Published in final edited form as:

Commun Stat Theory Methods. 2008 January ; 37(12): 1855–1866. doi:10.1080/03610920801893731.

Expected Power for the False Discovery Rate with Independence

D. H. GLUECK¹, K. E. MULLER², A. KARIMPOUR-FARD¹, and L. HUNTER¹

¹Department of Preventive Medicine and Biometrics, University of Colorado Denver and Health Sciences Center, Denver, Colorado, USA

²Department of Epidemiology and Health Policy Research, Division of Biostatistics, University of Florida, Gainesville, Florida, USA

Abstract

The Benjamini–Hochberg procedure is widely used in multiple comparisons. Previous power results for this procedure have been based on simulations. This article produces theoretical expressions for expected power. To derive them, we make assumptions about the number of hypotheses being tested, which null hypotheses are true, which are false, and the distributions of the test statistics under each null and alternative. We use these assumptions to derive bounds for multiple dimensional rejection regions. With these bounds and a permanent based representation of the joint density function of the largest p-values, we use the law of total probability to derive the distribution of the total number of rejections. We derive the joint distribution of the total number of rejections and the number of rejections when the null hypothesis is true. We give an analytic expression for the expected power for a false discovery rate procedure that assumes the hypotheses are independent.

Keywords

Benjamini–Hochberg procedure; Distribution of number of rejections; Multiple comparisons; Rejection region bounds

1. Introduction

Our goal is to provide analytical formulas for the expected power for the Benjamini and Hochberg (1995) false discovery rate procedure. Benjamini and Hochberg (1995) controlled the False Discovery Rate, the expected ratio of the number of rejections of hypotheses where the null is true (false rejections) to the number of total rejections. With a single hypothesis, power is the probability of rejecting the null hypothesis. With multiple hypotheses, we study average, or expected power (Benjamini and Liu, 1999). Expected power is the expectation of the ratio of the number of rejections to the true number of hypotheses for which the alternative holds.

Many articles describe extensions of the theory for false discovery rate, or related statistics (see, e.g., Benjamini and Liu, 1999; Benjamini and Yekutieli, 2001; Curran-Everett, 2000; Efron et al., 2001; Finner and Roters, 2002; Genovese and Wasserman, 2002, ²⁰⁰⁴; Sarkar, 2002, 2004, 2006; Storey, 2002, 2003). Power has been studied almost entirely via simulation (Benjamini and Liu, 1999; Keselman et al., 2002; Lee and Whitmore, 2002; Storey, 2002).

In analytic power analysis, it is usual to make assumptions about the underlying state of nature. We assume that we know how many hypotheses we are testing and what statistics we will use to test them. We even assume that we know when each hypothesis is true or false, and the total number of true and false hypotheses. Thus, we know the true distribution of each test statistic. We also know the distribution of the p -value, whether uniform if the null hypothesis is true, or a different distribution, calculated under the alternative.

Using these assumptions, the expected power can be calculated by using the law of total probability. The steps in the calculation are as follows.

1. Find the cumulative distribution function and the probability density function of the individual p -values when the null hypothesis is true or false.
2. Show that the probability of rejection depends only on the largest p -values.
3. Use a computational form for the joint density function of the largest order statistics (Balakrishnan, 2007; Vaughan and Venables, 1972).
4. Enumerate and delimit the rejection regions implicitly defined by the Benjamini and Hochberg (1995) procedure in terms of p -value space.
5. Give an explicit recursive algorithm for generating the rejection regions.
6. Use the law of total probability to calculate the probability distribution function of the total number of rejections.
7. Derive the joint distribution of the total rejections and false rejections.
8. Finally, derive formulas for the expected power from this joint distribution.

In Sec. 2, we define the notation and adapt some previous known results to our problem. In Sec. 3, we prove lemmas on ordered lists. In Sec. 4, we derive the distributions of the number of total rejections, and false rejections, and give our main results on expected power. Technical proofs are in the Appendix.

2. Notation and Known Results

Suppose one plans to use the Benjamini and Hochberg (1995) false discovery rate procedure for m hypotheses and m decisions. We take a frequentist, parametric view, and envision each decision as being between a null and an alternative hypothesis, which may differ for each decision. Let $n \in \{0, \dots, m\}$ be the number of decisions for which a null hypothesis holds in the population, while an alternative hypothesis holds for the remaining $m - n$ decisions. For $i \in \{1, \dots, m\}$, index the hypotheses by H_i , with associated absolutely continuous, independent, but not necessarily identically distributed real valued test statistic T_i (with realization t_i) and random p -value, P_i .

