A method for robust and rigorous control of the false discovery rate. Preliminary report.

The analysis of mega-dimensional data often involves performing a very large number of hypothesis tests. Multiple-testing analyses can produce many false discoveries (Type I errors) if no adjustments are performed. Benjamini and Hochberg (1995; BH95) and Storey (2002; St02) introduced methods to adjust p-values for multiple-testing. They also developed formal mathematical proofs that rigorously establish the false discovery rate (FDR) control properties of those methods under certain conditions. However, BH95 and St02 have exhibited empirical instability in some simulation studies and applications. Thus, other methods have been proposed that fit curves to the observed distribution of p-values in order to stabilize results. However, the FDR control properties have not been rigorously established by formal mathematical proof for any of these curve-fitting methods. Here, we propose cmFDR as a method that uses curve-fitting to empirically stabilize results and formally prove that it has similar FDR control properties similar to BH95 and St02. In simulation studies, cmFDR exhibits similar FDR control and less variability than BH95 and St02. These results suggest that cmFDR may be the preferred FDR method for some applications. (Received August 24, 2015)