2013

A Bayesian decision theoretic approach to directional multiple hypotheses problems

Naveen Bansal, *Marquette University*

A Bayesian Decision Theoretic Approach to Directional Multiple Hypotheses Problems

Naveen K. Bansal

Department of Mathematics, Statistics, and Computer Science, P.O. Box 1881, Marquette University, Milwaukee, WI 53201, U.S.A.

naveen.bansal@marquette.edu

and

Klaus J. Miescke

Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL 60680, U.S.A.

klaus@uic.edu

Abstract

A multiple hypothesis problem with directional alternatives is considered in a decision theoretic framework. Skewness in the alternatives is considered, and it is shown that this skewness permits the Bayes rules to possess certain advantages when one direction of the alternatives is more important or more probable than the other direction. Bayes rules subject to certain constrains on the directional false discovery rates are obtained, and their performances are compared with a traditional FDR rule through simulation. We also analyzed a gene expression data using methodology developed, and compare the results to that of FDR method.

AMS Subject Classification: 62C10; 62H15

Key Words: Directional Alternatives; False Discovery Rate; Microarray data analysis; Multiple hypotheses.

1 Introduction

Multiple hypotheses problems have received a significant amount of attention in the recent literature due to its use in microarray data analysis, imaging analysis, and other biological applications. In these applications, typically, data can be formulated in the form of a $m \times n$ matrix $\mathbf{X}$ whose ith row $\mathbf{x}_i$ is a sample from a probability model $P(\cdot|\theta_i,\ \eta)$, where θ_i is a parameter of interest, and η is a nuisance parameter. Most of the work in literature focuses on the two-tail hypotheses $H_0^i : \theta_i = 0$ vs.

$H_a^i : \theta_i \neq 0,\ i = 1, 2, \cdots, m$. In this paper, we consider the following hypotheses with directional alternatives

$$H_0^i : \theta_i = 0 \quad vs \quad H_-^i : \theta_i < 0 \ \ or \quad H_+^i : \theta_i > 0,\ i = 1, 2, ..., m. \tag{1}$$

Earlier work on this problem was based on the familywise error rate, FWER [4, 5]. However, since this approach controls false discovery of even one null, such an approach is not practical for higher dimensional problem, i.e., when m is very large. Schaffer [9], Lewis and Thayer [6] and Sarkar and Zhou [10] focussed on directional false discovery rate (DFDR), an analogous version of the false discovery rate (FDR) [2] for the directional hypotheses. We focus on a Bayesian decision theoretical formulation of this problem with an emphasis on the directional false discoveries with skewed alternatives as described in Section 2.

The loss function, we consider, is of the form

$$L(\boldsymbol{\theta}, \mathbf{d}^X) = \sum_{i=1}^{m} L_i(\theta_i, d_i^X), \tag{2}$$

where $L_i(\theta_i, d_i^X)$ is the loss for each individual hypothesis $H_0^i\ vs.\ H_-^i$ or H_+^i, and $\mathbf{d}^X = (d_1^X, d_2^X, ..., d_m^X)$, with $d_i^X \in \{-1, 0, 1\}$. Here, $d_i^X = -1$ means that H_-^i is selected, $d_i^X = 0$ means that H_0^i is selected, and $d_i^X = 1$ means that H_+^i.

One of the advantages of the Bayesian decision theoretic approach is that it allows us to incorporate the prior information on the direction of $\theta_i s$ as it may be relevant in many applications. We consider the prior of the form

$$\pi(\theta) = p_- \pi_-(\theta) + p_0 I(\theta = 0) + p_+ \pi_+(\theta), \tag{3}$$

were $p_-,\ p_0$, and p_+ are some preassigned or estimated probabilities with $p_- + p_0 + p_+ = 1$, and $\pi_-(\cdot)$ and $\pi_+(\cdot)$ are densities with support $(-\infty, 0)$ and $(0, \infty)$, respectively. This prior allows skewness in the distribution of $\theta_i s$. If, for example, $p_+ > p_-$, then chances are higher that more of $\theta_i s$ are generated with H_+ than with H_-. This may be appropriate in many applications. In particular, it is important in genetic experiments in which a microRNA (a non-coding gene) is transfected in cells of interest to test if it suppresses the gene expressions of mRNAs. The biological theory is that certain non-coding genes get attched to a subsequence of specific coding genes and thus suppress their expressions (citation). We will consider the data cited in (citation) in section 7, and analyze it using methodology developed in this paper.

The theorem below shows the importance of the prior (3), which is a straightforward generalization of Theorem 1 of [1], and thus stated without proof.

Theorem 1 *Let $\pi(\theta)$ and $\pi'(\theta)$ be two priors given by*

$$\begin{aligned} \pi(\theta) &= p_-\pi_-(\theta) + p_0 I(\theta = 0) + p_+\pi_+(\theta) \\ \pi'(\theta) &= p'_-\pi_-(\theta) + p_0 I(\theta = 0) + p'_+\pi_+(\theta), \end{aligned}$$

where $p_- < p'_-$ and thus $p_+ > p'_+$, and let δ_B and δ'_B denote the Bayes rules under the loss (2) with respect to π and π' respectively. Then

$$r_+^{\delta_B} \leq r_+^{\delta'_B}, \tag{4}$$

but

$$r_-^{\delta_B} \geq r_-^{\delta'_B}, \tag{5}$$

where r_-^{δ} and r_+^{δ} are the average Bayes risk of the rule δ with respect to π_- and π_+ respectively.

If $\pi(\theta)$ is the true prior with $p_+ > p_-$, but the prior $\pi'(\theta)$ is considered instead with $p'_- = p'_+$. Then the Theorem 1 implies that the Bayes rule so obtained performs poor in the right tail of $\theta_i s$ which occurs more frequently. If $L_i(\theta_i, d_i)$ is the "0-1" loss, then this implies that the number of false discoveries by δ'_B in the positive region of $\theta_i s$ will be higher than by δ_B. For the non-coding example, this would mean that there will be more false discoveries if the true skewness in the prior is not taken into account.

Most of the research work on multiple hypothesis problems focus on the false discovery rate (FDR). However, in this paper, we focus on the optimality of decision rule in a Bayesian decision theoretic framwork as advocated by Muller, Parmigiani and Rice [7]. We further impose different notions of directional false discovery rates to obtain constrained Bayes decision rule. The main advantage of this approach is that a specific prior information, in particular the skewness in the prior, can be utilize to obtain better decision rules. Generally, if the true skewness in the prior is not taken into account, then there is a greater chance of false discovery as we noted in the remarks following Theorem 1.

The rest of the paper is as organized as follows. In Section 2, we present the definitions of the false discovery rates for the directional multiple hypotheses. We present a general methodology of Bayes rule in Section 3. Specifically how to compute the Bayes rule is discussed in Section 4, including an example of normal populations. Simulation results of comparing the Bayes rules and the FDR type rules are discussed in Section 5. We end with some concluding remarks in Section 6.

