Modeling Dose-response Microarray data: The IsoGene libary

Dan Lin[1], Setia Pramana[1], Ziv Shkedy[1]and Tobias Verbeke[2]

1. I-Biostat, Hasselt University
Agoralaan 1, B-3590 Diepenbeek, Belgium
2. OpenAnalytics BVBA
Kortestraat 30A, 2220 Heist-op-den-Berg, Belgium

dan.lin@uhasselt.be

January, 2009

Contents

1 Introduction

Investigation of a dose-response relationship is of primary interest in many drug-development studies. Typically, in dose-response experiments the outcome of interest is measured at several (increasing) dose levels, and the aim of the analysis is to establish the form of the dependence of the response on dose (Agresti 1997). The response can be either the efficacy of a treatment or the risk associated with the exposure to the treatment (in toxicology studies). In a typical dose-response study subjects are randomized to several dose groups, among which there is usually a control group. Ruberg (1995a, 1995b) and Chuang-Stein and Agresti (1997) formulated four main questions usually asked in dose-response studies: (1) Is there any evidence of the drug effect? (2) For which doses is the response different from the response in the control group? (3) What is the nature of the dose-response relationship? and (4) What is the optimal dose?

Within the microarray setting, a dose-response experiment has the same structure as described above. The response is the gene-expression at a certain dose level. The dose-response curve, similarly to the dose-response studies, is assumed to be monotone, i.e., the gene activity increases or decreases as the dose level increases. The direction of the relationship is usually unknown in advance.

In this chapter we focus on the first question: is there any evidence of the drug effect? To answer this question, we test for the null hypothesis of homogeneity of means (no dose effect) against an ordered alternative. We compare several testing procedures, that take into account the order restriction of the means with respect to the increasing doses and that adjust for multiple testing. In particular, we discuss the testing procedures of Williams (Williams 1971 and 1972), Marcus (Marcus 1976), the global likelihood ratio test (LRT, Barlow *et al.* 1972, and Robertson *et al.* 1988), and the M (Hu *et al.* 2005) statistic. Moreover, we propose a novel procedure based on a modification of the estimator of standard error of the M statistic.

Williams (1971, 1972) proposed a step-down procedure to test for the dose effect. The tests are performed sequentially from the comparison between the isotonic mean of the highest dose and the sample mean of the control, to the comparison between the isotonic mean of the lowest dose and the sample mean of the control. The procedure stops at the dose level where the null hypothesis (of no dose effect) is not rejected. Marcus (1976) proposed a modification of the Williams procedure, in which the sample mean of the control was replaced by the isotonic mean of the control. A global

likelihood ratio test, discussed by Bartholomew *et al.* (1961), Barlow *et al.* (1972), and Robertson *et al.* (1988), uses the ratio between the variance calculated under the null hypothesis and the variance calculated under an ordered alternative. Recently, Hu *et al.* (2005) proposed a test statistic that was similar to Marcus' statistic, but with the variance estimator calculated under the ordered alternative. The degrees of freedom of the M statistic (the difference between the number of observations and the number of dose levels) were fixed for all the genes and all the arrays. We propose a modification for the variance estimator of the M statistic. Namely, the difference between the number of observations and the unique number of isotonic means is used as the degrees of freedom for the variance estimator.

`IsoGene` is an R package for the analysis of dose-response microarray experiments. The package can be used in order to identify differentially expressed genes, which are, within the framework of dose-response microarray experiments, a subset of genes, for which a monotone relationship between the gene expression and doses can be detected. Inference is based on resampling methods (both permutations and the significance analysis of Microarray (SAM)), in which the multiplicity issue is addressed by adjustment techques controlling for the false discovery rate (FDR). This guide provides a tutorial to the features of the package. It illustrates the capability of the `IsoGene` package and provides some background information about the methodology used for the analysis. In Chapter 2, we review different testing procedures; while in Chapter 3, we illustrate how the methodology discussed in Chapter 2 can be implemented using the `IsoGene` package.

2 Testing for Trend in Dose-response Microarray Experiments

3 Testing for Homogeneity of the Means Under Restricted Alternatives

In this section, we review several procedures for testing the homogeneity of the means against order restricted alternatives. In particular we focus on four existing procedures: Williams' (Williams 1971 and 1972), Marcus' (Marcus 1976), the global likelihood ratio test (Bartholomew 1961, Barlow *et al.* 1972, and Robertson *et al.* 1988),

and the M (Hu *et al.* 2005) statistic. Additionally, we introduce a modification to the degrees of freedom of the M statistic.

In the microarray experiment, for each gene, the following ANOVA model is considered:

$$Y_{ij} = \mu(d_i) + \varepsilon_{ij},\ i = 0, 1, \ldots, K,\ j = 1, 2, \ldots, n_i, \tag{1}$$

where Y_{ij} is the jth gene-expression at the ith dose level, d_i $(i = 0, 1, \ldots, K)$ are the $K+\mathbf{1}$ dose levels, $\mu(d_i)$ is the mean gene-expression at each dose level, and $\varepsilon_{ij} \sim N(0, \sigma^2)$.

The null hypothesis of no dose effect is given by

$$H_0 : \mu(d_0) = \mu(d_1) = \cdots = \mu(d_K). \tag{2}$$

A one-sided alternative hypothesis of a positive dose effect for at least one dose level (i.e., an increasing trend) is specified by

$$H_1^{Up} : \mu(d_0) \leq \mu(d_1) \leq \cdots \leq \mu(d_K), \tag{3}$$

with at least one strict inequality. When testing the effect of a drug for a positive outcome the researcher can specify a positive effect as the desirable alternative. However, in the current microarray setting, it seems reasonable to assume that the gene-expression levels may increase or decrease in response to increasing doses, but with the direction of the trend not known in advance. Thus, we must also consider an additional alternative:

$$H_1^{Down} : \mu(d_0) \geq \mu(d_1) \geq \cdots \geq \mu(d_K), \tag{4}$$

with at least one strict inequality. Testing H_0 against H_1^{Down} or H_1^{Up} requires estimation of the means under both the null and the alternative hypotheses. Under the null hypothesis, the estimator for the mean response $\hat{\mu}$ is the sample mean. Let $\hat{\mu}_0^\star, \hat{\mu}_1^\star, \ldots, \hat{\mu}_K^\star$ be the maximum likelihood estimates for the means (at each dose level) under the ordered alternative. Barlow *et al.* (1972) and Robertson *et al.* (1998) showed that $\hat{\mu}_0^\star, \hat{\mu}_1^\star, \ldots, \hat{\mu}_K^\star$ are given by the isotonic regression of the observed means.

3.1 Williams' (1971, 1972) and Marcus' (1976) Test Statistics

Williams' procedure defines H_0 as the null hypothesis, and H_1^{Up} or H_1^{Down} as the one-sided alternative. Williams' (1971, 1972) test statistic was suggested for a setting, in

which n_i observations are available at each dose level. As all dose levels are compared with the control level, the test statistic is given by

$$t_i = \frac{\hat{\mu}_i^\star - \bar{y}_0}{\sqrt{2s^2/r}}. \tag{5}$$

Here, $\bar{y}_0$ is the sample mean at the first dose level (control), $\hat{\mu}_i^\star$ is the estimate for the mean at the ith dose level under the ordered alternative, r is the number of replications at each dose level, and s^2 is an estimate of the variance. For $\hat{\mu}_i^\star$, Williams (1971, 1972) used the isotonic regression of the observed response with respect to dose (Barlow *et al.* 1972). Williams' test procedure is a sequential procedure. In the first step, $\hat{\mu}_K^\star$ is compared to $\bar{y}_0$. If the null hypothesis is rejected, $\hat{\mu}_{K-1}^\star$ is compared to $\bar{y}_0$, etc.

Marcus (1976) proposed a modification to Williams' test statistic that replaced $\bar{y}_0$ with $\hat{\mu}_0^\star$, the estimate of the first dose (control) mean under ordered restriction. Marcus' test statistic performs closely to Williams' in terms of power (Marcus 1976). Note that, for $K = 1$, Williams' and Marcus' test statistics reduce to the two-sample t-test.

3.2 Likelihood Ratio Test Statistic for Monotonicity (Barlow *et al.* 1972, and Robertson *et al.* 1988)

Williams' and Marcus' procedures are step-down procedures, i.e., the comparison between a lower dose and control is tested only if the test of a higher dose vs. control is significant. The underlying assumption is that there is a monotone dose-response relationship with a known direction.

