

Testing multiple hypotheses

The classical method of adjusting for testing multiple hypotheses is the so-called Bonferroni correction, given in beginning statistics courses. Recall that it works as follows. Suppose we are testing m hypotheses H_{0j} , $j = 1, \dots, m$. The overall null hypothesis H_0 is that all m hypotheses H_{0j} are true. Suppose we want to test H_0 at level $\alpha > 0$. For each j , we have a test of H_{0j} which gives a p -value q_j . The overall test procedure will be, to reject H_0 at level α if and only if we reject H_{0j} for at least one value of j at level α/m , in other words if $q_j \leq \alpha/m$. Let A_j be the event that we reject H_{0j} under this procedure and A the event that we reject H_0 at level α . Then $A = \bigcup_{j=1}^m A_j$. Under H_0 , all H_{0j} are true, so for each j , $P(A_j) \leq \alpha/m$ and so

$$P(A) = P\left(\bigcup_{j=1}^m A_j\right) \leq \sum_{j=1}^m P(A_j) \leq m \left(\frac{\alpha}{m}\right) = \alpha$$

by Boole's inequality (Bonferroni actually gave further inequalities, of which the next one after Boole's is a lower bound for the probability of a union). So, the probability of rejecting H_0 when it is true is at most α , as desired. If H_0 is rejected, then each H_{0j} with $q_j \leq \alpha/m$ would be rejected. No assumption on the joint distribution of the random variables 1_{A_j} such as independence is needed for the method, the A_j can be arbitrarily dependent.

In the later 20th century, there were applications, notably in genomics, with large numbers, e.g. thousands, of tests. Given two tissue samples, e.g. from different species, or different organs, one might have measured the expression levels of m different genes, and hypothesis H_{0j} might be that the expression levels are the same in the two tissue types for the j th gene measured. In such situations, the Bonferroni correction could be inappropriate for at least two reasons. One is that for tests with discrete test statistics, such as a Kolmogorov–Smirnov two-sample test, there will be for given sample sizes a smallest possible p -value, and if this is larger than α/m , rejection is not possible. Another reason is that the hypothesis H_0 that all H_{0j} are true is no longer interesting as one believes that only some of them are true.

Benjamini and Hochberg (1995) gave a procedure that can deal with such situations. Suppose one has a test procedure, giving a criterion for when H_{0j} will be rejected for each j . Let R be the (random) number of H_{0j} that will be rejected. Let V be the number of j such that H_{0j}

is true but rejected and $Q = (V/R)1_{R>0}$. Then EQ is called the *false discovery rate* (FDR). One would like this not to be too large. Suppose we would like it to be $\leq q$ for some q with $0 < q < 1$. A typical value of q used, aside from $q = 0.05$, might be $q = 0.1$. Beside keeping the FDR fairly small, one also wants to reject H_{0j} in those cases where it is not true, e.g. discover interesting genes for differences between tissue samples.

The Benjamini–Hochberg (1995) (BH) procedure is as follows. Do the m tests, giving p -values q_j , $j = 1, \dots, m$. Take the order statistics of the q_j , giving $q_{(1)} \leq q_{(2)} \leq \dots \leq q_{(m)}$. Let R be the largest i such that $q_{(i)} \leq iq/m$, or $R = 0$ if there is no such i . Then reject H_{0j} if and only if $q_j \leq Rq/m$. So, R will be the number of hypotheses rejected.

It is true under some hypotheses (but not so generally as for the Bonferroni procedure) that for the BH procedure, the false discovery rate is indeed $\leq q$. To prove this, Benjamini and Hochberg (1995) assumed that the p -values for the true null hypotheses are mutually independent, and also independent of the set of p -values of the untrue hypotheses (which may depend on one another). Benjamini and Hochberg also assumed implicitly in their proofs, although not explicitly in statements, that the test statistics had continuous distributions and so, that the p -values of the true null hypotheses are i.i.d. $U[0, 1]$. Their theorem (extended to possibly discrete test statistics) is given as Corollary 1 below.

The p -values are viewed as test statistics, so that hypotheses are rejected for small values of these, by the BH rule.

The assumption of continuous test statistics is not convenient, as we already saw from examples such as the Kolmogorov–Smirnov (or Mann–Whitney–Wilcoxon) statistics for comparing two samples. Benjamini and Yekutieli (2001) gave an improvement where the independence hypothesis was weakened and the test statistics need not be continuous, so the p -values of the true null hypotheses need not be $U[0, 1]$. Since their result subsumes the main theorem of Benjamini and Hochberg, an exposition of the proof in the 2001 paper will be given here.

