FDR Controlling Step-up-down Tests Related to the Asymptotically Optimal Rejection Curve

By Helmut Finner

Institute of Biometrics & Epidemiology
German Diabetes Center
Leibniz-Institute at the Heinrich-Heine-University, Duesseldorf, Germany

\[ \mu \text{TOSS 2010 Berlin - Multiple comparisons from theory to practice} \]

February 15-16, 2010
Co-Workers

Main Co-Workers:

Thorsten Dickhaus (Berlin Institute of Technology)
Veronika Gontscharuk (DDZ)
Markus Roters (Omnicare Clinical Research)

Further:

Marsel Scheer (DDZ)
Klaus Strassburger (DDZ)
References

Personal view on FDR

- **1995:** Very sceptical, FDR allows cheating; Proof in B&H ‘incomplete’; no reference to Eklund (1961-1963), Eklund & Seeger (1965)
- **1996:** First talk on FDR, Title: Controlling the false discovery rate: a doubtful concept in multiple comparisons
  But statement: There will be hundreds of papers on FDR in the near future
- **1997-1999:** Draft with the same title
- **2001:** Publication (Biom. J.) with a more relaxed title
- **2002:** Theoretical paper (Ann. Statist.) on expected errors (FWER and FDR), Asymptotics
- **1998-2003:** Work on the Partitioning Principle (FWER)
- **2005-2007:** FDR, dependency, asymptotics, optimal rejection curve
Acknowledgment

We should all be very grateful to

Prof. Yoav Benjamini and Prof. Yosef Hochberg

for their 1995 paper which can be viewed as the initial point for a new era in multiple testing!

Overview

Introduction

Warm-up: Linear step-up test (LSU)

Critical value functions and rejection curves

FDR bounds for step-up procedures

Asymptotically optimal rejection curve

Finite FDR control

Recursion for the computation of critical values: Exact solving

Simple adjustments for the AORC

Iterative method

A general results on FDR bounding curves and rejection curves
Notation & Definition of FDR

Θ parameter space

\( H_1, \ldots, H_n \) null-hypotheses; \( p_1, \ldots, p_n \) p-values

\( \varphi = (\varphi_1, \ldots, \varphi_n) \) multiple test procedure

\( V_n = |\{ i : \varphi_i = 1 \text{ and } H_i \text{ true } \}| = \) number of true hypotheses rejected

\( R_n = |\{ i : \varphi_i = 1 \}| = \) number of hypotheses rejected

\[
\text{FDR}_\vartheta(\varphi) = \mathbb{E}_\vartheta \left[ \frac{V_n}{R_n \lor 1} \right] \quad \text{actual false discovery rate given } \vartheta \in \Theta
\]

**Definition 1.** Let \( \alpha \in (0, 1) \) be fixed.

\( \varphi \) controls the false discovery rate (FDR) at level \( \alpha \) if

\[
\text{FDR}(\varphi) = \sup_{\vartheta \in \Theta} \text{FDR}_\vartheta(\varphi) \leq \alpha.
\]
False Discovery Proportion (FDP)

Define $FDP = 0$ if $R_n = 0$ and for $R_n > 0$

$$FDP = \frac{V_n}{R_n}$$

$= \frac{\text{number of true hypotheses rejected}}{\text{number of hypotheses rejected}}$

$= \frac{\text{number of false significances}}{\text{number of significances}}$

$$FDR = \text{expectation of } FDP = E[FDP]$$
A Note on a Method for the Analysis of Significances en masse

PAUL SEEGER

University of Uppsala, Sweden

This note concerns the derivation of the \( p \)-mean significance levels, in the case of independent tests, for a mass-significance method developed by Eklund [1]. The solution is reached by formulating and solving an urn problem. Some comparisons are made with the \( p \)-mean significance levels of Duncan's multiple range test.