In frequentist statistics, the null hypothesis is rejected when the p -value is smaller than a specified bound. In a multiple comparisons situation, there are many p -values. The Benjamini and Hochberg procedure uses the sorted p -values to decide which null hypotheses are rejected, and which are not. Let $\{P_{(1)} \leq \dots \leq P_{(m)}\}$ be the set of order statistics for the p -values, with $\{P_{(1)} \leq \dots \leq P_{(m)}\}$ a realization. Let $\alpha^* \in [0, 1]$ and $b_i = i\alpha^*/m$. This set of numbers, $b_1, b_2, \dots, b_m$, is a set of bounds. We compare the smallest p -value, $p_{(1)}$ to b_1 , the next largest p -value, $p_{(2)}$ to b_2 , and so on, until we find the largest i for which $p_{(i)} \leq i\alpha^*/m$. Call that subscript k . Thus, k is the largest number so that $p_k \leq k\alpha^*/m$, but $p_{(k+1)} > [(k+1)\alpha^*]/m$. The Benjamini and Hochberg (1995) testing procedure rejects the k hypotheses that correspond to the k smallest p -values. The event $\mathcal{R}_k$, the rejection of k hypotheses, is:

$$\mathcal{R}_k = \begin{cases} \left\{ \bigcap_{i=1}^m (p_{(i)} \geq b_i) \right\} & k=0 \\ \left\{ \bigcap_{i=k+1}^m (p_{(i)} \geq b_i) \cap (p_{(k)} \leq b_k) \right\} & 1 \leq k \leq m-1 \\ \{p_{(m)} \leq b_m\} & k=m. \end{cases} \quad (1)$$

Now consider repeating the experiment over and over again. Each time, because the data are stochastic, the number of rejections would vary. Thus, the number of rejections is a random variable. Let $K \in \{0, \dots, m\}$ be the random variable denoting the number of rejections of null hypotheses in an experiment, and k its realization. Rejecting a null hypothesis is a decision which may be correct, if the alternative holds, or incorrect, if the null actually holds. Rejections of hypotheses where the null is true are Type I errors. Let J be the number of rejections for which the null does hold, and j its realization. In general, $j \in \{\max[0, k - (m - n)], \dots, \min(n, k)\}$. Table 1 illustrates the relationships.

To find the distribution of K and J , we need to understand the probability density of the largest p -values. This depends on the distribution of the test statistics. Typically, the distribution of T_i differs under the null and alternative hypotheses. Let $h = 0$ if the null hypothesis holds, and $h = 1$ otherwise. Let $\theta_i(0)$ be the vector of parameters for the distribution function under the null, and $\theta_i(1)$ be the vector of parameters for the distribution function under the alternative. The number of rows of $\theta_i(0)$ may be different from the number of rows of $\theta_i(1)$. For example, the null may be a central chi-square with one parameter, the degrees of freedom, while the alternative may be a non central chi-square, with two parameters, the degrees of freedom, and a non centrality parameter. With N_i the sample size for T_i , the cumulative distribution and probability density functions are $F_{T_i}[t_i; N_i, \theta_i(0)]$ and $f_{T_i}[t_i; N_i, \theta_i(0)]$ under the null, and $F_{T_i}[t_i; N_i, \theta_i(1)]$ and $f_{T_i}[t_i; N_i, \theta_i(1)]$ under the alternative.

Using this notation, we can define the distribution and density function for each p -value. Since the test statistic T_i is absolutely continuous under the null and the alternative, the distribution function is smooth and monotone increasing, and the inverse distribution function exists. With $p_i = 1 - F_{T_i}[t_i; N_i, \theta_i(h)] \in [0, 1]$, $h \in \{0, 1\}$, yields

$$t_i = F_{T_i}^{-1} [1 - p_i; N_i, \theta_i(h)]. \quad (2)$$

Let $q_i = F_{T_i}^{-1} [1 - p_i; N_i, \theta_i(0)]$. For a one-tailed test for which larger values lead to rejection, the distribution and density functions for P_i are:

$$F_{P_i} [p_i; N_i, \theta_i(h)] = 1 - F_{T_i} [q_i; N_i, \theta_i(h)], \quad (3)$$

$$f_{P_i} [p_i; N_i, \theta_i(h)] = \frac{f_{T_i} [q_i; N_i, \theta_i(h)]}{f_{T_i} [q_i; N_i, \theta_i(0)]}. \quad (4)$$