2 False Discovery Rates

Let $m_0,\ m_-$, and m_+ be true numbers of $H_0,\ H_-$, and H_+ hypotheses, respectively, and let $U,\ V$, and W be the number of $H_0,\ H_-$, and H_+ selected by a decision rule $\mathbf{d}$. The Table 1 shows different possibilities of the decisions. Different types of false discoveries are made; for example, V_0 and V_+ together represent the total number of falsely selected H_- hypotheses. Similarly, W_0 and W_+ together represent the total number of falsely selected H_+. Following Shaffer (2002), we define the directional false discovery rate $DFDR$ as

$$DFDR = E\left[\frac{V_0 + V_+ + W_0 + W_-}{V + W + I(V=0)I(W=0)}\right].$$

Table 1

	Accept H_0	Accept H_-	Accept H_+	Total
H_0 is true	U_0	V_0	W_0	m_0
H_- is true	U_-	V_-	W_-	m_-
$H_+$is true	U_+	V_+	W_+	m_+
	U	V	W	m

A different variation of this false discovery rate can be defined. Suppose, one is interested in controlling false discoveries of H_- and H_+ separately. In that case, one may like to control the following left and right false discovery rates, respectively:

$$LFDR = E\left[\frac{V_0 + V_+}{V + I(V=0)}\right] \quad \text{and} \quad RFDR = E\left[\frac{W_0 + W_-}{W + I(W=0)}\right]$$

The quantities $DFDR,\ LFDR$, and $RFDR$ above are defined in a frequentist manner. In other words, the expectations are with respect to $\mathbf{X}$ given $\boldsymbol{\theta}$ and η. A Bayesian analogue of this can be defined if the expectation is with respect to both $\mathbf{X}$ given $\boldsymbol{\theta}$ and η, and with respect to $\boldsymbol{\theta}$ and η. We will call these error rates as $BDFDR, BLFDR$, and $BRFDR$, respectively. Posterior versions of these with respect to the posterior distribution will be denoted by $PDFDR$, $PLFDR$, and $PRFDR$. Note that $V_0 + V_+ = \sum_{i=1}^m I(\theta_i \geq 0) I(d_i^X = -1)$ and $W_0 + W_- = \sum_{i=1}^m I(\theta_i \leq 0) I(d_i^X = 1)$. Thus, it is easy to see that

$$PDFDR = \frac{\sum(\nu_i^0 + v_i^+) I(d_i^X = -1) + \sum(\nu_i^0 + v_i^-) I(d_i^X = 1)}{V + W + I(V = 0) I(W = 0)}. \tag{6}$$

and

$$PLFDR = \frac{\sum(\nu_i^0 + v_i^+) I(d_i^X = -1)}{V + I(V = 0)}, \quad PRFDR = \frac{\sum(\nu_i^0 + v_i^-) I(d_i^X = 1)}{W + I(W = 0)}, \tag{7}$$

where

$$\nu_i^- = P(\theta_i < 0|\mathbf{x}), \qquad \nu_i^+ = P(\theta_i > 0|\mathbf{x}), \qquad \nu_i^0 = P(\theta_i = 0|\mathbf{x}) \tag{8}$$

3 Statistical Methodology

3.1 Bayes Rules under the "0-1" Loss Function

Many special cases of the loss (2) can be considered to reflect the loss in terms of number of false discoveries or false discovery rates. The simplest loss is the "0-1" loss

$$\begin{aligned} L_0(\boldsymbol{\theta}, d) = \sum I(d_i^X \geq 1) I(\theta_i < 0) + &\sum I(d_i^X \leq 1) I(\theta_i > 0) \\ &+ \sum I(d_i^X \neq 0) I(\theta_i = 0). \end{aligned} \tag{9}$$

Note that the expected loss $E[L(\boldsymbol{\theta}, d(\mathbf{X}))]$ is the expected number of total false decisions, $(V_0 + W_0) + (V_+ + W_-) + (U_- + U_+)$. The Bayes rule can be obtained by minimizing the posterior expected loss,

$$E[L_0(\theta, d)|\mathbf{x}] = E[(m_- + m_+)|\mathbf{x}] - \sum(v_i^- - v_i^0) I(d_i^X = -1) - \sum(v_i^+ - v_i^0) I(d_i^X = 1), \tag{10}$$

From (10), it is easy to see that the Bayes rule selects H_0^i if $\max(\nu_i^-, \nu_i^+) \leq \nu_i^0$, selects H_-^i if $\nu_i^- > \nu_i^0$ and $\nu_i^- > \nu_i^+$, and selects H_+^i if $\nu_i^+ > \nu_i^0$ and $\nu_i^+ > \nu_i^-$. Thus with the notations

$$D_1^- = \{i : \nu_i^- > \nu_i^0, \;\; \nu_i^- > \nu_i^+\}, \qquad D_1^+ = \{i : \nu_i^+ > \nu_i^0, \;\; \nu_i^+ > \nu_i^-\}, \tag{11}$$

the Bayes rule can be stated as

$$\delta_B^{(0)} = \begin{cases} \text{Select } H_-^i \text{ for } i \in D_1^- \\ \text{Select } H_+^i \text{ for } i \in D_2^+ \\ \text{Select } H_0^i \text{ for } i \notin D_1^- \cup D_1^+ \end{cases} \tag{12}$$

Note that the Bayes rule $\delta_B^{(0)}$ does not control any false discovery rates. We define the constrained Bayes rule $\delta_B^{(1)}$ as the rule that minimizes the posterior expected loss (10) subject to constrain that $PDFDR \leq \alpha$, where $PDFDR$ is given by (6). Note that $PDFDR \leq \alpha$ also implies that $BDFDR \leq \alpha$.

