False discovery and false non-discovery rates in single-step multiple testing procedures

Sanat K. Sarkar
Temple University, Philadelphia, USA

[FIRST DRAFT]
Tuesday, August 5, 2003

Single-step multiple testing procedures where each hypothesis is rejected for a value of the corresponding test statistic exceeding a certain fixed critical value is considered. Some results on false discovery rate (FDR) and false non-discovery rate (FNR) are developed for dependent test statistics both under a model where the number of true null hypotheses is assumed fixed and a mixture model where different configurations of true and false null hypotheses are assumed to have certain probabilities. These results verify some desirable properties of FDR and FNR as measures of error rates and extend some previously known results, providing further insights into the notions of FDR and FNR and related measures under dependence. One of these results also provides an inequality that explains how a single-step procedure, like Bonferroni or Sidak procedure, can potentially be improved by modifying them using an estimate of the number of true null hypotheses. Numerical examples are presented showing that such modification can often yield much less conservative FDR-controlling procedure. A simulation study is also conducted investigating how the modified Bonferroni or Sidak procedure performs compared to its unmodified version in terms of a measure of power involving both FDR and FNR.

Keywords: Single-step multiple testing; Bonferroni procedure; Sidak procedure; unbiased FDR-controlling procedure; mixture model.

1 Introduction

The false discovery rate (FDR) and related measures have been receiving considerable attention due to their relevance as measures of overall error rate in multiple testing problems arising in many scientific investigations, particularly in the context of DNA microarray analysis. Consider the following table presenting all possible outcomes when a multiple testing procedure is applied to simultaneously test n null hypotheses $H_1, \dots, H_n$:

Table 1. The outcomes in testing n null hypotheses

	Rejected	Accepted	Total
True Null	V	U	n_0
False Null	S	T	n_1
Total	R	A	n

The research is supported by 2003 Summer Research Fellowship Awarded by Temple University.

Let $Q = V/R$ if $R > 0$, and $= 0$ if $R = 0$, i.e, the proportion of false positives (Type I errors) among the rejected null hypotheses. Genovese and Wasserman (2002) called it the realized FDR. The FDR is defined by $E(Q)$. It was first introduced in multiple testing by Benjamini and Hochberg (1995). They also gave a procedure that controls the FDR with independent test statistics. Let H_i be tested using a right-tailed test based on the statistic X_i, $i = 1, \dots, n$, and that the X_i's have a common cdf F under the null hypotheses. Then, with $X_{(1)} \leq \cdots \leq X_{(n)}$ denoting the ordered components of the X_i's, the Benjamini-Hochberg procedure controlling the FDR at α rejects all H_i for which $X_i \geq X_{(i^*)}$, where

$$i^* = \min_{1 \leq j \leq n} \{j : X_{(j)} \geq c_j\}, \tag{1.1}$$

and $F(c_i) = 1 - [(n - i + 1)\alpha/n]$. This is a step-up procedure as it performs the tests in a step-up manner starting with the least significant hypothesis. Later, Benjamini and Liu (1999) offered the step-down procedure that rejects the H_i's for which $X_i > X_{(i^{**})}$, where

$$i^{**} = \max_{1 \leq j \leq n} \{j : X_{(j)} \leq d_j\}, \tag{1.2}$$

and $F(d_i) = [1 - \min(1, \frac{n}{i}\alpha)]^{\frac{1}{i}}$, which also controls the FDR at α under independence. The FDR-controlling property of the Benjamini-Hochberg procedure is extended by Benjamini and Yekutieli (2001) to some positively dependent multivariate distributions. Sarkar (2002b) strengthened the above results. He proved that the critical values of the Benjamini-Hochberg procedure can be used in a more general stepwise procedure providing a control of the FDR not only under independence but also when the test statistics have the same type positive dependence property as considered by Benjamini and Yekutieli (2001). In addition, he established the FDR-controlling property of the Benjamini-Liu step-down procedure for some positively dependent test statistics. Genovese and Wasserman (2002) investigated some operating characteristics of the Benjamini-Hochberg procedure asymptotically under independence.

A different approach to FDR has been taken in a number of recent articles. Rather than finding rejection regions for the hypotheses providing a control of the FDR at a specified level, these papers devote to the idea of estimating FDR or pFDR (positive false discovery rate) for fixed rejection regions. The pFDR was proposed by Storey (2002), which is the conditional FDR given at least one rejection, that is, pFDR $= E(V/R \mid R > 0)$. It has an interpretation of a Bayesian Type I error rate under a mixture model involving iid p-values when a single-step multiple testing procedure is used; see also Storey (2003). This is a measure which is appropriate for estimation and not meant to be controlled the way FDR can be controlled through a testing scheme. Storey (2002) provides estimates of FDR and pFDR under the above mixture model for a single-step procedure. These estimates are related to the empirical Bayes FDR of Efron, Tibshirani, Storey, and Tusher (2001); see also Efron (2003).

An analog of FDR in terms of false negatives (Type II errors) has been developed by Genovese and Wasserman (2002) and Sarkar (2002a) to compare FDR-controlling procedures. It is the FNR, called false non-discovery rate by Genovese and Wasserman (2002) and the false negatives rate by Sarkar (2002a), and is defined by $E(N)$, where $N = T/A$ if $A > 0$, and $= 0$ if $A = 0$. The quantity N is the proportion of false negatives (Type II errors) among the accepted hypotheses and is called the realized FNR by Genovese and Wasserman (2002). Storey (2002) introduced the pFNR (positive false non-discovery rate), the conditional expectation $E(T/A \mid A > 0)$, as an analog of his pFDR.

In this article, we concentrate on single-step multiple testing procedures and develop some new results on FDR and FNR with dependent test statistics both under a model where n_0 and n_1 are assumed fixed and the so called mixture model where different configurations of true and false null hypotheses are assumed to have certain probabilities. The intent of these results is to verify some desirable properties of FDR and FNR and extend some previously known results, thereby providing further insights into the notions of FDR and FNR and related measures under dependence.