Let $q_{vi} = F_{T_i}^{-1} [1 - p_i/2; N_i, \theta_i(0)]$ and $q_{li} = F_{T_i}^{-1} [p_i/2; N_i, \theta_i(0)]$. For a two-tailed test with equal tail probabilities, the distribution for P_i is

$$F_{P_i} [p_i; N_i, \theta_i(h)] = 1 - F_{T_i} [q_{Ui}; N_i, \theta_i(h)] + F_{T_i} [q_{Li}; N_i, \theta_i(h)]. \quad (5)$$

Using Theorem 12.4.4, (Leithold, 1968, p. 410), the density function is:

$$f_{P_i} [p_i; N_i, \theta_i(h)] = \frac{1}{2} \cdot \frac{f_{T_i} [q_{Ui}; N_i, \theta_i(h)]}{f_{T_i} [q_{Ui}; N_i, \theta_i(0)]} + \frac{1}{2} \cdot \frac{f_{T_i} [q_{Li}; N_i, \theta_i(h)]}{f_{T_i} [q_{Li}; N_i, \theta_i(0)]}. \quad (6)$$

For clarity, abbreviate the distribution function and density for P_i by $F_{P_i}(p_i)$ and $f_{P_i}(p_i)$. Notice that the p -values are absolutely continuous, and independent, but not necessarily identically distributed. By assumption, each decision has a potentially different null and alternative hypothesis. This can lead to different distributions for each p -value.

We need to look at the joint density of the p -values to understand the power for the Benjamini and Hochberg procedure. But which p -values? Suppose $m = 10$ and $k = 4$, so 4 hypotheses are rejected by the Benjamini and Hochberg procedure. The idea here is that the four smallest p -values are less than their bounds. We need to look also at the 5th smallest p -value, to make sure that it is bigger than its bound. If it were smaller than its bounds, we would have rejected five, not four. We also need to look at the 6th smallest one, and the 7th, 8th, 9th, and 10th. In fact, we need the joint density of the largest $(m - k) + 1$ of m of the p -values to figure out the probability of rejecting k hypotheses.

A convenient form of the joint density of independent, but not necessarily identically distributed random variables is in terms of a permanent (Balakrishnan, 2007; Vaughan and Venables, 1972). The permanent of a square matrix is defined like the determinant, except that all signs are positive (Aitken, 1999, p. 30). If A is a square matrix, we write its permanent by $\text{per}[A]$. For $k \in \{1, \dots, m\}$, the marginal density of $p_{(k)} \leq \dots \leq p_{(m)}$ is

$$\begin{aligned} & f_{P_{(k)}, \dots, P_{(m)}}(p_{(k)}, \dots, p_{(m)}) \\ &= [(k-1)!]^{-1} \times \text{per} \begin{bmatrix} F_{P_1}(p_{(k)}) & F_{P_2}(p_{(k)}) & \cdots & F_{P_m}(p_{(k)}) \\ \vdots & \vdots & & \vdots \\ F_{P_1}(p_{(k)}) & F_{P_2}(p_{(k)}) & \cdots & F_{P_m}(p_{(k)}) \\ \text{---} & \text{---} & \text{---} & \text{---} \\ f_{P_1}(p_{(k)}) & f_{P_2}(p_{(k)}) & \cdots & f_{P_m}(p_{(k)}) \\ \vdots & \vdots & & \vdots \\ f_{P_1}(p_{(m)}) & f_{P_2}(p_{(m)}) & \cdots & f_{P_m}(p_{(m)}) \end{bmatrix}, \end{aligned} \quad (7)$$

where the first block contains $k - 1$ rows, and the second block contains $(m - k) + 1$ rows.

When the test statistics are identically distributed, the result reduces (David, 1981, p. 10) to

$$f_{P_{(k)}, \dots, P_{(m)}}(p_{(k)}, \dots, p_{(m)}) = \frac{m! [F(p_{(k)})]^{k-1}}{(k-1)!} \prod_{i=k}^m f(p_{(i)}). \quad (8)$$

3. Bounds for Integration

In Eq. (1), we described the rejection regions corresponding to the Benjamini and Hochberg (1995) procedure. Although these are exact definitions of the rejection regions, they do not

describe all the ways the p -values could be arranged with respect to the bounds. If we understand all the ways that the p -values can be arranged with respect to the bounds, we can define a set of rejection regions in p -value space. If we integrate the joint density of the p -values over this rejection region, we will get the probability distribution of the number of rejections.