In three seminar papers from 1961–1963 [1] Eklund suggested the following solution to what he called the mass-significance problem: In large exploratory investigations it is desirable to keep the proportion of false significances low, at most equal to a small value \( k \). Consider therefore the variable

\[
y = \frac{\text{number of false significances}}{\text{number of significances}}
\]

where the denominator is observed but the numerator has to be predicted. Both numerator and denominator are functions of the level of significance \( \alpha' \), which is supposed to be used for each of \( N \) tests. Eklund's method consists in determining \( \alpha' \) so that \( y \leq k \), where \( k \) is predetermined. The observed number of
The FDR-Theorem: Benjamini & Hochberg (1995)

$H_i$ true for $i \in I_{n,0}$, $H_i$ false for $i \in I_{n,1}$

$I_{n,0} + I_{n,1} = I_n = \{1, \ldots, n\}$, $n_0 = |I_{n,0}|$

Basic independence assumptions (BIA):

- $p_i \sim \text{U}([0, 1])$, $i \in I_{n,0}$, independent
- $(p_i : i \in I_{n,0})$, $(p_i : i \in I_{n,1})$ independent

$p_{1:n} \leq \cdots \leq p_{n:n}$ ordered p-values

Linear step-up procedure (LSU) $\varphi_{\text{LSU}}$ based on Simes’ crit. val. $\alpha_{i:n} = i\alpha/n$:

Reject all $H_i$ with $p_i \leq \alpha_{m:n}$, where $m = \max\{i : p_{i:n} \leq \alpha_{i:n}\}$.

Then

$$\text{FDR}_{\varphi}(\varphi_{\text{LSU}}) = \mathbb{E}_{\varphi} \left[ \frac{V_n}{R_n \vee 1} \right] = \frac{n_0}{n} \alpha$$
Linear step-up in terms of ecdf

Reformulate the LSU test in terms of the rejection curve $r(t) = t/\alpha$. We call $r(t) = t/\alpha$ Simes’ line.

Let $F_n$ denote the empirical cdf (ecdf) of the p-values, that is,

$$
\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^{n} I_{[0,t]}(p_i).
$$

Define

$$
t^* = \sup \{ t \in [0, \alpha] : \hat{F}_n(t) \geq t/\alpha \}
$$

and reject all $H_i$ with $p_i \leq t^*$.

Call $t^*$ the largest crossing point (LCP).

Note: $t^* \in \{0, \alpha_{1:n}, \ldots, \alpha_{n:n}\}$ with $\alpha_{i:n} = i\alpha/n = r^{-1}(i/n)$. 
Remark

It is sometimes convenient to write

\[ V_n(t) = \{ i \in I_n,0 : p_i \leq t \}, \]
\[ R_n(t) = \{ i \in I_n : p_i \leq t \}, \]

where \( t \) may be replaced by a random threshold \( \tau \).

Many multiple testing procedures can be defined in terms of a single (random) threshold \( \tau \) resulting in the decision rule

\[ \text{Reject } H_i \text{ if and only if } p_i \leq \tau. \]

Here we give a proof for the LSU-FDR-Theorem based on Optional Stopping.

Re-Definition of LSU Threshold

\( t^* \) in the definition of LSU can be replaced by

\[
\tau = \max \{ t : \hat{F}_n(t) \lor \frac{1}{n} = \frac{t}{\alpha} \}
\]

\[
= \max \{ t : R_n(t) \lor 1 = \frac{n}{\alpha} t \}
\]

\[
= \max \{ \alpha_{i:n} : R_n(\alpha_{i:n}) \lor 1 = i, \ i \in I_n \},
\]

hence,

\[
\text{FDR}_\vartheta (\varphi^{\text{LSU}}) = E_\vartheta \left[ \frac{V_n(\tau)}{R_n(\tau) \lor 1} \right] = \frac{\alpha}{n} E_\vartheta \left[ \frac{V_n(\tau)}{\tau} \right].
\]

It remains to show that \( E_\vartheta \left[ \frac{V_n(\tau)}{\tau} \right] = n_0. \)
Martingale and Stopping Time

**Fact 1:** \( \tau \) is a stopping time with respect to the backward filtration \( \mathcal{F}_t = \sigma(I\{p_i \leq s\}, t \leq s \leq 1, i \in I_n), 0 < t \leq 1. \)