Suppose that $\mathbf{X} = (X_1, \ldots, X_n)$ has a joint distribution indexed by parameters (θ, η), where $\theta = (\theta_1, \ldots, \theta_n)$. Let $H_i : \theta_i = \theta_{i0}$ which is being tested against $K_i : \theta_i > \theta_{i0}$, for some given θ_{i0}, $i = 1, \ldots, n$. Let $\{H_i : i \in J_0\}$ and $\{H_i : i \in J_1\}$ be the sets of true and false null hypotheses respectively. It will be assumed that J_0 is non-empty. An important property that is desired of both FDR and FNR is that, if the distribution of $\mathbf{X}$ is stochastically increasing in each θ_i, which is typically the case in many multiple testing problems, then, for fixed n_0 and n_1, both FDR and FNR should decrease as θ_i increases from θ_{i0} for at least one $i \in J_1$. We will prove these monotonicity properties for a single-step procedure using a more precise definition of the stochastic increasing property. These monotonicity properties of FDR and FNR yield the following inequalities:

$$FDR \leq \frac{n_0}{n} P\{R > 0\} \quad \text{and} \quad FNR \leq \frac{n_1}{n} P\{A > 0\}, \tag{1.3}$$

where the probabilities are evaluated at $\theta = (\theta_{10}, \ldots, \theta_{n0})$ and the random variables are assumed exchangeable under these null hypothetical values. When the common critical value t of such a procedure provides a level α test for the overall null $\cap_{i=1}^{n} H_i$, we have

$$FDR \leq \frac{n_0}{n} \alpha. \tag{1.4}$$

This is interesting in that the same inequality holds for some stepwise procedure with the critical values providing a level α test for the overall null [Benjamini and Hochberg (1995); Benjamini and Yekutieli (2001); Sarkar (2002b)]. It represents a more precise version of the result stating that these single-step procedures control the FDR at α, and provides a way to potentially improve these procedures in terms of having a better control of the FDR by using an estimate of n_0 in the spirit of Storey (2002), Benjamini and Hochberg (2000) and Yekutieli and Benjamini (1999). For instance, a single-step procedure using Bonferroni critical value or Sidak critical value assuming independence or positive dependence of the test statistics control the FDR. While this is not a surprising result, as they are known to control the familywise error rate (FWER), what is more important, however, is that that they can be improved with the help of the first inequality in (1.3). We suggest improving these procedures by modifying the critical values using an estimate of n_0, and present numerical evidence showing that such modification can often yield much less conservative FDR-controlling procedures.

Since the FDR measures expected proportion of incorrect decisions, a good multiple testing procedure must ensure that it does not exceed expected proportion of correct decisions [Sarkar (2002a)]. The quantity 1 - FNR, which Genovese and Wasserman (2002) called the correct non-discovery rate, is a measure of correct decisions. In situations where controlling false negatives is of primary importance, the FNR provides a measure of incorrect decisions with the corresponding measure of correct decisions being 1 - FDR. Whether we have a multiple testing procedure designed to control FDR or FNR, the inequality FDR + FNR $\leq$ 1 represents a desirable property for any such multiple testing procedure. This is referred to as the unbiasedness condition of an FDR- or FNR-controlling multiple testing procedure by Sarkar (2002a). He established this property for FDR-controlling generalized stepwise procedures of Sarkar (2002b) under independence and assuming that the level at which the FDR is controlled is less than or equal to 0.5, and left it open for other cases. The inequalities in (1) prove this property for single-step procedures under dependence.

Next, we will derive certain results extending Storey's (2002, 2003) work to dependence case. Storey obtained expressions for the FDR and FNR of a single-step procedure under a mixture model where given any configuration of true and false null hypotheses the X_i's are assumed independent, providing useful Bayesian interpretations to his notions of pFDR and pFNR. More specifically, he proved: $pFDR = P\{H_1 \text{ is accepted} \mid X_1 \geq t\}$ and $pFNR = P\{H_1 \text{ is rejected} \mid X_1 < t\}$, irrespective of the number of tests. Assuming a more general mixture model in which the X's are assumed dependent with a location-family of distributions and having a certain type of positive dependence structure, we prove in this article that

$$pFDR \leq \max_{1 \leq i \leq n} P\{H_i \text{ is accepted} \mid X_i \geq t\} \tag{1.5}$$

and

$$pFNR \leq \max_{1 \leq i \leq n} P\{H_i \text{ is rejected} \mid X_i < t\}, \tag{1.6}$$

with the equalities holding under independence. An important implication of (1.5) is that Storey's (2002) q-value of a single-step multiple test under the independence assumption works conservatively, as we would want, even under certain commonly encountered types of dependence.

The paper is organized as follows. In Section 2, we define the kind of stochastically increasing distributions we need to obtain the inequalities in (1.3). Section 3 reports the results related to FDR assuming fixed number of true and false null hypotheses, and the numerical study investigating the amount of improvement achieved by the modified Bonferroni procedure over its unmodified version. The results related to FNR, again assuming fixed n_0 and n_1, are presented in Section 4. The unbiasedness property is discussed in Section 5, along with a numerical study comparing the strength of unbiasedness, we call power, of the FDR-controlling Bonferroni procedure with its modified version. Section 6 presents the results on FDR and FNR under the aforementioned mixture model with dependent $\mathbf{X}$. We have considered 100 tests in our numerical studies. Since the critical values of the Bonferroni and Sidak procedures are practically same for large n, the numerical results for these procedures are almost identical. We have, therefore, reported the numerical findings only for the Bonferroni procedure, taking a view that whatever conclusions we make for the Bonferroni procedure based on these numerical studies, they can be similarly made for the Sidak procedure, at least for large number of hypotheses. The paper concludes with some final remarks in Section 7. Most of the technical proofs are deferred to the Appendix.

2 Stochastically increasing family of distributions conditional on a subset.

This section defines a type of stochastic increasing property of a family of distributions that will be required to establish main results of this paper on FDR and FNR. We will use the notation $J_{(-i)} = J - \{i\}$ in the definition. Also, whenever an increasing or decreasing condition or property in terms of $\mathbf{X}$ or θ is mentioned, it is to be understood as being coordinatewise. Regarding the families of distributions, we will assume that, for any $M \subset J$, the marginal distribution of $\{X_j : j \in M\}$ depends on θ only through $\{\theta_j : j \in M\}$.

Definition 1. An n-dimensional random vector $\mathbf{X} = (X_1, ..., X_n)$ or the corresponding family of distributions $\{P_\theta\}$, where $\theta = (\theta_1, \ldots, \theta_n)$, is said to be stochastically increasing in θ conditional on a subset $\{X_i : i \in M\}$, where $M \subset J$, if $P_\theta\{\mathbf{X} \in \mathbf{C} | X_i\}$ is increasing in $\{\theta_j : j \in J_{(-i)}\}$, for each $i \in M$ and any set $\mathbf{C}$ which is increasing in $\{X_j : j \in J_{(-i)}\}$ for fixed X_i.

The above property is a less restrictive version of the usual stochastic increasing property of $\mathbf{X}$ in θ, where the probability $P_\theta\{\mathbf{X} \in \mathbf{C}\}$ is required to be increasing in θ for any set $\mathbf{C}$ which is increasing in $\mathbf{X}$. If $\mathbf{X}$ is stochastically increasing in θ, then $P_\theta\{\mathbf{X} \in \mathbf{C} | X_i, i \in N\}$ is increasing in $\{\theta_i : i \in J - N\}$ for every $N \subset J$ and increasing set $\mathbf{C}$ in $\mathbf{X}$. The property here, however, requires the increasing property to hold conditionally given only one X_i out of a subset of $\mathbf{X}$ and for a set which is increasing in $\{X_j : j \in J_{(-i)}\}$ for fixed X_i.