There are many ways the p -values can be arranged with respect to the bounds and still cause a rejection. First, let us consider a small example, and demonstrate how we can proceed from a pictorial representation, to a set of ordered lists, to a set of integration bounds. Suppose that we are considering two hypotheses, so $m = 2$. How many ways can we reject both hypotheses? Let $b_1 = \alpha^*/2$, and $b_2 = \alpha^*$. The rejection region is $\{p_{(2)} \leq b_2\}$. Notice that the rejection region can be satisfied in three different ways, shown as number lines in Fig. 1.

We can represent the number lines shown in Fig. 1 abstractly as ordered lists. The ordered lists corresponding to Fig. 1 are shown in Fig. 2. We still need to define the bounds of the rejection region. For this small example, the bounds are shown in Fig. 3.

For the general case, we need to formally define ordered lists. Suppose q is a finite, positive integer, and $a_1, \dots, a_q$ are real numbers, with $a_1 \leq \dots \leq a_q$. $\mathcal{A}$ is an ordered list if $\mathcal{A} = \{a_1, \dots, a_q\}$. If $\mathcal{A}$ and $\mathcal{B}$ are ordered lists with $a_j \leq b_{j'}$, $\forall j$ and j' , then their concatenation is

$$C = \mathcal{A} \& \mathcal{B}, \quad (9)$$

with C an ordered list whose elements are the entries in $\mathcal{A}$ followed by those in $\mathcal{B}$.

Now we will give two lemmas that describe how to generate and use ordered lists for more general cases. We will use these ordered lists to define the rejection regions for Benjamini and Hochberg (1995). The proofs are given in the Appendix. The first lemma shows how to generate these ordered lists for a specific number of hypothesis tests.

The second lemma shows that these ordered lists correspond to ordered sets of ordered pairs that are bounds for integration. Integrating the marginal density of the $(m - k) + 1$ smallest order statistics of the p -values over these bounds will give the probabilities of rejection.

Lemma 3.1

Suppose $k \in \{0, \dots, m\}$ indexes the number of rejections. Let

$$c_k = [2 \cdot (m - k)]! / [(m - k)! (m - k)! (m - k + 1)]. \quad (10)$$

Let $p \in \{1, \dots, c_k\}$, and let $q \in \{1, \dots, 2(m - k) + 2\}$ be index variables. Let $e_{k,p,q} \in \{p_{(k)}, \dots, p_{(m)}; b_k, \dots, b_m\}$ be the entries of an ordered list, so that

$$0 \leq e_{k,p,1} \leq \dots \leq e_{k,p,2k+2} \leq 1. \quad (11)$$

Notate the list itself by letting

$$\mathcal{L}_{k,p} = \{0, e_{k,p,1}, \dots, e_{k,p,2k+2}, 1\}. \quad (12)$$

Let $\mathcal{L}_k$ be the set of such ordered lists so that

$$\mathcal{L}_k = \{\mathcal{L}_{k,1}, \mathcal{L}_{k,2}, \dots, \mathcal{L}_{k,c_k}\}. \quad (13)$$

Then the set $\mathcal{L}_{k-1}$ can be generated from the set $\mathcal{L}_k$. The number of entries in $\mathcal{L}_k$ is c_k , a Catalan number.

Lemma 3.2

Let $k \in \{0, \dots, m\}$ index the number of rejections and $p \in \{1, \dots, c_k\}$ be an index variable for the number of ordered lists in the set $\mathcal{L}_k$. Let $o \in \{1, \dots, k+1\}$ index the variable of integration. Suppose $l_{k,p,o}$ is the lower bound of the integral with respect to $dp_{(o)}$, and $u_{k,p,o}$ the corresponding upper bound. For each ordered list $\mathcal{L}_{k,p}$, define the ordered set of ordered pairs

$$\mathcal{B}_{k,p} = \begin{cases} \{(0, u_{m,p,1}), (l_{m,p,2}, u_{m,p,2}), \dots, (l_{m,p,k}, u_{m,p,k})\} & k=0 \\ \{(0, u_{k,p,1}), (l_{k,p,2}, u_{k,p,2}), \dots, (l_{k,p,k}, u_{k,p,k}), (b_{k+1}, 1)\} & 1 \leq k \leq m-1 \\ \{(0, b_m)\} & k=m. \end{cases} \quad (14)$$

