Comment: Partially Observed Markov Chains and Genetic Demography

Joel E. Cohen

1. INTRODUCTION

Stochastic population models based on branching processes have progressed from single-type models to multi-type models. Jagers gives a marvellously elegant, succinct and insightful synthesis of this progress. To stimulate further progress, I wish to draw attention to two (of the many) questions that remain open. First, how should a population's dynamics be described when the population has unobserved heterogeneity, that is, when it contains more types than an observer distinguishes? Second, how can one understand the long-run behavior of simple nonlinear models that accommodate genetic and demographic processes? Specifically, how can one determine when the asymptotic or stable composition of a multi-type population does or does not depend on the initial composition of the population?

2. UNOBSERVED HETEROGENEITY

To take a simple case, suppose a population contains just two types each with its own distinct life law, in Jagers' terminology (demographers may interpret *life law* to mean the net fertility function). If the observer does not distinguish the two types, but instead counts all individuals as belonging to a single type, the total population size is no longer Markovian. Future change depends, not only on the present total population size, but also on the (unobserved) numbers of individuals of each type.

More generally, suppose a population contains a large number of types with distinct life laws, which an observer crudely partitions into a small number of distinguishable types. For example, a typical human being's DNA has three billion or so base pairs. Except for identical twins, any two individuals are genetically distinct and could have distinct life laws. The genetic markers currently available to distinguish among genotypes are relatively few.

Joel E. Cohen is Professor of Populations at Rockefeller University, 1230 York Avenue, Box 20, New York, New York 10021-6399.

The situation where the life law depends on many individual characteristics that are inaccessible to the observer is probably the generic situation in biological population growth (as opposed to neutron cascades, for example, where the homogeneity of the reproducing particles is plausible). Hence it seems highly desirable to develop theory for this situation.

A partially observed Markov chain or process, sometimes also called a hidden Markov chain or a lumped Markov chain, is a special case of a random system with complete connections (RSCC). There is a well-developed theory of RSCCs, which is at last available in an up-to-date, detailed exposition in English (Iosifescu and Grigorescu, 1990). I believe the development and application of the theory of RSCCs for population growth with unobserved heterogeneity remains for the future. What are the analogs of Jagers' theorems in this situation?

3. GENETIC DEMOGRAPHY

Two of the main branches of biological population modeling are stable population theory (the demographic theory of age-structured populations) and Mendelian population genetics. In stable population theory, as in the models Jagers describes, only the female population is modeled. Interactions between the sexes are ignored. In the Mendelian genetics of diploid organisms, the genetic contribution of each parent is crucial.

Even for very large populations, where stochastic effects play no role or are ignored, the demographic and Mendelian models differ strikingly. Under reasonable assumptions, in stable population theory, the age-composition of a population approaches a limiting composition that is independent of initial (demographic) conditions, whereas the genotypic composition of a Mendelian population approaches a limiting composition that depends on initial (genetic) conditions. How are these two modes of behavior reconciled when the two models are combined (e.g., Norton, 1928; Charlesworth, 1980)?

To be specific (Charlesworth, 1980; Orzack, 1985), consider a hypothetical population with two age groups, young (group 1) and old (group 2),

genetically differentiated at a single locus with two alleles, B and b. An individual receives one allele from each parent, so there are three genotypes BB, Bb and bb, and we use the dummy variable g to denote any of these. Let $L(g)$ be the so-called Leslie matrix of genotype g:

$$L(g) = \begin{pmatrix} m_1(g) & m_2(g) \\ l_1(g) & 0 \end{pmatrix}, \tag{3.1}$$

where $m_i(g)$ is the effective fertility (measured in gametes, not in zygotes) of the ith age group and $l_1(g)$ is the probability that an individual of age 1 survives one unit of time to become age 2, for genotype g. The vital rates (i.e., the elements of $L(g)$) are assumed to be identical for both sexes of genotype g. For t = 0, 1, 2, . . . , let $X(g, t)$ be the population vector, that is, the number of individuals by age group, of genotype g:

$$X(g,t) = \begin{pmatrix} X_1(g,t) \\ X_2(g,t) \end{pmatrix}. \tag{3.2}$$

Let the total number of gametes produced between t and $t + 1$ be

$$\begin{aligned} G(t+1) = {} & (L(BB)X(BB,t))_1 \\ & + (L(Bb)X(Bb,t))_1 \\ & + (L(bb)X(bb,t))_1. \end{aligned} \tag{3.3}$$

Then the frequencies $P(B, t + 1)$ and $P(b, t + 1)$ of alleles B and b, respectively, among these gametes are

$$\begin{aligned} P(B,t+1) &= \frac{2(L(BB)X(BB,t))_1 + (L(Bb)X(Bb,t))_1}{2G(t+1)}, \\ P(b,t+1) &= 1 - P(B,t+1). \end{aligned} \tag{3.4}$$

Assuming that the gametes mate to form zygotes at random, and that there are no stochastic deviations from expected values, the numbers of young of genotype g at time $t + 1$ is taken as

$$\begin{aligned} X_1(BB,t+1) &= P^2(B,t+1)G(t+1), \\ X_1(Bb,t+1) &= 2P(B,t+1)P(b,t+1)G(t+1), \\ X_1(bb,t+1) &= P^2(b,t+1)G(t+1). \end{aligned} \tag{3.5}$$

Age group 2 at $t + 1$ consists of the survivors of the young at time t:

$$\begin{aligned} X_2(g,t+1) &= (L(g)X(g,t))_2 \\ &= l_1(g)X_1(g,t), \end{aligned} \tag{3.6}$$

for all genotypes g. Thus the population vectors $X(g, t + 1)$ are completely specified by the population vectors $X(g, t)$ at the preceding time and by the Leslie matrices $L(g)$.

