Abstract

Most recent common ancestors, genetic inheritance, stochastic gene transcription, and Brownian motion on disconnected sets: a probabilistic analysis of a few models.

by

Peter Lochhead Ralph
Doctor of Philosophy in Statistics

University of California, Berkeley

Professor Steven N. Evans, Chair

In Part I we build from certain Poisson processes a general class of processes that are analogous to the temporal dynamics of the time back to the most recent common ancestor in an asexually reproducing population. For these processes we derive transition probabilities and other detailed information about their sample paths. Special cases of this construction describe the time evolution of family ages in $(1+\beta)$-stable continuous-state branching processes conditioned on nonextinction, which we find to have a stationary description related to the GEM distribution.

In Part II, we study gene flow and inheritance in a continuous-time model of a large, randomly mating population of sexually reproducing organisms. We define a mechanism that tracks inheritance patterns for a general class of recombination mechanisms both forwards and backwards in time, construct measure-valued processes in both time directions, and identify the deterministic large-population-size limits, which are described in terms of dual branching processes. We also provide a joint construction of both processes that displays an interesting duality between the two.

The transcription of genes into RNA is a complex process that can depend on many different regulatory compounds in complex ways, and often must display precise, reliable behavior, despite inherent stochasticity in the molecular processes. In Part III, we introduce a framework for modeling the time until transcription occurs as the first-passage time in a continuous-time Markov chain. To provide practical tools for comparison of different regulatory network configurations when the network is large, we develop a decomposition method based on separating the chain into a sequence of smaller chains, from which the distribution of the first passage time can be reconstructed.

In Part IV, motivated by Lévy's characterization of one-dimensional Brownian motion on the line, we study an analogue of Brownian motion whose state space is an arbitrary unbounded closed subset of the line. We then consider the case where the state space is the self-similar set $\{\pm q^k : k \in \mathbb{Z}\} \cup \{0\}$ for some $q > 1$, and using continued fractions appearing in Ramanujan's "lost"

notebook, find explicit formulae in terms of basic hypergeometric functions for many properties of the process.

Contents