Statistics Seminar

Department of Mathematical Sciences

DATE:	Thursday, February 21, 2019
TIME:	1:15pm – 2:15pm
LOCATION:	WH 100E
SPEAKER:	Fang Yuan, Binghamton University
TITLE:	Variable selection in clustering via Dirichlet process mixture models

Abstract

Variable selection in clustering via Dirichlet process mixture models Abstract: The increased collection of high-dimensional data in various fields has raised a strong interest in clustering algorithms and variable selection procedures. In this paper, we propose a model-based method that addresses the two problems simultaneously. We introduce a latent binary vector to identify discriminating variables and use Dirichlet process mixture models to define the cluster structure. We update the variable selection index using a Metropolis algorithm and obtain inference on the cluster structure via a split-merge Markov chain Monte Carlo technique. We explore the performance of the methodology on simulated data and illustrate an application with a DNA microarray study.