ABSTRACT:

Two major research tasks lie at the heart of high dimensional data analysis: accurate parameter estimation and correct variable selection. Existing literature mostly aims for either the best parameter estimation or the best model selection behavior, however little has been done to understand the potential interaction between the estimation precision and the selection behavior. In this work, our minimax result shows that an estimator’s $L_2$ error rate critically depends on its performance of type I error control, and reveals a trade-off phenomenon between the rate of convergence and the false discoveries control: better estimation accuracy leads to more false discoveries. In particular, we characterize the false discovery control behavior of rate-optimal and rate-suboptimal estimators under different sparsity regimes, and discover a rigid dichotomy between these two estimators under near-linear and linear sparsity settings. In addition, this work provides a rigorous explanation to the incompatibility phenomenon between selection consistency and rate-minimaxity which has been frequently observed in the high dimensional literature.

ABOUT THE SPEAKER:

Qifan Song, Assistant Professor, Department of Statistics, Purdue University. Dr. Song obtained his Ph.D. in Statistics from Texas A&M University in 2014, B.S. in Statistics from Peking University in 2009. Dr. Song’s research interests include high dimensional data inference; variable selection from both Bayesian and frequentist perspectives; big data analysis: parallel Bayesian analysis for massive data set; Bayesian Computation: convergence and ergodicity studies for adaptive MCMC. Dr. Song has published in top journals.