



The University of Chicago
Department of Statistics

Ph.D. Seminar

JIE YANG

Department of Statistics
The University of Chicago

**“Infinite Exchangeability & Partitions
and Permanent Process & Classification Model”**

**TUESDAY, July 18, 2006 at 3:00 PM
110 Eckhart Hall, 5734 S. University Avenue**

ABSTRACT

In Part I: Infinite Exchangeability and Partitions, we develop a partition model with applications to multiple comparisons and cluster analysis. Unlike an ordinary Bayesian setup, we construct an infinitely exchangeable variety process and assign positive probability to each partition of the varieties. Using this process as a prior in a Gaussian model, we obtain inferences in the form of a posterior distribution on partitions. For typical multiple comparison applications, we suggest the Ewens family as a class of prior distributions on partitions with parameter in the range roughly 1-4. We also give inference for variety contrasts from the partition model, which allows positive probabilities for the events that two or more varieties are equal. For application to cluster analysis, we develop MCMC algorithms to estimate summary statistics, especially the similarity matrix.

In Part II: Permanent Process and Classification Model, we develop a classification model based on the permanent process introduced recently. In the model, there are only 2-3 estimable parameters, regardless of the number of classes or the dimension of the feature space. The model works well even if the classes occupy non-convex regions or disconnected regions in the feature space. Under the model, we express the conditional distribution of the class of a subsequent unit given the training data in terms of ratios of weighted permanents. We propose an analytic approximation for the weighted permanent ratios based on the cycle expansion of the weighted permanent. Our experience is that the approximation usually has acceptable error for typical classification problems.