GENERALIZED WEIGHTED CHINESE RESTAURANT PROCESSES
FOR SPECIES SAMPLING MIXTURE MODELS

Abstract

This talk describes a class of “species sampling” mixture models which serve as an extension of nonparametric mixture models based on the Dirichlet processes to models based on a much more general family of prior distributions. This is inspired by the recent work of Lo, Brunner, and Chan (1999) on what they call a weighted Chinese restaurant process, a sequential seating algorithm which generates random partitions of the data. Their method can be used to provide an iid Monte Carlo approximation to posterior quantities based on the Dirichlet process.

One of our contributions is an extension of their method to encompass all priors based on species sampling prediction rules. These schemes potentially have utility in population genetics, mathematical biology, and other areas where random partition models of this type have been used. To rigorously justify these results, the theoretical work of Lo (1984) on Bayesian density and mixture model estimation is extended to this more general setting. Our approach relies on the work of Pitman (1995, 1996) on exchangeable species sampling sequences, their prediction rules, and the de Finetti characterizations of their corresponding laws.

(This is joint work with Hemant Ishwaran of the Cleveland Clinic Foundation and is supported by the Acheson J. Duncan Fund for the Advancement of Research in Statistics.)