Optimal Control of a Birth-and-Death Process Population Model

W. M. GETZ*
Department of Applied Mathematics, University of the Witwatersrand, Johannesburg, Republic of South Africa

Communicated by R. Bellman

ABSTRACT

A birth-and-death process population model 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 to reach the target value 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. INTRODUCTION

The formulation of a birth-and-death type process, as pioneered by Yule, Feller and Kendall among others, has been used successfully to model the behaviour of stochastic populations. Recent examples of this technique applied to finding the probability distribution of the size of the population being modelled are many and varied and include modelling of multiple births [1], competing species [2] and birth-death-and-migration processes [3], as well as models in which differentiation between the sexes is incorporated [4].

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

*Supported by : AE & CI Post-graduate research fellowship.

will vary with time. However, it appears that no attempt has been made to analyze controlling a population in order to steer its distribution towards a desired target at a given time and further minimize a cost performance criterion. It should be noted, though, that there are a number of papers in the ecological field, tackling this problem from an operations research approach [5, 6].

The purpose of this paper is to analyze the control of a birth-and-death process by introducing two parameters. These parameters will be termed positive-control and negative-control, depending on whether they increase or decrease, respectively, the value of the mean of the population distribution from what it would have been had no control been applied.

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 $0,1,2,3,\ldots$ Let $P_n(t)$ denote the probability that the population is of size n 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 relationship:

$$\begin{aligned}P_n(t+\delta t)=&\{(n-1)[\lambda(t)\delta t+o(\delta t)]+\nu(t)\delta t+o(\delta t)\}P_{n-1}(t)\\&+\{1-n[\lambda(t)\delta t+\mu(t)\delta t+o(\delta t)]-[\nu(t)\delta t+\alpha(t)\delta t+o(\delta t)]\}P_n(t)\\&+\{(n+1)[\mu(t)\delta t+o(\delta t)]+\alpha(t)\delta t+o(\delta t)\}P_{n+1}(t)+o(\delta t),\end{aligned}\tag{2.1}$$

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

$$P_n'(t) = [(n+1)\mu(t) + \alpha(t)]P_{n+1}(t) - [n(\mu(t) + \lambda(t)) + \alpha(t) + \nu(t)]P_n(t)$$
$$+ [(n-1)\lambda(t) + \nu(t)]P_{n-1}(t),$$
$$n = 1, 2, 3, \ldots, \tag{2.2a}$$

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

$$P_0'(t) = [\mu(t) + \alpha(t)]P_1(t) - \nu(t)P_0(t). \tag{2.2b}$$

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

$$P_i(0) = \gamma_i, \qquad i = 0, 1, 2, \ldots, \tag{2.3a}$$

$$\sum_{i=0}^{\infty} \gamma_i = 1 \tag{2.3b}$$

(because $P_i(0)$ is a distribution),

$$\text{mean} = n_0, \tag{2.3c}$$

$$\text{variance} = \sigma_0^2, \tag{2.3d}$$

we can solve (2.2) subject to (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 [7], in which only $\alpha(t)$ of the four parameters is zero.

3. GENERAL SOLUTION DISTRIBUTION

Define a probability generating function

$$\phi(S,t) = \sum_{n=0}^{\infty} P_n(t) S^n. \tag{3.1}$$

Then

$$\frac{\partial\phi(S,t)}{\partial t}=\sum_{n=0}^{\infty}P_n'(t)S^n, \tag{3.2}$$

$$\frac{\partial\phi(S,t)}{\partial t}=\sum_{n=0}^{\infty}nP_n(t)S^{n-1}, \tag{3.3}$$

$$\frac{\partial^2\phi(S,t)}{\partial S^2}=\sum_{n=0}^{\infty}n(n-1)P_n(t)S^{n-2}. \tag{3.4}$$

Obviously (3.1) converges for $|S|\leqslant 1$, since $\sum_{n=0}^{\infty}P_n(t)=1$. In addition (3.3) and (3.4) exist for $|S|\leqslant 1$ 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 interested reader can refer to Ref. 9 for a rigorous mathematical account of general queueing processes.

Multiplying Eq. (2.2a) by S^n and summing over all n, we have, using (3.1), (3.2) and (3.3),

$$\frac{\partial\phi}{\partial t}=(\lambda(t)S-\mu(t))(S-1)\frac{\partial\phi}{\partial S}+\left(\nu(t)-\frac{\alpha(t)}{S}\right)(S-1)\phi. \tag{3.5}$$

where for tractability the arguments of $\phi(S,t)$ have been dropped. The method of solution to (3.5) is sketched below for completeness. The full theory can be found in Piaggio [8].

