FDR adjustments of Microarray Experiments (FDR-AME)

Yoav Benjamini, Effi Kenigsberg, Anat Reiner, Daniel Yekutieli

October 28, 2009

Department of Statistics and O.R., Tel Aviv University

Purpose This R package adjusts p-values generated in multiple hypotheses testing of gene expression data obtained by a microarray experiment. The software applies multiple testing procedures that control the False Discovery Rate (FDR) criterion introduced by Benjamini and Hochberg (1995). It applies both theoretical-distribution-based and resampling-based multiple testing procedures, and presents as output adjusted p-values and p-value plots, as described in Reiner et al (2003). It goes beyond Reiner et al in offering adjustments according to the adaptive two stage FDR controlling procedures in Benjamini et al (2001, submitted), and in addressing differences in expression between many classes using one-way ANOVA.

The False Discovery Rate (FDR) Criterion The FDR is the expected proportion of erroneously rejected null hypotheses among the rejected ones. Consider a family of m simultaneously tested null hypotheses of which m_0 are true. For each hypothesis H_i a test statistic is calculated along with the corresponding p-value P_i. Let R denote the number of hypotheses rejected by a procedure, V the number of true null hypotheses erroneously rejected, and S the number of false hypotheses rejected. Now let Q denote V/R when $R{>}0$ and 0 otherwise. Then the FDR is defined as

$$\text{FDR=E(Q)}.$$

The Linear Step-Up Procedure (BH)

This procedure makes use of the ordered p-values $P_{(1)} \leq \ldots \leq P_{(m)}$. Denote the corresponding null hypotheses $H_{(1)}, \ldots, H_{(m)}$. For a desired FDR level q, the ordered p-value $P_{(i)}$ is compared to the critical value $q{\cdot}i/m$. Let $k = max\ \{\ i : P_{(i)} \leq q{\cdot}i/m\ \}$. Then reject $H_{(1)}, \ldots, H_{(k)}$, if such a k exists.

Benjamini and Hochberg (1995) show that when the test statistics are independent, this procedure controls the FDR at the level q. Actually, the FDR is controlled at level $FDR \leq q{\cdot}m_0/m \leq q$.

Benjamini and Yekutieli (2001) further show that $FDR \leq q \cdot m_0/m$ for positively dependent test statistics as well. The technical condition under which the control holds is that of positive regression dependency on each test statistic corresponding the true null hypotheses. Reiner et al (2003) and Reiner (unpublished thesis) shows $FDR \leq q$ for two sided tests under positive and negative correlations.

The Adaptive Procedures

Since the BH procedure controls the FDR at a level too low by a factor of m_0/m, it is natural to try to estimate m_0 and use $q^* = q \cdot \frac{m}{m_0}$ instead of q to gain more power. Benjamini et al (2001) suggest a simple two-stage procedure: use BH once to reject r1 hypotheses; then use the BH at the second stage at level $q^* = q \cdot \frac{m}{(m-r_1) \cdot (1+q)}$ This two stage procedure has proven FDR controlling properties under independence and simulation support for its controlling properties under positive dependence.

Resampling FDR Adjustments

For data containing high inter-correlations, generally designed multiple comparisons may be over-conservative in specific dependency structures. Resampling-based multiple testing procedures utilize the empirical dependency structure of the data to construct more powerful FDR controlling procedures.

In p-value resampling, the data is repeatedly resampled under the complete null hypotheses, and a vector of resample-based p-values is computed. The underlying assumption is that the joint distribution of p-values corresponding to the true null hypotheses, which is generated through the p-value resampling scheme, represents the real joint distribution under the null hypothesis. Thus, for each value of p, the number of resampling-based p-values less than p, denoted by $V^*(p)$, is an estimated upper bound to the expected number of p-values corresponding to true null hypotheses less than p.

Yekutieli and Benjamini (1999) introduce resampling-based FDR control, while taking into account that the FDR is also a function of the number of false null hypotheses rejected. Therefore, for each value of p, they first conservatively estimate the number of false null hypotheses less than p, denoted by $\hat{s}(p)$, and then estimate the FDR adjustment by

$$FDR^{est}(p) = E_{V^*(p)} \frac{V^*(p)}{V^*(p) + \hat{s}(p)}$$

Two estimation methods are suggested differing by their strictness level. The FDR local estimator is conservative on the mean, and the FDR upper limit bounds the FDR with probability 95%.

A third alternative uses the BH procedure to control the FDR, but rather than using the raw p-values, it estimates the p-values by resampling from the marginal distribution and collapsing over all hypotheses, assuming exchangeability of the marginal distributions: For the k-th gene, with an observed test statistics t_k, the estimated p-value is

$$P_k^{est} = \frac{1}{I} \sum_{i=1}^{I} \left[\frac{1}{N} \# \left(|t_i^{*j}| \geq |t_k| \right) \right]$$

2

We next use the estimated p-values in the BH procedure to easily obtain the BH point estimate for the k-th gene:

$$P_{(k)}^{BH} = \min_{k \le i} \frac{P_{(i)}^{est} \cdot m}{i}$$

Plots of p-values   In addition to output of significant genes in a file, the program offers plots of p-value. The plot of p-values versus rank for all genes is a diagnostic plot that allows researchers to examine the adequacy of the preprocessing stage as well as of the assumptions on which the distribution of the test statistics are based. The plot of the adjusted p-values versus rank (or versus estimated difference) allows researchers to pick their desired FDR level by comparing simply comparing the adjusted p-value to the desired level, and then view the consequence in terms of the pool of genes thereby identified as significant. Each FDR controlling method results in its corresponding set of adjusted p-values.

References

1. Benjamini,Y. and Hochberg,Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. Roy. Stat. Soc. B., **57**, 289-300.

2. Benjamini,Y., Krieger,A. and Yekutieli,D. (2001) Two-Staged Linear Step-Up FDR Controlling Procedure, Department of Statistics and Operation Research, Tel-Aviv University, and Department of Statistics, Wharton School, University of Pennsylvania, Technical Report. (Submitted)

3. Benjamini,Y. and Yekutieli,D. (2001) The Control of the False Discovery Rate Under Dependency. Ann Stat, **29**, 1165-1188.

4. Reiner,A., Yekutieli,D. and Benjamini,Y. (2003) Identifying Differentially Expressed Genes Using False Discovery Rate Controlling Procedures. Bioinformatics, 19(3), 368-375.

5. Yekutieli,D. and Benjamini,Y. (1999) Resampling-Based False Discovery Rate Controlling Multiple Test Procedures for Correlated Test Statistics. J Stat Plan Infer, **82**, 171-196.