

Using observed p-values to estimate the number of true null hypotheses when conducting many tests

**Daniel S Nettleton PhD**

Department of Statistics

Iowa State University

Abstract:

Mosig et al. (2001, *Genetics* 157, 1683-1698) proposed an intuitively appealing method for estimating the number of true null hypotheses in a multiple test situation. They presented an iterative algorithm that relies on the distribution of observed p-values to obtain their estimator. In this talk, I will characterize the limit of their iterative algorithm and describe how their estimator can be computed directly without iteration. I will compare the performance of the resulting simple estimator with other procedures for estimating the number of true null hypotheses from a collection of observed p-values.

Estimation of the number of true null hypotheses plays a direct role in the estimation of false discovery rate (FDR). The FDR concept has been used in many modern applications involving hundreds or thousands of tests. I will show how the proposed estimator can be used to approximate FDR for some microarray experiments conducted at Iowa State University.