Scandinavian Journal of Statistics Theory and Applications

doi: 10.1111/j.1467-9469.2012.00791.x

False Discovery Rate Control of Step-Up-Down Tests with Special Emphasis on the Asymptotically Optimal Rejection Curve

HELMUT FINNER and VERONIKA GONTSCHARUK

Institute of Biometrics and Epidemiology, German Diabetes-Center

THORSTEN DICKHAUS

Department of Mathematics, Humboldt-University

ABSTRACT. This paper is concerned with exact control of the false discovery rate (FDR) for step-up-down (SUD) tests related to the asymptotically optimal rejection curve (AORC). Since the system of equations and/or constraints for critical values and FDRs is numerically extremely sensitive, existence and computation of valid solutions is a challenging problem. We derive explicit formulas for upper bounds of the FDR and show that under a well-known monotonicity condition, control of the FDR by a step-up procedure results in control of the FDR by a corresponding SUD procedure. Various methods for adjusting the AORC to achieve finite FDR control are investigated. Moreover, we introduce alternative FDR bounding curves and study their connection to rejection curves as well as the existence of critical values for exact FDR control with respect to the underlying FDR bounding curve. Finally, we propose an iterative method for the computation of critical values.

Key words: dirac-uniform configuration, false discovery rate, false discovery rate bounding curve, least favourable configuration, multiple hypotheses testing, step-up-down test

1. Introduction

In massive-multiple testing problems, strong control of the familywise error rate (FWER) is regarded as a much too conservative criterion leading to procedures with extremely low power. The false discovery rate (FDR) criterion, which bounds the (expected) type I error proportion among all rejections, allows for more type I errors but offers the possibility of more significances and higher power compared with FWER-controlling procedures. Under specific independence assumptions, the original linear step-up (LSU) procedure proposed in Benjamini & Hochberg (1995) controls the FDR at a prespecified level α, but only exhausts α in cases when all n hypotheses are true. If exactly $n_0 < n$ hypotheses are true, the actual FDR level of the LSU test degrades to $n_0\alpha/n$. This fact resulted in various attempts to improve the LSU test with respect to α-exhaustion and, consequently, power. Especially, adaptation techniques have recently attracted special attention. These methods try to incorporate information about the unknown quantity n_0 into the test procedure to exhaust the prespecified level α and to increase power. Explicit adaptation techniques utilize an estimate $\hat{n}_0$ of n_0 and then apply the LSU (or a related) procedure at the relaxed level $n\alpha/\hat{n}_0$, cf. e.g. Storey *et al.* (2004), Langaas *et al.* (2005), Benjamini *et al.* (2006), Sarkar (2008), Blanchard & Roquain (2009) and Celisse & Robin (2010). An alternative avoiding estimation of n_0 was presented in Finner *et al.* (2009). Based on asymptotic considerations, they derived a new nonlinear rejection curve that was termed asymptotically optimal rejection curve (AORC). The AORC leads asymptotically to exact FDR control for least favourable parameter configurations if $\liminf_{n\to\infty} n_0/n \geq \alpha$ and certain independence assumptions hold. Although AORC-based multiple tests control the FDR asymptotically, they can behave liberally for finite systems of hypotheses. The main purpose in this paper is to address the issue of modifying AORC-related stepwise tests such that the FDR is controlled in a prespecified sense for finite n.

1.1. Notation

Let H_i, $i \in \mathbb{N}_n = \{1, \ldots, n\}$, denote n null hypotheses with $\emptyset \neq H_i \subset \Theta$, where Θ denotes the underlying parameter space. The index set of true null hypotheses given $\vartheta \in \Theta$ is denoted by $I_0 = I_0(\vartheta) = \{i \in \mathbb{N}_n : \vartheta \in H_i\}$. Furthermore, n_0 denotes the number of true null hypotheses. Let $\varphi = (\varphi_i : i \in \mathbb{N}_n)$ denote a multiple test procedure for H_i, $i \in \mathbb{N}_n$. For a fixed $\vartheta \in \Theta$ and a given test φ, we define the number of false rejections by $V_n = |\{i \in \mathbb{N}_n : \varphi_i = 1 \text{ and } \vartheta \in H_i\}|$ and the number of all rejections by $R_n = |\{i \in \mathbb{N}_n : \varphi_i = 1\}|$. Then, the FDR of a multiple test φ given $\vartheta \in \Theta$ is defined by $\mathrm{FDR}_\vartheta(\varphi) = \mathbb{E}_\vartheta[V_n/(R_n \vee 1)]$, and the multiple test φ is said to control the FDR at level $\alpha \in (0, 1)$ if $\sup_{\vartheta \in \Theta} \mathrm{FDR}_\vartheta(\varphi) \leq \alpha$. It will be assumed that p-values p_i for testing H_i are at hand, and we denote the ordered p-values by $p_{1:n} \leq \cdots \leq p_{n:n}$. Moreover, we assume that p_i, $i \in I_0 = I_0(\vartheta)$, are independent and identically uniformly distributed on $[0, 1]$ and that $(p_i : i \in I_0)$ and $(p_i : i \in \mathbb{N}_n \setminus I_0)$ are independent random vectors. We refer to these assumptions as (GA).

1.2. Rejection curves and stepwise tests

Helpful tools for theoretical investigations and visualization of procedures are rejection curves, critical value curves and the empirical cumulative distribution function (cdf) $\hat{F}_n$ of all p-values. A rejection curve $r : [0, 1] \to [0, 1]$ is a non-decreasing (continuous) function, and its (generalised) inverse $\rho = r^{-1}$ is called the corresponding critical value curve. Critical values are defined by $\alpha_{i:n} = \rho(i/n)$, $i \in \mathbb{N}_n$. In general, r and ρ do not depend on n in asymptotic considerations while they may for finite n. Sen (1999) mentioned the nice relationship between the empirical cdf $\hat{F}_n$ of n distinct p-values $p_1, \ldots, p_n$, the ordered p-values, the critical values and the rejection curve r:

$$\hat{F}_n(p_{i:n}) \geq r(p_{i:n}) \quad \text{if and only if } p_{i:n} \leq \alpha_{i:n}.$$

The original LSU procedure rejects H_i if and only if $p_i \leq m\alpha/n$, where $m = \max\{i \in \mathbb{N}_n : p_{i:n} \leq \alpha_{i:n}\}$ with $\alpha_{i:n} = i\alpha/n$ for $i \in \mathbb{N}_n$. Hence, the LSU test can be defined in terms of the rejection curve $r(t) = t/\alpha$ called Simes-line.

For a fixed $\alpha \in (0, 1)$, the AORC with respect to asymptotic control of the FDR is defined by

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

This curve was motivated by the idea to exhaust the FDR-level α for extreme parameter configurations at least asymptotically. Various multiple tests related to the AORC with the latter property are investigated in Finner *et al.* (2009). However, the construction of AORC-related multiple tests controlling the FDR for finite n and exhausting the prespecified FDR level as sharp as possible remains a challenging problem and leads to interesting questions as will be seen in this paper. The critical values induced by f_α are given by

$$\alpha_{i:n} = f_\alpha^{-1}(i/n) = \frac{i\alpha}{n - i(1-\alpha)}, \quad i \in \mathbb{N}_n. \tag{1}$$

Unfortunately, a step-up (SU) procedure based on these AORC critical values does not work because of $\alpha_{n:n} = f_\alpha^{-1}(1) = 1$. An interesting alternative class of procedures are step-up-down (SUD) procedures. For $\lambda \in \mathbb{N}_n$, an SUD(λ) procedure $\varphi^\lambda = (\varphi_1, \ldots, \varphi_n)$ (say) of order λ based on some critical values $\alpha_{1:n} \leq \cdots \leq \alpha_{n:n}$ is defined as follows. If $p_{\lambda:n} \leq \alpha_{\lambda:n}$, set $m = \max\{j \in \{\lambda, \ldots, n\} : p_{i:n} \leq \alpha_{i:n}$ for all $i \in \{\lambda, \ldots, j\}\}$, whereas for $p_{\lambda:n} > \alpha_{\lambda:n}$, put $m = \sup\{j \in \{1, \ldots, \lambda - 1\} : p_{j:n} \leq \alpha_{j:n}\}$ ($\sup \emptyset = -\infty$). Define $\varphi_i = 1$ if $p_i \leq \alpha_{m:n}$ and $\varphi_i = 0$ otherwise ($\alpha_{-\infty:n} = -\infty$). Note that $\lambda = 1$ yields a step-down (SD) procedure and $\lambda = n$ yields an SU procedure. In the class of SUD procedures based on the same set of critical values, the SD procedure never leads to more rejections than all the other SUD procedures whereby the SU procedure is most powerful with respect to the number of rejections. However, even SUD procedures of order $\lambda < n$ with critical values (1) fail to control the FDR for finite n, cf. Finner *et al.* (2009).

