

Step-down FDR Procedures for Very Large Numbers of Hypotheses

Paul N. Somerville

University of Central Florida

OBJECTIVE:

Describe **FDR test procedures** for simultaneously testing a large number of hypotheses which:

1. are powerful
2. control the expected proportion of "false discoveries"
3. control probabilities of $\leq \mathbf{u}$ "false discoveries"
4. control probabilities of proportion of $\leq \mathbf{x}$ "false discoveries"
5. are easily implemented
6. there are tables available

METHODS FOR CONTROLLING ERRORS

H_1, H_2, \dots, H_m (r, q)

- i. **PCER** (**Per Comparison Error Rate**) $\leq r$
 r is the probability of **Type I Error** for **each H_i**
Many $H_i \rightarrow$ many Type I errors

- ii. **FWER** (**Family Wise Error Rate**) $\leq r$
 r probability at least one Type I Error in family of H_i 's
Step-down or step-up (**SD** or **SU**)

- iii. **FDR** (**False Discovery Rate**) $\leq q$
Critical values $d_1 \leq d_2 \leq \dots \leq d_m$
Step-down or step-up (**SD** or **SU**)

iii. **FDR (False Discovery Rate)**

R Total number of hypotheses rejected

V Number of TRUE hypotheses rejected

Q = V/R (If $R = 0$, $Q = 0$)

FDR = E(Q)

Choose critical values such that **E(Q) ≤ q**

$$d_1 \leq d_2 \leq \dots d_m$$

Benjamini and Hochberg (1995) Step Up FDR Procedure (Indep. hypotheses)

Benjamini and Liu (1999) Step Down FDR Procedure (Indep. hypotheses)

Troendle (2000) SD &SU FDR (Multivariate-t distn. of test statistics)

Benjamini and Yekutieli (2001) (Distribution free hypotheses)

Somerville (2003, 2004) SD &SU FDR (Multivariate-t distn. of test statistics common ρ)

Horne and Dunnett (2004) Power studies of FWER and FDR procedures

SOME CRITICISMS of FDR PROCEDURES

FDR IS AN EXPECTED VALUE (CONTROLLED)

but

Number of False Discoveries could be large

Proportion of false discoveries could be large

Critical values for the "most powerful" procedures are
"impractical" to calculate for very large m .

Power studies:

Most powerful **FDR** procedures are methods of Troendle (2000) and Somerville (2004)

ALTERNATE PROCEDURES:

i) KORN, TROENDLE ET AL (2004)

PROCEDURE A

Controls actual number of false discoveries

PROCEDURE B

Controls proportion of false discoveries

PROBLEM: Very computer intensive

ii) van der Laan, Dudoit and Pollard (2004)

1. generalized FWER gFWER

Controls u , the number of Type I errors in family

2. PFP

Controls γ the proportion of false positives " "

SOMERVILLE (2004) (Uses Minimum Critical Value)

Some FDR critical values can be negative (even $-\infty$)

Remedy: Select a Minimum Critical Value (MCV)

Cannot reject H_i unless test statistic or unadjusted p-value is less than some value - say 2.00 (test statistic) or .05 (p-val).

Using an MCV, calculate all critical values:

$$d_1 = \dots = d_{m-j} = \text{MCV} \leq d_{m-j+1} \leq \dots \leq d_m.$$

Note: large MCV means few cv s need to be calculated.

i.e. few "unique" critical values.

Empirical result: For "optimum" powers, MCV is inversely related to n_F where n_F is the number of false hypotheses. For **small** n_F "best" is **large** MCV. This is especially pertinent if n_F is small or we are satisfied with a few "discoveries".

"OPTIMUM" MCV IS ONE WHICH RESULTS IN APPROX. n_F "unique" critical values.

NOTE: We usually don't know n_F .

<u>m</u>	<u>50</u>	<u>100</u>	<u>250</u>	<u>500</u>	<u>1000</u>	<u>2500</u>	<u>5000</u>	<u>10000</u>
d_m	3.083	3.283	3.533	3.713	3.884	4.102	4.259	4.412
d_{m-1}	2.867	3.081	3.346	3.537	3.714	3.938	4.103	4.266
d_{m-2}	2.731	2.958	3.234	3.426	3.613	3.842	4.009	4.172
d_{m-3}	2.629	2.865	3.150	3.349	3.538	3.772	3.940	4.106
...
d_{m-29}	1.549	2.066	2.490	2.746	2.969	3.249	3.443	3.628
d_1-d_{m-30}	1.397	1.991	2.437	2.700	2.935	3.212	3.408	3.594

Table A.1

Step-down FDR Critical Values for $v = \infty$, $q = .05$, $\rho = 0$ (31 "unique")

(Tables for $v = \infty$, $q = .05$, $\rho = 0, .1, .5$ $v = 15$, $q = .05$, $\rho = 0, .1$)

INTERPOLATION: linear in $\ln(m)$; linear in $1/v$; quadratic in ρ .