The auxilliary equations corresponding to (3.5) are

$$\frac{dS}{(\lambda(t)S-\mu(t))(S-1)}=\frac{dt}{-1}=\frac{d\phi}{(\alpha(t)/S-\nu(t))(S-1)\phi}. \tag{3.6}$$

We first find a function $U(S,t)=$constant by solving

$$\frac{dS}{dt}=-(S-1)(\lambda(t)S-\mu(t)). \tag{3.7}$$

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

$$\rho(t)=\exp\left\{\int_0^t[\mu(\tau)-\lambda(\tau)]d\tau\right\}, \tag{3.8}$$

so that

$$Z\rho(t)=\int_0^t \lambda(\tau)\rho(\tau)\,d\tau+k, \tag{3.9}$$

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

$$U(S,t)=\frac{\rho(t)}{S-1}-\int_0^t \lambda(\tau)\rho(\tau)d\tau=k \tag{3.10}$$

as required.

Manipulating (3.10), we find that

$$S-1=\frac{\rho(t)}{k+\int_0^t \lambda(\tau)\rho(\tau)\,d\tau}, \tag{3.11}$$

$$S=\frac{k+\rho(t)+\int_0^t \lambda(\tau)\rho(\tau)\,d\tau}{k+\int_0^t \lambda(\tau)\rho(\tau)\,d\tau}. \tag{3.12}$$

We now find a function $V(\phi,t,k)=\text{const}$ by solving

$$\frac{d\phi}{dt}=\left(\nu(t)-\frac{\alpha(t)}{S}\right)(S-1)\phi \tag{3.13}$$

and substituting for S and $S-1$ in (3.13) from (3.11) and (3.12). Since

$$\int\frac{d\phi}{\phi}=\log\phi(S,t)-K,$$

where K is the constant of integration, we have

$$\log\phi(S,t)-\int_0^t\left\{\frac{\nu(\tau)\rho(\tau)}{k+\int_0^\tau \lambda(\omega)\rho(\omega)\,d\omega}-\frac{\alpha(\tau)\rho(\tau)}{k+\rho(\tau)+\int_0^\tau \lambda(\omega)\rho(\omega)\,d\omega}\right\}d\tau=K \tag{3.14a}$$

—i.e., we have found

$$V(\phi,t,k)=K, \tag{3.14b}$$

where k given by (3.10) is replaced after the integration has been performed.

The general solution to (3.5) is given by

$$V(\phi,t,k)=f(U(S,t)), \tag{3.15}$$

where f is an arbitrary function determined from the initial conditions. From (2.3a) and (3.1),

$$\phi(S,0)=\sum_{n=0}^{\infty}\gamma_n S^n,$$

and from (3.14a) and (3.14b) for $t=0$ we have

$$V(\phi,0,k)=\log\phi(S,0)$$

$$=\log\sum_{n=0}^{\infty}\gamma_n S^n. \tag{3.16}$$

From (3.8) and (3.1), since $\rho(0)=0$, we have

$$u(S,0)=\frac{1}{S-1}, \tag{3.17}$$

which leads from Eqs. (3.15) and (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

$$f(\omega)=\log\sum_{n=0}^{\infty}\gamma_n\left(1+\frac{1}{\omega}\right)^n, \tag{3.18}$$

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

So finally, from (3.15) for all t, using (3.10), (3.14) and (3.18), we have that

$$\phi(S,t)=\exp\left\{\int_0^t\left[\frac{\nu(\tau)\rho(\tau)}{k+\int_0^\tau\lambda(\omega)\rho(\omega)\,d\omega}-\frac{\alpha(\tau)\rho(\tau)}{k+\rho(\tau)+\int_0^\tau\lambda(\omega)\rho(\omega)\,d\omega}\right]d\tau\right\}$$

$$\times\left\{\sum_{n=0}^{\infty}\gamma_n\left(1+\frac{1}{k}\right)^n\right\}, \tag{3.19}$$

where k given by (3.10) is replaced after the integration has been performed.