**Proof.** It suffices to show that \( \{\tau \leq \alpha_{i:n}\} \in \mathcal{F}_{\alpha_{i:n}} \). Obviously,

\[
\{\tau \leq \alpha_{i:n}\} = \bigcap_{j=i+1}^{n} \{p_{j:n} > \alpha_{j:n}\} \in \mathcal{F}_{\alpha_{i+1:n}} \subseteq \mathcal{F}_{\alpha_{i:n}} \text{ for } i = 1, \ldots, n-1.
\]

For \( i = n \) we have \( \alpha_{n:n} = \alpha \) and \( \{\tau \leq \alpha\} \in \mathcal{F}_t \) for all \( t \in (0, 1] \).
Martingale and Stopping Time

**Fact 2:** $V(t)/t$ for $0 < t \leq 1$ is a martingale with time running backwards with respect to the filtration $\mathcal{F}_t$, $0 < t \leq 1$.

**Proof.** Easy to check:

1. \[ \frac{V(t)}{t} \text{ is } \mathcal{F}_t - \text{measurable for } 0 < t \leq 1, \]
2. \[ E_\vartheta \left[ \frac{V(s)}{s} \mid \mathcal{F}_t \right] = \frac{V(t)}{t} \text{ for } 0 < s \leq t \leq 1, \]
3. \[ E_\vartheta \left[ \frac{V(t)}{t} \right] = n_0 < \infty. \]

Finally, the **Optional Stopping Theorem** yields

\[ E_\vartheta \left[ \frac{V_n(\tau)}{R_n(\tau) \lor 1} \right] = \frac{\alpha}{n} E_\vartheta \left[ \frac{V_n(\tau)}{\tau} \right] = \frac{\alpha}{n} E_\vartheta \left[ \frac{V_n(1)}{1} \right] = \frac{n_0}{n} \alpha. \]
Can we improve LSU?

Since

$$\text{FDR}_\varphi(\varphi^{\text{LSU}}) = \frac{n_0}{n} \alpha < \alpha \text{ for } n_0 < n,$$

an obvious question is whether LSU can be improved.
Do there exit better rejection curves ?

What happens with the FDR if we use non-linear critical values / a non-linear rejection curve for an SU- or more generally a step-up-down procedure ?

Does there exist a rejection curve such that

$$E_{\vartheta} \left[ \frac{V_n(\tau)}{R_n(\tau) \lor 1} \right] = \alpha$$

for some $\vartheta$’s with $n_0 < n$ ?
Critical value functions

Let \( \rho : [0, 1] \rightarrow [0, 1] \) be non-decreasing and continuous with \( \rho(0) = 0 \) and positive values on \( (0, 1] \).

Define critical values

\[
\alpha_{i:n} = \rho(i/n) \in (0, 1], \ i = 1, \ldots, n.
\]

We call \( \rho \) a critical value function.

Define \( r \) by

\[
r(x) = \inf\{u \in [0, 1] : \rho(u) = x\} \text{ for } x \in [0, 1].
\]

We call \( r \) a rejection curve.

Often: \( r(x) = \rho^{-1}(x) \).
Least favorable configurations for SU based on $\rho$

From Benjamini and Yekutieli (2001) we get the following.

Assume BIA and consider an SU-procedure based on $\rho$.

If $\rho(x)/x$ is non-decreasing in $x \in (0, 1]$ then the FDR is **largest** if $p_i = 0$ a.s. for all alternative $p$-values $p_i$, $i \in I_{n,1}$.

i.e., Dirac-uniform configurations are least favorable (LFC) for the FDR.

**Note:** $\rho(x)/x$ non-decreasing in $x$ iff $\alpha_{i:n}/i$ non-decreasing in $i$

We refer to this monotonicity property as (M).
FDR bound for SU

Let $\vartheta \in \Theta$ such that $n_0 = |I_{n,0}| =$ number of true nulls,

let $\rho$ satisfy (M) and let $I'_{n,0} = I_{n,0} \setminus \{i_0\}$ for some $i_0 \in I_{n,0}$.

Then the FDR of the SU-test $\varphi_{SU,\rho}^{(n)}$ based on $\rho$ is bounded by

$$\text{FDR}_{\vartheta}(\varphi_{SU,\rho}^{(n)}) \leq \frac{n_0}{n} \mathbb{E}_{I'_{n,0}} \left[ \frac{\rho(R_n/n)}{R_n/n} \right],$$

with $\rho(\gamma\zeta) = \gamma\zeta$ for the Dirac-uniform configuration $DU(n,n_0)$.

