_{i=1}^{n} w_i A_i\Big)^{-1} \sum_{i=1}^{n} w_i b_i + \tfrac{1}{2}\sqrt{\Delta(\underset{\sim}{w})}\; \underset{\sim}{w}^T\underset{\sim}{w}/\lambda$$

where

$$\Delta(\underset{\sim}{w}) = (\lambda/\underset{\sim}{w}^T\underset{\sim}{w})\Big[\sum_{i=1}^{m} w_i b_i^T\Big(\sum_{i=1}^{m} w_i A_i\Big)^{-1} \sum_{i=1}^{m} w_i b_i - 4\sum_{i=1}^{n} w_i c_i\Big]$$

and $\lambda > 0$ is the largest eigenvalue of $\sum_{i=1}^{n} w_i A_i$. □

Note that without loss of generality A_i $i=1,\ldots,m$ can be taken as symmetric. Also explicit upper bounds for t_e are given in Getz and Jacobson (1975) to which paper the reader is referred for further details.

Clearly (5.2.1) can be written in the form of (5.2.5) by defining for $i=1,\ldots,m$

$$A_i = \begin{bmatrix} 0 & \cdots & 0 & \frac{1}{2}a_{i1} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \vdots & 0 & \cdots & 0 \\ \frac{1}{2}a_{i1} & \cdots & & a_{ii} & & \cdots & \frac{1}{2}a_{im} \\ 0 & \cdots & 0 & \vdots & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \frac{1}{2}a_{im} & 0 & \cdots & 0 \end{bmatrix} \qquad (5.3.2.1)$$

$$b_i^T = (0,\ldots,0,\varepsilon_i,0,\ldots,0) \qquad (5.3.2.2)$$

$$c_i = 0. \qquad (5.3.2.3)$$

Let A denote the matrix of elements a_{ij}, and $\underline{\varepsilon}$ the vector of elements ε_i $i,j=1,\ldots,n$ in systems (5.2.1). If A is non-negative then there is a unique solution, $\underline{\xi}$ say, to the linear algebraic equation

$$A\underline{\xi} = -\underline{\varepsilon}$$

i.e.

$$\underline{\xi} = -A^{-1}\underline{\varepsilon}\,. \qquad (5.3.3)$$

Clearly if the elements g_i of $\underline{g}$ are non-negative then $\underline{g}$ is an equilibrium point of (5.2.1) (see discussion preceding theorem 5.4.1).

Choose A_i, b_i and c_i as defined in (5.3.2.') and note that $\sum_{i=1}^{m} A_i$ is the symmetric part of the matrix A associated with system (5.2.1). Then the corollary below follows immediately by choosing $\underline{g}$ as the vector whose elements are all unity.

Corollary 5.3.2

The solution to system (5.2.1) explodes in a finite time if the system matrix A is positive definite and

$$\sum_{i=1}^{m} n_{i0} > -\tfrac{1}{2}\sum_{i=1}^{m} g_i + \tfrac{1}{2}\sqrt{(n/\lambda)\underline{g}^T\underline{g}} \qquad (5.3.4)$$

where λ is the largest eigenvalue of $\sum_{i=1}^{m} A_i$ and $\underline{g}$ is given by equation (5.3.3). □

Using the Cauchy-Schwarz inequality for inner products it can be shown (Getz and Jacobson 1975) that the right-hand side of (5.3.4) is non-zero for all $\underline{g}$.

If (5.2.1) is a model of predator-prey system then under a suitable transformation of the vector $\underline{n}$, the elements of the system matrix A satisfy

$$a_{ij} = -a_{ji} \qquad i \neq j \qquad i,j=1,\ldots,m$$

(Rescigno and Richardson 1973) and the symmetric part of A is the diagonal matrix of elements a_{ii}. Hence from corollary 5.3.2 if $a_{ii} > 0$ $i=1,\ldots,m$ and the transformed initial condition satisfies (5.3.4), the predator-prey model has an explosive solution.

