A Package to Study the Performance of Step-Up and Step-Down Test Procedures

1-1-2012

Paul N. Somerville

University of Central Florida

Find similar works at: https://stars.library.ucf.edu/facultybib2010

University of Central Florida Libraries http://library.ucf.edu

Recommended Citation

Somerville, Paul N., "A Package to Study the Performance of Step-Up and Step-Down Test Procedures" (2012). Faculty Bibliography 2010s. 3340. https://stars.library.ucf.edu/facultybib2010/3340

This Article is brought to you for free and open access by the Faculty Bibliography at STARS. It has been accepted for inclusion in Faculty Bibliography 2010s by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.
A Package to Study the Performance of Step-Up and Step-Down Test Procedures

Paul N. Somerville
University of Central Florida

Abstract

The package can be used to analyze the performance of step-up and step-down procedures. It can be used to compare powers, calculate the "false discovery rate", to study the effects of reduced step procedures, and to calculate $P[U \leq k]$, where $U$ is the number of rejected true hypotheses. It can be used to determine the maximum number of steps that can be made and still guarantee (with a given probability) that the number of false rejections will not exceed some specified number. The test statistics are assumed to have a multivariate-$t$ distribution. Examples are included.

Keywords: step-up and step-down procedures, performance analysis, reduced step procedures, maximum steps to guarantee true rejection probabilities.

1. Introduction

There are many situations where a researcher is interested in the outcome of a family of tests. A classical example is the case where $m$ experimental drugs are compared with a standard with respect to an outcome. Analysts are increasingly confronting large-mining environments, due to advances in computing facilities and data collection strategies. Multiple testing has increased in importance and micro-array experiments are common in such diverse areas as neuro-imaging, genomics and astronomy.

In multiple testing, it is important to control the probability of rejecting true hypotheses. A standard procedure has been to control the family-wise error rate (FWER), the probability of rejecting at least one true null hypotheses. For large numbers of hypotheses, using FWER can result in very low power for testing single hypotheses. Recently powerful multiple step methods have been developed for controlling the false discovery rate (FDR), that is, the expected proportion of type I errors. More recently Van der Laan, Dudoit, and Pollard (2004) proposed controlling a generalized family error rate $k$-FWER (also called $g$FWER($k$)).
2. Step-down and step-up procedures

Step-down test procedures may be described as follows. Let \( t_1, t_2, \ldots, t_m \) be the test statistics corresponding to the null hypotheses \( H_1, H_2, \ldots, H_m \). Denote by \( T_i \) the random variable associated with \( t_i \).

Let \( T_{(1)} \leq T_{(2)} \leq \cdots \leq T_{(m)} \) be the ordered values for the test statistics and denote the corresponding hypotheses \( H_{(1)}, H_{(2)}, \ldots, H_{(m)} \). If \( T_{(m)} \geq d_m \), reject \( H_{(m)} \) and continue for \( i = m - 1, m - 2, \ldots, \) comparing \( T_{(i)} \) with \( d_i \), rejecting \( H_{(i)} \) if \( T_{(i)} \geq d_i \) and continuing until for the first time \( T_{(i)} < d_i \). If \( T_{(m)} < d_m \), reject no hypotheses. If \( T_{(i)} \geq d_i \), for all values of \( i \), reject all the hypotheses. The values \( d_1 \leq d_2 < \cdots \leq d_m \) are constants (called critical values).

For step-up procedures \( T_{(i)} \) is compared with \( d_i \), beginning with \( i = 1 \), continuing for \( i = 2, 3, \ldots \) until, for the first time, \( T_{(i)} > d_i \). \( H_{(i)}, H_{(i+1)}, \ldots, H_{(m)} \) are then rejected.

3. Reduced step procedures

An \( s \) step, step-down procedure, is defined as follows: Compare \( T_{(i)} \) with \( d_i \) for \( i = m, m - 1, \ldots, m - s + 1 \) until for the first time \( T_{(i)} < d_i \), in which case reject \( H_{(m)}, \ldots, H_{(i+1)} \). If \( T_{(m)} < d_m \), reject no hypotheses. If \( i = m - s + 1 \), reject all hypotheses for which \( T_i \geq d_i \).

For the \( s \) step, step-up procedure, begin with the comparison of \( T_{(m-s+1)} \) with \( d_{m-s+1} \). If \( T_{(m-s+1)} \geq d_{m-s+1} \), reject all hypotheses \( H_i \) for which \( T_{(i)} \geq d_{m-s+1} \), and otherwise compare \( T_{(m-s+2)} \) with \( d_{m-s+2} \), \( T_{(m-s+3)} \) with \( d_{m-s+3} \), \ldots , until for the first time \( T_{(i)} \geq d_i \). \( H_{(i)}, H_{(i+1)}, \ldots, H_{(m)} \) are then rejected.

An \( s \) step procedure is equivalent to an \( m \) step procedure where the \( m - s \) smallest values are replaced with the value \( d_{m-s+1} \). The critical value \( d_{m-s+1} \) is the minimum critical value (MCV). The concept of an MCV was introduced by Somerville (2004).

4. Defining the problem

Assume there are \( m \) multivariately distributed test statistics. The \( n_T \) test statistics corresponding to the true hypotheses have zero means, and the \( n_F \) test statistics corresponding to the false hypotheses have means equal to \( \delta \). The test statistics corresponding to the true and false hypotheses have common correlations \( \varrho_T \) and \( \varrho_F \) respectively, with the common correlation between the true and false test statistics being \( \varrho_C \). Using Monte Carlo methods the program \( \text{FDR} \) generates \( n \) \( m \)-dimensional random vectors from the multivariate distribution, and for each uses the step-up or step-down procedure to determine the number of true and false hypotheses rejected using the designated test procedure. Powers, probabilities, etc., are thus estimated.
The following parameters are needed as input:

\( m \): Number of null hypotheses.

\( n_T \): Number of true hypotheses.

\( n_F \): Number of false hypotheses.

\( \delta \): Standardized mean of test statistics corresponding to false hypotheses.

\( df \): Degrees of freedom for the multivariate distribution.

\( d \): Vector of the \( m \) critical values.

\( \rho_T, \rho_F, \rho_C \): Correlation among true, among false and between true and false test statistics, respectively.

\( n_{F\text{beg}}, n_{F\text{end}}, n_{F\text{int}} \): For mode 1, for a series of \( n_F \) value, first, last and interval between.

\( n_{S\text{beg}}, n_{S\text{end}}, n_{S\text{int}} \): For mode 2, for a series of \# of steps, first, last and interval between.

\( k_{\text{beg}}, k_{\text{end}}, k_{\text{int}} \): For tables of \( P[U \leq k] \geq \text{GAMMA} \) first \( k \), last \( k \) and interval between \( k \) values.

\text{GAMMA}: Given by the user. Requirement is that \( P[U \leq k] \geq \text{GAMMA} \).

