JOURNAL OF REGIONAL SCIENCE, VOL. 3, NO. 2. 1961

ON THE CONSTRUCTION OF STOCHASTIC MODELS OF POPULATION GROWTH AND MIGRATION

by John S. de Cani*

1. INTRODUCTION

This paper has two purposes: first, to present three stochastic models of population growth and migration; and second, to show the mathematical development of two of these models in some detail in order to demonstrate that useful results can be obtained with relatively simple mathematics.

The processes of growth and migration can be viewed as stochastic processes because, in a very short time interval, there are finite probabilities that an individual in a given population will (a) remain unchanged, (b) become two individuals (i. e., give birth to a new individual), (c) die, or (d) migrate to a different area. As far as a given area is concerned, we can view migration as equivalent to birth or death. (This point of view is not provincial; it is merely mathematically convenient.) If we choose an extremely short time interval, the population will remain unchanged, increase by one individual (a birth or an immigration), or decrease by one individual (a death or an emigration), and we can attach probabilities to these events. The process of choosing an extremely short time interval is a limit process, and implicit in the whole argument is the system of difference-differential equations known as the Kolmogorov differential equations. We shall be primarily concerned with various ways of either solving such systems of differential equations or of avoiding the problem of solving the system when it looks too complicated.

In this paper we shall consider three models of increasing complexity. The first is a pure migration model. Consider a system consisting of two regions and containing a finite number, N, of individuals. This population cannot increase or decrease in size; the individuals in the system can only migrate from one region to the other. The problem is: given the initial distribution of the N individuals between the two regions, to express the probability distributions of the populations of the two regions as functions of time. We also want to examine the limiting distributions, i. e., the probability distributions after an infinite time has elapsed, if such distributions exist. In the first model these distributions turn out to be binomial and independent of the initial distribution of the N individuals between the two regions.

In the second model we again consider a system made up of two regions, but the individuals within the regions are give considerable freedom of choice; they can (a) remain where they are, (b) die or emigrate from the system, (c) give birth, or (d) move from one region to the other within the system. We also permit immigration into the system. As in the first model, the basic

* The author is Assistant Professor of Statistics in the University of Pennsylvania.

problem is to express the probability distributions of the regional populations as functions of time and to examine limiting distributions if they exist. In this model, however, we are confronted with a more difficult system of difference-differential equations than in the first model, and we content ourselves with finding the means and variances and the covariance of the joint distribution of the two populations. While the populations grow exponentially so that limiting distributions do not exist, the influence of the initial distribution of the population between the two regions decreases with time.

The third model is a form of "predator-prey" model. Here we consider a region which contains two groups of individuals, one of which, the "prey" population, is alienated by the other, the "predator" population. The larger the total population, the more likely are predators to be attracted to the region, while the larger the predator population, the more likely are members of the prey population to migrate from the region. Only the means of this model are expressed as functions of time, while implicit expressions for the variances and covariance are given. Under the assumptions made in this model, the predators will eventually have the region to themselves.

Only the very elementary aspects of certain branches of mathematics are required in order to follow the development of this paper. While the derivations involve systems of difference-differential equations, these are linear with constant coefficients and can be solved by the use of Laplace transforms. In fact, the only Laplace transforms required are those of simple exponential functions, and these are given in [1], although a more convenient source is [3]. Some knowledge of matrix algebra is required, but only enough to follow the computation of the inverses of two-by-two and three-by-three matrices. Anyone who has read Allen [1] and Feller [4] should find the going here quite easy. Much of this paper owes its inspiration to Bharucha-Reid [2], and the reader is referred to this excellent work if he is interested in pursuing the subject further.

2. A PURE MIGRATION MODEL

Consider first a system containing N individuals and consisting of two areas, A and B. The only choice open to the individuals is either to remain where they are or to move to the other area. At time $t = 0$ there are x_0 individuals in area A and $(N - x_0)$ individuals in area B. We denote by $P(x \to y \mid h)$ the probability that in a time interval of h units the population of area A changes from x individuals to y individuals, and we assume:

$$
\begin{aligned}
&P(x \to x + 1 \mid h) = (N - x)\lambda h + o(h) \\
&P(x \to x - 1 \mid h) = x\mu h + o(h) \qquad\qquad (2.1)\\
&P(x \to y \mid h) = o(h)\,, \qquad\qquad y \neq x + 1\,,\quad x - 1\,,
\end{aligned}
$$

where λ and μ are independent of t and $o(h)$ contains only powers of h higher than the first and goes to zero faster than h. The symbol $o(h)$ denotes any quantity with this property. Implicit in (2.1) is the somewhat unrealistic assumption that the individuals in the system tend to migrate to the area with the smaller population. It is this implicit assumption which permits the migration process to go on indefinitely so that there is a limiting probability distribution of the population of area A.

Let $P_x(t)$ denote the probability that the population of area A at time t is x individuals. The population of area A can be x at time $t+h$ because any of the following things happen:

a. The population was x at time t, and no migrations in either direction took place between time t and time $t+h$. The probability of this is $P_x(t)[(1-(N-x)\lambda h - x\mu h)+o(h)]$.

b. The population was $x+1$ at time t, and one individual migrated from A to B. The probability of this is $P_{x+1}(t)[(x+1)\mu h + o(h)]$.

c. The population was $x-1$ at time t, and one individual migrated from B to A. The probability of this is $P_{x-1}(t)[(N-x+1)\lambda h + o(h)]$.

d. Other events occurred whose probability is encompassed in the term $o(h)$.

