

The Medical College of Wisconsin

Division of Biostatistics

Proudly Presents:

A Seminar Talk

By:



Duo Jiang, PhD

Assistant Professor in Statistics, Oregon State University

Efficient Gene-Set Interpretation of Expression Data using Mixed Models

Competitive gene-set analysis, also called enrichment analysis, is a widely used tool for functional interpretation of high throughput biological data such as gene expression data. It aims at testing a known category (e.g. a pathway) of genes for enriched differential expression (DE) signals compared to genes not in the category. Most enrichment testing methods ignore the widespread correlations among genes, which has been shown to result in excessive false positives. We show, both theoretically and empirically, that existing methods to account for correlations, such as GSEA and CAMERA, can result in severely mis-calibrated type 1 error and/or substantial power loss due to the failure to properly accommodate the DE heterogeneity across genes. We propose MEACA, a new gene-set testing framework based on a mixed effects model. Our method flexibly incorporates the unknown distribution of DE effects, effectively adjusts for completely unknown, unstructured correlations among genes, and does not rely on time-consuming permutations. Compared to existing methods, MEACA enjoys robust type 1 error control in widely ranging scenarios and improves power.

Tuesday, March 27, 2018

3:30 PM – 4:30 PM

Medical College of Wisconsin

MEB M2050

Light snacks provided