



香港中文大學統計學系

Department of Statistics

THE CHINESE UNIVERSITY OF HONG KONG

## DISTINGUISHED LECTURE

# On Methods for Controlled Variable Selection in Linear, Generalized-linear, and Index Models



Professor **Jun S LIU**

Department of Statistics  
Harvard University

Elected Fellow of International Society of Computational Biology,  
2022

Distinguished Achievement Award, International Chinese Statistical  
Association, 2012

Elected Fellow of Institute of Mathematical Statistics, 2005

Elected Fellow of American Statistical Association, 2004

The 2002 COPSS Presidents' Award

Date: 29 May 2023 (Monday)

Time: 2:30 pm — 3:30 pm

Venue: LT2, Lady Shaw Building,  
The Chinese University of Hong Kong  
(Tea reception after the talk)

### Abstract

A classical statistical idea is to introduce data perturbations and examine their impacts on a statistical procedure. In the same token, the knockoff methods carefully create “matching” fake variables in order to measure how real signals stand out. I will discuss some recent investigations we made regarding both methodology and theory on a few related methods applicable to a wide class of regression models including the knock off filter, data splitting (DS), Gaussian mirror (GM), for controlling false discovery rate (FDR) in fitting *linear*, *generalized linear* and *index* models. We theoretically compare, under the weak-and-rare signal framework for linear models, how these methods compare with the oracle OLS method. We then focus on the DS procedure and its variation, *Multiple Data Splitting* (MDS), which is useful for stabilizing the selection result and boosting the power. DS and MDS are straightforward conceptually, easy to implement algorithmically, and applicable to a wide class of linear and nonlinear models. Interestingly, their specializations in GLMs result in *scale-free procedures* that can circumvent difficulties caused by non-traditional asymptotic behaviors of MLEs in moderate-dimensions and debiased Lasso estimates in high-dimensions. For index models, we had developed an earlier LassoSIR algorithm (Lin, Zhao and Liu 2019), which fits the DS framework quite well. I will also discuss some applications and open questions. The presentation is based on joint work with Chenguang Dai, Buyu Lin, Xin Xing, Tracy Ke, Yucong Ma, and Zhigen Zhao.

★★★★★ All are welcome ★★★★★

For enquiries please contact Miss Esther TAM (Tel: 3943 7931)  
<http://www.sta.cuhk.edu.hk/Events/UpcomingEvents.aspx>