Suppose (5.2.1) satisfies the conditions of corollary (5.3.1), i.e. A is positive definite and the initial condition satisfies (5.3.4). Then under this assumption, a very interesting result emerges in the comparison of systems (5.2.1) and (5.2.2). Before this comparison can be made, however, we must state the following result (Beckenbach and Bellman 1961).

Result 5.3.3

A necessary and sufficient condition that a matrix B be positive definite is that the inequality

$$\text{trace}(AB) > 0 \qquad (5.3.5)$$

holds for all $A > 0$ (positive definite).

In fact it follows by continuity of the matrix $B_\varepsilon = B+\varepsilon I$ in ε, that the strict inequality 5.3.5 can be weakened to hold with possible equality for B positive semi-definite.

Using this result (i.e. when B is positive semi-definite) we prove the following theorem.

Theorem 5.3.4

Suppose that the matrix A of system (5.2.1) is positive definite and B_0 satisfies inequality (5.3.4). Then the solution to system (5.2.2)

with $\bar{n}_0 = n_0$ has a finite escape time and escapes not later than the solution to system (5.2.1).

Proof

Summing the equations in system (5.2.2) we have

$$\sum_{i=1}^{m} \frac{d\bar{n}_i}{dt} = \underline{\varepsilon}^T \bar{\underline{n}} + \bar{\underline{n}}^T A \bar{\underline{n}} + \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}\sigma_{ij}(t).$$

Now the term $\sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}\sigma_{ij}(t)$ is just the trace of the matrix product $A^T S(t)$. Since $S(t)$ is a covariance matrix it is positive semi-definite for all t for which the solution to the system is defined. By assumption, A is positive definite so that trace $A^T S(t)) \geq 0$ for all t for which the solution is defined. Hence

$$\sum_{i=1}^{m} \frac{d\bar{n}_i}{dt} \geq \underline{\varepsilon}^T \bar{\underline{n}} + \bar{\underline{n}}^T A \bar{\underline{n}}. \qquad (5.3.6)$$

From (5.2.1) we see that $\sum_{i=1}^{m} \frac{dn_i}{dt}$ satisfies (5.3.6) with equality. Since $\bar{n}_0 = n_0$ the scalar function $\sum_{i=1}^{n} n_i(t)$ is a lower bound to $\sum_{i=1}^{m} \bar{n}_i(t)$ (Hale 1969, section 1.6). But by corollary 5.3.2 we have that system (5.2.1) escapes in a finite time, i.e. $\sum_{i=1}^{m} n_i(t) \to \infty$ in a finite time. Hence $\sum_{i=1}^{m} \bar{n}_i(t) \to \infty$ not later than $\sum_{i=1}^{m} n_i(t)$: since each $n_i(t)$ is positive, system (5.2.1) escapes when $\sum_{i=1}^{n} n_i(t) \to \infty$. Thus system (5.2.2) will escape not later than system (5.2.1). □

Finally, examining systems (5.2.3) and (5.2.4) we see that the system matrix A of elements α_{ij}, with $\alpha_{ii} = -a_{ii}$ and $\alpha_{ij} = a_{ij} - a_{ji} = -\alpha_{ji}$ is negative semi-definite (since in this case the a_{ij}'s are all non-negative) and hence the *finite escape time results cannot* be applied to this system. This is not surprising since, as discussed in section 5.2, the non-linear multiple migration process has limits built into the growth rates of each colony wherever $-a_{ii}$ is negative.

5.4 A COMPARISON BETWEEN SOLUTIONS TO BE DETERMINISTIC AND STOCHASTIC MODELS

In section 4.1 we saw that the deterministic model of the population size is identical to the stochastic model of the means of the linear multivariate population process. Hence the solution given by the deterministic model is, in the linear case, identical to the estimates of the means as given by the stochastic model. In section 5.2, however, we saw that for non-linear processes the correspondence is not identical since in the stochastic models the variance and covariance terms enter the equations for the means of the population. Thus in the non-linear case, the solutions given by the deterministic models will be different from the estimate of the population means as given by the stochastic models.