When we combine these probabilities, remembering that $o(h)$ denotes any quantity which goes to zero faster than h, we have

$$
\begin{aligned}
P_x(t+h) = P_x(t)[1-(N-x)\lambda h - x\mu h] \\
+ P_{x+1}(t)[(x+1)\mu h] \\
+ P_{x-1}(t)[(N-x+1)\lambda h] + o(h)\,.
\end{aligned}
\tag{2.2}
$$

(2.2) holds except at the boundary points, $x=0$ and $x=N$. For these two points, we introduce the convention

$$P_x(t) = 0\,, \qquad x<0\,, \qquad x>N\,. \tag{2.3}$$

With this convention, (2.2) becomes quite general. Next, we subtract $P_x(t)$ from both sides of (2.2) and divide by h. This gives

$$
\begin{aligned}
\frac{P_x(t+h)-P_x(t)}{h} = -N\lambda P_x(t) + x(\lambda-\mu)P_x(t) \\
+ (x+1)\mu P_{x+1}(t) + N\lambda P_{x-1}(t) \\
- (x-1)\lambda P_{x-1}(t) + o(h)/h\,.
\end{aligned}
\tag{2.4}
$$

When we take the limit of (2.4) as h goes to zero, we have the system of difference-differential equations known as the Kolmogorov differential equations for our pure migration process:

$$
\begin{aligned}
dP_0(t)/dt &= -N\lambda P_0(t) + \mu P_1(t) \\
dP_x(t)/dt &= -N\lambda P_x(t) + x(\lambda-\mu)P_x(t) + (x+1)\mu P_{x+1}(t) \\
&\quad + N\lambda P_{x-1}(t) - (x-1)\lambda P_{x-1}(t)\,, \qquad (x=1,\cdots,N-1)\,. \\
dP_N(t)/dt &= -N\mu P_N(t) + \lambda P_{N-1}(t)\,.
\end{aligned}
\tag{2.5}
$$

Readers familiar with the theory of queues will recognize our arguments here.

We now introduce the generating function: (see Feller [4], Chapter 11)

$$P(s,t) = \sum_{x=0}^{N} s^x P_x(t)\,. \tag{2.6}$$

To derive $P(s,t)$, we multiply $dP_x(t)/dt$ by s^x and sum as x varies from zero to N. We have

$$
\begin{array}{llllll}
s^0p_0' = -N\lambda p_0 & & +\mu p_1 & & \\
s^1p_1' = -N\lambda s p_1 & +(\lambda-\mu)sp_1 & +\mu 2sp_2 & +N\lambda sp_0 & \\
s^2p_2' = -N\lambda s^2 p_2 & +(\lambda-\mu)2s^2p_2 & +\mu 3s^2p_3 & +N\lambda s^2p_1 & -\lambda s^2 p_1 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
s^Np_N' = -N\lambda s^N p_N & +(\lambda-\mu)Ns^Np_N & +N\lambda s^N p_{N-1} & -\lambda(N-1)s^Np_{N-1}\,, &
\end{array}
\tag{2.7}
$$

where

$$p'_x = dP_x(t)/dt \qquad (2.8)$$
$$p_x = P_x(t)\,.$$

The reason for the rather complicated form of (2.7) is to facilitate addition in (2.9) below. Summing (2.7) gives

$$\begin{aligned}\sum_{x=0}^{N} s^x p'_x = &- N\lambda \sum_{x=0}^{N} s^x p_x + s(\lambda-\mu) \sum_{x=0}^{N} x s^{x-1} p_x \\ &+ \mu \sum_{x=1}^{N} x s^{x-1} p_x + N\lambda s\Big(\sum_{x=1}^{N} s^x p_x - s^N p_N\Big) \\ &- \lambda s^2\Big(\sum_{x=1}^{N} x s^{x-1} p_x - N s^{N-1} p_N\Big)\,,\end{aligned} \qquad (2.9)$$

or,

$$\begin{aligned}\partial P(s, t)/\partial t = &- N\lambda P(s, t) + s(\lambda - \mu)\partial P(s, t)/\partial s \\ &+ \mu\partial P(s, t)/\partial s + N\lambda s[P(s, t) - s^N p_N] \\ &- \lambda s^2[\partial P(s, t)/\partial s - N s^{N-1} p_N]\,.\end{aligned} \qquad (2.10)$$

(2.10) simplifies to

$$\partial P(s, t)/\partial t = N\lambda(s - 1)P(s, t) - (s - 1)(\lambda s + \mu)\partial P(s, t)/\partial s\,. \qquad (2.11)$$

Temporarily, we postpone the problem of solving (2.11) and consider instead the Kolmogorov differential equations when $N = 1$.

