



Is Benjamini-Hochberg correction more conservative as the number of comparisons increases?

Asked 9 years, 8 months ago Active 9 years, 8 months ago Viewed 25k times



20

How conservative is Benjamini-Hochberg multiple testing correction relative to the total number of comparisons? For example, if I have a list of 18,000 features for two groups and I perform a Wilcoxon test to get a p-value. I adjust that p-value using Benjamini-Hochberg and next to nothing comes out as significant.



7

I know that Bonferroni correction can be quite conservative as the number of comparisons increases, does Benjamini-Hochberg have the same property?



multiple-comparisons

p-value

Share Cite Edit Follow Flag

edited Jan 16 '12 at 23:30



cardinal

24.5k

8

92

127

asked Jan 16 '12 at 23:01



user4673

1,441

4

17

26

1 Do you truly mean conservative in the sense that the actual type I rate is much less than the nominal? Or do you mean 'low power'? – [shabbychef](#) Jan 17 '12 at 4:22

He said "conservative", no "too conservative". If a method is conservative it has low power... – [Elvis](#) Jan 17 '12 at 11:16

1 Answer

Active

Oldest

Votes



32

First, you need to understand that these two multiple testing procedures do not control the same thing. Using your example, we have two groups with 18,000 observed variables, and you make 18,000 tests in order to identify some variables which are different from one group to the other.



- Bonferroni correction controls the [Familywise error rate](#), that is the probability, assuming all the 18,000 variables have identical distribution in the two groups, that you are falsely claiming "here I have some significant differences". Usually, you decide that if this probability is $< 5\%$, your claim is credible.
- Benjamini-Hochberg correction controls the [False discovery rate](#), that is, the expected proportion of false positives among the variables for which you claim the existence of a difference. For example, if with FDR controlled to 5% 20 tests are positive, "in average" only 1 of these tests will be a false positive.

Now, when the number of comparison increases... well, it depends on the number of marginal null hypotheses that are true. But basically, with both procedures, if you have a few, let's say 5 or 10, truly associated variables, you have more chances to detect them among 100 variables than among 1,000,000 variables. That should be intuitive enough. There's no way to avoid this.

Share Cite Edit Follow Flag

edited Jan 17 '12 at 11:52

answered Jan 17 '12 at 11:18






Elvis

11.8k

35

55

2  That was a nice explanation. Thanks for the thoughtful response. – user4673 Jan 17 '12 at 23:25 

2  Nice answer (+1), maybe this might interest you stats.stackexchange.com/questions/166323/... – user83346 Sep 9 '15 at 9:02 