5. Using the package

The package is designed to study the performance of arbitrary step-up and step-down procedures. There are three modes for the Fortran program FDR. Mode 1 was designed to calculate three kinds of power (per pair, all pairs and any pair). See Horn and Dunnett (2004). In addition FDR is calculated. Using mode 1, the user specifies the number of steps to be used, and one or more series of ranges for \( n_F \). The output file PWR.out contains, for all the values for \( n_F \) specified, the three kinds of power, and the false discovery rate (FDR).

Somerville and Hemmelmann (2008) showed that, under fairly general conditions, when the means of test statistics corresponding to false hypotheses increase without limit, \( P[U \leq k] \) is minimized. In Appendix A, it is shown that \( P[U \leq k] \), as a function of \( n_F \), is minimized when \( n_F \) is less than or equal to the number of steps minus 1. Extensive calculations, using a Fortran program BHmax with extended precision, strongly suggest, at least for the Benjamini and Hochberg (1995, BH for short) procedure, and most practical cases, that the minimum occurs when \( n_F \) is exactly equal to the number of steps minus \((k + 1)\). Mode 1 can be used to verify this for particular sets of values of \( m, n_F, d \), etc.

While mode 1 uses a fixed value for the number of steps, and one or more series of ranges of values for \( n_F \), mode 2 uses \( n_F \) equal to the number of steps minus \((k + 1)\), and one or more series of ranges for the number of steps.

A file CV.in must exist before executing FDR. The file must contain in non-decreasing order of magnitude, the \( m \) critical values for the procedure. A Fortran program makeCV, which calculates the critical values for the BH procedure, and for the step-down procedure of Lehmann
and Romano (2005) is given in the supplementary materials. The user should be able to follow the instructions given in FDR.

The Fortran program FDR has been compiled using Lahey/Fujitsu Fortran 95 on Microsoft platforms XP, Vista and Windows 7, and using g95 on Microsoft platform Windows 7.

6. Selecting the value of kSTAR

If the object is to calculate the maximum number of steps such that \( P[U \leq k] \geq \text{GAMMA} \), then the proper value for kSTAR is \( k \). However the user may be interested in several values of \( k \) for several values of \( \text{GAMMA} \). The minimum value of \( s \) is usually overestimated if kSTAR is chosen otherwise. The error is less serious if a lesser value is chosen for kSTAR. Especially for large values of \( m \), kSTAR may be chosen slightly smaller with minimal error. Choosing kSTAR larger is usually more serious, with the error increasing much more rapidly than the choice of choosing the value slightly smaller. Example 3 in Appendix C gives a more detailed analysis.

7. Accuracy of estimated probabilities

The number of random \( m \)-dimensional vectors generated is given by \( n \), a user input. Values of \( n \) of 100, 1000, 10000 and 100000 for the estimation of a probability of 0.95 result in standard errors of 0.0218, 0.0069, 0.0022 and 0.0007 respectively.

8. Miscellaneous

The following procedure is suggested when both one-sided and two-sided hypotheses are included. Calculate the \( p \) values corresponding to each test statistic, and then use the test statistic value which corresponds to that \( p \) value as a one-sided hypotheses.

Clearly the more steps that are used in a procedure, the more powerful the procedure will be. Mode 2 was added to the predecessor package to assist in determining the maximum number of steps which will still guarantee \( P[U \leq k] \geq \text{GAMMA} \).

The value of \( n \), the number of generations of \( m \)-dimensional random vectors from a multivariate-\( t \) distribution, determines the accuracy of the estimates and the results. A useful technique is to begin a study with small values of \( n \), but very large numbers of steps or numbers of false hypotheses. These preliminary runs can determine the values for the number of steps or false hypotheses in future runs where increased accuracies are needed.

The package has been used with the number of hypotheses as large as 1,000,000. With such large values for \( m \), the value of \( n \), the number of random generations of the \( m \)-dimensional vectors must be reduced so as to result in reasonable running times.

The original version of the program FDR was used to produce tables for two papers published in “Recent Developments in Multiple Comparison Procedures”, Institute of Mathematical Statistics, one by Somerville (2004) and one by Horn and Dunnett (2004).
9. Summary and conclusions

The Fortran program FDR can be a very useful tool in the evaluation of step-up or step-down procedures. It can be used to compare competing procedures.

It is worth noting that, since single step procedures are special cases of multi-step procedures, all of the capabilities of the program apply to single step procedures. See Example 5 in Appendix C.

Acknowledgments

The author wishes to thank the referees, the associate editor and an editor for their helpful constructive criticisms which have materially strengthened the paper.

References

Benjamini Y, Hochberg Y (1995). “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.” Journal of the Royal Statistical Society B, 57(1), 289–300.

Horn M, Dunnett CW (2004). “Power and Sample Size Comparisons of Stepwise FWE and FDR Controlling Test Procedures in the Many-One Case.” In Y Benjamini, F Bretz, S Sarkar (eds.), Recent Developments in Multiple Comparison Procedures, volume 47 of IMS Lecture Notes – Monograph Series, pp. 48–64.

Lehmann EL, Romano JP (2005). “Generalizations of the Familywise Error Rate.” The Annals of Statistics, 33(3), 1138–1154.

Somerville PN (2004). “FDR Step-Down and Step-Up Procedures for the Correlated Case.” In Y Benjamini, F Bretz, S Sarkar (eds.), Recent Developments in Multiple Comparison Procedures, volume 47 of IMS Lecture Notes – Monograph Series, pp. 100–118.

Somerville PN, Hemmelmann C (2008). “FDR Procedures Controlling the Number and Proportion of False Positives.” Computational Statistics & Data Analysis, 52, 1323–1334.

Van der Laan MJ, Dudoit S, Pollard KS (2004). “Augmentation Procedures for Control of the Generalized Family-Wise Error Rate and Tail Probabilities for the Proportion of False Positives.” Statistical Applications in Genetics and Molecular Biology, 3, 1–25.
A. Corollary to Theorem 6.1, Somerville and Hemmelman

Somerville and Hemmelmann (2008) proved the following:

**Theorem:** Let \( d_1 \leq d_2 \leq \cdots \leq d_m \) be the critical values corresponding to a step-up or step-down procedure used to simultaneously test \( m \) hypotheses. Let the random variable \( T_i \), associated with the test statistic \( t_i \) be absolutely continuous or discrete, have mean \( \mu_i \), and a distribution such that

\[
P[T_i \geq a|\mu_i = \mu] = P[T_i \geq (a + \delta)\mu_i = (\mu + \delta)],
\]

where \( a, \mu \) and \( \delta \) are arbitrary constants.

Let \( m = n_T + n_F \), where \( n_T \) and \( n_F \), respectively are the number of true and false hypotheses. Suppose \( T_1, T_2, \ldots, T_m \) have means \( \mu_1, \mu_2, \ldots, \mu_{n_F}, 0, \ldots, 0 \), respectively. If \( U \) is the number of true hypotheses which are rejected, then, for the 1-sided case, \( P[U \leq k] \) is minimized for \( k \in [0, n_F] \), when \( \mu_i \to \infty \), for \( i = 1, \ldots, n_F \).

We show that the following corollary is true.