Example 1. Random variables with mixtures of independent MLR distributions. In multiple testing, the X_i's often have distributions that are mixtures of independent distributions belonging to monotone likelihood ratio (MLR) families. That is, the density of P_θ is often of the form $f_\theta(\mathbf{x}) = \int \prod_{i=1}^n f_{i\theta_i}(x_i, y) dG(y)$, with $f_{i\theta_i}(x, y)$ satisfying the condition that, for any $\theta_i < \theta_i'$, $f_{i\theta_i'}(x, y)/f_{i\theta_i}(x, y)$ is increasing in x, for each i, and G being a probability distribution independent of θ. The multivariate distribution of such random variables is stochastically increasing in θ, see, e.g, Lehmann (1986), and hence it is stochastically increasing in θ conditional on any subset.

Example 2. Multivariate location family of distributions. Let the density of P_θ be of the form $f_\theta(\mathbf{x}) \equiv f(\mathbf{x} - \theta)$. Distributions of this type are stochastically increasing conditional on any subset. This is because, for $\theta < \theta'$ with $\theta_i = \theta_i'$ for any fixed i, we have

$$\begin{aligned} P_{\theta'}\{\mathbf{X} \in C | X_i\} &= P_\theta\{\mathbf{X} \in C - (\theta' - \theta) | X_i\} \\ &\geq P_\theta\{\mathbf{X} \in C | X_i\}. \end{aligned} \tag{2.1}$$

Many of the distributions arising in multiple testing are of the type in Example 1 or 2. For instance, independent normals or absolute values of independent normals with θ_i's as the means or variances and their standardized versions. They are of the type in Example 1, arising in simultaneous testing of means or variances of independent normals against one or two-sided alternatives. Multivariate lnF that arises in many-to-one comparisons of variances against one-sided alternatives is another distribution of the type in Example 1. Multivariate normal and multivariate t are distributions of the type in Example 2. They arise, for instance, in Dunnett's many-to-one and Tukey's pairwise comparisons of means against one-sided alternatives in a one-way layout with a known or unknown common variance.

3 The FDR of a single-step procedure

Suppose that, for $i = 1, \dots, n$ and some fixed t, $X_i \ge t$ is the rejection region of H_i. Let $\theta^{(1)} = (\theta_1, \dots, \theta_n)$ with $\theta_i = \theta_{i0}$ for $i \in J_0$ and $\theta_i > \theta_{i0}$ for $i \in J_1$, and $\theta^{(0)} = (\theta_{10}, \dots, \theta_{n0})$. Define $X_{(1)}^{(-i)} \le \cdots \le X_{(n-1)}^{(-i)}$ as the ordered components of the subset $\{X_j : j \in J_{(-i)}\}$. We will ignore the subscript η in the probability distribution of $\mathbf{X}$ in what follows. The FDR of this single-step multiple testing procedure is then given in the following:

Lemma 1.

$$FDR_{\theta^{(1)}}(t; n_0, n_1) = \frac{1}{n} \sum_{i \in J_0} \Big[P_{\theta_{i0}}\{X_i \ge t\} + n \sum_{j=1}^{n-1} \frac{P_{\theta^{(1)}}\{X_{(j)}^{(-i)} < t, X_i \ge t\}}{(n-j)(n-j+1)} \Big]. \quad (3.1)$$

This lemma is proved in the Appendix.

Now, suppose that $\mathbf{X}$ is stochastically increasing in θ conditional on $\{X_i : i \in J_0\}$. Since the set $\{X_{(j)}^{(-i)} < t, X_i \ge t\}$ is decreasing in $\{X_j : j \in J_i\}$, the probability $P_{\theta^{(1)}}\{X_{(j)}^{(-i)} < t, X_i \ge t\}$ is decreasing in $\{\theta_j : j \in J_i\}$. Thus, we have

$$P_{\theta^{(1)}}\{X_{(j)}^{(-i)} < t, X_i \ge t\} \le P_{\theta^{(0)}}\{X_{(j)}^{(-i)} < t, X_i \ge t\}, \quad (3.2)$$

which yields

$$\sup_{\theta^{(1)}} FDR_{\theta^{(1)}}(t; n_0, n_1) = FDR_{\theta^{(0)}}(t; n_0, n_1). \quad (3.3)$$

If $\mathbf{X}$ is exchangeable when all the null hypotheses are true with the common marginal cdf F, the right-hand side in (3.1) reduces to

$$\begin{aligned} & \frac{n_0}{n} \Big[\bar{F}(t) + n \sum_{j=1}^{n-1} \frac{P_{\theta^{(0)}}\{X_{(j)}^{(-1)} < t, X_1 \ge t\}}{(n-j)(n-j+1)} \Big] \\ = & \frac{n_0}{n} FDR_{\theta^{(0)}}(t; n, 0) \\ = & \frac{n_0}{n} P_{\theta^{(0)}}\{R > 0\}, \end{aligned} \quad (3.4)$$

where $\bar{F} = 1 - F$. Thus, we have the following theorem as one of the main results of this article.

Theorem 1. If $\mathbf{X}$ is stochastically increasing in θ conditional on the subset $\{X_i : i \in J_0\}$, then $\mathrm{FDR}_{\theta^{(1)}}(t, n_0, n_1)$ is decreasing in $\{\theta_i : i \in J_1\}$. Furthermore, if $\mathbf{X}$ is exchangeable under the null hypotheses, then

$$\sup_{\theta^{(1)}} FDR_{\theta^{(1)}}(t; n_0, n_1) = \frac{n_0}{n} P_{\theta^{(0)}}\{R > 0\}. \quad (3.5)$$

As a corollary to Theorem 1, we see that

$$FDR_{\theta^{(1)}}(t; n_0, n_1) \le \frac{n_0}{n} \alpha, \quad (3.6)$$

implying that the FDR is controlled at α, if the critical value t provides a level α test for the overall null hypothesis $\cap_{i=1}^{n} H_i$, i.e., if t satisfies:

$$P_{\theta^{(0)}}\{\max_{i \in J} X_i < t\} \geq 1 - \alpha. \tag{3.7}$$

For instance, as discussed in Section 2, Dunnett's many-to-one and Tukey's pairwise comparisons of means against one-sided alternatives in a balanced one-way layout control the FDR. If one does not want to utilize the distributional form of $\mathbf{X}$ or if it is unknown, Bonferroni critical value satisfying

$$F(t) = 1 - \frac{\alpha}{n} \tag{3.8}$$

can be used. If $\mathbf{X}$ is known to be positively dependent under the null hypotheses yielding the inequality

$$P_{\theta^{(0)}}\{\max_{i \in J} X_i < t\} \geq F^n(t), \tag{3.9}$$

with the equality holding under independence, as in the case of many distributions arising in multiple testing, one could use Sidak's critical value t satisfying the equation

$$F(t) = (1 - \alpha)^{\frac{1}{n}}. \tag{3.10}$$

Of course, when n is large, there is not much difference between Bonferroni and Sidak's critical values.

