

Multiple Testing Methods for the Analysis of Microarray Data

3/9/2009

Copyright © 2009 Dan Nettleton

1

The Multiple Testing Problem

- Suppose one test of interest has been conducted for each of m genes in a microarray experiment.
- Let p_1, p_2, \dots, p_m denote the p -values corresponding to the m tests.
- Let $H_{01}, H_{02}, \dots, H_{0m}$ denote the null hypotheses corresponding to the m tests.

2

The Multiple Testing Problem (continued)

- Suppose m_0 of the null hypotheses are true and m_1 of the null hypotheses are false.
- Let c denote a value between 0 and 1 that will serve as a cutoff for significance:
 - Reject H_{0i} if $p_i \leq c$ (declare significant)
 - Fail to reject (or accept) H_{0i} if $p_i > c$ (declare non-significant)

3

Table of Outcomes

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	U	V	m_0
False Nulls	T	S	m_1
Total	W	R	m

4

Table of Outcomes

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	U	V	m_0
False Nulls	T	S	m_1
Total	W	R	m

U=number of true negatives

5

Table of Outcomes

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	U	V	m_0
False Nulls	T	S	m_1
Total	W	R	m

V=number of false positives
=number of false discoveries
=number of type 1 errors

6

Table of Outcomes

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	U	V	m_0
False Nulls	T	S	m_1
Total	W	R	m

T=number of type 2 errors

7

Table of Outcomes

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	U	V	m_0
False Nulls	T	S	m_1
Total	W	R	m

S=number of true positives
=number of true discoveries

8

Table of Outcomes

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	U	V	m_0
False Nulls	T	S	m_1
Total	W	R	m

W=number of non-rejections
(number of null hypotheses accepted)

9

Table of Outcomes

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	U	V	m_0
False Nulls	T	S	m_1
Total	W	R	m

R=number of rejections
(of null hypotheses)

10

Table of Outcomes

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	U	V	m_0
False Nulls	T	S	m_1
Total	W	R	m

↑

Random Variables

↑

Constants

11

Table of Outcomes

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	U	V	m_0
False Nulls	T	S	m_1
Total	W	R	m

↙

Unobservable

↑

Observable

12

Familywise Error Rate (FWER)

- Traditionally statisticians have focused on controlling FWER when conducting multiple tests.
- FWER is defined as the probability of one or more false positive results:
$$\text{FWER} = P(V > 0).$$
- Controlling FWER amounts to choosing the significance cutoff c so that FWER is less than or equal to some desired level α .

13

The Bonferroni Method

- The Bonferroni Method is the simplest way to achieve control of the FWER at any desired level α .
- Simply choose $c = \alpha / m$.
- With this value of c , the FWER will be no larger than α for any family of m tests.

14

Weak Control vs. Strong Control

- A method provides *weak control* of an error rate for a family of m tests if the error rate is controlled whenever all null hypotheses are true ($m = m_0$).
- A method provides *strong control* of an error rate for a family of m tests if the error rate is controlled no matter how many or which of the m tests have true null hypotheses.

15

Holm's Method

- Both the Bonferroni method and the Holm method provide strong control of the FWER for any family of m tests.
- Holm's method is less conservative than the Bonferroni method.
- The methods will provide the same results for many data sets, but sometimes Holm's method will result in more rejected null hypotheses.

16

Holm's Method for Controlling FWER at Level α

- Let $p_{(1)}, p_{(2)}, \dots, p_{(m)}$ denote the m p -values ordered from smallest to largest.
- Find the largest integer k so that
$$p_{(i)} \leq \alpha / (m - i + 1) \text{ for all } i = 1, \dots, k.$$
- If no such k exists, set $c = 0$ (declare nothing significant).
- Otherwise set $c = p_{(k)}$ (reject the nulls corresponding to the smallest k p -values).

17

An Example

- Suppose we conduct 5 tests and obtain the following p -values for tests 1 through 5.

Test	1	2	3	4	5
p -value	0.042	0.001	0.031	0.014	0.007

- Which tests' null hypotheses will you reject if you wish to control the FWER at level 0.05?
- Use both the Bonferroni method and the Holm method to answer this question.

18

Solution

Test	1	2	3	4	5
<i>p</i> -value	0.042	0.001	0.031	0.014	0.007

- The cutoff for significance is $c = 0.05/5 = 0.01$ using the Bonferroni method. Thus we would reject the null hypothesis for tests 2 and 5 with the Bonferroni method.