Now, define $\psi_i = \nu_i^+ + \nu_i^0$ if $i \in D_1^-$ and $\psi_i = \nu_i^- + \nu_i^0$ if $i \in D_1^+$, and rank $\{\psi_i,\ i \in D_1^- \cup D_1^+\}$ from the lowest to the highest. Suppose the ranked values are denoted $\psi_{[1]} \leq \psi_{[2]} \leq ... \leq \psi_{[|D_1^- \cup D_1^+|]}$, where notation $|\cdot|$ is used to denote the cardinality of a set. Denote

$$i_0 = \max\{j \leq |D_1^- \cup D_1^+| : \frac{1}{j}\sum_{i=1}^{j} \psi_{[i]} \leq \alpha\}. \tag{13}$$

If $D_{1\psi} \subseteq D_1^- \cup D_1^+$ denotes the set of indices corresponding to $[1], [2], \ldots, [i_0]$, then it is easy to see, from (6) and (12), that the constrained Bayes rule $\delta_B^{(1)}$ is given by

$$\delta_B^{(1)} = \begin{cases} \text{Select } H_-^i \text{ if } i \in D_1^- \cap D_{1\psi} \\ \text{Select } H_+^i \text{ if } i \in D_1^+ \cap D_{1\psi} \\ \text{Select } H_0^i \text{ if } i \notin (D_1^- \cup D_1^+) \cap D_{1\psi}. \end{cases} \tag{14}$$

In many applications, there may be a need for controlling the false discovery of left and right tail hypotheses H_-^i and H_+^i separately as it may be the case for the non-coding gene example described in the Introduction. The constrained Bayes rule $\delta_B^{(2)}$ in this case can be obtained by minimizing (10) subject to $PLFDR \leq \alpha_L$ and $PRFDR \leq \alpha_R$, where α_L and α_R are some pre-assigned error rates depending upon the risks associated with selecting H_-^i and H_+^i, respectively. To obtain $\delta_B^{(2)}$, first rank $\{\nu_i^{0+} = \nu_i^+ + \nu_i^0,\ i \in D_1^-\}$ from the lowest to the highest and rank $\{\nu_i^{0-} = \nu_i^- + \nu_i^0,\ i \in D_1^+\}$ from the lowest to the highest. Let the ranked values be, respectively, denoted by $v_{[1],-}^{0+} \leq v_{[2],-}^{0+} \leq ... \leq v_{[|D_1^-|],-}^{0+}$ and $v_{[1],+}^{0-} \leq v_{[2],+}^{0-} \leq ... \leq v_{[|D_1^+|],+}^{0-}$, respectively. Now, define

$$i_0^- = \max\{j \leq |D_1^-| : \frac{1}{j}\sum_{i=1}^{j} v_{[i],-}^{0+} \leq \alpha_L\}, \tag{15}$$

$$i_0^+ = \max\{j \leq |D_1^+| : \frac{1}{j}\sum_{i=1}^{j} v_{[i],+}^{0-} \leq \alpha_R. \tag{16}$$

Denoting $D_{1\nu}^- \subseteq D_1^-$ as the set of indices corresponding to $\nu_{[1],-}^+, \ldots, \nu_{[i_0^-],-}^+$ and $D_{1\nu}^+ \subseteq D_1^+$ as the set of indices corresponding to $v_{[1],+}^{0-} \geq ... \geq v_{[i_0^+],+}^{0-}$, it is easy to see from (12), (15), and (16), that the constrained Bayes rule $\delta_B^{(2)}$ can be written as

$$\delta_B^{(2)} = \begin{cases} \text{Select } H_-^i \text{ if } i \in D_{1\nu}^- \\ \text{Select } H_+^i \text{ if } i \in D_{1\nu}^+ \\ \text{Select } H_0^i \text{ if } i \notin D_{1\nu}^- \cup D_{1\nu}^+. \end{cases} \tag{17}$$

3.2 Bayes Rules under a General Loss Function

The "0-1" loss gives equal loss of 1 for misclassifying a true H_- as H_0 or H_+, and a true H_+ as H_0 or H_-. It may be more appropriate to give higher loss for selecting a true H_- as H_+ than selecting it as H_0, and likewise for selecting a true H_+ as H_- than selecting it as H_0. More generally, we may allow the loss to depend on the actual value of θ.

$$\begin{aligned} L_i(\theta_i, d_i^X = -1) &= \begin{cases} 0 & \text{if } \theta_i < 0 \\ l_0 & \text{if } \theta_i = 0 \\ l_0 + l(\theta_i) & \text{if } \theta_i > 0 \end{cases} \\ L_i(\theta_i, d_i^X = 0) &= \begin{cases} l_1 + l(\theta_i) & \text{if } \theta_i < 0 \\ 0 & \text{if } \theta_i = 0 \\ l_1 + l(\theta_i) & \text{if } \theta_i > 0 \end{cases} \\ L_i(\theta_i, d_i^X = 1) &= \begin{cases} l_0 + l(\theta_i) & \text{if } \theta_i < 0 \\ l_0 & \text{if } \theta_i = 0 \\ 0 & \text{if } \theta_i > 0 \end{cases}, \end{aligned} \tag{18}$$

where l_0 and l_1 are some some positive constants, and $l(\theta) \geq 0$ is a function that is symmetric around 0 and is increasing in $|\theta|$. Note that $l_0 = l_1 = 1$ and $l(\cdot) = 0$ lead to the "0-1" loss.

The posterior loss, in this case, is given by

$$E[\sum L_i(\theta_i, d_i)|\mathbf{x}] = \sum E[(l_1 + l(\theta_i))I(\theta_i \neq 0)|\mathbf{x}] \\ - \sum(w_i^- + l_0 v_i^- + l_1 - l_0 - l_1 v_i^0)d_i^- \\ - \sum(w_i^+ + l_0 v_i^+ + l_1 - l_0 - l_1 v_i^0)d_i^+,$$

where

$$w_i^- = E[l(\theta_i)I(\theta_i < 0)|\mathbf{x}] \quad \text{and } w_i^+ = E[l(\theta_i)I(\theta_i > 0)|\mathbf{x}] \tag{19}$$

The analogous versions of $\delta_B^{(0)}$, $\delta_B^{(1)}$ and $\delta_B^{(2)}$ can now be obtained, in a similar manner as discussed in susection 3.1, by replacing D_1^- and D_1^+, in (11), by

$$\begin{aligned} D_{1l}^- &= \{i : w_i^- + l_0\nu_i^- > l_0 - l_1 + l_1\nu_i^0, \;\; w_i^- + l_0\nu_i^- > w_i^+ + l_0\nu_i^+\}, \\ D_{1l}^+ &= \{i : w_i^+ + l_0\nu_i^+ > l_0 - l_1 + l_1\nu_i^0, \;\; w_i^+ + l_0\nu_i^+ > w_i^- + l_0\nu_i^-\}, \end{aligned} \tag{20}$$

respectively.

