$0) > 0$. 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

$\nu_1 + \nu_2 + \cdots + \nu_N = N$. Neutrality implies that the ν_i are independent of family sizes in previous generations, and we further assume that the distribution of offspring numbers does not change with time.

For neutral models, much of the interesting randomness is contained in the reproductive mechanism. This in turn is well, and it turns out, elegantly captured in a description of the genealogy or family tree structure induced by the model. With minor conditions on the moments of the offspring numbers ν_i, any model of the above type has, after a suitable rescaling of time, a genealogy that is well approximated for large N by a particular, nice, stochastic process called the coalescent. (See Kingman, 1982, for a description, and Donnelly and Joyce, 1992, for the general convergence result.)

To answer questions of genetic interest, one simply superimposes labels or genetic types on the "demography," or equivalently the genealogy, of the model. The nature of the labels will depend on the level of genetic detail being modeled, but enormous generality is possible. The key is that, because of neutrality, the assignment of labels does not interact with the genealogy.

The final component is a description of the way in which mutation works, and again great generality is possible. At an abstract level, this involves specifying a "mutation process," a Markov process on the set of possible labels. Think of this as describing the way in which the types of the individuals in a particular lineage change as we follow the lineage forward through (rescaled) time. It is conceptually easier, perhaps more realistic, and in fact effectively equivalent for these models, to assume that an individual's type may change throughout its life as a result of mutation, rather than only at birth. Superimposing the labels and mutation process onto the coalescent gives rise to an urn type mechanism, which, for example, specifies the joint distributions of the types in samples taken from the population (Ethier and Griffiths, 1987; Ethier and Kurtz, 1992; Donnelly and Kurtz, 1991).

A complementary approach describes the forward evolution of the population as a measure-valued process. Again, there is a particular measure-valued diffusion, called the Fleming-Viot process (Ethier and Kurtz, 1986), which arises as a robust description of any of the above exchangeable models (subject to the same minor conditions). The behavior of the Fleming-Viot process, of course, depends on the details of the mutation process and the type space (set of possible labels). There is a discrete representation of this process in terms of a particular interactive particle system in which genealogical processes and the urn mechanism are naturally embedded (Donnelly and Kurtz, 1991).

It is perhaps worth noting that all of the above also applies to most models with overlapping generations and variable population sizes provided that the variation in population size is not dependent on the genetic composition of the population.

This structure in turn raises a number of questions about the branching process models in the article. I wonder what can be said about the genealogy, or correlations in ancestry, of several individuals in stable branching populations. Genealogical ideas are proving useful in the measure-valued diffusions that arise from spatial branching processes. Is there a theory of stable populations in the spatial context? On more specifically genetic matters, how is the analysis affected if mutation rates depend on the individual's type? This is encompassed in the genetics framework above; it is dependence on features like age, as the author points out, that cannot be captured in the genetics models. While it is possible to analyze genetics models that incorporate selection and recombination, for example via a generalization of the Fleming-Viot process (Ethier and Kurtz, 1991), this is much less straightforward and their structure less transparent. The author notes that it is difficult to incorporate genuinely diploid features into the branching framework, but perhaps some forms of selection are easily modeled as multitype processes.

THE MOLECULAR CLOCK

In discussions of the molecular clock, there is a danger of confusing two different measures of the rate at which evolution occurs. Focus attention on a particular part of the genome in question—it may be a gene locus or possibly a single nucleotide site. If we were to watch the evolution of the population over time, we would observe many mutations, most of which (at least in the classical genetics models) are destined to be lost from the population relatively quickly. From time to time, however, one of these mutant types will sweep through the population and after some time the entire population will be of that type. The mutation is said to fix and a mutant substitution is said to have occurred. Define k_s, *the rate of evolution in terms of mutant substitutions* (Kimura, 1983), as the rate of such occurrences.