An *example of the* non-correspondence of the solutions given by the deterministic and stochastic models is contained in theorem 5.3.4. In this case there is an interesting relationship between the two solutions which arises as a result of the system matrix A being positive definite. In the proof of this theorem we saw that the

solution given by the deterministic system (5.2.1) is a lower bound (with respect to the sum of the elements) to the solution given by the stochastic system (5.3.4). If A is negative definite, however, the reverse is true and the deterministic solution is now a lower bound.

Define a region $D(A,\underline{\varepsilon})$ in R^m as

$$D(A,\underline{\varepsilon}) = \{\underline{n} \in R^m \mid \text{sign}(\underline{n}^T A \underline{n} + \underline{\varepsilon}^T \underline{n}) = \text{sign } \underline{n}^T A \underline{n}\}$$

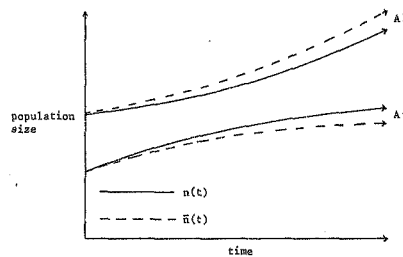
i.e. D is the region in which the quadratic term dominates the linear term. If A is definite then $D(A,\underline{\varepsilon})$ is the whole space R^m minus a region around the origen $\underline{n} = 0$ in which the linear term dominates the quadratic terr If A is positive (negative) definite, then obviously from (5.2.1), $\sum_{i=1}^{m} n_i(t)$ is an increasing (decreasing) function of time for $\underline{n} \in D(A,\underline{\varepsilon})$. In addition if A is positive (negative) definite we have seen that

$$\sum_{i=1}^{m} n_i(t) \leqslant (\geqslant) \sum_{i=1}^{m} \bar{n}_i(t).$$

Thus we can conclude that the solution given by the deterministic model tends to underestimate the rate at which the growth rate is speeding up or slowing down.

In the *scalar case* this phenomenon is illustrated by Diagram 3 below.

DIAGRAM 3. The scalar case



The same conclusion is reached by comparing systems (5.2.3) and (5.2.4). Summing the m equations in systems (5.2.3) and (5.2.4) we obtain

$$\sum_{i=1}^{m} \frac{dn_i}{dt} = \sum_{i=1}^{m} (\varepsilon_i n_i - a_{ii} n_i^2) \qquad (5.4.1.1)$$

and

$$\sum_{i=1}^{m} \frac{d\bar{n}_i}{dt} = \sum_{i=1}^{m} (\varepsilon_i \bar{n}_i - a_{ii} \bar{n}_i^2) - \sum_{i=1}^{m} a_{ii} \sigma_{ii}. \qquad (5.4.1.2)$$

All terms summed over the indeces i and j have dropped out, as these sums are anti-symmetrical. In this case only the variance terms σ_{ii} appear in (5.4.1.2) and are non-negative. Since the a_{ii} are

also non-negative we have that

$$\sum_{i=1}^{m} \frac{d\bar{n}_i}{dt} \leq \sum_{i=1}^{m} \frac{dn_i}{dt} .$$

As mentioned at the end of the last section the system matrix A is negative semi-definite so that we can again conclude that in the region where the quadratic term dominates the linear term the solution given by the deterministic model underestimates the rate at which the population growth rate slows down as the colonies *approach their growth saturation levels.*

These conclusions indicate that there are important limitations in using a deterministic system to estimate the means of a nonlinear population process.

Another important aspect to consider in comparing the solutions given by the deterministic and stochastic models is the relationship between the equilibrium points and their stability properties.

The stability properties of the predator-prey model

$$\frac{dn_i}{dt} = \beta_i n_i (\epsilon_i + \sum_{j=1}^{m} a_{ij} n_j)$$

where $\beta_i > 0$ and $a_{ij} = -a_{ji}$ $i \neq j$ $i,j=1,\ldots,m$, have been extensively *studied since* Volterra first turned his attention to this model in 1926 (Scudo 1971). A recent review of these results is given by Rescigno and Richardson (*1973*).

Some of these results have been generalized to hold for system (5.2.1) where no special relationship between the parameters of the

matrix has been assumed (Getz 1975b).

