A Bayesian Discovery Procedure

Abstract:

We discuss a Bayesian discovery procedure for multiple-comparison problems. We show that, under a coherent decision theoretic framework, a loss function combining true positive and false positive counts leads to a decision rule that is based on a threshold of the posterior probability of the alternative. Under a semiparametric model for the data, we show that the Bayes rule can be approximated by the optimal discovery procedure, which was recently introduced by Storey (2007). Improving the approximation leads us to a Bayesian discovery procedure, which exploits the multiple shrinkage in clusters that are implied by the assumed non-parametric model. We compare the Bayesian discovery procedure and the optimal discovery procedure estimates in a simple simulation study and in an assessment of differential gene expression based on microarray data from tumor samples. We extend the setting of the optimal discovery procedure by discussing modifications of the loss function that lead to different single-thresholding statistics. Finally, we provide an application of the previous arguments to spatial data. Most of this presentation stems from a joint work with Peter Muller and Song Zhang.