

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

I	Dynamics of the most recent common ancestor in a critical branching population	1
1	Introduction	2
2	General lifetime measures	10
3	Duality, time-reversal, and the peak and trough chains	16
3.1	Duality and time-reversal	16
3.2	The peak and trough jump chains	19
4	Genealogical properties of the $(1 + \beta)$ -stable CSBP	24
4.1	The $(1 + \beta)$ -stable MRCA process	24
4.2	The sequence of split times	27
II	Gene flow in biparental models of inheritance	33
5	Introduction	34
5.1	Previous work	37
6	Genetic descendants and ancestral splitting	41
6.1	Definition of the measure-valued processes	41
6.2	Trees and root systems	43
6.3	Limiting flows	44
6.4	Joint definition	46
7	Convergence results	52
7.1	Generators	52
7.2	Finite-time convergence	56
III	Stochastic transcription and traversal times of Markov chains	59
8	Introduction	60
8.1	Previous work	62
9	Markov models of protein-DNA interactions	64
9.1	Model construction	64
9.2	Computing first-passage times of Markov chains	65
9.3	Measuring noisiness of assembly	66

9.4	Application to the effects of gene regulation on noise	68
10	Detailed regulatory models	71
11	Decomposition into sequential modules	75
11.1	Computing system noise properties	77
11.2	Transition counts for the pinch-chain	81
11.3	Transition probabilities between modules	81
11.4	Traversal times within modules	83
11.5	Inverses and singular matrices	85
12	Noise properties in interacting parallel chains	87
12.1	The general case	87
12.2	A simple approximation for parallel two-state enhancers	89
12.3	Moments for parallel two-state chains	90
IV	Brownian motion on disconnected sets, basic hypergeometric functions, and some continued fractions of Ramanujan	92
13	Introduction	93
14	Brownian motion on a general unbounded closed subset of \mathbb{R}	97
14.1	Existence	97
14.2	Uniqueness	101
14.3	Reversibility	104
15	Hitting times of bilateral birth-and-death processes	106
15.1	Hitting times on a scattered subset of \mathbb{T}	108
15.2	Introducing the process on \mathbb{T}_q	110
15.3	Hitting time distributions for \mathbb{T}_q	112
15.4	Excursion theory for \mathbb{T}_q	117
15.5	Resolvent of the killed process on $\mathbb{T}_q \cap (0, \infty)$	120
15.6	Resolvent for \mathbb{T}_q	121
15.7	A remark on spectral representations	121
16	Background on basic hypergeometric functions, recurrence relations, and continued fractions	123
16.1	Background on basic hypergeometric functions	123
16.2	Background on recurrence relations and continued fractions	125
	Bibliography	129