

STATISTICS SEMINAR

Cen Wu, Ph.D.

Yale University

Tuesday, December 16, 2014

Bluemont Hall, Room 122, 4:00-5:30 pm

Refreshments: Dickens 108, 3:30-4:00 pm



A Robust Network-Constrained Penalization Approach for Integrative Analysis with Applications in TCGA Data

Integrative analysis of multiple genetic features has become increasingly important, partly due to the unique perspective that it offers to elucidate the complicated regulation mechanism among the features. Although heavy-tailed errors and outliers in response variables of the integrative analysis have been commonly observed, the issue has not been taken care of in existing studies. In this talk, I will present a robust network-constrained penalization approach for the integrative analysis of two genetic features with high dimensional multivariate measurements. Consider copy number alterations (CNAs) and gene expressions (GE). Both of them play pivotal roles in the development of complex diseases. A semi-parametric modelling strategy is taken to flexibly model the effects of multiple CNAs on multiple GEs. We choose the least absolute deviation (LAD) loss function to tackle data contamination and heavy-tailed errors in GEs. A network-constrained penalty is adopted to accommodate the network adjacency accounting for the correlations. Sparse and biologically meaningful associations are identified by the penalization procedure which can be effectively implemented within the coordinate descent framework. The proposed approach outperforms the alternatives in simulation study. We carry out the integrative analysis with TCGA (The Cancer Genome Atlas) data on melanoma, and demonstrate the advantage of the proposed approach.