**Corollary:** Let \( P^*(n_F, k, m) \) be the minimum value of \( P[U \leq k] \) when there are \( m \) hypotheses, of which \( n_F \) are false. Under the conditions of the above theorem, the minimum value of \( P^* \), as a function of \( n_F \), occurs for a value of \( n_F \) which is never greater than \( s - 1 \).

**Proof:** Assume \( \mu_i \to \infty \), for \( i = 1, \ldots, n_F \). Then \( T_{(m)}, \ldots, T_{(m-n_F+1)} \to \infty \). If \( n_F \geq s \), each of the \( s \) steps of an \( s \)-step procedure will result in the rejection of a false hypothesis and no true hypotheses can be rejected, i.e., \( P[U \leq k] = 1 \) for all possible \( k \) values. Thus a minimum value of \( P^*(n_F, k, m) \), if one exists, cannot occur for a value of \( n_F \) greater than \( s - 1 \).

**Comments:** The program FDR (mode 2) uses the value of \( n_F \) as \( s - k - 1 \) in any calculations requiring \( P^*(n_F, k, m) \). The program can also be used (mode 1) to check that \( n_F = s - k - 1 \) gives the smallest value of \( P[U \leq k] \) for arbitrary \( k, m, s \) and critical values. (See example 2 in Appendix C.)

Extensive calculations (primarily using critical values for the BH procedure) strongly suggest that, for most practical applications, \( P^*(n_F, k, m) \), as a function of \( n_F \), is minimized when \( n_F = s - k - 1 \). Using a Fortran program BHmax with extended precision, no cases were encountered where a value of \( n_F \) resulted in \( P[U \leq k] \) less than its value when \( n_F = s - k - 1 \).

B. Random multivariate normal vector generator

This appendix outlines the methodology used to construct an \( m \)-dimensional vector with the first \( n_T \) components having mutual correlation \( \varrho_T \), and the last \( n_F = m - n_T \) components having mutual correlation \( \varrho_F \), and with the correlation between elements of the two sets of components having mutual correlation \( \varrho_C \).

For the \( i \)-th \( m \)-dimensional random vector, generate \( z_T, z_F \) and \( z_C \) as follows:

\[
\begin{align*}
z_T & = nror(nz,nw) \cdot \sqrt{\varrho_T - \varrho_C} \\
z_F & = nror(nz,nw) \cdot \sqrt{\varrho_F - \varrho_C} \\
z_C & = nror(nz,nw) \cdot \sqrt{\varrho_C}
\end{align*}
\]

where \( nror(nz,nw) \) are randomly generated \( N(0,1) \) variates. (Note that the generation requires that \( \varrho_C \) can never be larger than the smaller of \( \varrho_T \) and \( \varrho_F \).)
Generate \( n_T \) random normal variates:
\[
z(i,j) = \text{rnorm}(n_z,n_w) \cdot \sqrt{1 - \rho_T} + z_T + z_C \quad (j = 1, \ldots, n_T).
\]
Generate \( n_F = m - n_T \) random normal variates:
\[
z(i,j) = \text{rnorm}(n_z,n_w) \cdot \sqrt{1 - \rho_F} + z_F + z_C \quad (j = n_T + 1, \ldots, m).
\]
It is not difficult to show that \( \text{Var}(z(i,j)) = 1 \) Also
\[
\text{Cov}(z(i,j),z(i,j')) = \rho_C \quad (j = 1, \ldots, n_T; j' = n_T + 1, \ldots, m)
\]
\[
\text{Cov}(z(i,j),z(i,j')) = \rho_T \quad (j, j' = 1, \ldots, n_T; j, j' \text{ unequal})
\]
\[
\text{Cov}(z(i,j),z(i,j')) = \rho_F \quad (j, j' = n_T + 1, \ldots, m; j, j' \text{ unequal})
\]
The \( i \)-th \( m \)-dimensional vector \((z(i,1), z(i,2), \ldots, z(i,m))\) has the required characteristics.

C. Worked examples (tutorial)

This appendix illustrates how to use the Fortran program FDR for 6 different objectives. (A seventh possible objective for the user is listed but not illustrated.
The program is interactive and its use begins with the user typing FDR and pressing ENTER for each example. All the program responses and the user actions are replicated. Upon completion of the calculations, FDR notifies the user of the output files which contain the results of the computations (e.g. The output file is AT1.out.) The appropriate output file is explicitly presented for the user.
The seven different objectives are:

1. Calculate powers and false discovery rate (FDR) for multiple step procedures.
2. Check that minimum of \( P[U \leq k] \) occurs when \( n_F = s - k - 1 \).
3. Determine the maximum number of steps so that \( P[U \leq k] \geq \text{GAMMA} \).
4. Check accuracies when \( k_{\text{STAR}} \) is not chosen to equal \( k \).
5. Calculate powers and false discovery rate (FDR) for single step procedures.
6. Examine effects of number of steps on powers and probabilities.
7. Compare two step-up or step-down procedures (not included).

In all cases, before executing the program FDR, the file CV.in must exist and contain the critical values for the step-up or step-down procedure to be analysed by FDR. The critical values must be arranged in non-descending order from smallest to largest.

A Fortran program makeCV, which can be used to give the critical values (in CV.in) for both the BH procedure and the step-down procedure of Lehmann and Romano (2005), is provided in the supplementary materials. Also included are the CV.in files needed for the examples in the appendix.
For all of the examples we assume the conditions for Theorem 6.1 of Somerville and Hemmelmann (2008), see Appendix A. Assume also that the means of the test statistics corresponding to a null hypothesis equal zero.

**Example 1**

Calculate powers and false discovery rate (FDR) for multiple step procedures.

(a) There are 50 null one-sided hypotheses to be simultaneously tested. Using the BH procedure with \( q = 0.05 = \alpha \), determine the per pair, all pairs and any pair power, and the actual false discovery rate (FDR) when there are 1, 11, 21, 31 or 41 false hypotheses, each of which are 3 standard errors greater than those of the true hypotheses (\( \text{delta} = 3 \)). Use \( n = 100000 \) and \( \text{seed} = 757 \). Assume the test statistics have a multivariate-\( t \) distribution with degrees of freedom equal to infinity and \( \rho_T = \rho_F = \rho_C = 0.0 \). We will need a file \( \text{CV.in} \), containing the critical values (ascending order) for the BH procedure.

(b) Would the powers be different with \( \rho_T = 0.2 \), and \( \rho_F = \rho_C = 0.0 \)?

(a) Executing FDR, we obtain:

```
TYPE 1 and press ENTER for MODE 1
TYPE 2 and press ENTER for MODE 2
TYPE 3 and press ENTER for MODE 3

1
```

In the next step user will need to give values for the following parameters.

- \( n \) number of random vectors used
- \( \text{iseed} \) seed used for generation, e.g. 757, 1234998
- \( m \) number of hypotheses
- \( \text{df} \) degrees of freedom for test statistics
  - use "-1" if infinity
- \( \text{nbrsided} \) use "1" for 1-sided tests
  - use "2" for two-sided tests
- \( \text{upordown} \) use pos. integer or zero for stepup
  - use neg. integer for down

GIVE the following PARAMETERS needed for FDRpwr5 (as integers) and TYPE ENTER

```
n  iseed  m  df  upordown  nbrsided
100000 757 50  -1  1  1
```

Type a value for delta, the common expected value of the test statistics corresponding to the false hypotheses.