We have

$$\begin{aligned} p'_0 &= - \lambda p_0 + \mu p_1 \\ p'_1 &= - \mu p_1 + \lambda p_0\,.\end{aligned} \qquad (2.12)$$

We consider two cases, $x_0 = 0$ and $x_0 = 1$. In the first case, taking Laplace transforms gives

$$\begin{aligned} z p_0(z) - 1 &= - \lambda p_0(z) + \mu p_1(z) \\ z p_1(z) \qquad &= - \mu p_1(z) + \lambda p_0(z)\,,\end{aligned} \qquad (2.13)$$

where $p_x(z)$ is the Laplace transform of p_x. Solving the system (2.13) gives

$$\begin{aligned} p_0(z) &= \frac{z + \mu}{z(z + \lambda + \mu)} = \frac{1}{\lambda + \mu}\left[\frac{\mu}{z} + \frac{\lambda}{z + \lambda + \mu}\right] \\ p_1(z) &= \frac{\lambda}{z(z + \lambda + \mu)} = \frac{\lambda}{\lambda + \mu}\left[\frac{1}{z} - \frac{1}{z + \lambda + \mu}\right].\end{aligned} \qquad (2.14)$$

In order to simplify the problem of finding the inverse transforms, we have expanded the Laplace transforms into partial fractions. Inversion can, in fact, be done directly from tables. The expressions in (2.14) each contain two terms, the first of which is the transform of a constant and the second the transform of an exponential function. Inverting, we have

$$\begin{aligned} P_0(t) &= [1/(\lambda + \mu)](\mu + \lambda e^{-(\lambda+\mu)t}) \\ P_1(t) &= [\lambda/(\lambda + \mu)](1 - e^{-(\lambda+\mu)t})\,.\end{aligned} \qquad (2.15)$$

In (2.15) note that $P_0(t) + P_1(t) = 1$. From (2.15) we can compute $P_1(s, t)$, the generating function for the case $x_0 = 0$.

$$P_1(s,t) = P_0(t) + sP_1(t) \tag{2.16}$$
$$= (1/(\lambda+\mu))[(\lambda s+\mu) - \lambda(s-1)e^{-(\lambda+\mu)t}] .$$

It is easy to show that (2.16) satisfies (2.11).

Next, consider the case when $x_0 = 1$. Taking Laplace transforms of (2.12) gives

$$\begin{aligned} zp_0(z) \quad &= -\lambda p_0(z) + \mu p_1(z) \\ zp_1(z) - 1 &= -\mu p_1(z) + \lambda p_0(z) . \end{aligned} \tag{2.17}$$

The solution to (2.17) is

$$\begin{aligned} p_0(z) &= \mu/[z(z+\lambda+\mu)] \\ p_1(z) &= (z+\lambda)/[z(z+\lambda+\mu)] , \end{aligned} \tag{2.18}$$

which, after expanding into partial fractions and inverting, gives

$$\begin{aligned} P_0(t) &= [\mu/(\lambda+\mu)](1-e^{-(\lambda+\mu)t}) \\ P_1(t) &= [1/(\lambda+\mu)](\lambda+\mu e^{-(\lambda+\mu)t}) . \end{aligned} \tag{2.19}$$

Finally, the generating function in this case is

$$P_2(s,t) = [1/(\lambda+\mu)][(\lambda s+\mu) + \mu(s-1)e^{-(\lambda+\mu)t}] , \tag{2.20}$$

which also satisfies (2.11).

Now, $P_2(s,t)$ describes the stochastic behavior of an individual who is in area A at time 0, and $P_1(s,t)$ describes the stochastic behavior of an individual who is in area B at time 0. In our original model, the number of persons in area A at time t is the sum of those who were there at time 0 and are still there and those who were in area B at time 0 and have since moved to area A. Since the generating function of a sum of independent random variables is the product of their generating functions, a logical candidate for a solution to (2.11) is

$$P(s,t) = P_2^{x_0}(s,t)P_1^{N-x_0}(s,t) . \tag{2.21}$$

It is a simple exercise to show that (2.21) is, in fact, a solution to (2.11) and meets the boundary condition

$$P(s,0) = s^{x_0} , \tag{2.22}$$

The generating function of the limiting probability distribution of the population of area A is obtained by taking the limit of $P(s,t)$ as t becomes infinite. We have

$$\lim_{t\to\infty} P(s,t) = [(\lambda s+\mu)/(\lambda+\mu)]^N , \tag{2.23}$$

(2.23) is the generating function for the binomial distribution. Hence

$$P_x(\infty) = \binom{N}{x}\left(\frac{\lambda}{\lambda+\mu}\right)^x\left(\frac{\mu}{\lambda+\mu}\right)^{N-x} . \tag{2.24}$$

$P_x(t)$ is the coefficient of s^x in the power series expansion of $P(s,t)$. It is most easily obtained in this case by repeated differentiation, i. e.,

$$P_x(t) = \frac{1}{x!}\left.\frac{d^xP(s,t)}{ds^x}\right|_{s=0} , \qquad (x = 0, 1, \cdots, N) \tag{2.25}$$

In many applications, the moments of the distribution $P_x(t)$ are more useful than the probabilities themselves. It is easy to show by differentiating (2.6) twice and setting $s = 1$ that

$$\frac{dP(s,t)}{ds}\bigg|_{s=1} = E(x \mid t) \qquad (2.26)$$
$$\frac{d^2P(s,t)}{ds^2}\bigg|_{s=1} = E(x^2 \mid t) + E(x \mid t)\,.$$

When we apply (2.26) to (2.21) and use (2.15), we obtain the trend in the population of area A and the variance about the trend.

$$\begin{aligned} E(x \mid t) &= (N\lambda + \mu)(1 - e^{-(\lambda+\mu)t}) + x_0 e^{-(\lambda+\mu)t} \\ \sigma^2(x \mid t) &= [(1 - e^{-(\lambda+\mu)t})/(\lambda + \mu)^2] \\ &\quad \times [N\lambda\mu + (x_0\mu^2 + (N - x_0)\lambda^2)e^{-(\lambda+\mu)t}]\,. \end{aligned} \qquad (2.27)$$