$0.001 \leq 0.05/(5-1+1) = 0.01$
 $0.007 \leq 0.05/(5-2+1) = 0.0125$
 $0.014 \leq 0.05/(5-3+1) = 0.0167$
 $0.031 > 0.05/(5-4+1) = 0.025$
 $0.042 \leq 0.05/(5-5+1) = 0.05$

These calculations indicate that Holm's method would reject null hypotheses for tests 2, 5, and 4.

19

A Conceptual Description of FWER

- Suppose a scientist conducts 100 independent microarray experiments.
- For each experiment, the scientist produces a list of genes declared to be differentially expressed by testing a null hypothesis for each gene.
- Each list that contains one or more false positive results is considered to be in error.
- The FWER is approximated by the proportion of 100 lists that contain one or more false positives.

20

FWER Too Conservative for Microarrays?

- Suppose that one of the 100 gene lists consists of 500 genes declared to be differentially expressed.
- Suppose that 1 of those 500 genes is not truly differentially expressed but that the other 499 are.
- This list is considered to be in error and such lists are allowed to make up only a small proportion of the total number of lists if FWER is to be controlled.
- However such a list seems quite useful from the scientific viewpoint. Perhaps it is not so important to control FWER for most microarray experiments.

21

False Discovery Rate (FDR)

- FDR is an alternative error rate that can be useful for microarray experiments.
- FDR was introduced by Benjamini and Hochberg (1995) and is formally defined as

$$E(Q) \text{ where } Q = V/R \text{ if } R > 0 \text{ and } Q = 0 \text{ otherwise.}$$
- Controlling FDR amounts to choosing the significance cutoff c so that FDR is less than or equal to some desired level α .

22

A Conceptual Description of FDR

- Suppose a scientist conducts 100 independent microarray experiments.
- For each experiment, the scientist produces a list of genes declared to be differentially expressed by testing a null hypothesis for each gene.
- For each list consider the ratio of the number of false positive results to the total number of genes on the list (set this ratio to 0 if the list contains no genes).
- The FDR is approximated by the average of the ratios described above.

23

FDR: The Appropriate Error Rate for Microarrays?

- The hypothetical gene list discussed previously with 1 false positive and 499 true positives would be a good list that would help to keep the FDR down.
- Some of the gene lists may contain a high proportion of false positive results and yet the method we are using may still control FDR at a given level because it is the average performance across repeated experiments that matters.
- The comment above applies to FWER control as well.

24

The Benjamini and Hochberg Procedure for Strongly Controlling FDR at Level α

- Let $p_{(1)}, p_{(2)}, \dots, p_{(m)}$ denote the m p -values ordered from smallest to largest.
- Find the largest integer k so that $p_{(k)} \leq k \alpha / m$.
- If no such k exists, set $c = 0$ (declare nothing significant).
- Otherwise set $c = p_{(k)}$ (reject the nulls corresponding to the smallest k p -values).

25

Our Example Revisited

- Suppose we conduct 5 tests and obtain the following p -values for tests 1 through 5.

Test	1	2	3	4	5
p -value	0.042	0.001	0.031	0.014	0.007

- Which tests' null hypotheses will you reject if you wish to control the FDR at level 0.05?
- Use the Benjamini and Hochberg (1995) method to answer this question.

26

Solution

Test	1	2	3	4	5
p -value	0.042	0.001	0.031	0.014	0.007

$$0.001 \leq 1 * 0.05 / 5 = 0.01$$

$$0.007 \leq 2 * 0.05 / 5 = 0.02$$

$$0.014 \leq 3 * 0.05 / 5 = 0.03$$

$$0.031 \leq 4 * 0.05 / 5 = 0.04$$

$$0.042 \leq 5 * 0.05 / 5 = 0.05$$

The B&H method would reject the null hypotheses for all 5 tests.

27

New Example (p_3 changed slightly)

- Suppose we conduct 5 tests and obtain the following p -values for tests 1 through 5.

Test	1	2	3	4	5
p -value	0.042	0.001	0.041	0.014	0.007

- Which tests' null hypotheses will you reject if you wish to control the FDR at level 0.05?
- Use the Benjamini and Hochberg (1995) method to answer this question.