The model specified by (3.1)-(3.6), together with initial conditions $X(g, 0)$, $g = BB, Bb, bb$, includes as special cases both the classical model of population genetics for one locus with two alleles in a large, randomly mating population with no migration, mutation or drift, and the classical model of stable population theory for a genetically undifferentiated, large, closed, age-structured population.

To reduce this model to the population genetics model, let $X_1(g, 0) > 0$, $X_2(g, 0) = l_1(g) = 0$, for all g. Then the population contains young individuals only, and the so-called fitness of genotype g is given by $m_1(g)$. Let all genotypes have positive fitness equal to some constant m, so that $m_1(BB) = m_1(Bb) = m_1(bb) = m > 0$. Then, from $t = 1$ onward, the population is in Hardy–Weinberg equilibrium. The gene frequencies $P(B, t)$ and $P(b, t)$ are constant,

$$P(B,t) = \frac{2X_1(BB,0) + X_1(Bb,0)}{2X(0)}, \qquad t = 1,2,\ldots, \tag{3.7}$$

where $X(t) = \sum_g [X_1(g, t) + X_2(g, t)]$ is the total population size (in this case, young individuals only) at time t, and the proportions of all young that are of each genotype are constant as the absolute numbers and total population size change exponentially:

$$\begin{aligned} X_1(BB,t) &= P^2(B,1)m^tX(0), \\ X_1(Bb,t) &= 2P(B,1)P(b,1)m^tX(0), \\ X_1(bb,t) &= P^2(b,1)m^tX(0), \\ X(t) &= m^tX(0), \qquad t = 1,2,\ldots. \end{aligned} \tag{3.8}$$

The key point is that the proportions of the population belonging to each genotype stabilize from $t = 1$ onward at an equilibrium that depends on the initial genetic composition of the population through (3.7).

To reduce the model (3.1)-(3.6) to the classical model of stable population theory with two age classes, take $L(Bb) = L(bb) = 0$ or $X(Bb) = X(bb) = 0$ or both, thereby eliminating genetic heterogeneity from $t = 1$ onward, and take $m_1(BB) > 0$, $m_2(BB) > 0$, $l_1(BB) > 0$, $X_1(BB, 0) > 0$, $X_2(BB,$

$0) > 0$. Then $P(B, t) = 1$ for $t = 1, 2, \ldots$ and the theory of stable populations (the Perron–Frobenius theory of primitive matrices in demographic disguise) guarantees that

$$\lim_{t\to\infty} \frac{X_i(BB, t)}{X(t)} = y_i, \qquad \lim_{t\to\infty} \frac{X(t)}{\rho^t} = \text{constant}, \qquad i = 1, 2, \tag{3.9}$$

where $\rho > 0$ is the eigenvalue of $L(BB)$ of maximal modulus, and y is the corresponding eigenvector, with positive elements y_1 and y_2 normalized so that $y_1 + y_2 = 1$. Asymptotically the absolute numbers of young and old and the total population size change exponentially, all at the same rate.

Here the key point is that the equilibrial fractions y_1 and y_2 of young and old depend only on $L(BB)$ and are independent of the initial demographic composition of the population (provided the initial population is not zero).

What happens when all the parameters of the full model are nonzero? It appears that nobody knows. In numerical simulations that allowed the Leslie matrices $L(g, t)$ to vary randomly in time, Orzack (1985, page 559) assumed that the model "represents an ergodic process, [so that] numerical analysis consisted of examining the long-run behavior of a single sample path of the process." However, it is clear from the genetic submodel that, even with constant Leslie matrices, the model may not be ergodic in Orzack's sense, in that the long-run behavior may depend on initial conditions. My own numerical calculations of the full model with time-invariant parameters show that sometimes the asymptotic composition of the population depends on the initial conditions, and sometimes is independent of initial conditions. Still other forms of behavior are not yet excluded. Specifying the regions of the parameter space that give the various forms of behavior seems to be a challenging task.

Small-population versions of this model would describe the production and the pairing of gametes and the survival of young as stochastic processes. Similar questions arise, in addition to the problem of characterizing the probabilities of extinction.

ACKNOWLEDGMENTS

The author acknowledges the support of U.S. National Science Foundation Grant BSR-87-05047, the hospitality of Mr. and Mrs. William T. Golden, and the comments of M. Ionifescu and S. H. Orzack.

Comment

Peter Donnelly

What a pleasure it is to see outlined one of the principal goals in applied probability, the elucidation of the *structure* common to a range of models that enjoy certain basic properties, followed by an exhilarating tour through that structure in the case in which the basic property is that of branching.

The application of these models in the context of genetics serves several purposes. On one level, it broadens our understanding of evolution, in this case through the illumination of a collection of conditions that are consistent with the molecular clock hypothesis. More generally, the contrast between the structure of the branching process models and that of more traditional population genetics models highlights the features of the latter which are fundamental consequences of the correlations in offspring numbers that arise through constraints on total population sizes.

Peter Donnelly is Professor of Mathematical Statistics and Operational Research in the School of Mathematical Sciences, Queen Mary and Westfield College, University of London, Mile End Road, London, E1 4NS, United Kingdom.

THE STRUCTURE OF GENETICS MODELS

In the neutral case, the structure of population genetics models is now well understood. In a population of fixed size N, which evolves in nonoverlapping generations, we could describe a specific model for the way in which the population reproduces by randomly labeling the individuals in a particular generation and specifying the joint distribution of the random variables $\nu_1, \nu_2, \ldots, \nu_N$, where ν_i is the number of offspring born to the ith individual. The random variables $\{\nu_i\}$ will be exchangeable, and