1.3. Outline of the paper

In section 2, we first introduce an important condition on critical values called feasibility and provide effective formulas for the computation of upper FDR bounds for SUD tests with feasible critical values. Then, we prove that SUD procedures with feasible critical values that provide FDR control for an SU test control the FDR as well. Moreover, we show that FDR control for an SD test may lead to FDR control for SUD tests based on the same set of critical values as the SD procedure. Section 3 deals with general computational issues concerning critical values and FDR. Among others, we provide a recursive scheme for the computation of critical values with respect to so-called FDR bounding curves. Then, in section 4, we discuss alternative FDR bounding curves and investigate the solvability of the recursive scheme introduced in section 3. In section 5, we investigate some simple adjustments of the AORC to obtain sets of critical values ensuring finite FDR control. Finally, an iterative method for adjusting AORC-related critical values is presented in section 6, and some concluding remarks are given in section 7.

2. SUD tests and upper FDR bounds

We consider SUD(λ) tests based on p-values and defined in terms of critical values $0 < \alpha_{1:n} \leq \cdots \leq \alpha_{n:n} \leq 1$ with $\alpha_{i:n} = \rho(i/n)$ for some critical value curve ρ. With respect to bounds for the FDR and least favourable parameter configurations, the condition

$$q(x) = \rho(x)/x \text{ is non-decreasing in } x \tag{2}$$

is extremely helpful. This condition implies that $\alpha_{i:n}/i$ is non-decreasing in i. Such critical values will be called feasible. It is well known that feasible critical values imply that the FDR of SU procedures becomes larger if p-values under alternatives decrease stochastically, cf. Benjamini & Yekutieli (2001). As shown below, feasibility offers the possibility to compute upper bounds for the FDR not only for SU tests, but for all SUD procedures too. In fact, upper bounds for the FDR of SUD procedures with feasible critical values are obtained in one of the so-called Dirac-uniform (DU) configurations (cf. Finner *et al.*, 2009), that is all p-values p_i, $i \in I_0$, are independent and uniformly distributed on [0, 1], and all p_i, $i \in \mathbb{N}_n \setminus I_0$, follow a Dirac distribution with point mass 1 at 0. We refer to this setting as DU(n, n_0) and replace $\mathbb{P}_\vartheta$, FDR$_\vartheta$ and $\mathbb{E}_\vartheta$ by $\mathbb{P}_{n,n_0}$, FDR$_{n,n_0}$ and $\mathbb{E}_{n,n_0}$. In case of SU, the upper bound is sharp if the corresponding DU configuration belongs to the model. For SUD procedures with parameter $\lambda < n$, DU configurations may not be least favourable for the FDR, cf. an example in Blanchard, G., Dickhaus, T., Roquain, E. & Villers, F. (2011), *On least favourable configurations for step-up-down tests*, unpublished manuscript. However, Roquain & Villers (2011) give specific conditions under which DU configurations are least favourable for an

SD test. Moreover, DU configurations yield an upper bound for the FDR in this case. The following result under (GA) for SUD(λ) tests φ^{λ} corresponds to the slightly more general theorem 4.3 in Finner *et al.* (2009). Below, P_{ϑ^i} refers to the situation where $(p_j : j \in \mathbb{N}_n \setminus \{i\})$ has the same distribution under ϑ^i as under ϑ but $p_i \equiv 0$ under ϑ^i.

Theorem 1. *Let* $\vartheta \in \Theta$ *and* $i \in I_0$. *Then, for an* SUD(λ) *test with* $\lambda \in \mathbb{N}_n$, *it holds under* (GA) *and* (2)

$$\mathrm{FDR}_{\vartheta}(\varphi^{\lambda}) \leq \frac{n_0}{n} \sum_{j=1}^{n} q(j/n) \mathbb{P}_{\vartheta^i}(R_n/n = j/n) = \frac{n_0}{n} \mathbb{E}_{\vartheta^i} q(R_n/n) \tag{3}$$

$$\leq \frac{n_0}{n} \mathbb{E}_{n,n_0-1} q(R_n/n), \tag{4}$$

with equality in (3) *for an* SU *test, i.e. for* $\lambda = n$.

This result yields an explicit formula for a ϑ-free upper bound for the FDR of the SUD(λ) test for all $n_0 \in \mathbb{N}_n$ given by

$$b(n, n_0 \mid \lambda) = \frac{n_0}{n} \mathbb{E}_{n,n_0-1} q(R_n/n) = n_0 \sum_{j=1}^{n_0} \frac{\alpha_{n_1+j:n}}{n_1+j} \mathbb{P}_{n,n_0-1}(V_n = j-1). \tag{5}$$

Hence, the FDR of φ^{λ} is bounded by $\max_{1 \leq n_0 \leq n} b(n, n_0 \mid \lambda)$. For SU, that is $\lambda = n$, the bound equals $\mathrm{FDR}_{n,n_0}(\varphi^n)$ which results in the alternative formula

$$b(n, n_0 \mid n) = \sum_{j=1}^{n_0} \frac{j}{n_1+j} \mathbb{P}_{n,n_0}(V_n = j). \tag{6}$$

Moreover, for the SU procedure, we get the nice recursive formula

$$\mathbb{P}_{n,n_0}(V_n = j) = \frac{n_0}{j} \alpha_{n_1+j:n} \mathbb{P}_{n,n_0-1}(V_n = j-1) \text{ for } j \in \mathbb{N}_{n_0}, \tag{7}$$

which can easily be verified by considering lemma 3.2 in Finner & Roters (2002).

Formulas for the probability mass function (pmf) of V_n under DU configurations are essential for evaluating $b(n, n_0 \mid \lambda)$. They can be obtained in terms of the joint cdf of order statistics, cf. lemma 3.2 in Gontscharuk (2010). However, computation of the pmf of V_n becomes numerically cumbersome for larger values of n. Computations of $b(n, n_0 \mid n)$, i.e. the upper bound for the SU test, are much easier and faster due to the efficient recursive formula (7). Anyhow, as long as we are able to compute the pmf of V_n for the SUD(λ) procedure with fixed critical values, we can easily compute the bounds given by (5) for the FDR.

Remark 1. Consider a sequence of SUD(λ_n) tests based on some rejection curve r with $\rho = r^{-1}$ satisfying (2) and $\lambda_n/n \to \kappa \in [0, 1]$. Moreover, consider a sequence of DU(n, n_0) models with $n_0/n \to \zeta \in [0, 1]$ and suppose that R_n/n converges almost surely to some fixed value. Then, the bound given in (5) converges to the limiting FDR, that is $\lim_{n\to\infty} b(n, n_0) = \lim_{n\to\infty} \mathrm{FDR}_{n,n_0}$ for all $\zeta \in [0, 1]$ if $\kappa \in (0, 1]$ and for all $\zeta \in [0, 1)$ if $\kappa = 0$ (which includes SD procedures). For $\kappa = 0$ and $\zeta = 1$, bound and FDR may not be equal in the limit. For example, for $n_0 = n$ the FDR of the SD test based on f_{α} equals $1 - (1 - \alpha_{1:n})^n$ which converges to $1 - \exp(-\alpha) < \alpha = \lim_{n\to\infty} b(n, n \mid 1)$.

A question of general interest is whether FDR control of the SU procedure implies FDR control of a corresponding SUD procedure with the same set of critical values. An interesting result in this direction can be found in Blanchard & Roquain (2008), where a specific dependency condition is used. The dependency condition in Blanchard & Roquain (2008) results in very restrictive conditions on the critical values. Theorem 2 states the desired result for SUD tests with (2) under (GA).

Theorem 2. *Consider an* SU *test* φ^n *and an* SUD(λ) *test* φ^λ, $\lambda < n$, *with the same critical value curve* ρ *satisfying* (2). *Then, under* (GA) *it holds*

$$\mathrm{FDR}_\vartheta(\varphi^\lambda) \le \mathrm{FDR}_\vartheta(\varphi^n) \quad \textit{for all } \vartheta \in \Theta. \tag{8}$$

Hence, if the FDR *is controlled by the* SU *test* φ^n, *then the* SUD(λ) *test* φ^λ *also controls the* FDR. *Moreover, the bounds* $b(n, n_0 \mid \lambda)$ *given in* (5) *are non-decreasing in* $\lambda \in \mathbb{N}_n$.