Remarkably, there is no assumption on the marginal distributions of the p -values of the false null hypotheses such as, for continuous test statistics, being stochastically smaller than $U[0, 1]$.

Here, the two theorems of Benjamini and Yekutieli (2001) will be stated and proved. The first one implies that of Benjamini and Hochberg. The second one has no dependence hypotheses and a different conclusion.

First we need some terminology. For two points x, y of \mathbb{R}^m , $x \leq y$ will mean $x_j \leq y_j$ for all $j = 1, \dots, m$. Benjamini and Yekutieli call a set $D \subset \mathbb{R}^m$ an *increasing set* if whenever $x \in D$ and $x \leq y$ then also $y \in D$. (Equivalently, $\{-x : x \in D\}$ is what has been called a *lower layer* in the probability literature, so an increasing set might be called an “upper layer.”)

Definition. For a Borel probability measure P on \mathbb{R}^m and a subset $I_0 \subset \{1, 2, \dots, m\}$, P will be called *positively regression dependent with respect to I_0* , or $\text{PRDS}(I_0)$, if for each increasing set D and each $i \in I_0$, conditional probabilities can be chosen such that $P(x \in D | x_i = u)$ is a nondecreasing function of u .

Here is the main theorem of Benjamini and Yekutieli (2001), namely their Theorem 1.2:

Theorem 1. *Suppose m null hypotheses are being tested, with corresponding p -values p_1, \dots, p_m , suppose m_0 of them are true, let I_0 be the set of j such that the j th null hypothesis is true, and suppose the joint distribution of (p_1, \dots, p_m) is $\text{PRDS}(I_0)$. Then if the BH procedure is used for some $0 < q < 1$, we have $E(Q) \leq m_0 q / m \leq q$.*

Example. This artificial example shows an extreme dependence allowed by PRDS. Suppose m null hypotheses are all true and are equivalent. All are tested with the same test having a continuous test statistic, so that the p -values $p_1 = \dots = p_m$, each with a $U[0, 1]$ marginal distribution. The joint distribution of $\{p_j\}_{j=1}^m$ satisfies the PRDS condition with $I_0 = \{1, 2, \dots, m\}$. Thus for given q with $0 < q < 1$ by Theorem 1, $E(Q) \leq q$. In fact one sees directly that $E(Q) = q$. In this case if one true null hypothesis is rejected, all m will be. Storey (2002) proposed considering the “positive false discovery rate” $\text{pFDR} = E(V/R | R > 0)$ as contrasted with the $\text{FDR} = EQ$. In this example $E(Q) = q < 1$ which equals the pFDR , showing a wide difference between the two. If on the other hand $P(R > 0)$ is close to 1, as is often true in applications where a number of null hypotheses are false, then there is little difference between pFDR and FDR .

Suppose B is the gain from discovering a false null hypothesis and C the cost of rejecting a true one. Let $G := (R - V)B - VC$ be the net gain. One would like to maximize EG . For the BH procedure with a given q where $0 < q < 1$, so that $E(Q) \leq q$, we would have in the Example $EG = -qmC$ which is as bad as possible.

Suppose on the other hand that all null hypotheses are true but now their p -values are jointly independent. In the BH procedure we have the bound, by Stirling's formula, if $0 < q \leq 0.1$,

$$\begin{aligned} EV = ER &\leq \sum_{k=1}^m \binom{m}{k} \left(\frac{kq}{m}\right)^k \leq \sum_{k=1}^m k \frac{(kq)^k}{k!} \\ &\leq \sum_{k=1}^m k(eq)^k / \sqrt{2\pi} < \frac{eq}{\sqrt{2\pi}} \cdot \frac{1}{(1-eq)^2} < 2.05q, \end{aligned}$$

so $EG > -2.05qC$, a lower bound not depending on m .

It seems then that to apply the PRDS hypothesis, one should not only check it but guard against excessive correlation of the p -values of the true null hypotheses.

Before the proof of either theorem, here is the other theorem of Benjamini and Yekutieli (2001), their Theorem 1.3. In their paper, the assumption of continuity, i.e. $U[0, 1]$ distribution, of the p -values of the true null hypotheses is not stated in the theorem, but it is used in the proof.

