MICHIGAN STATE UNIVERSITY
Department of Statistics and Probability

JAMES FRANCIS HANNAN LECTURE SERIES

Peter Bühlmann
ETH Zürich

Inhomogenous Large-Scale Data: New Opportunities for Causal Inference and Prediction

Thursday, November 17, 2016
3:00 p.m. – 3:50 pm
B122 Wells Hall

Abstract

Large-scale or "big" data usually refers to scenarios with potentially very many variables (large dimension) and large sample size. Such data is most often of "inhomogeneous" nature, i.e., neither being random samples from a common population nor being generated from a stationary distribution. We discuss how to exploit the advantage of heterogeneity in large datasets. A key ingredient is an invariance principle that leads to new approaches for causal inference and novel prediction methods, which exhibit "robustness" even for scenarios not present in the observed data. As a concrete application, we discuss large-scale gene knock-down experiments in yeast (Saccharomyces Cerevisiae) where computational and statistical methods have an interesting potential for prediction and prioritization of new experimental interventions.

To request an interpreter or other accommodations for people with disabilities, please call the Department of Statistics and Probability at 517-355-9589.