Proof. Set $R_n^\lambda = R_n$ for the SUD(λ) test. The SUD(λ_2) test rejects at least as many hypotheses as the SUD(λ_1) test for any $1 \le \lambda_1 \le \lambda_2 \le n$, which implies that $R_n^{\lambda_1}$ is stochastically not greater than $R_n^{\lambda_2}$. Under (2), we obtain that $\rho(R_n^\lambda/n)/(R_n^\lambda/n)$ is stochastically non-decreasing in λ; hence, the bounds $b(n, n_0 \mid \lambda)$ defined in (5) and $\mathbb{E}_{\vartheta^i} q(R_n^\lambda/n)$ are non-decreasing in λ. Theorem 1 finally yields

$$\mathrm{FDR}_\vartheta(\varphi^\lambda) \le \frac{n_0}{n}\mathbb{E}_{\vartheta^i} q(R_n^\lambda/n) \le \frac{n_0}{n}\mathbb{E}_{\vartheta^i} q(R_n^n/n) = \mathrm{FDR}_\vartheta(\varphi^n).$$

Lemma 1 is a partial reverse of theorem 2 and shows that FDR control of the SD test φ^1 sometimes implies FDR control of the corresponding SUD test φ^λ for certain values of λ, if the corresponding SU test φ^n controls the FDR in DU configurations for large values of n_0.

Lemma 1. *Under* (GA), *let* φ^λ, $\lambda \in \mathbb{N}_n$, *denote* SUD(λ) *tests with a critical value curve* ρ *satisfying* (2) *such that* $b(n, n_0 \mid 1) \le \alpha$ *for all* $n_0 = 1, \ldots, n$. *Define*

$$n_0^* = \min\{k \in \mathbb{N}_n : \mathrm{FDR}_{n,n_0}(\varphi^n) \le \alpha \quad \textit{for all} \quad n_0 = k+1, \ldots, n\} \tag{9}$$

with the convention $\min \emptyset = n$. *Then,* $\mathrm{FDR}_{n,n_0}(\varphi^\lambda) \le \alpha$ *for all* $n_0 \in \mathbb{N}_n$ *and all* $\lambda \le n - n_0^* + 1$, *that is the* SUD($\lambda$) *test controls the* FDR *at level* α *if* $\lambda \le n - n_0^* + 1$.

Proof. Suppose that $n_0^* \le n$. Theorem 2 yields that $\mathrm{FDR}_{n,n_0}(\varphi^\lambda) \le \alpha$ for $n_0 = n_0^* + 1, \ldots, n$ and $\lambda \in \mathbb{N}_n$. A look at lemma 3.2 in Gontscharuk (2010) and formula (5) immediately yields for $\lambda \in \mathbb{N}_n$ that $b(n, n_0 \mid \lambda) = b(n, n_0 \mid 1)$ for all $n_0 \le n - \lambda + 1$. Hence, for $\lambda \le n - n_0^* + 1$ we obtain $\mathrm{FDR}_{n,n_0}(\varphi^\lambda) \le b(n, n_0 \mid \lambda) = b(n, n_0 \mid 1) \le \alpha$ for all $n_0 \le n_0^*$, which completes the proof.

3. General computational issues

The formulas derived in section 2 offer various ways to deal with FDR control of SUD tests of order $\lambda \in \mathbb{N}_n$. Thereby, it suffices to check FDR control for all DU configurations. Noting that any SUD procedure rejects all false hypotheses with probability 1 under DU(n, n_0), we have to check that the FDR is less than or equal to $g^*(n_0/n)$ in this case, where the function g^* is defined by $g^*(\zeta) = \min\{\alpha, \zeta\}$ for $\zeta \in [0, 1]$ and plays an important role in what follows. If

$$b(n, n_0 \mid \lambda) \le g^*(n_0/n) \quad \text{for all } n_0 \in \mathbb{N}_n, \tag{10}$$

then the FDR is controlled at level α.

Clearly, it would be attractive to find critical values close to (1) for SUD procedures as described in the previous sections such that the FDR is strictly controlled at level α and as close as possible to $g^*(n_0/n)$ for $1 \le n_0 \le n$. We call any $g : [0, 1] \to [0, \alpha]$ satisfying the natural restrictions $g(0) = 0$ and $0 < g(\zeta) \le g^*(\zeta)$ for all $\zeta \in (0, 1]$ and some $\alpha \in (0, 1)$ an FDR bounding curve. Suppose for a moment that for each $n_0 \in \mathbb{N}_n$, the FDR under a DU(n, n_0) configuration shall be bounded by $g(n_0/n)$ for some fixed FDR bounding curve g. Then, it is tempting with respect to (5) to require

$$n_0 \sum_{j=1}^{n_0} \frac{\alpha_{n_1+j:n}}{n_1+j} \mathbb{P}_{n,n_0-1}(V_n = j-1) = g(n_0/n) \quad \text{for all } n_0 \in \mathbb{N}_n. \tag{11}$$

Setting $h_{n_0}(\alpha_{n-n_0+2:n}, \ldots, \alpha_{n:n})$

$$= \frac{n-n_0+1}{n_0 \mathbb{P}_{n,n_0-1}(V_n=0)} \left[g(n_0/n) - n_0 \sum_{j=2}^{n_0} \frac{\alpha_{n_1+j:n}}{n_1+j} \mathbb{P}_{n,n_0-1}(V_n = j-1) \right],$$

we formally obtain

$$\alpha_{n:n} = ng(1/n) \quad \text{and} \quad \alpha_{n-n_0+1:n} = h_{n_0}(\alpha_{n-n_0+2:n}, \ldots, \alpha_{n:n}) \quad \text{for } 2 \leq n_0 \leq n, \tag{12}$$

i.e. we get a recursive scheme for the determination of critical values. However, the $\alpha_{i:n}$s determined by (12) may fail to yield a valid set of critical values, that is they may be not monotone or may have values outside of $[0, 1]$. Hence, we have to check whether the resulting solution, i.e. $\alpha_{i:n}$s in (12), is feasible. Unfortunately, for $g \equiv g^*$, the recursive scheme (12) only leads to feasible critical values for very small values of n. For example, for $\alpha = 0.05$ and SU tests, we only get feasible solutions for $n \leq 6$, cf. Kwong & Wong (2002).

A question of general interest is to find functions g such that condition (11) leads to feasible critical values for SUD(λ) tests for all $n \in \mathbb{N}$. There exists at least one such function, namely, $g(\zeta) = \zeta\alpha$, $\zeta \in [0, 1]$, which leads to the LSU procedure introduced in Benjamini & Hochberg (1995). Further candidates will be presented in section 4.

To find feasible critical values close to the ones induced by the AORC, one may relax (11) (and consequently (12)) as follows. In a first step, one may choose $m \in \mathbb{N}_{n-1}$ and initial values $\alpha_{n-i+1:n} \leq \cdots \leq \alpha_{n:n}$, $i = 1, \ldots, m$, satisfying all constraints required for a feasible solution and

$$b(n, i \mid \lambda) \leq g^*(i/n) \quad \text{for } i = 1, \ldots, m, \tag{13}$$

where some of the inequalities may be strict. In a second step, one can try to examine whether recursive computation of the remaining critical values via (12) leads to a feasible solution with

$$b(n, i \mid \lambda) = g^*(i/n) \quad \text{for } i = m+1, \ldots, n. \tag{14}$$

Although this proposal sounds attractive, it turns out to be a balancing act and extremely sensitive with respect to the initial critical values. Our experience is that one needs to be in luck to find a feasible solution with this method for larger values of n. The main reason for the sensitivity of this method seems to be that the new critical value to be calculated via (12) is the smallest value in the support of the distribution of V_n and typically has very small impact on the actual FDR.

4. Alternative FDR bounding curves and exact solving for SU tests

We investigate the question whether there exist further FDR bounding curves g and SU procedures φ^n with a critical value curve ρ_n satisfying (2) such that for all $n \in \mathbb{N}$

$$\text{FDR}_{n,n_0}(\varphi^n) = g(n_0/n) \quad \text{for all } n_0 \in \mathbb{N}_n. \tag{15}$$

If (15) holds for some g and φ^n with (2), the critical values of φ^n are given by (12). Given a fixed g, we say that (15) is solvable if there exists the SU test φ^n with critical value curve ρ_n satisfying (2) such that (15) is fulfilled. Hence, (15) is solvable if (12) leads to feasible critical values. For SUD(λ) tests, (15) may be replaced by $b(n, n_0 \mid \lambda) = g(n_0/n)$ for all $n_0 \in \mathbb{N}_n$. Moreover, we investigate conditions such that $\lim_{n\to\infty} \text{FDR}_{n,n_0}(\varphi^\lambda) = \lim_{n\to\infty} b(n, n_0 \mid \lambda) = g(\zeta)$ for all ζ if $n_0/n \to \zeta$.

