

## Statistics and Data Science Seminar

### *Distributions of pattern statistics in sparse Markov models*

Donald E.K. Martin (North Carolina State University)

**Abstract:** Higher-order Markov models provide a good approximation to probabilities associated with many categorical time series, and thus they are applied extensively. However, a major drawback associated with them is that the number of model parameters grows exponentially in the order of the model, and thus only very low-order models are considered in applications. Another drawback is lack of flexibility, in that higher-order Markov models give relatively few choices for the number of model parameters. Sparse Markov models are Markov models where transition probabilities are lumped into classes comprised of invariant probabilities. The contexts for conditioning may be either hierarchical (as in variable length Markov chains) or non-hierarchical. This supplies a model that helps with the two problems given above, and which thus gives a better handling of the trade-off between bias associated with having too few model parameters and variance associated with having too many. In this work, methods for efficient computation of pattern distributions through Markov chains with minimal state spaces are extended to the sparse Markov framework.

Wednesday, August 29 at 4:00 PM in 636 SEO