Testing the equality of ordered means using likelihood ratio tests (when response is assumed to be normally distributed) was discussed by Barlow *et al.* (1972) and Robertson *et al.* (1988). Both authors considered the likelihood ratio test, in which the variance under the null and the alternative were compared. The likelihood ratio test statistic is given by

$$\Lambda_{01}^{\frac{2}{N}} = \frac{\hat{\sigma}_{H_1}^2}{\hat{\sigma}_{H_0}^2} = \frac{\sum_{ij}(y_{ij} - \hat{\mu}_j^\star)^2}{\sum_{ij}(y_{ij} - \hat{\mu})^2}, \tag{6}$$

where $\hat{\sigma}_{H_0}^2$ and $\hat{\sigma}_{H_1}^2$ are the estimates for the variance under the null and the alternative hypothesis, respectively. And $\hat{\mu} = \sum_{ij} y_{ij} / \sum_i n_i$ is the overall mean. The null

hypothesis is rejected for a "small" value of $\Lambda_{01}^{\frac{2}{N}}$. Equivalently, H_0 is rejected for large value of $\bar{E}_{01}^2$, where

$$\bar{E}_{01}^2 = 1 - \Lambda_{01}^{\frac{2}{N}} = \frac{\sum_{ij}(y_{ij} - \hat{\mu})^2 - \sum_{ij}(y_{ij} - \hat{\mu}_j^\star)^2}{\sum_{ij}(y_{ij} - \hat{\mu})^2}. \tag{7}$$

Estimating the parameters using isotonic regression requires the knowledge of the direction of the trend. In practice, the direction of the trend is often not known in advance. In such a case one can maximize the likelihood twice: for a monotone decreasing trend and for a monotone increasing trend, and choose the trend with a higher likelihood. In practice, we can calculate $\bar{E}_{01}^2$ for each direction and choose the higher value of $\bar{E}_{01}^2$ (Barlow *et al.* 1972). A resampling-based approach, as described in Section 3.2.2, can be used to approximate the null distribution for the test statistic, so that two-sided p-values are obtained for inference.

3.3 The M Test Statistic of Hu *et al.* (2005)

Recently, Hu *et al.* (2005) proposed the following test statistic M to test for a monotonic trend:

$$M = \frac{\hat{\mu}_K^\star - \hat{\mu}_0^\star}{\sqrt{\sum_{i=0}^{K}\sum_{j=1}^{n_i}(y_{ij} - \hat{\mu}_i^\star)^2/(n-K)}}. \tag{8}$$

where n is the total number of arrays.

Hu *et al.* (2005) discussed a setting, in which the comparison of primary interest is the difference between the highest dose level (K) and the control dose. The numerator of the M test statistic is the same as that of Marcus' statistic, while the denominator is an estimate of the standard error under an ordered alternative. This is in contrast to Williams' and Marcus' approaches that use the unrestricted means to derive the estimate for the standard error.

Hu *et al.* (2005) evaluated the performance of the $\bar{E}_{01}^2$ and M test statistics by comparing the ranks of genes obtained by using both statistics, and reported similar findings for simulated and real-life data sets.

3.4 A Modification to the M Test Statistic (Lin *et al.* 2007)

For the variance estimate, Hu *et al.* (2005) used $n - K$ degrees of freedom (see equation (8)). However, the unique number of isotonic means is not fixed, but changes

across the genes. For that reason, we propose a modification to the standard error estimator used in the M statistic by replacing it with $\sqrt{\sum_{i=0}^{K}\sum_{j=1}^{n_i}(y_{ij}-\hat{\mu}_i^{\star})^2/(n-I)}$, where I is the unique number of isotonic means for a given gene. Such a modification is expected to improve the standard error estimates across all the genes.

The five test statistics are implemented in the *R* `IsoGene` package, which is discussed in detail in the next chapter.

4 Directional Inference

4.1 Directional Inference in Isotonic Regression

The five test statistics discussed in Section 3 should be calculated assuming a particular direction of the ordered alternative. However, the direction of the test is unknown in advance. In this section, we address the issue of how to obtain the two-sided p-value from the five testing procedures, and how to determine the direction of the trend from two-sided p-value afterwards.

We focus on the two possible directions of the alternatives: H_1^{Up} defined in equation (3) and H_1^{Down} defined in equation (4). Let p^{Up} and T^{Up} denote the p-value and the corresponding test statistic computed to test H_0 vs. H_1^{Up}, and let p^{Down} and T^{Down} denote the p-value and the corresponding test statistic computed to test H_0 vs. H_1^{Down}. Barlow *et al.* (1972) showed that, for $K > 2$, a $\bar{\chi}^2$ statistic for testing H_0 may actually yield $p^{Up} < \alpha$ and $p^{Down} < \alpha$. However, $p = 2\min(p^{Up}, p^{Down})$ is always a conservative p-value for the two-sided test of H_0 vs. either H_1^{Up} or H_1^{Down}.

Hu *et al.* (2005) adapted the approach by taking the larger of the likelihoods of H_1^{Up} or H_1^{Down}, i.e., the larger of T^{Up} and T^{Down} is used as the test statistic for the two-sided inference. In contrast to Hu *et al.* (2005), we obtain two-sided p-values by taking $p = \min(2\min(p^{Up}, p^{Down}), 1)$, where p^{Up} and p^{Down} are calculated for T^{Up} and T^{Down} using permutations to approximate the null distribution of these test statistics. We use p^{Up} and p^{Down} to determine the direction of the trend, as described below.

After rejecting the null hypothesis against the two-sided test there is still a need to determine the direction of the trend. The direction can be inferred by the following procedure. If $p^{Up} \leq \alpha/2$, then reject H_0 and declare H_1^{Up}; if $p^{Down} \leq \alpha/2$, then reject H_0 and declare H_1^{Down}. The validity of this directional inference is based on the following property: under H_1^{Up}, p^{Down} is stochastically larger than $U[0,1]$; and

under H_1^{Down}, p^{Up} is stochastically larger than $U[0,1]$ (proof not given here). Thus, the probability of falsely rejecting H_0 is $\leq \alpha$, and the probability of declaring a wrong direction for the trend is $\leq \alpha/2$. It is also important to note that the event $p^{Up} < \alpha/2$ and $p^{Down} < \alpha/2$ may be observed. Under H_0, H_1^{Up}, or H_1^{Down}, this event is unlikely. However, it is likely if the treatment has a large and non-monotone effect.

In order to illustrate whether the property needed for directional inference applies to the five test statistics, we conduct a simulation study to investigate the distribution of the p^{Up} and p^{Down} values. For each simulation, data are generated under H_1^{Up}: the means are assumed to be equal to $(1,\ 2,\ 3,\ 4)/\sqrt{5}$ for the four doses, respectively, and the variance is equal to $\sigma^2 = 1$. The test statistics T^{Up} and T^{Down} are calculated for the two possible alternatives H_1^{Up} and H_1^{Down}. Their corresponding p^{Up}- and p^{Down}-values are obtained using 10,000 permutations.

Figure 1 shows the cumulative distribution of p^{Up} and p^{Down}. Clearly, the simulations show that the cumulative distribution of p^{Down} (the p-value of the test statistics calculated assuming the wrong direction, dotted line in Figure 1) is stochastically higher than $U[0,1]$ (solid line in Figure 1), which is the distribution of the p-values under the null hypothesis. Moreover, the distribution of p^{Up} (the p-value for the test statistics calculated assuming the right direction, dashed line in Figure 1) is, as expected, stochastically smaller than $U([0,1]$. Similar results (not shown) are obtained when the data are generated under H_1^{Down}. The results imply that all the five test statistics possess the property required for the directional inference: under H_1^{Up} the distribution of p^{Down} is stochastically greater than $U[0,1]$.

Figure 2 shows the values of test statistics, which were calculated under H_1^{Up} and H_1^{Down}, for data generated under H_1^{Up}. The five test statistics are calculated for testing H_0 vs. H_1^{Down} (the x-axis of each test statistic in Figure 2). The behavior of Marcus', M, and the modified M statistics is similar as they all use the difference between the highest and the lowest isotonic mean. The maximum value of the test statistics (when calculated assuming the wrong direction) is equal to zero. In contrast, Williams' test statistic for testing H_0 vs. H_1^{Down} (shown on the x-axis of the panel b) can be positive or negative, because the sample mean of control group is used instead of the isotonic mean. Note that we reject the null hypothesis in favor of H_1^{Down} for negative values of the test statistic. Further, the value of the test statistics for testing H_0 vs. H_1^{Up} (the y-axis of Figure 2) is higher than the value of the test statistics calculated for testing H_0 vs. H_1^{Down} (the x-axis of Figure 2).