Let $\mathcal{B}_k$ be the set of such ordered lists so that

$$\mathcal{B}_k = \{\mathcal{B}_{k,1}, \dots, \mathcal{B}_{k,c_k}\}. \quad (15)$$

Then:

1. $\mathcal{B}_{k,p}$ can be formed from the ordered lists $\mathcal{L}_{k,p}$ in Lemma 3.1 by construction.
2. $\mathcal{B}_k$ can be formed from the set $\mathcal{L}_k$.
3. The number of entries in $\mathcal{B}_k$ is c_k .
4. For $0 \leq k \leq m-1$, the $k+1$ ordered pairs in $\mathcal{B}_{k,p}$ are bounds of integration for a $k+1$ dimensional integral. For $k=m$, the ordered pair in $\mathcal{B}_{m,p}$ is the bounds of integration for a one dimensional integral.
5. The ordered pairs in $\mathcal{B}_{k,p}$ delineate regions in $\mathcal{R}^{k+1}$, and the ordered pair in $\mathcal{B}_{m,p}$ delineates a region in $\mathcal{R}$. These are the rejection regions from Eq. (1).

4. Exact Distributions and Expected Power

We now have expressions for the joint density of the order statistics of the p -values (Eq. (7)), and an expression for the bounds of the rejection regions (Lemma 3.1). Using the law of total probability, in this section we derive the probability distribution for the total number of rejections, and the joint distribution of the total number of rejections and the number of false rejections. We use these results to find expected power.

Result 4.1

The distribution of the number of rejections for independent, but not necessarily identically distributed random variables, and corresponding p -values is given by:

$$\Pr \{K=k\} = \begin{cases} \sum_{p=1}^{c_0} \int_{\mathcal{B}_{0,p}} m! \prod_{i=1}^m f_{p_i}(p_{(i)}) dp_{(1)} \dots dp_{(m)} & k=0 \\ \sum_{p=1}^{c_k} \int_{\mathcal{B}_{k,p}} f_{p_{(k)}, \dots, p_{(m)}}(p_{(k)}, \dots, p_{(m)}) dp_{(k)} \dots dp_{(m)} & 1 \leq k \leq m-1 \\ \prod_{i=1}^m F_{p_i}(b_m) & k=m. \end{cases} \quad (16)$$

Result 4.2

For independent and identically distributed random variables, and corresponding p -values

$$\Pr \{K=k\} = \begin{cases} \sum_{p=1}^{c_0} \int_{\mathcal{B}_{0,p}} m! \prod_{i=1}^m f(p_{(i)}) dp_{(1)} \dots dp_{(m)} & k=0 \\ \sum_{p=1}^{c_k} \int_{\mathcal{B}_{k,p}} \frac{m! [F(p_{(k)})]^{k-1}}{(k-1)!} \prod_{i=k}^m f(p_{(i)}) dp_{(k)} \dots dp_{(m)} & 1 \leq k \leq m-1 \\ [F(b_m)]^m & k=m. \end{cases} \quad (17)$$

Result 4.3

For independent and identically and uniformly distributed random variables,

$$\Pr \{K=k\} = \begin{cases} 1 - \alpha^* & k=0 \\ \sum_{p=1}^{c_k} \int_{\mathcal{B}_{k,p}} \frac{m! (p_{(k)})^{k-1}}{(k-1)!} dp_{(k)} \dots dp_{(m)} & 1 \leq k \leq m-1 \\ (b_m)^m & k=m. \end{cases} \quad (18)$$

Proof—Results 4.4–4.6 follow from the law of total probability. The probability of rejection is given by integrating the appropriate density over the multidimensional region. Result 4.3 occurs when the null holds for every hypothesis, and the p -values are then uniformly, and identically distributed.

Now, we need to derive the joint probability distribution of the total number of rejections and the number of false rejections. Without loss of generality, we can label the hypotheses H_i so that for $i \in \{1, \dots, n\}$, the null hypothesis holds, and for $i \in \{n+1, \dots, m\}$, the alternative hypothesis holds. If $n = m$, all of the null hypotheses hold, while if $n = 0$, then all alternative hypotheses hold.