(2.27) shows clearly that the effect of the initial distribution of the N individuals between areas A and B diminishes with the passage of time. Both the trend and the variance approach a finite limit.

$$\begin{aligned} \lim_{t\to\infty} E(x \mid t) &= N\lambda/(\lambda + \mu) \\ \lim_{t\to\infty} \sigma^2(x \mid t) &= N\lambda\mu/(\lambda + \mu)^2\,. \end{aligned} \qquad (2.28)$$

As we remarked at the outset, this model is based on the assumption that individuals in the system tend to migrate to the area with the smaller population. The long run distribution of the population between the two areas depends on how strong these tendencies are, that is, on the relative magnitudes of λ and μ. λ is a measure of the relative attractiveness of area A. The larger λ, the more likely it is that individuals will shift from area B to area A. Similarly, μ is a measure of the relative attractiveness of area B. Both areas become less attractive as they become more crowded. If we change our assumption so that both areas become more attractive as they become more crowded, the Kolmogorov equations do not have a solution. If we make the attractiveness of each area independent of its population, the solution to the Kolmogorov equations exists but is very difficult to derive. Readers familiar with the theory of queues will recognize this latter situation as that of a single server queue with a finite number of states.

3. A BIRTH-DEATH MODEL WITH MIGRATION

Our second model might be considered as an extension of our first. Again, we consider a system made up of two areas, A and B. In this case, however, we permit migration to and from the system and migration within the system. Migrations into and out of the system may occur because of birth, death, emigration or immigration. We let x be the population of Area A and y be the population of area B. Employing the same notation as in the previous section, we have

$$\begin{aligned} P(x \to x+1, y \to y \mid h) &= x\lambda_1 h + o(h) \\ P(x \to x, y \to y+1 \mid h) &= y\lambda_2 h + o(h) \\ P(x \to x-1, y \to y \mid h) &= x\mu_1 h + o(h) \end{aligned} \qquad (3.1)$$

$$P(x \to x, y \to y-1 \mid h) = y\mu_2 h + o(h)$$
$$P(x \to x-1, y \to y+1 \mid h) = x\Delta_1 h + o(h)$$
$$P(x \to x+1, y \to y-1 \mid h) = y\Delta_2 h + o(h)$$

All other events have probability $o(h)$, which goes to zero faster than h.

Our reasoning here is precisely the same as in the preceding section. We denote by $P_{x,y}(t)$ the probability that at time t the population of area A is x and the population of area B is y.

We have

$$\begin{aligned} P_{0,0}(t+h) &= P_{0,0}(t) + \mu_1 h P_{1,0}(t) + \mu_2 h P_{0,1}(t) + o(h) \\ P_{x,y}(t+h) &= P_{x,y}(t)[1 - (\lambda_1 + \mu_1 + \Delta_1)h \\ &\quad - (\lambda_2 + \mu_2 + \Delta_2)h + (x-1)\lambda_1 P_{x-1,y}(t) \\ &\quad + (x+1)\mu_1 P_{x+1,y}(t) + (x+1)\Delta_1 P_{x+1,y-1}(t) \\ &\quad + (y-1)\lambda_2 P_{x,y-1}(t) + (y+1)\mu_2 P_{x,y+1}(t) \\ &\quad + (y+1)\Delta_2 P_{x-1,y+1}(t) + o(h)\,. \qquad (x, y = 1, 2, \cdots) \end{aligned} \tag{3.2}$$

Subtraction of $P_{x,y}(t)$ from both sides of (3.2), division by h, and taking the limit as h goes to zero gives us our system of Kolmogorov differential equations.

$$\begin{aligned} p'_{0,0} &= \mu_1 p_{1,0} + \mu_2 p_{0,1} \\ p'_{x,y} &= -\,[(\lambda_1 + \mu_1 + \Delta_2)y]\,p_{x,y} + \lambda_1(x-1)p_{x-1,y} \\ &\quad + \mu_1(x+1)p_{x+1,y} + \Delta_1(x+1)p_{x+1,y-1} \\ &\quad + \lambda_2(y-1)p_{x,y-1} + \mu_2(y+1)p_{x,y+1} \\ &\quad + \Delta_2(y+1)p_{x-1,y+1} + o(h)\,. \qquad x, y = 1, 2, \cdots \end{aligned} \tag{3.3}$$

In this case, our generating function is the bivariate generating function,

$$P(r, s, t) = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} r^x s^y P_{x,y}(t)\,. \tag{3.4}$$

Multiplying (3.3) by $r^x s^y$ and summing gives, after some manipulation,

$$\begin{aligned} \frac{\partial P(r,s,t)}{\partial t} &= [\mu_1(1-r) - r\lambda_1(1-r) + \Delta_1(s-r)]\frac{\partial P(r,s,t)}{\partial r} \\ &\quad + [\mu_2(1-s) - s\lambda_2(1-s) + \Delta_2(r-s)]\frac{\partial P(r,s,t)}{\partial s} \end{aligned} \tag{3.5}$$

In spite of its very attractive symmetry, (3.5) is a rather difficult partial differential equation. We cannot solve it by a ruse similar to the one employed in the previous section. We can, however, obtain the means, and the covariance of the joint probability distribution of x and y if we resort to what one of the author's professors used to call "the method of brutal force." This is descriptive, as we shall see.

