Poisson-Kingman Partitions

Jim Pitman

Abstract

This paper presents some general formulas for random partitions of a finite set derived by Kingman's model of random sampling from an interval partition generated by subintervals whose lengths are the points of a Poisson point process. These lengths can be also interpreted as the jumps of a subordinator, that is an increasing process with stationary independent increments. Examples include the two-parameter family of Poisson-Dirichlet models derived from the Poisson process of jumps of a stable subordinator. Applications are made to the random partition generated by the lengths of excursions of a Brownian motion or Brownian bridge conditioned on its local time at zero.

Keywords: exchangeable; stable; subordinator; Poisson-Dirichlet; distribution

1 Introduction

This paper presents some general formulas for random partitions of a finite set derived by Kingman's model of random sampling from an interval partition generated by subintervals whose lengths are the points of a Poisson point process. Instances and variants of this model have found applications in the diverse fields of population genetics [17, 19], combinatorics [4, 48], Bayesian statistics [23], ecology [15, 37], statistical physics [11, 12, 13, 53, 55], and computer science [25].

Section 2 recalls some general results for partitions obtained by sampling from a random discrete distribution. These results are then applied in Section 3 to the Poisson-Kingman model. Section 4 discusses three basic operations on Poisson-Kingman models: scaling, exponential tilting, and deletion of classes. Section 5 then develops formulas for specific examples of Poisson-Kingman models. Section 6 recalls the two-parameter family of Poisson-Dirichlet models derived in [50] from the Poisson process of jumps of a stable(α) subordinator for $0 < \alpha < 1$. Section 7 reviews some results of [41, 46, 49, 50] relating the two-parameter family to the lengths of excursions of a Markov process whose zero set is the range of a stable subordinator of index α. Section 8 provides further detail in the case $\alpha = \frac{1}{2}$ which corresponds to partitioning a time interval by the lengths of excursions of a Brownian motion. As shown in [2, 3], it is this stable($\frac{1}{2}$) model which governs the asymptotic distribution of partitions derived in various ways from random forests, random mappings, and the additive coalescent. See also [5, 9] for further developments in terms of Brownian paths, and [10, 25] for