4.1. Asymptotic relation between FDR bounding curves and rejection curves

As in Finner *et al.* (2009), one can try to find the asymptotic rejection curve r and the asymptotic critical value curve ρ associated with the FDR bounding curve g. Since ρ should satisfy (2), this imposes further conditions on g as will be seen below. Assume for a moment that $\lim_{n\to\infty} n_0/n=\zeta\in(0,1)$. Then, for a fixed threshold t, the asymptotic FDR with respect to DU configurations is given by

$$\mathrm{FDR}_\zeta(t)=\frac{t\zeta}{(1-\zeta)+t\zeta}.$$

Solving $\mathrm{FDR}_\zeta(t)=g(\zeta)$ for t leads to

$$t_\zeta=\frac{g(\zeta)(1-\zeta)}{\zeta(1-g(\zeta))}.$$

Assuming on the other hand that the threshold for the p-values is determined by the asymptotic crossing point between the rejection curve r and the asymptotic empirical cdf $F_\infty(t\mid\zeta)=\zeta t+(1-\zeta)$ of p-values with respect to DU configurations, this results in an implicit definition of the asymptotic rejection curve r given by $r(t_\zeta)=F_\infty(t_\zeta\mid\zeta)$, or equivalently,

$$r\left(\frac{g(\zeta)(1-\zeta)}{\zeta(1-g(\zeta))}\right)=\frac{1-\zeta}{1-g(\zeta)},\quad \zeta\in(0,1). \tag{16}$$

Analogously, the asymptotic critical value function $\rho\equiv\rho(\cdot\mid\eta)=r^{-1}$ is implicitly defined by

$$\rho\left(\frac{1-\zeta}{1-g(\zeta)}\right)=\frac{g(\zeta)(1-\zeta)}{\zeta(1-g(\zeta))},\quad \zeta\in(0,1). \tag{17}$$

Lemma 2 shows that r and ρ are well defined for suitable FDR bounding curves g.

Lemma 2. *Let g be a continuous FDR bounding curve such that $g(\zeta)/\zeta$ is non-increasing in $\zeta\in(0,1]$ and let $b=\lim_{\zeta\to 0} g(\zeta)/\zeta\in(0,1]$. Then $r:[0,b]\to[0,1]$ and $\rho:[0,1]\to[0,b]$ are well defined via* (16) *and* (17)*, respectively, and by setting $\rho(0)=r(0)=0$ and $r(b)=1$, $\rho(1)=b$. Moreover, ρ fulfils condition* (2)*.*

Proof. Let $\underline{\zeta}=\sup\{\zeta\in[0,1]:g(\zeta)=\zeta\}$. Then, $g(\zeta)=\zeta$ for $\zeta\in[0,\underline{\zeta}]$ and $g(\zeta)<\zeta$ for $\zeta\in(\underline{\zeta},1]$. Moreover, if there exists a $\zeta\in(0,\underline{\zeta})$, then $b=1$ and (16) yields $r(1)=1$ and (17) yields $\rho(1)=1$. Setting $g_1(\zeta)=(1-\zeta)/(1-g(\zeta))$, $\zeta\in[0,1]$, $g_2(\zeta)=g(\zeta)/\zeta$, $\zeta\in(\underline{\zeta},1]$ and $g_2(\zeta)=b$ for $\zeta\in[0,\underline{\zeta}]$, (17) can be written as

$$\frac{\rho(g_1(\zeta))}{g_1(\zeta)}=g_2(\zeta).$$

Since g_2 is non-increasing and g_1 is strictly decreasing on $[\underline{\zeta},1]$, we obtain that $r:[0,b]\to[0,1]$ and $\rho:[0,1]\to[0,b]$ are well defined and ρ fulfils condition (2). From $g_1(0)=1$, $g_1(1)=0$ and $g_2(0)=b$, we obtain the remainder.

We note that if ζ_i denotes the solution of $(1-\zeta)/(1-g(\zeta))=i/n$ with respect to ζ, the asymptotic critical values can be computed by

$$\alpha_{i:n}=\rho(i/n)=\frac{g(\zeta_i)(1-\zeta_i)}{\zeta_i(1-g(\zeta_i))},\quad i\in\mathbb{N}_{n-1}\text{ and }\alpha_{n:n}=b.$$

Theorem 3. *Let g be an* FDR *bounding curve with the same properties as in lemma* 2*. Consider* SUD(λ_n) *tests φ^{λ_n} based on r defined in* (16) *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*

$$\lim_{n\to\infty} \mathrm{FDR}_{n,n_0}(\varphi^{\lambda_n}) = g(\zeta)$$

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* $\zeta \in [0,1)$.

Proof. Let g_1 and g_2 be defined as in the proof of lemma 2. Setting $t_\zeta = g_1(\zeta)g_2(\zeta)$, we obtain that t_ζ as a function of ζ is continuous for $\zeta \in [0,1]$ and strictly decreasing for $\zeta \in [\underline{\zeta}, 1]$ with $t_\zeta = b$ for $\zeta \le \underline{\zeta}$ and $t_\zeta = 0$ for $\zeta = 1$. For each $\zeta \in [0,1]$, it will be shown that $r(t) = F_\infty(t \mid \zeta)$ has at least one solution and at most two solutions in $[0, b]$. Note that from (16), we obtain $r(t_\zeta) = F_\infty(t_\zeta \mid \zeta)$, which implies that there exists at least one solution, namely t_ζ. Now suppose there exists a further solution $t' \neq t_\zeta$. The strict monotonicity of t_ζ in $\zeta \in [\underline{\zeta}, 1]$ yields that there exists a $\zeta' \in [\underline{\zeta}, 1]$ such that $t' = t_{\zeta'}$. Altogether we get $r(t_{\zeta'}) = F_\infty(t_{\zeta'} \mid \zeta') = F_\infty(t' \mid \zeta)$, hence $\zeta = \zeta'$ or $t' = 1$ which implies the existence of at most two solutions, namely $t_\zeta < 1$ and 1 or only t_ζ. Finally, we get $R_n/n \to F_\infty(t_\zeta \mid \zeta) = r(t_\zeta) = g_1(\zeta)$ and $\rho(R_n/n) \to \rho(r(t_\zeta)) = t_\zeta = g_1(\zeta)g_2(\zeta)$, hence $b(n, n_0 \mid \lambda_n) \to g(\zeta)$ for $\zeta \in [0,1]$ with formula (5). Remark 1 completes the proof.

To complete the picture concerning the relationship between asymptotic rejection curves, asymptotic critical value curves and asymptotic FDR bounding curves, the following remark covers the case that we start with an asymptotic rejection curve r.

Remark 2. Let $r:[0,b] \to [0,1]$ be continuous with $b \in (0,1]$ and $r(b) = 1$, and suppose there exists a $\zeta_0 \in [0,1)$ such that for each $\zeta \in (\zeta_0, 1]$ there exists a unique crossing point $t(\zeta)$ between $F_\infty(\cdot \mid \zeta)$ and r on $[0,b]$ if $b<1$ or on $[0,1)$ if $b=1$, while the unique crossing point $t(\zeta)$ on $[0,1]$ is b for $\zeta \in [0, \zeta_0]$. Moreover, suppose that $r(t)/t$ is non-increasing in $t \in (0,b]$. Consider a sequence of $\mathrm{DU}(n, n_0)$ models and a sequence of $\mathrm{SUD}(\lambda_n)$ tests based on r such that $R_n/n \to r(t(\zeta))$ for $n_0/n \to \zeta$. Then, the asymptotic FDR bounding curve on $[0,1)$ is given by

$$g(\zeta) = \frac{\zeta t(\zeta)}{1 - \zeta + \zeta t(\zeta)}$$

and $g(\zeta)/\zeta$ is non-increasing in $\zeta \in (0,1)$ with $\lim_{\zeta \to 0} g(\zeta)/\zeta = b$. Moreover, with $\rho = r^{-1}$ and $\rho(1 - \zeta + \zeta t(\zeta)) = t(\zeta)$, we get

$$\lim_{\zeta \to 1} g(\zeta) = \lim_{\zeta \to 1} \frac{\zeta t(\zeta)}{1 - \zeta + \zeta t(\zeta)} = \lim_{\zeta \to 1} \zeta \frac{\rho(1 - \zeta + \zeta t(\zeta))}{1 - \zeta + \zeta t(\zeta)} = \lim_{t \to 0} \frac{\rho(t)}{t} = q(0),$$

which is in line with the asymptotic results in Finner *et al.* (2009) for SUD procedures, where it is shown that under suitable assumptions, the asymptotic FDR for $n \to \infty$ and $\zeta \to 1$ (or $\zeta = 1$) is $q(0)$.