We should point out that while the inequality (3.6) for the Bonferroni critical value seems trivial, as it can be proved using the following simple argument: $FDR \leq E(V) = n_0[1 - F(t)]$, it is not so immediate in general for other critical values. The inequality (3.6) is actually a single-step analog of the same inequality known to hold for stepwise procedures with Simes (1986) critical values providing an α-level test for $\cap_{i=1}^{n} H_i$ [Benjamini and Hochberg (1995); Benjamini and Yekutieli (2001); Sarkar (2002b)]. It explicitly states how a procedure satisfying (3.7) can potentially be improved by utilizing an estimate of n_0. For instance, t satisfying

$$F(t) = 1 - \min\{1, \frac{\alpha}{\hat{n}_0}\}, \tag{3.11}$$

where $\hat{n}_0$ is an estimate of n_0, is expected to perform better than than the Bonferroni critical value satisfying (3.8) at least when n is large. Similarly, when $\mathbf{X}$ is positively dependent under the null hypotheses, the modified Sidak's critical value given by

$$F(t) = (1 - \min\{1, \frac{n}{\hat{n}_0}\alpha\})^{\frac{1}{n}} \tag{3.12}$$

is expected to be better than satisfying (3.10). Regarding estimation of n_0, one can consider, as discussed in Storey (2002), estimates of the type

$$\hat{n}_0(\tau) = \frac{1}{F(\tau)} \sum_{i=1}^{n} I(X_i < \tau), \tag{3.13}$$

for some arbitrarily chosen small τ. The idea is that $E\{\sum_{i=1}^{n} I(X_i < \tau)\} = n_0 F(\tau) + \sum_{i \in J_1} F_{\theta_i}(\tau) \approx n_0 F(\tau)$, for small τ.

We conduct a numerical study to investigate if the above modifications of the Bonferroni and Sidak procedures do indeed result in improvements over their unmodified versions in terms of having better control of the FDR. We generate $n = 100$ dependent random variables $X_i \sim N(\mu_i, 1)$, $i = 1, \ldots, 100$, with the same variance 1 and a common correlation ρ, and perform 100 hypothesis tests of $\mu = 0$ against $\mu > 0$, each using the Bonferroni critical value corresponding to $\alpha = 0.05$. The value of Q (the realized FDR) is then calculated by setting n_0 of the μ_i's to zero and the remaining μ_i's to a positive value δ. The FDR is then estimated by averaging over 5000 iterations. This gives us the simulated FDR of the Bonferroni procedure. The FDR of the modified Bonferroni procedure is similarly calculated using the critical value that incorporates $\hat{n}_0$ with $\tau = 0$. Table 2 compares the FDR of the Bonferroni procedure and its modification for $n_0 = 30, 50$ and 70, $\rho = 0$ (independent) and 0.5 (dependent), and for different values of δ. We notice that the modified Bonferroni controls the FDR, quite often much less conservatively than the corresponding unmodified procedure, especially when there is a dependence among the tests.

We have carried out similar calculations for the Sidak and its modified procedures. The numerical results for these procedures are almost identical to those for the procedures related to the Bonferroni critical value, which is, of course, expected since the Bonferroni and Sidak critical values are practically same even for moderately large n. So, our conclusion regarding the performance of the modified Sidak procedure in comparison with its unmodified version is same as what we made above for the Bonferroni procedures, even though the corresponding numerical findings are not being reported.

After having found more than one procedure that can control the FDR in a particular testing situation; for example, the Bonferroni and its modification, comparing them further in terms of *power* seems to be the next important objective. While the idea of power can be conceptualized in terms of Type II errors (false negatives) in several different ways, extending it from single testing to multiple testing, one particular concept, which is the average power, i.e., $\frac{1}{n_1}E(S)$, has been used in a number of recent papers to compare FDR-controlling procedures [Benjamini and Liu (1999); Storey (2002)]. However, it is argued in Sarkar (2002a) that, since the FDR is a measure of false positives, it seems more appropriate to compare different FDR-controlling procedures in terms of a similar measure in terms of false negatives, the FNR [Genovese and Wasserman (2002); Sarkar (2002a)]. It will be interesting to see how the different FDR-controlling procedures in this paper compare in terms of measures involving FNR under the same distributional setting. This will be carried out in Section 5 after deriving some results on FNR in the next section.

4 The FNR of a single-step procedure

The following lemma gives an expression for the FNR of the single-step multiple testing procedure where H_i is rejected when $X_i \geq t$, for $i = 1, \ldots, n$.

Table 2. Simulated values of the FDR of Bonferroni and its modified procedures

		Independent ($\rho = 0$)		Dependent ($\rho = 0.5$)	
n_0	δ	Unmodified	Modified	Unmodified	Modified
70	1.0	0.0298	0.0366	0.0371	0.0427
	1.5	0.0176	0.0230	0.0287	0.0374
	2.0	0.0101	0.0135	0.0207	0.0285
	2.5	0.0049	0.0066	0.0143	0.0180
	3.0	0.0031	0.0039	0.0091	0.0109
50	1.0	0.0214	0.0287	0.0187	0.0320
	1.5	0.0110	0.0169	0.0176	0.0279
	2.0	0.0047	0.0065	0.0140	0.0184
	2.5	0.0022	0.0036	0.0092	0.0113
	3.0	0.0012	0.0021	0.0040	0.0055
30	1.0	0.0081	0.0145	0.0109	0.0199
	1.5	0.0049	0.0088	0.0107	0.0188
	2.0	0.0019	0.0037	0.0058	0.0109
	2.5	0.0011	0.0022	0.0031	0.0062
	3.0	0.0006	0.0014	0.0016	0.0037

Lemma 2.

$$\begin{aligned} & FNR_{\theta^{(1)}}(t; n_0, n_1) \\ = \; & P_{\theta^{(1)}}\{A > 0\} - \frac{1}{n} \sum_{i \in J_0} \left[P_{\theta_{i0}}\{X_i < t\} + n \sum_{j=1}^{n-1} \frac{P_{\theta^{(1)}}\{X_{(j)}^{(-i)} \geq t, X_i < t\}}{j(j+1)} \right]. \end{aligned} \tag{4.1}$$

The proof is given in the Appendix

Under the assumption that $\mathbf{X}$ is stochastically increasing in θ conditional on $\{X_i : i \in J_0\}$, we have

$$\begin{aligned} P_{\theta^{(1)}}\{A > 0\} &= 1 - P_{\theta^{(1)}}\{X_i \geq t, i \in J\} \\ &\leq 1 - P_{\theta^{(0)}}\{X_i \geq t, i \in J\} \\ &= P_{\theta^{(0)}}\{A > 0\}, \end{aligned} \tag{4.2}$$

and

$$P_{\theta^{(1)}}\{X_{(j)}^{(-i)} \geq t, X_i < t\} \geq P_{\theta^{(0)}}\{X_{(j)}^{(-i)} \geq t, X_i < t\}, \tag{4.3}$$

which yield

$$\sup_{\theta^{(1)}} FNR_{\theta^{(1)}}(t; n_0, n_1) = FNR_{\theta^{(0)}}(t; n_0, n_1). \tag{4.4}$$