3.3 A Bayes Rule with Exact $BDFDR = \alpha$

Note that if $l_1 = l_0$ in the loss (18), then from (20), $i \in D_{1l}^- \cup D_{2l}^+$ if $w_i^- > l_0(\nu_i^0 - \nu_i^-)$ or $w_i^+ > l_0(\nu_i^0 - \nu_i^+)$. Thus, $i \in D_1^- \cup D_2^+$ implies $i \in D_{1l}^- \cup D_{2l}^+$, which implies that the Bayes rule based on the loss (18) enlarges the rejection region (of rejecting H_0^i) in comparison to the "0-1" loss. Moreover, if $l(\theta) = l_a$ (some fixed positive value), then it can be seen from (19) and (20) that

$$D_{1l}^- = \{i : \nu_i^- > c_0\nu_i^0, \;\; \nu_i^- > \nu_i^+\}, \qquad D_{1l}^+ = \{i : \nu_i^+ > c_0\nu_i^0, \;\; \nu_i^+ > \nu_i^-\},$$

where $c_0 = l_0/(l_0 + l_a)$. This c_0 now can be used as an instrument so that the Bayes rule is of particular false discovery rate. For example, we can find c_0 $(0 < c_0 < 1)$ such that

$$BDFDR = E[\frac{\sum(\nu_i^0 + v_i^+)d_i^- + \sum(\nu_i^0 + v_i^-)d_i^+}{V + W + I(V = 0)I(W = 0)}] = \alpha, \tag{21}$$

where the expectation is with respect to the marginal distribution of $\mathbf{X}$. It may be interesting to know the property of $BDFDR$ as a function of c_0. Of particular interests would be to know whether $BDFDR$ is a monotonic function of c_0. If so, c_0 can be found through simulation as we illustrate in the next section. This yields a new Bayes rule with exact $BDFDR = \alpha$, which selects H_-^i if $i \in D_{1l}^-$, selects H_+^i if $i \in D_{1l}^+$, and selects H_0^i for all other i.

4 Computation of the Bayes rules

In this section, we discuss the computation of the Bayes rules under the prior (3). We will assume that $\mathbf{X}$ can be reduced to a sufficient statistics $(\mathbf{Y}, S)$ such that $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_m)^T$ is a vector of independent variates with $Y_i \sim f(y_i; \theta_i, \eta)$, $i = 1, \ldots, m$, the distribution of S is independent of $\boldsymbol{\theta}$, and where $\mathbf{Y}$ and S are independently distributed. A typical example, where this can be done, is $X_{ij} = \theta_i + \varepsilon_{ij},\ j = 1, 2, ..., n,\ i = 1, 2, ..., m$, where $\varepsilon_{ij} \sim N(0, \sigma^2)$. In this case, $Y_i = \bar{X}_i$, and $S = \sum\sum(X_{ij} - \bar{X}_i)^2$.

As discussed in Section 2, the Bayes rules are determined by the posterior distribution of $\boldsymbol{\theta}$ given $\mathbf{X} = \mathbf{x}$. When there is a nuisance parameter η, the posterior expected loss can be computed by first computing the expectation with respect to the posterior distribution of $\boldsymbol{\theta}$ given $\mathbf{X} = \mathbf{x}$, η and then with respect to the posterior distribution of η given $\mathbf{X} = \mathbf{x}$. Note that based on the assumption on $(\mathbf{Y}, S)$, the posterior distribution of $\boldsymbol{\theta}$ given $\mathbf{X} = \mathbf{x}$, η requires only the consideration of the distribution $Y_i \sim f(y_i; \theta_i, \eta)$, $i = 1, \ldots, m$ and the prior distribution (3). We show in the Appendix that a good approximation can be obtained by only considering the distribution $Y_i \sim f(y_i; \theta_i, \eta)$, $i = 1, \ldots, m$ with the prior (3), and then by replacing η by the mode of the posterior distribution of η given $S = s$. Thus, from now on, we will assume that η is known with the understanding that, in the case of unknown η, η can be replaced by the posterior mode of η given $S = s$.

Conditionally, given $\mathbf{X} = \mathbf{x}$, the $\theta_1, \theta_2, \cdots, \theta_m$ are independently distributed with marginal densities (suppressing η for simplicity)

$$\pi(\theta_i|y_i) = \pi(H_i^-|y_i)\pi(\theta_i|H_i^-, y_i) + \pi(H_i^0|y_i)I(\theta_i = 0) + \pi(H_i^+|y_i)\pi(\theta_i|H_i^+|y_i), \tag{22}$$

where $\pi(\theta_i|H_i^-, y_i)$ and $\pi(\theta_i|H_i^+, y_i)$ are the posterior distributions with respect to the priors $\pi_-(\theta_i)$ and $\pi_+(\theta_i)$ respectively, and

$$\pi(H_i^-|y_i) \propto p_- f(y_i|H_i^-),\ \pi(H_i^0|y_i) \propto p_0 f(y_i|H_0^i),\ \pi(H_i^+|y_i) \propto p_+ f(y_i|H_i^+),$$

where $f(y_i|H_i^0) = f(y_i|0)$, and $f(y_i|H_i^-)$, $f(y_i|H_i^+)$ are the marginal densities under the priors π_- and π_+, respectively, keeping η fixed. Note that the proportionality constant above is the inverse of $[p_- f(y_i|H_i^-) + p_0 f(y_i|H_i^0) + p_+ f(y_i|H_i^+)]$.

Clearly,

$$v_i^- = \pi(H_i^-|y_i), \quad v_i^0 = \pi(H_i^0|y_i), \quad v_i^+ = \pi(H_i^+|y_i). \tag{23}$$

The Bayes decision rule and the constrained Bayes rules $\delta_B^{(0)}, \delta_B^{(1)}$, and $\delta_B^{(2)}$, as defined in (12), (14), and (17), can now be obtained from (23). For the loss function (18), the Bayes rule requires, in addition to (23), the computation of $w_i^- = E[l(\theta)I(\theta < 0)|\mathbf{X}]$ and $w_i^+ = E[l(\theta)I(\theta > 0)|\mathbf{X}]$, which are given by

$$w_i^- = \pi(H_i^-|y_i)E[l(\theta_i)|H_i^-, y_i], \; w_i^+ = \pi(H_i^+|y_i)E[l(\theta_i)|H_i^+, y_i]. \tag{24}$$

Denoting $T_+(y_i) = f(y_i|H_i^-)/f(y_i|0)$ and $T_-(y_i) = f(y_i|H_i^+)/f(y_i|0)$, it can be seen that

$$\begin{aligned} \nu_i^- &= \frac{p_- T_-(y_i)}{p_- T_-(y_i) + p_+ T_+(y_i) + p_0}, \quad \nu_i^+ = \frac{p_+ T_+(y_i)}{p_- T_-(y_i) + p_+ T_+(y_i) + p_0} \\ \nu_i^0 &= \frac{p_0}{p_- T_-(y_i) + p_+ T_+(y_i) + p_0}. \end{aligned}$$

Thus, from (11),

$$\begin{aligned} D_1^- &= \{i : T_-(y_i) > p_0/p_-, \; p_- T_-(y_i) > p_+ T_+(y_i)\} \\ D_1^+ &= \{i : T_+(y_i) > p_0/p_+, \; p_+ T_+(y_i) > p_- T_-(y_i)\}. \end{aligned} \tag{25}$$