Theorem 2. *For any joint distribution of p -values (p_1, \dots, p_m) for m tests such that the marginal distributions of p_i for the true null hypotheses are continuous and so $U[0, 1]$, applying the BH procedure for $0 < q < 1$, we have*

$$E(Q) \leq q \sum_{j=1}^m \frac{1}{j}.$$

Thus, to get $E(Q) \leq q$, one could apply the procedure with q replaced by $q / \sum_{j=1}^n 1/j$.

Proof of Theorem 1. Let

$$(1) \quad q_j := jq/m, \quad j = 1, \dots, m.$$

Suppose that m_0 of the null hypotheses are true and without loss of generality, that they are the j th for $j = 1, \dots, m_0$. Thus $m_1 := m - m_0$ is the number of false null hypotheses. Let $A_{v,s}$ be the event that the BH procedure rejects exactly v true and s false null hypotheses. Then the false discovery rate is

$$(2) \quad \text{FDR} := E(Q) = \sum_{s=0}^{m_1} \sum_{v=1}^{m_0} \frac{v}{v+s} \Pr(A_{v,s}).$$

The proof uses the following, in which p_i is the p -value of the i th test:

Lemma 1. For any $v = 1, \dots, m_0$ and $s = 0, \dots, m_1$, we have

$$(3) \quad \Pr(A_{v,s}) = \frac{1}{v} \sum_{i=1}^{m_0} \Pr(\{p_i \leq q_{v+s}\} \cap A_{v,s}).$$

Proof of lemma. For a given $v = 0, 1, \dots, m_0$ and $s = 0, 1, \dots, m_1$ let $V \subset \{1, \dots, m_0\}$ be a set with cardinality $|V| = v$. Let $A_{v,s}^V$ be the subset of $A_{v,s}$ on which the v true null hypotheses rejected are exactly those in V . Then for $i = 1, \dots, m_0$, $\Pr(\{p_i \leq q_{v+s}\} \cap A_{v,s}^V) = \Pr(A_{v,s}^V)$ if $i \in V$ and is 0 otherwise. It follows that for each v and s

$$(4) \quad \begin{aligned} \sum_{i=1}^{m_0} \Pr(\{p_i \leq q_{v+s}\} \cap A_{v,s}) &= \sum_{i=1}^{m_0} \sum_{|V|=v} \Pr(\{p_i \leq q_{v+s}\} \cap A_{v,s}^V) \\ &= \sum_{|V|=v} \sum_{i=1}^{m_0} \Pr(\{p_i \leq q_{v+s}\} \cap A_{v,s}^V) \\ &= \sum_{|V|=v} \sum_{i=1}^{m_0} \mathbf{1}_{i \in V} \Pr(A_{v,s}^V) \\ &= \sum_{|V|=v} v \Pr(A_{v,s}^V) = v \Pr(A_{v,s}), \end{aligned}$$

which proves the Lemma. \square

Now returning to the proof of Theorem 1, from the lemma just proved and (2), we get

$$(5) \quad \begin{aligned} E(Q) &= \sum_{s=0}^{m_1} \sum_{v=1}^{m_0} \frac{v}{v+s} \left\{ \sum_{i=1}^{m_0} \frac{1}{v} \Pr(\{p_i \leq q_{v+s}\} \cap A_{v,s}) \right\} \\ &= \sum_{i=1}^{m_0} \left\{ \sum_{s=0}^{m_1} \sum_{v=1}^{m_0} \frac{1}{v+s} \Pr(\{p_i \leq q_{v+s}\} \cap A_{v,s}) \right\}. \end{aligned}$$

For each m -tuple $p = \{p_i\}_{i=1}^m$ (of possible p -values) and each $i = 1, \dots, m_0$ let $p^{(i)} := \{p_j\}_{j \neq i}$. Let $P^{(i)}$ be the set of all possible $(m-1)$ -tuples $\{\pi_j\}_{j \neq i}$ of possible p -values for $j \neq i$. Let $\text{ran}(p_i)$ be the range (set of possible values) of p_i , which is $[0, 1]$ for a continuous test statistic or some subset, typically a finite subset, for a discrete test statistic. For each possible p_i and each $\pi^{(i)} := \{\pi_j\}_{j \neq i} \in P^{(i)}$ let $(p_i, \pi^{(i)}) := \{r_j\}_{j=1}^m$ where $r_i = p_i$ and $r_j = \pi_j$ for $j \neq i$. For each possible $v \geq 1$ and s , let $B_{v,s}$ be the projection of $\{p_i \leq q_{v+s}\} \cap A_{v,s}$ into $P^{(i)}$, in other words, the set of all $\pi^{(i)} \in P^{(i)}$ such that for some possible value of $p_i \leq q_{v+s}$ we have $(p_i, \pi^{(i)}) \in A_{v,s}$. This set is empty if no possible value of p_i is