If you wish to have all false hypotheses rejected, TYPE "10"

Press ENTER
rhoT is the correlation among TRUE test statistics
rhoF is the correlation among FALSE test statistics
rhoC is the correlation between TRUE and FALSE
Give rhoT, rhoF and rhoC
0.0 0.0 0.0

Give number of steps for procedure
50

GIVE values for nFbeg, nFend, nFint.
TYPE ENTER
P[U <= k] will be calculated for values of nF
from nFbeg to nFend, in intervals of nFint
U is the number of false positives

1 41 10

Do you want to repeat with new numbers of false Hypotheses?
If YES, type any positive number, press ENTER
If NO, type 0 or any negative number, press ENTER

0

The file PWR.IN is complete.
PRE completed
Program calculates P[U <= k] for all k <= m
and all the values used in FDR
For which values of k would you like tables?
The program will make tables of P[U <= k] for values
of k from kBEG to kEND in intervals of kINT
Maximum number of tabulated k values in a run is 110!
Give kBEG, kEND and press "ENTER"
Warning: kEND must not be larger than m, the number of hypotheses to be
tested.

1 11 1

instructPOST completed
date and time 20110811 153257.199
Finished calculating for nF = 1
...
Finished calculating for nF = 41
Number of steps is 50
date and time 201110811 15336.543
BODY completed
POST has been successfully executed
Output file is AT1.out
Output file PWR.out was completed in BODY
Date and time 20110811 053306.559

FDR completed

The file PWR.out is as follows:

| nF | Per Pair Power | All Pairs Power | Any Pair Power | Actual FDR |
|----|----------------|-----------------|---------------|------------|
| 1  | 0.473          | 0.473           | 0.473         | 0.049      |
| 11 | 0.730          | 0.053           | 1.000         | 0.039      |
| 21 | 0.815          | 0.023           | 1.000         | 0.029      |
| 31 | 0.859          | 0.015           | 1.000         | 0.019      |
| 41 | 0.887          | 0.011           | 1.000         | 0.009      |

Number of steps is 50
m = 50 df = -1
rhoT = 0.00000000 rhoF = 0.00000000 rhoC = 0.00000000
seed = 757 n = 100000 1-sided test

STEP UP testing delta = 3.00000000 mode = 1

(b) Executing FDR with rhoT = 0.2, rhoF = 0.0, rhoC = 0.0, PWR.out is as follows:

| nF | Per Pair Power | All Pairs Power | Any Pair Power | Actual FDR |
|----|----------------|-----------------|---------------|------------|
| 1  | 0.474          | 0.474           | 0.474         | 0.046      |
| 11 | 0.730          | 0.054           | 1.000         | 0.038      |
| 21 | 0.815          | 0.023           | 1.000         | 0.029      |
| 31 | 0.859          | 0.015           | 1.000         | 0.019      |
| 41 | 0.887          | 0.011           | 1.000         | 0.009      |

Number of steps is 50
m = 50 df = -1
rhoT = 0.200000003 rhoF = 0.00000000E+00 rhoC = 0.00000000E+00
seed = 757 n = 100000 1-sided test

STEP UP testing delta = 3.00000000 mode = 1

*Example 2*

Check that minimum of \( P[U \leq k] \) occurs when \( nF = s - k - 1 \)

Check that \( P[U \leq k] \) is smallest when \( nF = s - k - 1 \) for the BH procedure with \( m = 1000 \), \( q = \alpha = 0.05 \) and \( s = 143 \). (Assume that means of test statistics increase without limit.) Using mode 1, \( \text{delta} = 10 \), and \( nF \) first ranging from 0 to 900 in intervals of 100, and second, ranging from 949 to 999 in intervals of 50, \( k \) values from 1 to 11, execute FDR and obtain \( P[U \leq k] \) from AT1.out. Observe that \( P[U \leq k] \) is \( \cup \)-shaped as a function of \( nF \).

Execute again, as previous, but this time with \( nF \) ranging from 135 to 150, again obtaining \( P[U \leq k] \) from AT1.out. Observe that the smallest value of \( P \) occurs when \( nF = s - k - 1 \).
(Using \( n = 100,000 \), an estimate of a value with a true probability of 0.95 would have a standard error of 0.00069.)

Execute FDR (first execution for Example 2) results in:

TYPE 1 and press ENTER for MODE 1
TYPE 2 and press ENTER for MODE 2
TYPE 3 and press ENTER for MODE 3

1

In the next step user will need to give values for the following parameters.

- \( n \) number of random vectors used
- \( \text{iseed} \) seed used for generation, e.g. 757, 1234998
- \( m \) number of hypotheses
- \( \text{df} \) degrees of freedom for test statistics
  - use "-1" if infinity
- \( \text{nbrsided} \) use "1" for 1-sided tests
  - use "2" for two-sided tests
- \( \text{upordown} \) use pos. integer or zero for stepup
  - use neg. integer for down

GIVE the following PARAMETERS needed for FDR (as integers)

\[
\begin{align*}
n & = 100000 \\
\text{iseed} & = 757 \\
m & = 1000 \\
\text{df} & = -1 \\
\text{upordown} & = 1 \\
\text{nbrsided} & = 1
\end{align*}
\]

Type a value for delta, the common expected value of the test statistics corresponding to the false hypotheses.
If you wish to have all false hypotheses rejected, TYPE "10"
Press ENTER

10

\( \rho_T \) is the correlation among TRUE test statistics
\( \rho_F \) is the correlation among FALSE test statistics
\( \rho_C \) is the correlation between TRUE and FALSE
Give \( \rho_T, \rho_F \) and \( \rho_C \)

0.0 0.0 0.0

Give number of steps for procedure

143
GIVE values for \( nF_{beg} \), \( nF_{end} \), \( nF_{int} \).
TYPE ENTER
\( P[U \leq k] \) will be calculated for values of \( nF \)
from \( nF_{beg} \) to \( nF_{end} \), in intervals of \( nF_{int} \)
\( U \) is the number of false positives

0 900 100

Do you want to repeat with new numbers of false Hypotheses?
If YES, type any positive number, press ENTER
If NO, type 0 or any negative number, press ENTER

9

GIVE values for \( nF_{beg} \), \( nF_{end} \), \( nF_{int} \).
TYPE ENTER
\( P[U \leq k] \) will be calculated for values of \( nF \)
from \( nF_{beg} \) to \( nF_{end} \), in intervals of \( nF_{int} \)
\( U \) is the number of false positives

949 999 50

Do you want to repeat with new numbers of false Hypotheses?
If YES, type any positive number, press ENTER
If NO, type 0 or any negative number, press ENTER

0

The file PWR.IN is complete.
PRE completed
Program calculates \( P[U \leq k] \) for all \( k \leq m \)
and all the values used in FDR
For which values of \( k \) would you like tables?
The program will make tables of \( P[U \leq k] \) for values
of \( k \) from \( k_{BEG} \) to \( k_{END} \) in intervals of \( k_{INT} \)
Maximum number of tabulated \( k \) values in a run is 110!
Give \( k_{BEG} \), \( k_{END} \) and press "ENTER"
Warning: \( k_{END} \) must not be larger than \( m \), the number of hypotheses to be tested.