4.2. *A class of FDR bounding curves*

We now introduce a promising class of FDR bounding curves g which allow to approximate g^* in a smooth way. Let $E = [\underline{\eta}, \infty)$ or $E = (\underline{\eta}, \infty)$ for some $\underline{\eta} \in \mathbb{R}$ and let $G_\eta : [0,1] \to [0, \alpha]$, $\eta \in E$, be continuous and non-decreasing functions such that $G_\eta(x)/x$ is non-increasing in $x \in [0,1]$ with $\lim_{x \downarrow 0} G_\eta(x)/x = b_\eta \in (0, \infty)$, $G_\eta(0) = 0$ for all $\eta \in E$ and $\lim_{\eta \to \infty} G_\eta(x) = \alpha$ for all $x \in (0,1]$. Moreover, G_η is assumed to satisfy one of the following two conditions: (G1) $\exists\, \gamma \in (0, 1-\alpha)$ such that $G_\eta(\gamma) = \alpha$ for all $\eta \in E$ and $G_\eta(x)$ is strictly increasing in $\eta \in E$ for all $x \in (0, \gamma)$, (G2) $G_\eta(x)$ is strictly increasing in $\eta \in E$ for all $x \in (0,1]$. In case of (G2), we formally set $\gamma = 1$. We denote the set of all these $(G_\eta)_{\eta \in E}$ by $\mathcal{G}$.

Now define h_η by $h_\eta(x)=x+G_\eta(x)$ and $g(\cdot\mid\eta):[0,1]\to[0,\alpha]$ by

$$g(\zeta\mid\eta)=G_\eta(h_\eta^{-1}(\zeta)),\quad \zeta\in[0,1]. \tag{18}$$

A little analysis yields that

$$g(\zeta\mid\eta)\le g^*(\zeta)\quad \forall\,\eta\in E \text{ and } \forall\,\zeta\in[0,1],$$
$$g(\zeta\mid\eta)< g^*(\zeta)\quad \forall\,\eta\in E \text{ and } \forall\,\zeta\in(0,\min\{\gamma+\alpha,1\}),$$
$$\lim_{\eta\to\infty} g(\zeta\mid\eta)=g^*(\zeta)\quad \forall\,\zeta\in[0,1],$$
$$\lim_{\zeta\to 0} g(\zeta\mid\eta)/\zeta=b_\eta/(1+b_\eta)\quad \forall\,\eta\in E.$$

If (G1) applies, we have $g(\zeta\mid\eta)=\alpha$ for $\zeta\in[\alpha+\gamma,1]$.

Lemma 3. *Let* $(G_\eta)_{\eta\in E}\in\mathcal{G}$ *and let* $g(\cdot\mid\eta)$ *be defined by* (18). *Then, the asymptotic rejection curve* $r\equiv r(\cdot\mid\eta)$ *defined via* (16) *is strictly increasing on* $[0,b_\eta/(1+b_\eta)]$ *with* $\lim_{\eta\to\infty} r(t\mid\eta)=f_\alpha(t)$, $t\in[0,1]$. *If* (G1) *applies, i.e.* $\gamma+\alpha<1$, *then* $r(t\mid\eta)=f_\alpha(t)$, $t\in[0,t_\gamma]$, *where* $t_\gamma=\{\alpha(1-\alpha-\gamma)\}/\{(1-\alpha)(\gamma+\alpha)\}$. *The asymptotic critical value function* $\rho\equiv\rho(\cdot\mid\eta)$ *defined via* (17) *satisfies the monotonicity condition* (2).

Proof. For $\min\{\gamma+\alpha,1\}<\zeta\le 1$, the asymptotic rejection curve r implicitly defined by (16) coincides with the AORC which has all desired properties. Therefore, it suffices to show the assertions of the lemma for $0\le\zeta\le\min\{\gamma+\alpha,1\}$. In view of lemma 2, we have to show that $g(\zeta\mid\eta)/\zeta$ is continuous (which is trivial) and non-increasing in ζ. Substituting $\zeta=h_\eta(y)$ in $g(\zeta\mid\eta)/\zeta=G_\eta(h_\eta^{-1}(\zeta))/\zeta$, we see that $g(\zeta\mid\eta)/\zeta$ is non-increasing if

$$\frac{G_\eta(y)/y}{G_\eta(y)/y+1}$$

is non-increasing, which is implied by the assumptions.

Clearly, there are uncountable choices of G_η to approach g^* in a smooth way. For example, we can choose $G_\eta=\alpha H_\eta I_{[0,1]}$ for a suitable family of cdf's H_η on $[0,\infty)$ such that G_η has the desired properties, see the following example.

Example 1. (Families of probability distributions for generating FDR bounding curves) Let $\alpha\in(0,1)$.

(a) (Beta distributions) Let $E=[1,\infty)$ and consider the family of beta distributions with cdfs $H_\eta(u)=(1-(1-u)^\eta)I_{[0,1]}(u)+I_{(1,\infty)}(u)$ for $\eta\in E$. Setting $G_\eta=\alpha H_\eta$ and $x=u\gamma$ for some $\gamma\in(0,1-\alpha]$, this leads to

$$G_\eta(x)=\alpha(1-(1-x/\gamma)^\eta)I_{[0,\gamma)}(x)+\alpha I_{[\gamma,1]}(x),\quad \eta\in E.$$

Then, $(G_\eta)_{\eta\in E}\in\mathcal{G}$, hence lemma 3 applies. For convenience, we denote the resulting FDR bounding curves by $g(\cdot\mid\eta,\gamma)$. Note that $g(\cdot\mid\eta,\gamma)$ is non-increasing in $\gamma\in(0,1-\alpha]$ for $\zeta\in[0,1]$. Moreover, $g(\zeta\mid 1,1-\alpha)=\alpha\zeta$, which is the FDR bounding curve of the LSU procedure.

(b) (Exponential distributions) Let $E=(0,\infty)$ and consider the family of exponential distributions with parameter $\eta\in E$ and cdf H_η (say) and define again $G_\eta=\alpha H_\eta$. Then, we have

$$G_\eta(x)=\alpha(1-\exp(-\eta x))I_{[0,1]}(x),\quad \eta\in E,$$

and $(G_\eta)_{\eta\in E}\in\mathcal{G}$ with $\gamma=1$, hence lemma 3 applies here, too.

It seems that one can choose FDR bounding curves of the type introduced in example 1 being close to g^* and leading to feasible critical values in (12) for large values of n. For suitable choices of η and γ in (a) and (b) in example 1, we obtain approximately identical FDR bounding curves and critical value functions (rejection curves). Moreover, in example 1(a) with $\gamma = 0.5$ and example 1(b), we can find suitable ηs for $\alpha = 0.01, 0.05, 0.1$ and $n \leq 1000$ (and probably for much larger values of n) such that (15) is solvable. For instance, if $\alpha = 0.05$, then for $\eta = 16$, $\gamma = 0.5$ in example 1(a) and $\eta = 35$ in example 1(b) there exist feasible critical values with (15) for at least $n \leq 1000$. All in all this approach (as long as it works) yields an attractive possibility to obtain a feasible set of critical values which should not differ too much from the AORC-based critical values (1). Anyhow, it remains completely unclear whether for each n there exists an η such that (15) is solvable.

Of course, for SUD procedures, it is also possible to apply the recursive scheme (12) such that the upper bound is equal to one of the FDR bounding curves considered in example 1. But computations for SUD tests take typically a longer time.

