Available online at www.sciencedirect.com
ScienceDirect
Statistics and Probability Letters 78 (2008) 2024–2030
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

On optimality of the Benjamini–Hochberg procedure for the false discovery rate

Wenge Guo*, M. Bhaskara Rao

Department of Environmental Health, University of Cincinnati, Cincinnati, OH 45267-0056, United States

Received 16 February 2007; accepted 9 January 2008
Available online 1 February 2008

Abstract

The Benjamini–Hochberg step-up procedure controls the false discovery rate (FDR) provided the test statistics have a certain positive regression dependency. We show that this procedure controls the FDR under a weaker property and is optimal in the sense that its critical constants are uniformly greater than those of any step-up procedure with the FDR controlling property.

MSC: 62J15; 62G30

1. Introduction

In this article, we consider the problem of simultaneously testing a finite number of null hypotheses H_i $(i = 1, \ldots, m)$. A main concern in multiple testing is the multiplicity problem. A traditional approach for solving this problem is to control the familywise error rate (FWER), which is the probability of one or more false rejections, at a desired level. However, when the number m of null hypotheses is large, very few false null hypotheses are rejected when one uses a multiple testing procedure that controls FWER. Consequently, alternative measures of error rates have been considered in the literature. Control of these measures purportedly leads to rejection of more false null hypotheses. One well-known measure is the false discovery rate (FDR), which is the expected proportion of type I errors among the rejected hypotheses, proposed by Benjamini and Hochberg (1995).

In this paper, discussion is focused on multiple testing procedures controlling FDR. Benjamini and Hochberg (1995) proposed a simple linear step-up procedure with critical constants $\alpha_i = \frac{i}{m}\alpha,\ 1 \leq i \leq m$, and showed that $FDR = \frac{m_0}{m}\alpha$, where m_0 is the number of true null hypotheses, under the assumption of independence of the underlying test statistics. Finner and Roters (2001) and Storey et al. (2004) offered different proofs of the above equality result under independence. Subsequently, Benjamini and Liu (1999) constructed a step-down procedure with the FDR controlling property for independent test statistics. Benjamini and Yekutieli (2001) extended the FDR controlling property of the Benjamini–Hochberg procedure to the case in which the test statistics have positive regression dependency on each of the test statistics corresponding to the true null hypotheses (the PRDS property). Sarkar (2002)

* Corresponding address: Biostatistics Branch, NIEHS, PO Box 12233, MD A3-03111 TW Alexander Drive, 27709-2233RTP, NC, United States.

E-mail address: wenge.guo@gmail.com (W. Guo).

0167-7152/$ - see front matter
doi:10.1016/j.spl.2008.01.069

strengthened the result of Benjamini and Yekutieli by showing that a more general step-down–step-up procedure with the same critical values as those of the Benjamini–Hochberg procedure controls the FDR under the PRDS property. In addition, he also showed that the Benjamini–Liu step-down procedure has the FDR controlling property under certain positive dependence requirements. Genovese and Wasserman (2002, 2004) investigated some operating characteristics of the Benjamini–Hochberg procedure asymptotically under independence by using a stochastic process method. Storey (2002) and Storey et al. (2004) derived a new family of FDR procedures based on estimates of FDR. Sarkar (2006) provided an FDR controlling single-step procedure under a certain dependence property. Along with these theoretical developments, the FDR has also been extensively used in many applications such as microarray data analysis (Reiner et al., 2003), clinical trials (Mehrotra and Heyse, 2004), model selection (Abramovich et al., 2005), and educational evaluation (Williams et al., 1999).

In this article, we mainly investigate optimality of the Benjamini–Hochberg procedure in terms of its critical constants. We first introduce a property of the underlying test statistics and prove that it is strictly weaker than the PRDS property. Under this new property, we provide an upper bound of the FDR for the step-up procedure with any nondecreasing critical constants. We also show that, under the reign of the new property, the critical constants of any FDR controlling procedure are uniformly smaller than those of the Benjamini–Hochberg procedure. That is, the Benjamini–Hochberg procedure is optimal in terms of the critical constants, which offers new insights into the Benjamini–Hochberg procedure. We need to point out that most of the techniques used in the literature in deriving an upper bound for the FDR rely on probability inequalities, but the upper bound of the FDR that we have obtained uses optimization methods stemming from the knapsack problem in analysis of algorithms.