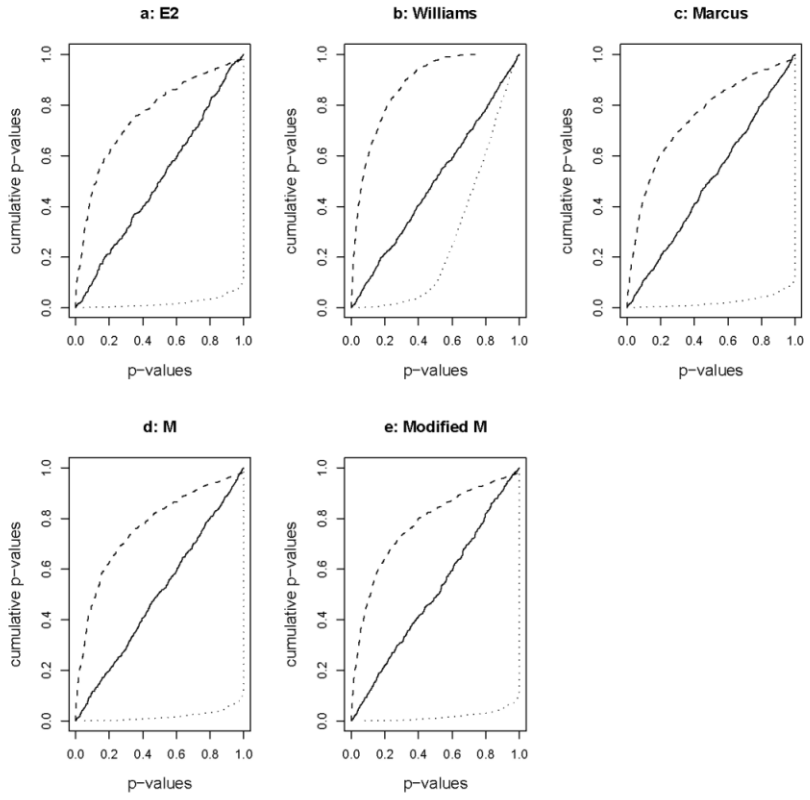

Figure 1: *The cumulative distribution of p^{Up}-values (dashed line) and p^{Down}-values (dotted line) for the five test statistics. Data are generated under H_1^{Up} with isotonic means (1, 2, 3, 4)/$\sqrt{5}$ for the four doses. Solid line: cumulative distribution of $H_0 \sim U[0,1]$.*

4.2 Control of the Directional FDR

When the FDR controlling procedures are used to adjust for multiple testing in the microarray setting, the set of two-sided p-values computed for each gene is adjusted by using the BH-FDR or BY-FDR procedure described in Section 3.2.1. A discovery

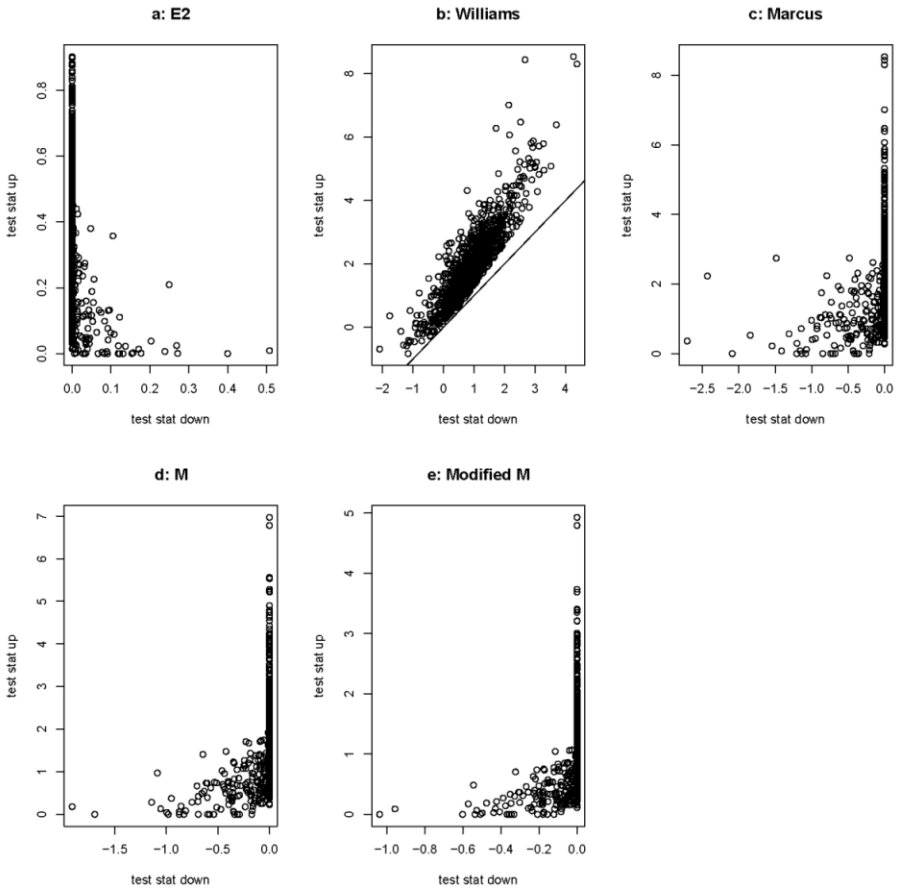

Figure 2: *The five test statistics calculated for* H_0 *vs.* H_1^{Up} *(y-axis) and* H_0 *vs.* H_1^{Down} *(x-axis).*

in this case is a rejection of H_0 for some gene; a false discovery is to reject H_0 when H_0 is true. As mentioned before, in a microarray dose-response experiment we are also interested in the direction of the dose-response trend.

Benjamini and Yekutieli (2005) provide a framework for addressing the multiplicity problem when attempting to determine the direction of multiple parameters: a discovery is to declare the sign of a parameter as either being positive or negative.

Three types of false discoveries are possible: declaring a zero parameter either as negative or as positive, declaring a negative parameter as positive, and declaring a positive parameter as negative. The FDR corresponding to these discoveries is termed the Mixed Directional FDR (MD-FDR). In the current setting, the MD-FDR is the expected value of the number of genes, for which H_0 is true, that are erroneously declared to have either a positive or negative trend plus the genes with a monotone trend but with a wrong direction of the declared trend, divided by the total number of genes declared to have a trend. Benjamini and Yekutieli (2005) prove that if p-values pose the directional property described in Section 4.1, then applying the BH procedure at level q to the the set of two-sided p-values computed for each gene, and declaring the direction of the trend corresponding to the smaller one-sided p-value, controls the MD-FDR at level $q/2 \cdot (1 + m_0/m)$, where m is the total number of genes and m_0 is the number of genes, for which H_0 holds.

In general, directional inference is a more general setting than hypotheses testing (Benjamini and Yekutieli, 2005). Nevertheless, as a false discovery is made based on the p-value that is stochastically larger than $U[0, 1]$, then the resampling-based methods that control the FDR (Yekutieli and Benjamini, 1999) also control the MD-FDR. This is achieved by simply applying the resampling-based procedure to test H_0, and if H_0 is rejected, declaring the direction of the trend according to the minimum one-sided p-value. For each rejected null hypothesis it is also advisable to examine if the larger p-value is $\leq \alpha$. If this is the case, this may serve as an indication of a non-monotone dose-response relationship.

5 Introduction to `IsoGene` Package

The main IsoGene package functions are `IsoRawp()` and `IsoTestBH()`, which calculate the raw p-values using permutations and adjust them using the BH- and BY-FDR procedures. The supporting functions `IsoGene1()` and `IsoGenem()` are used to calculate the five test statistics from isotonic regression for one gene and all the genes, respectively. On the other hand, the SAM procedure is also implemented to reduce some computational time as compared to the permutation method. The main function of the SAM is `IsoTestSAM()`, with supporting functions `Isofudge(), IsoGenemSAM(), Isoqqstat(), Isoallfdr(), Isoqval()`. The remaining functions `IsopvaluePlot(), IsoBHPlot(), IsoSAMPlot() IsoPlot()` is used to display the data and to show the results of testing procedures.

6 Testing for Trends: Testing Procedures, Multiplicity and Resampling-based Inference

6.1 Resampling-based Multiple Testing

For adjusting for multiple testing, only the BH-FDR procedure (Benjamini and Hochberg 1995) and BY-FDR procedure (Benjamini and Yekutieli 2001) are considered in package `IsoGene()`. The matrix of the values of the test statistics for each gene and permutation is referred as the permutation matrix under the null distribution (see Section 3.2.2).

This matrix is used to calculate the one-sided p-values for the inference. In the first step the one-sided raw (unadjusted for multiple testing) p^{Up}-values are calculated using (9) or (10) based on the test statistic T^{Up}.

$$P_i = \frac{\#(b : t_{ib} \geq t_i)}{B-1}, \tag{9}$$

where t_i is the observed test statistic for gene i.

$$P_i = \frac{\sum_{b=1}^{B}\sum_{j=1}^{m}(t_{jb} \geq t_i)}{(B-1)\times m}. \tag{10}$$

For p^{Down}-values, expect of $\bar{E}_{01}^2$, for which the test statistic value t_i is always between 0 and 1 and can be obtained in the same way as p^{Up}-values,

$$p^{Down} = \#(b : t_{ib} \leq t_i)/B \text{ or } p^{Down} = \sum_{b=1}^{B}\sum_{j=1}^{m}(t_{jb} \leq t_i)/(B \times m)$$

should be used with t_{ib} and t_{jb} the test statistic values obtained for gene i and j from permutation b. This is because under the decreasing trend, we reject the four test statistics (namely, Williams', Marcus', the M and modified M) with large negative values.

