THE UNIVERSITY OF TEXAS AT AUSTIN

Mathematics & Statistics Colloquium

When: Wednesday, April 30, 2:00 pm - 2:50 pm

Where: Lee Drain Building 216

Size-biased permutation of a finite i.i.d sequence

Ngoc Tran
Simons postdoctoral fellow
Department of Mathematics
University of Texas at Austin

Line up n blocks of lengths X_1 ... X_n, where the X_i's are independent and identically distributed (i.i.d) positive random variables. Throw a ball uniformly at random on the interval 0 to X_1 + ... + X_n, and record length of the block X_i that the ball falls in, and then remove this block. This is called size-biased sampling, since the bigger blocks are more likely to be discovered earlier. Do this recursively n times. This yields a size-biased permutation of the finite i.i.d sequence (X_1, ... X_n).

This model is known as Kingman's paintbox, put forward by Kingman in 1976 to study random infinite partitions. In the infinite version of the above model, the X_i's are jumps of a subordinator. In 1992, Perman, Pitman and Yor derived various distributional properties of this infinite size-biased permutation. Their work found applications in species sampling, oil and gas discovery, topic modeling in Bayesian statistics, amongst many others.

In this talk, we will derive distributional properties of finite i.i.d size-biased permutation, both for fixed and asymptotic n. We have multiple derivations using tools from Perman-Pitman-Yor, as well as the induced order statistics literature. Their comparisons lead to new results, as well as simpler proofs of existing ones. Our main contribution describes the asymptotic distribution of the last few terms in a finite i.i.d size-biased permutation via a Poisson coupling with its few smallest order statistics. For example, we will answer the question: what is the asymptotic probability that the smallest block is discovered last? This is joint work with Jim Pitman.