Examining *equation* (5.2.1) *we see* that $\underline{n} = 0$ is an equilibrium point and solutions of the linear algebraic system $A\underline{n} = -\underline{\varepsilon}$ are equilibrium points. We shall assume that A is non-singular, so that the solution

$$\underline{\bar{n}} = -A^{-1}\underline{\varepsilon} \tag{5.4.2}$$

is unique. Also since n_i is a factor of the i-th equation, i=1,...,m, no solution to (5.2.1) can cross the hyperplanes $n_i = 0$ i=1,...,m. Hence any solution of (5.2.1) *is confined to* the *orthant of* R^m *in which its* initial value is defined. Physically speaking all solutions in the positive orthant are interpretable. Thus choosing $\underline{n}(t_0) > 0$ ensures the physical sense of the solution. Clearly $\underline{n} = 0$ is a critical (equilibrium) point for a solution in the positive orthant but $\underline{\bar{n}}$, given by (5.4.2), is a critical point for a solution in the positive orthant only *if its elements* are non-negative, *i.e. only* if $\underline{\bar{n}} \geq 0$.

Bearing thus in mind we can prove the following theorem.

Theorem 5.4.1

If A is negative definite and the elements of $\underline{\bar{n}} = -A^{-1}\underline{\varepsilon}$ are non-negative then $\underline{\bar{n}}$ is an asymptotically stable *solution of*

$$\frac{dn_i}{dt} = \varepsilon_i n_i + \sum_{j=1}^{m} a_{ij} n_i n_j \qquad n_i(0) = n_{i0} > 0 \qquad i=1,\ldots,m \tag{5.4.3}$$

for all $n_0 > 0$.

Proof

We know that a unique solution $\underline{n}(t;\underline{n}_0)$ exists to (5.4.3) on an interval $[t_0,t_e)$ (where $t_e < \infty$ only if the solution explodes), that its elements $n_i(t)$ are all positive on this interval and that $\underline{g}$ is an equilibrium point of (5.4.3). Hence on $[t_0,t_e)$ we can divide the i-th equation of (5.4.3) by n_i and replace e_i by $-\sum_{j=1}^{m} a_{ij}g_j$ to obtain

$$\frac{1}{n_i(t)}\frac{dn_i(t)}{dt} = \sum_{j=1}^{n} a_{ij}(n_j(t) - g_j) \qquad n_{i0} > 0 \qquad i=1,\ldots,m.$$

Multiplying the i-th equation by $(n_i(t) - g_i)$ and then summing the m equations we obtain the relation

$$\sum_{i=1}^{m}\left(\frac{dn_i(t)}{dt} - g_i\frac{d\ \log\ n_i(t)}{dt}\right) = (\underline{n}(t;\underline{n}_0) - \underline{g})^T A(\underline{n}(t;\underline{n}_0) - \underline{g}).$$

This relation, after integrating over the interval t_0,t for any $t < t_e$ and taking exponentials on both sides of the equation, becomes

$$\prod_{i=1}^{m}\left[e^{n_i(t)}/n_i(t)^{g_i}\right] = \prod_{i=1}^{m}\left[e^{n_{i0}}/n_{i0}^{g_i}\right]\exp\left\{\int_{\tau_0}^{t}(\underline{n}(\tau;\underline{n}_0) - \underline{g})^T A(\underline{n}(\tau;\underline{n}_0) - \underline{g})d\tau\right\}. \qquad (5.4.4)$$

Suppose λ is the largest real eigenvalue of $(A + A^T)/2$. Then λ is real and negative and the inequality

$$x^T A x \leq \lambda \sum_{i=1}^{m} x_i^2 < 0$$

holds for all $x \in R^m$. Also the functions $(e^{n_i(t)}/n_i(t)^{g_i})$ have a positive global minimum at $n_i(t) = g_i$ for $g_i \geq 0$. (Note if $g_i = 0$ then the function is just $e^{n_i(t)}$ which assumes its minimum at $n_i = 0$). Hence the following inequality

$$0 < \prod_{i=1}^{m} [e/g_i]^{g_i} \leq \prod_{i=1}^{m} \left(e^{n_i(t)}/n_i(t)^{g_i}\right)$$