Since, under the null hypotheses, $\mathbf{X}$ is exchangeable, the right-hand side in (4.1) is less than or equal to

$$\begin{aligned} & P_{\theta^{(0)}}\{A > 0\} - \frac{n_0}{n}\left[F(t) + n\sum_{j=1}^{n-1}\frac{P_{\theta^{(0)}}\{X_{(j)}^{(-1)} \geq t, X_1 < t\}}{j(j+1)}\right] \\ = \; & P_{\theta^{(0)}}\{A > 0\} - \frac{n_0}{n}P_{\theta^{(0)}}\{A > 0\} \\ = \; & \frac{n_1}{n}P_{\theta^{(0)}}\{A > 0\}. \end{aligned} \tag{4.5}$$

The first equality in (4.5) follows from (4.1) as $FNR_{\theta^{(0)}}(t; n, 0) = 0$; see also Sarkar (1998). This gives our next main result of this article.

Theorem 2. Under the conditions stated in Theorem 1,

$$\sup_{\theta^{(1)}} FNR_{\theta^{(1)}}(t; n_0, n_1) = \frac{n_1}{n}P_{\theta^{(0)}}\{A > 0\}. \tag{4.6}$$

Clearly, the FNR of a single-step procedure can be controlled at a level β under the condition stated in the above theorem by choosing t subject to the condition:

$$P_{\theta^{(0)}}\{\min_{i \in J} X_i \geq t\} \geq 1 - \beta. \tag{4.7}$$

If the dependence structure of $\mathbf{X}$ is not utilized, the equation $F(t) = \beta/n$ provides a Bonferroni-type choice for t. When $\mathbf{X}$ is known to be positively dependent so that the inequality $P_{\theta^{(0)}}\{\min_{i\in J} X_i \geq t\} \geq \bar{F}^n(t)$ is true, with the equality holding under independence, Sidak-type t can be determined from the equation $F(t) = 1-(1-\beta)^{\frac{1}{n}}$. These procedures can potentially be improved in terms of having better control of FNR by replacing β by $\min\{1, \beta/\hat{n}_1\}$, where $\hat{n}_1 = n - \hat{n}_0$ is an estimate of n_1.

5 Comparing FDR-controlling procedures

Having proved in the last two sections that there are more than one single-step procedure in a given multiple testing situation that can control FDR, we now proceed to our next goal of comparing different FDR-controlling procedures in terms of some meaningful criteria. Towards that goal, we first have the following result establishing the unbiasedness property of a single-step multiple testing procedure.

Theorem 3. Under the conditions stated in Theorem 1,

$$FDR_{\theta^{(1)}}(t; n_0, n_1) + FNR_{\theta^{(1)}}(t; n_0, n_1) \leq 1. \tag{5.1}$$

Proof. The proof follows from Theorems 1 and 2 by noting that

$$\begin{aligned} & FDR_{\theta^{(1)}}(t; n_0, n_1) + FNR_{\theta^{(1)}}(t; n_0, n_1) \\ \leq \; & \frac{n_0}{n}P_{\theta^{(0)}}\{R > 0\} + \frac{n_1}{n}P_{\theta^{(0)}}\{A > 0\}. \end{aligned} \tag{5.2}$$

According to this theorem, the single-step procedures with fixed t discussed in Section 3 that are designed to control FDR are unbiased. In other words, these procedures maintain, on the average, a higher proportion of correct decisions than proportion of incorrect decisions. How would one compare these procedures? A natural way to compare would be to see how they perform in terms of a measure that reflects the strength of unbiasedness. This leads us to the consideration of the following quantity

$$\pi_{\theta^{(1)}}(t; n_0, n_1) = 1 - FDR_{\theta^{(1)}}(t; n_0, n_1) - FNR_{\theta^{(1)}}(t; n_0, n_1). \tag{5.3}$$

It is also related to the idea of Genovese and Wasserman (2002) who suggested using $1 - \pi_{\theta^{(1)}}(t; n_0, n_1)$ as a risk function to compare multiple testing procedures. This is the concept of power that Sarkar (2002a) introduced. It assesses how good a multiple testing procedure is in maintaining a high proportion of correct decisions and/or a low proportion of incorrect decisions. It exhibits a monotonicity property with respect to the parameters of interest that is desired of any concept of power of a test. That is, $\pi_{\theta^{(1)}}(t; n_0, n_1)$ increases as $\theta^{(1)}$ moves further away from $\theta^{(0)}$, of course, under the stochastic increasing property assumed in Theorems 1 and 2 and for fixed (n_0, n_1).

We will now investigate how different FDR-controlling procedures, like the Bonferoni and Sidak procedures, that are unbiased, perform in terms of the aforementioned concept of power. Since, as mentioned before, the critical values of the Bonferroni and Sidak procedures are practically the same even for moderately large n, we will simulate the power of only the Bonferroni procedure and compare it with its modified version discussed in Section 3. Keep in mind that we have only numerically verified the FDR-controlling property of the modified Bonferroni procedure. The simulated power of this procedure will serve the dual purpose of numerically verifying its unbiasedness property as well as checking if the modification improves its performance in terms of power. We compute the FNR and then the power 1-FNR-FDR for both Bonferroni and its modified version (with $\tau = 0$) based on the normal data that have been simulated before for FDR calculations. These simulated powers are displayed in Figure 1. Numerically, we see that the modified Bonferroni procedure is unbiased. Also, it is observed that modifying the Bonferroni procedure using an estimate of n_0 can often yield more power than the unmodified procedure. This improvement seems to work well when the statistics are not much dependent. The same conclusion can be made for the modified Siadk procedure when the number of tests is large.

It is important to point out that the FNR-controlling procedures in Section 4 are also unbiased, and that the same concept of power could be used to further compare them.