5. AORC adjustments

5.1. AORC with β-adjustment

A simple *ad hoc* possibility to obtain a valid set of critical values for an AORC-related SUD(λ) procedure guaranteeing strict FDR control consists in adjusting the critical values given in (1) in a suitable way. For example, as already mentioned in Finner *et al.* (2009), we can try to find a $\beta_n > 0$ such that the set of critical values

$$\alpha_{i:n} = \frac{i\alpha}{n + \beta_n - i(1-\alpha)}, \quad i \in \mathbb{N}_n, \tag{19}$$

which are always feasible, yields an SUD(λ) test controlling the FDR at level α. Another adjustment of the critical values induced by the AORC can be found in Blanchard & Roquain (2009). The critical values (19) correspond to the rejection curve $f_{\alpha,\beta_n}(t) = \left(1 + \beta_n/n\right) f_\alpha(t)$, $t \in [0, \alpha/(\alpha + \beta_n/n)]$. Formally, β_n is chosen as small as possible such that $b(n, n_0 \,|\, \lambda) \leq \alpha$ for all $n_0 \in \mathbb{N}_n$ with equality for at least one n_0. We call such a β_n b-optimal. A nice result in Gavrilov *et al.* (2009) shows that the SD test with $\beta_n \equiv 1$ always controls the FDR at level α for $\alpha \in (0,1)$ and $n \in \mathbb{N}$.

It follows from the arguments in the proof of theorem 2 together with formula (5) and the monotonicity of β_n-adjusted critical values in β_n that b-optimal β_n-values are non-decreasing in λ_n for SUD(λ_n) tests. In other words, choosing a larger value for λ results in smaller critical values. On the other hand, for fixed critical values, the SUD($\tilde{\lambda}$) test rejects at least as many hypotheses as the SUD(λ) test if $\tilde{\lambda}$ is larger than λ. Hence, there is a trade-off between the conservativeness of critical values, quantified by β_n, and the conservativeness of the test structure, quantified by the parameter λ of the SUD test.

Now, we apply the result in lemma 1 for β_n-adjusted critical values (19). Although the SU test with critical values (19) and β_n b-optimal for the corresponding SD test does not control the FDR for certain values of n_0, we observed in all our calculations that the prechosen FDR level α is exceeded only for a certain set of small values of n_0, that is for each $n \in \mathbb{N}$ there seems to exist an $n_0^* < n$ defined by (9) such that lemma 1 applies. Moreover, the b-optimal β_n of the SD test is always smaller than 2 (cf. Gontscharuk, 2010, p. 63) so that one can take $\beta_n \equiv 2$ instead of the b-optimal β_n for a suitable SUD test.

For convenience, we restrict attention to the case $\alpha = 0.05$. For $n = 100$, 500, 1000, 2000, the b-optimal β_n-values of the corresponding SD tests are given by $\beta_n = 1.34, 1.47, 1.53, 1.58$. Exact calculations showed that due to lemma 1, this results in $n_0^* = 29, 134, 271, 565$. Hence, the corresponding SUD(λ_n) tests with $\lambda_n \leq 72, 367, 730, 1436$ (i.e. $\lambda_n/n \leq 0.72, 0.734, 0.73, 0.7185$) control the FDR, cf. Fig. 1.

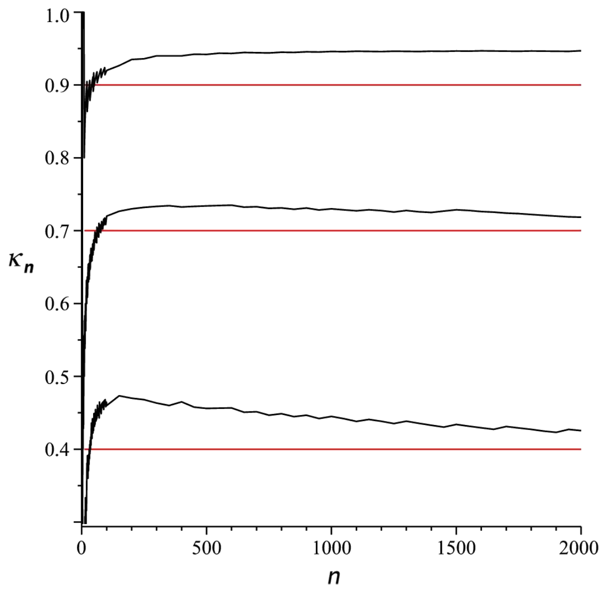

Fig. 1. Maximal values of κ_n such that the SUD(λ_n) test with $\lambda_n/n \le \kappa_n$ and β_n optimized with respect to the SD test controls the FDR at level $\alpha = 0.01, 0.05, 0.1$ (from top to bottom).

Figure 1 shows the exactly calculated maximal possible values of λ_n/n denoted by κ_n for $\alpha = 0.01, 0.05, 0.1$ and $n \le 2000$. Thereby, we observed that κ_n increases if α decreases. It also seems that κ_n slightly decreases for $\alpha = 0.01$ and increases for $\alpha = 0.05, 0.1$ as n increases for $n \ge 200$. Figure 1 implies that the SUD(λ_n) test with $\lambda_n \approx 0.9n, 0.7n, 0.4n$ and β_n b-optimal for the SD test controls the FDR at level $\alpha = 0.01, 0.05, 0.1$ for larger values of n. Moreover, our simulation study for larger values of n and $\alpha = 0.05$ shows that $b(n, n_0 \mid \lambda_n)$ with $\lambda_n = 0.7n$ and $\beta_n \equiv 1$ exceeds the α-level only slightly (for $n = 5000, 10{,}000, 50{,}000$ the maximum upper bound is about $0.05022, 0.05020, 0.05008$), while $b(n, n_0 \mid \lambda_n)$ with $\lambda_n = 0.7n$ and $\beta_n \equiv 2$ seems to be always smaller than α.

5.2. *Another type of AORC adjustment*

Typically, realised FDR-values for SU tests based on (19) under DU configurations with varying number n_0 of true nulls have a maximum peak which is attained for smaller values of n_0, cf. Fig. 2. We briefly mention a possibility to flatten this peak. Since large critical values have most impact on FDR-values for small values of n_0 (cf. the discussion around (11) and (12)), we can shrink the largest critical values and in turn enlarge the smaller ones to reduce the peaking behaviour of the realised FDR-values. For some fixed $k \in \mathbb{N}_n$, we therefore replace the largest critical values $\alpha_{i:n}$, $i \ge k$, in (19) such that $\alpha_{i:n}/i$ become constant for $i \ge k$. This corresponds to the adjusted SU procedures proposed in example 3.2 in Finner *et al.* (2009). Then, we search for a suitable constant $\beta_n^* > 0$ such that the critical values

$$\alpha_{i:n} = \begin{cases} \dfrac{i\alpha}{n + \beta_n^* - i(1-\alpha)}, & 1 \le i \le k-1, \\ i\alpha_{i-1:n}/(i-1), & k \le i \le n \end{cases} \tag{20}$$

yield FDR control at level α.

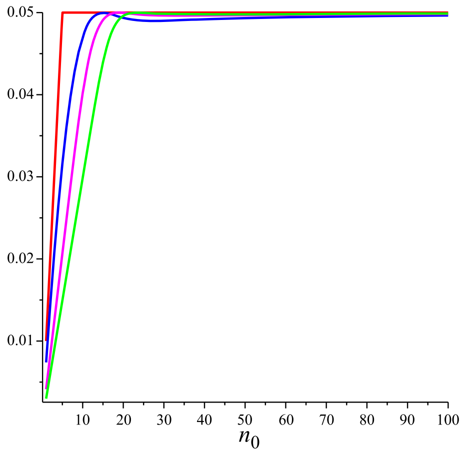

Fig. 2. FDR bounding curve $g^*(n_0/n)$ (the upper curve in both graphs) and FDR-values for SU tests with $\alpha = 0.05$ and $n = 100$ under DU configurations based on simultaneously β_n-adjusted critical values and critical values (20) with $\beta_n^* = 1.41$, $k = 95$ and $\beta_n^* = 1.3$, $k = 90$ (left graph: three lower curves, from top to bottom in $n_0 = 10$). The right graph is zoomed.

To make the procedure not unnecessarily conservative, k should be chosen somewhat larger than a guess for $n - n_0$. If there is no *a priori* information about n_0, the choice of k is a question of taste. For example, k equal to $n(1-\alpha)$ or $n(1-2\alpha)$ may be a good choice. The critical values (20) are always feasible and the b-optimal β_n^* is smaller than the corresponding b-optimal β_n of the simultaneous adjustment method with critical values (19). This results in a flatter FDR curve, such that n_0 is not too small and the values FDR_{n,n_0} are closer to α than the corresponding FDR-values of the simultaneous β_n-adjustment. For $n = 100$ and $\alpha = 0.05$, Fig. 2 shows the FDR bounding curve $g^*(n_0/n)$ and realised FDR-values for the simultaneous β_n-adjusted method with $\beta_{100} = 1.76$ and for β_n^*-adjustment methods with $k = 95$, $\beta_n^* = 1.41$, and, $k = 90$, $\beta_n^* = 1.3$.

