Indiana University-Purdue University Indianapolis
Department of Mathematical Sciences

STATISTICS SEMINAR

12:15pm—1:15pm, Tuesday, November 09, 2021
Zoom Meeting: Meeting ID: 845 0989 4694

Speaker: Tao He
Department of Mathematics, San Francisco State University

Title: **An Adaptive Multivariate Kernel-based Test for Association with Multiple Quantitative Traits in High-dimensional Data**

Abstract:

Identifying genetic variants associated with a complex disease has benefited from recent advances in set-based and multi-trait testing methods. Jointly testing sets of variants (e.g., those corresponding to gene sets or pathways) for association with a set of phenotypic disease traits can be more powerful than testing individually with single variants or single traits. A common challenge in set-based testing is that sample size may be small and vastly exceeded by the number of variants. To overcome this, we develop an adaptive kernel-based test that uses a supervised method to filter out noncausal variants and reduce dimension. Our test does not rely on assuming a particular functional form for the association or a particular distribution for the traits, and allows the genetic effect on each trait to be captured using a different kernel function, employing a fast kernel selection method based on asymptotic results under a high-dimensional setting. We demonstrate the speed of our method written in C++ and implemented in an R package, and we compare its effectiveness against other kernel-based tests in both simulations and applications to real-world data.

Bio:

Tao He is an Assistant Professor of Statistics in the Department of Mathematics at San Francisco State University. She received his PhD in Statistics and dual PhD in Quantitative Biology from Michigan State University in 2015. Currently she also serves as the Past President of San Francisco Bay Area Chapter of American Statistical Association. Dr. Hes research interests include statistical genetics/genomics, high-dimensional statistical inference, non- and semi-parametric models, statistical learning and their applications in biomedical research.