28

Solution

Test	1	2	3	4	5
p -value	0.042	0.001	0.041	0.014	0.007

$$0.001 \leq 1 * 0.05 / 5 = 0.01$$

$$0.007 \leq 2 * 0.05 / 5 = 0.02$$

$$0.014 \leq 3 * 0.05 / 5 = 0.03$$

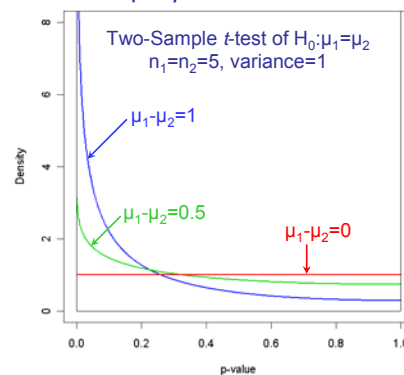
$$0.041 > 4 * 0.05 / 5 = 0.04$$

$$0.042 \leq 5 * 0.05 / 5 = 0.05$$

The B&H method would still reject the null hypotheses for all 5 tests even though $0.041 > 0.04$.

29

Example p -value Distributions



30

An Example

- Suppose 10,000 genes are tested for differential expression between two treatments.
- Suppose 200th smallest p -value is 0.001.
- If no genes were truly differentially expressed, how many of the 10,000 p -values would be expected to be less than or equal to 0.001?
- Use the calculations above to provide an estimate of the proportion of false positive results among the list of 200 genes with p -values no larger than 0.001.

31

Solution

- If all 10,000 null hypotheses were true, we would expect $0.001 \cdot 10,000 = 10$ tests to yield p -values less than 0.001
- A simple estimate of the proportion of false positive results among the list of 200 genes with p -values less than 0.001 is $0.001 \cdot 10,000 / 200 = 0.05$.
- Recall that the B&H FDR procedure involves finding the largest integer k so that $p_{(k)} \leq k \alpha / m$.
- This is equivalent to finding the largest integer k such that $p_{(k)} m / k \leq \alpha$.

32

Other Methods for Estimating or Controlling FDR

Rather than finding the largest integer k such that

$$p_{(k)} m / k \leq \alpha,$$

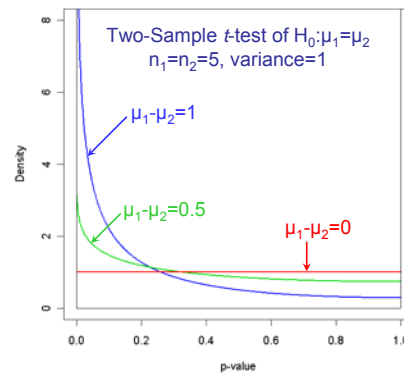
consider finding the largest integer k such that

$$p_{(k)} \hat{m}_0 / k \leq \alpha,$$

where \hat{m}_0 is an estimate of the number of true null hypotheses among the m tests.

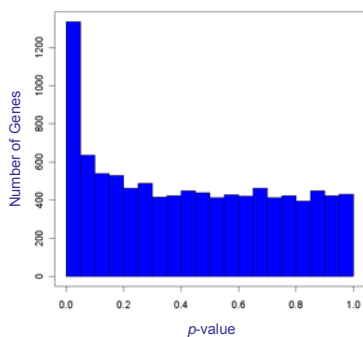
33

Example p -value Distributions



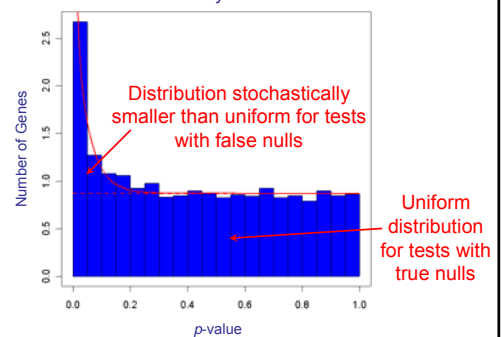
34

Histogram of p -values for a Test of Interest



35

Mixture of a Uniform Distribution and a Distribution Stochastically Smaller than Uniform



36

Recall that estimating the number of true null hypotheses m_0 is important for estimating FDR.

Rather than finding the largest integer k such that

$$p_{(k)} m / k \leq \alpha,$$

consider finding the largest integer k such that

$$p_{(k)} \hat{m}_0 / k \leq \alpha,$$

where \hat{m}_0 is an estimate of the number of true null hypotheses among the m tests.