Moreover, if

$$\lim_{n \to \infty} \frac{n_0}{n} = \zeta \quad \text{and} \quad \lim_{n \to \infty} \frac{R_n}{n} = \gamma\zeta \quad \text{a.s. (w.r.t. } DU(n,n_0)),$$

then the FDR is asymptotically bounded by

$$\zeta \frac{\rho(\gamma\zeta)}{\gamma\zeta}.$$
Idea: Optimize $\rho$ for an SU procedure

Try to find a $\rho$ such that

$$\zeta \frac{\rho(\gamma \zeta)}{\gamma \zeta} \equiv \min\{\zeta, \alpha\}$$

for as many $\zeta$’s as possible!

In other words, try to find a critical value function (or rejection curve) such that the FDR-level $\alpha$ is (asymptotically) exhausted under DU-configurations.

Solution:

$$\rho(t) = \frac{\alpha t}{1 - (1 - \alpha)t}$$

$$r(t) = f_\alpha(t) = \frac{t}{\alpha + (1 - \alpha)t}$$

$f_\alpha$: Asymptotical Optimal Rejection Curve (AORC)
Simes’ line and AORC for $\alpha = 0.1$
A further heuristic for the AORC

Consider DU\((n, n_0)\)-models with \(\lim_{n \to \infty} n_0/n = \zeta\).

Reject all \(H_i\) with \(p_i \leq t\) \((t \in (0, 1))\). Then the asymptotic FDR (depending on \(\zeta\) and \(t\)) is given by

\[
\text{FDR}_\zeta(t) = \frac{t\zeta}{(1 - \zeta) + t\zeta}.
\]

Aim:

\[\text{FDR} \equiv \alpha\] for possibly all \(\zeta \in (0, 1)\).

We have:

\[
\text{FDR}_\zeta(t_\zeta) = \alpha \iff t_\zeta = \frac{\alpha(1 - \zeta)}{\zeta(1 - \alpha)} \quad (\implies \zeta \in (\alpha, 1))
\]

\[
\iff \zeta = \frac{\alpha}{t_\zeta(1 - \alpha) + \alpha}.
\]
A further heuristic for the AORC

Ansatz: $t_\zeta$ crossing point between rejection curve $r$ and the limiting cdf $G_\zeta(t) = 1 - \zeta + \zeta t$,

i.e., $r(t_\zeta) = G_\zeta(t_\zeta)$.

Plugging in $\zeta = \frac{\alpha}{t_\zeta(1-\alpha)+\alpha}$ yields

$$r(t_\zeta) = G_\zeta(t_\zeta) = (1 - \zeta) + \zeta t_\zeta = \frac{t_\zeta}{t_\zeta(1-\alpha)+\alpha}.$$  

Hence, $r = f_\alpha$ defined by

$$f_\alpha(x) = \frac{x}{x(1-\alpha)+\alpha}, \ x \in [0, 1],$$

is the curve we are looking for!
Problem: SU does not work with $f_\alpha$

Critical values induced by AORC:

$$\alpha_{i:n} = f_\alpha^{-1}(i/n) = \frac{i \alpha}{1 - \frac{i}{n}(1 - \alpha)} = \frac{i \alpha}{n - i(1 - \alpha)}, \quad i = 1, \ldots, n,$$

are not valid for SU because of $\alpha_{n:n} = 1$.

Ways out of this dilemma:

- Adjust the AORC slightly
- Step-up-down procedures
Step-down (SD), step-up (SU), step-up-down (SUD)

(1) SUD(\(\lambda\)) test \(\varphi^\lambda\) with parameter \(\lambda \in \{1, \ldots, n\}\):

\[ p_{\lambda:n} \leq \alpha_{\lambda:n} \Rightarrow \text{SD-part:} \]
\[ m = \max\{j \in \{\lambda, \ldots, n\} : p_{i:n} \leq \alpha_{i:n} \text{ for all } i \in \{\lambda, \ldots, j\}\} , \]

\[ p_{\lambda:n} > \alpha_{\lambda:n} \Rightarrow \text{SU-part:} \]
\[ m = \max\{j \in \{1, \ldots, \lambda - 1\} : p_{j:n} \leq \alpha_{j:n}\} , (\max \emptyset = -\infty). \]

Reject all \(H_i\) with \(p_i \leq \alpha_{m:n}\).

