The block containing 1 for exchangeable coalescent processes

Sprecher: Dr. Fabian Freund

Eingeladen von Prof. Dr. Markus Pauly

23.06.2017 | 14:30 Uhr | R220 HeHo 18

We consider exchangeable coalescent processes, which are Markovian processes \( \{ \Pi_t \}_{t \geq 0} \). Their state space is the set of the partitions of \( \mathbb{N} \) and all transitions are mergers of partition blocks. We focus on the behavior of the partition block \( B_1(t) \) at time \( t \) which contains \( 1 \in \mathbb{N} \). Let \( f_1 = (f_1(t))_{t \geq 0} \) be the asymptotic frequency process of 1 defined by \( f_1(t) := \lim_{n \to \infty} n^{-1} |B_1(t) \cap [n]| \), where \([n] := \{1, \ldots, n\} \) and \(|A|\) denotes the cardinality of set \( A \). For coalescent processes with dust, i.e. where \( t \) has a positive fraction of \( i \in \mathbb{N} \) as singleton blocks \( \{i\} \) with positive probability, \( f_1 \) is a jump-hold process which can be described via a stick-breaking procedure with uncorrelated stick lengths with common mean. We provide a closer look at the law of the first jump of \( f_1 \) for a subclass of coalescents with dust.

Consider the restriction of the coalescent process to \([n]\) by intersecting all partition blocks with \([n]\). The restriction, the (Markovian) \( n \)-coalescent, can be interpreted as a random tree and used as a model for a gene genealogy of a sample of DNA sequences \([n]\), whose mutations are modeled by a homogeneous Poisson point process with rate \( \theta \) on the branches of this tree. We consider the size \( O_n \) of the partition block containing 1 in the \( n \)-coalescent at a random time \( T_n \), given by the sum of the waiting time for the first merger of 1 and an independent exponential time with rate \( \theta \). \( O_n \) can be interpreted as the size of the smallest (non-trivial) family of sequence 1 which can be distinguished by the sampled sequences. We analyze the distribution of \( O_n \) for different \( n \)-coalescents and its asymptotics for \( n \to \infty \). We end with a discussion about the use of \( O_n \) as a test statistic to distinguish between differently distributed \( n \)-coalescents with mutation, modeling different evolutionary scenarios.