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 (3.8),

$$\rho(t)=e^{(\mu-\lambda)t}, \tag{4.1}$$

and

$$\int_0^t \lambda(\tau)\rho(\tau)\,d\tau=\frac{\lambda}{\mu-\lambda}\left(e^{(\mu-\lambda)t}-1\right). \tag{4.2}$$

From (3.10),

$$k=\left[\frac{1}{S-1}-\frac{\lambda}{\mu-\lambda}\right]e^{(\mu-\lambda)t}+\frac{\lambda}{\mu-\lambda}, \tag{4.3}$$

and

$$k+1=\left[\frac{1}{S-1}-\frac{\lambda}{\mu-\lambda}\right]e^{(\mu-\lambda)t}+\frac{\mu}{\mu-\lambda}. \tag{4.4}$$

Thus (3.19) becomes

$$\phi(S,t)=\exp\left\{\nu\int_0^t \frac{e^{(\mu-\lambda)\tau}}{k+[\lambda/(\mu-\lambda)](e^{(\mu-\lambda)\tau}-1)}\,d\tau\right.$$
$$\left.-\alpha\int_0^t \frac{e^{(\mu-\lambda)\tau}}{k+e^{(\mu-\lambda)\tau}+[\lambda/(\mu-\lambda)](e^{(\mu-\lambda)\tau}-1)}\,d\tau\right\}\left\{\sum_{n=0}^{\infty}\gamma_n\left[\frac{k}{k+1}\right]^n\right\} \tag{4.5}$$

Using the identity

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

we can integrate (4.5), which after simplification and substitution for k

becomes

$$\phi(S,t)=\left[\frac{(\mu-\lambda S)+\lambda(S-1)e^{(\lambda-\mu)t}}{\mu-\lambda}\right]^{-\nu/\lambda}$$

$$\times\left[\frac{(\mu-\lambda S)+\mu(S-1)e^{(\lambda-\mu)t}}{S(\mu-\lambda)}\right]^{\alpha/\mu}$$

$$\times\left\{\sum_{n=0}^{\infty}\gamma_n\left[\frac{\mu-\lambda S+\mu(S-1)e^{(\lambda-\mu)t}}{\mu-\lambda S+\lambda(S-1)e^{(\lambda-\mu)t}}\right]^n\right\}. \tag{4.6}$$

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 (4.6), poses a difficult problem.

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 (3.3) and (3.4) we obtain

$$\frac{\partial\phi}{\partial S}(1,t)=\sum_{n=0}^{\infty}nP_n(t)\triangleq E(n), \tag{5.1}$$

$$\frac{\partial^2\phi}{\partial S^2}(1,t)=\sum_{n=0}^{\infty}n(n-1)P_n(t)\triangleq E[n(n-1)], \tag{5.2}$$

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

$$\sigma^2(t)=E[n(n-1)]+\bar{n}(t)-\bar{n}^2(t). \tag{5.3}$$

To facilitate the algebra involved in differentiating (4.6) once and then twice and setting $S=1$ after each differentiation, we can approach the problem step by step as follows: Let

$$f(S)=(\mu-\lambda S)+\lambda(S-1)e^{(\lambda-\mu)t}, \tag{5.4}$$

$$g(S)=(\mu-\lambda S)+\mu(S-1)e^{(\lambda-\mu)t}. \tag{5.5}$$

Then

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

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

$$f''(S)=f''(1)=0,$$

$$g(1)=\mu-\lambda,$$

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

$$g''(S)=g''(1)=0.$$

Let

$$\psi(S,t)=\frac{f(S)}{\mu-\lambda}, \tag{5.6}$$

$$\kappa(S,t)=\frac{g(S)}{S(\mu-\lambda)}, \tag{5.7}$$

$$r(S,t)=\frac{g(S)}{f(S)}. \tag{5.8}$$

Then

$$\psi(1,t)=1,$$

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

$$\psi''(1,t)=0,$$

$$\kappa(1,t)=1,$$

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

$$\kappa''(1,t)=-\frac{2\mu}{\mu-\lambda}(e^{(\lambda-\mu)t}-1),$$

$$r(1,t)=1,$$

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

$$r''(1,t)=-\frac{2\lambda}{\mu-\lambda}(e^{(\lambda-\mu)t}-1).$$

Rewriting (4.6) in terms of (5.6), (5.7) and (5.8), using (5.4) and (5.5), we have

$$\phi(S,t)=\psi(S,t)^{-\nu/\lambda}\kappa(S,t)^{\alpha/\mu}\left\{\sum_{n=0}^{\infty}\gamma_n r^n(S,t)\right\} \tag{5.9}$$

Recalling Eqs. (2.3 a, b, c and d) and using (5.3) we have that

$$\sum_{n=0}^{\infty} n\gamma_n = n_0, \tag{5.10}$$

$$\sum_{n=0}^{\infty} n(n-1)\gamma_n=\sigma_0^2-n_0+n_0^2. \tag{5.11}$$

We can now proceed to find $\bar{n}(t)$ and $\sigma^2(t)$ by differentiating (5.6) once and then twice and setting $S=1$. Using the preceding analysis we finally obtain

$$\bar{n}(t)=n_0e^{(\lambda-\mu)t}+\frac{\alpha-\nu}{\mu-\lambda}\left(e^{(\lambda-\mu)t}-1\right), \tag{5.12}$$

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

$$+\frac{\lambda+\mu}{\lambda-\mu}n_0e^{(\lambda-\mu)t}\left(e^{(\lambda-\mu)t}-1\right). \tag{5.13}$$

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.