\(\lambda = 1 \Rightarrow \) step-down test

\(\lambda = n \Rightarrow \) step-up test.

(2) SUD(\(\lambda\)) based on \(\beta\)-adjusted AORC:

\[ f_{\alpha,\beta_n} = (1 + \beta_n/n)f_\alpha \text{ for a suitable } \beta_n > 0. \]
SUD($\lambda$) test based on $f_\alpha$ with $\lambda_1 = 20$ and $\lambda_2 = 40$
($n = 50, \alpha = 0.1$)
Upper FDR bounds for SUD($\lambda_n$) tests

**Theorem:** (Finner et al (2009))

Let $\vartheta \in \Theta$ such that $n_0$ hypotheses are true.

Under (BIA) and (M) we have for an SUD($\lambda$) test $\varphi^{\lambda}$ based on $\rho$

$$\text{FDR}_{\vartheta}(\varphi^{\lambda}) \leq \frac{n_0}{n} \mathbb{E}_{n,n_0-1} \left[ \frac{\alpha_{R_n:n}}{R_n/n} \right]$$

with 

with "" = "" for a step-up test $\varphi^{SU}$ under DU$(n, n_0)$.

**Note:** $\alpha_{R_n:n} = \frac{n_0}{n} \rho(R_n/n)$.

**Important:** $b(n, n_0|\lambda_n) := \mathbb{E}_{n,n_0-1} \left[ \frac{\alpha_{R_n:n}}{R_n/n} \right]$ is computable!
Upper FDR bounds for SUD($\lambda_n$) tests: Asymptotics

(A) For SUD($\lambda_n$) tests with $\lambda_n/n \to \kappa \in (0, 1]$ and $n_0/n \to \zeta \in [0, 1]$: 

$$\lim_{n \to \infty} b(n, n_0|\lambda_n) = \lim_{n \to \infty} \text{FDR}_{n,n_0}.$$  

(B) For $\kappa = 0$ (e.g. SD tests) and $\zeta \in [0, 1)$, (1) holds too.

(C) But: For $\kappa = 0$, $\zeta = 1$ it is possible that 

$$\lim_{n \to \infty} b(n, n_0|\lambda_n) > \lim_{n \to \infty} \text{FDR}_{n,n_0}.$$  

Example: Let $n_0 = n$. Then the FDR (=FWER) of an SD test based on $f_\alpha$ is equal to $1 - (1 - \alpha_{1:n})^n$. Moreover, 

$$\lim_{n \to \infty} \text{FDR}_{n,n} = 1 - \exp(-\alpha) < \alpha = \lim_{n \to \infty} b(n, n|1).$$
Upper FDR bounds for SUD(\(\lambda\)) tests

Given (BIA) and (M), the upper bounds for the FDR of SUD(\(\lambda_n\)) tests do not depend on the specific \(\vartheta \in \Theta\). Set

\[
b(n, n_0|\lambda) = n_0 \sum_{j=1}^{n_0} \frac{\alpha_{n-n_0+j:n}}{n - n_0 + j} \mathbb{P}_{n,n_0-1}(V_n = j - 1)
\]

and

\[
b^*_n = \max_{1 \leq n_0 \leq n} b(n, n_0|\lambda).
\]

Then

\[
\sup_{\vartheta \in \Theta} \text{FDR}_{\vartheta}(\varphi) \leq b^*_n.
\]

**Computing time** for the upper bounds \(b(n, n_0|\lambda)\) of a SUD(\(\lambda\)) test can be much longer than for an SU test.

