

Berkeley Algebraic Statistics Seminar

Organizer(s): and Andrew Critch and Shaowei Lin

Wednesday, 2:00–3:00pm, 939 Evans

Sep 28 **Andrew Critch**, UC Berkeley

How to find a hidden Markov process

For statisticians: sometimes nature hides variables from us... how can a human or artificial intelligence check whether that's happening? One way is to use algebraic geometry! In this talk I'll introduce hidden Markov models as algebraic varieties, and how to express their defining ideals using the *cumulant coordinates* of Sturmfels-Zwiernik (2011). Generators for the ideals of these models provide empirical tests for whether a data source is a actually hidden Markov process. HMMs are used in temporal pattern recognition such as speech and gesture recognition, musical score following, machine translation, alignment of bio-sequences, and metamorphic virus detection, and it would be nice to find out where else!

For algebraic geometers: the n th binary HMM is a particular rational projective variety in \mathbb{P}^{2^n-1} . It is always 5 dimensional, and its rational parametrization comes from a generative statistical process I'll describe in the talk. Implicitizing these varieties is a difficult task, but a non-linear birational map $\mathbb{P}^{2^n-1} \rightarrow \mathbb{P}^{2^n-1}$ called the *cumulant map* greatly simplifies problem, which I'll show for $n = 3$ and 4.