6. DIRECT DERIVATION OF MEAN AND VARIANCE

It is possible to find the mean and variance directly from (3.5) without first solving for the distribution. Differentiating (3.5) with respect to S and (5.1) with respect to t, we have when $S=1$ that

$$\frac{\partial^2\phi}{\partial S\,\partial t}(1,t)=\frac{d\bar{n}}{dt}(t). \tag{6.1}$$

Similarly using (5.2)

$$\frac{\partial^3\phi}{\partial S^2\partial t}(1,t)=\frac{d}{dt}E(n(n-1))$$

$$=\frac{d}{dt}E(n^2)-\frac{d}{dt}\bar{n}(t) \tag{6.2}$$

Differentiating (3.5) with respect to S and setting $S=1$ will lead directly to the set of equations

$$\frac{d\bar{n}}{dt}(t)-[\lambda(t)-\mu(t)]\bar{n}(t)=\nu(t)-\alpha(t), \tag{6.3}$$

$$\frac{d}{dt}E(n^2)-2[\lambda(t)-\mu(t)]E(n^2)=[2\nu(t)-2\alpha(t)+\lambda(t)+\mu(t)]\bar{n}(t)$$
$$+\nu(t)+\alpha(t). \tag{6.4}$$

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

$$\sigma^2(t)=E(n^2)-\bar{n}(t)^2 \tag{6.5}$$

Equations (6.3) and (6.4) are thus subject to the initial conditions

$$\bar{n}(0)=n_0 \tag{6.6a}$$

$$E(n^2)|_{t=0}=\sigma_0^2+n_0^2 \tag{6.6b}$$

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

$$\int_0^t[\lambda(\tau)-\mu(\tau)]\,d\tau=f(t). \tag{6.7}$$

The integrating factor for (6.3) is then $e^{-f(t)}$, so that

$$\bar{n}(t)=e^{f(t)}\left[\int_0^t[\nu(\tau)-\alpha(\tau)]e^{-f(\tau)}\,d\tau+C\right], \tag{6.8}$$

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

$$C=n_0e^{-f(0)}. \tag{6.9}$$

To evaluate (6.8) the form of the controls $\nu(t)$ and $\alpha(t)$ must be given, and 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

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

so

$$\bar{n}(t)=n_0e^{(\lambda-\mu)t}+\frac{\nu-\alpha}{\mu-\lambda}e^{(\lambda-\mu)t}(e^{-(\lambda-\mu)t}-1), \tag{6.10}$$

which is identical to (5.12). Similarly, for constant parameters, solving (6.4) subject to (6.6b) and using the relationship (6.5), we will obtain an expression for $\sigma^2(t)$ which will be identical to (5.13). When only the mean and variance of the distribution are required, it significantly reduces the amount of work involved to use (6.1) and (6.2) directly, rather than solving (3.5) first and then finding the mean and variance afterwards.

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 (5.12)

$$N = n_0 e^{(\lambda-\mu)t_f} + \frac{\alpha-\nu}{\mu-\lambda}\left(e^{(\lambda-\mu)t_f}-1\right). \tag{7.1}$$

Rearranging (7.1), we get the form of the control as

$$\alpha-\nu = \frac{(\mu-\lambda)\left(N-n_0 e^{(\lambda-\mu)t_f}\right)}{e^{(\lambda-\mu)t_f}-1}. \tag{7.2}$$

For all $t_f > 0$ the factor

$$\frac{\mu-\lambda}{e^{(\lambda-\mu)t_f}-1} < 0 \tag{7.3}$$

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

$$N - n_0 e^{(\lambda-\mu)t_f} \tag{7.4}$$

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 (7.3) and (7.2) $\alpha-\nu<0$, and as α and ν are both non-negative we have $\alpha<\nu$. 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, and from (7.2) the level of constant positive control required to achieve the mean N must be

$$\nu=\frac{(\lambda-\mu)(N-n_0e^{(\lambda-\mu)t_f})}{e^{(\lambda-\mu)t_f}-1}. \tag{7.5}$$

(B) If $N<n_0e^{(\lambda-\mu)t_f}$, we need to apply a net negative control, and similarly to A we find that

$$\alpha=\frac{(\mu-\lambda)(N-n_0e^{(\lambda-\mu)t_f})}{e^{(\lambda-\mu)t_f}-1} \tag{7.6}$$

and

$$\nu=0.$$

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

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

Adapting Eq. (7.2), we have

$$\alpha-\nu=\frac{(\mu-\lambda)(N-Ne^{(\lambda-\mu)(t-t_f)})}{e^{(\lambda-\mu)(t-t_f)}-1} \tag{8.1}$$

