CASE WESTERN RESERVE UNIVERSITY EST. 1826

Department of Epidemiology and Biostatistics
Biostatistics Seminar

Thursday, September 25, 2014
12:00pm - 1:00pm -- WG73

"Modeling relational data using nested infinite relational models"

Dr. Kaushik Ghosh, Associate Professor
Department of Mathematical Sciences, University of Nevada Las Vegas

Abstract: We introduce a flexible class of models for relational data based on hierarchical extension of the two-parameter Poisson-Dirichlet process. The model is constructed by partitioning the original data matrix into nested blocks corresponding to objects with similar behavior, with the number of blocks estimated from the data. The partitions so generated provide a sparse representation of the data and allow for sharing of information across rows and columns, leading to improved estimation. The model is illustrated with two applications: 1) a study of cancer mortality rates in the U.S., where rates for different cancer types are available for each state, and 2) analysis of microarray data, where expression levels for a large number of genes are available in a sample of subjects.

(Joint work with Abel Rodriguez, UC Santa Cruz.)