Instead of the generating function, (3.4), we introduce the moment generating function,

$$M(u, v, t) = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} e^{ux+vy} P_{x,y}(t)\,. \tag{3.6}$$

This implies the substitutions

$$
\begin{aligned}
&r = e^u\,, \quad s = e^v \\
&\frac{\partial P(r, s, t)}{\partial t} = \frac{\partial M(u, v, t)}{\partial t} \\
&\frac{\partial P(r, s, t)}{\partial r} = e^{-u}\frac{\partial M(u, v, t)}{\partial u} \\
&\frac{\partial P(r, s, t)}{\partial s} = e^{-u}\frac{\partial M(u, v, t)}{\partial v}
\end{aligned} \tag{3.7}
$$

Further, when we expand e^{ux+vy} as a power series in u and v and sum, (3.6) becomes

$$
\begin{aligned}
M(u, v, t) = 1 &+ uE(x \mid t) + vE(y \mid t) \\
&+ (u^2/2!)E(x^2 \mid t) + uvE(xy \mid t) \\
&+ (v^2/2!)E(y^2 \mid t) + \cdots .
\end{aligned} \tag{3.8}
$$

Making the substitutions of (3.7) into (3.5) and using the expansion (3.8) gives us a very unwieldy expression:

$$
\begin{aligned}
&uE'(x) + vE'(y) + (u^2/2!)E'(x^2) \\
&\quad + uvE'(xy) + (v^2/2!)E'(y^2) + \cdots \\
&= u[(\lambda_1 - \mu_1 - \Delta_1)E(x) + \Delta_2 E(y)] \\
&\quad + v[\Delta_1 E(x) + (\lambda_2 - \mu_2 - \Delta_2)E(y)] \\
&\quad + (u^2/2!)[2(\lambda_1 - \mu_1 - \Delta_1)E(x^2) \\
&\quad + (\lambda_1 + \mu_1 + \Delta_1)E(x) + 2\Delta_2 E(xy) + \Delta_2 E(y)] \\
&\quad + uv[\Delta_1 E(x^2) - \Delta_1 E(x) \\
&\quad + (\lambda_1 - \mu_1 - \Delta_1 + \lambda_2 - \mu_2 - \Delta_2)E(xy) \\
&\quad + \Delta_2 E(y^2) - \Delta_2 E(y)] + (v^2/2!)[\Delta_1 E(x) \\
&\quad + 2\Delta_1 E(xy) + 2(\lambda_2 - \mu_2 - \Delta_2)E(y^2) \\
&\quad + (\lambda_2 + \mu_2 + \Delta_2)E(y)] + \cdots ,
\end{aligned} \tag{3.9}
$$

where

$$
\begin{aligned}
E'(\cdot) &= dE(\cdot \mid t)dt \\
E(\cdot) &= E(\cdot \mid t) .
\end{aligned} \tag{3.10}
$$

Because (3.9) is a power series, we can equate coefficients of like powers of u and v. Ignoring powers of u and v higher than the second gives us two systems of simultaneous differential equations:

$$
\begin{aligned}
E'(x) &= (\lambda_1 - \mu_1 - \Delta_1)E(x) + \Delta_2 E(y) \\
E'(y) &= \Delta_1 E(x) + (\lambda_2 - \mu_2 - \Delta_2)E(y)
\end{aligned} \tag{3.11}
$$

$$
\begin{aligned}
E'(x^2) &= 2(\lambda_1 - \mu_1 - \Delta_1)E(x^2) + 2\Delta_2 E(xy) \\
&\quad + (\lambda_1 + \mu_1 + \Delta_1)E(x) + \Delta_2 E(y) \\
E'(xy) &= \Delta_1 E(x^2) + (\lambda_1 - \mu_1 - \Delta_1 + \lambda_2 - \mu_2 - \Delta_2)E(xy) \\
&\quad + \Delta_2 E(y^2) - \Delta_1 E(x) - \Delta_2 E(y) \\
E'(y^2) &= 2\Delta_1 E(xy) + 2(\lambda_2 - \mu_2 - \Delta_2)E(y^2) \\
&\quad + \Delta_1 E(x) + (\lambda_2 + \mu_2 + \Delta_2)E(y)
\end{aligned} \tag{3.12}
$$

We can write the variances and covariance directly into (3.12) if we remember their definitions:

$$\begin{aligned} \sigma^2(x \mid t) &= E(x^2 \mid t) - E^2(x \mid t) \\ \sigma(xy \mid t) &= E(xy \mid t) - E(x \mid t)E(y \mid t) \\ \sigma^2(y \mid t) &= E(y^2 \mid t) - E^2(y \mid t) \end{aligned} \tag{3.13}$$

Substituting the expressions in (3.13) and their derivatives into (3.12) gives

$$\begin{aligned} \sigma'^2(x) &= 2(\lambda_1 - \mu_1 - \Delta_1)\sigma^2(x) + 2\Delta_2\sigma(xy) \\ &\qquad + (\lambda_1 + \mu_1 + \Delta_1)E(x) + \Delta_2 E(y) \\ \sigma'(xy) &= \Delta_1\sigma^2(x) + (\lambda_1 - \mu_1 - \Delta_1 + \lambda_2 - \mu_2 - \Delta_2)\sigma(xy) \\ &\qquad + \Delta_2\sigma^2(y) - \Delta_1 E(x) - \Delta_2 E(y) \\ \sigma'^2(y) &= 2\Delta_1\sigma(xy) + 2(\lambda_2 - \mu_2 - \Delta_2)\sigma^2(y) \\ &\qquad + \Delta_1 E(x) + (\lambda_2 + \mu_2 + \Delta_2)E(y), \end{aligned} \tag{3.14}$$

where, again the "prime" notation indicates differentiation with respect to time. We can solve (3.11) for $E(x)$ and $E(y)$ and use these results in (3.14) to obtain the variances and covariance.