Based on the p-values, various methods adjusting the type I error can be applied, such as the Bonferroni, Holm, Hochberg, and BH-FDR and BY-FDR (Reiner *et al.* 2003 and Ge *et al.* 2003).

6.2 Significance Analysis of Microarray (SAM)

SAM (Tusher *et al.* 2001, Lin *et al.* 2008) is a procedure widely used in the microarray setting. SAM is a testing procedure, which estimates the FDR by using permutations

under the assumption that all null hypotheses are true. The procedure consists of three components: (1) the adjusted test statistics, (2) an approximation of the distribution of the test statistics based on permutations, and (3) the control of the FDR.

For a two-group setting, the modified test statistic in SAM is given by,

$$t_k^{SAM} = \frac{\bar{x}_k - \bar{x}_l}{s_k + s_0}, \tag{11}$$

where

$$\bar{x}_l = \frac{\Sigma_{j=1}^{n_l} x_{jl}}{n_l}, \quad \bar{x}_k = \frac{\Sigma_{j=1}^{n_k} x_{jk}}{n_k},$$

$$s_k = \sqrt{\left(\frac{1}{n_k} + \frac{1}{n_l}\right) \frac{\Sigma_{j=1}^{n_k}(x_{jk} - \bar{x}_{jk})^2 + \Sigma_{j=1}^{n_l}(x_{jl} - \bar{x}_{jl})^2}{n_k + n_l - 2}},$$

and s_0 is the fudge factor which is estimated from the data and is discussed later, k and l are the index of the two groups of array, and j is the index of the array.

For F-type test statistic, such as $\bar{E}_{01}^2$, the modified test statistic is given by,

$$\bar{E}_{01}^{2SAM} = \frac{\sqrt{\hat{\sigma}_{H_0}^2 - \hat{\sigma}_{H_1}^2}}{\sqrt{\hat{\sigma}_{H_0}^2} + s_0}, \tag{12}$$

SAM requires that the test statistic for each permutation is sorted for all the genes, such that the first row of the sorted matrix is the minimum test statistic across permutations, and the last row is the maximum, i.e.,

$$\boldsymbol{T}^{SAM} = \begin{pmatrix} t_{(1)1} & t_{(1)2} & \dots & t_{(1)B} \\ t_{(2)1} & t_{(2)2} & \dots & t_{(2)B} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ t_{(m)1} & t_{(m)2} & \dots & t_{(m)B} \end{pmatrix}.$$

In $\boldsymbol{T}^{SAM}$, each element $t_{(i)b}$ is the sorted test statistic for gene i in permutation b. The expected values of the observed ordered statistics are approximated by the means of the rows of $\boldsymbol{T}^{SAM}$, given by $\bar{t}_{(1)}^{SAM}, \bar{t}_{(2)}^{SAM}, \dots, \bar{t}_{(m)}^{SAM}$ that are constructed in

the following way:

$$\boldsymbol{T}^{SAM} = \left(\begin{array}{cccc} t_{(1)1} & t_{(1)2} & \dots & t_{(1)B} \\ \hline t_{(2)1} & t_{(2)2} & \dots & t_{(2)B} \\ \hline . & . & . & . \\ \hline . & . & . & . \\ \hline . & . & . & . \\ \hline t_{(m)1} & t_{(m)2} & \dots & t_{(m)B} \end{array}\right) \Rightarrow \left(\begin{array}{c} \frac{1}{B}\sum_{b=1}^{B} t_{(1)b} \\ \frac{1}{B}\sum_{b=1}^{B} t_{(2)b} \\ . \\ . \\ . \\ \frac{1}{B}\sum_{b=1}^{B} t_{(m)b} \end{array}\right) = \left(\begin{array}{c} \bar{t}_{(1)}^{SAM} \\ \bar{t}_{(2)}^{SAM} \\ . \\ . \\ . \\ \bar{t}_{(m)}^{SAM} \end{array}\right).$$

The SAM procedure proposed by Tusher *et al.* (2001) is as follows:

1. Compute order statistics $t_{(1)}^{SAM} \leq t_{(2)}^{SAM} \leq \cdots \leq t_{(m)}^{SAM}$.

2. Compute the permutation matrix $\boldsymbol{T}^{SAM}$.

3. Calculate the expected test statistics $\bar{t}_{(1)}^{SAM}, \bar{t}_{(2)}^{SAM}, \ldots, \bar{t}_{(m)}^{SAM}$.

4. Plot the $t_{(1)}^{SAM}, t_{(2)}^{SAM}, \ldots, t_{(m)}^{SAM}$ values versus the $\bar{t}_{(1)}^{SAM}, \bar{t}_{(2)}^{SAM}, \ldots, \bar{t}_{(m)}^{SAM}$ values (SAM plot).

5. For a fixed threshold Δ, starting at the origin, and moving up to the right, find the first $i = i_1$ such that $t_i^{SAM} - \bar{t}_i^{SAM} > \Delta$. All genes, for which $t_i^{SAM} > t_{i1}^{SAM}$, are called "significant positive". Similarly, start at origin, move down to the left and find the first $i = i_2$ such that $\bar{t}_2^{SAM} - t_i^{SAM} > \Delta$. All genes, for which $t_i^{SAM} < t_{i2}^{SAM}$, are called "significant negative". For each Δ define the upper cut-point $\mathrm{C}_{up}(\Delta)$ as the smallest t_i^{SAM} among the significant positive genes, and similarly define the lower cut-point $\mathrm{C}_{low}(\Delta)$.

6. For a grid of Δ values, compute the total number of significant genes (from step 5), and the median number of falsely called genes, i.e., the median number of values among each of the B sets of t_{ib} , $i = 1, 2, \ldots, m$ that fall above cut-point $\mathrm{C}_{up}(\Delta)$ and below cut-point $\mathrm{C}_{down}(\Delta)$. Similarly, compute the 90th percentile of the number of falsely called genes.

7. Estimate π_0, the proportion of truly non-differentially expressed genes in the data set, as follows:

   (a) Compute the first and third quantiles of the permuted t^{SAM} values, denoted as $q25$ and $q75$ (the t_i^{SAM} are the values for the original data set; there are m such values).

   (b) Compute $\hat{\pi}_0 = \#\{t_i \in (q25, q75)\}/(.5m)$.

   (c) Let $\hat{\pi}_0 = min(\hat{\pi}_0, 1)$.

8. The median and the 90th percentile of the number of falsely called genes from step 6, are multiplied by $\hat{\pi}_0$,

9. Pick a Δ and the corresponding number of significant genes.

10. The FDR is estimated by the median (or the 90th percentile) of the number of falsely called genes divided by the number of significant genes.

Estimation of the SAM Fudge Factor s_0

In the procedure described above, a fudge factor s_0 in the denominator of the test statistic (12) is used. It is calculated as the percentile of the gene-wise standard error distribution that minimizes the coefficient of variation (CV) of the test statistics. This modification is used to overcome bias for genes with expressions close to zero, which have a large value of the test statistic due to a small sample variance. By using an inflated standard error, SAM addresses the problem of the dependence of the value of the test statistic on the variance of expression levels for a particular gene. The calculation of s_0 is as follows:

1. Let s_α be the $\alpha \cdot 100\%$ percentile of s_i values. Let $t_i^\alpha = (\bar{X}_1 - \bar{X}_0)/(s_i + s_\alpha)$.

2. Compute the 100 centiles of the s_i values, denoted by $q_1 < q_2 \cdots < q_{100}$.

3. For $\alpha \in (0, 0.05, 0.10, \ldots, 1.0)$

   (a) compute $\nu_j = \text{MAD}(t_i^\alpha | s_i \in [q_j, q_{j+1})), j = 1, 2, \ldots, m$, where MAD is the median absolute deviation from the median, divided by .64;

   (b) compute $cv(\alpha) =$ coefficient of variation of the ν_j values.

4. Choose $\hat{\alpha} = argmin[cv(\alpha)]$, i.e., $\hat{\alpha}$ is the quantile of the standard error that minimizes the coefficient of variation of the SAM test statistics.

5. Compute $\hat{s}_0 = s_{\hat{\alpha}}$.

7 Using the IsoGene Package

7.1 Data Example

The data used for the analysis presented below are outcome of a dose-response microarray experiment consisting of four dose levels. Three microarray samples are available at each dose level (hence, in total gene expression was measured for 12 arrays). Each array consists of 16,998 genes.