37

Methods of Estimating m_0

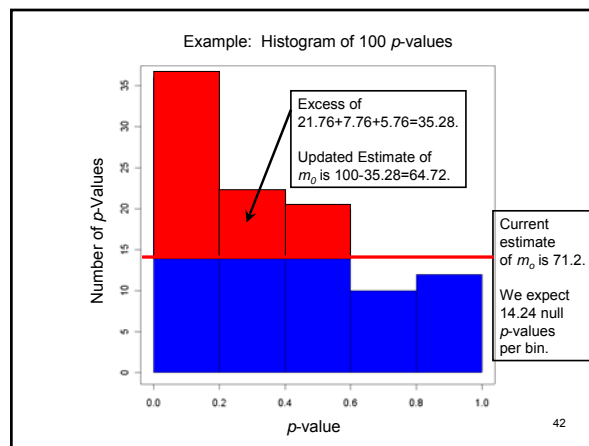
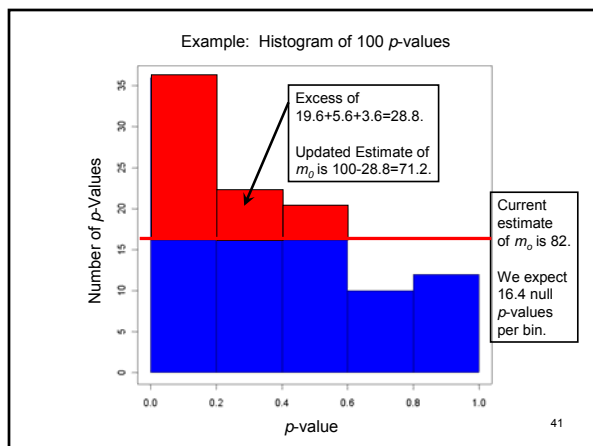
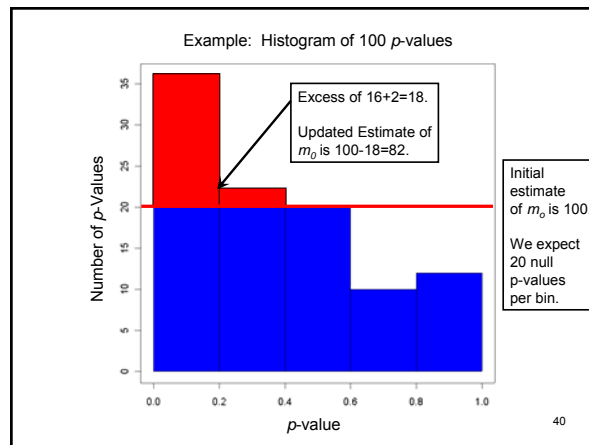
- Lowest Slope: Schweder and Spjøtvoll (1982), Hochberg and Benjamini (1990), Benjamini and Hochberg (2000).
- Mixture of Uniform and Beta: Allison et al. (2002), Pounds and Morris (2003).
- λ Threshold: Storey (2002), Storey and Tibshirani (2003)
- Density Estimation: Langaas, Ferkingstad, Lindqvist (2004)
- Histogram Based Estimation: Mosig et al. (2001), Nettleton, Hwang, Caldo, Wise (2006)

