

STATISTICS COLLOQUIUM

FRIDAY, FEBRUARY 23, 2007

TALK: 4:00 PM — SCIENCE CENTER RM. 705

RECEPTION: 5:15 PM — SCIENCE CENTER, 7TH FLOOR

“The nested Dirichlet process”

Abel Rodriguez

Institute of Statistics and Decision Sciences

Duke University

ABSTRACT

In multicenter studies, subjects in different centers may have different outcome distributions. This article is motivated by the problem of nonparametric modeling of these distributions, borrowing information across centers while also allowing centers to be clustered. Starting with a stick-breaking representation of the Dirichlet process (DP), we replace the random atoms with random probability measures drawn from a DP. This results in a nested Dirichlet process (nDP) prior, which can be placed on the collection of distributions for the different centers, with centers drawn from the same DP component automatically clustered together. Theoretical properties are discussed, and an efficient MCMC algorithm is developed for computation. The methods are illustrated using a simulation study and an application to quality of care in US hospitals.