A dataframe with the log2 transformed gene intensities is loaded into *R* environment. The first ten genes and first six samples are displayed, where the row names of the genes show the probe ID, `X1, X1.1 and X1.2` are the three arrays for dose zero, while `X2, X2.1 and X2.2` are the arrays for the first dose. The dataframe is loaded suing the function `load()`,

```
> load("data.Rdata")
```

A printout of the first 10 lines is given below.

```
> data[1:10,1:6]
                X1     X1.1     X1.2       X2     X2.1     X2.2
g1 6.923109 7.024719 7.170328 7.219297 7.076908 7.404949
g2 5.107275 5.092935 5.255918 5.312913 4.893855 4.596591
g3 5.913526 6.026197 5.141728 5.828770 5.269202 5.461664
g4   4.919469 4.908159 3.500307 4.814068 4.139949 4.278321
g5   6.002091 5.878718 5.777668 6.214799 5.895586 6.163291
g6 7.162715 7.294693 6.903935 7.223069 6.972928 7.412160
g7   4.049696 4.748409 3.845498 4.780287 4.076589 4.300242
g8   3.191931 4.326571 3.771206 3.570291 2.179324 3.988911
g9   6.487708 6.285804 6.229814 6.109103 6.340837 5.931840
g10   6.695870 6.687039 6.652153 6.503670 6.387794 6.698711
```

7.2 Loading the Package

To load the `IsoGene` package into *R*, a binary zip-package of `IsoGene` program (for Windows) needs to be installed. `IsoGene` package requires *R* packages `Multtest` and `ff`, which need to be installed as well. Once the packages are installed, they are available for use after being loaded in memory, which is usually done by the user:

```
> library(IsoGene)
```

```
Iso 0.0-8
Note: This package now has a NAMESPACE.
Loading package ff 2.1-2
- getOption("fftempdir")=="C:/DOCUME~1/lucp1898/LOCALS~1/Temp/RtmpTKLDbm"
- getOption("ffextension")=="ff"
- getOption("ffdrop")==TRUE
- getOption("fffinonexit")==TRUE
- getOption("ffpagesize")==65536
- getOption("ffcaching")=="mmnoflush"  -- consider "ffeachflush" if your system stalls on large
- getOption("ffbatchbytes")==16095641 -- consider a different value for tuning your system
Attaching package ff
```

The functions included in the package can be listed using the R help system:

```
> help(package = IsoGene)
```

First, `IsoPlot()` can be used to explore the data. Second, `IsoGene1()` and `IsoGenem()` can be used to calculate the test statistics. Third, `IsoRawp()` provides the output for two-sided or one-sided p-values (p^{Up} or p^{Down}). Based on the p-values obtained, one can choose one test statistic and multiplicity adjustment method for inference by using `IsoTestBH()`. Finally, `IsopvaluePlot()` can be useful for examining both of p^{Up}- or p^{Down}-values, and in particularly, as a post hoc procedure it can be used to examine genes with both small p^{Up}- and p^{Down}-values.

8 The IsoGene Functions

8.1 Quick Start

The first stage of the analysis (which is also the time consuming stage) consists of permutations under the null hypothesis in order to obtain the distribution of the test statistic under the null hypothesis. Note that, by default, all five test statistics discussed above are calculated. The function `IsoRawp()` is used to perform the permutation. A general call of the function `IsoRawp()` has the form of

```
IsoRawp(x, data, niter=1000)
```

Here, `x` is a vector which contains the dose levels and `data` is the R object, which contains the information about gene expression and genes names. Once the permutation stage is completed, the FDR adjusted p-values can be obtained using function `IsoTestBH()`. The function calculates the adjusted p values for each statistic using either the BH-FDR or BY-FDR for multiplicity adjustment. The user can specify one of the five test statistics discussed above, or use the default call, in the later case adjusted p-values for all test statistics will be calculated. A general form of the `IsoTestBH()` has the form

```
IsoTestBH(rp, FDR=c(0.05,0.1), type=c("BH","BY"),
stat=c("E2","Williams","Marcus","M","ModifM"))
```

Note that `rp` is an R object, which contains all the output produced by the function `IsoRawp()`.
In what follows we illustrate in more details the use of the functions of the `IsoGene` package.

8.2 Exploring the Data

`IsoPlot()` can be used to explore the data. Scatterplots for the second gene in the dataset (`data[2,]`) can be produced by

```
> data(exampleData)

> x <- c(rep(1, 3), rep(2, 3), rep(3, 3), rep(4, 3))
> gene1 <- as.matrix(exampleData[2, ])
> par(mfrow = c(1, 2))
> IsoPlot(x, y = gene1)
> IsoPlot(x, y = gene1, type = "ordinal", add.curve = TRUE)
```

The left panel in Figure 3 shows the original data points (as circles) and sample means (as pluses) for each dose. The right panel in Figure 3 shows the increasing isotonic regression model (blue solid line) fitted on the data. The fitted monotonic line does not indicate the significance of the test, but simply shows a more likely increasing (or decreasing) trend.

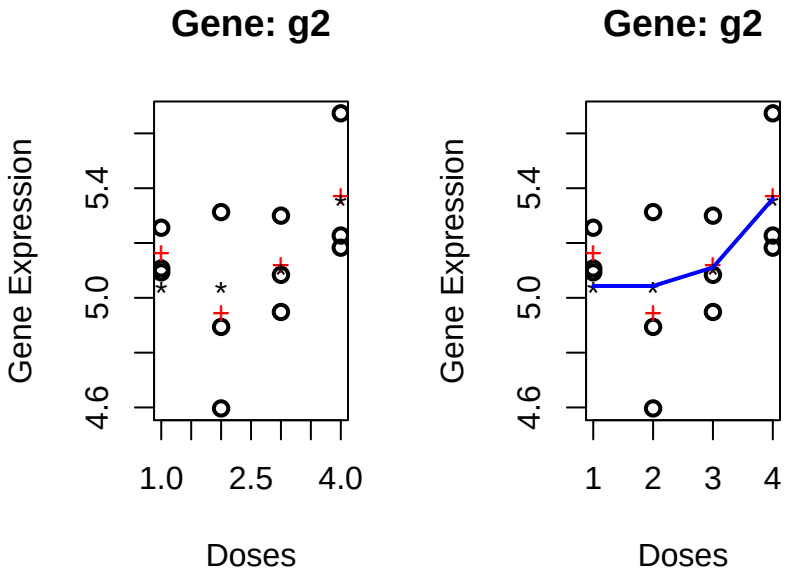

Figure 3: *The data points are plotted as circles, while sample means as pluses. The right panel additionally plots the fitted increasing isotonic regression model (blue solid line).*

8.3 Calculating the Test Statistics

The five test statistics described in Chapter 2 can be obtained by using the function `IsoGene1()` for a single gene and using the function `IsoGenem()` for all the genes simultaneously. The following *R* codes illustrate the input and output generated by these two functions:

```
> stat1 <- IsoGene1(x, gene1)
```

The object `stat1` contains the information about the five test statistics and the direction for which the likelihood is maximizes.

```
> stat1

$E2.up
[1] 0.2697894
```

```
$Williams.up
[1] 1.040134

$Marcus.up
[1] 1.581191

$M.up
[1] 1.205802

$ModM.up
[1] 1.278946

$E2.dn
[1] 0.0008106545

$Williams.dn
[1] -0.08238646

$Marcus.dn
[1] -0.08238646

$M.dn
[1] -0.05370908

$ModM.dn
[1] -0.06004858

$direction
[1] "u"
```

The first 10 objects are the values calculated for the five test statistics under increasing and decreasing trends. The last object indicates the higher likelihood of isotonic regression with “u” meaning a increasing trend or “d” meaning a decreasing trend.