Index the $m!$ terms in the expansion of the permanent in Eq. (7) by $v \in \{1, \dots, m!\}$, and denote term v by a_v . Each term is a product of m factors. All entries in the last $[m-k+1]$ rows of the permanent in Eq. (7) are of the form $f_{p_d}(p_{(e)})$. Describe any such entry with $d \in \{1, \dots, n\}$ and $e \in \{k+1, \dots, m\}$ as satisfying the (d, e) condition. These entries correspond to a null hypothesis underlying one of the $m-k$ largest p -values, or a null hypothesis that is not rejected. Let C_0 be the set of terms which have no factors which fulfill the (d, e) condition. If n is constrained to be zero, then C_0 is all the terms in the density, and all other sets C_{n-j} are empty. If $n \neq 0$, but $(n-j) = 0$, then C_0 is the set of all terms in the density that have 0 factors that fulfill the (d, e) condition. In general, let C_{n-j} be the set of terms that have exactly $n-j$ factors that fulfill the (d, e) condition. We can then rewrite the density of the largest $(m-k)+1$ p -values as a sum over these sets.

$$f_{p_{(k)}, \dots, p_{(m)}}(p_{(k)}, \dots, p_{(m)}) = \sum_j \sum_{v: a_v \in C_{n-j}} a_v. \quad (19)$$

Result 4.4

If the number of hypotheses for which the alternative holds is zero, then $n = m$. Then the number of rejections of null hypotheses is equal to the number of total rejections and $k = j$. Then $\Pr\{J = j, K = k\} = \Pr\{J = j\} = \Pr\{K = k\}$, the distribution in Result 4.1.

We also give some special cases that may be useful, depending on the testing situation. They can be derived from Results 4.2 and 4.3.

Result 4.5

For independent, but not necessarily identically distributed random variables, and with $j \in \{\max[0, k - (m - n)], \dots, \min(n, k)\}$,

$$\Pr\{J=j, K=k\} = \begin{cases} \sum_{p=1}^{c_0} \int_{\mathcal{B}_{0,p}} m! \prod_{i=1}^m f_{p_i}(p_{(i)}) dp_{(1)} \dots dp_{(m)} & k=0, j=0 \\ \sum_{p=1}^{c_k} \int_{\mathcal{B}_{k,p}} \sum_{v: a_v \in C_{n-j}} a_v dp_{(k)} \dots dp_{(m)} & 1 \leq k \leq m-1 \\ \prod_{i=1}^m F_{p_i}(b_m) & k=m, j=n. \end{cases} \quad (20)$$

Result 4.6

With ${}_a C_b = a!/[b!(a-b)!]$ and for independent and identically distributed random variables, the joint probability distribution function of the number of total and false rejections is given by:

$$\Pr\{J=j, K=k\} = {}_n C_{j \cdot (m-n)} {}_m C_{(k-j)} \cdot \Pr\{K=k\} / {}_m C_k. \quad (21)$$

Proof—Results 4.4–4.6 follow from direct applications of the law of total probability.

We now have the joint distribution of the number of true rejections, K , and the number of rejections of nulls, J . Power for one realization of the experiment is $(k-j)/(m-n)$. The numerator is the total number of rejections minus the number of nulls which are rejected. Thus, it is the number of hypotheses for which the alternative is true that are rejected. The denominator is the number of hypotheses for which the alternative is true. This quantity is also known as sensitivity. Benjamini and Liu (1999) called the expected value of this quantity expected power, and used it in simulations as measure of the success of an experiment that tested multiple hypotheses. We can now give an explicit, analytic formula for expected power.

Result 4.7

For $m \neq n$, i.e., when the null is not true for every hypothesis, the expected power is given by:

$$\mathcal{E}[(K - J) / (m - n)] = \sum_{k=0}^m \sum_j [(k - j) / (m - n)] \Pr\{K=k; J=j\}. \quad (22)$$

Result 4.8

A formula of Lindgren (1976, p. 116) gives the expected value of a function of two random variables in terms of the joint distribution. The result then follows from the definitions of the various estimators.

