



The Chinese University of Hong Kong
Department of Statistics

Seminar

Zero-Inflated Generalized Dirichlet Multinomial (ZIGDM) Regression Model for Microbiome Compositional Data

By

Dr. Zhengzheng TANG
Department of Biostatistics and Medical Informatics
University of Wisconsin Madison

Abstract

There is heightened interest in using high-throughput sequencing technologies to quantify abundances of microbial taxa and linking the abundance to human diseases and traits. Proper modeling of multivariate taxon counts is essential to the power of detecting this association. Existing models are limited in handling excessive zero observations in taxon counts and in flexibly accommodating complex correlation structures and dispersion patterns among taxa. We develop a new probability distribution, Zero-Inflated Generalized Dirichlet Multinomial (ZIGDM), that overcomes these limitations in modeling multivariate taxon counts. Based on this distribution, we propose a ZIGDM regression model to link microbial abundances to covariates (e.g. disease status), and develop a fast Expectation-Maximization (EM) algorithm to efficiently estimate parameters in the model. The derived tests enable us to reveal rich patterns of variation in microbial compositions including differential mean and dispersion. The advantages of the proposed methods are demonstrated through simulation studies and an analysis of a gut microbiome dataset.

Date: April 2, 2019 (Tuesday)
Time: 2:30 p.m. - 3:30 p.m.
Venue: Science Centre L2
The Chinese University of Hong Kong

ALL INTERESTED ARE WELCOME !!