The paper is organized as follows. In Section 2, we describe our basic setting and terminology. In Section 3, we present a property on the underlying test statistics strictly weaker than the PRDS property. Under this property, we then investigate the upper bound of the FDR for the step-up procedure with any nondecreasing critical constants and prove that the Benjamini–Hochberg procedure is optimal in terms of the critical constants. Further extensions and possible problems are discussed in Section 4.

2. Basic setting

Consider the problem of testing simultaneously m null hypotheses $H_1, H_2, \ldots, H_m$, of which m_0 are true and $m_1 = m - m_0$ are false. Let I_0 and I_1 be the index sets of true and false null hypotheses respectively, and $I = I_0 \bigcup I_1$, the index set of all null hypotheses.

Suppose R is the total number of hypotheses rejected and V the number of true null hypotheses rejected. The proportion of false discoveries is defined to be $Q = \frac{V}{R}$ (and equal to 0 if $R = 0$) and the false discovery rate (FDR) is defined to be the expectation of Q, i.e.,

$$FDR = E(Q) = E\left(\frac{V}{R}\right). \tag{1}$$

When testing $H_1, \ldots, H_m$, the corresponding p-values $P_1, \ldots, P_m$ are available to us. Multiple testing procedures are usually built on the p-values. Theoretically, the p-value associated with a true null hypothesis has a distribution stochastically dominating the uniform distribution $U(0, 1)$. Here, we assume that each of the p-values corresponding to true null hypotheses has a uniform distribution $U(0, 1)$. Let the ordered p-values be denoted by $P_{(1)} \leq \cdots \leq P_{(m)}$, and the associated hypotheses by $H_{(1)}, \ldots, H_{(m)}$. Suppose $\alpha_1 \leq \cdots \leq \alpha_m$ is a nondecreasing sequence of critical constants.

The step-up procedure based on the critical constants proceeds as follows. If $P_{(m)} \leq \alpha_m$, then reject all null hypotheses; otherwise, reject hypotheses $H_{(1)}, \ldots, H_{(r)}$ where r is the smallest index satisfying $P_{(m)} > \alpha_m, \ldots, P_{(r+1)} > \alpha_{r+1}$. If, for all r, $P_{(r)} > \alpha_r$, then reject none of the hypotheses. A step-up procedure begins with the least significant hypothesis and continues accepting hypotheses as long as their corresponding p-values are greater than the corresponding critical constants. In particular, the Benjamini–Hochberg procedure is a step-up procedure with critical constants $\alpha_i = \frac{i}{m}\alpha,\ i \in I$.

The following expression for the FDR is fundamental in deriving upper bounds for FDR.

Lemma 2.1 (*Benjamini and Yekutieli 2001, Sarkar 2002*). *The FDR of the step-up procedure based on any nondecreasing critical constants $\alpha_i, i \in I$, is given by*

$$FDR = \sum_{i \in I_0} \sum_{k=1}^{m} \frac{1}{k} Pr(P_i \leq \alpha_k, R = k). \tag{2}$$

For convenience, let $\alpha_0 = 0$, and define $p_{ijk} = Pr(P_i \in (\alpha_{j-1}, \alpha_j], R = k)$. Then, from Lemma 2.1, the FDR can be expressed as follows:

$$FDR = \sum_{i \in I_0} \sum_{k=1}^{m} \sum_{j=1}^{k} \frac{1}{k} p_{ijk} = \sum_{i \in I_0} \sum_{j=1}^{m} \sum_{k=j}^{m} \frac{1}{k} p_{ijk}. \tag{3}$$

3. Step-up procedure

Benjamini and Yekutieli (2001) defined a dependency property of test statistics, which they called positive regression dependency on each one from a subset I_0, or PRDS on I_0. Recall that a set D is called increasing if $x \in D$ and $y \geq x$ imply that $y \in D$ as well. Denote by X the vector of test statistics, i.e., $X = (X_1, X_2, \ldots, X_m)$. The PRDS property is defined as follows.

