

Title:

Simultaneous Inference: When Should Hypothesis Testing Problems Be Combined?

Author(s):

Bradley Efron

Technical Report number (Dept. of Statistics, Stanford Univ.):

2007-07B/243

Date:

December 2007

Abstract:

Modern statisticians are often presented with hundreds or thousands of hypothesis testing problems to evaluate at the same time, generated from new scientific technologies such as microarrays, medical and satellite imaging devices, or flow cytometry counters. The relevant statistical literature tends to begin with the tacit assumption that a single combined analysis, for instance a False Discovery Rate assessment, should be applied to the entire set of problems at hand. This can be a dangerous assumption, as the examples in the paper show, leading to overly conservative or overly liberal conclusions within any particular subclass of the cases. A simple Bayesian theory yields a succinct description of the effects of separation or combination on false discovery rate analyses. The theory allows efficient testing within small subclasses, and has applications to “enrichment”, the detection of multi-case effects.