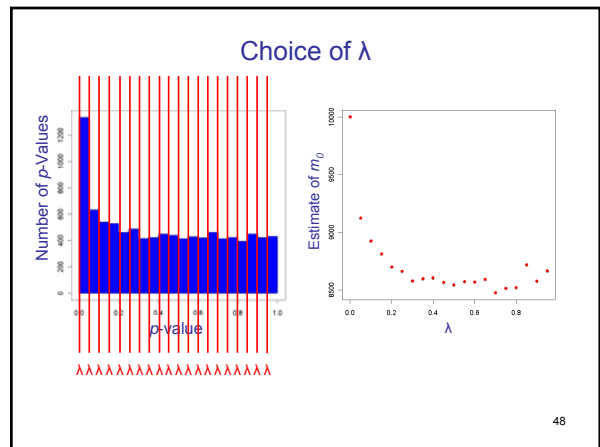
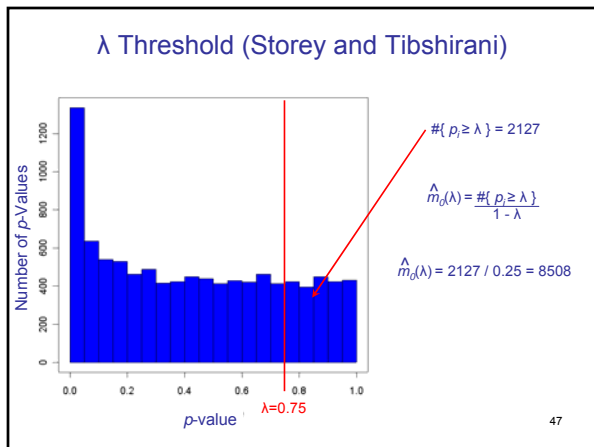
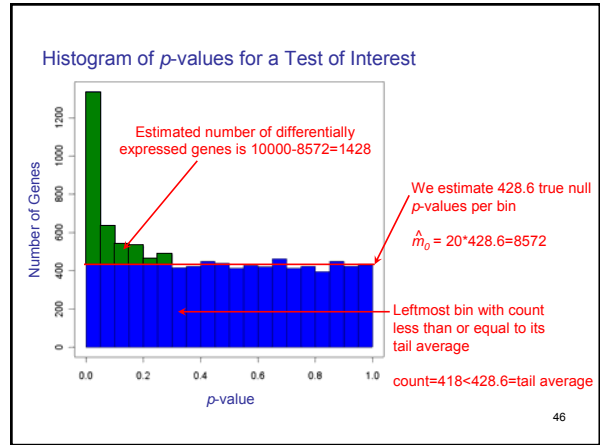
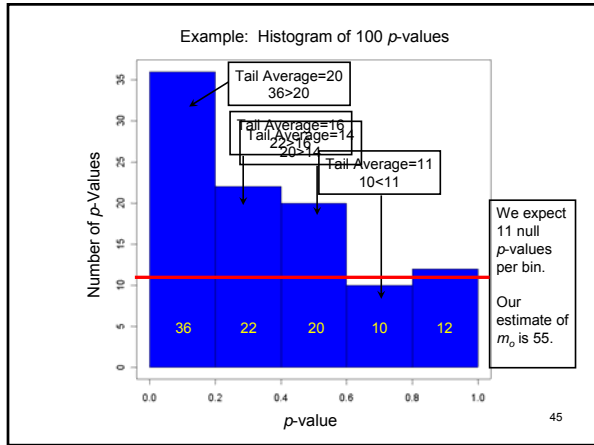
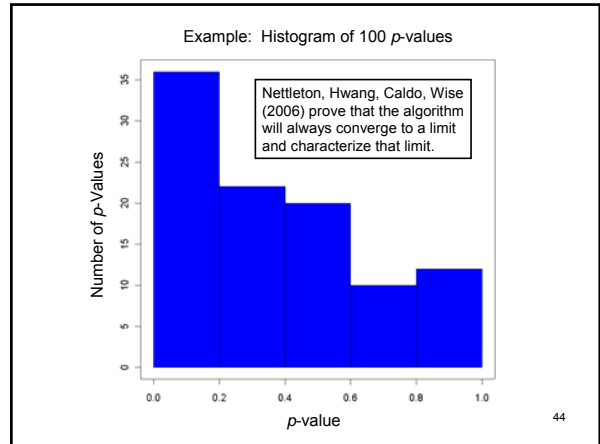
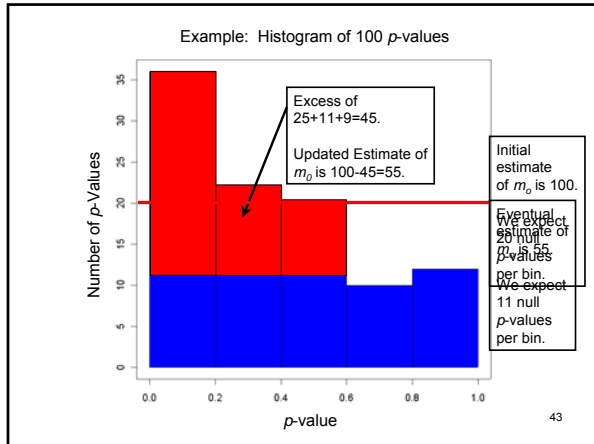
38

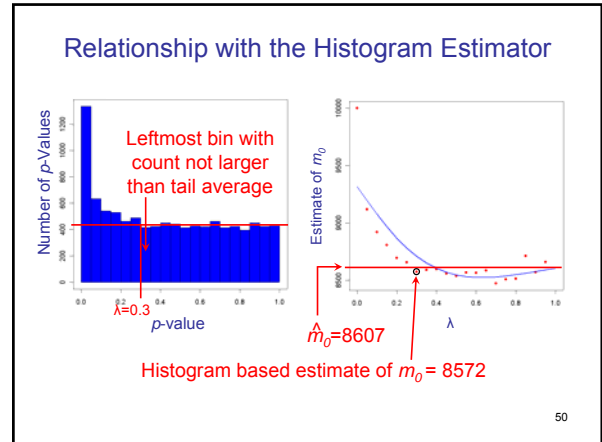
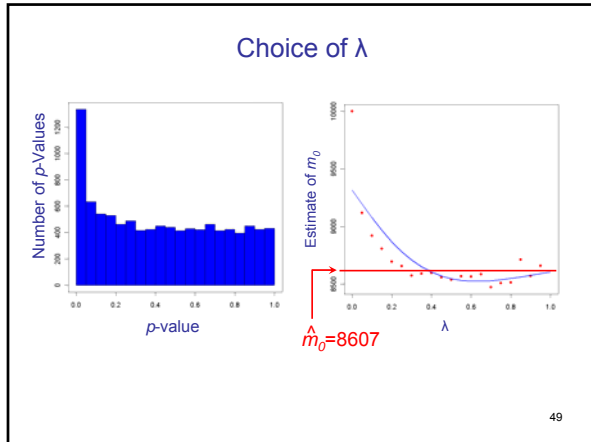
Histogram Based Estimation of m_0