In order to simplify the solution of (3.11), we assume that $\lambda_1 - \mu_1 = \lambda_2 - \mu_2 = \alpha$. Further, we let $X(z)$ be the Laplace transform of $E(x \mid t)$ and $Y(z)$ be the Laplace transform of $E(y \mid t)$; that is,

$$\begin{aligned} X(z) &= \int_0^\infty e^{-zt}E(x \mid t)\,dt \\ Y(z) &= \int_0^\infty e^{-zt}E(y \mid t)\,dt \end{aligned} \tag{3.15}$$

When we transform (3.11), we obtain

$$\begin{aligned} zX(z) - x_0 &= (\alpha - \Delta_1)X(z) + \Delta_2 Y(z) \\ zY(z) - y_0 &= \Delta_1 X(z) + (\alpha - \Delta_2)Y(z) \end{aligned} \tag{3.16}$$

Rearranging (3.16) gives us a matrix equation:

$$\begin{bmatrix} z - (\alpha - \Delta_1) & -\Delta_2 \\ -\Delta_1 & z - (\alpha - \Delta_2) \end{bmatrix} \begin{bmatrix} X(z) \\ Y(z) \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \tag{3.17}$$

The solution to (3.17) is

$$\begin{aligned} X(z) &= \frac{x_0(z - (\alpha - \Delta_2)) + y_0\Delta_2}{(z - a)[z - (\alpha - \Delta_1 - \Delta_2)]} \\ Y(z) &= \frac{x_0\Delta_1 + y_0[z - (\alpha - \Delta_1)]}{(z - \alpha)[z - (\alpha - \Delta_1 - \Delta_2)]} \end{aligned} \tag{3.18}$$

We can either expand the expressions in (3.18) into partial fractions and invert from the very short table of inverse transforms in [1], or we can obtain the inverses directly from the rather extensive table in [3]. In either case, we obtain

(3.19)
$$E(x \mid t) = \frac{\Delta_2}{(\Delta_1 + \Delta_2)}(x_0 + y_0)e^{\alpha t} + \left[\frac{(\Delta_1 x_0 - \Delta_2 y_0)}{(\Delta_1 + \Delta_2)}\right] e^{(\alpha - \Delta_1 - \Delta_2)t}$$
$$E(y \mid t) = \frac{\Delta_1}{(\Delta_1 + \Delta_2)}(x_0 + y_0)e^{\alpha t} - \left[\frac{(\Delta_1 x_0 - \Delta_2 y_0)}{(\Delta_1 + \Delta_2)}\right] e^{(\alpha - \Delta_1 - \Delta_2)t}$$

In (3.19) we notice two things immediately. First, the population of th- system (areas A and B combined) grows exponentially.

(3.20)
$$E(x + y \mid t) = (x_0 + y_0)e^{\alpha t} .$$

Second, the influence of the initial distribution of the population between the two areas diminishes with time. That is,

(3.21)
$$E(x \mid t) \approx \frac{\Delta_2}{(\Delta_1 + \Delta_2)}(x_0 + y_0)e^{\alpha t}$$
$$E(y \mid t) \approx \frac{\Delta_1}{(\Delta_1 + \Delta_2)}(x_0 + y_0)e^{\alpha t} ,$$

where the symbol $\approx$ means that the ratio of the two sides tends to unity. In the long run, the distribution of the population within the system depends on the relative attractiveness of the two areas that comprise the system.

In order to obtain the variances and covariance, we make further simplifying assumption:

(3.22)
$$\lambda_1 = \lambda_2 = \lambda$$
$$\mu_1 = \mu_2 = \mu$$
$$\Delta_1 = \Delta_2 = \Delta$$

We also note that, initially, the variances and covariances are zero and introduce the Laplace transforms

(3.23)
$$\sigma_x^2(z) = \int_0^\infty e^{-zt}\sigma^2(x \mid t)dt$$
$$\sigma_{xy}(z) = \int_0^\infty e^{-zt}\sigma(xy \mid t)dt$$
$$\sigma_y^2(z) = \int_0^\infty e^{-zt}\sigma^2(y \mid t)dt$$

When we take Laplace transforms and rearrange (3.14), we obtain the matrix equation

(3.23)
$$\begin{bmatrix} z - 2(\lambda - \mu - \Delta) & & 0 \\ -\Delta & z - 2(\lambda - \mu - \Delta) & -\Delta \\ 0 & -2\Delta & z - 2(\lambda - \mu - \Delta) \end{bmatrix} \begin{bmatrix} \sigma_x^2(z) \\ \sigma_{xy}(z) \\ \sigma_y^2(z) \end{bmatrix} = \begin{bmatrix} (\lambda + \mu + \Delta)X(z) + \Delta Y(z) \\ -\Delta(X(z) + Y(z)) \\ \Delta X(z) + (\lambda + \mu + \Delta)Y(z) \end{bmatrix}$$