$$\leq \prod_{i=1}^{m} \left(e^{n_{i0}}/n_{i0}^{g_i}\right) \exp\{\lambda \int_{t_0}^{t} \sum_{i=1}^{m} (n_i(\tau;n_0) - g_i)^2 d\tau\} \tag{5.4.5}$$

holds for all $t \in [t_0, t_e)$. Since $\lambda < 0$ and the integrand is non-negative the term

$$\exp\{\lambda \int_{t_0}^{t} \sum_{i=1}^{m} (n_i(\tau;n_0) - g_i)^2 d\tau\} \tag{5.4.6}$$

is a monotonic decreasing function of t, decreasing from unity and bounded below by the positive constant $\prod_{i=1}^{m} [(n_{i0}e)^{g_i}/e^{n_{i0}g_i}]$. Also the functions $(e^{n_i(t)}/n_i(t)^{g_i})$ besides having positive global minima, tend to infinity if and only if $n_i(t)$ tends to zero or infinity. Thus, since the product of these functions is bounded above and below by positive constants independent of t, each $n_i(t)$ is bounded above and away from zero uniformly on $[t_0, t_e)$. Hence $t_e = \infty$ and n(t) is stable on $[t_0, \infty)$.

Since (5.4.6) is a continuous monotonic function bounded below by a positive constant for all $t\in[t_0,t_e)$ we have that

$$\int_{t_0}^{\infty} \sum_{i=1}^{n} (n_i(\tau;\underline{n}_0) - g_i)^2 d\tau$$

is finite. But each term in the integrand is a non-negative continuous function of t. Thus we have for i=1,...,n that

$$n_i(t;\underline{n}_0) \to g_i \qquad \text{as} \qquad t \to \infty,$$

i.e. $\underline{g}$ is asymptotically stable. Since this result holds for all $n_0 > 0$, the theorem is proved. □

The stability of system (5.2.2) unlike that of system (5.2.1), is difficult to analyze owing to its coupling to the highly nonlinear system of equations in the variance and covariance terms. If, however, we assume that the stochastic Lotka-Volterra model has a stationary distribution with mean $\tilde{\underline{g}}$ and covariance matrix P then we can compare $\bar{\underline{g}}$ and $\underline{g}$ as follows.

From (5.2.2) we have that

$$\tilde{g}_i(\varepsilon_i + \sum_{j=1}^{n} a_{ij}\tilde{g}_j) + \sum_{i=1}^{n} a_{ij}p_{ij} = 0 \qquad i=1,\ldots,n$$

which on adding the n equations becomes

$$\tilde{\underline{g}}^T(\underline{\varepsilon} + A\tilde{\underline{g}}) = -\text{ trace } A^T P \leq 0,$$

i.e. $\underline{g}$ satisfies the algebraic inequality

$$\bar{\underline{g}}^T A \bar{\underline{g}} \leqslant -\bar{\underline{g}}^T \underline{g}.$$

Using (5.4.2) to replace $\underline{g}$ we finally obtain the following relationship between $\bar{\underline{g}}$ and $\underline{g}$

$$\bar{\underline{g}}^T A \bar{\underline{g}} < \bar{\underline{g}}^T A \underline{g}$$

or

$$\bar{\underline{g}}^T A (\bar{\underline{g}} - \underline{g}) \leqslant 0. \qquad (5.4.7)$$

A *simple geometric interpretation* of (5.4.7) is only possible in the univariate (scalar) case, whence we can conclude that the stable equilibrium point of the deterministic system is never less than the equilibrium point of the stochastic system.

Thus once more we see that the deterministic model gives a biased solution in estimating the means of non-linear multivariate processes; a limitation that can be overcome by solving the stochastic model formulated directly from the probability transition rates defining the process.

CHAPTER 6

CONCLUSION

In the introduction we remarked that this thesis is concerned with developing techniques for modelling the dynamic behaviour of populations and the use of population models as management tools for formulating optimal exploitation policies.