$\leq q_{v+s}$. If there is a possible value p'_i of p_i such that $p'_i \leq q_{v+s}$, then $B_{v,s}$ is the set of all $\pi^{(i)} := \{\pi_j\}_{j \neq i} \in P^{(i)}$ such that:

- (i) $p_j \leq q_{v+s}$ for exactly $v - 1$ values of $j \neq i$ with $1 \leq j \leq m_0$ (true null hypotheses);
- (ii) $p_j \leq q_{v+s}$ for exactly s values of j with $m_0 < j \leq m$ (false null hypotheses); and
- (iii) For each $k > v + s$ and the order statistics $\pi_{(j)}^{(i)}$ for $j = 1, \dots, m - 1$ of the coordinates of $\pi^{(i)}$, $\pi_{(k-1)}^{(i)} > q_k$.

Here (iii) holds because, for $p_i \leq q_{v+s}$ and $p \in A_{v+s}$, exactly $v + s$ null hypotheses, including the i th, are rejected.

Let $P^{(0)}$ be the set of all possible m -tuples of p -values $\{\pi_j\}_{j=1}^m$, which is the unit m -cube $[0, 1]^m$ if the test statistics are continuous and independent, but otherwise may be a proper subset. For $i, v = 1, \dots, m_0$ and $s = 0, 1, \dots, m_1$ define $C_{v,s}^{(i)}$ as the Cartesian product $\text{ran}(p_i) \times B_{v,s}$ or in other words as $\left\{ \{\pi_j\}_{j=1}^m \in P^{(0)} : \{\pi_j\}_{j \neq i} \in B_{v,s} \right\}$. [This is selected from possible interpretations of Benjamini and Yekutieli's definition(s) as working in what follows.] Then we have for each $v \geq 1$ and s

$$(6) \quad \{p_i \leq q_{v+s}\} \cap A_{v,s} = \{p_i \leq q_{v+s}\} \cap C_{v,s}^{(i)}.$$

For $k = 1, \dots, m$ let

$$C_k^{(i)} := \bigcup \{C_{v,s}^{(i)} : 1 \leq v \leq m_0, 0 \leq s \leq m_1, v + s = k\}.$$

Let $p_{(r)}^{(i)}$, $r = 1, \dots, m - 1$, be the order statistics of p_j for $j \neq i$. Then using (iii) and (6),

$$(7) \quad C_k^{(i)} = \text{ran } p_i \times \{p^{(i)} : p_{(k-1)}^{(i)} \leq q_k \text{ and for all } K > k, p_{(K-1)}^{(i)} > q_K\}.$$

It follows that for each i , the sets $C_k^{(i)}$ are disjoint, and so by (5) and (6)

$$(8) \quad E(Q) = \sum_{i=1}^{m_0} \sum_{k=1}^m \frac{1}{k} \Pr \left(\{p_i \leq q_k\} \cap C_k^{(i)} \right).$$

So we have an expression for $E(Q)$ in which v and s no longer appear.

Next some sets will be defined which will be shown to be increasing sets (upper layers), so that the hypothesis can be applied to them. Let $D_k^{(i)} := \bigcup \{C_j^{(i)} : j \leq k\}$ for $k = 1, \dots, m$. Then

$$(9) \quad D_k^{(i)} = \left\{ p : q_k < p_{(K-1)}^{(i)} \text{ for } K = k + 1, \dots, m \right\}$$

for $k = 1, \dots, m-1$, and $D_m^{(i)}$ is the whole set $P^{(0)}$ of possible m -tuples of p -values. It follows from (9) that for each k and i , $D_k^{(i)}$ is an increasing set (upper layer).