Mosig, M. O., Lipkin, E., Galina, K. Tchourzyna, E., Soller, M., and Friedmann, A. (2001). A whole genome scan for quantitative trait loci affecting milk protein percentage in Israeli-Holstein cattle, by means of selective milk DNA pooling in a daughter design, using an adjusted false discovery rate criterion. *Genetics*, **157**, 1683-1698.

39







Estimating FDR Using Estimates of m_0

- 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.
- Mosig, M. O., Lipkin, E., Galina, K. Tchourzyna, E., Soller, M., and Friedmann, A. (2001). A whole genome scan for quantitative trait loci affecting milk protein percentage in Israeli-Holstein cattle, by means of selective milk DNA pooling in a daughter design, using an adjusted false discovery rate criterion. *Genetics*, **157**, 1683-1698.
- Storey, J. D., and Tibshirani, R. (2001). Estimating false discovery rates under dependence, with applications to DNA microarrays. Technical Report 2001-28, Department of Statistics, Stanford University.
- Storey J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B*, **64**, 479-498.

51

Estimating FDR Using Estimates of m_0

- Genovese, C. and Wasserman, L. (2002). Operating characteristics and extensions of the false discovery rate procedure, *Journal of the Royal Statistical Society, Series B*, **64**, 499-517.
- Storey J. D. (2003). The positive false discovery rate: A Bayesian interpretation and the q-value. *Annals of Statistics*, **31**, 2013-2035.
- Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences* **100**, 9440-9445

52

Estimating FDR Using Estimates of m_0

- Storey J. D., Taylor JE, and 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.
- Fernando, R. L., Nettleton, D., Southey, B. R., Dekkers, J. C. M., Rothschild, M. F., and Soller, M. (2004). Controlling the proportion of false positives (PFP) in multiple dependent tests. *Genetics*. **166**, 611-619.
- Genovese, C. and Wasserman, L. (2004). A stochastic process approach to false discovery control. *The Annals of Statistics*, **32**, 1035-1061.

53

A method for obtaining a list of genes that has an estimated FDR $\leq \alpha$

- Find the largest integer k such that

$$p_{(k)} \hat{m}_0 / k \leq \alpha,$$
 where \hat{m}_0 is an estimate of the number of true null hypotheses among the m tests.
- If no such k exists, declare nothing significant. Otherwise, reject the null hypotheses corresponding to the smallest k p -values.

54

q-values

- Recall that a p -value for an individual test can be defined as the smallest significance level (tolerable type 1 error rate) for which we can reject the null hypothesis.
- The q -value for one test in a family of tests is the smallest FDR for which we can reject the null hypothesis for that one test and all others with smaller p -values.

55

The q -value for a given test fills the blanks in the following sentences:

- "If I set my cutoff for significance c equal to this p -value, I must be willing to accept a false discovery rate of _____."
- "To reject the null hypothesis for this test and all others with smaller p -values, I must be willing to accept a false discovery rate of _____."
- "To include this gene on my list of differentially expressed genes, I must be willing to accept a false discovery rate of _____."

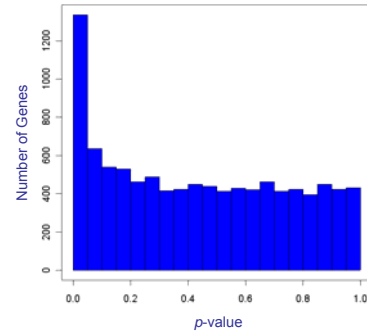
56

Computation and Use of q -values

- Let $q_{(i)}$ denote the q -value that corresponds to the i^{th} smallest p -value $p_{(i)}$.
- $q_{(i)} = \min \{ p_{(k)} \hat{m}_0 / k : k = i, \dots, m \}$.
- To produce a list of genes with estimated $\text{FDR} \leq \alpha$, include all genes with q -values $\leq \alpha$.