Now it is a simple but laborious task to solve (3.24), expand the solutions into partial fractions, and to find the inverse transforms. The net result of all the work involved is

$$\begin{aligned}
\sigma^2(x \mid t) &= [(x_0 + y_0)/2]e^{2\alpha t}\left[\frac{\lambda+\mu}{2\alpha} + \frac{(\lambda+\mu+4\Delta)}{2(\alpha-4\Delta)}e^{-4\Delta t}\right. \\
&\qquad\qquad \left. - \frac{(\lambda+\mu)\alpha - 4\Delta\mu}{(\alpha-4\Delta)}e^{-\alpha t}\right] \\
&\quad + [(x_0 - y_0)/2][(\lambda+\mu)/\alpha]e^{2(\alpha-\Delta)t}(1-e^{-\alpha t}) \qquad (3.25)\\
\sigma(xy \mid t) &= [(x_0 + y_0)/2]e^{2\alpha t}\left[\frac{\lambda+\mu}{2\alpha} - \frac{(\lambda+\mu+4\Delta)}{2(\alpha-4\Delta)}e^{-4\Delta t}\right. \\
&\qquad\qquad \left. + \frac{4\lambda\Delta}{(\alpha-4\Delta)\alpha}e^{-\alpha t}\right]
\end{aligned}$$

Because of symmetry, $\sigma'^2(y \mid t)$ can be obtained by interchanging x_0 and y_0 in the expression for $\sigma^2(x \mid t)$ in (3.25); this merely changes the sign of the last term in the expression for $\sigma^2(x \mid t)$. We can use (3.25) to obtain the variance in the population of the system:

$$\begin{aligned}
\sigma^2(x+y \mid t) &= \sigma^2(x \mid t) + \sigma^2(y \mid t) + 2\sigma(xy \mid t) \qquad (3.26)\\
&= (x_0 + y_0)[(\lambda+\mu)/\alpha]e^{2\alpha t}(1-e^{-\alpha t})
\end{aligned}$$

Equations (3.20) and (3.26) give the mean and variance of a pure birth-death process when the initial population is $x_0 + y_0$. Thus, implicit in this section are two models, a pure birth-death model and a mixed birth-death-and-migration model. The pure birth-death model is discussed completely in [2].

4. A PREDATOR-PREY MODEL

Our third and final model is a form of predator-prey model. It might be used to describe the stochastic behavior of what is often euphemistically called a "changing" neighborhood. We consider an area whose population is of two types, A and B. As the number of B-type individuals increases, the A-type individuals become more and more inclined to leave the area, while as the total population of the area increases, more and more B-type individuals are attracted to the area. In the ordinary biological predator-prey model, the B-type individuals eat the A-type individuals; here they merely drive them away. We might call the A-type "old residents" and the B-type "newcomers."

For transition probabilities we have

$$\begin{aligned}
&P(x \to x+1,\ y \to y \mid h) = x\lambda_1 h + o(h) \\
&P(x \to x-1,\ y \to y \mid h) = x\mu_{11} h + y\mu_{12} h + o(h) \\
&P(x \to x,\ y \to y+1 \mid h) = y\lambda_{21} h + x\lambda_{22} h + o(h) \\
&P(x \to x,\ y \to y-1 \mid h) = y\mu_2 h + o(h)
\end{aligned} \qquad (4.1)$$

Thus, $(\lambda_1 - \mu_{11})$ describes the "natural" growth rate of the A-type population, $(\lambda_{21} - \mu_2)$ describes the "natural" growth rate of the B-type population, and μ_{12} and λ_{22} describe the way in which the two populations interact.

After manipulations similar to the ones we have described in the preceding section, we arrive at the following partial differential equation for the bivariate moment generating function of our process:

$$\frac{\partial M(u, v, t)}{\partial t} = [\mu_{11}(e^{-u} - 1) + \lambda_1(e^u - 1) + \lambda_{22}(e^v - 1)]\frac{\partial M(u, v, t)}{\partial u} + [\mu_2(e^{-v} - 1) + \lambda_{21}(e^v - 1) + \mu_{12}(e^{-u} - 1)]\frac{\partial M(u, v, t)}{\partial v} \tag{4.2}$$

As in the preceding model, the symmetry of the equation is deceptive, so we expand (4.2) into power series in u and v and equate coefficients of like powers. After considerable manipulation, we have

$$\begin{aligned} E'(x) &= (\lambda_1 - \mu_{11})E(x) - \mu_{12}E(y) \\ E'(y) &= \lambda_{22}E(x) + (\lambda_{21} - \mu_2)E(y) \end{aligned} \tag{4.3}$$

$$\begin{aligned} \sigma'^2(x) &= (\lambda_1 - \mu_{11})\sigma^2(x) - 2\mu_{12}\sigma(xy) + (\lambda_1 + \mu_{11})E(x) + \mu_{12}E(y) \\ \sigma'(xy) &= \lambda_{22}\sigma^2(x) + [(\lambda_1 - \mu_{11}) + (\lambda_{21} - \mu_2)]\sigma(xy) - \mu_{12}\sigma^2(y) \\ \sigma'^2(y) &= 2\lambda_{22}\sigma(xy) + 2(\lambda_{21} - \mu_2)\sigma^2(y) + \lambda_{22}E(x) + (\lambda_{21} + \mu_2)E(y) \end{aligned} \tag{4.4}$$

Even under very restrictive simplifying assumptions, (4.4) involves some rather brutal algebra, although the solution is quite easy once we assign numerical values to the coefficients. Hence, we will give only the solution to (4.3). The method of solution is the same as in the preceding section, so we omit all but the initial and final steps. As in the preceding section, we make the simplifying assumption that $\lambda_1 - \mu_{11} = \lambda_{21} - \mu_2 = \alpha$. When we take Laplace transforms and rearrange (4.3), we have the matrix

$$\begin{bmatrix} z - \alpha & \mu_{12} \\ -\lambda_2 & z - \alpha \end{bmatrix}\begin{bmatrix} X(z) \\ Y(z) \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \tag{4.5}$$

