**Statistical power of multiplicity adjustment strategies for correlated binary endpoints**

（Andrew C. Leon， 2007）

The objective of this manuscript is to examine the statistical power of three procedures (Bonferroni, Hochberg, and James) when used for two-tailed 2 tests that compare two groups on correlated binary endpoints.

The Hochberg approach was chosen over the Holm approach because by definition, it has statistical power that is greater than or equal to that of Holm.

The James p-value adjustment is based on the standard multivariate normal.

**Identifying differentially expressed genes using false discovery rate controlling procedures**

(Anat Reiner, 2003)

In this paper we address this very large multiplicity problem by adopting the false discovery rate (FDR) controlling approach. In order to address the dependency problem, we present three resampling-based FDR controlling procedures, that account for the test statistics distribution, and compare their performance to that of the na¨ıve application of the linear step-up procedure in Benjamini and Hochberg (1995).

In terms of power, using resampling of the marginal distribution of each test statistics substantially improves the performance over the na¨ıve one.

The Westfall and Young step-down algorithm, herein WY (Westfall and Young, 1989), a permutationbased procedure, is used to adjust for multiplicity by controlling the FWE, without assuming t distribution of the test statistics of each gene’s differential expression.

Tusher et al. (2001): A two-stage p-value adjustment is applied. The estimated FDR is computed using permutations of the data, allowing the possibility of dependent tests. Therefore, as pointed out by the authors, it seems plausible that this estimated FDR approximates the strongly controlled FDR when any subset of null hypotheses is true. However, the authors noted that due to the limited number of possible distinct permutations, the number of distinct values that the p-value can take is limited. Consequently, the FDR estimate turns out to be too ‘granular’, so that either zero or 300 significant genes are identified, depending on how the p-value was defined. A similar result was obtained using the adaptation to dependent tests suggested by Benjamini and Yekutieli (2001b).

Since the BH procedure controls the FDR at a level too low by a factor of m0/m, it is natural to try to estimate m0 and use q∗ = q \*m /m0 instead of q to gain more power.

Adaptive methods offer better performance only by utilizing the difference between m0/m and 1. If the difference is small, i.e. when the potential proportion of differentially expressed genes is small, they offer little advantage in power while their properties are not well established.

For data containing high inter-correlations, generally designed multiple comparisons may be over-conservative in specific dependency structures.

**The effect of correlation in false discovery rate estimation**

（ARMIN SCHWARTZMAN 2011）

Correlation may greatly inflate the variance of both the number of false discoveries (Owen, 2005) and common false discovery rate estimators (Qiu & Yakovlev, 2006).

While there exist procedures that control the false discovery rate under arbitrary dependence (Benjamini & Yekutieli, 2001), they have substantially less power than procedures that assume independence (Farcomeni, 2008) and the latter are often preferred.

There are several ways to estimate p0 (Genovese & Wasserman, 2004; Storey et al., 2004; Efron, 2007b; Jin & Cai, 2007). In applications p0 is often close to 1 and setting ˆp0 =1 biases the estimate only slightly and in a conservative fashion (Efron, 2004).

**The Control of the False Discovery Rate in Multiple Testing under Dependency**

（Yoav Benjamini and Daniel Yekutieli 2001）

We prove that this same procedure also controls the false discovery rate when the test statistics have positive regression dependency on each of the test statistics corresponding to the true null hypotheses.

THEOREM 1.2. If the joint distribution of the test statistics is PRDS on the subset of test statistics corresponding to true null hypotheses, the Benjamini Hochberg procedure controls the FDR at level less than or equal to m0/m\*q.

A mul- tivariate distribution is said to have positive regression dependency if for any increasing set D, P(X E D I X1 = xl, ..., Xi = xi) is nondecreasing in (Xi, . \* \*, Xi).

CASE 1 (Multivariate normal test statistics). Consider X - N(,t, 1) a vec- tor of test statistics each testing the hypothesis Aui = 0 against the alternative ,ui > 0, for i = 1, ..., m. For i E Io, the set of true null hypotheses, t = 0. Otherwise 1xi > 0. Assume that for each i E Io, and for each j 0 i, tU > 0, then the distribu- tion of X is PRDS over Io.

THEOREM 1.3. When the Benjamini Hochberg procedure is conducted with q/(∑mi=1 1/i) taking the place of q in (1), it always controls the FDR at level less than or equal to m0/m\*q.

**A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion**

(Alessio Farcomeni 2008)

The most pressing advance is in our opinion the discovery of a multiple testing procedure that controls the FDR under arbitrary dependence but it is competitive with BH method in terms of power.

There are many other open questions for research in multiple testing: until now for

instance the literature on FDP and multiple testing in general does not seem to be

interested in extensions to composite null hypotheses. It is well known that when the null hypothesis is composite the interpretation of p-values is more complex (see for instance138), and furthermore the distribution under the null hypothesis need not be uniform. The only practical solution at this point seems to be estimation of p-values through resampling.

Among other open problems, there is the derivation of a framework for power analysis; and a closely related problem, that is, a method to choose the sample size for each test.

**Some Results on the Control of the False Discovery Rate under Dependence**

（ALESSIO FARCOMENI 2006）