We use the first 10 genes as an example to illustrate the use of function IsoGenem():

```
> statm <- IsoGenem(x, exampleData[1:10, ])
> statm

$E2.up
        g1         g2         g3         g4         g5         g6         g7
0.81527841 0.26978939 0.81226244 0.04625381 0.02356596 0.01270386 0.00000000
        g8         g9        g10
0.05438655 0.02259598 0.99060602

$Williams.up
 [1]  5.706246398  1.040133583  4.532537020  0.485215821  0.546648138
 [6]  0.357615051 -0.134476246 -0.009111866 -0.693505488 28.489753149

$Marcus.up
 [1]  5.7062464  1.5811905  5.0052320  0.5279040  0.5466481  0.3576151
 [7]  0.0000000  0.5170096  0.4308062 29.4025214

$M.up
 [1]  4.6591307  1.2058018  3.9221693  0.4308356  0.2929374  0.2138936
 [7]  0.0000000  0.3916276  0.2867030 20.1706306

$ModM.up
 [1]  4.6591307  1.2789459  4.3851186  0.4569702  0.3275140  0.2391403
 [7]  0.0000000  0.4378530  0.3205437 21.3941846

$E2.dn
          g1           g2           g3           g4           g5           g6
0.0000000000 0.0008106545 0.0000000000 0.0000000000 0.3328630108 0.3126307156
          g7           g8           g9          g10
0.0902853975 0.0194293995 0.2236936672 0.0000000000

$Williams.dn
 [1]  2.47954450 -0.08238646  0.77861303  0.15969984 -1.08594004 -1.06231999
```

```
 [7] -0.45383921 -0.35682222 -1.35548019  7.55960274

$Marcus.dn
 [1]  0.00000000 -0.08238646  0.00000000  0.00000000 -2.16867670 -1.77404170
 [7] -0.63872592 -0.35682222 -1.35548019  0.00000000

$M.dn
 [1]  0.00000000 -0.05370908  0.00000000  0.00000000 -1.40596911 -1.27167021
 [7] -0.51444688 -0.26542632 -1.01219460  0.00000000

$ModM.dn
 [1]  0.00000000 -0.06004858  0.00000000  0.00000000 -1.49125543 -1.42177051
 [7] -0.57516910 -0.29675565 -1.13166796  0.00000000

$direction
 g1  g2  g3  g4  g5  g6  g7  g8  g9 g10
"u" "u" "u" "u" "d" "d" "d" "u" "d" "u"
```

The output from `IsoGenem()` has the same structure as the one for `IsoGene1()`, but each object contains the values of the test statistics and the likely direction of the isotonic regression model for all the genes.

8.4 Obtaining Raw p-values

As discussed above, we use permutations to obtain the raw p-values for the five test statistics. The function `IsoRawp()` can be used in the following way:

```
> rawp <- IsoRawp(x = x, y = exampleData, niter = 2)
```

The four arguments in this function need to be specified, with no default pre-specified values. `x` is the explanatory variable indicating the dose levels for all the samples in the data. `data` is the data frame of gene expression values. `niter` defines the number of permutations used to approximate the null distribution. The output item `rawp` contains four objects with p-values for the five test statistics: the first one contains the two-sided p-values, the second contains the one-sided p-values, the third contains p^{Up}-values, and the last one contains p^{Down}-values. Below we print a part of the object with two-sided p-values for illustration:

```
> rawp.twosided <- rawp[[1]]
```

The first 10 rows in of the object `rawp.twosided` are

```
> rawp.twosided[1:10, ]
```

```
   Probe.ID  E2 Williams Marcus   M ModM
1        g1 0.0      0.0    0.0 0.0  0.0
2        g2 0.5      0.5    0.5 0.5  0.5
3        g3 0.0      0.0    0.0 0.0  0.0
4        g4 0.0      0.0    0.0 0.0  0.0
5        g5 0.0      0.0    0.0 0.0  0.0
6        g6 0.0      0.5    0.5 0.5  0.0
7        g7 0.5      0.5    0.5 0.5  0.5
8        g8 0.0      0.0    0.0 0.0  0.0
9        g9 0.0      0.0    0.0 0.0  0.0
10      g10 0.0      0.0    0.0 0.0  0.0
```

The first output object from `rawp` is a matrix with six columns, where the first column indicates the probe ID. Columns from the second to the sixth are p-values for each of the five test statistics, respectively. The remaining three output objects (`rawp`[[2]], `rawp`[[3]], `rawp`[[4]]) are structured in the same way.

8.5 Plot of p-values for a Single Gene

For a single gene, the function `IsopvaluePlot()` can be used to show the p^{Up} and p^{Down}-values for a given test statistic:

```
> IsopvaluePlot(x, y, niter, stat = c("E2", "Williams", "Marcus",
+     "M", "ModifM"))
```

We use one gene as an example to illustrate how p^{Up} and p^{Down}-values (in the upper and lower panels of Figure 4) are obtained. In Figure 4, the observed test statistics are drawn as the dashed line, and the values of the test statistics obtained from permutations are spread over the x-axis. For this gene, the p^{Up} is much smaller as compared to the p^{Down} since $T^{Up} \gg T^{Down}$, which implies a possible increasing trend in the data.

```
> IsopvaluePlot(x, gene1, niter = 1000, stat = "E2")
```

1 . 2 . 3 . 4 . 5 . 6 . 7 . 8 . 9 . 10 . 11 . 12 . 13 . 14 . 15 . 16 . 17 . 18 . 19 . 20 . 21 .

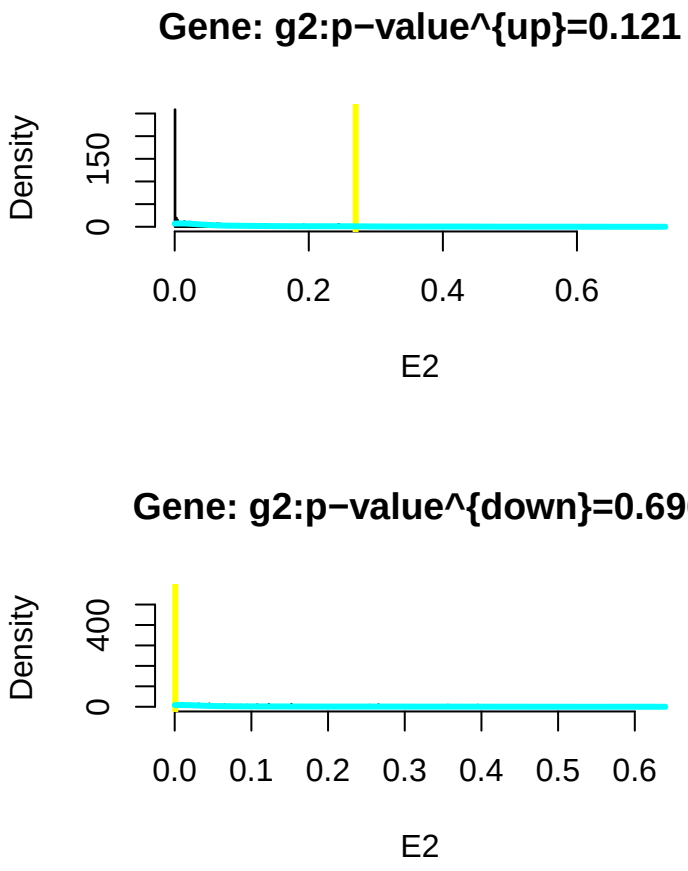

Figure 4: *The p^{Up} and p^{Down}-values using $\bar{E}^2_{01}$ for an example gene. The dashed line is the observed test statistic value. In the upper panel, the dashed line (at the right) is larger than most of the test statistics from permutations, which results in a small p^{Up}-value. In the lower panel, the dashed line (close to zero) is smaller than most of the test statistics from permutations, which results in a large p^{Down}-value.*

8.6 BH/BY-FDR Procedures for Adjusting for Multiple Testing

With the two-sided p-values, the user needs to select one of the five test statistics, the FDR level, and the type of multiplicity adjustment (BH-FDR or BY-FDR) to obtain the list of significant genes:

```
> IsoTestBH(rp, FDR = c(0.05, 0.1), type = c("BH", "BY"), stat = c("E2",
+     "Williams", "Marcus", "M", "ModifM"))
```

The following example shows the use of the global likelihood ratio test $\bar{E}_{01}^2$, the FDR level of 0.05 and the BH-FDR procedure controlling the FDR:

```
> E2.BH <- IsoTestBH(rawp.twosided, FDR = 0.05, type = "BH", stat = "E2")
```

The first 10 rows in the object `E2.BH` list the sorted row and adjusted p values for the $\bar{E}_{01}^2$ statistic.

```
> E2.BH[1:10, ]
```

	Probe.ID	row.name	raw p-values	BH adjusted p values
1	g1	1	0	0
2	g3	3	0	0
3	g4	4	0	0
4	g5	5	0	0
5	g6	6	0	0
6	g8	8	0	0
7	g9	9	0	0
8	g10	10	0	0
9	g11	11	0	0
10	g12	12	0	0

Here we show only the first ten genes declared significant by using $\bar{E}_{01}^2$ test. The output results in a matrix of five columns: the first column indicates the probe ID, the second column is the corresponding row number of significant genes in the original dataset, the third column is the unadjusted/raw p-value, and the last column is the

adjusted p-value using the requested "BH" procedure. The order of the list of genes found significant is based on the row number. Moreover, the function `IsoBHPlot()` can be used to visualize the number of significant findings for the BH-FDR and BY-FDR procedures for the specified test statistic:

```
> IsoBHPlot(rp, FDR = c(0.05, 0.1), stat = c("E2", "Williams",
+     "Marcus", "M", "ModifM"))
```

Figure 6 shows the unadjusted (solid blue line) and the BH-FDR (dotted and dashed red line) and BY-FDR (dashed green line) adjusted p-values for $\bar{E}^2_{01}$. It is obtained using the function `IsoBHPlot()`:

```
> IsoBHPlot(rawp.twosided, FDR = 0.05, stat = "E2")
```

9 Significance Analysis of Dose-response Microarray Data (SAM)

In this package, we also implement the significance analysis of microarray (SAM) for testing for the dose-response relationship under order restricted alternatives. The SAM procedure was proposed by Tusher *et al.* (2001) for finding differentially expressed genes by using permutations while controlling for the FDR.