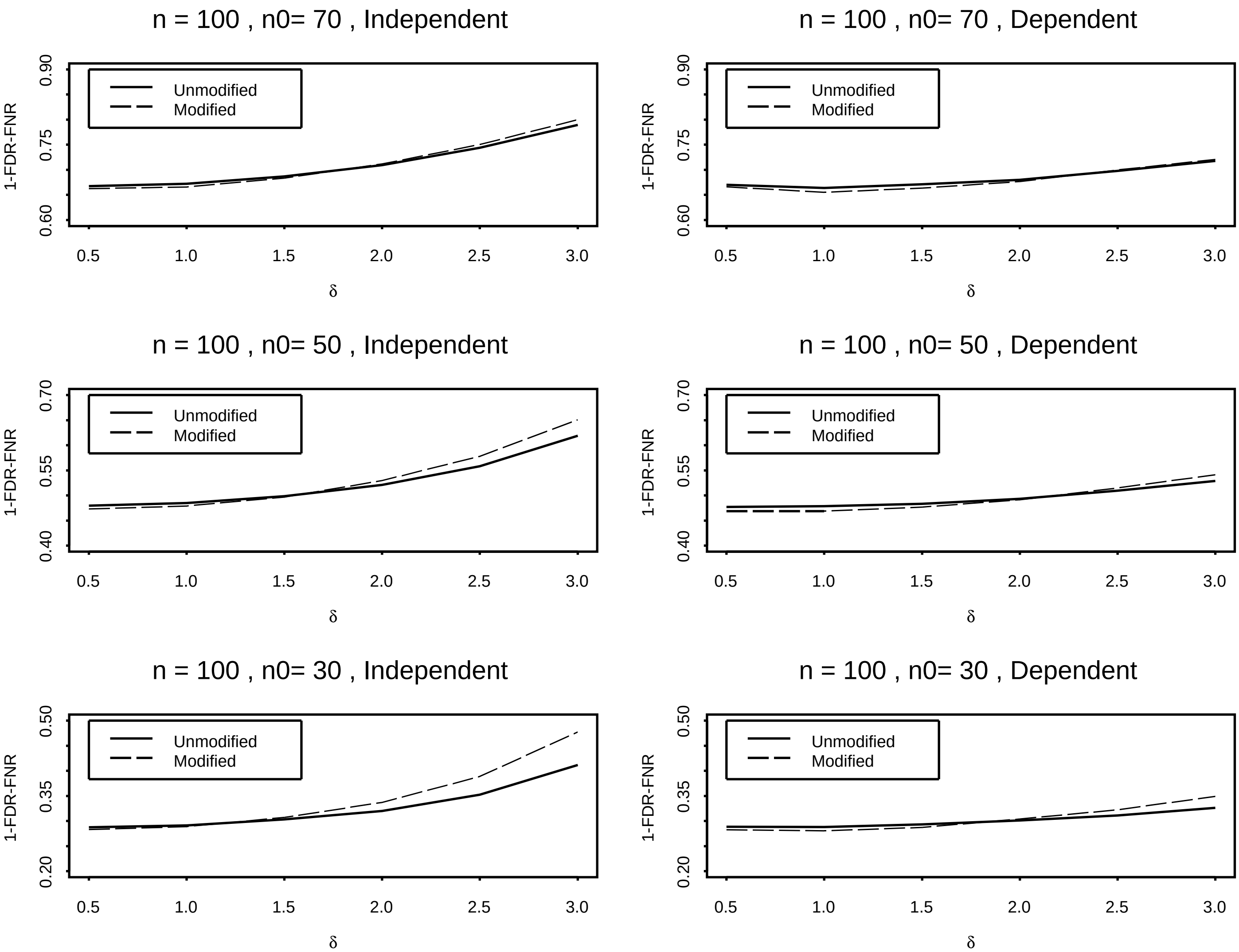

Figure 2. Comparison of Bonferroni procedure with its modified version in terms of 1-FDR-FNR

6 Mixture model

In this section, we present appropriate modifications to Lemmas 1 and 2 when mixture approach is taken as in Efron et al. (2001) and Storey (2002). We will, however, assume a more general mixture model in the sense that it does not assume independence of the test statistics. More specifically, we first let $\mathbf{H} = (H_1, \ldots, H_n)$, with $H_i = 0$ indicating that H_i is true and $H_i = 1$ indicating that it is false. Then, we assume that (X_i, H_i), $i = 1, \ldots, n$, have the following distribution: $\mathbf{X} \mid \mathbf{H} = \mathbf{h} \sim f(\mathbf{x} - \theta_{\mathbf{h}})$, where $\theta_{\mathbf{h}} = (\theta_{h_1}, \ldots, \theta_{h_n})$, $\theta_{h_i} = (1 - h_i)\theta_{i0} + h_i\theta_i$, with $\theta_i > \theta_{i0}$, for $1 = 1, \ldots, n$, and $P\{\mathbf{H} = \mathbf{h}\} = \pi_{\mathbf{h}}$, for some probabilities $\pi_{\mathbf{h}}$ defined on $\mathcal{H} = \{\mathbf{h} = (h_1, \ldots, h_n) : h_i = 0 \text{ or } 1\}$. Regarding dependence structure of $\mathbf{X}$, we assume that it is PRDS (positively regression dependent on subset) on $\{X_i : H_i = 0\}$, conditionally given any $\mathbf{H}$. That is, for any increasing (or decreasing) function ϕ of $\mathbf{X}$, the expectation $E\{\phi(\mathbf{X}) \mid X_i, \mathbf{H} \text{ with } H_i = 0\}$ is increasing (or decreasing) in X_i. This is true if, for instance, $\mathbf{X}$ is PRDS under the density $f(\mathbf{x})$, as in the case of multivariate normal with positive correlations and many other multivariate distributions encountered in multiple testing; see, for example, Benjamini and Yekutieli (2001) and Sarkar (2002b). Of course, when (X_i, H_i), $i = 1, \ldots, n$, are independent, we will assume no particular form for the density f; i.e., we simply assume that $X_i \mid H_i = h_i \sim f(x, \theta_{h_i})$. Also, (X_i, H_i), $i = 1, \ldots, n$, will have an exchangeable distribution if both $f(\mathbf{x} - \theta^{(0)})$ and $\pi_{\mathbf{h}}$ are exchangeable, and the θ_i's are assumed to have some common prior.

Theorem 4. Under the above mixture model and the conditions assumed therein,

$$FDR(t, n) \leq \frac{1}{n} \sum_{i=1}^{n} \delta_i P\{H_i = 0 \mid X_i \geq t\}, \tag{6.1}$$

where

$$\delta_i = P\{X_i \geq t\} + n \sum_{j=1}^{n-1} \frac{P\{X_{(j)}^{(-i)} < t, X_i \geq t\}}{(n-j)(n-j+1)} \text{ and } \frac{1}{n} \sum_{i=1}^{n} \delta_i = P\{R > 0\}, \tag{6.2}$$

with the equality holding when the (X_i, H_i)'s are independent.

The theorem is proved in the Appendix.

When (X_i, H_i), $i = 1, \ldots, n$, are exchangeable, Theorem 4 reduces to

$$FDR(t, n) \leq P\{H_1 = 0 \mid X_1 \geq t\} P\{R > 0\}. \tag{6.3}$$

The equality in ( ) holds when (X_i, H_i), $i = 1, \ldots, n$ are iid, which is Storey's (2002, 2003) result, providing a "Bayesian Type I error rate" interpretation to his notion of $pFDR = FDR/P\{R > 0\}$. Thus, the following corollary to Theorem 4 is an extension of his result to dependence case.