In the first chapter we developed a stochastic model of a population whose behaviour is governed by the general simple linear birth and death process. We introduced the concept of positive and negative control parameters and provided a control strategy for driving the population mean towards a given target value. A performance criterion for use in optimal control studies was discussed in detail and the novel idea introduced of using the initial variance as a control parameter.

In the second chapter we extended our techniques to modelling populations whose behaviour is governed by single-event multivariate birth and death processes. As discussed in the introduction, very little progress has hitherto been made in analyzing multivariate birth and death processes of any generality. Hence the system of equations in the first and second order moments, as derived in the second chapter, represents an important method of obtaining the essential statistics of populations governed by multivariate birth and death processes. The extension, in chapter 3, of these modelling techniques to multiple-event processes, allows such phenomena as multiple births and the migration of family units to be accounted for.

In the fourth chapter we demonstrated the use of the multivariate linear model for the optimal control of a population structured into classes (age, size, etc.). We saw that these models made it possible to solve a stochastic optimization problem without going beyond deterministic theory. The only concession we had to make was to increase the number of equations in our system by an order of magnitude.

Finally in chapter 5 we showed that in the non-linear case the solution to the deterministic model gives a biased estimate of the population mean size. This indicates that one cannot obtain an accurate stochastic model of the population's behaviour by adding a Guassian noise term with zero mean to the non-linear deterministic model. Rather, a thorough examination should be undertaken of the processes governing the populations, and this will, in fact, lead to the type of stochastic model discussed in this thesis.

We thus propose that systems of ordinary differential equations derived from the forward Kolmogorov equations of stochastic population processes should be used to model stochastic population, in order to avoid the biased estimates of population size as given by stochasticized deterministic models and to facilitate the application of deterministic optimal control theory in management studies of stochastic populations.

In a sense the results contained in this thesis are presented in the form of theoretical "case studies". Each case study was motivated to demonstrate the methodology for dealing with a particular modelling or management problem. We hope that the scope of this thesis is sufficiently comprehensive to cover a large class of stochastic population processes and that the methods described will allow us to model many

real-world systems more accurately than before. Clearly an essential part of future research in this area should be the application of the formulations discussed in this thesis to real-world problems and the subsequent evaluation of their performance and the validity of the results obtained.

R E F E R E N C E S

AKSLAND, N. (1975) A birth, death and migration process with immigration. Adv. Appl. Prob. vol. 7, pp. 44-60.

ANDERSON, B.D.O. & MOORE, J.B. (1971) Linear optimal control. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

ARLEY, N. (1967) On the general birth-and-death-with-immigration stochastic processes. Skandinavisk Aktuarietidskift, vol. 12, pp. 175-182.

ATHANS, M. & FALB, P.L. (1966) Optimal Control. McGraw-Hill, Inc., New York.

BAILEY, N.T.J. (1968) Stochastic birth, death and migration processes for spatially distributed populations. Biometrika, vol. 55, pp. 189-198.

BECKENBACH, E.F. & BELLMAN, R. (1961) Inequalities. Springer-Verlag, Berlin-Göttingen-Heidelberg.

BECKER, N.G. (1970) Control of a pest population. Biometrics, vol. 26, pp. 365-375.

BELL, D.J. & JACOBSON, D.H. (1975) Singular optimal control problems. Academic Press, London-New York-San Francisco.

BELLMAN, R. (1970) Introduction to matrix analyses, second edition. McGraw-Hill, Inc., New York.

BELLMAN, R. & DREYFUS, S.E. (1962) Applied dynamic programming. Princeton University Press, Princeton.

BEYER, J.E. & KEIDING, N. (1975) Remarks on the paper, a stochastic bivariate ecology model for competing species, by C.P. Tsokos and S.W. Hinkley. Math. Biosci., vol. 24, pp. 351-352.

BILLARD, L. (1974) Competition between two species. Stochastic Processes and their Applications, vol. 2, pp. 391-198.

BOSWELL, M.T. & PATIL, G.P. (1972) Certain chance mechanisms related to birth-and-death processes which have classical discrete distribu- tions as special cases. In: P.A.W. LEWIS, ed. Stochastic point processes: statistical analysis, theory and applications. Wiley-Interscience, New York.