1 11 1

instructPOST completed
date and time 20110817 104014.322
Finished calculating for \( nF = 0 \)
...
Finished calculating for \( nF = 999 \)
Number of steps is 143
date and time 20110817 104904.729
BODY completed
POST has been successfully executed
Output file is AT1.out
Output file PWR.out was completed in BODY
date and time 20110817 104905.634
FDR completed

The file AT1.out is as follows:

\[ P[U \leq k] \] U is # of false rejections
\[
\begin{array}{cccccccccccc}
  nF & k & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\
  0 & 0.995 & 0.999 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 \\
  100 & 0.055 & 0.156 & 0.310 & 0.488 & 0.787 & 0.880 & 0.937 & 0.969 & 0.985 & 0.994 & & \\
  200 & 0.022 & 0.075 & 0.174 & 0.320 & 0.491 & 0.649 & 0.782 & 0.875 & 0.933 & 0.968 & 0.985 & \\
  300 & 0.040 & 0.121 & 0.261 & 0.437 & 0.616 & 0.761 & 0.867 & 0.932 & 0.968 & 0.986 & 0.994 & \\
  400 & 0.071 & 0.195 & 0.376 & 0.571 & 0.738 & 0.858 & 0.930 & 0.968 & 0.988 & 0.995 & 0.998 & \\
  500 & 0.126 & 0.305 & 0.518 & 0.712 & 0.848 & 0.929 & 0.970 & 0.989 & 0.996 & 0.999 & 1.000 & \\
  600 & 0.218 & 0.452 & 0.677 & 0.839 & 0.930 & 0.974 & 0.991 & 0.997 & 0.999 & 1.000 & 1.000 & \\
  700 & 0.365 & 0.635 & 0.830 & 0.934 & 0.978 & 0.994 & 0.998 & 1.000 & 1.000 & 1.000 & 1.000 & \\
  800 & 0.579 & 0.827 & 0.944 & 0.984 & 0.997 & 0.999 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & \\
  900 & 0.839 & 0.964 & 0.994 & 0.999 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & \\
  950 & 0.949 & 0.994 & 0.999 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & \\
  999 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 \\
\end{array}
\]

Observe from AT1.out that \( P[U \leq k] \) as a function of nF is \( \cup \)-shaped.

Executing the second time with nF ranging from 130 to 145, the file AT1.out is as follows:

\[ P[U \leq k] \] U is # of false rejections
\[
\begin{array}{cccccccccccc}
  nF & k & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\
  130 & 0.022 & 0.071 & 0.164 & 0.299 & 0.457 & 0.611 & 0.743 & 0.842 & 0.910 & 0.952 & 0.976 & \\
  131 & 0.021 & 0.069 & 0.160 & 0.294 & 0.451 & 0.605 & 0.737 & 0.839 & 0.908 & 0.950 & 0.975 & \\
  132 & 0.020 & 0.067 & 0.157 & 0.289 & 0.445 & 0.600 & 0.732 & 0.835 & 0.905 & 0.948 & 0.975 & \\
  133 & 0.020 & 0.065 & 0.154 & 0.284 & 0.439 & 0.594 & 0.727 & 0.831 & 0.902 & 0.949 & 0.975 & \\
  134 & 0.019 & 0.064 & 0.150 & 0.279 & 0.434 & 0.588 & 0.721 & 0.827 & 0.902 & 0.949 & 0.975 & \\
  135 & 0.018 & 0.062 & 0.147 & 0.274 & 0.428 & 0.582 & 0.717 & 0.828 & 0.903 & 0.950 & 0.976 & \\
  136 & 0.018 & 0.060 & 0.145 & 0.269 & 0.421 & 0.576 & 0.718 & 0.829 & 0.904 & 0.950 & 0.976 & \\
  137 & 0.017 & 0.059 & 0.141 & 0.264 & 0.416 & 0.577 & 0.719 & 0.830 & 0.904 & 0.950 & 0.976 & \\
  138 & 0.017 & 0.057 & 0.138 & 0.260 & 0.417 & 0.578 & 0.720 & 0.830 & 0.905 & 0.951 & 0.976 & \\
  139 & 0.016 & 0.056 & 0.135 & 0.261 & 0.418 & 0.580 & 0.721 & 0.831 & 0.905 & 0.951 & 0.976 & \\
  140 & 0.015 & 0.054 & 0.136 & 0.262 & 0.419 & 0.581 & 0.722 & 0.832 & 0.906 & 0.951 & 0.977 & \\
  141 & 0.015 & 0.054 & 0.137 & 0.263 & 0.421 & 0.582 & 0.723 & 0.832 & 0.906 & 0.951 & 0.977 & \\
  142 & 0.015 & 0.055 & 0.137 & 0.263 & 0.422 & 0.583 & 0.724 & 0.833 & 0.907 & 0.952 & 0.977 & \\
  143 & 0.015 & 0.055 & 0.138 & 0.264 & 0.423 & 0.584 & 0.725 & 0.834 & 0.907 & 0.952 & 0.977 & \\
  144 & 0.015 & 0.055 & 0.138 & 0.265 & 0.424 & 0.585 & 0.726 & 0.835 & 0.908 & 0.952 & 0.977 & \\
  145 & 0.015 & 0.056 & 0.139 & 0.266 & 0.426 & 0.586 & 0.727 & 0.836 & 0.908 & 0.952 & 0.977 & \\
\end{array}
\]
From this AT1.out file, we observe that the value of \( P[U \leq k] \) is never less than its value at \( nF = s - k - 1 \). (The values are mined from the file MISC.out where \( P[U \leq k] \) is given to 4 decimal places. An estimate of a probability of 0.95 would have a standard error of 0.00069.)

**Example 3**

Determine the maximum number of steps so that \( P[U \leq k] \geq \text{GAMMA} \).

(a) Given 1000 hypotheses, make tables for the largest number of steps so that \( P[U \leq k] \geq \text{GAMMA} \). Use the BH procedure (one-sided) with \text{GAMMA} values of 0.95 and 0.99. Assume all correlations are zero and degrees of freedom infinity. Use \text{kSTAR} value of 1.

(b) Having completed 3a, and having the external file MISC.out, we can now use mode 3 to determine the maximum number of steps such that \( P[U \leq k] \geq \text{GAMMA} \) for a different set of \( k \) values, and, or different values of \text{GAMMA} for the same number of hypotheses and the same values of \( \alpha \) and degrees of freedom. Find the maximum number of steps such that \( P[U \leq k] \geq \text{GAMMA} \), for \( k < 12 \) when \text{GAMMA} equals 0.50 and 0.90.

(a) Two approaches can be used. We can make a preliminary study using numbers of steps from 3 to 400 in intervals of 10, (for example) using a small value of \( n \) (say 1000). The subroutine \text{POST} will use linear interpolation to give preliminary values, and then we can make a second study using one or more series of values for the number of steps and a larger value of \( n \), and all numbers of steps from 3 to 400 in intervals of 1, and use \( n = 100000 \). This is much more time consuming.

