

**Indiana University-Purdue University
Indianapolis**
Department of Mathematical Sciences

STATISTICS SEMINAR

12:15pm—1:15pm, Tuesday, March 30, 2021
Zoom Meeting: Meeting ID: 751 025 519

Speaker: Liangliang Zhang

Department of Biostatistics, University of Texas MD Anderson Cancer Center

Title: Bayesian variable selection for high dimensional data
with complex structures

Abstract:

Due to the emergence of next-generation sequencing techniques, which enable comprehensive profiling of high-throughput sequencing data, there is growing interest in developing novel methods of variable selection for high dimensional data with complex structures. The applicability of standard variable selection methods is complicated by the main challenging characteristics of sequencing data, including compositionality and dependency. To enable sparse regression modeling with sequencing predictors, I proposed a novel zero-constrained prior to handle the compositionality. The zero-constrained prior consistently controls the summation of random coefficients to zero across the model space. To account for the dependency between variables, the model utilized the Ising prior to encourage the joint selection of variables that are similar to each other. I applied this model into a real data set to link patients' body mass index to their gut microbiome sequences. In the end, I extended the Bayesian linear regression framework to Bayesian logistic regression and survival models.

Bio:

Liangliang Zhang is currently a postdoctoral fellow working in the Department of Biostatistics at The University of Texas MD Anderson Cancer Center. He received his PhD degree in Statistics from the Department of Statistics and Probability at Michigan State University. His research interests focus on Bayesian inference and prediction, model and feature selection, image data analysis and microbiome data analysis.