AADITYA RAMDAS
Department of Statistics
University of California, Berkeley

STAR: A General Interactive Framework for FDR Control
Under Structural Constraints

MONDAY, February 19, 2018 at 4:30 PM
Eckhart 133, 5734 S. University Avenue
Refreshments before the seminar at 4:00PM 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 $R_0 = [n]$ and shrinking with each step. At step $t$, the analyst estimates the false discovery proportion (FDP) of the current rejection set $R_t$, and halts and rejects every $H_i$ with $i \in R_t$ if $\text{FDP}_t \leq \alpha$. Otherwise, the analyst may shrink the rejection set to $R_{t+1} \subseteq R_t$ however she wants, provided the choice depends only on partially masked p-values $g(p_i)$ for $i \in R_t$, as well as unmasked p-values $p_i$ for $i \not\in R_t$. Typically, the choice will be based on eliminating the “least promising” hypothesis from $R_t$, 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 $\alpha$ 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