Property PRDS. For any increasing set D, and for each $i \in I_0$, $Pr\{X \in D \mid X_i = x\}$ is nondecreasing in x.

The PRDS property captures the positive dependency structure of the test statistics, and is a relaxed form of the positive regression dependency property. Benjamini and Yekutieli (2001) showed that the Benjamini–Hochberg step-up procedure controls the FDR under the PRDS property. For the step-up procedure with any nondecreasing critical values, the following inequality (4) holds under the PRDS property.

Lemma 3.1. *Under the PRDS property, the following inequality holds for the step-up procedure with any nondecreasing critical values α_i, $i \in I$:*

$$\sum_{k=1}^{m} Pr(R = k | P_i \leq \alpha_k) \leq 1, \quad \textit{for } i \in I_0. \tag{4}$$

Specifically, if the underlying test statistics are independent, the above inequality becomes an equality.

Proof. Note that for $i \in I_0$ and $k = 1, \ldots, m-1$,

$$\begin{aligned} Pr(R = k \mid P_i \leq \alpha_k) &= Pr(R \geq k \mid P_i \leq \alpha_k) - Pr(R \geq k+1 \mid P_i \leq \alpha_k) \\ &\leq Pr(R \geq k \mid P_i \leq \alpha_k) - Pr(R \geq k+1 \mid P_i \leq \alpha_{k+1}). \end{aligned}$$

The inequality follows from the PRDS property and the fact that the set $\{R \geq k+1\}$ is decreasing in the p-values. Then,

$$\begin{aligned} \sum_{k=1}^{m} Pr(R = k \mid P_i \leq \alpha_k) &\leq \sum_{k=1}^{m-1} \{Pr(R \geq k \mid P_i \leq \alpha_k) - Pr(R \geq k+1 \mid P_i \leq \alpha_{k+1})\} \\ &\quad + Pr(R = m \mid P_i \leq \alpha_m) = Pr(R \geq 1 \mid P_i \leq \alpha_1) = 1. \end{aligned}$$

If the underlying test statistics are independent, for each $i \in I_0$, let $R_{m-1}^{(-i)}$ denote the number of rejections of null hypotheses when the step-up procedure based on the subset of p-values $\{P_1, \ldots, P_n\} \setminus \{P_i\}$ and the critical values $\alpha_2 \leq \cdots \leq \alpha_n$ are used. Then

$$\sum_{k=1}^{m} Pr(R = k \mid P_i \leq \alpha_k) = \sum_{k=1}^{m} Pr(R_{m-1}^{(-i)} = k-1) = 1. \quad \blacksquare$$

Remark 3.1. If the underlying test statistics are negative regression dependent on each one from I_0, by an argument similar to Lemma 3.1, the direction of inequality (4) is reversed.

We now construct a joint distribution of the p-values, under which the inequality (4) is satisfied, but the PRDS property does not hold. This implies that the inequality (4) is a property of the underlying test statistics strictly weaker than the PRDS property. For convenience of notation, define $k_1 \wedge m_0 = \min\{k_1, m_0\}$.

Theorem 3.1. *Consider the step-up procedure with nondecreasing critical values $\alpha_k, k \in I$. Let $k_1 = \arg\max_{1\le k\le m} \frac{\alpha_k}{k}$ (there could be several k_1's satisfying $k_1 = \arg\max_{1\le k\le m} \frac{\alpha_k}{k}$, and we choose the maximum of these k_1's here). If $\frac{m_0}{k_1\wedge m_0}\alpha_{k_1} \le 1$, then there exists a joint distribution for the p-values that satisfies the inequality (4), but does not satisfy the PRDS property.*