The assumed PRDS property implies that for any such set D , conditional probabilities can be chosen so that for each $i \leq m_0$ and $0 \leq r \leq r' \leq 1$ we have

$$(10) \quad \Pr(D|p_i = r) \leq \Pr(D|p_i = r').$$

As Benjamini and Yekutieli say, it's easy to see that for any $j \leq l$, since $q_j \leq q_l$, we have

$$(11) \quad \Pr(D|p_i \leq q_j) \leq \Pr(D|p_i \leq q_l)$$

for any increasing set D ; Lehmann (1966) is cited also. So for $k = 1, \dots, m-1$,

$$(12) \quad \frac{\Pr(\{p_i \leq q_k\} \cap D_k^{(i)})}{\Pr(p_i \leq q_k)} \leq \frac{\Pr(\{p_i \leq q_{k+1}\} \cap D_k^{(i)})}{\Pr(p_i \leq q_{k+1})}$$

where either fraction is defined as 0 if its denominator is 0. By this and the fact that $D_{j+1}^{(i)} = D_j^{(i)} \cup C_{j+1}^{(i)}$, a disjoint union by (7), we have for all $k \leq m-1$

$$(13) \quad \begin{aligned} & \frac{\Pr(\{p_i \leq q_k\} \cap D_k^{(i)})}{\Pr(p_i \leq q_k)} + \frac{\Pr(\{p_i \leq q_{k+1}\} \cap C_{k+1}^{(i)})}{\Pr(p_i \leq q_{k+1})} \\ & \leq \frac{\Pr(\{p_i \leq q_{k+1}\} \cap D_k^{(i)})}{\Pr(p_i \leq q_{k+1})} + \frac{\Pr(\{p_i \leq q_{k+1}\} \cap C_{k+1}^{(i)})}{\Pr(p_i \leq q_{k+1})} \\ & = \frac{\Pr(\{p_i \leq q_{k+1}\} \cap D_{k+1}^{(i)})}{\Pr(p_i \leq q_{k+1})}. \end{aligned}$$

Since $C_1^{(i)} = D_1^{(i)}$, using (13) iteratively gives

$$(14) \quad \sum_{k=1}^m \frac{\Pr(\{p_i \leq q_k\} \cap C_k^{(i)})}{\Pr(p_i \leq q_k)} \leq \frac{\Pr(\{p_i \leq q_m\} \cap D_m^{(i)})}{\Pr(p_i \leq q_m)} = 1,$$

where the last equality holds because $D_m^{(i)}$ is the whole space $P^{(0)}$.

For any $i = 1, \dots, m_0$ and k ,

$$(15) \quad \Pr(p_i \leq q_k) \leq q_k = qk/m$$

because p_i is the p -value of a true null hypothesis. We then have by (8)

$$\begin{aligned}
(16) \quad E(Q) &= \sum_{i=0}^{m_0} \sum_{k=1}^m \frac{1}{k} \Pr \left(\{p_i \leq q_k\} \cap C_k^{(i)} \right) \\
&\leq \sum_{i=0}^{m_0} \sum_{k=1}^m \frac{q}{m} \cdot \frac{\Pr \left(\{p_i \leq q_k\} \cap C_k^{(i)} \right)}{\Pr(p_i \leq q_k)},
\end{aligned}$$

and by (14),

$$(17) \quad \frac{q}{m} \sum_{i=0}^{m_0} \sum_{k=1}^m \frac{\Pr \left(\{p_i \leq q_k\} \cap C_k^{(i)} \right)}{\Pr(p_i \leq q_k)} \leq \frac{m_0}{m} q,$$

which proves Theorem 1. \square

The main theorem of Benjamini and Hochberg (1995), extended to cover possibly discrete test statistics, now follows:

Corollary 1. *Suppose that for m tests, where the i th null hypothesis is true if and only if $i \leq m_0 \leq m$, the p -values p_i for $i = 1, \dots, m_0$ are jointly independent, and independent of the vector of p -values $\{p_j\}_{j>m_0}$ for the false null hypotheses (which can be arbitrarily dependent on one another). Then the FDR satisfies*

$$\begin{aligned}
(18) \quad FDR = E(Q) &= \sum_{i=1}^{m_0} \sum_{k=1}^m \frac{1}{k} \Pr \left(\left\{ p_i \leq \frac{k}{m} q \right\} \cap C_k^{(i)} \right) \\
&= \sum_{i=1}^{m_0} \sum_{k=1}^m \frac{1}{k} \Pr \left(p_i \leq \frac{k}{m} q \right) \cdot \Pr \left(C_k^{(i)} \right) \\
&\leq \sum_{i=1}^{m_0} \frac{q}{m} \sum_{k=1}^m \Pr \left(C_k^{(i)} \right) = \frac{m_0}{m} q,
\end{aligned}$$

where the inequality in the last line becomes an equality if the test statistics are continuous for the true null hypotheses, and so $U[0, 1]$.

