THE UNIVERSITY OF CHICAGO

Department of Statistics

STATISTICS COLLOQUIUM

Joint Seminar with The University of Chicago Booth School of Business

---

AADITYA RAMDAS

Department of Statistics
University of California, Berkeley

Multiple Hypothesis Testing with a Human In-The-Loop

MONDAY, February 19, 2018, at 4:30 PM

133 Eckhart Hall, 5734 S. University Avenue

Refreshments before the seminar at 4:00 PM in Jones 111

ABSTRACT

We propose a general framework based on *selectively traversed accumulation rules* (STAR) for interactive "human-in-the-loop" multiple testing with generic structural constraints on the rejection set. STAR combines accumulation tests from ordered multiple testing with data-carving ideas from post-selection inference, allowing for highly flexible adaptation to generic structural information. Given independent p-values for each of n null hypotheses, STAR defines an iterative protocol for gradually pruning a candidate rejection set, beginning with $\mathcal{R}_0 = [n]$ and shrinking with each step. At step t, the analyst estimates the false discovery proportion (FDP) of the current rejection set $\mathcal{R}_t$, and halts and rejects every H_i with $i \in \mathcal{R}_t$ if $\widehat{FDP}_t \leq \alpha$. Otherwise, the analyst may shrink the rejection set to $\mathcal{R}_{t+1} \subseteq \mathcal{R}_t$ however she wants, provided the choice depends only on partially masked p-values $g(p_i)$ for $i \in \mathcal{R}_t$, as well as unmasked p-values p_i for $i \notin \mathcal{R}_t$. Typically, the choice will be based on eliminating the "least promising" hypothesis from $\mathcal{R}_t$, after estimating a model from the observable data. By restricting the information available to the analyst, our iterative protocol guarantees exact false discovery rate (FDR) control at level α in finite samples, for any data-adaptive update rule the analyst may choose. We suggest heuristic update rules for a variety of applications with complex structural constraints, show that STAR performs well for problems ranging from convex region detection and bump-hunting to FDR control on trees and DAGs, and show how to extend STAR to regression problems where knockoff statistics are available in lieu of p-values.

Keywords: interactive multiple testing, data carving, knockoffs, FDR control

Joint work with Lihua Lei, Will Fithian, and Eugene Katsevich

---

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