This is due to much more complicated recursive formulas for \(\mathbb{P}_{n,n_0-1}(V_n = j - 1)\) for SUD(\(\lambda\)) tests.
SU FDR control implies SUD FDR control

**Theorem:** Given (BIA) and (M). Consider an SU test $\phi^n$ and an SUD test $\phi^\lambda$ with $\lambda \in \{1, \ldots, n-1\}$ and critical values $(\alpha_{i:n})_{i=1}^n$. Then

$$\text{FDR}_{\vartheta}(\phi^\lambda) \leq \text{FDR}_{\vartheta}(\phi^n) \text{ for all } \vartheta \in \Theta$$

and

$$b(n, n_0 | \lambda) \text{ is non-decreasing in } \lambda.$$
Violation of the FDR level for finite $n$

SD tests based on $f_\alpha$ with $\alpha = 0.05$ and $n = 100, 150, 200$

Aim: Tests closely related to AORC with FDR controlled at level $\alpha$. 
Stepwise search for suitable critical values

It always holds \( b(n, n_0|\lambda) \leq n_0/n \). Therefore we reformulate our aim.

**Aim:** Set \( g^*(\zeta) = \min(\zeta, \alpha) \) and try to find critical values such that

\[
b(n, n_0|\lambda) = g^*(n_0/n) \quad \text{for} \quad n_0 \in \{1, \ldots, n\}.
\]

**Recursion:**

\[
\alpha_{n:n} = n g^*(1/n) \quad \text{and} \quad \alpha_{n-n_0+1:n} = h_{n_0}(\alpha_{n-n_0+2:n}, \ldots, \alpha_{n:n})
\]

with

\[
h_{n_0}(\alpha_{n-n_0+2:n}, \ldots, \alpha_{n:n}) =
\]

\[
\frac{n - n_0 + 1}{n_0 \mathbb{P}_{n,n_0-1}(V_n = 0)} \left[ g^*(n_0/n) - n_0 \sum_{j=2}^{n_0} \frac{\alpha_{n-n_0+j:n}}{n - n_0 + j} \mathbb{P}_{n,n_0-1}(V_n = j - 1) \right].
\]

Does not work even for small \( n \) !!!
Recursion for SU

**Theorem.** For an SU test, that is $\lambda = n$, $b(n, n_0|n)$ can alternatively be calculated by

\[
b(n, n_0|n) = \sum_{j=1}^{n_0} \frac{j}{n - n_0 + j} \mathbb{P}_{n,n_0}(V_n = j) \tag{2}
\]

and it even holds

\[
b(n, n_0|n) = \text{FDR}_{n,n_0}(\varphi^n)
\]

and

\[
\mathbb{P}_{n,n_0}(V_n = j) = \frac{n_0}{j} \alpha_{n-n_0+j:n} \mathbb{P}_{n,n_0-1}(V_n = j - 1) \text{ for } j = 1, \ldots, n_0. \tag{3}
\]

Equation (3) makes computation much faster for SU than for SUD.
Why does it not work?
Recursion with new FDR bounding curves

**Question:** Does there exist a $g : [0, 1] \rightarrow [0, 1]$ with $g \leq g^*$ such that

$$b(n, n_0 | \lambda) = g(n_0 / n) \text{ for } n_0 \in \{1, \ldots, n\}$$

results in admissible critical values $(\alpha_{i:n})_{i=1}^{n}$ satisfying (M)?

The **only known solution** is $g(\zeta) = \alpha \zeta$ which results in the **LSU test** with

$$\text{FDR}_{n,n_0} = g(n_0 / n) = \frac{n_0}{n} \alpha.$$
Define $g(\zeta|\gamma, \eta) = \alpha (1 - (1 - \zeta/\gamma)^\eta)I_{[0,\gamma]} + \alpha I_{(\gamma,1]}$, $0 \leq \zeta < 1$ with $1 \leq \eta \leq \gamma/\alpha$ and $\alpha \leq \gamma \leq 1$.

