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),