Proof. The main idea is to construct a joint distribution under which the event $P_i \le \alpha_k$ is same as the event $R = k_1$ for $i \in I_0$ and $k \le k_1$. That is, $Pr\{R = k | P_i \le \alpha_k\} = 1$ if $k = k_1$; otherwise it is equal to 0. So,

$$\sum_{k=1}^{m} Pr(R = k | P_i \le \alpha_k) = 1. \tag{5}$$

The construction of the joint distribution proceeds as follows. Let $U_1, \ldots, U_{m+1}$ be $m + 1$ uniformly distributed random variables such that $U_i \sim U(\alpha_{i-1}, \alpha_i)$, $i = 1, \ldots, m$, and $U_{m+1} \sim U(\alpha_m, 1)$. Let N be a random variable taking values $1, \ldots, k_1$, and $m + 1$ with respective probabilities $\pi_1, \ldots, \pi_{k_1}$, and $1 - \sum_{i=1}^{k_1} \pi_i$, and n be its realized value. The π_i's are predetermined as follows:

$$\pi_i = \begin{cases} \dfrac{m_0(\alpha_i - \alpha_{i-1})}{k_1} & \text{if } k_1 \le m_0, \\ \alpha_i - \alpha_{i-1} & \text{elsewhere.} \end{cases} \tag{6}$$

Given $n = 1, \ldots, k_1$, if $k_1 \le m_0$, then randomly pick k_1 indices from I_0 without replacement; if $k_1 > m_0$, also randomly pick k_1 indices, which consist of m_0 indices from I_0 and $k_1 - m_0$ indices from I_1. Let the k_1 p-values associated with these k_1 indices all be equal to U_n and the p-values associated with the remaining $m - k_1$ indices all be equal to U_{m+1}. Given $n = m + 1$, let the p-values associated with the indices from I_0 and I_1 be equal to U_{m+1} and 0, respectively. It is easy to verify that, for each $i \in I_0$, the corresponding p-value $P_i \sim U(0, 1)$ and the event $P_i \le \alpha_k$ is the same as the event $R = k_1$ for $k \le k_1$. Thus, the inequality (4) holds. Note that $\{R \le k_1\}$ is an increasing set; $Pr\{R \le k_1 | P_i \le \alpha_k\} = 1$ if $k \le k_1$ and it equals 0 if $k \ge k_1$. So the p-values do not satisfy the PRDS property. ■

Remark 3.2. Yekutieli (2002) constructed an interesting FDR controlling procedure for pairwise comparisons in which the inequality (4) holds, but the p-values do not satisfy the PRDS property.

We now prove our main result assuming that the inequality (4) holds for the p-values.

Theorem 3.2. *Consider the step-up procedure with nondecreasing critical values α_k, $k \in I$, for which the inequality (4) holds. Let $k_1 = \arg\max_{1\le k\le m} \frac{\alpha_k}{k}$; then FDR $\le \frac{m_0}{k_1}\alpha_{k_1}$. Furthermore, FDR $\le \frac{m_0}{m}\alpha$ if and only if $\alpha_k \le \frac{k}{m}\alpha$ for every $k \in I$, provided $\frac{m_0}{k_1\wedge m_0}\alpha_{k_1} \le 1$.*

Proof. Note that, for any $i \in I_0$ and $j \in I$,

$$\begin{aligned} \sum_{k=j}^{m} p_{ijk} &= Pr\{P_i \in (\alpha_{j-1}, \alpha_j], j \le R \le m\} \\ &\le Pr\{P_i \in (\alpha_{j-1}, \alpha_j]\} = \alpha_j - \alpha_{j-1}. \end{aligned} \tag{7}$$

The inequality (4) can be rewritten as

$$\begin{aligned} \sum_{k=1}^{m} Pr(R = k | P_i \le \alpha_k) &= \sum_{k=1}^{m} \frac{1}{\alpha_k} Pr(P_i \le \alpha_k, R = k) \\ &= \sum_{k=1}^{m} \frac{1}{\alpha_k} \sum_{j=1}^{k} p_{ijk} = \sum_{j=1}^{m} \sum_{k=j}^{m} \frac{1}{\alpha_k} p_{ijk} \le 1. \end{aligned} \tag{8}$$

Combining (3), (7) and (8), we consider the following optimization problem:

$$\text{maximize } FDR = \sum_{i\in I_0}\sum_{j=1}^{m}\sum_{k=j}^{m}\frac{1}{k}p_{ijk},$$

with respect to p_{ijk}'s ≥ 0 and subject to the constraints

$$\sum_{k=j}^{m} p_{ijk} \leq \alpha_j - \alpha_{j-1}, \quad \text{for } i \in I_0,\ j \in I, \text{ and}$$

$$\sum_{j=1}^{m}\sum_{k=j}^{m}\frac{1}{\alpha_k}p_{ijk} \leq 1, \quad \text{for } i \in I_0. \tag{9}$$