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

$$\alpha-\nu=(\lambda-\mu)N. \tag{8.2}$$

If $N>Ne^{(\lambda-\mu)(t-t_f)}$, which implies $\lambda<\mu$, then (7.3) becomes

$$\nu=(\mu-\lambda)N. \tag{8.3}$$

If $N<Ne^{(\lambda-\mu)(t-t_f)}$, which implies $\lambda>\mu$, then (7.4) becomes

$$\alpha=(\lambda-\mu)N. \tag{8.4}$$

The results (8.3) and (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 (8.3), or to cancel the net increase in the total population by applying negative control at the same level, as in (8.4).

9. VARIANCE FOR CONSTANT PARAMETERS AND CONTROL

The variance at any time t can be easily found using Eq. (5.13) and the results obtained in Secs. 7 and 8 for the control levels. Since

$$\frac{e^{(\lambda-\mu)t}}{\lambda-\mu}>0 \tag{9.1}$$

for all λ,μ non-negative, all the terms in (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. However when α is non-zero, the greatest value α can have over the interval $[t_0,t_f]$ is such as to make $N=0$; i.e., from (7.6),

$$\alpha_{\max}=n_0e^{(\lambda-\mu)t_f}\left(\frac{\lambda-\mu}{e^{(\lambda-\mu)t_f}-1}\right). \tag{9.2}$$

Considering the term

$$-\alpha\mu\left(\frac{e^{(\lambda-\mu)t_f}-1}{\lambda-\mu}\right)^2+\frac{\lambda+\mu}{\lambda-\mu}n_0e^{(\lambda-\mu)t_f}\left(e^{(\lambda-\mu)t_f}-1\right) \tag{9.3}$$

in $\sigma^2(t_f)$ we find that, on substituting (9.2) in (9.3), we are left with

$$\lambda n_0e^{(\lambda-\mu)t_f}\left(\frac{e^{(\lambda-\mu)t_f}-1}{\lambda-\mu}\right),$$

which by (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 control parameters and no doubt for all feasible controls of the system. Also from (5.13), $\sigma^2(t)$ is not necessarily monotonic, because of the appearance of the factor $e^{(\lambda-\mu)t}-1$. However, if the mean is decreasing (i.e., $\lambda<\mu$ and/or $\nu<\alpha$), the variance will tend to decrease after a certain time, and vice-versa.

The variance at any time t for $N<n_0e^{(\lambda-\mu)t_f}$, using (5.13), (7.6) and (8.4), is as follows:

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

$$\sigma^2(t)=\sigma_0^2e^{2(\lambda-\mu)t}+\frac{\mu}{\lambda-\mu}\,\frac{N-n_0e^{(\lambda-\mu)t_f}}{e^{(\lambda-\mu)t_f}-1}\left(e^{(\lambda-\mu)t}-1\right)^2$$

$$+\frac{\lambda+\mu}{\lambda-\mu}n_0e^{(\lambda-\mu)t}\left(e^{(\lambda-\mu)t}-1\right). \tag{9.4}$$

(ii) $t=t_f$:

$$\sigma^2(t_f)=\sigma_0^2e^{2(\lambda-\mu)t_f}+\frac{\lambda}{\lambda-\mu}n_0e^{(\lambda-\mu)t_f}(e^{(\lambda-\mu)t_f}-1)$$

$$+\frac{\mu}{\lambda-\mu}N(e^{(\lambda-\mu)t_f}-1). \tag{9.5}$$

(iii) $t>t_f$:

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

$$+\frac{\lambda+\mu}{\lambda-\mu}n_0e^{(\lambda-\mu)(t-t_f)}(e^{(\lambda-\mu)(t-t_f)}-1)$$

$$-\frac{\mu}{\lambda-\mu}N(e^{(\lambda-\mu)(t-t_f)}-1)^2. \tag{9.6}$$

Similarly, we can obtain a set of expressions for $\sigma^2(t)$ when $N>n_0e^{(\lambda-\mu)t_f}$ using (5.13), (7.5) and (8.3).