Corollary 1. Under the above mixture model and the conditions assumed therein,

$$pFDR(t, n) \leq \max_{1 \leq i \leq n} P\{H_i = 0 \mid X_i \geq t\}. \tag{6.4}$$

When the (X_i, H_i)'s are exchangeable, we have

$$pFDR(t, n) \leq P\{H_1 = 0 \mid X_1 \geq t\}, \tag{6.5}$$

with the equality holding when the (X_i, H_i)'s are iid.

Storey (2002) introduced a pFDR analogue of the p-value, called the q-value, providing a measure of the strength of the tests in a multiple testing procedure with respect to pFDR. For a single-step multiple testing procedure of n hypotheses with rejection region of the form $X_i \geq t$ for each H_i, it is defined as

$$q_n(t) = \inf_{x \leq t} pFDR(x, n). \tag{6.6}$$

Storey (2002), however, considered this quantity when (X_i, H_i), $i = 1, \ldots, n$ are iid, which is

$$q(t, H_1) = \inf_{x \le t} P\{H_1 = 0 \mid X_1 \ge x\}. \tag{6.7}$$

Corollary 1 says that when the (X_i, H_i)'s are dependent, in the sense assumed in that corollary, we have

$$q_n(t) \quad \le \quad \max_{1 \le i \le n} q(t, H_i). \tag{6.8}$$

That is, the q-value of a single-step multiple test procedure obtained under the iid assumption of (X_i, H_i) works conservatively, as we would want, even under certain commonly encountered types of dependence.

Theorem 5. Under the conditions stated in Theorem 4,

$$FNR \le \frac{1}{n} \sum_{i=1}^{n} \gamma_i P\{H_i = 1 \mid X_i < t\}, \tag{6.9}$$

where

$$\gamma_i = P\{X_i < t\} + n \sum_{j=1}^{n-1} \frac{P\{X_{(j)}^{(-i)} \ge t, X_i < t\}}{j(j+1)} \text{ and } \frac{1}{n} \sum_{i=1}^{n} \gamma_i = P\{A > 0\}, \tag{6.10}$$

with the equality holding when the (X_i, H_i)'s are independent.

This theorem can be proved following arguments similar to those used in proving Theorem 4 and with the help of an identity for $P(A > 0)$ given in Sarkar (1998).

Corollary 2. Under the conditions stated in Theorem 4,

$$pFNR \le \max_{1 \le i \le n} P\{H_i = 1 \mid X_i < t\}. \tag{6.11}$$

When the (X_i, H_i)'s are exchangeable, we have

$$pFNR \le P\{H_1 = 1 \mid X_1 < t\}, \tag{6.12}$$

with the equality holding when the (X_i, H_i)'s are iid.

7 Concluding Remarks

In this paper, we have focused primarily on single-step procedures with the idea of providing further insight into the notions of FDR and FNR and related measures under dependence. We have developed some theoretical results that extend previous work done under the assumption of independent tests and explain precisely how an FDR- or FNR-controlling single-step procedure can potentially be improved using an estimate of n_0. We have also shed some light through numerical investigations on how some commonly used single-step FDR-controlling procedures, like the Bonferroni and Sidak procedures, perform when they are modified using an estimate of n_0. It is in fact observed numerically that modification of the Bonferroni or Sidak procedure provides a better control of the FDR, at least for some choice of τ. Coming up with an analytical justification of this numerical finding would be an important undertaking, though it is going to be a theoretically challenging problem, as the Bonferroni or Sidak procedure, once modified, reduces to a more complicated two-step procedure. To see this, notice that the modified Bonferroni procedure rejects H_i, for some fixed τ, if

$$X_i \geq F^{-1}(1 - \min\{1, \frac{\alpha F(\tau)}{K_\tau}\}), \tag{7.1}$$

where $K_\tau = \sum_{i=1}^n I(X_i \leq \tau)$. Let τ be such that $F(\tau) \leq 1/2$, and define $t_k(\tau) = F^{-1}(1 - \min\{1, \frac{\alpha F(\tau)}{k}\})$. Since, $t_k(\tau) = -\infty$ for $k = 0$, and $\geq \tau$ for all $k = 1, \ldots, n$, the modified Bonferroni procedure in terms of the ordered X_i's works as follows: First, find $k = \max_{0 \leq i \leq n}\{i : X_{(i)} \leq \tau\}$ (recall that $X_0 = -\infty$); then, find $j = \max_{k \leq i \leq n}\{i : X_{(i)} \leq t_k(\tau)\}$; reject all H_i for which $X_i > X_{(j)}$. Clearly, developing the theory for the FDR of this modified procedure would be a more challenging task than what is presented here for the unmodified procedure, although we strongly believe that the ideas in this paper will play a key role in an attempt toward that development. We will address this problem in a different communication.

An important related issue we have not addressed in this article is the estimation of the parameter τ. While the value of τ we have chosen in this article seems to work in the particular normal example, a statistically more meaningful choice would be the one that provides, in a certain sense, the *best* control of the FDR within all possible values of τ.

8 Appendix

Proof of Lemma 1.

First, note that the realized FDR is given by

$$Q(t; n_0, n_1) = \sum_{i \in J_0} \sum_{j=0}^{n-1} \frac{1}{n-j} I\{R = n - j, X_i \geq t\}. \tag{8.1}$$

Since $\{R = n - j\} = \{X_{(j)} < t \leq X_{(j+1)}\}$, with $X_{(0)} = -\infty$ and $X_{(n+1)} = \infty$, we have

$$\begin{aligned} & \{R = n - j, X_i \geq t\} \\ = \; & \{X_{(j)}^{(-i)} < t \leq X_{(j+1)}^{(-i)}, X_i \geq t\}. \end{aligned}$$

Therefore,

$$
\begin{aligned}
& \sum_{i\in J_0}\sum_{j=0}^{n-1}\frac{1}{n-j}I\{R=n-j,X_i\geq t\} \\
= & \sum_{i\in J_0}\sum_{j=0}^{n-1}\frac{1}{n-j}I\{X_{(j)}<t\leq X_{(j+1)},X_i\geq t\} \\
= & \sum_{i\in J_0}\sum_{j=0}^{n-1}\frac{1}{n-j}[I\{X_{(j)}^{(-i)}<t,X_i\geq t\}-I\{X_{(j+1)}^{(-i)}<t,X_i\geq t\}] \\
= & \frac{1}{n}\sum_{i\in J_0}I\{X_i\geq t\}+\sum_{i\in J_0}\sum_{j=1}^{n-1}\frac{1}{(n-j)(n-j+1)}I\{X_{(j)}^{(-i)}<t,X_i\geq t\}. \qquad (8.2)
\end{aligned}
$$

The lemma is then proved by taking the expectation.