Be sure that CV.in has the critical values (non-decreasing order) for the desired procedure. The required file for the BH procedure can be found in the supplementary materials. The first step is to execute \text{FDR}, using mode 2, and \text{delta} = 10.

Executing \text{FDR}, using \text{n}=100000, we have:

\text{FDR}  

\text{TYPE 1} and PRESS ENTER for \text{MODE 1}  
\text{TYPE 2} and PRESS ENTER for \text{MODE 2}  
\text{TYPE 3} and PRESS ENTER for \text{MODE 3}  

2  

In the next step user will need to give values for the following parameters.  
\text{n} number of random vectors used  
\text{iseed} seed used for generation, e.g. 757, 1234998  
\text{m} number of hypotheses  
\text{df} degrees of freedom for test statistics  
use "-1" if infinity  
\text{nbrsided} use "1" for 1-sided tests  
use "2" for two-sided tests  
\text{upordown} use pos. integer or zero for stepup  
use neg. integer for down
GIVE the following PARAMETERS needed for FDRpwr (as integers) and TYPE ENTER
   n iseed m df upordown nbrsided

100000 757 1000 -1 1 1

Type a value for delta, the common expected value of the test statistics corresponding to the false hypotheses. If you wish to have all false hypotheses rejected, TYPE "10"
Press ENTER

10

rhoT is the correlation among TRUE test statistics
rhoF is the correlation among FALSE test statistics
rhoC is the correlation between TRUE and FALSE
Give rhoT,rhoF and rhoC

0.0 0.0 0.0

Type a value for kSTAR
To give the most accurate computation of P[U <= k] use kSTAR=k. However, especially for large values of m, the value of P is only very slightly overestimated using values of k which are moderately larger than kSTAR. Using values for k smaller than kSTAR usually results in large overestimates of P. Thus a value of kSTAR greater than contemplated values of k is not recommended
A value of kSTAR equal to 1 is often sufficient.

1

Type nsBEG nsEND nsINT, press ENTER

3 400 1

Do you want to repeat the process with a different value for the number of steps? If YES, type any positive number, press ENTER If NO, type 0 or any negative number, press ENTER

0

File PWR.in is complete.
PRE completed
Program calculates P[U <= k] for all k <= m and all the values of nF used in FDR
For which values of \( k \) would you like tables?
The program will make tables of \( P[U \leq k] \) for values
of \( k \) from \( k_{\text{BEG}} \) to \( k_{\text{END}} \) in intervals of \( k_{\text{INT}} \)
Maximum number of tabulated \( k \) values in a run is 110!
Give \( k_{\text{BEG}}, k_{\text{END}}, k_{\text{INT}} \) and press "ENTER"
WARNING: \( k_{\text{END}} \) must not be larger than \( m \), the number of hypotheses tested

\[ \begin{align*}
1 & 22 1
\end{align*} \]

The program can calculate the largest number of steps
such that \( P[U \leq k] \geq \Gamma \).
How many Gammas will you be using?

2

Give the 2 values and press ENTER.
For Example .99 .95 .90

\[ .95 .99 \]

instructPOST completed
date and time 20110813 125256.970
Finished calculating for steps = 3
...
Finished calculating for steps = 400
date and time 20110813 150842 .856
Finished forMAXsteps
POST has been successfully executed
Output files are AT1.out,AT2.out
Output file PWR.out was completed in BODY
See also file MAXsteps.out
date and time 20110813 150844.369
FDR completed

The file MAXsteps.out is as follows:

\[ \begin{align*}
m & = 1000 \text{ df} = -1 \\
rho_T & = 0.00000000 \text{ rhoF} = 0.00000000 \text{ rhoC} = 0.00000000 \\
seed & = 757 \ n = 100000 \text{ 1 -sided test} \\
STEP UP testing & \text{ 10.00000000 mode} = 2 \\
NOTE: A negative number (or zero) for the max number of steps (eg -xx)
indicates an undetermined max less than xx.
A number equal to \( n_{\text{END}} \) usually indicates an undetermined max number of
steps exceeding \( n_{\text{END}} \). However too small a value for \( n \) and or \( n_{\text{END}} \)
could be the cause. Examination of the appropriate AT file, or a rerun
with a larger \( n_{\text{INT}} \) (quickest) or a larger value of \( n \) is suggested.
\( n_{\text{END}} = 400 \).
We now use an alternate (quicker and less exact) procedure. Before proceeding, save the file \texttt{MISC.out} (use another name as \texttt{MISC.out} will be overwritten in the following). We shall again execute \texttt{FDR}, but this time use \( n = 1000 \). We will again obtain the exterior files \texttt{MISC.out}, \texttt{AT1.out}, \texttt{AT2.out} and \texttt{MAXsteps.out}. The resulting \texttt{MAXsteps.out} exterior file is given below:

\begin{verbatim}
Maximum number of steps to guarantee \( P[U \leq k] \geq \text{GAMMA} \). \( \text{kSTAR}= 1 \)
\text{GAMMA} 0.950
\begin{verbatim}
k 1 2 3 4 5 6 7 8 9 10 11
  7 16 28 41 54 70 86 104 123 144 166
k 12 13 14 15 16 17 18 19 20 21 22
  189 215 245 279 318 368 400 400 400 400 400
\end{verbatim}
\text{GAMMA} 0.990
\begin{verbatim}
k 1 2 3 4 5 6 7 8 9 10 11
  -3 8 16 26 36 48 61 75 89 106 123
k 12 13 14 15 16 17 18 19 20 21 20
  142 160 182 206 233 261 295 336 400 400 400
\end{verbatim}
\end{verbatim}

It is obvious that the maximum number of steps such that \( P[U \leq k] \geq \text{GAMMA} \) is a non-decreasing function of \( k \). However the file suggests otherwise. The problem is that we are using \( n = 1000 \), and the standard error of an estimate of a probability which is 0.95 is 0.007. An examination of the \texttt{AT1.out} file has probability estimates with obvious 'errors'. (See \texttt{GAMMA} = 0.95, and the sequence of probabilities for \texttt{nsteps} = 45, 46, 47 for example).

