

NYU Stern School of Business
Department of Information, Operations & Management Sciences
STATISTICS RESEARCH SEMINAR

TOPIC: CARS: A Covariate Assisted Approach to Large-Scale Two-Sample Inference

SPEAKER: Wenguang Sun (USC)

DATE: Friday, November 13, 2015

TIME: 11:30 AM - 12:30 PM

PLACE: KMC 4-80

Abstract

The conventional practice in two-sample multiple testing involves first reducing the data matrix to a vector of p-values and then choosing a cutoff along the rankings to adjust for multiple comparison. However, this inference framework often leads to suboptimal multiple testing procedures due to the loss of information in the data reduction step. In this talk I discuss how to construct an auxiliary variable from the original data matrix and incorporate the variable in the inference procedure to improve the power in multiple testing. We study the problem in a decision-theoretic framework and develop oracle and data-driven procedures for false discovery rate (FDR) control. The proposed oracle procedure employs a covariate-assisted ranking and screening (CARS) scheme, which is shown to be optimal in the sense that it has the smallest missed discovery rate among all valid FDR procedures. We then develop a data-driven procedure to mimic the oracle procedure and establish its asymptotic properties. Numerical studies show that the proposed method outperforms existing methods in the literature. Applications to satellite imaging and time-course data analysis will be discussed if time permits.

Bio

Wenguang Sun got his PhD from University of Pennsylvania in 2008. He is currently an associate professor in the department of data sciences and operations at Marshall School of Business, University of Southern California. His research interests include large-scale multiple testing, high-dimensional statistics and statistical decision theory.