Stanford University
Department of Statistics

DEPARTMENTAL SEMINAR

4:15pm, Tuesday, February 24, 2015
*** Note Special Location ***
McCullough Building 04-490 Room 115
Cookies served at 3:45pm, 1st floor Lounge.

Speaker: Rachel Wang
Department of Statistics,
UC Berkeley

Title: Problems in network modeling: Inferring edges and community detection

Abstract:

Networks pervade many disciplines of science as a way of analyzing complex systems with interacting components. The problem of network modeling is often two-fold. First, the relationships between pairs of nodes, if not directly observed, have to be estimated from data. Based on the estimated (or given) network topology, various statistical and computational tools can then be applied to extract interesting patterns such as the presence of communities. In this talk I will present two studies related to both parts of the problem. I will first discuss a study in the context of gene regulatory networks, where the goal is to infer gene interactions using expression data with large and heterogeneous samples. We propose two gene coexpression statistics based on counting patterns of expression ranks to account for local gene associations. Moving on to general networks, I will then discuss model selection for the stochastic block model (SBM), which is a popular tool for community detection. We consider an approach based on the log likelihood ratio statistic and analyze its asymptotic properties under model misspecification. The results enable us to derive the correct order of the penalty term for model complexity and arrive at a likelihood-based model selection criterion that is asymptotically consistent and valid also in the semi-sparse regime.