The solution to (4.5) is

$$\begin{aligned} X(z) &= [x_0(z - \alpha) - y_0\mu_{12}]/[(z - \alpha)^2 + \lambda_{22}\mu_{12}] \\ Y(z) &= [x_0\lambda_{22} + y_0(z - \alpha)]/[(z - \alpha)^2 + \lambda_{22}\mu_{12}] \end{aligned} \tag{4.6}$$

for which the inverse transforms are

$$\begin{aligned} E(x \mid t) &= e^{\alpha t}[x_0 \cos (\lambda_{22}\mu_{12})^{1/2}t - y_0(\mu_{12}/\lambda_{22})^{1/2} \sin (\lambda_{22}\mu_{12})^{1/2}t] \\ E(y \mid t) &= e^{\alpha t}[y_0 \cos (\lambda_{22}\mu_{12})^{1/2}t + x_0(\lambda_{22}/\mu_{12})^{1/2} \sin (\lambda_{22}\mu_{12})^{1/2}t) \end{aligned} \tag{4.7}$$

In spite of appearances, (4.7) does not imply periodic variation in the trends of the two populations. Instead, (4.7) implies that $E(x \mid t)$ will go to zero in finite time, after which $E(y \mid t)$ grows exponentially. In other words, the "newcomers" will have the area to themselves in finite time. How long this takes depends upon the magnitudes of λ_{22} and μ_{12}. We have

$$\begin{aligned} &E(x \mid t^*) = 0 \\ &t^* = [\tan^{-1}(x_0/y_0)(\lambda_{22}/\mu_{12})^{1/2}]/(\lambda_{22}\mu_{12})^{1/2} \end{aligned} \tag{4.8}$$

If α is large relative to λ_{22} and μ_{12}, the A-type population may even increase for some time before it diminishes and eventually vanishes.

To touch briefly on the question of the variances and the covariance, when we introduce Laplace transforms and rearrange (4.4), we obtain

$$\begin{bmatrix} z-2\alpha & 2\mu_{12} & 0 \\ -\lambda_{22} & z-2\alpha & \mu_{12} \\ 0 & -\lambda_{22} & z-2\alpha \end{bmatrix} \begin{bmatrix} \sigma_x^2(z) \\ \sigma_{xy}(z) \\ \sigma_y^2(z) \end{bmatrix} = \begin{bmatrix} (\lambda_1+\mu_{11})X(z)+\mu_{12}Y(z) \\ 0 \\ \lambda_{22}X(z)+(\lambda_{21}+\mu_2)Y(z) \end{bmatrix} \tag{4.9}$$

If we use the symbol A to designate the matrix on the left hand side of (4.9), we have for its determinant and its inverse

$$|A| = (z-2\alpha)[(z-2\alpha)^2 + 4\mu_{12}\lambda_{22}]$$

$$A^{-1} = \frac{1}{|A|}\begin{bmatrix} (z-2\alpha)^2+2\lambda_{22}\mu_{12} & -2\mu_{12}(z-2\alpha) & 2\mu_{12}^2 \\ \lambda_{22}(z-2\alpha) & (z-2\alpha)^2 & -\mu_{12}(z-2\alpha) \\ 2\lambda_{22}^2 & 2\lambda_{22}(z-2\alpha) & (z-2\alpha)^2+2\lambda_{22}\mu_{12} \end{bmatrix} \tag{4.10}$$

It is easy to see that, unless we assign numerical values to the parameters, the result of applying (4.10) and (4.6) to the system (4.9) is extremely complicated. Once numerical values are assigned, however, the variances and covariance of the process are quite easy to compute. Hence, we will not derive general expressions for the variances and covariance. They are given implicitly by (4.6), (4.9), and (4.10).

5. SUMMARY AND CONCLUSIONS

In this paper we have presented three stochastic models of birth, death, and migration. The first model was a pure migration model, and we were able to derive the generating function of the stochastic process as well as the trend and the variance about the trend of the process. The second model was a mixed birth, death, and migration model. Using fairly simple mathematics, we were unable to derive the generating function of the process, but we did derive the trend and the variance of the mixed model and, implicitly, of a pure birth-death model. Our third model was a predator-prey model. In the development of this model, the mathematics soon become rather complicated, and we exhibited only the trend of the process. Once numerical values are assigned to the parameters, however, the variances and covariance of the predator-prey model can be computed from the implicit expressions provided.

While our major purpose has been to exhibit these models and to discuss some of their implications, we have also tried to show that interesting and useful results can be obtained using relatively elementary aspects of advanced mathematics. A thorough study of stochastic processes, however, requires a comprehensive grasp of some rather advanced mathematics. The interesting and useful results of applying the theory, however, are well worth the effort of learning the mathematics.

REFERENCES

[1] Allen, R. G. D. *Mathematical Economics*. London: Macmillan, 1956.

[2] Bharucha-Reid, A. T. *Elements of the Theory of Markov Processes and their Application*. New York: McGraw-Hill, 1954.

[3] Erdelyi, A. (ed.) *Tables of Integral Transforms, Vol. 1*. New York: McGraw-Hill, 1954.

[4] Feller, W. *An Introduction to Probability Theory and its Applications*. New York: John Wiley, 1957.