57

We will convert these p -values to q -values using the method of Storey and Tibshirani.



58

p -values q -values

1.	0.0000011114	0.009565612
2.	0.0000208581	0.056578703
3.	0.0000252334	0.056578703
4.	0.0000283551	0.056578703
5.	0.0000328686	0.056578703
6.	0.0000427945	0.057791900
7.	0.0000481212	0.057791900
8.	0.0000594893	0.057791900
9.	0.0000642639	0.057791900
10.	0.0000700471	0.057791900
11.	0.0000738614	0.057791900
12.	0.0000920778	0.059338371
13.	0.0000990813	0.059338371
14.	0.0001113717	0.059338371
15.	0.0001125189	0.059338371
16.	0.0001155592	0.059338371
17.	0.0001172041	0.059338371
18.	0.0001338667	0.060982276
19.	0.0001347057	0.060982276
20.	0.0001556119	0.060982276

If we want FDR to be 1%, we can declare only one gene differentially expressed.

59

p -values q -values

1.	0.0000011114	0.009565612
2.	0.0000208581	0.056578703
3.	0.0000252334	0.056578703
4.	0.0000283551	0.056578703
5.	0.0000328686	0.056578703
6.	0.0000427945	0.057791900
7.	0.0000481212	0.057791900
8.	0.0000594893	0.057791900
9.	0.0000642639	0.057791900
10.	0.0000700471	0.057791900
11.	0.0000738614	0.057791900
12.	0.0000920778	0.059338371
13.	0.0000990813	0.059338371
14.	0.0001113717	0.059338371
15.	0.0001125189	0.059338371
16.	0.0001155592	0.059338371
17.	0.0001172041	0.059338371
18.	0.0001338667	0.060982276
19.	0.0001347057	0.060982276
20.	0.0001556119	0.060982276

If we want FDR to be 5%, we can declare only one gene differentially expressed.

60

	<i>p</i> -values	<i>q</i> -values	
1.	0.0000011114	0.009565612	
2.	0.0000208581	0.056578703	
3.	0.0000252334	0.056578703	
4.	0.0000283551	0.056578703	
5.	0.0000328686	0.056578703	
6.	0.0000427945	0.057791900	
7.	0.0000481212	0.057791900	
8.	0.0000594893	0.057791900	
9.	0.0000642639	0.057791900	
10.	0.0000700471	0.057791900	
11.	0.0000738614	0.057791900	
12.	0.0000920778	0.059338371	
13.	0.0000990813	0.059338371	
14.	0.0001113717	0.059338371	
15.	0.0001125189	0.059338371	
16.	0.0001155592	0.059338371	
17.	0.0001172041	0.059338371	
18.	0.0001338667	0.060982276	
19.	0.0001347057	0.060982276	
20.	0.0001556119	0.060982276	

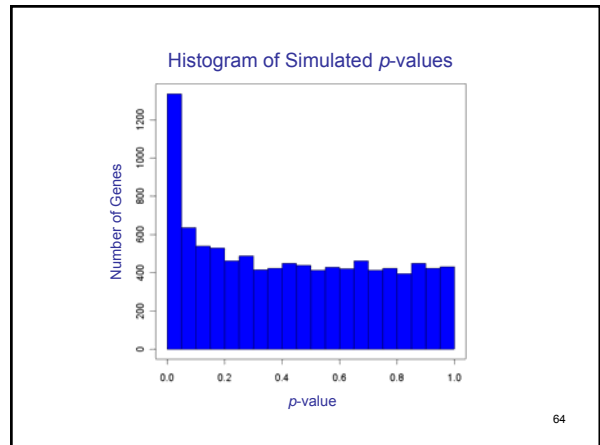
If we want FDR to be 6%, we can declare 17 genes differentially expressed.

Some Established Properties of FDR Procedures

- Benjamini and Hochberg (1995) shows that their FDR procedure will strongly control FDR below a specified level α whenever the tests corresponding to true null hypotheses are independent.
- Benjamini and Yekutieli (2001) show that the B&H FDR procedure will strongly control FDR below a specified level α whenever the tests corresponding to true null hypotheses satisfy a *positive regression dependency* property.
- In both cases, if the B&H FDR procedure is set up for control at level α , FDR will be controlled at level $\left(\frac{m_0}{m}\right)\alpha$.