Problem (9) can be decomposed into a family of sub-problems indexed by $i \in I_0$ as follows:

$$\text{maximize } Q_i = \sum_{j=1}^{m}\sum_{k=j}^{m}\frac{1}{k}p_{ijk},$$

with respect to p_{ijk}'s ≥ 0 and subject to the constraints

$$\sum_{j=1}^{m}\sum_{k=j}^{m}\frac{1}{\alpha_k}p_{ijk} \leq 1, \quad \text{and} \tag{10}$$

$$p_{ijk} \leq \alpha_j - \alpha_{j-1}, \quad \text{for } 1 \leq j \leq k \leq m.$$

This falls in the realm of the classical fractional knapsack problem (Martello and Toth, 1990). Let $k_1 = \arg\max_{1\leq k\leq m}\{\frac{\alpha_k}{k}\}$, and define p^*_{ijk} as below.

$$p^*_{ijk} = \begin{cases} \alpha_j - \alpha_{j-1}, & \text{if } k = k_1 \text{ and } 1 \leq j \leq k_1, \\ 0 & \text{elsewhere.} \end{cases} \tag{11}$$

Note that

$$\sum_{j=1}^{m}\sum_{k=j}^{m}\frac{1}{\alpha_k}p^*_{ijk} = \sum_{j=1}^{k_1}\frac{\alpha_j - \alpha_{j-1}}{\alpha_{k_1}} = 1. \tag{12}$$

Thus, p^*_{ijk}, $1 \leq j \leq k \leq m$, is the optimal solution of the problem (10) for each $i \in I_0$. It is easy to verify that

$$\sum_{k=j}^{m} p^*_{ijk} \leq \alpha_j - \alpha_{j-1}, \quad \text{for } i \in I_0 \text{ and } j \in I. \tag{13}$$

Consequently, p^*_{ijk}, $i \in I_0$ and $1 \leq j \leq k \leq m$, is the optimal solution of the problem (9). Thus,

$$FDR \leq \sum_{i\in I_0}\sum_{j=1}^{m}\sum_{k=j}^{m}\frac{1}{k}p^*_{ijk} = \sum_{i\in I_0}\sum_{j=1}^{k_1}\frac{1}{k}(\alpha_j - \alpha_{j-1}) = \frac{m_0}{k_1}\alpha_{k_1}. \tag{14}$$

Note that in Theorem 3.1, the constructed joint distribution of the p-values satisfies $p_{ijk} = p^*_{ijk}$ for $i \in I_0$ and $1 \leq j \leq k \leq m$, and then $FDR = \frac{m_0}{k_1}\alpha_{k_1}$. So, to make $FDR \leq \frac{m_0}{m}\alpha$, we must have $\frac{m_0}{k_1}\alpha_{k_1} \leq \frac{m_0}{m}\alpha$; that is, for each $k \in I$, $\alpha_k \leq \frac{k}{m}\alpha$. Conversely, if $\alpha_k \leq \frac{k}{m}\alpha$, for each $k \in I$, then $FDR \leq \frac{m_0}{k_1}\alpha_{k_1} \leq \frac{m_0}{m}\alpha$. ■

Theorem 3.2 implies that, in the setting of p-values where the inequality (4) is satisfied, the Benjamini–Hochberg procedure is optimal. That is, for any step-up procedure with nondecreasing critical values α_k, $k \in I$, if it can control the FDR at α, then $\alpha_k \leq \frac{k}{m}\alpha$ for each $k \in I$.

Remark 3.3. The first result of Theorem 3.2 also holds for the step-down procedure with any nondecreasing critical values α_k, $k \in I$.