Alternatively, one could choose an individual at random from the current population and trace its ancestral lineage backwards in time. (All the models in question are haploid so that the individual will have exactly one parent, grandparent, great grandparent and so on). Define k_a, *the rate of*

evolution along an ancestral lineage, to be the rate at which mutations occur along this lineage. Symmetry considerations ensure that k_a does not depend on the individual chosen and, in fact, in genetics models, the whole population will share a common ancestor after a relatively small period of time (in evolutionary terms) in the past, after which the *actual lineages* of different individuals will be the same. (The period of time is of the order of the population size).

As they have been defined, it appears that the two evolutionary rates, k_s and k_a, are measuring different things. Let us explore their relationship further. Any mutation that fixes in the population must occur on the ancestral lineage in question, so that $k_a \geq k_s$. How many of the mutations on the ancestral line will actually fix in the population? Consider such a mutation, in the distant past. Its probability of fixing is the probability that a type initially present in one individual fixes in the population conditional on the fact that that individual is the one whose forward lineage survives the longest. Since fixation here implies that the individual's descendants outlive those of all its contemporaries,

$$\text{P(fixation of mutation on ancestral lineage)} = \frac{\text{P}\left(\begin{array}{c}\text{fixation of a type initially}\\ \text{present in one individual}\end{array}\right)}{\text{P}\left(\begin{array}{c}\text{a particular individual's}\\ \text{lineage survives the longest}\end{array}\right)}.$$

Symmetry arguments ensure that, in the neutral case, the denominator is N^{-1}, where N is the number of haploid individuals in the population. If (as is usual) it is assumed that mutations occur at a constant rate u to all types, then an analysis of the diffusion approximation (see Ewens, 1979, equations 4.17, 4.18, and 5.60) shows that the numerator above is $N^{-1}(1 - \theta)$ for $\theta < 1$ and 0 for $\theta \geq 1$, where θ is a scaled mutation rate. (θ is of the order of Nu with the exact scaling depending on the specifics of the reproductive mechanism being modeled). Thus

$$\text{P(fixation of a mutation on the ancestral lineage)} = \max(1 - \theta, 0).$$

In a study of molecular evolution, we would have focused on a single (or possibly three) nucleotide sites, and θ is likely to be very small, say of the order of 10^{-3}, so that $k_s \approx k_a$. (There is also dependence between the process of mutations along the ancestral lineage and the fate in the population of each mutation, but this too will be negligible for small θ.) In this context, then, the interchanging of k_s and k_a is not serious. Inferences concerning the latter require information only on the ancestors of sampled individuals rather than on the entire evolutionary history of the population. Of course, even this information is not directly available, and current inferential techniques rely on rather unrealistic additional assumptions. In studying the evolution of whole genes, however, θ may be of order 1, and so k_s and k_a are in fact measuring quite different things.

Observe also that while the individuals in the ancestral lineage are special—they are individuals whose (forward) lineage survives a long time—this is entirely for "demographic" reasons. If the mutations are neutral, the inclusion of an individual in the ancestral lineage is independent of its genetic type and of the process of mutations. It follows that $k_a = u$, the rate at which mutations occur to an individual. In particular, k_a *does not depend on either the population size or on the details of the reproductive mechanism of the population.* This conclusion remains valid even if mutation rates differ for different types (provided neutrality still obtains). Further, if the (maximum) mutation rate is small relative to the population size, $k_s \approx k_a$, and the same will be approximately true for k_s. This fact, that $k_s \approx u$ independent of the population size, is the original form of the molecular clock hypothesis. The usual derivation is different (see Kimura, 1983).

In branching process models, it is clear that k_s and k_a do measure different things. Indeed, fixation of a mutant type requires the extinction of all the subpopulations descending from the other individuals alive at the time of the mutant's birth. Late in the (supercritical) population's history, this will effectively never happen, and k_s will be zero. It is not unnatural then for Jagers' article to reinterpret the molecular clock hypothesis in terms of k_a.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge several helpful discussions with Warren Ewens and Geoff Watterson concerning the molecular clock hypothesis.