Note that

$$T_-(y_i) = \int_{-\infty}^{0} \frac{f(y_i|\theta_i)}{f(y_i|0)} \pi_-(\theta_i) d\theta_i \text{ and } \; T_+(y_i) = \int_0^{\infty} \frac{f(y_i|\theta_i)}{f(y_i|0)} \pi_+(\theta_i) d\theta_i. \tag{26}$$

Thus if $f(y_i|\theta_i)$ has monotone likelihood ratio (MLR) property in y_i, then $T_-(y_i)$ ($T_+(y_i)$) is a decreasing (increasing) function of y_i. In that case, from (12) and (25), it is easy to see that $\delta_B^{(0)}$ selects H_0^i if $T_-^{-1}(p_0/p_-) \le y_i \le T_+^{-1}(p_0/p_+)$, selects H_-^i if $y_i < T_-^{-1}(p_0/p_-)$, and selects H_+^i if $y_i > T_+^{-1}(p_0/p_+)$. Also note that in this case ν_i^- is a decreasing function of y_i, and ν_i^+ is an increasing in y_i. This property makes the computation of $\delta_B^{(1)}$ and $\delta_B^{(2)}$ simpler in the sense of less computational need for $T_-(y_i)$ and $T_+(y_i)$. The following Theorem summarizes the results.

Theorem 2 *Suppose the pdf of y_i, $f(y_i|\theta)$, is MLR in y_i. Let k_- be the largest integer i such that $y_{[i]} < T_-^{-1}(p_0/p_-)$ and $p_- T_-(y_{[i]}) > p_+ T_+(y_{[i]})$, and let k_+ be the smallest integer i such that $y_{[i]} > T_+^{-1}(p_0/p_+)$ and $p_+ T_+(y_{[i]}) > p_- T_-(y_{[i]})$, where T_- and*

T_+ are defined by (26). Then the Bayes rule $\delta_B^{(0)}$ selects H_-^i for all i associated with $y_{[1]}, \ldots, y_{[k_-]}$, selects H_+^i for all i associated with $y_{[k_+]}, \ldots, y_{[m]}$, and selects H_0^i for all other i.

Additionally, suppose the $\max\{\frac{1}{j_1+j_2}(\sum_{i_1=1}^{j_1} \nu_{[i_1]}^{0+} + \sum_{i_2=1}^{j_2} \nu_{[i_2]}^{0-}) : j_1 \le k_-, \ j_2 \le m - k_+ + 1\}$, subject to the constrain that it is $\le \alpha$, is attained by the indices $j_1 = i_{10}^-$ and $j_2 = m - i_{20}^+ + 1$. Then the constrained Bayes rule $\delta_B^{(1)}$ selects H_-^i for all i associated with $y_{[1]}, \ldots, y_{[i_{10}^-]}$, selects H_+^i for all i associated with $y_{[i_{20}^+]}, \ldots, y_{[m]}$, and selects H_0^i for all other i.

Also, let $i_0^- = \max\{j : \frac{1}{j}\sum_{i=1}^{j} \nu_{[i]}^{0+} \le \alpha_L, \ i \le k_-\}$ and $i_0^+ = \min\{m - j + 1 : \frac{1}{j}\sum_{i=1}^{j} \nu_{[j]}^{0-} \le \alpha_R, \ j \le m - k_+ + 1\}$. Then the constrained Bayes rule $\delta_B^{(2)}$ selects $H_-^{(i)}$ for all i associated with $y_{[1]}, \ldots, y_{[i_0^-]}$, selects H_+^i for all i associated with $y_{[i_0^+]}, \ldots, y_{[m]}$, and selects H_0^i for all other i.

Some more simplification in the computation is possible if $\pi_-(\theta_i)$ and $\pi_+(\theta_i)$ are such that $\pi_+(\theta_i) = \pi_-(-\theta_i)$, and if $f(y_i|\theta_i) = f(y_i'|-\theta_i)$ for some y_i'. Then it can be seen that $T_-(y_i) = T_+(y_i')$. For example, in the binomial case, where $y_i \sim b(n, p_i), \ i = 1, 2, ..., m$, consider the problem of testing $H_0^i : \theta_i = 0$ *vs.* $H_-^i : \theta_i < 0$ *or* $H_+^i : \theta_i > 0$, where $\theta_i = \log p_i/(1 - p_i)$. If $\pi_+(\theta_i) = \pi(-\theta_i)$, then $T_-(y_i) = T_+(n - y_i)$. A similar point can be made about the normal distributions as shown later.

Remark 3 *For the general loss as defined in Section 3, the above Theorem can be easily modified with the initial cut-off k_- as the largest integer i such that $w_i^- + l_0 v_i^- > \max(w_i^+ + l_0 v_i^+, l_0 - l_1 + l_1 v_i^0\}$, and the cut-off k_+ as the smallest integer i such that $w_i^+ + l_0 v_i^+ > \max(w_i^- + l_0 v_i^-, l_0 - l_1 + l_1 v_i^0\}$, where $w_i^- = E[l(\theta_i) I(\theta_i < 0)|\mathbf{x}]$ and $w_i^+ = E[l(\theta_i) I(\theta_i > 0)|\mathbf{x}]$.*

4.1 Determination of p_-, p_+, and p_0

The Bayes rules as described above are sensitive to the choice of $(p_-, \ p_0, \ p_+)$. If these probabilities are not available subjectively, they need to be estimated from the data itself. The EM algorithm can be used to estimate $(p_-, \ p_0, \ p_+)$ by maximizing the marginal likelihood function. Note that the marginal likelihood (fixing η) is given by

$$L(p_-, p_0, \ p_+|y) = \prod_{i=1}^{m} [p_- f(y_i|H_i^-) + p_0 f(y_i|0) + p_+ f(y_i|H_i^+)].$$

It is easy to see that the iterative solution of the EM algorithm is given by

$$
\begin{aligned}
\hat{p}_-^{(j+1)} &= \frac{1}{m}\sum_{i=1}^{m}\frac{\hat{p}_-^{(j)}T_-(y_i)}{\hat{p}_-^{(j)}T_-(y_i)+\hat{p}_+^{(j)}T_+(y_i)+\hat{p}_0^{(j)}}\\
\hat{p}_0^{(j+1)} &= \frac{1}{m}\sum_{i=1}^{m}\frac{\hat{p}_0^{(j)}}{\hat{p}_-^{(j)}T_-(y_i)+\hat{p}_+^{(j)}T_+(y_i)+\hat{p}_0^{(j)}}\\
\hat{p}_+^{(j+1)} &= \frac{1}{m}\sum_{i=1}^{m}\frac{\hat{p}_+^{(j)}T_+(y_i)}{\hat{p}_-^{(j)}T_-(y_i)+\hat{p}_+^{(j)}T_+(y_i)+\hat{p}_0^{(j)}}
\end{aligned}
$$