The left graph in Fig. 3 shows b-optimal β_n-values for SU and SD based on (19) and β_n^*-values for SU based on (20) for $k = \lceil n(1-2\alpha) \rceil$, $\alpha = 0.05$ and $n = 2, \ldots, 2000$. The right graph of Fig. 3 shows the corresponding values β_n/n and β_n^*/n. Thereby, the curves for β_n/n for SD and β_n^*/n for SU are nearly identical (lower curves).

5.3. Asymptotics of AORC adjustments

A complete characterization of the asymptotic behaviour of the b-optimal β_n- and β_n^*-values remains an interesting open question for the considered β-adjusted procedures. It is not entirely clear for SUD tests whether β_n- and β_n^*-values are bounded or diverge for $n \to \infty$. In any case, we have $\beta_n^* \le \beta_n$. For the SD test with critical values (19), Gontscharuk (2010, p. 63) showed that $\beta_n = 2$ yields control of the bound (5) and the result in Gavrilov *et al.* (2009) even shows that $\beta_n = 1$ works for FDR control. As mentioned before, for the SUD(λ_n) test the b-optimal β_n-values are non-decreasing in λ_n. The same is true for the β_n^*-values. For b-optimal β_n-values of SU tests, we have that $\lim_{n\to\infty} \beta_n = \infty$, $\lim_{n\to\infty} \beta_n/n = 0$ and $\lim_{n\to\infty} f_{\alpha,\beta_n} = f_\alpha$, cf. Gontscharuk (2010, pp. 68–69). Moreover, for SUD(λ_n) tests based on (20) with $\lambda_n \in \mathbb{N}_n$ and $k/n \to b > 0$, there are some indications that β_n^* may be bounded. For SUD(λ_n) tests based on (19) with $\lambda_n/n \to \kappa \in [0, 1)$ it seems also possible that β_n is bounded. Anyhow, these are speculations and need further investigation.

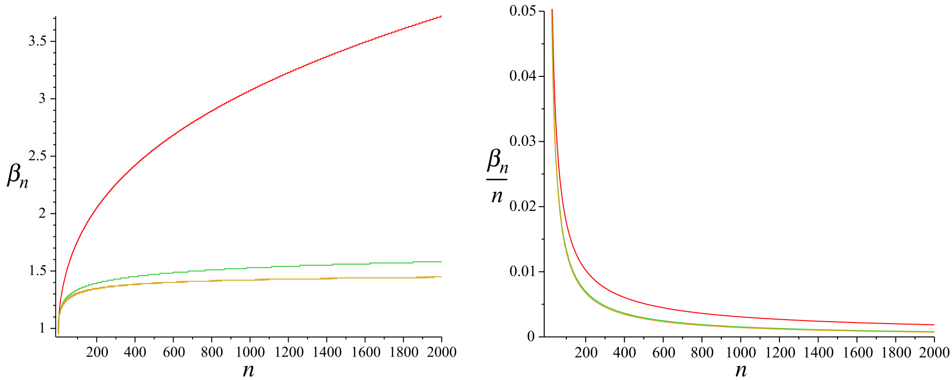

Fig. 3. Left graph: b-optimal β_n-values for SU and SD based on (19) and β_n^*-values for SU based on (20) for $k = \lceil n(1-2\alpha) \rceil$, $\alpha = 0.05$ and $n = 2, \ldots, 2000$. The curves can be distinguished by noticing that for any displayed n, β_n for SU is larger than β_n for SD and β_n for SD is in turn larger than β_n^* for SU. Right graph: Corresponding values β_n/n and β_n^*/n. Thereby, the curves displaying β_n/n for SD and β_n^*/n for SU are nearly identical (lower curves).

6. Iterative method

A further refinement of critical values for the SU test is possible by means of iterative modification. Suppose $\mathrm{FDR}_{n,n_0}(\varphi^n) \approx \alpha$ for all $n_0 \geq k$ for some integer $k \equiv k(n, \alpha) \geq n\alpha$, where φ^n is defined in terms of feasible critical values $\alpha_{1:n}, \ldots, \alpha_{n:n}$. For example, in case of $\alpha = 0.05$ and $n = 100$, the right graph in Fig. 2 suggests that β_n-adjusted AORC-based critical values fulfil this requirement for $k = 15$ and can therefore be taken as initial values. Now, we can try to iteratively modify certain critical values to reduce the corresponding distances $|\alpha - \mathrm{FDR}_{n,n_0}(\varphi^n)|$ even further.

To this end, we first have to identify which critical values have the most impact on $\mathrm{FDR}_{n,n_0}(\varphi^n)$ for a given value of n_0. We recall that, at least for $\zeta < 1$, the ratio V_n/R_n converges to the limiting FDR in DU models with $n_0/n \to \zeta$. Since p-values under alternatives are equal to 0 with probability 1 in DU models, we conclude that $V_n/(V_n + n_1)$ should approximately equal α leading to $V_n \approx n_1\alpha/(1-\alpha)$. Consequently, approximately $R_n = V_n + n_1 \approx n_1/(1-\alpha)$ hypotheses get rejected by an α-exhausting SU test in DU models. Therefore, the critical values $\alpha_{i:n}$ with i close to $n_1/(1-\alpha)$ are crucial for $\mathrm{FDR}_{n,n_0}(\varphi^n)$, and accordingly for a given $i \in \mathbb{N}_n$, the critical value $\alpha_{i:n}$ has the most impact on $\mathrm{FDR}_{n,n_0}(\varphi^n)$ with $n_0 \approx n - i(1-\alpha)$. To modify $\mathrm{FDR}_{n,n_0}(\varphi^n)$-values for all $n_0 = k, \ldots, n$, we have to modify critical values with indices ranging from 1 to $i^* \equiv i^*(n, k, \alpha)$, which is an integer close to $(n-k)/(1-\alpha)$. For the derivation of an appropriate iteration scheme, we rewrite the initial critical values in the form

$$\alpha_{i:n} = \frac{ic_i}{n - i(1-c_i)} = f_{c_i}^{-1}(i/n), \quad i \in \mathbb{N}_n, \tag{21}$$

which formally equals (1) with a vector of 'local FDR levels' $c = (c_1, \ldots, c_n)$. Moreover, we make use of the notation $\mathrm{FDR}_{n,n_0}(c)$ for $\mathrm{FDR}_{n,n_0}(\varphi^n)$, where φ^n is defined via the critical values given in (21). Now, let $n_0(i)$ be an integer closest to $n - i(1-\alpha)$ and consider the mapping $c \mapsto u(c) = (u_1(c), \ldots, u_{i^*}(c), c_{i^*+1}, \ldots, c_n)$, where

$$u_i(c) = \alpha \frac{c_i}{\mathrm{FDR}_{n,n_0(i)}(c)}, \quad i = 1, \ldots, i^*.$$

We note that $\mathrm{FDR}_{n,n_0}(u(c)) = \mathrm{FDR}_{n,n_0}(c)$ for $n_0 = 1, \ldots, n - i^*$. Assume for the moment that for fixed, given constants c_i, $i = i^* + 1, \ldots, n$, there exist c_i^*, $i = 1, \ldots, i^*$, such that $c^* =$

$(c_1^*, \ldots, c_{i^*}^*, c_{i^*+1}, \ldots, c_n)$ fulfils the fixed point property $c^* = u(c^*)$. Then, we have $c_i^* = u_i(c^*) = \alpha c_i^* / \mathrm{FDR}_{n,n_0}(c^*)$, $i = 1, \ldots, i^*$, which is equivalent to $\mathrm{FDR}_{n,n_0}(c^*) = \alpha$, $n_0 = k, \ldots, n$. Therefore, an iteration scheme for the vector of local FDR levels c, i.e. setting $c^{(j)} = u(c^{(j-1)})$, seems to be a promising approach. Clearly, there is no fixed-point theorem at hand guaranteeing convergence. Moreover, as mentioned in section 3, for a given FDR-bounding curve (or prespecified FDR-values) there are not necessarily corresponding feasible critical values, i.e. the formal solution of the target equation (15) is not necessarily feasible. Finally, resulting FDR-values can slightly exceed the given α-level, because the mapping $u(c)$ does not guarantee that c^* is approached from below. However, the method seems to work well and the distances $|\alpha - \mathrm{FDR}_{n,n_0}(\varphi^n)|$, $n_0 = k, \ldots, n$, on average get reduced by the outlined iteration method for a suitable number of iterations $J \in \mathbb{N}$ (say).