The main function for the SAM procedure is `IsoTestSAM()`. Within the main function, `Isofudge()` calculates the fudge factor in the SAM test statistic, `IsoGenemSAM()` is used to obtain the values of SAM test statistics, `Isoqqstat()` calculates the SAM test statistic for the required number of permutations specified by users, `Isoallfdr()` obtains the delta table in the SAM procedure, `Isoqval()` computes q-values of the SAM.

The syntax of these function is as follows,

```
> Isofudge(x, y)
> IsoGenemSAM(x, y, fudge.factor)
> Isoqqstat(x, y, fudge = c(0, "pooled"), niter = 100)
> Isoallfdr(qqstat, ddelta, stat = c("E2", "Williams", "Marcus",
+     "M", "ModifM"))
> Isoqval(delta, allfdr, qqstat, stat)
```

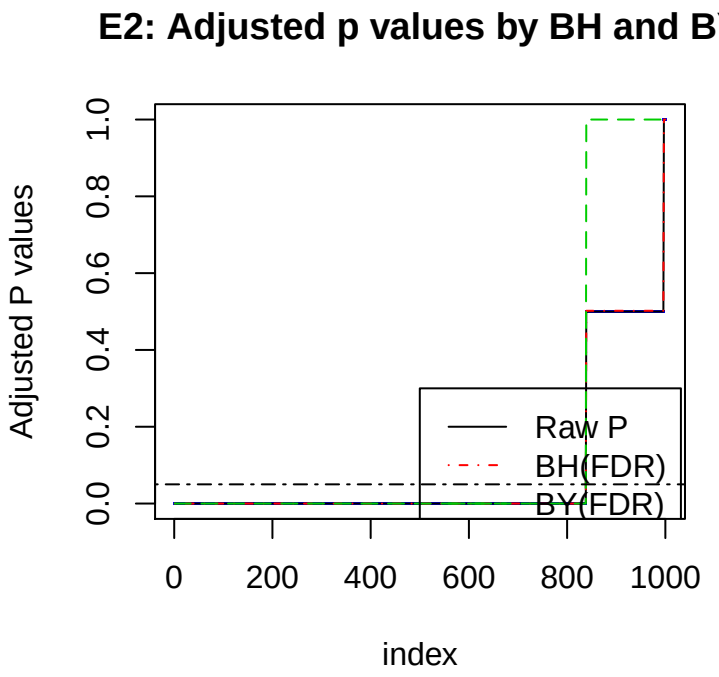

Figure 5: *The unadjusted (solid blue line), and the BH-FDR (dotted and dashed red line) and BY-FDR (dashed green line) adjusted p-values for* $\bar{E}_{01}^2$.

We use the same data as above to obtain the fudge factor of the SAM procedure for the five test statistics by using function `Isofudge()`.

```
> fudge.factor <- Isofudge(x, exampleData)
```

The output of this function gives a vector of five fudge factors for each of the test statistics.

```
> fudge.factor

[1] 0.07794229 0.16687744 0.10486056 0.19962253 0.12201373
```

Note that the fudge factor of the $\bar{E}_{01}^2$ is obtained based on the algorithm for F test statistics given by Tusher *et al.* (2001) and should be used with cautions. The

performance of using the fudge factor as compared to the t-type test statistics has not yet investigated in term of power and control of the FDR. Therefore, it's advisable to use the fudge factor in the t-type test statistics.

```
> SAMtest.stat <- IsoGenemSAM(x, exampleData, fudge.factor)
```

The output of the function gives

```
> names(SAMtest.stat)

[1] "E2"        "Williams"  "Marcus"    "M"         "ModM"      "direction"
```

The following codes produce the values of the five test statistics for the first ten genes,

```
> SAMtest.stat[[1]][1:10]

        g1         g2         g3         g4         g5         g6         g7
0.76307188 0.24474561 0.79981561 0.04428623 0.25513904 0.26372825 0.08792619
        g8         g9        g10
0.05322140 0.16347105 0.98418205

> SAMtest.stat[[2]][1:10]

 [1]  2.524719216  0.568405925  2.795837275  0.335299641 -0.393142350
 [6] -0.477446410 -0.333583454 -0.006801912 -0.529556199  9.327805999

> SAMtest.stat[[3]][1:10]

 [1]  3.1845748  1.0392371  3.6000422  0.4121190 -1.0291185 -1.0024228
 [7] -0.5207607  0.4260847 -0.6845731 12.8347836

> SAMtest.stat[[4]][1:10]

 [1]  2.0885434  0.6862569  2.4788191  0.2999204 -0.5940818 -0.6202021
 [7] -0.3818276  0.2994731 -0.4229476  7.5100918

> SAMtest.stat[[5]][1:10]

 [1]  2.6588739  0.8578876  3.1369183  0.3561781 -0.7907050 -0.8276609
 [7] -0.4648386  0.3617761 -0.5797297 10.2222441
```

```
> SAMtest.stat[[6]][1:10]

 g1  g2  g3  g4  g5  g6  g7  g8  g9 g10
"u" "u" "u" "u" "d" "d" "d" "u" "d" "u"
```

To obtain the SAM test statistics for one of five test statistic values, for example, the modified M test, with the required number of permutations specified by users and compute the delta table in the SAM procedure, we can use function `Isoqqstat()` and `Isoallfdr()` as follows,

```
> qqstat <- Isoqqstat(x, exampleData, fudge = "pooled", niter = 2)
> dtable <- Isoallfdr(qqstat, , stat = "ModifM")

> dim(dtable)

[1] 121   6

> head(dtable)

     Ddelta FalsePositive50% FalsePositive90% Called FDR50% FDR90%
[1,]   0.01          547.932         558.0344    929 0.5898 0.6007
[2,]   0.11          470.008         470.5008    866 0.5427 0.5433
[3,]   0.21          409.024         410.0096    800 0.5113 0.5125
[4,]   0.31          369.292         373.4808    754 0.4898 0.4953
[5,]   0.41           69.608          77.4928    404 0.1723 0.1918
[6,]   0.51           40.040          46.4464    339 0.1181 0.1370
```

Note that in `Isoallfdr()`, ddelta is left blank, with default value taken from the data, i.e., all the percentiles of the standard errors. By fixing the 50% FDR at 0.05, the corresponding delta value is 0.83 (marked in-between the dashed lines) as we obtain from the delta table above, the number of differentially expressed genes are 872 with potential 42 genes as false positives.

```
> qval <- Isoqval(delta = 0.83, allfdr = dtable, qqstat = qqstat,
+     stat = "ModifM")
> dim(qval[[1]])

[1] 1000    3
```

```
> head(qval[[1]])

     Row.names    t.stat q.val
g987       987 -6.742283     0
g409       409 -6.717422     0
g374       374 -6.602694     0
g229       229 -5.569383     0
g963       963 -4.952691     0
g962       962 -4.813606     0

> dim(qval[[2]])

[1] 186   3

> head(qval[[2]])

     Row.names    t.stat q.val
g987       987 -6.742283     0
g409       409 -6.717422     0
g374       374 -6.602694     0
g229       229 -5.569383     0
g963       963 -4.952691     0
g962       962 -4.813606     0
```

By specifying the desired delta value, delta table, and the user-defined test statistic in function `Isoqval()`, we can obtain the q value of each gene from the SAM procedure. The first object of the output is the list of q values for all the genes, ranking from the smallest test statistic value to the largest; while the second object is the list of q values for the 872 differentially expressed genes, ranking from the smallest test statistic value to the largest. The first column of the output matrices is the row number of genes in the data set, the second column is the observed modified M test statistic value, and the last column is the q value of the SAM procedure for both objects.