The expected power is given in Eq. (22). How does the power depend on the sample size? The distribution of each test statistic, T_i , includes the sample size for the hypothesis H_i , N_i , as a parameter. Thus, the distribution for each p -value, P_i , also includes the sample size. The joint density of the ordered p -values includes the distribution and density functions for the p -values, and is thus a function of set of the sample sizes for each hypothesis, $\{N_1, \dots, N_m\}$. The probability distribution function of the total number of rejections, and the number of false rejections is expressed as an integral of the joint probability density function, and thus is also a function of the set of sample sizes. Finally, the expected power is a function of the set of sample sizes, since it involves an expectation over the joint distribution of total and false rejections. Thus, one can calculate the power for a fixed set of sample sizes.

Acknowledgments

The authors thank Gary Grunwald for his close reading and numerous suggestions, which greatly improved the manuscript.

Glueck was supported by NCI K07CA88811. Muller was supported by NCI P01 CA47 982-04, NCI R01 CA095749-01A1, and NIAID 9P30 AI 50410. Hunter was supported by NLM 5R01LM008111-03 and NCI 5 P30 CA46934-15.

5. Appendix

Proof of Lemma 3.1

By construction.

1. With $k = m$, let

$$\mathcal{L}_m = \mathcal{L}_{m,1} = \{0, p_{(m)}, b_m, 1\}. \quad (23)$$

2. If $k = 0$, then do the following:

- a. In $\mathcal{L}_{k,p}$, for all $p \in \{1, \dots, c_k\}$ delete $p_{(0)}$ and b_0 .
- b. Stop generating lists.

3. Otherwise, repeat the following steps for $p \in \{1, \dots, c_k\}$. Consider $\mathcal{L}_{k,p}$.

- a. Insert the ordered list $\{p_{(k-1)}, b_{k-1}\}$ after 0 in each ordered list that corresponds to k rejections.
- b. Remove $p_{(k)}$ from the list.
- c. From the resulting list, remove the prefix, which is the ordered list $\{0, \dots, b_k\}$.
- d. From the resulting list, remove the suffix, which is the ordered list $\{p_{(k+1)}, 1\}$. If $k = m$, remove the suffix, which is the ordered list $\{1\}$.
- e. What remains is the core. If the core has nothing in it, insert $p_{(k)}$. Otherwise, insert $p_{(k)}$ sequentially before and after every entry of the core. The resulting ordered lists are the new cores.

- f. For each core, insert the prefix at the beginning, and add the suffix onto the end. The result set of ordered lists are the elements of the set $\mathcal{L}_{k-1}$, the set of ordered lists that correspond to $k - 1$ rejections.
4. Let $k = k - 1$, and go to Step 2.

Proof of Lemma 3.2

There are five separate statements in Lemma 3.2. We present the proofs of the five statements in order.

1. By construction, $\mathcal{B}_{m,p} = \{(0, b_1)\}$ by definition. For $k \in \{0, \dots, m - 1\}$, $\mathcal{B}_{k,p}$ can be formed from $\mathcal{L}_{k,p}$ using the following algorithm.
 - a. Write the first three elements of $\mathcal{L}_{k,p}$ as an ordered list. Call it the trio. The fourth through last elements of $\mathcal{L}_{k,p}$ call the *remainder*.
 - i. If the middle element is of the form $p_{(o)}$, define $l_{k,p,1}$ to be the first element of the trio, and define $u_{k,p,1}$ to be the third element of the trio, and then do the following. Otherwise, proceed to Step ii.
 - A. Add the ordered pair $(l_{k,p,1}, u_{k,p,1})$ to the ordered list $\mathcal{B}_{k,p}$.
 - B. Form a new trio with the same first element. The middle element of the new trio is the third element of the original trio. The third element of the new trio is the first element of the remainder.
 - C. Remove the first element of the remainder. The new remainder is the old remainder with the first element removed.
 - ii. If the middle element of the trio is of the form b_j , then do the following.
 - A. Delete the first element of the trio.
 - B. Form a new trio. Let the middle element be the new first element, the new second element be the old third element, and let the new third element be the first item in the remainder.
 - C. Remove the first element of the remainder. The new remainder is the old remainder with the first element removed.
 - b. Using the new trio and the new remainder defined above, repeat Steps i and ii until all the elements of $\mathcal{L}_{k,p}$ are exhausted.
2. $\mathcal{B}_k$ can be formed from the set $\mathcal{L}_k$ in the following manner. First, form all the $\mathcal{B}_{k,p}$ from $\mathcal{L}_{k,p}$ in $\mathcal{L}_k$ as shown above. Second, $\mathcal{B}_k$ is simply the set of all the $\mathcal{B}_{k,p}$.
3. This holds since the number of elements in $\mathcal{L}_k$ is c_k and there are the same number of elements in $\mathcal{L}_k$ and $\mathcal{B}_k$.

