Indiana University-Purdue University Indianapolis

Department of Mathematical Sciences

Statistics Seminar

12:15pm—1:15pm, Tuesday, January 17, 2023
Zoom Meeting: Meeting ID: 845 0989 4694

Speaker: **Jinghang Lin**
Quantitative Analytics Specialist at Wells Fargo

Title: **Expectile Neural Networks for Genetic Data Analysis of Complex Disease**

Abstract:

The genetic etiologies of common diseases are highly complex and heterogeneous. Classic methods, such as linear regression, have successfully identified numerous variants associated with complex diseases. Nonetheless, for most diseases, the identified variants only account for a small proportion of heritability. Challenges remain to discover additional variants contributing to complex diseases. Expectile regression is a generalization of linear regression and provides complete information on the conditional distribution of a phenotype of interest. While expectile regression has many nice properties, it has been rarely used in genetic research. In this paper, we develop an expectile neural network (ENN) method for genetic data analyses of complex diseases. Similar to expectile regression, ENN provides a comprehensive view of relationships between genetic variants and disease phenotypes and can be used to discover variants predisposing to sub-populations. We further integrate the idea of neural networks into ENN, making it capable of capturing non-linear and non-additive genetic effects (e.g., gene-gene interactions). Through simulations, we showed that the proposed method outperformed an existing expectile regression when there exist complex genotype-phenotype relationships. We also applied the proposed method to the data from the Study of Addiction: Genetics and Environment(SAGE), investigating the relationships of candidate genes with smoking quantity.

Bio:
Dr. Jinghang Lin got his PhD in statistics at Michigan State University in 2021. After he completed his one-year postdoctoral training at Yale University, he joined Wells Fargo as quantitative analytics specialist. His research field is statistical genetics, machine learning and high dimensional statistics.