We shall try to circumvent the problem by repeating the execution of \texttt{FDR} using \texttt{nsINT} = 10 instead of 1. The output of the file \texttt{MAXsteps.out} is now:

\begin{verbatim}
Maximum number of steps to guarantee \( P[U \leq k] \geq \text{GAMMA} \). \( \text{kSTAR}= 1 \)
\text{GAMMA} 0.950
\begin{verbatim}
k 1 2 3 4 5 6 7 8 9 10 11
  7 16 28 41 54 70 86 104 123 144 166
k 12 13 14 15 16 17 18 19 20 21 22
  400 400 400 400 400 400 400 400 400 400 400
\end{verbatim}
\text{GAMMA} 0.990
\begin{verbatim}
k 1 2 3 4 5 6 7 8 9 10 11
  -3 8 18 24 34 40 56 400 87 400 400
k 12 13 14 15 16 17 18 19 20 21 22
  400 400 400 400 400 400 400 400 400 400 400
\end{verbatim}
\end{verbatim}
(b) Use mode 3 (and existing external file MISC.out) to find the maximum number of steps such that \( P[U \leq k] \geq \text{GAMMA} \) for \( k \) less than or equal to 11 when GAMMA equals 0.50 and 0.90. Execution of FDR results in the following exterior file MAXsteps.out:

| GAMMA 0.500 | k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------------|---|---|---|---|---|---|---|---|---|---|----|----|
|             |   | 34| 56| 79|103|129|157|188|222|261|307|368|

| GAMMA 0.900 | k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------------|---|---|---|---|---|---|---|---|---|---|----|----|
|             |   | 10| 22| 36| 51| 67| 84|103|123|145|168|194|

\( \text{nsEND}= 393 \)

This has removed the inconsistency, but reducing \( n \) from 100000 to 1000 has decreased the accuracy materially.

Example 4

Check accuracies when \( k\text{STAR} \) is not chosen to equal \( k \).

We have demonstrated that \( P[U \leq k] \) as a function of \( nF \) is \( U \)-shaped, and have proved that, under the conditions of Theorem 6.1 of Somerville and Hemmelmann (2008), the minimum occurs when \( nF \) is not greater than \( s-1 \), and assert that, in most practical cases, the minimum occurs when \( nF = s-k-1 \).

We have observed that the slope of \( P[U \leq k] \) is steepest when \( nF \) is less than \( s-k \), and is nearly flat for a large proportion of the range from \( nF = s-k \) to \( m \).

Table 1 illustrates the differences between the maximum number of steps such that \( P[U \leq k] \geq \text{GAMMA} \) when \( k\text{STAR} = k \), and when \( k\text{STAR} \) is 5.

Note that, as previously indicated, using \( k\text{STAR} \) larger or smaller, in FDR results in an overestimate of the number of steps which can be used, and still maintain \( P[U \leq k] \geq \text{GAMMA} \). If \( k\text{STAR} \) is less than \( k \), the overestimate is usually slight, particularly if \( m \) or \( k \) or both are large. However, when \( k\text{STAR} \) is greater than \( k \) the overestimate is much more serious, even when the difference is small.
Table 1: Table comparing $P[U \leq k]$ when $k_{STAR} = 5$, and $k_{STAR} = k$, $m = 1000$, $q = 0.05$, BH procedure, $\Gamma = 0.95$, $n = 100000$.

Example 5

Calculate powers and false discovery rate (FDR) for single step procedures.

A single step procedure can be regarded as a multi-step procedure with $s = 1$. The following example uses FDR to determine the powers and false discovery rate (FDR) for the Bonferroni procedure. Note that, using $s = 1$, the file CV.in can use the critical values for the Holm procedure, the Lehman-Romano procedure with $k = 1$, or the BH procedure.

Problem: For $m = 1000$, $q = 0.05$ and $\delta = 3$, find the powers and false discovery rate (FDR) of the Bonferroni procedure.

Use $s = 1$, and assume all correlations equal to 0.

Executing FDR we obtain the following:

1

In the next step user will need to give values for the following parameters.

- **n** number of random vectors used
- **iseed** seed used for generation, e.g. 757, 1234998
- **m** number of hypotheses
- **df** degrees of freedom for test statistics
  - use "-1" if infinity
- **nbrsided** use "1" for 1-sided tests
  - use "2" for two-sided tests
- **upordown** use pos. integer or zero for stepup
  - use neg. integer for down

GIVE the following PARAMETERS needed for FDRpwr (as integers) and TYPE ENTER

- **n**
- **iseed**
- **m**
- **df**
- **nbrsided**

100000 757 1000 -1 1 1

Type a value for $\delta$, the common expected value of the test statistics corresponding to the false hypotheses.

If you wish to have all false hypotheses rejected, TYPE "10"

Press ENTER
rhoT is the correlation among TRUE test statistics
rhoF is the correlation among FALSE test statistics
rhoC is the correlation between TRUE and FALSE
Give rhoT, rhoF, and rhoC

0.0 0.0 0.0

Give number of steps for procedure

1

GIVE values for nFbeg, nFend, nFint.
TYPE ENTER

1 20 1

P[U <= k] will be calculated for values of nF from nFbeg to nFend, in intervals of nFint
U is the number of false positives

Do you want to repeat with new numbers of false Hypotheses?
If YES, type any positive number, press ENTER
If NO, type 0 or any negative number, press ENTER

0

File PWR.IN is complete.
PRE completed
Program calculates P[U <= k] for all k <= m and all the values of nF used in FDR
For which values of k would you like tables?
The program will make tables of P[U <= k] for values of k from kBEG to kEND in intervals of kINT.
Maximum number of tabulated k values in a run is 110!
Give kBEG, kEND, kINT and press "ENTER"
WARNING: kEND must not be larger than m, the number of hypotheses.

1 22 1

InstructPOST completed
date and time 20111012 113006.187
Finished calculating for nF = 1
...
Finished calculating for nF = 20
Number of steps is  1
date and time 20111012 113621.062
BODY completed

POST has been successfully executed
Output files are AT1.out, AT2.out
Output file PWR.out was completed in BODY
date and time 20111012 113621.234
FDR completed

The following is the output in PWR.out:

| nF | Per Pair Power | All Pairs Power | Any Pair Power | Actual FDR |
|----|----------------|----------------|---------------|------------|
| 1  | 0.189          | 0.189          | 0.189         | 0.044      |
| 2  | 0.188          | 0.036          | 0.339         | 0.040      |
| 3  | 0.187          | 0.007          | 0.462         | 0.037      |
| 4  | 0.187          | 0.001          | 0.564         | 0.033      |
| 5  | 0.188          | 0.000          | 0.645         | 0.031      |
| 6  | 0.187          | 0.000          | 0.711         | 0.028      |
| 7  | 0.187          | 0.000          | 0.766         | 0.026      |
| 8  | 0.187          | 0.000          | 0.811         | 0.024      |
| 9  | 0.187          | 0.000          | 0.844         | 0.023      |
| 10 | 0.187          | 0.000          | 0.874         | 0.021      |
| 11 | 0.187          | 0.000          | 0.898         | 0.020      |
| 12 | 0.187          | 0.000          | 0.916         | 0.019      |
| 13 | 0.187          | 0.000          | 0.932         | 0.017      |
| 14 | 0.187          | 0.000          | 0.945         | 0.017      |
| 15 | 0.187          | 0.000          | 0.955         | 0.016      |
| 16 | 0.187          | 0.000          | 0.964         | 0.015      |
| 17 | 0.187          | 0.000          | 0.970         | 0.014      |
| 18 | 0.187          | 0.000          | 0.976         | 0.013      |
| 19 | 0.187          | 0.000          | 0.981         | 0.013      |
| 20 | 0.187          | 0.000          | 0.984         | 0.012      |

Number of steps is  1
m = 1000 df= -1
rhoT= 0.00000000 rhoF= 0.00000000 rhoC= 0.00000000
seed= 757 n= 100000 1 -sided test
STEP UP testing delta= 3.00000000 mode= 1

Example 6

Examine effects of number of steps on powers and probabilities.