BROCKETT, R.W. (1970) Finite dimensional linear systems. John Wiley and Sons, Inc., New York.

CLIFF, E.M. & VINCENT, T.L. (1973) An optimal policy for a fish harvest. JOTA, vol. 12, no. 5, pp. 485-496.

CORMACK, R.M. & NEWTON, J. (1975) A test example of an interactive two-dimensional stochastic process. Math. Biosci., vol. 24, pp. 345-349.

DOUBLEDAY, W.G. (1973) On linear birth-death processes with multiple births. Math. Biosci., vol. 17, pp. 43-56.

DOUBLEDAY, W.G. (1975) Harvesting in matrix populations. Biometrics, vol. 31, pp. 189-200.

DREYFUS, S.E. (1972) The main results of optimal control theory made simple. In: T.N.E. GREVILLE, ed. Population dynamics. Academic Press, New York and London.

DURROCH, J.N. (1958) The multiple-recapture census. I. Estimation of a closed population. Biometrika, vol. 45, pp. 343-359.

DURROCH, J.N. (1959) The multiple-recapture census. II. Estimation when there is immigration or birth. Biometrika, vol. 45, pp. 336-351.

FRAYMAN, M. (1974) Quadratic systems: A study in non-linear systems theory. Ph.D. Thesis, University of Maryland.

GETZ, W.M. (1975a) Optimal control of a birth and death population model. Math. Biosci., vol. 23, pp. 87-111.

GETZ, W.M. (1975b) Some stability results for a system of quadratic differential equations in population dynamics. In: Proceedings of a symposium on differential equations, presented by the National Research Institute for Mathematical Sciences, CSIR Special Report, WISK 161, Pretoria.

GETZ, W.M. (1975c) A non-linear multiple migration process, CSIR Special Report, WISK 182, Pretoria.

GETZ, W.M. (1976) Stochastic equivalents of the linear and Lotka-Volterra systems of equations - a general birth and death process. Math. Biosci. (to appear)

GETZ, W.M. & JACOBSON, D.H. (1975) Sufficiency conditions for the existence of solutions of systems of quadratic differential equations, CSIR Special Report, WISK 169, Pretoria.

GOEL, N.S., MAITRA, S.C. & MONTROL, E.W. (1971) On the Lotka-Volterra and other non-linear models of interacting populations. Revs. Mod. Phy., vol. 43, no. 3, part I, pp. 231-276.

GOEL, N.S. & RICHTER-DYN, N. (1974) Stochastic models in biology. Academic Press, Inc., New York.

GOH, B.S., LEITMANN, G. & VINCENT, T.L. (1974) Optimal control of a predator-prey system. Math. Biosci., vol. 19, pp. 263-286.

GRAY, J.R. (1971) Probability. University Mathematical Texts, Oliver and Boyd, Edinburgh and London.

HADLEY, G. (1964) Non-linear and dynamic programming, Addison Wesley, Inc., New York.

HALE, J. (1969) Ordinary differential equations. Wiley-Interscience, New York.

HIMMELBLAU, D.M. (1972) Applied non-linear programming. McGraw-Hill, Inc., New York.

HINKLEY, W.S. (1971) Stochastic processes in multivariate population dynamics. Ph.D. Thesis, Virginia Polytechnic Institute and State University.

INGLEHART, D.I. (1964) Multivariate competition processes. Annals of Math. Stat., vol. 35, no. 1, pp. 350-361.

IOSIFESCU, M. & TAUTU, P. (1973a) Stochastic processes and applications in biology and medicine I: theory, Springer-Verlag, Berlin-Heidelberg-New York.

IOSIFESCU, M. & TAUTU, P. (1973b) Stochastic processes and applications in biology and medicine II: models. Springer-Verlag, Berlin-Heidelberg-New York.

JACOBSON, D.H. & MAYNE, D.Q. (1970) Differential dynamic programming. American Elsevier Publishing Company, Inc., New York.

JOLLY, G.M. (1965) Explicit estimate for capture-recapture data with both death and immigration stochastic model. Biometrika, vol. 52, pp. 225-247.