Alternatively, we can use function `IsoTestSAM()` to summarize all the steps above and give results of a list of significant findings, which is the same second output of function `Isoqval()`.

```
> IsoTestSAM(x, y = data, fudge = c(0, "pooled"), niter = 100,
+     FDR = 0.05, stat = c("E2", "Williams", "Marcus", "M", "ModifM"))
```

Specifying the same options as above in this function, we can obtain the list of significant genes as follows: the first column is the Probe.ID, the second column is the corresponding row numbers of the genes in the data set, the third column is the observed modified M SAM test statistics, the fourth column is the q values of genes by using the SAM procedure. The last two columns gives additional information by calculating the p values based on the SAM permutation matrix and adjusting these p values using the BH-FDR procedure.

```
> IsoSAM.obj <- IsoTestSAM(x, y = exampleData, fudge = "pooled",
+     niter = 2, FDR = 0.05, stat = "ModifM")
```

The resulting object `IsoSAM.obj`, contains three components:

1. `sign.genes1` contains a list of genes declared significant using the SAM procedure.
2. `qqstat` gives the SAM regularized test statistics obtained from permutations.
3. `allfdr` provides a delta table in the SAM procedure for the specified test statistic.

To extract the list of significant gene, one can do:

```
> IsoSAM.sig <- IsoSAM.obj[[1]]
> dim(IsoSAM.sig)

[1] 323   6

> head(IsoSAM.sig)

  Probe.ID row.number  stat.val qvalue pvalue adj.pvalue
1     g987        987 -6.742283      0      0          0
2     g409        409 -6.717422      0      0          0
3     g374        374 -6.602694      0      0          0
4     g229        229 -5.569383      0      0          0
5     g963        963 -4.952691      0      0          0
6     g962        962 -4.813606      0      0          0
```

Finally, the graphic output of the SAM procedure can be produced using function `IsoSAMPlot()`.

```
> IsoSAMPlot(qqstat, allfdr, FDR = 0.05, stat = c("E2", "Williams",
+     "Marcus", "M", "ModifM"))
```

This function requires the use of output from `Isoqqstat()` and `Isoallfdr()`, given a user-defined test statistic and the FDR level to control. We still take the modified M test statistic for example, at the FDR of 0.05. There are four plots yielded from the SAM procedure. Panel *a* shows the FDR (either 50% or 90% (more stringent)) vs. Δ, from which, user can choose the delta value with the corresponding desired FDR. Panel *b* shows the number of significant genes vs. Δ, and panel *c* shows the number of false positives (either 50% or 90%) vs. Δ. Finally panel *d* shows the observed vs. the expected (obtained from permutations) test statistics, in which the red dots are those genes called differentially expressed.

```
> IsoSAMPlot(qqstat = qqstat, allfdr = dtable, FDR = 0.05, stat = "ModifM")
```

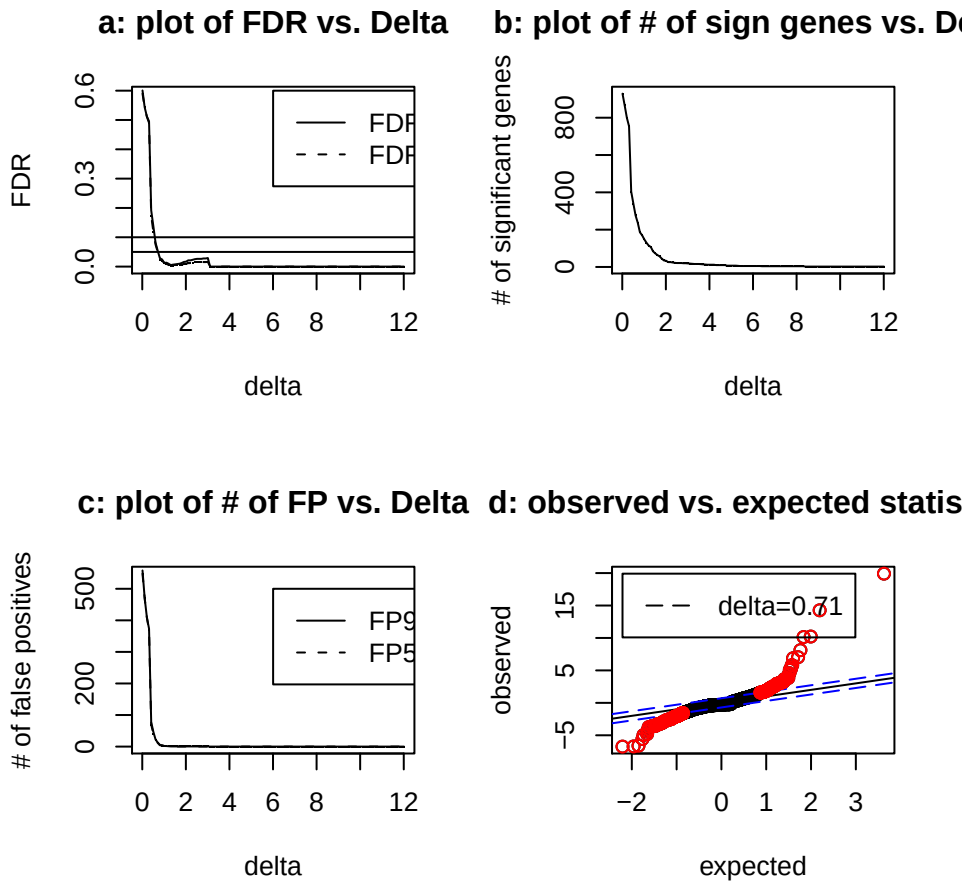

Figure 6: *The SAM plots: a.Plot of the FDR vs. delta; b. Plot of number of significant genes vs. delta; c. Plot of number of false positives vs. delta; d. Plot of observed vs. expected test statistics.*

References

Agresti, A. (1997) *Statistical Methods for the Social Sciences*, Finlay.

Barlow, R.E., Bartholomew, D.J., Bremner, M.J. and Brunk, H.D. (1972) *Statistical inference under order restriction*, New York: Wiley.

Bartholomew, D.J. (1961) Ordered tests in the analysis of variance, *Biometrika*, **48**, 325–332.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of Royal Statistical Soceity, Biostatistics*, **57**, 289–300.

Benjamini, Y. and Yekutieli, D. (2001) The control of the false discovery rate in multiple testing under dependency. *Annal of Statistics*, **29(4)**, 1165–1188.

Benjamini, Y. and Yekutieli, D. (2005) False Discovery Rate-Adjusted Multiple Confidence Intervals for Selected Parameters. *Journal of the American Statistical Association*, **100**, 71–81.

Chuang-Stein, C. and Agresti, A. (1997) Tutorial in biostatistics: A review of tests for detecting a monotone dose-response relationship with ordinal response data, *Statistics in Medicine*, **16**, 2599–2618.

Ge, Y., Dudoit, S. and Speed, P.T. (2003) Resampling based multiple testing for microarray data analysis, *technical report*, **633**, University of Berkeley.

Hu, J., Kapoor, M., Zhang, W., Hamilton, S.R. and Coombes, K.R. (2005) Analysis of dose response effects on gene expression data with comparison of two microarray platforms, *Bioinformatics*, **21(17)**, 3524–3529.

Lin, D., Shkedy, Z., Yekutieli, D., Burzykowki, T., Göhlmann, H.W.H., De Bondt, A., Perera, T., Geerts, T., Bijnens, L. (2007a) Testing for trend in dose-response microarray experiments: comparison of several testing procedures, multiplicity, and resampling-based inference. *Statistical Application in Genetics and Molecular Biology*, **6(1)**, article 26.

Lin, D., Shkedy, Z., Burzykowki, R., Ion, T., Göhlmann, H.W.H., De Bondt, A., Perera, T., Geerts, T., Bijnens, L. (2008) An investigation on performance of

Significance Analysis of Microarray (SAM) for the comparisons of several treatments with one control in the presence of small variance genes. *Biometrical Journal*, Multiple Comparison Problem, Special Issue, **50(5)**, 801–823.

Marcus, R. (1976) The powers of some tests of the quality of normal means against an ordered alternative, *Biometrika*, **63**, 177–83.

Reiner, A., Yekutieli, D. and Benjamini, Y. (2003) Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics*, **19(3)**, 368–375.

Robertson, T., Wright, F.T. and Dykstra, R.L. (1988) *Order restricted statistical inference*, Wiley.

item Ruberg, S.J. (1995a) Dose response studies. I. Some design considerations. *Journal of Biopharmaceutical Statistics*, **5(1)**, 1–14.

Ruberg, S.J. (1995b) Dose response studies. II. Analysis and interpretation. *Journal of Biopharmaceutical Statistics*, **5(1)**, 15–42.

Tusher, V.G., Tibshirani, R. and Chu, G. (2001) Significance analysis of microarrys applied to the ionizing radiation response, *Proceedings of the National Academy of Sciences*, **98** , 5116–5121.

Williams, D.A. (1971) A test for differences between treatment means when several dose levels are compared with a zero dose control, *Biometrics*, **27**, 103–117.

Williams, D.A. (1972) The comparision of several dose levels with a zero dose control, *Biometrics*, **28**, 519–531.