4. Discussion

For the Benjamini–Hochberg procedure, if we do not impose any assumption on the joint distribution of the underlying test statistics, the value of $\sum_{k=1}^{m} Pr\{R = k|P_i \leq \alpha_k\}$ will become much larger. Using the same approach as was used in Theorem 3.2, we consider the following optimization problem:

$$\text{maximize} \sum_{k=1}^{m} Pr\{R = k|P_i \leq \alpha_k\} = \sum_{j=1}^{m} \sum_{k=j}^{m} \frac{1}{\alpha_k} p_{ijk},$$

$$\text{with respect to } p_{ijk}\text{'s } \geq 0 \text{ and subject to the constraints} \tag{15}$$

$$\sum_{k=j}^{m} p_{ijk} \leq \alpha_j - \alpha_{j-1}, \quad \text{for } i \in I_0,\ j \in I.$$

It is easily solved that the maximum value of $\sum_{k=1}^{m} Pr\{R = k|P_i \leq \alpha_k\}$ equals $\sum_{k=1}^{m} 1/k$. So, for arbitrary joint distribution of the underlying test statistics, the following inequality:

$$\sum_{k=1}^{m} Pr\{R = k|P_i \leq \alpha_k\} \leq c \tag{16}$$

holds for some given constant c satisfying $1 \leq c \leq \sum_{k=1}^{m} 1/k$.

It is easy to see that, for a different value of c, we can obtain a different upper bound of the FDR for the Benjamini–Hochberg procedure. We now consider an optimization problem similar to (9), in which the only difference is that the constraint (4) is replaced by (16). Following the technique used in Theorem 3.2, we obtain that the maximum value of the FDR is $cm_0\alpha/m$. So, the smaller the value of c is, the more powerful the FDR step-up procedure constructed is.

We note that, in the setting of the underlying test statistics satisfying (4), an undesirable result arises for step-up procedures. For example, consider quasi-single-step procedures with critical constants $\alpha_i = 0, i = 1, \ldots, k-1$, and $\alpha_i = \beta, i = k, \ldots, m$, for some $2 \leq k \leq m$. By Theorem 3.2, the step-up procedure with above critical constants cannot control the FDR at α when $\frac{k}{n}\alpha < \beta < \frac{k+1}{n}\alpha$ under the inequality (4), although the critical constants of the Benjamini–Hochberg procedure are almost uniformly greater than those of the step-up procedure. A question arises: Can we find a joint distribution of the p-values satisfying the PRDS property, under which the FDR of the above procedure is also greater than α? Specifically, under stronger assumptions like independence of the p-values, does the undesirable result still hold? Sarkar (2006) had extensively investigated the single-step control of the FDR. In future, we will discuss the control of FDR for the quasi-single-step procedures and answer the above questions using techniques of Sarkar (2006).

References

Abramovich, F., Benjamini, Y., Donoho, D., Johnstone, I.M., 2006. Adapting to unknown sparsity by controlling the false discovery rate. The Annals of Statistics 34, 584–653.

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

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

Finner, H., Roters, M., 2001. On the false discovery rate and expected type I errors. Biometrical Journal 43, 985–1005.

Genovese, C., Wasserman, L., 2002. Operating characteristics and extensions of the false discovery rate procedure. Journal of the Royal Statistical Society, Series B 64, 499–517.

Genovese, C., Wasserman, L., 2004. A stochastic process approach to false discovery control. The Annals of Statistics 32, 1035–1061.

Martello, S., Toth, P., 1990. Knapsack Problems: Algorithms and Computer Implementations. Wiley, New York.

Mehrotra, D.V., Heyse, J.F., 2004. Use of the false discovery rate for evaluating clinical safety data. Statistical Methods in Medical Research 13, 227–238.

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

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

Sarkar, S.K., 2006. False discovery and false nondiscovery rates in single-step multiple testing procedures. The Annals of Statistics 34, 394–415.

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., Taylor, J.E., Siegmund, D., 2004. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. Journal of the Royal Statistical Society, Series B 66, 187–205.

Williams, V., Jones, L., Tukey, J.W., 1999. Controlling error in multiple comparisons with examples from state-to-state differences in educational achievement. Journal of Educational and Behavioral Statistics 24, 42–69.

Yekutieli, D., 2002. A false discovery rate procedure for pairwise comparisons. Research Paper, Department of Statistics and Operation Research, Tel Aviv University.