4.2 Normal Populations

Let $y_i \backsim N(\theta_i, \sigma^2/n)$. We consider the priors π_- and π_+ as the left and the right truncated $N(0, \sigma^2/\omega)$ priors, where ω is some positive constant. Then $\pi_+(\theta_i) = \pi_-(-\theta_i)$, and thus it is easy to see that $y_i' = -y_i$ and $T_-(y_i) = T_+(-y_i)$, where

$$
T_+(y_i) = 2\sqrt{\frac{\omega}{n+\omega}}\exp\{\frac{n^2 y_i^2}{2(n+\omega)\sigma^2}\}\Phi(\frac{n y_i}{\sigma\sqrt{n+\omega}}) \tag{27}
$$

Since the density of Y_i is MLR in y_i, Theorem 2 can be applied. Thus, from Theorem 2, the Bayes rule $\delta_B^{(0)}$ can be stated as follows:

Let

$$
\begin{aligned}
k_- &= \max\{i : y_{[i]} < -c_p\frac{\sigma\sqrt{n+\omega}}{n},\ p_- T_+(-y_{[i]}) > p_0\} \\
k_+ &= \max\{i : y_{[i]} > -c_p\frac{\sigma\sqrt{n+\omega}}{n},\ p_+ T_+(y_{[i]}) > p_0\},
\end{aligned} \tag{28}
$$

where c_p is the $p_+/(p_+ + p_-)$th percentile of $N(0,1)$. Then select H_i^- for all i associated with $y_{[1]}, \ldots, y_{[k_-]}$; select H_+^i for all i associated with $y_{[k_+]}, \ldots y_{[m]}$; and select H_0^i for all other i.

The constrained Bayes rules $\delta_B^{(1)}$ and $\delta_B^{(2)}$ can also be computed as described in Theorem 2 with

$$
\nu_{[i]}^{0-} = \frac{p_- T_+(-y_{[i]}) + p_0}{p_- T_+(-y_{[i]}) + p_+ T_+(y_{[i]}) + p_0},\ \text{for } i \geq k_+
$$

and

$$
\nu_{[i]}^{0+} = \frac{p_+ T_+(y_{[i]}) + p_0}{p_- T_+(y_{[i]}) + p_+ T_+(y_{[i]}) + p_0},\ \text{for } i \leq k_-,
$$

Note that, because of monotonicity of $T_+(\cdot)$, $\nu^{0+}_{[1]} \leq \ldots \leq \nu^{0+}_{[k_-]}$ and $\nu^{0-}_{[m]} \leq \ldots \leq \nu^{0-}_{[k_+]}$.

As we mentioned in Section 3.3, the following Bayes rule attains exact $BDFDR = \alpha$:

Select H^i_0 *if* $\nu^-_i > c_0\nu^0_i$ *and* $\nu^-_i > \nu^+_i$*; select* H^i_+ *if* $\nu^+_i > c_0\nu^0_i$ *and* $\nu^+_i > \nu^-_i$*; otherwise, select* H^i_0.

Here the constant c_0 is such that the (21) is satisfied. Using the same arguments as in the proof of Theorem 2, it can be seen that the Bayes rule selects H^i_- for all i associates with $y_{[1]}, \ldots, y_{[k_-(c_0)]}$; selects H^i_+ for all i associated with $y_{[k_+(c_0)]}, \ldots y_{[m]}$; and selects H^i_0 for all other i, where

$$
\begin{aligned}
k_-(c_0) &= \max\{i : y_{[i]} < -c_p\frac{\sigma\sqrt{n+\omega}}{n},\ p_-T_+(-y_{[i]}) > c_0p_0\} \\
k_+(c_0) &= \max\{i : y_{[i]} > -c_p\frac{\sigma\sqrt{n+\omega}}{n},\ p_+T_+(y_{[i]}) > c_0p_0\}.
\end{aligned}
$$

The constant c_0 does not depend on σ, and thus a general table can be created for a specific sample size for this cut-off point. For example, when sample size $n = 10$ for all i with $\omega = 1$, Figure 1 shows the graph of $BDFDR$ (computed through simulation) as a function of c_0 for $p_+ = 0.4$ and $p_- = 0.1$. From this, for $\alpha = 0.05$, we obtain $c_0 = 2.896$.

[Figure 1]

5 Simulation Results

We now use simulation to illustrate the performance of $\delta^{(0)}_B, \delta^{(1)}_B, \delta^{(2)}_B$ and compare them to the Benjamini-Hochberg (BH) procedure. We generated y_i from $N(\theta_i, 1/n)$, and θ_i from (3) with π_- and π_+ as left and right truncated standard normal densities, respectively, for different choices of $(p_-,\ p_0,\ p_+)$. Note that the Benjamini-Hochberg procedure is based on two-tailed test. Thus, to adapt it to the directional hypothesis (1), upon rejecting H^i_0, we select H^-_i if $y_i < 0$, and H^+_i if $y_i > 0$. The following quantities will be used to compare the results: left-tailed correct discovery rate ($LCDR$), right-tailed correct discovery rate ($RCDR$), left-tailed false non-discovery rate ($LFNDR$), and right-tailed false non-discovery rate ($RFNDR$), defined by

$$
\begin{aligned}
LCDR &= E[\frac{V_-}{V}],\ RCDR = E[\frac{W_+}{W}], \\
LFNDR &= E[\frac{U_-}{U}],\ RFNDR = E[\frac{U_+}{U}].
\end{aligned} \tag{29}
$$

Note that the expectations above are with respect to both $\mathbf{X}$ and $\boldsymbol{\theta}$. $LCDR$ and $RCDR$ reflects the power of the tests for the proportion of correct discoveries of left tail and right tail hypotheses, respectively. $LFNDR$ and $RFNDR$ reflects the false non-discoveries in left and right directions, respectively. The reason the latter rates are important is that a poor test might have a high proportion of correct discoveries but too many of true left and right tails might be declared null. A large error rates of $LFNDR$ and $RFNDR$ will clearly reflect this.

Note that all quantities in (29) depend on the k_- and k_+ of (28) and $T_+(y_i)$ as defined in (27). It is easy to see that k_-, k_+ and $T_+(y_i)$ are all invariant with respect to σ. Thus, without loss of generality, all comparisons made here can be stated for any σ.

Tables 2-4 are based on the simulation of $5,000$ repetitions with $m = 1000$ and $n = 10$. In Table 2, the BH procedure is based on FDR of the same level as the FDR of level 0.05. In Table 3, the constrained Bayes rule $\delta_B^{(1)}$ is based on $BDFDR$ $(= \alpha)$ of 0.05, and in Table 4, the constrained Bayes rule $\delta_B^{(2)}$ is based on $BLFDR$ $(= \alpha_L)$ of 0.025 and $BRFDR$ $(= \alpha_R)$ of 0.025.