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. However, in a large class of problems 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 Eq. (7.2), which in fact shows that there is effectively only one control, namely $\alpha-\nu$. However, an interesting point to this model is 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 an 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 $N<n_0e^{(\lambda-\mu)t_f}$. As in Sec. 7 (B), we take $\nu=0$. However, in optimizing a cost performance index, 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 (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 (15.12) we have

$$x_1(\alpha)=n_0e^{(\lambda-\mu)t_f}+\alpha\frac{e^{(\lambda-\mu)t_f}-1}{\mu-\lambda}-N \tag{10.1}$$

Let x_2 denote the variance at time t_f; then from (5.13) we have

$$\sigma^2(t_f)\triangleq x_2(\alpha,\sigma_0)=\sigma_0^2e^{2(\lambda-\mu)t_f}-\alpha\mu\left(\frac{e^{(\lambda-\mu)t_f}-1}{\mu-\lambda}\right)^2$$
$$+\frac{\lambda+\mu}{\lambda-\mu}n_0(e^{(\lambda-\mu)t_f}-1)e^{(\lambda-\mu)t_f}. \tag{10.2}$$

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

$$J(\alpha,\sigma_0)=f(x_1(\alpha),x_2(\alpha,\sigma_0),\alpha,\sigma_0). \tag{10.3}$$

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

$$f(x_1,0,0,0)=a_1|x_1(\alpha)|^{k_1} \tag{10.4}$$

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 twice as far 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 (10.2) we have $\sigma(t_f)=x_2(\alpha,\sigma_0)^{1/2}$. Since both x_1 (α) and σ (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

$$f(0,x_2,0,0)=a_2x_2(\alpha,0)^{k_1/2}. \tag{10.5}$$

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.,

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

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

$$a_1x_1(\alpha)^{k_1}=a_1\beta^{k_1}x_2(\alpha,\sigma_0)^{k_1/2}, \tag{10.6}$$

which from (10.5) implies that

$$a_2=a_1\beta^{k_1}. \tag{10.7}$$

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 (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 will be the weighting factor a_1, and the more effect the variance will have in 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

$$f(0,0,\alpha,0)=a_3\alpha^{k_2}, \tag{10.8}$$

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

$$f(0,0,0,\sigma_0)=\frac{a_4}{(\sigma_0{}^{k_3}+a_5)^{k_4}}, \tag{10.9}$$

where $a_4>0$, $a_5\geqslant 0$, $k_3k_4>0$. Since $k_3k_4>0$, (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 $a_5>0$ it will cost $a_4a_5^{-k_4}$.

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 $\tilde{n}$ be a best estimatate of the population size. The variance associated with $\tilde{n}$, denoted var $(\tilde{n})$, has the property that the smaller var $(\tilde{n})$ is the more we know about $\tilde{n}$. Let us define information I on $\tilde{n}$ as

$$I=\frac{1}{\operatorname{var}(\tilde{n})}. \tag{10.10}$$

Clearly, zero information implies infinite variance and vice versa.

Consider a multiple recapture census comprising a sequence of samples $S_1, S_2, \ldots, S_s$, where the members of $S_1,\ldots,S_{s-1}$ are all tagged before being returned to the population, while the members $S_2,\ldots,S_s$ 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 var $(\tilde{n})$, 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

$$q=\exp(-\alpha e) \qquad \alpha>0 \tag{10.11}$$

J. N. Darroch [10] proves that if e is increased to ke, $k>1$, the information I is enlarged to more than k^2I, while if e is held constant and the number of samples is increased from s to $s+1$, the information is increased by more than $(s+1)/(s-1)$, i.e.,

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

and

$$I_{2s}>\frac{(2s)(2s-1)}{(s)(s-1)}I_s$$

$$\approx 4I_s,$$

where I_e denotes the amount of information obtained with effort e for fixed s, and I_s is the amount of information obtained with s samples for fixed e.

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

$$I_{2e}=4I_e \tag{10.12}$$

$$I_{2s}=4I_s \tag{10.13}$$

will be good. This condition will invariably hold, since pn/n will usually be of the order of a fraction of a percent, implying that q will be very close to 1, and from (10.11) this gives αe 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 (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 (10.8) becomes

$$f(0,0,0,\sigma_0)=\frac{a_4}{\sigma_0}. \tag{10.14}$$

Clearly (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 (10.8) can be determined as previously discussed by fitting a curve to these points.

The reader interested in capture-recapture statistics should consult Refs. 11 and 12.

