

Controlling FDR while highlighting selected discoveries

Marina Bogomolov

Technion - Israel Institute of Technology

E-mail: marinabo@technion.ac.il

Abstract: "Modern scientific investigations often start by testing a large number of hypotheses by a False Discovery Rate controlling procedure, in order to identify the hypotheses that are promising for follow-up. In many cases, the set of discoveries is somewhat redundant, and it is subject to a second round of selection, where researchers identify the discoveries that better represent distinct findings for reporting and follow-up. For example, in genetic studies, if several genetic variants in a certain locus are identified as associated with the phenotype of interest, typically only the "lead" variant is reported, representing the entire locus. The guarantees of the FDR control for the initial set do not translate to this subset of reported discoveries. We show that if the rule defining how the discoveries will be filtered can be specified in advance, the Benjamini-Hochberg procedure can be modified to result in a focused set of discoveries with FDR guarantees. The proposed method allows researchers to curate rejections not only by subsetting, but also by prioritizing. We illustrate our methodology on a phenome-wide association study, where the hypotheses are structured as a tree.

Joint work with Eugene Katsevich and Chiara Sabatti"