The Coalescent; BBI, Singapore

Chris Cannings,University of Sheffield

19.05,2011

The Wright-Fisher model of "genetic drift" considers a population of fixed size N, with non-overlapping generations. Each individual in the population at time $t+1$ is the offspring of any particular individual in the population at time t with probability $1/N$, (mutually) independently of the parentage of other individuals. It is possible in this and in a wider class of models (Cannings,1974) to specify the eigenvalues and eigenvectors in a fairly full fashion and thus to study various aspects of the process.

Kingman(1982) introduced a major insight for the study of such processes, the Coalescent. Instead of looking at whole generations with time running forward the coalescent runs time backward. Since the number of individuals in generation t who are actually, rather than potentially, parents of some set of k individuals in generation $t+1$, is $\leq k$ (with non-zero probability for $< k$), the ancestry of any set of individuals is a tree running backwards in time to a MRCA (most recent common ancestor of the set). Now the study of the genetic drift process reduces to the study of this tree. A brief survey this process will include the probabilities of various tree topologies, and the time to the MRCA.

The coalescent approach will then be used to tackle two problems.

(1) Suppose that individuals have a type specified by an integer, and that a parent of type x produces an offspring of type $x-1$, x or $x+1$ with probabilities $\mu/2$, $1-\mu$ and $\mu/2$. We wish to specify aspects of the distribution of the types in the population at some time n. This model correspond to genetic situations in which an individual has some number of copies of a genetic unit, and that due to errors in the copying process required for producing an offspring, that offspring may have a slightly different number of copies. We derive certain results for the moments of this and a related normalised process.

(2) The branch from an individual backwards in time to the coalescent tree is called an external edge. Caliebe et al (2007) derived some asymptotic results regarding the distribution of the length of an external edge. Here some further results regarding the distribution of the lengths of such edges for trees of finite size will be derived.

References

[1] Caliebe A,Neininger R, Krawczak M, Rosler U. (2007) ON the length distribution of external branches in coalescent trees: genetic diversity within species. Theoret Popul Biol. 72, 245-252.

[2] Cannings,C. (1974) The latent roots of certain Markov chains arising in genetics; a new approach. 1. Haploid models. Adv.Appl.Prob., 6, 260–290.

[3] Kingman,JFC.(1982) The coalescent. Stochastic processes and their applications,13,235-248.