Finally synthesizing (10.3), (10.4), (10.5), (10.7), (10.8) and (10.9), we have

$$J(\alpha,\sigma_0)=a_1|x_1(\alpha)|^{k_1}+a_1\beta^{k_1}x_2(\alpha,\sigma_0)^{k_1/2}$$

$$+a_3\alpha^{k_2}+\frac{a_4}{(\sigma_0{}^{k_3}+a_5)^{k_4}}. \tag{10.15}$$

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

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 (10.15) becomes

$$J(\alpha,\sigma_0) = a_1 x_1(\alpha)^2 + a_1 \beta^2 x_2(\alpha,\sigma_0) + a_3\alpha + \frac{a_4}{\sigma_0}. \tag{11.1}$$

We can minimize (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 (10.1) and (10.2), and letting

$$\frac{e^{(\lambda-\mu)t_f} - 1}{\mu - \lambda} = r(t_f), \tag{11.2}$$

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

$$\frac{\partial J}{\partial \alpha}=2a_1x_1(\alpha)\frac{\partial x_1}{\partial \alpha}+a_1\beta^2\frac{\partial x_2}{\partial \alpha}+a_3$$

$$=2a_1\left[n_0e^{(\lambda-\mu)t_f}-N+\alpha r(t_f)\right]r(t_f)$$

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

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

$$\alpha=\frac{(\mu-\lambda)(N-n_0e^{(\lambda-\mu)t_f})}{(e^{(\lambda-\mu)t_f}-1)}+\frac{\mu\beta^2}{2}$$

$$-\frac{a_3}{2a_1}\left(\frac{\mu-\lambda}{e^{(\lambda-\mu)t_f}-1}\right), \tag{11.3}$$

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

$$=2a_1\beta^2\sigma_0e^{2(\lambda-\mu)t_f}-\frac{a_4}{\sigma_0^2}.$$

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

$$\sigma_0=\left(\frac{a_4}{2a_1\beta^2e^{2(\lambda-\mu)t_f}}\right)^{1/3}. \tag{11.4}$$

Since

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

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

(because a_1,β,a_4,σ_0 are all positive) and

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

it follows that the matrix $\begin{vmatrix}\dfrac{\partial^2 J}{\partial \alpha^2} & \dfrac{\partial^2 J}{\partial\sigma_0\partial\alpha}\\ \dfrac{\partial^2 J}{\partial\alpha\partial\sigma_0} & \dfrac{\partial^2 J}{\partial \sigma_0^2}\end{vmatrix}$

is positive semi-definite, so that the values of α and σ_0 given by (11.3) and (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, (11.3) is identical to (7.6). Thus in (11.3) we have found the optimal level at which to apply constant negative control on the system, and in (11.4) we have found how accurately we must determine our initial mean by indicating the size of the initial variance, so as to minimize the cost performance index (11.1) defined at the final time t_f.

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 ith 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 (5.12) and (5.13) to the new notation we have

$$n_i=n_{i-1}e^{(\lambda_i-\mu_i)\Delta t_i}+\frac{\alpha_i-\nu_i}{\mu_i-\lambda_i}\left(e^{(\lambda_i-\mu_i)\Delta t_i}-1\right), \tag{12.1}$$

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

$$+\frac{\nu_i\lambda_i-\alpha_i\mu_i}{(\mu_i-\lambda_i)^2}\left(e^{(\lambda_i-\mu_i)\Delta t_i}-1\right), \tag{12.2}$$

and thus (10.1) and (10.2) become

$$x_1(n_{i-1},\alpha_i)=n_i-N \tag{12.3}$$

$$x_2(n_{i-1},\sigma_{i-1},\alpha_i)=\sigma_i{}^2 \tag{12.4}$$

The cost of running the system over the ith interval will be, by (10.15),

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

$$+a_3\alpha_i{}^{k_1}+\frac{a_4}{(\sigma_{i-1}{}^{k_3}+a_5)^{k_4}}, \tag{12.5}$$

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

$$F(n_0, \sigma_0, \alpha_1, \alpha_2, \ldots, \alpha_m) = \sum_{i=1}^{m} J(n_{i-1}, \sigma_{i-1}, \alpha_i). \tag{12.6}$$

As before, n_0 is fixed, but we can choose the $m+1$ parameters $\sigma_0, \alpha_1, \ldots, \alpha_m$ optimally so as to minimize (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 $J(n_0, \sigma_0, \alpha_1)$ will not in general minimize F, since α_i and σ_{i-1} are dependent on $\sigma_0, \alpha_1, \ldots, \alpha_{i-1}$.

There are two important computational procedures to find a

$$u \triangleq (\sigma_0, \alpha_1, \alpha_2, \ldots, \alpha_m) \tag{12.7}$$

that will minimize (12.6).