$\alpha = 0.05$, $\gamma = 0.5$ and $\eta = 6, 8, 10$
FDR bounding curves: Transformation

**Idea:** Apply the linear transformation with $(1, 0) \rightarrow (1, 0)$ and $(0, 1) \rightarrow (1, 1)$.

\[ g(\zeta | \gamma, \eta) \] and transformed \[ g(\zeta | \gamma, \eta) \] with $\alpha = 0.1$, $\gamma = 1$, $\eta = 50$
FDR bounding curves

Examples:

- Choose cdf of a $\beta$-distribution:
  \[ G_\eta(x) = \alpha(1 - (1 - x/\gamma)^\eta)I_{[0,\gamma)}(x) + \alpha I_{[\gamma,1]}(x), \]

- Choose cdf of an exponential distribution:
  \[ G_\eta(x) = \alpha(1 - \exp(-\eta x))I_{[0,1]}(x). \]

and transform $G_\eta$ to $g(\cdot | \eta) \leq g^*$ as described before.
FDR bounding curves

The recursion with a suitable FDR bounding curve often yields admissible critical values with (M).

**Advantage:** The FDR (or the upper bound) is exactly known.

**Disadvantage:** The recursion may fail.

**General problem:** For larger values of $n$ the computation takes a long time.
Adjusted AORC

Adjusted AORC: \( f_{\alpha, \beta_n}(t) = (1 + \beta_n/n)f_{\alpha}(t) \) with \( \beta_n > 0 \) or a 'snapped off' version.

\[ \alpha = 0.05, \ n = 10, 30, 100, \ \beta_{10} = 1.23, \ \beta_{30} = 1.41, \ \beta_{100} = 1.76 \]

and \( n = 100 \) with \( \beta^*_{100} = 1.41, \ \beta^*_{100} = 1.30 \)
FDR curves for adjusted versions of AORC

\[ \alpha = 0.05, \ n = 100, \ \beta_n = 1.76, \ \beta^*_n = 1.41, \ k = 95, \ \beta^*_n = 1.3, \ k = 90 \]
Behaviour of $\beta_n$ and $\beta^*_n$

- $\beta_n \geq 1 \implies FDR_{n,n_0} \leq \alpha$ for SD tests (Gavrilov et al (2009)).
- $\beta_n \geq 2 \implies b(n, n_0|1) \leq \alpha$ for SD tests.

Red: SU; Green: SD; Yellow: Snapped-off SU with $\beta^*_n$
Behaviour of $\beta_n$  

Summary:

For the best possible choice of $\beta_n$ in $f_{\alpha, \beta_n}$ it holds:

- $\beta_n$ non-decreasing in $\lambda$
- $\beta_n$ bounded for SD tests ($\beta_n \leq 1$)
- $\lim_{n \to \infty} \beta_n/n = 0$ for all types of SUD Tests
- $\lim_{n \to \infty} \beta_n = \infty$, for SU tests ($\lambda = n$)
- $\beta_n \leq (1 - \alpha)n$ (not very helpful).

Conjecture:

$\beta_n$ is bounded for SUD($\lambda_n$) tests with $\lambda_n/n \to \kappa \in [0, 1)$. 
\[\beta\text{-adjustment}\]

\[\beta\text{-adjusted methods always yield admissible critical values.}\]

Corresponding FDR curves (resp. upper FDR bounds) converge to \(\alpha\) or a well defined FDR bounding curve.
Critical values with most influence on FDR under DU
Fix Point Ansatz with respect to critical values

Suppose \((\alpha_{i:n})_{i=1}^n\) satisfies (M).

**Idea:** Re-write critical values \((\alpha_{i:n})_{i=1}^n\) as critical values of an AORC \(f_\alpha\), that is, write

\[
\alpha_{i:n} = \frac{ic_i}{n - i(1 - c_i)} = f^{-1}_{c_i}(i/n), \ i \in I_n.
\]

\(c = (c_1, \ldots, c_n)\): Vector of "local FDR levels".

**Mapping:**

- \(c_i \rightarrow \alpha_{FDR_{n,n_0(i)}(c)}\), \(i = 1, \ldots, i^*\),
- \(c_i \rightarrow c_i, \ i = i^* + 1, \ldots, n\).

\(n_0(i)\) integer closest to \(n - i(1 - \alpha)\),

\(i^* = i^*(n, k, \alpha) = \lfloor (n - k)/(1 - \alpha) \rfloor\),

only FDR\(_{n_0,n}\)-values for \(n_0 = k, \ldots, n\) are iterated.
Iterative method with $\beta$-adjustment: FDR curves

SU test with $n = 100$, $\alpha = 0.05$

Start with $\beta_{100} = 1.76$-adjusted critical values.

