

HARVARD UNIVERSITY
DEPARTMENT OF STATISTICS



Tel. 617-495-5496
Fax. 617-496-8057
statdept@stat.harvard.edu
<http://www.stat.harvard.edu/>

Science Center 7th Fl.
One Oxford Street
Cambridge, MA 02138

STATISTICS COLLOQUIUM

MONDAY, NOVEMBER 17, 2014

TALK: 4:15 PM — SCIENCE CENTER RM. 705

RECEPTION: 3:50 PM — SCIENCE CENTER, 7TH FLOOR

“Controlling False Discovery Rate via Knockoffs”

Rina Foygel Barber
Department of Statistics
University of Chicago

ABSTRACT

In many fields of science, we observe a response variable together with a large number of potential explanatory variables, and would like to be able to discover which variables are associated with the response, while controlling the false discovery rate (FDR) to ensure that our results are reliable and replicable. The knockoff filter is a variable selection procedure for linear regression, proven to control FDR exactly under any type of correlation structure in the regime where $n > p$ (sample size > number of variables). As the name suggests, the method operates by manufacturing knockoff variables that are cheap — their construction does not require any new data — and are designed to mimic the correlation structure found within the existing variables, in a way that allows for accurate FDR control, beyond what is possible with permutation-based methods. Empirical results show that the knockoff method has far more power than existing selection rules when the proportion of null variables is high. We also apply the knockoff filter to HIV data with the goal of identifying those mutations associated with a form of resistance to treatment plans.

This is joint work with Emmanuel Candès..