EXAMPLE

Korn, Troendle et al (2000)

Dataset: gene expression for 8029 genes before and after chemotherapy on 20 breast cancer patients. Test null hypotheses that mean pre and post chemotherapy expression of genes was the same for each gene.

PROCEDURE A:

Identified 28 genes. $P[u \leq 2]$ with confidence .95.

PROCEDURE B:

Identified 28 genes. $P[\gamma \leq .10]$ with approx. confidence .95.

	$v = \infty$			$v = 15$	
<u>MCV</u>	<u>$\rho=.0$</u>	<u>$\rho=.1$</u>	<u>$\rho=.5$</u>	<u>$\rho=.0$</u>	<u>$\rho=.1$</u>
d_{m-7}	20	21	29	24	26
d_{m-11}	23	23	33	27	29
d_{m-15}	24	24	35	29	31
d_{m-19}	27	27	38	31	31
d_{m-23}	28	29	40	33	35
d_{m-27}	29	29	>50	33	36
d_{m-30}	29	33	>50	36	41

Number of genes identified by procedures ($q = .05$)

<u>MCV</u>	$v = \infty$			$v = 15$	
	<u>$\rho=.0$</u>	<u>$\rho=.1$</u>	<u>$\rho=.5$</u>	<u>$\rho=.0$</u>	<u>$\rho=.1$</u>
d_{m-7}	.99 (40)	.96 (70)	.95 (50)	.90 (50)	.94 (80)
d_{m-11}	.98 (50)	.94 (70)	.93 (55)	.90 (50)	.93 (90)
d_{m-15}	.96 (70)	.91 (70)	.92 (55)	.89 (80)	.91 (100)
d_{m-19}	.92 (80)	.88 (70)	.91 (60)	.89 (80)	.91 (100)
d_{m-23}	.88 (80)	.85 (70)	.91 (60)	.88 (90)	.89 (300)
d_{m-27}	.83 (80)	.82 (80)	.90 (60)	.88 (90)	.88 (300)
d_{m-30}	.78 (90)	.78 (90)	.87 (70)	.88 (100)	.85 (300)

Minimum values of $P[u \leq 2]$ ($m=8029$, $q=.05$, $U=3.464$)

Value of n_F at minimum in parentheses

Δ is the mean value of the t-statistic for the FALSE H.

<u>MCV</u>	$v = \infty$			$v = 15$	
	<u>$\rho=.0$</u>	<u>$\rho=.1$</u>	<u>$\rho=.5$</u>	<u>$\rho=.0$</u>	<u>$\rho=.1$</u>
d_{m-7}	.78 (19)	.82 (14)	.89 (8)	.94 (80)	.91 (30)
d_{m-11}	.77 (19)	.82 (15)	.90 (10)	.93 (100)	.90 (40)
d_{m-15}	.77 (19)	.82 (18)	.90 (16)	.91 (200)	.90 (50)
d_{m-19}	.77 (19)	.82 (18)	.90 (17)	.89 (300)	.90 (50)
d_{m-23}	.77 (19)	.82 (18)	.90 (17)	.88 (300)	.89 (50)
d_{m-27}	.77 (19)	.82 (25)	.90 (17)	.88 (300)	.89 (70)
d_{m-30}	.77 (19)	.82 (25)	.90 (17)	.83 (300)	.88 (80)

Minimum values of $P[x \leq .10]$ ($m=8029$, $q=.05$, $U=3.464$)

Value of n_F at minimum in parentheses

Δ is the mean value of the t-statistic for the FALSE H.

Calculation of Critical Values and Powers

Fortran 90 programs **SEQDN** and **SEQUP** sequentially calculate the critical values d_2 to d_m for step-down and step-up FDR for arbitrary values of m , q , ρ , and v . $N(10^5, 10^6$ or $10^7)$ random normal multivariate vectors of size m are used. Very computer intensive especially for large m .

Fortran 90 **FDRPWRDN** and **FDRPWRUP** calculate powers, $P[u \leq 1, 2, \dots, 7]$ and $P[\gamma \leq .05, .10, .15]$. Inputs are m , q , ρ , Δ , v , n_F and a set of m critical values. N random normal multivariate vectors are used. Per pair, all pairs and any pairs powers are always calculated.

SUMMARY AND CONCLUSIONS

- i) If common ρ is underestimated, # of "discoveries" decreases, i.e. a loss of power results. Also fewer "false discoveries".
- ii) If n_F (# of false hypotheses) is **known** "optimum" MCV is inversely related to n_F . (for **small** n_F use **large** MCV).
- iii) As MCV decreases, (n_F **unknown**), power increases **and** number of "false discoveries" increases.
- iv) If n_F is small, or a few "discoveries" is acceptable, one can choose MCV sufficiently large such that $P[u \leq 2] = .95$. Preliminary studies indicate that for all but the largest values of MCV that choice of MCV has little effect on $P[\gamma \leq .10]$
- iv) Tables are available (including interpolation guides) for $50 \leq m \leq 10,000$ or somewhat larger, $v \geq 15$, and moderate ρ for step-down FDR.