



THE UNIVERSITY OF  
CHICAGO

Department of Statistics

STATISTICS COLLOQUIUM

---

ARUN KUMAR KUCHIBHOTLA

Statistic Department  
The Wharton School, The University of Pennsylvania

Valid Post-selection Inference: Why and How

THURSDAY, February 6, 2020 at 4:00 PM  
Jones 303, 5747 S. Ellis Avenue

### ABSTRACT

An important component of most data analyses is data exploration for the purpose of finding "good models" on which to base statistical inference. The problem is that data exploration tends to invalidate statistical inference. We show by way of example how a simple variable screening exercise can reduce confidence interval coverage from nominally 95% to nearly 0%. We further illustrate with published examples how data exploration in practice tends to be informal and ad hoc, both in research and in teaching. In this talk we examine two forms of data exploration that are common in regression contexts: selection of covariates, and selection of transformations. We provide a solution to the problem of valid post-selection inference through a reduction to large-scale family-wise error control. The main novelty of the solution is that it applies (1) without any model assumptions (2) to random covariates and (3) to a diverging number of covariates. Furthermore, the solution can be generalized to many other post-selection inference problems. Justifications are asymptotic with rates similar to those of the literature on high-dimensional regression but without common assumptions of linearity, homoskedasticity, Gaussianity and sparsity.

This talk is based on joint work with "Larry's group," including the late Larry Brown, Andreas Buja, Junhui Cai, Edward George, and Linda Zhao.

---

For further information and inquiries about building access for persons with disabilities, please contact Jonathan Rodriguez at 773.702.8333 or send him an email at [jgrodriquez@galton.uchicago.edu](mailto:jgrodriquez@galton.uchicago.edu). If you wish to subscribe to our email list, please visit the following website:  
<https://lists.uchicago.edu/web/subscribe/statseminars>.