Choose $k = 15$, $i^* = 89$ and $J = 50$ iterations.
Iterative Method

Iterative Method yields FDRs close to $\alpha$.

**Advantage:**
Critical values always admissible.

**Disadvantage:**
FDRs sometimes slightly larger than $\alpha$.

**General problem:**
Convergence properties of the algorithm not clear.
Summary of methods

(M1) Critical values based on alternative FDR curves and exact solving;

(M2) Adjusted critical values based on $f_\alpha$ with $\beta_n$ or $\beta^*_n$;

(M3) Critical values based on iterative method.

Recommendation for large $n$:
For example, for $\alpha = 0.05$ and $n \geq 2000$:

SUD tests with $\lambda_n = 0.7n$ and $\beta \in [\beta_{2000}, 2] = [1.58, 2]$  
or SU tests with $\beta^*_n \in [\beta_{2000}, 2] = [1.45, 2]$ for $k \approx n(1 - 2\alpha)$.  


FDR bounding curves and rejection curves

Remember the heuristic for AORC.

Let $g$ denote an FDR bounding curve and consider DU-models with $n_0/n \to \zeta$. Try to find a rejection curve such that

$$\text{FDR}_{\zeta}(t) = g(\zeta).$$

This leads to an implicit definition of $r$ and $\rho = r^{-1}$:

$$r \left( \frac{g(\zeta)(1 - \zeta)}{\zeta(1 - g(\zeta))} \right) = \frac{1 - \zeta}{1 - g(\zeta)}, \quad \zeta \in (0, 1)$$

(4)

$$\rho \left( \frac{1 - \zeta}{1 - g(\zeta)} \right) = \frac{g(\zeta)(1 - \zeta)}{\zeta(1 - g(\zeta))}, \quad \zeta \in (0, 1).$$

(5)
The following lemma shows that $r$ and $\rho$ are well defined for suitable FDR bounding curves $g$.

**Lemma 1.** Let $g$ be a continuous FDR bounding curve such that $g(\zeta)/\zeta$ is non-increasing in $\zeta \in (0, 1]$ and let $b = \lim_{\zeta \to 0} g(\zeta)/\zeta$. Then $r : [0, b] \to [0, 1]$ and $\rho : [0, 1] \to [0, b]$ are well defined via (4) and (5), respectively, and by setting $\rho(0) = r(0) = 0$ and $r(b) = 1$, $\rho(1) = b$. Moreover, $\rho$ fulfills (M).
FDR bounding curves and rejection curves

**Theorem.** Let $g$ be an FDR bounding curve with the same properties as in Lemma 1. Consider SUD($\lambda_n$) tests $\varphi_n$ based on $r$ defined in (4) with $\lambda_n/n \to \kappa$. Then we obtain for the limiting FDR in DU($n, n_0$) models with $n_0/n \to \zeta$ that

$$\lim_{n \to \infty} \text{FDR}_{n,n_0} = g(\zeta)$$

for (i) $\kappa \in (0, 1]$ and $\zeta \in [0, 1]$ if $b < 1$, (ii) $\kappa \in (0, 1)$ and $\zeta \in [0, 1]$ if $b = 1$ and (iii) $\kappa = 0$ and $\zeta \in [0, 1)$. 
Open Problems

- Adjusted AORC: $\beta_n$ bounded for SUD?
- We know that DU is LFC for SU.
- Is DU LFC for SUD($\lambda$) for $\lambda < n$?
- No counterexample is known to the speaker.
- Bounds are not sharp for SUD.
- Do there exist FDR bounding curves working for all $n$ (others than $g(\zeta) = \alpha\zeta$) ?
- FDR bounds for plug-in methods (e.g. Storey’s procedure) are not sharp.
- Is DU LFC for plug-in methods ?
- Do there exist better formulas for the computations of SUD tests ?
References


Thanks for your patience!

Helau!

Altweiber 2010: Jecke feiern auf dem Rathausplatz in Düsseldorf