4. By inspection.
5. They correspond to the rejection regions by construction.

References

- Aitken, AC. Determinants and Matrices. Oliver and Boyd; Edinburgh: 1999.
- Balakrishnan N. Permanents, order statistics, outliers and robustness. *Revista Matematica Complutense* 2007;20:7–107.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B Statist. Methodol* 1995;57:289–300.
- Benjamini Y, Liu W. A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence. *J. Statist. Plann. Infer* 1999;82:163–170.
- Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann. Statist* 2001;29:1165–1188.
- Curran-Everett D. Multiple comparisons: philosophies and illustrations. *Am. J. Physiol. Regul. Integr. Comp. Physiol* 2000;279:R1–R8. [PubMed: 10896857]
- David, HA. Order Statistics. 2nd ed.. Wiley; New York: 1981.
- Efron B, Storey J, Tibshirani R. Microarrays, empirical Bayes methods, and false discovery rates. *J. Amer. Statist. Assoc* 2001;96:1151–1160.
- Finner H, Roters M. Multiple hypothesis testing and expected number of Type 1 errors. *Ann. Statist* 2002;30(1):220–238.
- Genovese CR, Wasserman L. Operating characteristics and extensions of the false discovery rate procedure. *J. Roy. Stat. Soc. Ser. B Statist. Methodol* 2002;64:499–517.
- Genovese CR, Wasserman L. A stochastic process approach to false discovery control. *Ann. Statist* 2004;32(3):1035–1061.
- Keselman HJ, Cribbie R, Holland B. Controlling the rate of Type 1 error over a large set of statistical tests. *Bri. J. Math. Statist. Psych* 2002;55:27–39.
- Lee M, Whitmore G. Power and sample size for DNA microarray studies. *Statist. Med* 2002;21:3543–3570.
- Leithold, L. The Calculus with Analytic Geometry. Harper and Row; New York: 1968.
- Lindgren, BW. Statistical Theory. 3rd ed.. Macmillan Publishing; New York: 1976.
- Sarkar SK. Some results on false discovery rate in stepwise multiple testing procedures. *Ann. Statist* 2002;30(1):239–257.
- Sarkar SK. FDR-controlling stepwise procedures and their false negatives rates. *J. Statist. Plann. Infer* 2004;125:119–137.
- Sarkar SK. False discovery and false nondiscovery rates in single-Step multiple testing procedures. *Ann. Statist* 2006;34(1):394–415.
- Storey J. A direct approach to the false discovery rate. *J. Roy. Statist. Soc. Ser. B Statist. Methodol* 2002;64:479–598.
- Storey J. The positive false discovery rate: a Bayesian interpretation and the q -value. *Ann. Statist* 2003;31(6):2013–2035.
- Vaughan RJ, Venables WN. Comments and queries: permanent expressions for order statistic densities. *J. Roy. Statist. Soc. Ser. B Statist. Methodol* 1972;34:308–310.

Figure 1.
Testing and rejecting two hypotheses.

$$\begin{aligned} &\{0, b_1, p_{(1)}, p_{(2)}, b_2, 1\} \\ &\{0, p_{(1)}, b_1, p_{(2)}, b_2, 1\} \\ &\{0, p_{(1)}, p_{(2)}, b_1, b_2, 1\} \end{aligned}$$

Figure 2.

Ordered lists for testing and rejecting two hypotheses.

$$\begin{aligned}
 p_{(1)} &\in (b_1, p_{(2)}), p_{(2)} \in (0, b_2) \\
 p_{(1)} &\in (0, b_1), p_{(2)} \in (b_1, b_2) \\
 p_{(1)} &\in (0, p_{(2)}), p_{(2)} \in (0, b_1)
 \end{aligned}$$

Figure 3.

Integration bounds for a rejection for rejecting two hypotheses.

Table 1

Types of decisions for a particular realization of an experiment decision

True state	Reject	Do not reject	
Null	j	$n - j$	n
Alternative	$k - j$	$(m - n) - (k - j)$	$(m - n)$
	k	$m - k$	m