[Tables 2-4]

Note that in all cases we have chosen $p_+ > p_-$. The results of BH are highly non-symmetrical depending upon the values of p_- and p_+. Generally for the BH rules, the correct discovery rates $LCDRs$ for the left tails are much lower than correct discovery rates $RCDRs$ for the right tails. However, for both $\delta_B^{(1)}$ and $\delta_B^{(2)}$, both rates are balanced and very high. For example when $(p_-,\ p_+) = (0.05,\ 0.3)$, the BH rule has a correct discovery rate of 0.7429 in the left tail as compared to the correct discovery rate of 0.9547 for $\delta_B^{(1)}$ and 0.9789 for $\delta_B^{(2)}$. In the context of microarray data analysis this would mean that if a particular cell-line generate more of over-expressed genes than under-expressed gene, then the BH rule would falsely select under-expressed genes with high proportion in comparison to $\delta_B^{(1)}$ or $\delta_B^{(2)}$. A similar point can be made about false non-discoveries.

When $p_+ > p_-$, $RFNDR$ is relatively large for BH, meaning a high percentage of right tails are declared null. In the context of microarray data analysis, this means that a higher percentage of overexpressed genes are declared null under BH rule when compared to $\delta_B^{(1)}$ or $\delta_B^{(2)}$. Overall conclusion is that if the selection of over-expressed genes are more likely, then $\delta_B^{(1)}$ or $\delta_B^{(2)}$ outperform BH rule. When comparing $\delta_B^{(1)}$ and $\delta_B^{(2)}$, $\delta_B^{(2)}$ has a better correct discovery rates than the $\delta_B^{(1)}$, but false non-discover rates are mostly higher for $\delta_B^{(2)}$.

6 Concluding Remarks

In this paper, we provide a general framework of computing Bayes decision rule for the directional multiple hypothesis problem when the left and the right hypotheses are asymmetrically generated. The decision rules we derived attain Bayesian optimality with a control directional false discovery rates. The false discovery rates we considered can be controlled in left and right directions combined, or separately. The methodologies presented here are useful in many practical situations where it is expected that one direction is more probable than the other. We show through simulation that taking this information into account yields better decision rules. We also noted that the type of the loss function makes a difference. Through simulation, we found that If the proportion of correct discoveries is more desirable, then a non - "0-1" loss function is better choice than the "0-1" loss. We also showed that if the densities have monotone likelihood ration (MLR) property, then the Bayes rule takes a very simple form. In that case, the Bayes rule can be obtained in the form of the ranked values of the sufficient statistics; and additionally, only a simple non-linear computation is required to find the cut-off points.

A Appendix

A.1 Estimation of $E[g(\theta_i, \eta)|\mathbf{x}]$ for a non-linear function g

Here, we assume that $\mathbf{X}$ can be reduced to a sufficient statistic $(\mathbf{Y}, S)$ such that $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_m)^T$ is a vector of independent variates with $Y_i \sim f(y_i; \theta_i, \eta)$, $i = 1, \ldots, m$, and where $\mathbf{Y}$ and S are independently distributed, and the distribution of S is independent of $\boldsymbol{\theta}$. When η is kept fixed, the computation of $E[g(\theta_i, \eta)|\mathbf{x}]$ is based on the posterior distribution of θ_i given y_i and η. Suppose the resulting expression is $h(\eta|y_i)$.

It can be seen that the posterior expectation of $h(\eta|y_i)$ can be written as

$$E[h(\eta|y_i)|\mathbf{x}] = \frac{f_*(s)}{f_*(y_i, s)} \int h(\eta|y_i) f_*(y_i|\eta)\pi(\eta|s)d\eta, \tag{30}$$

where $f_*(s),\ f_*(y_i, s)$ are the marginal densities of S and (Y_i, S) respectively; $f_*(y_i|\eta)$ is the marginal density of y_i given η; and $\pi(\eta|s)$ is the posterior density of η given $S = s$. Note that $\pi(\eta|s)$ does not depend on $\boldsymbol{\theta}$. Now, suppose $\pi(\eta|s) \propto \exp(-mg_m(\eta))$, where g_m is of order $O(1)$. In this case, using Laplace approximation, it can be shown (Severini [11], section 2.11)) that

$$\int h(\eta|y_i) f_*(y_i|\eta)\pi(\eta|s)d\eta \doteq h(\hat{\eta}|y_i) f_*(y_i|\hat{\eta}) \frac{\sqrt{2\pi}}{\sqrt{mg''_m(\hat{\eta})}} \exp\{-mg''_m(\hat{\eta})\}$$

with the remainder term of order $O(1/m)$, where $\hat{\eta}$ the posterior mode. Thus, from (30), we have

$$E[h(\eta|y_i)|\mathbf{x}] \doteq \frac{f_*(s)}{f_*(y_i, s)} f_*(y_i|\hat{\eta}) h(\hat{\eta}|y_i) \frac{\sqrt{2\pi}}{\sqrt{mg''_m(\hat{\eta})}} \exp\{-mg''_m(\hat{\eta})\} \tag{31}$$

A.2 Bayes rule in the presence of the nuisance parameter η

Since $f_*(s),\ f_*(y_i, s)$ and $f_*(y_i|\hat{\eta})$ does not depend on the function h and since the Bayes rules only depend on the ratios of the expressions (31) for different h functions, the Bayes rule can be expressed approximately in terms of $h(\hat{\eta}|y_i)$ alone for large m. Regarding the assumption that $\pi(\eta|s) \propto \exp(-mg_m(\eta))$, this would typically be the case if S is composed of $i.i.d.$ copies $(S_i,\ i = 1, 2, \cdots, m)$. In many practical cases, this would be the case; for example, when the observes variables are $X_{ij} = \theta_i + \varepsilon_{ij}$, where ε_{ij} are independently and identically distributed with distribution independent of $\theta_i,\ i = 1, 2, \cdots m$. In this case $S_i = \sum_j (X_{ij} - \bar{X}_i)^2$, and in this case the posterior distribution $\pi(\eta|s)$ satisfies the desired condition.

References

[1] Bansal, N.K. and Sheng, R. (2010). "Bayesian decision theoretic approach to hypotheses problems with skewed alternatives," Journal of Statistical Planning and Inference, Vol. 140, 2894-2903.

[2] Benjamini, Y., and Hochberg, Y. (1995). “Controlling the false discovery rate: a practice and powerful approach to multiple testing”, J. R. Statist. Soc. B, Vol. 57, No. 1, 289-300.

[3] Berger, J.O. (1985). Statistical Decision Theory and Bayesian Analysis, 2nd Edition. Springer, New York.

[4] Lehmann, E. L. (1957a). “A theory of some multiple decision problems I”, The Annals of Mathematical Statistics, Vol. 28, 1-25.