To describe the method more formally, we put without restriction $i^* \equiv i^*(n, k, \alpha) = \lfloor (n-k)/(1-\alpha) \rfloor$, and let $\alpha_{1:n}^{(0)}, \ldots, \alpha_{n:n}^{(0)}$ be feasible starting critical values and $\alpha_{i:n}^{(j)} \equiv \alpha_{i:n}^{(0)}$ for $i = i^* + 1, \ldots, n$, $j = 1, \ldots, J$. For modification of critical values with indices ranging from 1 to i^*, we proceed as given in the following algorithm.

For j from 1 to J do

1 For i from i^* to 1 by -1 do:
  (a) Determine $c_i^{(j-1)}$ from $\alpha_{i:n}^{(j-1)} = i c_i^{(j-1)} / (n - i(1 - c_i^{(j-1)}))$.
  (b) Put $c_i^{(j)} = \alpha c_i^{(j-1)} / \mathrm{FDR}_{n,n_0(i)}(c^{(j-1)})$.
  (c) Calculate $\alpha_{i:n}^{(j)} = i c_i^{(j)} / (n - i(1 - c_i^{(j)}))$.
  (d) If $\alpha_{i:n}/i > \alpha_{i+1:n}/(i+1)$, then put $\alpha_{i:n}^{(j)} = i \alpha_{i+1:n}^{(j)} / (i+1)$.

2 Calculate $\mathrm{FDR}_{n,n_0}(c^{(j)})$, $n_0 = n - i^* + 1, \ldots, n$.

Notice that in the latter algorithm, the number $n_0(i)$ in the expression $\mathrm{FDR}_{n,n_0(i)}(c^{(j-1)})$ is only loosely defined by setting $n_0(i)$ as the integer 'closest to $n - i(1-\alpha)$'. To be more precise, one can replace $\mathrm{FDR}_{n,n_0(i)}(c^{(j-1)})$ by a linear interpolation of the two adjacent values $\mathrm{FDR}_{n,\lfloor n-i(1-\alpha) \rfloor}(c^{(j-1)})$ and $\mathrm{FDR}_{n,\lceil n-i(1-\alpha) \rceil}(c^{(j-1)})$.

We tested the iterative method for a variety of values for n and α. As initial critical values, we took simultaneous β_n-adjusted and β_n^*-adjusted critical values, cf. sections 5.1 and 5.2. For example, for $n = 100, 300, 1000$, $J = 20, 10, 10$ iterations based on initial simultaneous β_n-adjustment and $J = 10, 2, 1$ iterations based on initial β_n^*-adjustment gave satisfying results. Although the resulting realised FDR-values under DU configurations typically exceed the given FDR level α for some $n_0 \geq k$, the actual differences $|\alpha - \mathrm{FDR}_{n,n_0}(\varphi^n)|$ for $n_0 \geq k$ seem to be of negligible magnitude, i.e. for a suitable number of iterations the observed differences were never larger than 5×10^{-5}. Clearly, in a final step we can decrease the resulting critical values by a suitable small amount such that all FDRs are smaller than α.

7. Concluding remarks

We investigated and implemented different approaches to construct critical values for SUD procedures close to AORC-based critical values to exhaust the given FDR level as much as possible. There seems to be no silver bullet for the computation of a set of critical values coming close to AORC critical values. Typically, different methods lead to different sets of critical values, and no set uniformly dominates the others. Finally, preference of any method is a matter of taste and also depends on computational resources and complexity.

For computation of critical values, we provide Maple worksheets under the URL http://www.helmut-finner.de, which can be executed in reasonable time on a standard desktop computer for $n \leq 2000$. Moreover, for SU and SD tests with critical values (19) (cf. section

5.1) and SU tests based on (20) (cf. section 5.2), we tabulated the constants β_n (β_n^*, respectively) for $n \leq 2000$, $\alpha = 0.01, 0.05, 0.1$. At least for $n \geq 100$, these SD β_n values can be used in SUD tests with $\lambda_n \leq 0.9n, 0.7n, 0.4n$.

Finally, we would like to give a recommendation for practical application. For $n \leq 2000$ the easiest way is to apply the SU test based on the tabulated b-optimal constants β_n or to apply the SUD(λ_n) test based on the tabulated b-optimal for SD constants β_n with a favourite and suitable value for λ_n or to run the provided Maple programs for other values of α than $0.01, 0.05, 0.1$ to obtain valid critical values. If SU tests based on (20) are preferred, one can use the tabulated values for β_n^* or the corresponding Maple program. For those who like the iterative approach (cf. section 6), we recommend this method with simultaneous β_n-adjusted critical values as starting values for smaller values of n (e.g. $n \leq 300$). For $300 \leq n \leq 2000$, we recommend to use the iterative method in connection with β_n^*-adjusted initial critical values, because only a few iterations are needed in this case. Furthermore, the method described in section 4.2 with suitable FDR bounding curves introduced in example 1 yields a reasonable alternative for at least $n \leq 2000$ as long as it results in a feasible solution. In addition, it offers the possibility to choose an alternative FDR bounding curve from a large class of different curves.

For larger n-values ($n > 2000$), computing time for all considered methods can be enormous, such that we recommend to use a β-adjustment with some fixed parameter β (or β^*). For example, for $\alpha = 0.05$ one may choose $\beta_n \in [\beta_{2000}, 2] = [1.58, 2]$ and $\lambda_n \approx 0.7n$ for the SUD(λ_n) test and $\beta_n^* \in [\beta_{2000}, 2] = [1.45, 2]$ for the SU test with critical values (20) for $k \approx n(1 - 2\alpha)$. Although the upper FDR bound can exceed α for these tests for some DU configurations, the possible exceedance should be negligible. As mentioned before, the FDR is asymptotically controlled such that the possible exceedance of the α-level converges to 0 as n increases.

References

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B* **57**, 289–300.

Benjamini, Y. & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.* **29**, 1165–1188.

Benjamini, Y., Krieger, A. & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* **93**, 491–507.

Blanchard, G. & Roquain, E. (2008). Two simple sufficient conditions for FDR control. *Elect. J. Statist.* **2**, 963–992.

Blanchard, G. & Roquain, E. (2009). Adaptive false discovery rate control under independence and dependence. *J. Mach. Learn. Res.* **10**, 2837–2871.

Celisse, A. & Robin, S. (2010). A cross-validation based estimation of the proportion of true null hypotheses. *J. Statist. Plann. Inference* **140**, 3132–3147.

Finner, H. & Roters, M. (2002). Multiple hypotheses testing and expected number of type I errors. *Ann. Statist.* **30**, 220–238.

Finner, H., Dickhaus, T. & Roters, M. (2009). On the false discovery rate and an asymptotically optimal rejection curve. *Ann. Statist.* **37**, 596–618.

Gavrilov, Y., Benjamini, Y. & Sarkar, S. K. (2009). An adaptive step-down procedure with proven FDR control. *Ann. Statist.* **37**, 619–629.

Gontscharuk, V. (2010). *Asymptotic and exact results on FWER and FDR in multiple hypotheses testing*. Dissertation, Heinrich-Heine Universität, Düsseldorf. Available on: http://docserv.uni-duesseldorf.de/servlets/DerivateServlet/Derivate-18133.

Kwong, K.-S. & Wong, E.-H. (2002). A more powerful step-up procedure for controlling the false discovery rate under independence. *Statist. Probab. Lett.* **56**, 217–225.

Langaas, M., Lindqvist, B. H. & Ferkingstad, E. (2005). Estimating the proportion of true null hypotheses, with application to DNA microarray data. *J. Roy. Statist. Soc. Ser. B* **67**, 555–572.

Roquain, E. & Villers, F. (2011). Exact calculations for false discovery proportion with application to least favorable configurations. *Ann. Statist.* **39**, 584–612.

Sarkar, S. K. (2008). On methods controlling the false discovery rate. *Sankya* **70-A**, 135–168.

Sen, P. K. (1999). Some remarks on Simes-type multiple tests of significance. *J. Statist. Plann. Inference* **82**, 139–145.

Storey, J. D., Taylor, J. E. & Siegmund, D. (2004). Strong control, conservative point estimation, and simultaneous conservative consistency of false discovery rates: a unified approach. *J. Roy. Statist. Soc. Ser. B* **66**, 187–205.

Received December 2009, in final form November 2011

Helmut Finner, Institute of Biometrics and Epidemiology, German Diabetes-Center, Auf'm Hennekamp 65, D-40225 Düsseldorf, Germany.
E-mail: finner@ddz.uni-duesseldorf.de