Proof of Lemma 2.

$$
\begin{aligned}
FNR_\theta(t;n_0,n_1) &= P_\theta\{A>0\}-E_\theta\{UI(A>0)/A\} \\
&= P_\theta\{A>0\}-\sum_{i\in I_0}\sum_{j=1}^{n}\frac{1}{j}P_\theta\{A=j,X_i<t\}. \qquad (8.3)
\end{aligned}
$$

The lemma then follows by noting that

$$
\begin{aligned}
& \sum_{j=1}^{n}\frac{1}{j}I\{A=j,X_i<t\} \\
= & \sum_{j=1}^{n}\frac{1}{j}I\{X_{(j)}<t\leq X_{(j+1)},X_i<t\} \\
= & \sum_{j=1}^{n}\frac{1}{j}[I\{X_{(j)}^{(-i)}\geq t,X_i<t\}-I\{X_{(j-1)}^{(-i)}\geq t,X_i<t\}] \\
= & \frac{1}{n}I\{X_i<t\}+\sum_{j=1}^{n-1}\frac{1}{j(j+1)}I\{X_{(j)}^{(-i)}\geq t,X_i<t\}. \qquad (8.4)
\end{aligned}
$$

Proof of Theorem 4. Since $V=\sum_{i=1}^{n}I(X_i\geq t)I(H_i=0)$, we first note from Lemma 1 that the FDR under the mixture model is given by

$$\begin{aligned} FDR(t,n) &= \frac{1}{n}\sum_{i=1}^{n} E_{\mathbf{H}}\Big[P\{X_i \geq t \mid H_i = 0\} + \qquad (8.5)\\ &\qquad n\sum_{j=1}^{n-1} \frac{P\{X_{(j)}^{(-i)} < t, X_i \geq t \mid \mathbf{H} \text{ with } H_i = 0\}}{(n-j)(n-j+1)}\Big] \\ &= \frac{1}{n}\sum_{i=1}^{n}\Big[P\{X_i \geq t, H_i = 0\} + n\sum_{j=1}^{n-1}\frac{P\{X_{(j)}^{(-i)} < t, X_i \geq t, H_i = 0\}}{(n-j)(n-j+1)}\Big] \\ &= \frac{1}{n}\sum_{i=1}^{n} P\{X_i \geq t, H_i = 0\}\Big[1 + n\sum_{j=1}^{n-1}\frac{P\{X_{(j)}^{(-i)} < t \mid X_i \geq t, H_i = 0\}}{(n-j)(n-j+1)}\Big]. \qquad (8.6) \end{aligned}$$

We will now prove that

$$P\{X_{(j)}^{(-i)} < t \mid X_i \geq t, H_i = 0\} \leq P\{X_{(j)}^{(-i)} < t \mid X_i \geq t\}, \qquad (8.7)$$

under the assumed positive dependence condition of the density f of $\mathbf{X}$.

Let $\psi(X_i) = P\{X_{(j)}^{(-i)} < t \mid X_i, \theta_i = 0\}$. Then, the conditional probability $P\{X_{(j)}^{(-i)} < t \mid X_i \geq t, \theta_i\}$ can be written as

$$\frac{E\{\psi(X_i)I(X_i \geq t - \theta_i)\}}{E\{I(X_i \geq t - \theta_i)\}}, \qquad (8.8)$$

with the expectations taken with respect to X_i under $\theta_i = 0$. Note that $\psi(x)$ is a decreasing function of x under the assumed positive dependence condition of f. Also, $I(x \geq t)$ is a TP_2 (totally positive of order two) function of (x,t) [see, e.g., Karlin (1968)]. Therefore, the ratio

$$\frac{E\{\psi(X_i)I(X_i \geq t)\}}{E\{I(X_i \geq t)\}} \qquad (8.9)$$

is deceasing in t, as it is the expectation of a decreasing function of a random variable whose distribution is stochastically increasing in t. This proves that $P\{X_{(j)}^{(-i)} < t \mid X_i \geq t, H_i = 0\} \leq P\{X_{(j)}^{(-i)} < t \mid X_i \geq t, H_i = 1\}$, implying that the probability $P\{X_{(j)}^{(-i)} < t \mid X_i \geq t\}$, being a convex combination of $P\{X_{(j)}^{(-i)} < t \mid X_i \geq t, H_i = 0\}$ and $P\{X_{(j)}^{(-i)} < t \mid X_i \geq t, H_i = 1\}$, is $\geq P\{X_{(j)}^{(-i)} < t \mid X_i \geq t, H_i = 0\}$. Thus the required inequality follows.

Applying (7.7) to (7.6), we get the inequality (6.1) to be proved in the theorem. The fact that

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n}\delta_i &= P\{\max_{1\leq i\leq n} X_i \geq t\} \\ &= P\{R > 0\} \qquad (8.10) \end{aligned}$$

follows from Sarkar (1998). Furthermore, it is clear that the equality in (6.1) holds under independence of (X_i, H_i). Thus, the theorem is proved

Acknowledgement. The author thanks Viswanath Iyer for doing the calculations.

References

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

Benjamini, Y. and Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics 25*, 60–83.

Benjamini, Y. and Liu, W. (1999). A step-down multiple hypothesis testing procedure that controls the false discovery rate under independence. *Journal of Statistical Planning and Inference 82*, 163–170.

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

Efron, B. (2003). Robbins, empirical Bayes and microarrays. *Annals of Statistics 31*, 366–378.

Efron, B., Tibshirani, R., Storey, J. D., and Tusher, V. (2001). Empirical Bayes analysis of a microarray experiment. *Journal of the American Statistical Association 96*, 1151–1160.

Genovese, C. and Wasserman, L. (2002). Operarting characteristics and extensions of the false discovery rate procedure. *Journal of Royal Statistical Society, Ser B 64*, 499–517.

Karlin, S. (1968). *Total Positivity.* CA: Stanford University Press.

Lehmann, E. (1986). *Testing Statistical Hypotheses* (2nd Ed. ed.). New York: Wiley.

Sarkar, S. K. (1998). Some probability inequalities for ordered MTP_2 random variables: A proof of the Simes conjecture. *Annals of Statistics 26*, 494–504.

Sarkar, S. K. (2002a). FDR-controlling stepwise procedures and their false negatives rates (Invited talk: The 3rd International Conference on Multiple Comparisons, Bethesda, MD, August, 2002.). Technical Report, Temple University.

Sarkar, S. K. (2002b). Some results on false discovery rate in stepwise multiple testing procedures. *Annals of Statistics 30*, 239–257.

Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika 73*, 751–754.

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B 64*, 479–498.

Storey, J. D. (2003). The positive false discovery rate: A Bayesian interpretation and the q-value. *to appear in Annals of Statistics*.

Yekutieli, D. and Benjamini, Y. (1999). Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *Journal of Statistical Planning and Inference 82*, 171–196.