[5] Lehmann, E. L. (1957b). “A theory of some multiple decision problems I”, The Annals of Mathematical Statistics, Vol. 28, 547-572.

[6] Lewis, C., and Thayer, D.T. (2004). Journal of Statistical Planning and Inference, Vol.125, 49-58.

[7] Muller, P., Parmigiani, G., and Rice, K. (2006). "FDR and Bayesian Multiple Comparisons Rules". Johns Hopkins University, Dept. of Biostatistics Working Papers. Working Paper 115. http://www.bepress.com/jhubiostat/paper115.

[8] Rowntree and Lee (2006). "Mapping of DNA replications origins to noncoding genes of the X-inactivation center", Molecular and Cellular Biology, Vol. 26, 3707-3717.

[9] Shaffer, J.P. (2002). “Multiplicity, directional(Type III) errors, and the null hypothesis”, Psychological Methods, Vol. 7, No. 3, 356-369.

[10] Sarkar. S.K., and Zhou T. (2008). “Controlling Bayes directional false discovery rate in random effects model”, J. Statist. Plann. Inference, Vol. 138, No. 3, 682-693.

[11] Severini, T.A. (2000). Likelihood Methods in Statistics. Oxford, New York.

Table 2: Left- and right-tailed correct discovery rates and false non-discovery rates of BH

(p_-, p_+)	$LCDR$	$RCDR$	$LFNDR$	$RFNDR$
$(0.1, 0.8)$	0.9260	0.9938	0.0899	0.7185
$(0.1, 0.6)$	0.8403	0.9735	0.0734	0.4423
$(0.1, 0.4)$	0.8337	0.9548	0.0652	0.2611
$(0.05, 0.4)$	0.7162	0.9553	0.0318	0.2551
$(0.1, 0.3)$	0.8456	0.9441	0.0626	0.1879
$(0.05, 0.3)$	0.7429	0.9471	0.0308	0.1853
$(0.1, 0.2)$	0.8670	0.9298	0.0612	0.1223
$(0.05, 0.2)$	0.7846	0.9373	0.0302	0.1217
$(0.1, 0.1)$	0.8987	0.8990	0.0607	0.0607
$(0.05, 0.1)$	0.7831	0.9362	0.0304	0.1215
$(0.05, 0.05)$	0.8926	0.8928	0.0313	0.0313

Table 3: Left- and right-tailed correct discovery rates and false non-discovery rates of $\delta_B^{(1)}$

(p_-, p_+)	$LCDR$	$RCDR$	$LFNDR$	$RFNDR$
$(0.1, 0.8)$	0.9568	0.9496	0.1807	0.5445
$(0.1, 0.6)$	0.9560	0.9495	0.0992	0.3920
$(0.1, 0.4)$	0.9539	0.9496	0.0794	0.2513
$(0.05, 0.4)$	0.9547	0.9502	0.0414	0.2465
$(0.1, 0.3)$	0.9535	0.9500	0.0734	0.1873
$(0.05, 0.3)$	0.9547	0.9503	0.0386	0.1841
$(0.1, 0.2)$	0.9520	0.9506	0.0690	0.1263
$(0.05, 0.2)$	0.9532	0.9509	0.0360	0.1242
$(0.1, 0.1)$	0.9512	0.9518	0.0656	0.0656
$(0.05, 0.1)$	0.9542	0.9503	0.0362	0.1240
$(0.05, 0.05)$	0.9538	0.9531	0.0339	0.0339

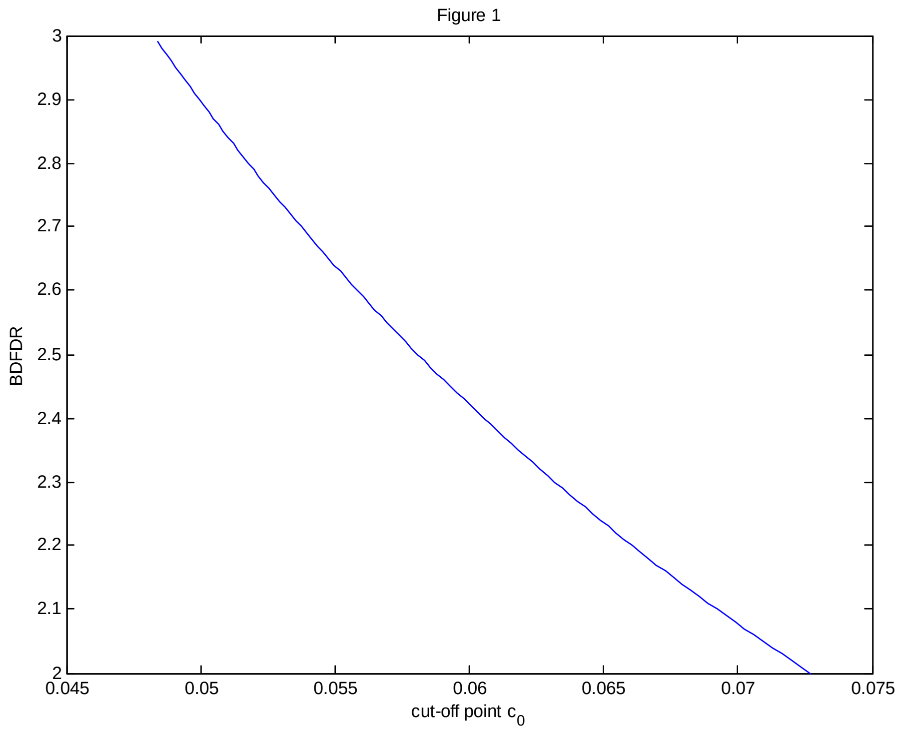

Table 4: Left- and right-tailed correct discovery rates and false non-discovery rates of $\delta_B^{(2)}$

(p_-, p_+)	$LCDR$	$RCDR$	$LFNDR$	$RFNDR$
$(0.1, 0.8)$	0.9766	0.9750	0.1459	0.6177
$(0.1, 0.6)$	0.9769	0.9752	0.0968	0.4293
$(0.1, 0.4)$	0.9768	0.9754	0.0816	0.2728
$(0.05, 0.4)$	0.9788	0.9753	0.0425	0.2679
$(0.1, 0.3)$	0.9769	0.9756	0.0766	0.2023
$(0.05, 0.3)$	0.9789	0.9756	0.0401	0.1989
$(0.1, 0.2)$	0.9771	0.9760	0.0728	0.1354
$(0.05, 0.2)$	0.9793	0.9760	0.0379	0.1334
$(0.1, 0.1)$	0.9771	0.9771	0.0697	0.0697
$(0.05, 0.1)$	0.9805	0.9755	0.0381	0.1331
$(0.05, 0.05)$	0.9805	0.9799	0.0359	0.0359