THE UNIVERSITY OF CHICAGO

Department of Statistics

STATISTICS COLLOQUIUM

RUTH HELLER

Department of Statistics and Operations Research
Tel-Aviv University, Israel

Designing Optimal Multiple Testing Procedures

MONDAY, May 3, 2021 at **12:00 PM**
Via Zoom (session information will be e-mailed to subscribers)

ABSTRACT

For a single hypothesis testing problem, the optimal policy maximizes the probability to reject the null (i.e., the power), subject to controlling the type I error probability at a predefined . Nowadays, conducting a study with a single hypothesis is rare. Even when designing a clinical trial, it is often the case that at least two hypothesis testing problems are simultaneously considered. We formulate the problem of finding the optimal policy for K>1 hypothesis testing problems for various notions of power, subject to controlling an overall measure of false discovery, like family-wise error rate (FWER) or false discovery rate (FDR). We start by describing a complete solution for deriving optimal policies for K=2 hypotheses, which have some desired monotonicity properties, and are computationally simple. For some of the optimization formulations this yields optimal policies that are identical to existing policies (such as Hochberg's procedure), while for others it yields completely novel policies that are more powerful than existing alternatives. We demonstrate the utility of our novel policies in reanalyzing a clinical trial which aims to infer on the treatment effect of a new drug in two distinct subgroups. Next, we consider the high dimensional setting, where the two-group model is often assumed to hold. For marginal false discovery rate (mFDR) control, the optimal policy is known to be to threshold the local false discovery rate at a fixed level. We address the challenge of controlling optimally the popular false discovery rate (FDR) or positive FDR (pFDR) in the general two-group model, which also allows for dependence between the test statistics. These criteria are less conservative than the mFDR criterion, so

they make more rejections in expectation. We derive their optimal policies, which turn out to be thresholding the local false discovery rate at a level that is a function of the entire set of statistics. We develop an efficient algorithm for finding these policies, and use it on gene expression studies.

Joint work with Abba Krieger and Saharon Rosset