Proof. We apply (8), then the assumed independence, and then (15), which becomes an equation for continuous test statistics. Then, as mentioned before (8), for each i , the sets $C_k^{(i)}$ are disjoint for different k . As stated around (9), their union $D_m^{(i)}$ has probability 1. The Corollary then follows.

For the inequality $E(Q) \leq m_0 q/m$, one could simply note that the assumed independence clearly implies the PRDS property for $I_0 = \{1, \dots, m_0\}$. \square

To apply Corollary 1 for equality when the true null hypotheses have continuous test statistics, one would need in practice that all the null hypotheses do.

Next, Theorem 2 will be proved.

Proof of Theorem 2. Keeping the notation of the proof of Theorem 1, for $i = 1, \dots, m_0$ and $j, k = 1, \dots, m$ let

$$p_{ijk} := \Pr \left(\left\{ p_i \in \left(\frac{j-1}{m}q, \frac{j}{m}q \right] \right\} \cap C_k^{(i)} \right).$$

Then for each $i \leq m_0$ and j ,

$$(19) \quad \sum_{k=1}^m p_{ijk} := \Pr \left(\left\{ p_i \in \left(\frac{j-1}{m}q, \frac{j}{m}q \right] \right\} \cap \left(\bigcup_{k=1}^m C_k^{(i)} \right) \right) = \frac{q}{m}$$

because the $C_k^{(i)}$ are disjoint for different k and their union is the whole space $P^{(0)}$, while since p_i is for a true null hypothesis with continuous test statistic, it has a $U[0, 1]$ distribution. From (8) we have

$$(20) \quad \begin{aligned} E(Q) &= \sum_{i=1}^{m_0} \sum_{k=1}^m \frac{1}{k} \sum_{j=1}^k p_{ijk} = \sum_{i=1}^{m_0} \sum_{j=1}^m \sum_{k=j}^m \frac{1}{k} p_{ijk} \\ &\leq \sum_{i=1}^{m_0} \sum_{j=1}^m \sum_{k=j}^m \frac{1}{j} p_{ijk} \leq \sum_{i=1}^{m_0} \sum_{j=1}^m \frac{1}{j} \sum_{k=1}^m p_{ijk} = m_0 \sum_{j=1}^m \frac{1}{j} \frac{q}{m}, \end{aligned}$$

using (19) in the last step, which proves Theorem 2. \square

Remark. Although Benjamini and Yekutieli (2001) in their Theorem 1.3 do not make the assumption of continuous test statistics for the true null hypotheses, as was made in Theorem 2, equality (19) cannot necessarily be replaced by \leq in case of discrete test statistics, as (15) has intervals $\{p_i \leq q\}$, and the corresponding inequality about intervals $\{q_{j-1} < p_j \leq q_j\}$ is not necessarily true for discrete test statistics.

Remarks on citations. The Benjamini–Hochberg paper has been extremely influential: it has been cited over 9000 times by October 2010 and over 15,000 times by October 2012 according to Google Scholar. A large number of the citations are applications in biology.

The Benjamini–Yekutieli (2001) paper is, according to Google Scholar, Benjamini’s second-most-cited work, with “only” about 1200 citations through October 2010, 1950 through October 2012. According

to the *Web of Science* as of about March 10, 2008, of some 2789 total citations of Benjamini and Hochberg (1995) (BH) in their data base, 497 were in “genetics and heredity,” 414 in “statistics and probability,” with other mentioned fields being subfields of biology.

Although I had difficulty finding any papers by biologists citing the Benjamini–Hochberg or Benjamini–Yekutieli work which checked hypotheses sufficiently, it seems perhaps more fruitful to look at later work such as Reiner, Yekutieli and Benjamini (2003) (the third-most-cited work co-authored by Benjamini) where the hypotheses become still more general and the procedure itself may need to be revised.

REFERENCES

Benjamini, Y., and Hochberg, Y. (1995), Controlling the false discovery rate: A practical and powerful approach to multiple testing, *J. Roy. Statist. Soc., Ser. B* **57**, 289-300.

Benjamini, Y., and Yekutieli, D. (2001), The control of the false discovery rate in multiple testing under dependency, *Ann. Statist.* **29**, 1165-1188.

Lehmann, E. L. (1966), Some concepts of dependence, *Ann. Math. Statist.* **37**, 1137-1153.

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

Storey, J. D. (2002), A direct approach to false discovery rates, *J. Royal Statist. Soc. Ser. B* **64**, 479-498.