Some Established Properties of FDR Procedures (continued)

- Storey and co-authors have proved several properties related to their procedures for controlling FDR. The strongest results are found in Storey, Taylor, Siegmund (2004). A relatively non-technical summary of some of the basic results is as follows.
- If $\hat{m}_0(\lambda)$ is used to estimate m_0 for any fixed λ in $(0, 1)$, then using *q*-values to generate lists of significant results will strongly control FDR
 - when tests corresponding to true null hypotheses are independent, or
 - when tests are *weakly dependent* and $m \rightarrow \infty$.



Control of FWER at 0.05 Using the Bonferroni Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8500	0	8500
False Nulls	1499	1	1500
Total	9999	1	10000

Control of FWER at 0.05 Using Holm's Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8500	0	8500
False Nulls	1499	1	1500
Total	9999	1	10000

Control of FDR at 0.05
Using the B&H Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8500	0	8500
False Nulls	1499	1	1500
Total	9999	1	10000

67

Control of FDR at 0.05
Using the Storey and Tibshirani Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8500	0	8500
False Nulls	1499	1	1500
Total	9999	1	10000

68

Control of FWER at 0.10
Using the Bonferroni Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8500	0	8500
False Nulls	1499	1	1500
Total	9999	1	10000

69

Control of FWER at 0.10
Using Holm's Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8500	0	8500
False Nulls	1499	1	1500
Total	9999	1	10000

70

Control of FDR at 0.10
Using the B&H Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8497	3	8500
False Nulls	1463	37	1500
Total	9960	40	10000

71

Control of FDR at 0.10
Using the Storey and Tibshirani Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8487	13	8500
False Nulls	1368	132	1500
Total	9855	145	10000

72

Control of FWER at 0.20
Using the Bonferroni Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8500	0	8500
False Nulls	1499	1	1500
Total	9999	1	10000

73

Control of FWER at 0.20
Using Holm's Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8500	0	8500
False Nulls	1499	1	1500
Total	9999	1	10000

74

Control of FDR at 0.20
Using the B&H Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8402	98	8500
False Nulls	1044	456	1500
Total	9446	554	10000

75

Control of FDR at 0.20
Using the Storey and Tibshirani Method

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	8370	130	8500
False Nulls	947	553	1500
Total	9317	683	10000

76

Note that all our methods missed many differentially expressed genes

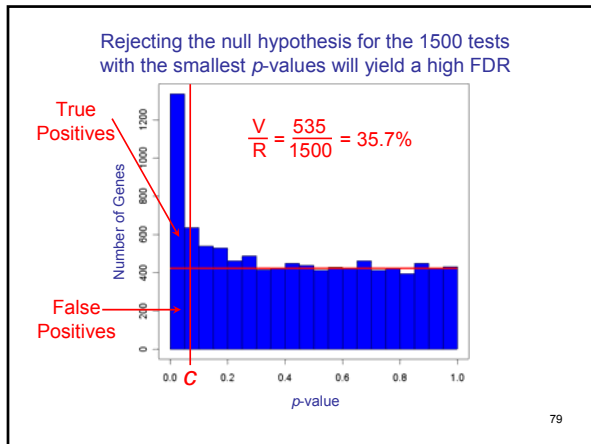
- Our methods of estimating m_0 suggest that around 1400 or 1500 genes are differentially expressed.
- In this case 1500 genes are truly differentially expressed.
- What if we declare the 1500 genes with the smallest p -values to be differentially expressed?

77

Rejecting the Null Hypotheses for the Tests
with the 1500 Smallest p -values

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	7965	535	8500
False Nulls	535	965	1500
Total	8500	1500	10000

78



Can we get a list of genes that contains all 1500 of the truly differentially expressed genes?

- It turns out that we would have to include 9917 genes on our list to get all 1500 truly differentially expressed genes in this case.
- The list would have all the truly differentially expressed genes, but about 85% of the genes on the list would be false positives.

80

Performance of the Shortest List that Contains All Truly Differentially Expressed Genes

	Accept Null Declare Non-Sig. No Discovery Negative Result	Reject Null Declare Sig. Declare Discovery Positive Result	
True Nulls	83	8417	8500
False Nulls	0	1500	1500
Total	83	9917	10000

81

Concluding Remarks

- In many cases, it will be difficult to separate the many of the differentially expressed genes from the non-differentially expressed genes.
- Genes with a small expression change relative to their variation will have a p -value distribution that is not far from uniform if the number of experimental units per treatment is low.
- To do a better job of separating the differentially expressed genes from the non-differentially expressed genes we need to use good experimental designs with more replications per treatment.

82

Concluding Remarks (continued)

- We have looked at only one simulated example.
- The behavior of the methods will vary from data set to data set.
- Simulations suggest that the methods control their error rates at nominal levels for a variety of practical situations.

83