

Department of
Biostatistics and Medical Informatics

SEMINAR



Hongzhe Li, PhD

Professor of Biostatistics and Statistics
Chair, Graduate Program in Biostatistics
Director, Center for Statistics in Big Data (CSBD)
University of Pennsylvania

Friday, September 29, 2017

12:00 - 1:00 p.m.

Biotech Auditorium

Methods for High Dimensional Compositional Data Analysis with Applications in Microbiome Studies

Abstract: Human microbiome studies using high throughput DNA sequencing generate compositional data with the absolute abundances of microbes not recoverable from sequence data alone. In compositional data analysis, each sample consists of proportions of various organisms with a unit sum constraint. This simple feature can lead traditional statistical treatments when naively applied to produce errant results and spurious correlations. In addition, microbiome sequence data sets are typically high dimensional, with the number of taxa much greater than the number of samples. These important features require further development of methods for analysis of high dimensional compositional data. This talk presents several latest developments in this area, including two-sample test for compositional vectors, regression analysis with compositional covariates and covariance estimation based on compositional data. Several real microbiome studies are used to illustrate these methods and several open questions will be discussed.



School of Medicine
and Public Health

UNIVERSITY OF WISCONSIN-MADISON