I. NON-LINEAR PROGRAMMING

There are a number of algorithms, including the various gradient methods [13, 14], that are designed to calculate

$$u^* = (\sigma_0^*, \alpha_1^*, \ldots, \alpha_m^*)$$

such that

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

for all admissible u defined by (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

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

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

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

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

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

$$k = m-1, m-2, \ldots, 1, \tag{12.8}$$

forms a recurrence relation to solve for

$$F_0(n_0) = \min_{\sigma_0} F_1(n_0, \sigma_0)$$

$$= \min_u F(n_0, u), \tag{12.9}$$

as required with the given initial condition

$$F_m(n_{m-1}, \sigma_{m-1}) = \min_{\alpha_m} J(n_{m-1}, \sigma_{m-1}, \alpha_m). \tag{12.10}$$

The problem of solving for (12.9) is complicated by the fact that (12.1), (12.2) and (12.8) are coupled and (12.1) and (12.2) are solved forwards in time while (12.8) is solved backwards in time.

Some optimization problems lend themselves readily to a dynamic programming approach, when 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 [15]. It is not possible to obtain a reproducing form for F_k in the above problem, as F_m contains terms in σ_{m-1}^2 and $1/\sigma_{m-1}$; F_{m-1} contains terms in σ_{m-2}^4, $1/\sigma_{m-2}^2$, $\sigma_{m-2}^2\, n_{m-2}$ and $\sigma_{m-2}^2 \alpha_{m-1}$; and F_{m-2} will contain even more nonlinear and cross product terms.

There are a number of general texts dealing with dynamic programming techniques and related problems [14, 16, 17, 18], so that it is not worth while deriving further general results for the above problem.

13. CONCLUSION

Results that are obtained using a general systems theory approach are usually applicable to a varied range of problems. The analysis in this paper applies generally to controlling a stochastic population. However a major region of application is in resource management. This includes problems encountered in the fishing industry, agriculture and wildlife management. In particular, in Africa, the control of large herbivorous mammal herds is important, as they can destroy their habitat through over-grazing.

There is definitely a need to use a more scientific approach in handling these problems in the future, and mathematics will have an important role to play. However, many problems would require a tremendous amount of fieldwork and data collection before satisfactory values could be given to the various parameters and cost functions of the system. In addition these factors may alter as the system is being controlled, and in fact it may be necessary to have some form of feedback process linked to the system if the desired goals are to be achieved.

The contribution of this paper is that it provides a general yet reasonably realistic stochastic population model together with a control strategy which optimizes a generally sensible, closely reasoned performance criterion and includes the novel idea of making the initial variance a control variable. It is felt that the results provide practically useful strategies for the control of certain types of populations.

REFERENCES

1 W. G. Doubleday, On linear birth-death processes with multiple births, *Math. Biosci.* 17, 43–56 (1973).
2 C. P. Tsokos and W. S. Hinkley, A stochastic bivariate model for competing species, *Math. Biosci.* 16, 191–208 (1973).
3 E. Renshaw, Birth, death and migration processes, *Biometrika* 59 (1), 49–60 (1972).
4 N. Keyfitz, *Introduction to the Mathematics of Population*, Addison Wesley, New York, 1968, Chapter 16.
5 N. G. F. Sancho, Optimal policies in resource management, *Math. Biosci.* 17, 35–41 (1973).
6 C. Shoemaker, Optimization of agricultural pest management II: formulation of a control model, *Math. Biosci.* 17, 357–365 (1973).
7 N. Arley, On the general birth-and-death-with-immigration stochastic process, *Skandinavisk Aktuarietidskrift* 12, 175–182 (1967).
8 H. T. H. Piaggio, *An Elementary Treatise on Differential Equations and their Applications*, Bells Mathematical Series, London (1956).
9 L. Takàcs, *Introduction to the theory of Queues*, Oxford University Press, New York, 1962.
10 J. N. Darroch The multiple-recapture census. I. estimation of a closed population, *Biometrika* 45, 356 (1958).
11 J. N. Darroch The multiple-recapture census. II. estimation when there is immigration or birth, *Biometrika* 46, 336–351 (1959).
12 G. M. Jolly Explicit estimate from capture-recapture data with both death and immigration stochastic model, *Biometrika* 52, 225–247 (1965).
13 D. M. Himmelblau *Applied Nonlinear Programming*, McGraw-Hill, New York, 1972.
14 G. Hadley *Nonlinear and Dynamic Programming*, Addison Wesley, New York, 1964.
15 J. S. Meditch, *Stochastic Optimal Linear Estimation and Control*, McGraw-Hill, New York, 1964.
16 G. L. Nemhauser, *Introduction to Dynamic Programming*, Wiley, New York, 1966.
17 R. E. Bellman and S. E. Dreyfus, *Applied Dynamic Programming*, Princeton Univ. Press, Princeton, 1962.
18 D. H. Jacobson and D. Q. Mayne, *Differential Dynamic Programming*, Elsevier, New York, 1970, Chapter 4.