Compare the per pair power and \( P[U \leq 1] \) for \( s = 1, 11, \ldots, 110 \) and 800 for values of \( nF = 1, 11, \ldots, 110, \) and 800. Use \( q = 0.05, m = 1000, n = 100000, \) \( \text{delta} = 3, \) \( \text{rhoT} = \text{rhoF} = \text{rhoC} = 0.0 \) for the BH procedure.

We will need a run for each value of \( s \). We illustrate for \( s = 1 \). Executing FDR, we obtain:
TYPE 1 and press ENTER for MODE 1
TYPE 2 and press ENTER for MODE 2
TYPE 3 and press ENTER for MODE 3

1

In the next step user will need to give values for the following parameters.
n number of random vectors used
iseed seed used for generation, e.g. 757, 1234998
m number of hypotheses
df degrees of freedom for test statistics
use "-1" if infinity
nbrsided use "1" for 1-sided tests
use "2" for two-sided tests
upordown use pos. integer or zero for stepup
use neg. integer for down

GIVE the following PARAMETERS needed for FDR (as integers)
and TYPE ENTER
n iseed m df upordown nbrsided

100000 757 1000 -1 1 1

Type a value for delta, the common expected value of the
test statistics corresponding to the false hypotheses.
If you wish to have all false hypotheses rejected, TYPE "10"
Press ENTER

3

rhoT is the correlation among TRUE test statistics
rhoF is the correlation among FALSE test statistics
rhoC is the correlation between TRUE and FALSE
Give rhoT,rhoF and rhoC

0.0 0.0 0.0

Give number of steps for procedure

1

GIVE values for nFbeg, nFend, nFint.
TYPE ENTER
P[U.le.k] will be calculated for values of nF
from nFbeg to nFend, in intervals of nFint
U is the number of false positives
1 101 10

Do you want to repeat with new numbers of false Hypotheses?
If YES, type any positive number, press ENTER
If NO, type 0 or any negative number, press ENTER

9

GIVE values for nFbeg, nFend, nFint.
TYPE ENTER
P[U \leq k] will be calculated for values of nF
from nFbeg to nFend, in intervals of nFint
U is the number of false positives

800 800 1

Do you want to repeat with new numbers of false Hypotheses?
If YES, type any positive number, press ENTER
If NO, type 0 or any negative number, press ENTER

0

The file PWR.IN is complete.
PRE completed
Program calculates P[U \leq k] for all k \leq m
and all the values of nF used in FDR.
For which values of k would you like tables?
The program will make tables of P[U \leq k] for values
of k from kBEG to kEND in intervals of kINT
Maximum number of tabulated k values in a run is 110!
Give kBEG, kEND, kINT and press "ENTER"
Warning: kEND must not be larger than m, the number of hypotenuses

1 22 1

INSTRUCTpost COMPLETED
Date and time 20111012 130407.281
Finished calculating for nF = 1
...
Finished calculating for nF = 101
Finished calculating for nF = 800
Number of steps is 1
Date and time 20111012 130407.281
BODY completed

POST has been successfully executed
Output files are AT1.out, AT2.out
Output file PWR.out was completed in BODY
Date and time 20111012 130409.562
FDR completed

The output for $P[U \leq k]$ and per pair power can be found in AT1.out and PWR.out and are given below:

| nF\k  | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   |
|-------|------|------|------|------|------|------|------|------|------|------|------|
|       | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 11    | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 21    | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 31    | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 41    | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 51    | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 61    | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 71    | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 81    | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 91    | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 101   | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| 800   | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|

| nF    | Per Pair Power | All Pairs Power | Any Pair Power | Actual FDR |
|-------|----------------|-----------------|---------------|------------|
| 1     | 0.189          | 0.189           | 0.189         | 0.044      |
| 11    | 0.187          | 0.000           | 0.898         | 0.020      |
| 21    | 0.187          | 0.000           | 0.987         | 0.012      |
| 31    | 0.187          | 0.000           | 0.998         | 0.008      |
| 41    | 0.187          | 0.000           | 1.000         | 0.006      |
| 51    | 0.186          | 0.000           | 1.000         | 0.005      |
| 61    | 0.186          | 0.000           | 1.000         | 0.004      |
| 71    | 0.186          | 0.000           | 1.000         | 0.003      |
| 81    | 0.186          | 0.000           | 1.000         | 0.003      |
| 91    | 0.186          | 0.000           | 1.000         | 0.002      |
| 101   | 0.186          | 0.000           | 1.000         | 0.000      |
| 800   | 0.187          | 0.000           | 1.000         | 0.000      |

We use column 2 of each table to produce the line for $s = 1$ for the tables of $P[U \leq k]$, Table 2 and per pair power, Table 3. Additional runs for $s = 11, 21, ...$ are needed for the respective lines in Tables 2 and 3 given below.
Table 2: Table of $P[U\leq 1]$.  

| $s\backslash nF$ | 1   | 11  | 21  | 31  | 41  | 51  | 61  | 71  | 81  | 91  | 101 | 800 |
|----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1              | .999 | .999 | .999 | .999 | .999 | .999 | .999 | .999 | .999 | .999 | .999 | 1.000 |
| 11             | .994 | .965 | .919 | .902 | .902 | .902 | .902 | .902 | .902 | .902 | .902 | .994 |
| 21             | .994 | .965 | .905 | .824 | .757 | .739 | .741 | .745 | .749 | .752 | .756 | .980 |
| 31             | .994 | .965 | .905 | .822 | .724 | .632 | .585 | .578 | .582 | .587 | .593 | .960 |
| 41             | .994 | .965 | .905 | .822 | .724 | .623 | .526 | .461 | .444 | .449 | .935 |
| 51             | .994 | .965 | .905 | .822 | .724 | .623 | .524 | .433 | .364 | .334 | .332 | .905 |
| 61             | .994 | .965 | .905 | .822 | .724 | .623 | .524 | .432 | .353 | .288 | .263 | .872 |
| 71             | .994 | .965 | .905 | .822 | .724 | .623 | .524 | .432 | .353 | .284 | .228 | .839 |
| 81             | .994 | .965 | .905 | .822 | .724 | .623 | .524 | .432 | .353 | .284 | .227 | .804 |
| 91             | .994 | .965 | .905 | .822 | .724 | .623 | .524 | .432 | .353 | .284 | .227 | .768 |
| 101            | .994 | .965 | .905 | .822 | .724 | .623 | .524 | .432 | .353 | .284 | .227 | .730 |
| 800            | .994 | .965 | .905 | .822 | .724 | .623 | .524 | .432 | .353 | .284 | .227 | .007 |

Table 3: Table of per pair power.

Affiliation:
Paul N. Somerville  
University of Central Florida  
Department of Statistics and Actuarial Science  
4000 Central Florida Boulevard  
Orlando, FL 32816, United States of America  
Telephone: +43/1/321-773-1854  
E-mail: somervil@knights.ucf.edu  
URL: http://pegasus.cc.ucf.edu/~somervil/