KEYFITZ, N. (1968) Introduction to the mathematics of populations. Addison-Wesley, Inc., Menlo Park, California.

LEE, E.B. & MARKUS, L. (1967) Foundations of optimal control theory. John Wiley and Sons Inc., New York.

LEFKOVITCH, L.P. (1967) A theoretical evaluation of population growth after removing individuals from some age groups. Bull. Ent. Res., vol. 57, pp. 437-445.

LEITMANN, G. (1966) An introduction to optimal control. McGraw-Hill, Inc., New York.

LEITMANN, G. (1968) Sufficiency theorems for optimal control. JOTA, vol. 2, no. 5, pp. 285-292.

LESLIE, P.H. (1945) On the use of matrices in certain population mathematics. Biometrika, vol. 33, pp. 183-212.

LESLIE, P.H. (1948) Some further notes on the use of matrices in population mathematics. Biometrika, vol. 35, pp. 213-245.

LEWIS, E.G. (1942) On the generation and growth of a population. Sankya, vol. 6, pp. 93-96.

MEDITCH, J.S. (1969) Stochastic optimal control and linear estimation. McGraw-Hill.

MOHLER, R.R. (1973) Bilinear control processes: with applications to engineering, ecology and medicine. Academic Press, New York and London.

PIAGGIO, H.T.H. (1956) An elementary treatise on differential equations and their applications. Bells Mathematical Series, London.

POLLARD, J.H. (1973) Mathematical models for the growth of human populations. Cambridge University Press.

PRABHU, N.U. (1965) Stochastic processes: basic theory and its applications. MacMillan, New York.

PURI, P.S. (1968) Interconnected birth and death processes. J. Appl. Prob., vol. 5, pp. 334-349.

PURI, P.S. (1975) A linear birth and death process under the influence of another process. J. Appl. Prob., vol. 12, pp. 1-17.

REID, W.T. (1972) Riccati differential equations. Academic Press, Inc., New York and London.

RENSHAW, E. (1972) Birth, death and migration processes. Biometrika, vol. 59, no. 1, pp. 49-60.

RESCIGNO, A. RICHARDSON, I.W. (1973) The deterministic theory of population dynamics. In: R. ROSEN, ed. Foundations of mathematical biology, vol. III; supercellular systems. Academic Press, Inc., New York and London.

RORRES, C. & FAIR, W. (1975) Optimal harvesting policy for an age specific population. Math. Biosci., vol. 24, pp. 31-47.

SANCHO, N.G.F. (1973) Optimal policies and resource management. Math. Biosci., vol. 17, 35-41.

SCUDO, F.M. (1971) Vito Volterra and theoretical ecology. Theor. Pop. Biol., vol. 2, pp. 1-23.

SHOEMAKER, C. (1973) Optimization of agricultural pest management II: formulation of a control model. Math. Biosci., vol. 17, pp. 357-365.

SMITH, T. (1973) Variable population projection matrix models: theory and application to the evaluation of harvesting strategy. Ph.D. Thesis, University of Washington.

TAKACS, L. (1962) Introduction to theory of Queues. Oxford University Press, New York.

TSOKOS, C.P. (1975) Excerpts from a letter by C. Tsokos to D. Perkel. Math. Biosci., vol. 24, pp. 353-354.

TSOKOS, C.P. & HINKLEY, W.S. (1973) A stochastic bivariate model for competing species. Math. Biosci., vol. 26, pp. 191-208.

USHER, M.B. (1972) Developments in the Leslie matrix model. In: J.N.R. JEFFERS, ed. Mathematical models in ecology. Blackwell Scientific publications, Oxford.

WHITTLE, P. (1967) Non-linear migration processes. Proc. 36th Session of the International Statistical Institute, pp. 642-647.

WHITTLE, P. (1968) Equilibrium distributions for an open migration process. J. Appl. Prob., vol. 5, pp. 567-571.

MdeV.

Author Getz Wayne Marcus
Name of thesis Modelling And Control Of Birth And Death Processes. 1976

PUBLISHER:
University of the Witwatersrand, Johannesburg
©2013

LEGAL NOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed, or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page) for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.